



**QM/MM Investigations  
on the  
Hydroxylation Reactions of  
Molybdopterin-containing Enzymes**

Inaugural-Dissertation

zur Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

**Sebastian Metz**  
aus Hagen i.W.

Mülheim an der Ruhr/Düsseldorf 2009



Aus dem Institut für Theoretische Chemie und Computerchemie der  
Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen  
Fakultät der Heinrich-Heine-Universität Düsseldorf

Referent:	Prof. Dr. Walter Thiel
Korreferenten:	Prof. Dr. Christel M. Marian
	Prof. Dr. Bernd Engels

Tag der mündlichen Prüfung: 17.12.2009





## List of papers included in this thesis

### Paper I

*Reductive Half-Reaction of Aldehyde Oxidoreductase toward Acetaldehyde: A Combined QM/MM Study*

**Sebastian Metz**, Dongqi Wang, and Walter Thiel, *J. Am. Chem. Soc.*, 131 (2009) 4628-4640.

Three different system setups were used in this paper: The one with an overall charge of -12e was provided by Dongqi Wang, while the two neutral systems were defined by myself. I carried out and evaluated all the QM/MM calculations on the latter two.

### Paper II

*Reductive Half-Reaction of Aldehyde Oxidoreductase toward Acetaldehyde: Ab Initio and Free Energy QM/MM Calculations*

Johannes M. Dieterich, Hans Joachim Werner, Ricardo A. Mata, **Sebastian Metz**, and Walter Thiel, *J. Chem. Phys.*, submitted.

The setup of this paper was taken from Paper I. The high-level single-point calculations were done by Johannes M. Dieterich and Ricardo A. Mata. I performed the free-energy calculations and prepared the manuscript.

### Paper III

*A Combined QM/MM Study on the Reductive Half-reaction of Xanthine Oxidase: Substrate Orientation and Mechanism.*

**Sebastian Metz** and Walter Thiel, *J. Am. Chem. Soc.*, 131 (2009) 14885-14902.

### Paper IV

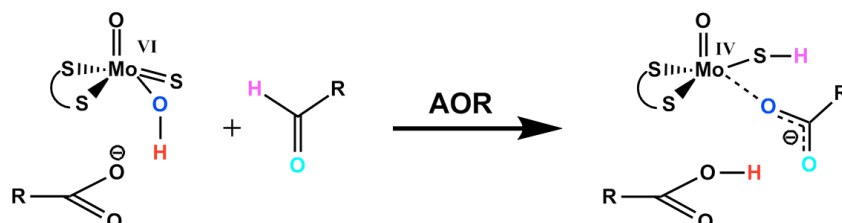
*QM/MM Studies of Xanthine Oxidase: Variations of Cofactor, Substrate, and Active-site Glu802*

**Sebastian Metz** and Walter Thiel, *J. Phys. Chem. B*, submitted.

## Zusammenfassung

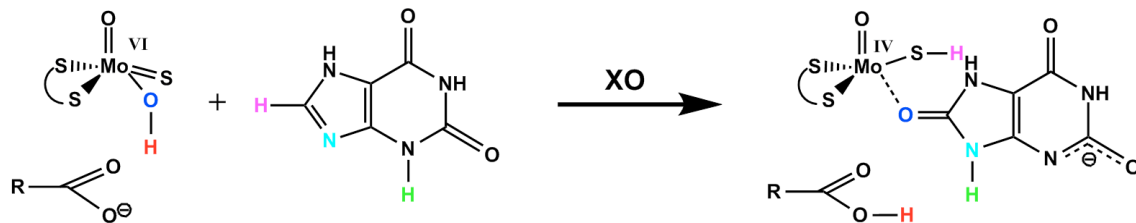
Die vorliegende Arbeit untersucht die Hydroxilierungsreaktionen von zwei Enzymen (Aldehyd-Oxidoreduktase und Xanthinoxidase), die leicht verschiedene Molybdopterin-Kofaktoren enthalten, mit Hilfe eines quantenmechanischen/molekülmechanischen (QM/MM) Ansatzes. Hierbei wird ein kleiner Teil des Gesamtmodells aus Enzym, Substrat und Solvens mit quantenmechanischen Methoden (zumeist Dichtefunktionaltheorie) und der Rest mit molekularmechanischen Methoden (dem CHARMM-Kraftfeld) beschrieben.

Für die durch Aldehyd-Oxidoreduktase (AOR) katalysierte Reaktion von Acetaldehyd zu Essigsäure wurden fünf unterschiedliche Mechanismen untersucht, die bis dato nur teilweise in der Literatur diskutiert wurden. Anhand der erhaltenen QM/MM-Energien konnte ein Mechanismus bevorzugt werden, der nicht nur die aufgrund experimenteller Ergebnisse postulierte Funktion von Glu869 bestätigt, sondern auch deutlich geringere Barrieren als die übrigen Mechanismen aufweist. Um diese Ergebnisse abzusichern, wurden zusätzlich Freie-Energie-Barrieren für die jeweiligen geschwindigkeitbestimmenden Schritte der verschiedenen Mechanismen berechnet. Die hierbei erhaltenen Ergebnisse führen zwar zu merklichen Korrekturen an den Barrieren, jedoch ergeben sich hieraus keine qualitativen Änderungen am vorgeschlagenen Mechanismus.



Bezüglich der Umwandlung von Xanthin zu Harnsäure in Xanthinoxidase (XO) wurden insgesamt sieben verschiedene System-Varianten mit unterschiedlichen Xanthin-Tautomeren, verschiedenen Protonierungszuständen und unterschiedlicher Substratorientierung untersucht. Aufgrund der sehr viel spezifischeren Koordination des Substrats (verglichen mit AOR) ist eine Umorientierung in der Bindungstasche nicht möglich. In allen System-Varianten erhält man für XO einen mehrstufigen Mechanismus mit einer anfänglichen Aktivierung, gefolgt von einem nucleophilen Angriff des Kofaktors auf das Substrat und einem abschließendem Hydridtransfer. Die sieben untersuchten Varianten unterscheiden sich in ihren Möglichkeiten, die auftretenden Intermediate zu stabilisieren. Zusätzliche Rechnungen wurden durchgeführt, um intrinsische Reaktivitäten und eine gemeinsame Energieskala zu etablieren.

Aufbauend auf diesen Ergebnissen wurde durch zusätzliche Rechnungen untersucht, wie sich Veränderungen am Kofaktor (Oxo-, Sulfido- und Selenido-Form), am Substrat (2-Oxo-6-methylpurin statt Xanthin) und an einer spezifischen Aminosäure in der Bindungstasche (Glu802→Gln Mutation) auswirken. Die hierbei erhaltenen Resultate stehen allesamt im Einklang mit den verfügbaren experimentellen Daten.

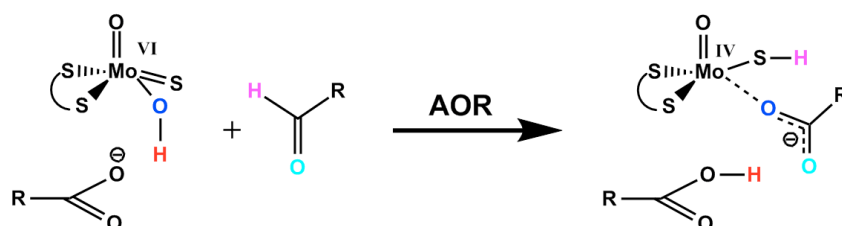


Aus den Untersuchungen beider Enzyme, AOR und XO, ergibt sich zusammen mit den vorliegenden experimentellen Ergebnissen ein konsistentes Bild, mit dem sich die Reaktivität der untersuchten Enzymfamilie umfassend erklären lässt.

## Summary

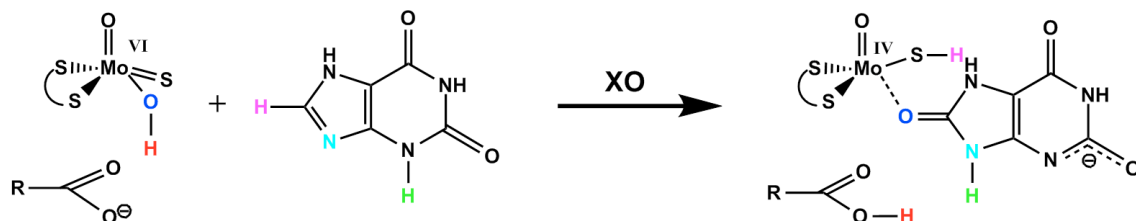
The present work investigates the hydroxylation reaction of two enzymes (aldehyde oxidoreductase and xanthine oxidase), which contain slightly different molybdopterin cofactors, using a quantum-mechanical/molecular-mechanical (QM/MM) approach. Herein, a small part of the complete model consisting of enzyme, substrate, and solvent, is described by quantum-mechanical methods (mostly density functional theory) and the rest is treated with molecular-mechanical methods (the CHARMM force field).

For the reaction from acetaldehyde to acetic acid, catalyzed by aldehyde oxidoreductase (AOR), five different reaction mechanisms were investigated, which had only been partially discussed in the literature so far. Based on the calculated QM/MM energies, a Lewis base catalyzed mechanism could be favored, not only confirming the function of Glu869 postulated on the basis of experimental results, but also coming up with significantly lower reaction barriers than the alternative pathways. To corroborate these results, free-energy barriers for the rate-determining steps of each mechanism were computed. This leads to appreciable corrections of the barriers, which however do not cause any qualitative changes in the proposed reaction mechanism.



Concerning the conversion of xanthine to uric acid by xanthine oxidase (XO), seven different setups were studied, covering different tautomers of xanthine, different protonation states, and different orientations of the substrate. Due to the much more specific binding of the substrate (compared to AOR), its rearrangement within the binding pocket is impossible. All setups therefore follow a multi-step mechanism in XO, consisting of an initial activation, followed by a nucleophilic attack of the cofactor toward the substrate and a final hydride transfer. The seven investigated variants differ by their ability to stabilize the occurring intermediates. Additional calculations were performed to establish an intrinsic reactivity and a common energy scale.

Based on these results, the consequences of variations in the cofactor (oxo, sulfido and selenido form), in the substrate (2-oxo-6-methylpurin instead of xanthine) and in one specific active-site residue of the binding pocket (Glu802 →Gln mutation) were explored through additional calculations. The results obtained are compatible with the corresponding experimental findings.



The investigations on both enzymes, AOR and XO, together with the available experimental evidence, provide a consistent mechanistic picture, enabling a thorough explanation of the reactivity of this family of enzymes.

## Contents

1 Introduction	1
1.1 General remarks	2
1.2 Hydroxylation reactions of molybdopterin-containing enzymes	3
1.3 Cofactor	3
1.4 Active site	5
1.5 Catalytic cycle	5
1.6 Theoretical investigations	8
2 Theoretical Methods	11
2.1 The QM/MM approach	12
2.2 Free-energy perturbation	16
3 Results and Discussion	19
3.1 Mechanism of aldehyde oxidoreductase	20
3.2 Mechanism of xanthine oxidase	27
3.3 Influence of a modified cofactor	33
4 Concluding remarks	37
5 Appendix	39
5.1 List of Abbreviations	40
5.2 List of Figures	41
5.3 List of Tables	42
5.4 Acknowledgements	43
5.5 Bibliography	44

# 1 Introduction

## 1.1 General remarks

The aim of this work was the elucidation of the detailed reaction mechanism of aldehyde oxidoreductase (AOR) with acetaldehyde as substrate and of xanthine oxidase (XO) with xanthine as substrate. These investigations have provided valuable insights into mechanistic details that are hard to unravel by experimental methods. This holds especially true with respect to AOR and the ongoing discussion on the nature of the catalytically active cofactor.<sup>1</sup> The reaction mechanism of XO is of special interest in the biochemical and pharmaceutical field: XO is involved in the nitrogen catabolism as a catalyst for the conversion of hypoxanthine to xanthine and on to uric acid, which is then excreted. A dysfunction of XO causes severe diseases; on the other hand, increase in the concentration of uric acid in the blood, due to over-activity, may cause gout.<sup>2</sup> An increasing number of potent inhibitors have been found for XO<sup>3-11</sup> and may overcome the side effects of the "gold standard" alloxanthine. Beside this, there is increasing effort to use the hydroxylation potential of the molybdopterin cofactor in the area of drug metabolism and drug cleavage.<sup>12-16</sup> An accurate insight into the underlying reactivity of molybdopterin-containing enzymes is thus highly desirable.

The key questions to be answered here are:

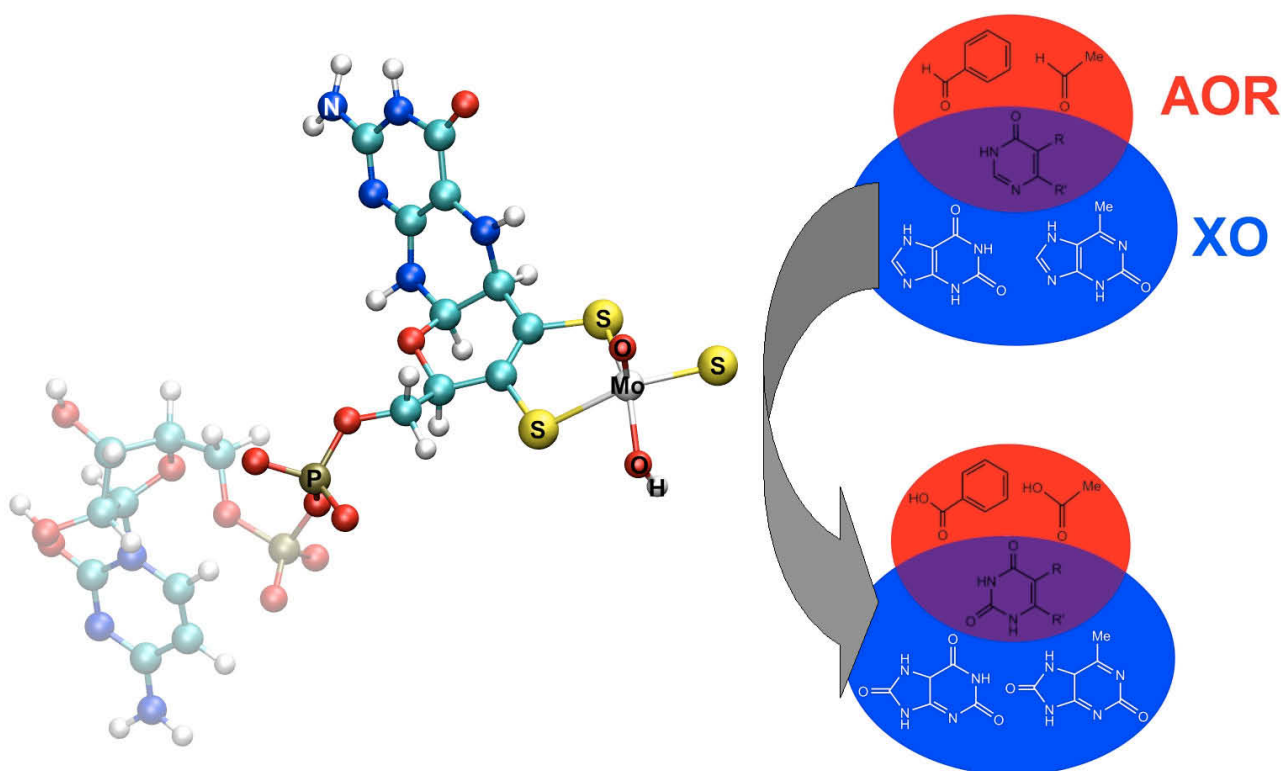
- What is the reaction mechanism of AOR, and how is it influenced by the enzymatic environment or modifications of the cofactor?
- Are the computed potential energies sufficiently accurate to distinguish between competing reaction paths in the AOR system?
- What is the reaction mechanism of XO, the orientation of the substrate, and the nature of the reactive species?
- How can one understand the catalytic effect of XO, and how is it influenced by individual active-site residues?
- Can one explain all the available experimental data for XO with the proposed reaction mechanism?

To answer these questions, this thesis is organized as follows: Sections 1.2 to 1.4 provide a detailed overview of the structural features of both enzymes. A succinct summary of the reaction mechanism as described in the literature is given in Section 1.5. Previous theoretical investigations on model systems are discussed in Section 1.6. Chapter 2 presents a short overview of the theoretical methods used throughout this study, namely the QM/MM and the Free-energy perturbation methods. Section 3.1 summarizes the results for the AOR system (papers I and II) and Section 3.2 those obtained for XO (papers III and IV). Unpublished results revealing the influence of a modified cofactor are presented separately in Section 3.3, and concluding remarks will be given in Chapter 4.



## 1.2 Hydroxylation reactions of molybdopterin-containing enzymes

Molybdenum-containing enzymes catalyze metabolic reactions in the nitrogen, sulfur, and carbon cycles.<sup>17</sup> With the exception of the nitrogenase cofactor, containing a  $\text{MoFe}_7\text{S}_9\text{N}$  cluster,<sup>18</sup> the molybdenum cofactors contain a mononuclear molybdenum atom coordinated by the sulfur atoms of a pterin derivative, named molybdopterin. Three families with this kind of cofactor are distinguished by the structural homology of their active sites, namely xanthine oxidase, sulfite oxidase, and dimethyl sulfoxide reductase.<sup>19</sup> The members of the xanthine oxidase family catalyze the reversible transfer of an oxygen atom, derived from water<sup>20</sup> after prior coordination to the molybdenum cofactor as a hydroxo ligand,<sup>21</sup> to a substrate in a two-electron redox reaction.<sup>22</sup> In our computational investigation, we will focus on two particular members, xanthine oxidase (XO) and aldehyde oxidoreductase (AOR), which differ slightly in their cofactor and significantly in their activity towards different substrates,<sup>23</sup> although some substrates are known to be converted by both enzymes, see Figure 1.1.

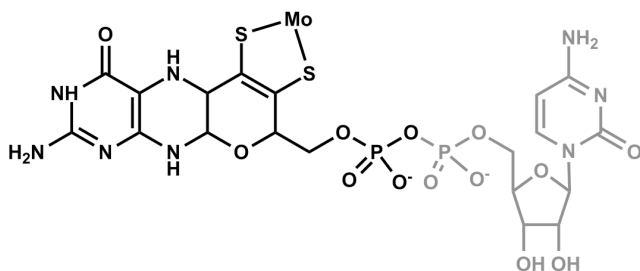


**Figure 1.1:** Substrate specificity of AOR and XO: Aldehydes (red) are naturally catabolized by AOR, whereas purine derivatives (blue) are converted by XO. There are, however, substrates (purple) that can be oxidized by either enzyme. Note that the shaded part of the depicted cofactor is not present in XO.

## 1.3 Cofactor

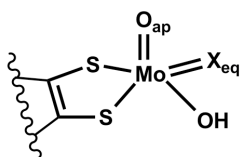
In different enzymes, the molybdopterin cofactor may differ in its nucleotidic part and therefore

interact differently with the corresponding apo-enzymes. A detailed analysis of these interactions for the cytidine (= cytosine + ribose) monophosphate molybdopterin cofactor within AOR has been published by Romão et al.<sup>24</sup> Beyond this form, guanosine (= guanine + ribose) monophosphate, adenosine (= adenine + ribose) monophosphate and inosine (= hyperxanthine + ribose) monophosphate variants were identified.<sup>25</sup> Nevertheless, the function of the different nucleotides is not yet understood, but due to the conservation of the molybdopterin core structure, a common function of the cofactor is assumed.



**Figure 1.2:** Shown is the cytidine monophosphate molybdopterin cofactor as present in AOR. The nucleotidic part is drawn in grey and differs between the several cofactor species, see text above.

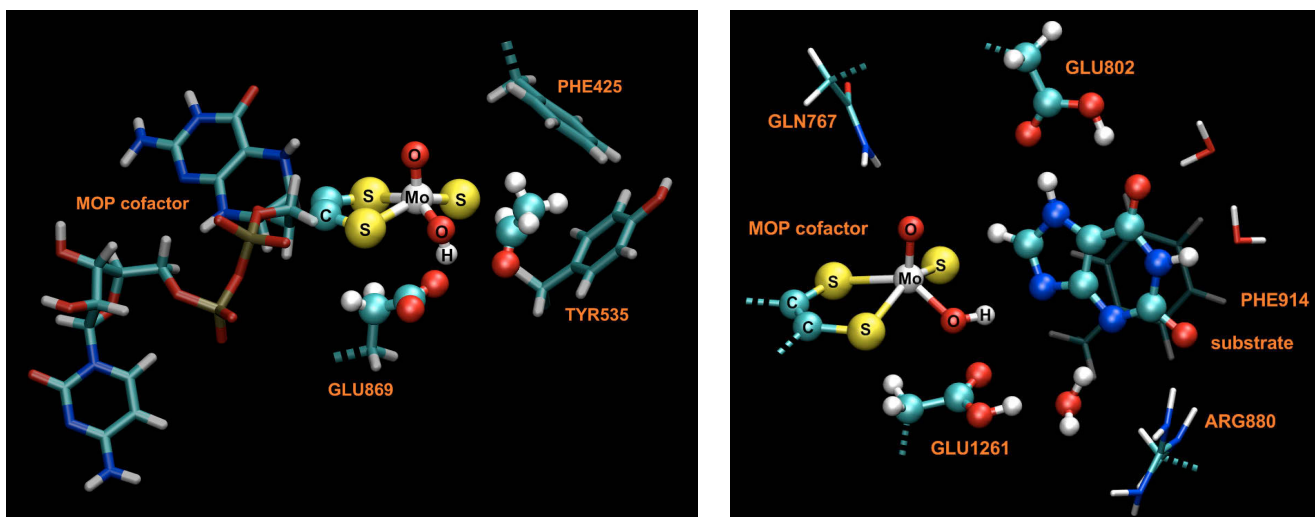
It has been shown by EXAFS experiments that the central atom of the molybdopterin cofactor in XO carries two thiolate ligands (S-Mo) from the connection to the pterin subunit, one terminal oxo ligand ( $\text{Mo}=\text{O}_{\text{ap}}$ ), one hydroxy ( $\text{Mo}-\text{OH}$ ), and one terminal sulfido ligand ( $\text{Mo}=\text{S}_{\text{eq}}$ ).<sup>21</sup> This assignment was accepted as common structural feature within the xanthine oxidase family<sup>17,19,26</sup> and used in nearly all theoretical investigations.<sup>27-34</sup> However, for AOR, in the first crystal structure published for the xanthine oxidase family, the hydride acceptor ligand positioned in the equatorial plane ( $\text{X}_{\text{eq}}$ , see Figure 1.3) was reported to be an oxygen atom ( $\text{O}_{\text{eq}}$ ).<sup>35</sup> As the analogous desulfo form of XO was known to be inactive,<sup>36</sup> it was assumed that this crystal structure of AOR may represent an inactive oxo form of the naturally active sulfido form. Attempts to restore activity by resulfuration with  $\text{H}_2\text{S}$  failed as the sulfur atom was added to the apical position ( $\text{S}_{\text{ap}}$ ) whereas the equatorial position was still occupied by an oxygen atom,  $\text{O}_{\text{eq}}$ . Recent results, however, indicate that AOR may work as well with an equatorial oxo ligand as a variant of the commonly accepted sulfido form.<sup>1</sup> Experimentally observed is a selenido form (with  $\text{X}_{\text{eq}} = \text{Se}$ ) of the cofactor for a nicotinate dehydrogenase system.<sup>37</sup> The different variants of the cofactor are investigated with theoretical methods in Section 3.3.



**Figure 1.3:** General sketch of the coordination sphere of the central molybdenum atom.

### 1.4 Active site

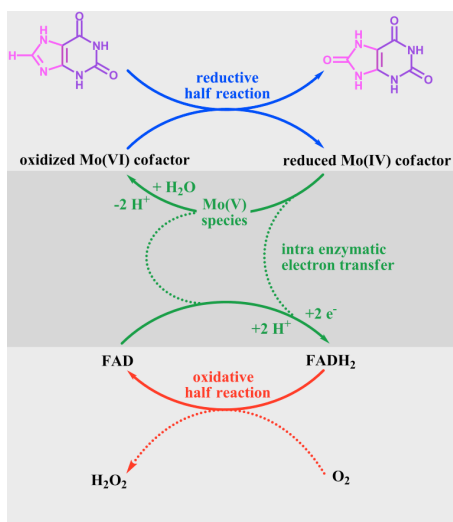
Despite the similarity of the cofactor and the conserved Glu869/1261 (in AOR/XO), the active sites of AOR and XO differ in how they are adapted to their substrates: In the active site of AOR, the hydrophobic nonpolar phenyl rings of Phe425 and Tyr535 interact with the substrate and shield it from the polar solvent, but no significant charge-charge interactions stabilize the substrate complex. By contrast, the interactions of the nonpolar residues within XO (Phe914 behind the substrate, Phe1009 and Ala1078 in front [not shown in Figure 1.4]) are less important compared with the charge-charge interactions of the polar/charged side-chain residues Glu802 and Arg880. However, they are contributing to the proper orientation of the substrate within the binding pocket.



**Figure 1.4:** Active site of AOR (left) and XO with the substrate in an "upside down" orientation (right). In AOR, the interactions of the substrate with enzymatic side chains are dominated by nonpolar interactions, whereas in XO charge-charge interactions prevail.

### 1.5 Catalytic cycle

The full catalytic cycle of XO consists of the reductive half-cycle, an intra-enzymatic electron transfer, and the oxidative half-cycle, in which reduced flavin adenine dinucleotide ( $\text{FADH}_2$ ) is re-oxidized to FAD, see Figure 1.5.

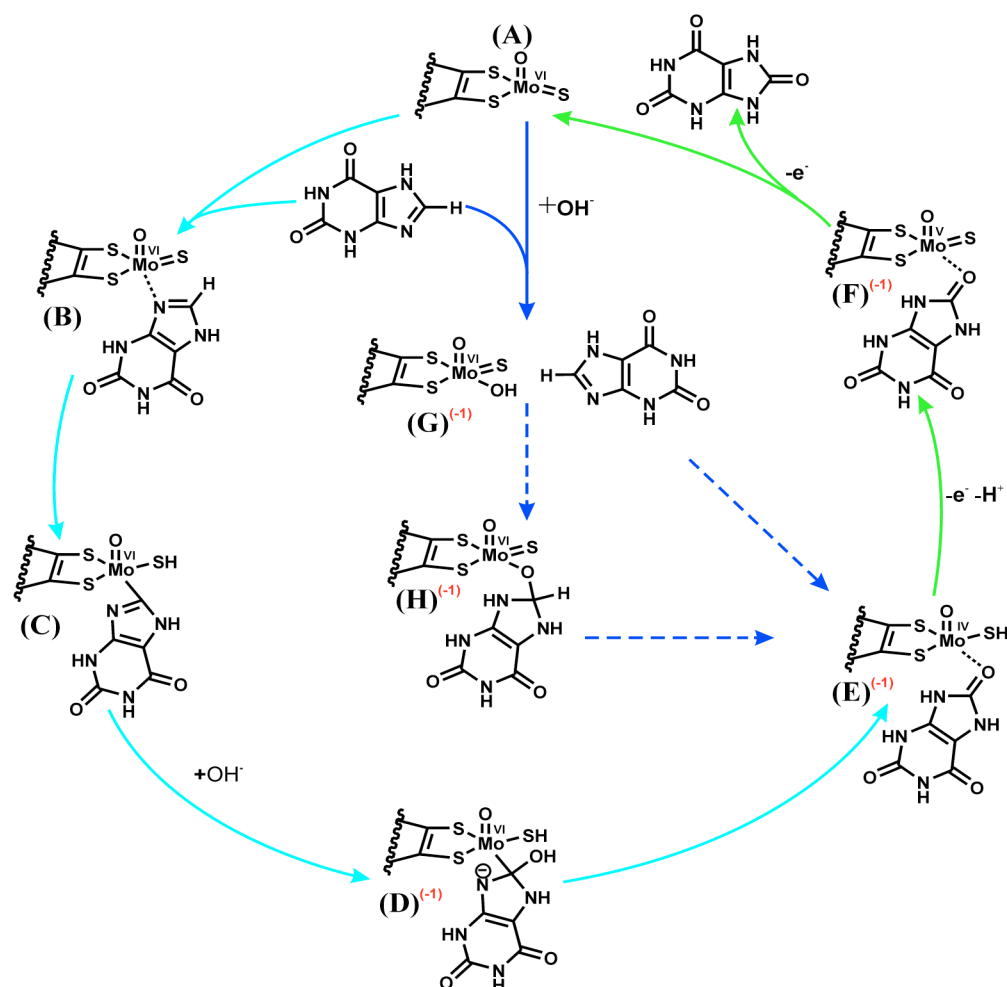


**Figure 1.5:** Complete catalytic cycle of XO.

Within this study, we will focus on the reductive half-cycle, i.e. the conversion of xanthine to uric acid. For this, two substantially different types of activation have been postulated, as shown in Figure 1.6.

For the first kind (light blue) a direct interaction of the substrate's nitrogen atom with the central molybdenum atom (B) has been claimed<sup>38</sup> prior to a concerted formation of a Mo-C bond and transfer of a hydrogen to the sulfido group (in C). The activated substrate can then be attacked by a hydroxy group either from a deprotonated water molecule or from a hydroxy group that was attached to the molybdenum center prior to formation of the carbon-bound tetrahedral intermediate (D). This species reacts to the reduced Mo(IV) species (in E). By transferring one electron from the molybdenum center via the iron-sulfur cluster to the FAD cofactor (see also Figure 1.5) and loss of one proton, the presumed ESR-active intermediate (F) is formed. Depending on the investigation, a more pronounced Mo-C or Mo-O interaction is claimed.<sup>38,39</sup> A significant Mo-C interaction is supported by recent X-ray data, showing a similar coordination of an ethylene glycol and glycerol inhibitor in AOR.<sup>1</sup>

The second pathway (dark blue) is favored in the current literature,<sup>39-43</sup> but controversial in its details. Herein, the active form of the cofactor bears the catalytically active hydroxy group attacking the substrate (G). Beside the question whether there is a tetrahedral intermediate (H) or whether the reaction follows a one-step mechanism to form the product-bound form (E), the orientation of the substrate within the binding pocket and its tautomeric form in the active pathway are unknown.



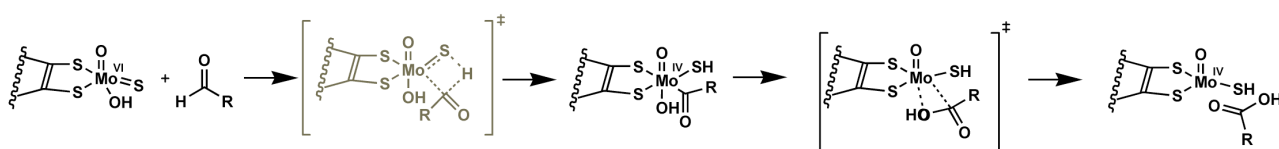
**Figure 1.6:** Reductive half-cycle of XO with different suggested pathways, see text. One representation of possible protonation states and substrate orientations was selected for this figure. The overall charge of the depicted species is given in parentheses, colored in red.

Similar reaction mechanisms and intermediates have been discussed for AOR and can easily be obtained by simply replacing the  $H-C=N$  moiety of xanthine by a  $H-C=O$  moiety of an aldehyde in Figure 1.6. It should be pointed out that although the direct Mo-C interaction seems to be unlikely in XO, especially due to the very specific binding of the substrate, this kind of interaction may be possible in AOR. In particular, if the oxo form is the active form of the cofactor, there may be enough space to extend the coordination number of molybdenum from five to six and obtain an additional interaction with the substrate, as suggested on the basis of recent X-ray data.<sup>1</sup>

## 1.6 Theoretical investigations

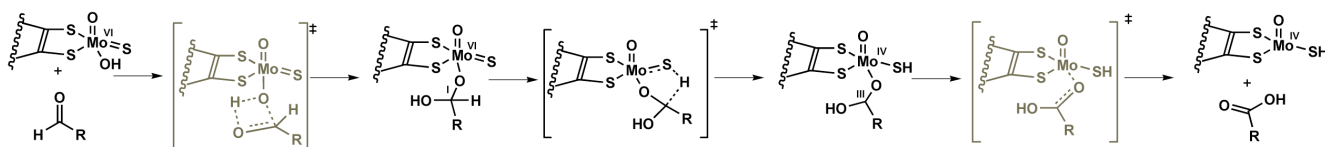
The two enzymatic systems considered have not been studied before by others using QM/MM methods, but there are some QM studies on simple model systems in the literature.

First mechanistic details were calculated at the B3LYP/Lan12DZ level of theory in 1997.<sup>27</sup> In the suggested mechanism, the substrate transfers its hydrogen atom to the sulfido group when binding to the molybdenum atom in a first step. The barrier of the subsequent C-O bond formation between the hydroxy and the carbonyl group that is accompanied by the Mo(VI)  $\rightarrow$  Mo(IV) transition was found to be 8.4 kcal mol<sup>-1</sup>, whereas the barrier of the first step was not reported. As in the light-blue colored pathway in Figure 1.6, a direct Mo-C interaction is assumed in this pathway (see Figure 1.7).



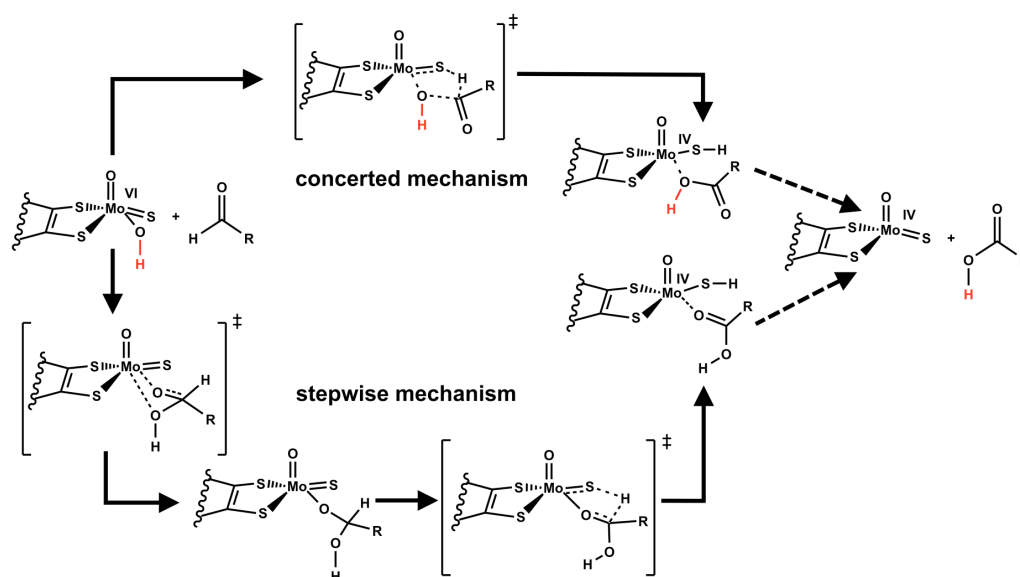
**Figure 1.7:** Reaction sequence reported by Bray et al.<sup>27</sup> The gray colored transition state was not localized.

The pathways colored in dark-blue in Figure 1.6 show mechanisms without direct Mo-C interaction, retaining the original coordination number of five. This can be achieved by either a stepwise or a one-step mechanism. The former was addressed by Voityuk et al.<sup>29</sup> using the BP86 functional, who proposed a two-step reaction with formation of a tetrahedral intermediate and subsequent hydride transfer. However, only results for the hydride transfer were reported, see Figure 1.8.



**Figure 1.8:** Reaction sequence reported by Voityuk et al.<sup>29</sup> The gray colored transition states were not localized.

The one-step mechanism was first considered by Illich et al.,<sup>30</sup> but due to the chosen level of theory (UMP2), the computed barrier is much too high and not comparable to the previous results. The most complete study at the DFT(B3LYP) level of theory was reported by Zhang and Wu,<sup>31</sup> who gave a direct comparison between the concerted mechanism and the stepwise pathway via a tetrahedral intermediate (see Figure 1.9) using formaldehyde as substrate. Their computed barriers were 11.9 kcal mol<sup>-1</sup> for the concerted pathway, and 17.8 and 5.4 kcal mol<sup>-1</sup> for the stepwise mechanism. However, there is experimental evidence against the rate limiting step of the presented stepwise mechanism in the case of XO.<sup>20</sup>



**Figure 1.9:** Mechanisms considered in previous theoretical work. (a) Protonated cofactor: Concerted and stepwise pathway reaction sequences were reported.<sup>29-31</sup> (b) Deprotonated cofactor (without proton shown in red): only a concerted and no stepwise mechanism was found.<sup>32</sup>

Barriers obtained by highly correlated *ab initio* methods were reported by Amano et al.,<sup>32</sup> using the less active formamide as substrate and taking a deprotonated cofactor into account. However, the computed overall barriers of 35-42 kcal mol<sup>-1</sup> are prohibitively high. The only mechanistic study using xanthine as substrate suggested a concerted mechanism,<sup>34</sup> without considering a stepwise mechanism.

All these QM model calculations are limited in scope. They suffer either from incompleteness of the calculated reaction paths,<sup>27,29</sup> from unrealistic high barriers,<sup>30,32</sup> or from inclusion of mechanistic steps that are not compatible with the available experimental results.<sup>20</sup> In addition, all investigations are limited to QM model systems that consist of only a part of the cofactor and a (model) substrate. They do not include any active-site residue, and therefore cannot account for the steric and electronic influence of the enzymatic environment which proves to be of crucial importance, as we will show in this work.





## 2 Theoretical Methods

## 2.1 The QM/MM approach

*QM treatments:*

Most of the presented QM/MM calculations employ density functional theory (DFT). Its foundation is the Hohenberg-Kohn theorem,<sup>44</sup> which states that the electronic energy of a system in its ground state is determined by the electron density:

$$\begin{aligned} E_{elec}[\rho] &= T[\rho] + E_{ne}[\rho] + E_{ee}[\rho] \\ &= T[\rho] + \int V_n(r) \rho(r) dr + E_{ee}[\rho] \end{aligned} \quad (2.1)$$

Here,  $T[\rho]$  describes the kinetic energy,  $E_{ne}[\rho]$  the nuclei-electron attraction, and  $E_{ee}[\rho]$  the electron-electron interaction.  $E_{ne}[\rho]$  can be obtained from the interaction of the external potential  $V_n$  from the nuclei with the electron density  $\rho(r)$ .  $E_{ee}$  contains the classical Coulomb interactions  $J[\rho]$ :

$$J[\rho] = \frac{1}{2} \iint \frac{1}{r-r'} \rho(r) \rho(r') dr dr' \quad (2.2)$$

Early functional forms to obtain  $T[\rho]$  performed poorly. This was overcome in modern DFT with the introduction of orbitals  $\varphi_i$  to represent the density (see 2.3) by Kohn and Sham.<sup>45</sup> They are used to approximate the kinetic energy  $T[\rho]$  by  $T_S[\rho(\varphi_i)]$ .

$$\rho(\varphi_i) = \sum_i |\varphi_i|^2 \quad (2.3)$$

$$T_S[\rho(\varphi_i)] = \sum_i \langle \varphi_i | -\frac{1}{2} \nabla^2 | \varphi_i \rangle \quad (2.4)$$

The total electronic energy can therefore be expressed as in equation 2.5:

$$\begin{aligned} E_{elec}[\rho] &= T_S[\rho(\varphi_i)] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \\ &= \sum_i \langle \varphi_i | -\frac{1}{2} \nabla^2 | \varphi_i \rangle + \int V_n(r) \rho(r) dr + \frac{1}{2} \iint \frac{1}{r-r'} \rho(r) \rho(r') dr dr' + E_{xc}[\rho] \end{aligned} \quad (2.5)$$

From equations 2.1 and 2.5, the exchange-correlation energy  $E_{xc}[\rho]$  is given as

$$E_{xc}[\rho] = T[\rho] - T_S[\rho(\varphi_i)] + E_{ee}[\rho] - J[\rho] \quad (2.6)$$

Under the constraint of orthonormality of the orbitals  $\varphi_i$ , one finally has to solve the Kohn-Sham equations  $\hat{H}_{KS} \varphi_i = \varepsilon_i \varphi_i$ , which are analogous to the Fock equations in Hartree-Fock theory, in an iterative manner. The Kohn-Sham Hamiltonian  $\hat{H}_{KS}$  is given as:

$$\hat{H}_{KS} = -\frac{1}{2} \nabla^2 + V_{ne} + \int \frac{\rho(r')}{r-r'} dr' + V_{xc} \quad \text{with } V_{xc} = \frac{\partial E_{xc}[\rho]}{\partial \rho} \quad (2.7)$$

Herein, the first term is the operator of the kinetic energy, the second and third terms represent the classical nuclei-electron and electron-electron interaction, and the last one the exchange-correlation potential, covering all non-classical contributions and corrections.

The exchange-correlation functional  $E_{xc}$  can take different forms which generally have a complicated form that does not allow simple interpretation of its components. It is customary to split the functional into a pure exchange part  $E_x$  and a pure correlation part  $E_c$ . Popular combinations of exchange and correlation functionals used within this thesis are BP86<sup>46-50</sup> and the generally more accurate B3LYP functional.<sup>46-48,51-53</sup> B3LYP is a so called hybrid functional involving the admixture of a certain amount of Fock exchange to improve the functional performance for many molecular properties.<sup>54</sup> This better performance is achieved at higher computational cost compared to pure functionals (without the inclusion of Hartree-Fock exchange). When compared to experiment, B3LYP calculations show a remarkable accuracy in the simulation of biologically relevant metal centers<sup>55-59</sup> and enzymatic reactions.<sup>60,61</sup> For example, in cytochrome P450, B3LYP predicts the structures and energetics of the species in the catalytic cycle quite accurately and correctly assigns the relative order of the various spin states.<sup>61-63</sup>

The use of local correlated ab initio methods such as LCCSD(T0) in QM/MM studies can improve the computed DFT/MM barriers for enzymatic reactions up to chemical accuracy.<sup>64,65</sup> This method has been applied in paper II to the demanding case of a metallo-enzyme (in cooperation with the Werner group in Stuttgart), and will therefore be introduced briefly.

Starting from the Hartree-Fock wave function  $\Psi_{HF}$ , the coupled cluster wave function is obtained from an exponential ansatz, including the cluster operator  $\hat{T}$ :

$$\Psi_{CC} = e^{\hat{T}} \Psi_{HF} \quad \text{with} \quad \hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots \quad (2.8)$$

For CCSD,  $\hat{T}$  simplifies to  $\hat{T} = \hat{T}_1 + \hat{T}_2$  with the individual terms defined as:

$$\hat{T}_1 \Psi_{HF} = \sum_i^{occ} \sum_a^{virt} t_i^a \Psi_i^a \quad \text{and} \quad \hat{T}_2 \Psi_{HF} = \sum_{i<j}^{occ} \sum_{a<b}^{virt} t_{ij}^{ab} \Psi_{ij}^{ab} \quad (2.9)$$

$\hat{T}_1$  ( $\hat{T}_2$ ) generates all singly (doubly) excited states  $\Psi_i^a$  ( $\Psi_{ij}^{ab}$ ) and the corresponding expansion coefficients  $t_i^a$  ( $t_{ij}^{ab}$ ). Since there are many less important excitations in large systems, one can apply two local approximations: First, in the *domain approximation*,<sup>66,67</sup> the excitations for a given localized molecular orbital (LMO) are restricted to a subset of spatially close projected atomic orbitals (PAOs), grouped together according to the centers of the original AOs. Secondly, in the *pair approximation*, the orbital pairs are classified according to their distance, and the pair correlation energies are obtained by different methods: For very distant pairs they can be neglected; for medium distances, lower-level perturbational correlation methods are assumed to be sufficiently accurate, and only for close pairs, the CCSD(T0) method is used.

*MM treatments:*

QM methods provide the electronic wave function and energy for given nuclear positions (Born-Oppenheimer approximation). Hence, they can be used to compute the potential-energy surface, i.e., the energy as a function of the nuclear coordinates. In molecular-mechanics methods, this “ab initio” potential is replaced by an empirical potential, in most force fields at an atomistic level. The potential energy of a diatomic molecule like O<sub>2</sub> can be approximated by a quadratic potential, see equation 2.8:

$$E_{bond} = k_{bond}(r - r_0)^2 \quad (2.8)$$

$k_{bond}$  is the force constant of the bond and  $(r - r_0)$  is the deviation of the actual bond length from the predefined equilibrium value. For more complicated molecules, the potential contains several terms, which are given for the CHARMM force field,<sup>68-70</sup> used throughout this study, in equation 2.9.

$$E_{MM} = \sum_{bonds} E_{bond} + \sum_{angles} E_{angle} + \sum_{dihedrals} E_{dihedral} + \sum_{impropers} E_{improper} + \sum_{nonbonded} E_{vdW} + \sum_{charges} E_{electrostatic} \quad (2.9)$$

The van der Waals (vdW) interactions (normally described with the Lennard-Jones potential) and the electrostatic interactions (described by the Coulomb potential) can be categorized as nonbonded interactions, whereas the bonded interactions include energy terms for all bonds, angles, dihedrals and impropers (dihedral terms describing the deviation from planarity, e.g., in aromatic rings). All four bonded terms represent similar kinds of energy penalties due to a deviation of the actual value from an equilibrium value. The latter is predefined and cannot change during a reaction, and therefore chemical reactions cannot be described with a standard force field.

*QM/MM methods:*

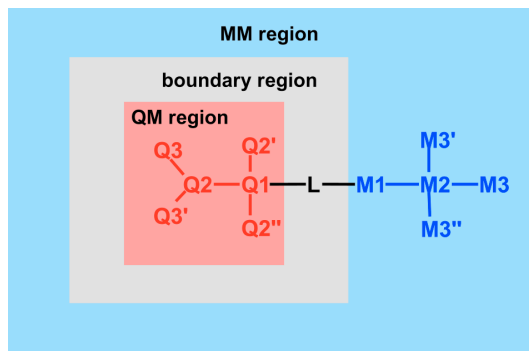
If interested in the reaction mechanism of large systems, one may therefore turn to a QM/MM hybrid approach. We give a very condensed introduction, referring the reader to a recent review by Senn and Thiel<sup>71</sup> and references therein for a detailed introduction to the QM/MM methodology.

The energy of a system consisting of two components, a QM and a MM region, can be expressed as a sum, see equation 2.10.

$$E_{QM/MM} = E_{QM} + E_{MM} + E_{QM-MM} \quad (2.10)$$

Here  $E_{QM}$  and  $E_{MM}$  are the energies of the isolated subsystems, and  $E_{QM-MM}$  results from the interactions between both regions. There are different approaches to handle the QM-MM interactions; the *electrostatic embedding* method<sup>72</sup> was used exclusively within this work. Within this method, the MM point charges are incorporated as one-electron terms in the QM Hamiltonian, such that the electronic structure of the QM region can adapt to changes in the charge distribution of the environment and is thus automatically polarized.

The non-bonded interactions between the QM and the MM part consist of the electrostatic interactions that are included by incorporation of the point charges into the QM Hamiltonian (see above) and the vdW terms. The repulsive part of the latter describe steric effects and are able to prevent clashes between atoms with charges of opposite sign, which can be handled easily within the force field code.



**Figure 2.1:** QM/MM scheme and nomenclature.

In cases where one or several bonds exist between the QM and MM regions, one must augment the QM/MM approach with an adequate treatment of the boundary region. Within this work, we use the link-atom approach,<sup>73</sup> in which an unphysical atom or group (normally a hydrogen atom) is added to unsaturated valencies to saturate the QM region. By adding link atoms, we are facing a new problem, as the link atom is normally positioned very close to the M1 point charge, which would lead to an overpolarization of the wavefunction. To prevent this, the charge from M1 is distributed evenly over the M2 atoms. The dipole created by shifting the charges is compensated for by a pair of point charges placed near each M2 atom, which generate dipoles of the same magnitude and opposite direction, such that the dipole moment along the M1-M2 bond vector reproduces the dipole moment of the original group. This method is called the *charge shift scheme*.<sup>74,75</sup>

To describe the Q1-M1 interaction, the MM bond-stretch potential is applied to that bond. The introduction of each link atom introduces three artificial degrees of freedom, which is undesirable in optimizations or MD simulations, and which are removed using a constraint that forces the link atom to remain at a certain position on the line connecting Q1 and M1. Link atom forces are distributed over on the frontier atoms (Q1, M1) by application of the chain rule. Schematically:

$$\frac{\partial \tilde{E}}{\partial r_{M1}} = \frac{\partial E}{\partial r_{M1}} + \frac{\partial E}{\partial r_L} \cdot \frac{\partial r_L}{\partial r_{M1}} \quad (2.11)$$

In addition to the Q1-M1 bond, all angle, dihedral angle, and improper angle terms contribute to the QM/MM energy that include at least one QM atom and one MM atom. The exceptions to this general rule are the M1-Q1-Q2 angle and the dihedrals of the form M1-Q1-Q2-Q3.

## 2.2 Free-energy perturbation

Reaction-path calculations based on QM/MM geometry optimizations provide potential energy profiles, and hence the possibility to discriminate between competing reaction mechanisms. It is, however, the free-energy barrier that determines the reaction rate, and which we therefore aim to obtain. There are several sampling-based free-energy calculation methods available, such as umbrella sampling<sup>76</sup> and thermodynamic integration,<sup>77</sup> whereby the latter can be regarded as limiting case of umbrella integration.<sup>78</sup> Since efficiently implemented within the ChemShell package,<sup>79</sup> the computationally less demanding free-energy perturbation (FEP) method<sup>80</sup> was used to investigate AOR within this thesis.

The FEP method is often attributed to Zwanzig<sup>81</sup> who derived the free-energy difference of two states  $a$  and  $b$  as an ensemble average:

$$\Delta A^{a \rightarrow b} = -k_B T \ln \left\langle \exp \left\{ -\Delta E_{pert}^{a \rightarrow b} / (k_B T) \right\} \right\rangle^a \quad (2.12)$$

$\Delta E_{pert}^{a \rightarrow b}$  denotes the energy difference between the actual state  $a$  and state  $a$  perturbed with state  $b$ ,  $k_B$  is the Boltzmann constant and  $T$  the absolute temperature in Kelvin. Zhang et al. introduced the FEP method in the framework of QM/MM methodology.<sup>80</sup> Here we present a brief overview adopting the nomenclature of Senn et al.<sup>82</sup> More details can be found in a recent review.<sup>83</sup>

The QM/MM FEP method focuses on the contributions of the environmental part ( $\Delta A_{env}$ ), and not on those of the QM region itself. The contribution of the latter ( $\Delta A_{QM}$ ) is added later, based on the harmonic-oscillator approximation, such that the total free energy is obtained as sum of the two contributions:

$$\Delta A_{FEP} = \Delta A_{QM} + \Delta A_{env} \quad (2.13)$$

In most cases, the free-energy difference to be calculated differs by more than approximately  $2k_B T$ , and therefore the computation between two states has to be split into several intermediate steps. The intermediate structures are normally obtained from a reaction path, optimized either with respect to an appropriately defined reaction coordinate or by applying the nudged elastic band method.<sup>84</sup> The general idea in QM/MM FEP is now to obtain the complete free energy of a process as sum over the sampled contributions of two sequential structures,  $\Delta A_{env}^{i \rightarrow i+1}$ , respectively.

$$\Delta A_{env} = \sum_{i=a}^{b-1} \Delta A_{env}^{i \rightarrow i+1} \quad (2.14)$$

The free-energy difference between state  $i$  and the subsequent intermediate  $i+1$  ( $\Delta A_{env}^{i \rightarrow i+1}$ ) is obtained by sampling at the MM level the energy difference induced by the perturbation of the QM region of intermediate  $i$  by the subsequent intermediate  $i+1$ .

$$\Delta A_{env}^{i \rightarrow i+1} = -k_B T \ln \left\langle \exp \left\{ -\Delta E_{pert}^{i \rightarrow i+1} / (k_B T) \right\} \right\rangle_{MM}^i \quad (2.15)$$

As the free-energy contribution of the QM part is treated separately,  $\Delta E_{pert}^{i \rightarrow i+1}$  simplifies to:

$$\Delta E_{pert}^{i \rightarrow i+1} = E_{QM-MM} \left( R_{QM}^{i+1}, R_{MM}(t) \right) - E_{QM-MM} \left( R_{QM}^i, R_{MM}(t) \right) \quad (2.16)$$

The contributions of the QM part (i.e. the QM atoms and all MM atoms sharing a bond with them) are contained in  $\Delta A_{QM}$ :

$$\Delta A_{QM} = \Delta E_{QM}^{\rho(QM/MM)} + \Delta E_{QM}^{ZPE} + \Delta U_{QM}^{th} - T \Delta S_{QM} \quad (2.17)$$

The first summand represents the change in energy of the QM region obtained by a gas phase calculation (i.e. without surrounding point charges), but using the density  $\rho(QM/MM)$  from the QM/MM calculation (with the surrounding point charges). The latter three terms are the zero-point vibrational energy  $\Delta E_{QM}^{ZPE}$ , the thermal contributions to the internal energy  $\Delta U_{QM}^{th}$  and the entropic contributions  $-T \Delta S_{QM}$ , which are determined at stationary points by applying the harmonic-oscillator approximation.

---

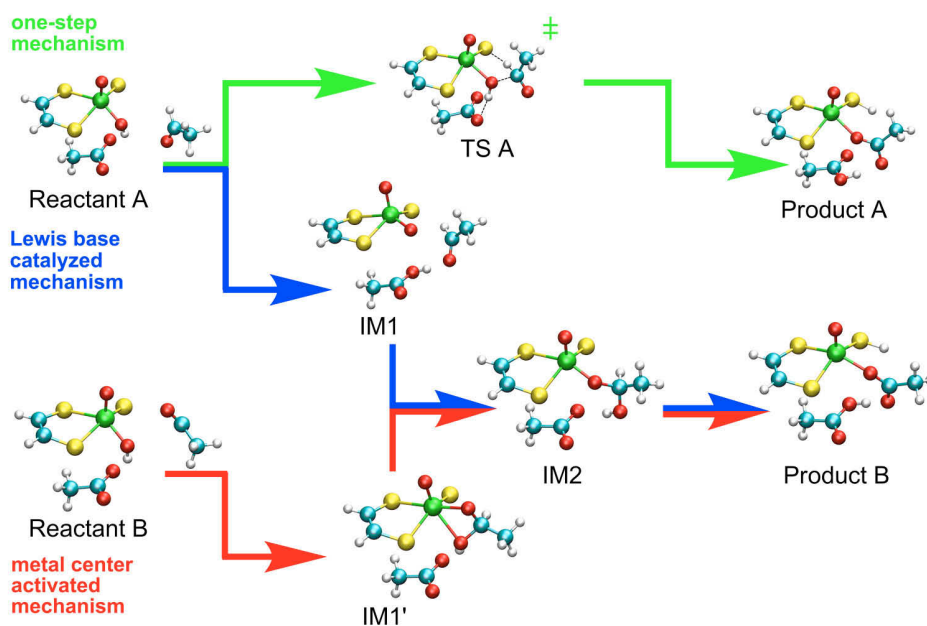


## 3 Results and Discussion

In this section, the results published in the attached papers are briefly reviewed. Sections 3.1 and 3.2 summarize the mechanistic results obtained from QM/MM calculations on AOR and XO, respectively. In Section 3.3, unpublished results for a modified cofactor (oxo-, sulfido- and selenido-form) in AOR are presented and compared to previously reported results for XO.

### 3.1 Mechanism of aldehyde oxidoreductase

The reductive half-reaction of AOR was investigated first using QM/MM methods. Five possible pathways were explored concerning the binding mode of acetaldehyde and the catalytic effect of the nearby glutamic acid (Glu869), taking both possible protonation states of the latter into account. The setup involving a protonated Glu869 precludes the activation of the cofactor by Glu869 acting as a Lewis base, and the computed effective barrier for this pathway is thus prohibitively high. For the setup involving a deprotonated Glu869, we obtain three distinctive reaction mechanisms, see Figure 3.1. A variant of the kinetically favored Lewis base catalyzed mechanism was considered by including one additional water molecule between Glu869 and acetaldehyde in an enlarged QM region. This water molecule may help to position the substrate and may move to the Mo center after the oxidation of the substrate to act as oxygen source for the next turnover. However, this variant did not improve the energetics (compared to the system without the extra water molecule) and will thus be disregarded in the following (for details see paper I<sup>85</sup>). We thus concentrate on the three reaction mechanisms shown in Figure 3.1.



**Figure 3.1:** Scheme of three competing reaction mechanisms.

For each of the three reaction mechanisms, there are three elementary transformations involved: The hydrogen atom of the molybdenum-bound hydroxy group is transferred to Glu869, a bond is

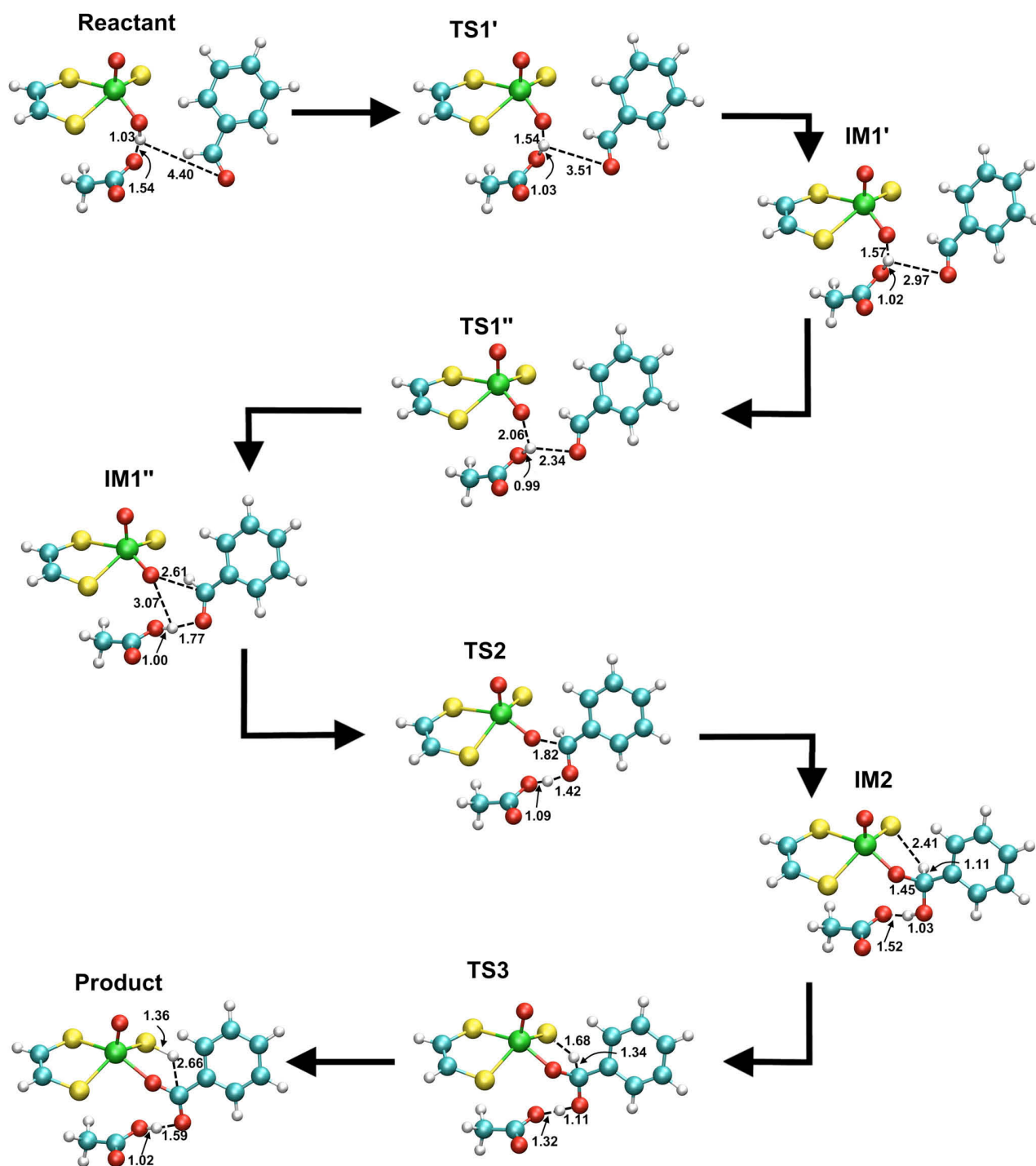
formed between the equatorial oxygen atom and the carbon atom of the substrate's carbonyl group, and a hydrogen atom (attached to the attacked carbon atom) has to be transferred to the equatorial sulfido ligand.

In the one-step mechanism, the reaction proceeds through a five membered transition state structure, see Figure 3.1. Due to the substrate orientation, this one-step pathway is not accessible to bulkier substrates, such as benzaldehyde. The species attacking the substrate contains a hydroxy group in the transition state, and hence there is no activating effect of the nearby Glu869, but only a stabilization of the product complex.

In the Lewis base activated mechanism, Glu869 participates actively in the mechanism and deprotonates the hydroxide group of the cofactor to form IM1, see Figure 3.1. This proton transfer is essential for the high activity of the enzyme towards the substrate, and the benefits from this are two-fold: First, the oxygen atom of the former hydroxy group becomes more nucleophilic. Second, the newly formed H-bond from the proton at Glu869 perturbs the carbonyl group of the substrate and induces an electron flow to the carbonyl oxygen atom, which strengthens the electrophilicity of the target carbon atom in the substrate. In the next step, the thus activated cofactor and substrate form a C-O bond yielding a tetrahedral intermediate (IM2). Judging from our QM/MM data, the catalytic effect of Glu869 mainly lowers the barrier of this nucleophilic attack so that the subsequent hydride transfer becomes the rate-determining step in the reductive half-reaction. More precisely, this step includes a coupled proton-hydride transfer in AOR, see Figures 3.1 and 3.2.

The metal center activated mechanism differs from the previous one how the intermediate IM2 is reached: The substrate first forms a carbon-oxygen bond by coordinating its carbonyl group to the metal center and at the same time forming a bidentate complex in IM1'. This rearranges to IM2, such that the final step is identical with the one of the Lewis base catalyzed mechanism.

Comparison of these three pathways confirms that Glu869 plays a crucial role as a Lewis base promoting the cofactor to a more active species. These results were confirmed for benzaldehyde as a substrate (as yet unpublished data). For this setup, no other pathway than the Lewis base catalyzed mechanism could be located. The only difference from the situation with acetaldehyde is that deprotonation of the cofactor and formation of a hydrogen bond to the substrate are now two distinctive steps, see Figure 3.2; all other steps are essentially identical.



**Figure 3.2:** Reaction mechanism of AOR with benzaldehyde as substrate.

The calculated energy values for the Lewis base activated pathways of acetaldehyde and benzaldehyde are presented in Tables 3.1 and 3.2. The other results for acetaldehyde are given in the original manuscript (paper I).

**Table 3.1:** Energetics of the Lewis base catalyzed mechanism, using acetaldehyde as substrate. Energy values are given in kcal mol<sup>-1</sup>, barriers relative to the preceding minima are given in parentheses.<sup>a</sup>

	BP86/B2	B3LYP/B1	B3LYP/B2
Reactant	0.0	0.0	0.0
TS1	1.8 (1.8)	4.7 (4.7)	4.6 (4.6)
IM1	-2.3	-2.8	-1.7
TS2	0.2 (2.5)	2.8 (5.6)	4.9 (6.6)
IM2	-8.4	-8.9	-5.6
TS3	-3.6 (4.8)	-0.4 (8.5)	2.1 (7.7)
Product	-13.3	-21.5	-19.0

<sup>a</sup> B1 denotes a combination of the 6-31G\* basis set for H, C, N and O<sup>86,87</sup> and a modified Lanl2DZ basis set for S<sup>88,89</sup> and Mo<sup>90,91</sup>, B2 the def2-TZVP basis.<sup>92,93</sup>

**Table 3.2:** Energetics of the Lewis base catalyzed mechanism, using benzaldehyde as substrate. Energy values are given in kcal mol<sup>-1</sup>, barriers relative to the preceding minima are given in parentheses.<sup>a</sup>

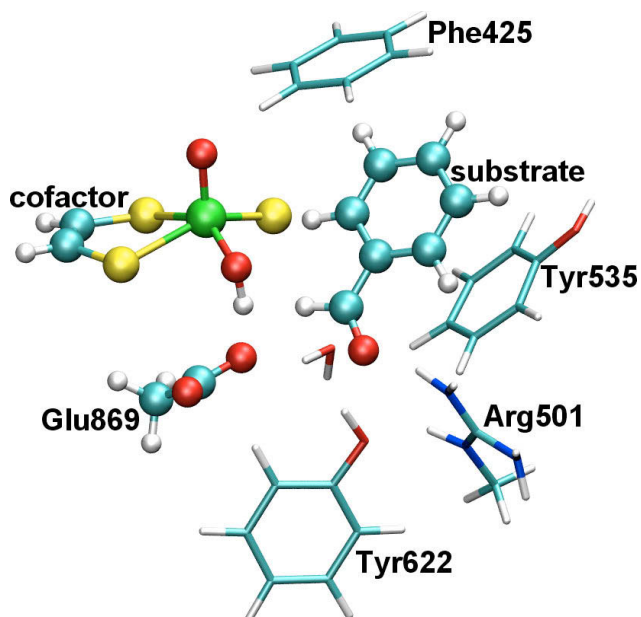
	BP86/B1		B3LYP/B1	
	small QM region	large QM region	small QM region	large QM region
Reactant	0.0	0.0	0.0	0.0
TS1'	5.7 (5.7)	1.8 (1.8)	5.8 (5.8)	3.1 (3.1)
IM1'	1.9	-0.5	1.8	0.3
TS1''	4.9 (3.0)	2.7 (3.2)	2.4 (0.6)	1.5 (1.2)
IM1''	1.2	0.1	-0.1	-1.1
TS2	1.6 (0.4)	2.0 (1.9)	6.0 (5.9)	4.5 (5.6)
IM2	-1.9	-4.8	-0.3	-2.6
TS3	2.8 (4.7)	-0.6 (4.2)	6.9 (7.2)	3.9 (6.5)
Product	-10.3	-13.2	-19.6	-21.9

<sup>a</sup> For basis sets definition, see footnote in Table 3.1; for the definition of the small/large QM region, see Figure 3.3.

In the following section, the B3LYP/B1 results from Tables 3.1 and 3.2 are compared with each

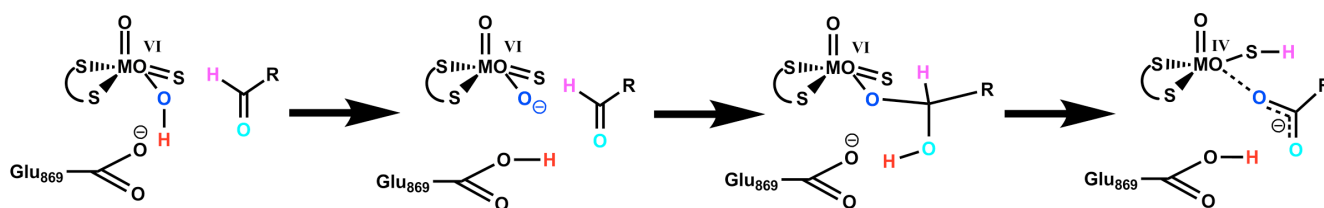
other. For Table 3.2, we will refer to the small QM region, which contains part of the cofactor, Glu869, and the substrate, and therefore (aside from the different substrates) the same residues as the system in paper I. The influence of the size of the QM region was investigated by reoptimizing all stationary points with a larger QM region, as depicted in Figure 3.3. This causes negligible differences for the rate determining barrier ( $0.7 \text{ kcal mol}^{-1}$ ), the main impact is the lowering of the initial deprotonation barrier by  $2.7 \text{ kcal mol}^{-1}$ .

The process of deprotonating the cofactor and forming a hydrogen bond to the substrate is split into two consecutive steps for benzaldehyde rather than being a single step as for acetaldehyde. However, the deprotonation step itself (described by TS1 for acetaldehyde and TS1' for benzaldehyde) has similar barriers for the two substrates ( $4.6 \text{ vs. } 5.8 \text{ kcal mol}^{-1}$ ). The additional barrier (TS1'') in the case of benzaldehyde is very low ( $0.6 \text{ kcal mol}^{-1}$ ) and does not affect the general shape of the energy profile. The nucleophilic attack is slightly less facile for acetaldehyde ( $6.6 \text{ vs. } 5.9 \text{ kcal mol}^{-1}$ ), as is the hydride transfer ( $8.5 \text{ vs. } 7.9 \text{ kcal mol}^{-1}$ ). The latter represents the rate-limiting step for benzaldehyde, consistent with the reaction mechanism from our previous study for acetaldehyde. Therefore, the activity for acetaldehyde should be slightly lower than that for benzaldehyde, which is supported by the published  $k_d$  values ( $1.14 \text{ vs. } 1.43 \text{ s}^{-1}$ ).<sup>23</sup>



**Figure 3.3:** Small (balls & sticks) and large (including licorice) QM region of AOR with benzaldehyde.

Figure 3.4 summarizes the structural features of the reaction mechanism for the oxidation of aldehydes by AOR including the adopted protonation states. Different substrates may introduce additional intermediates (as shown for benzaldehyde).



**Figure 3.4:** Scheme of the energetically preferred reaction mechanism of AOR.

All energies presented above are potential energies ( $\Delta E$ ). However, reaction rates are determined by changes in the free energy ( $\Delta A$ ) of the reacting system. Therefore, FEP theory was employed to generate free-energy profiles at the DFT(B3LYP)/MM level of theory for the most important reaction steps (see above) by sampling along the corresponding reaction paths using molecular dynamics. These results have been combined with accurate ab initio calculations including local correlation treatments [LMP2 and LCCSD(T0)] in combination with augmented triple- and quadruple-zeta basis sets (performed in the group of Prof. Werner at the University of Stuttgart) to derive best estimates of the free energy barriers,  $\Delta^\ddagger A_{FEP}^{best}$ .

At this level, the Lewis base catalyzed mechanism consisting of an initial proton transfer from the cofactor to the Glu869 residue, a subsequent nucleophilic attack that yields a tetrahedral intermediate (IM2), and a final hydride transfer remains the most favored pathway, with the hydride transfer being rate-limiting. The best estimate for the corresponding free-energy barrier is 10.8 kcal mol<sup>-1</sup>. The competing metal center activated pathway has the same final step, but needs to overcome a higher barrier in the initial step on the route to IM2 ( $\Delta^\ddagger A_{FEP}^{best} = 12.1$  kcal mol<sup>-1</sup>), whereas the potential barriers in the Lewis base catalyzed mechanism are significantly lower for the initial proton transfer from Mo-OH to Glu869 and the subsequent nucleophilic attack. The concerted mechanism has the highest free-energy barrier and can be ruled out ( $\Delta^\ddagger A_{FEP}^{best} = 17.1$  kcal mol<sup>-1</sup>). The free-energy barriers are presented in Table 3.3 and discussed in paper II in more detail.

While confirming the qualitative mechanistic scenario proposed previously on the basis of DFT(B3LYP)/MM energy profiles, the ab initio and FEP QM/MM calculations provide corrections to the barriers that are important when aiming at high accuracy. While it is reassuring that standard DFT/MM applications give qualitative mechanistic insights that withstand scrutiny at higher levels, it is clearly desirable to strive for enhanced accuracy both for QM/MM energies and free energies, especially in the case of competing mechanisms with similar barriers. The various corrections from the correlated ab initio QM/MM calculations and from the FEP treatment range up to several kcal mol<sup>-1</sup>, but they tend to compensate each other to some extent in the three reactions studied. The best estimates of the free energy barriers thus differ from the QM(B3LYP/B1)/MM barriers by only

about 3 kcal mol<sup>-1</sup>, and the Lewis base catalyzed mechanism as depicted in Figure 3.4 remains preferred over the concerted and the metal center activated pathways.

**Table 3.3:** Calculated barriers from QM(B3LYP/B1)/MM calculations given in kcal mol<sup>-1</sup>. For details on the nomenclature see paper II.

	one-step mechanism (R-A → P-A)	Lewis base catalyzed mechanism (IM2 → P-B)	metal center activated mechanism (R-B → IM1')
$\Delta^\ddagger E_{QM/MM}$	20.2	8.5	15.4
$\Delta^\ddagger E_{QM}^{ZPE}$	-0.8	-3.1	0.6
$\Delta^\ddagger E_{QM}^{th}$	-1.3	-0.5	-1.0
$-T\Delta^\ddagger S_{QM}$	3.2	1.1	3.4
$\Delta^\ddagger E_{QM/MM}^{corr}$	21.4	6.0	18.4
$\Delta^\ddagger E_{QM}$	9.2	7.4	20.4
$\Delta^\ddagger A_{QM}$	10.3	4.9	23.4
$\Delta^\ddagger A_{env}$	7.8	-0.7	-4.8
$\Delta^\ddagger A_{FEP}$	18.1	4.2	18.7
$\Delta^\ddagger A_{FEP}^{best}$	17.1	10.8	12.1



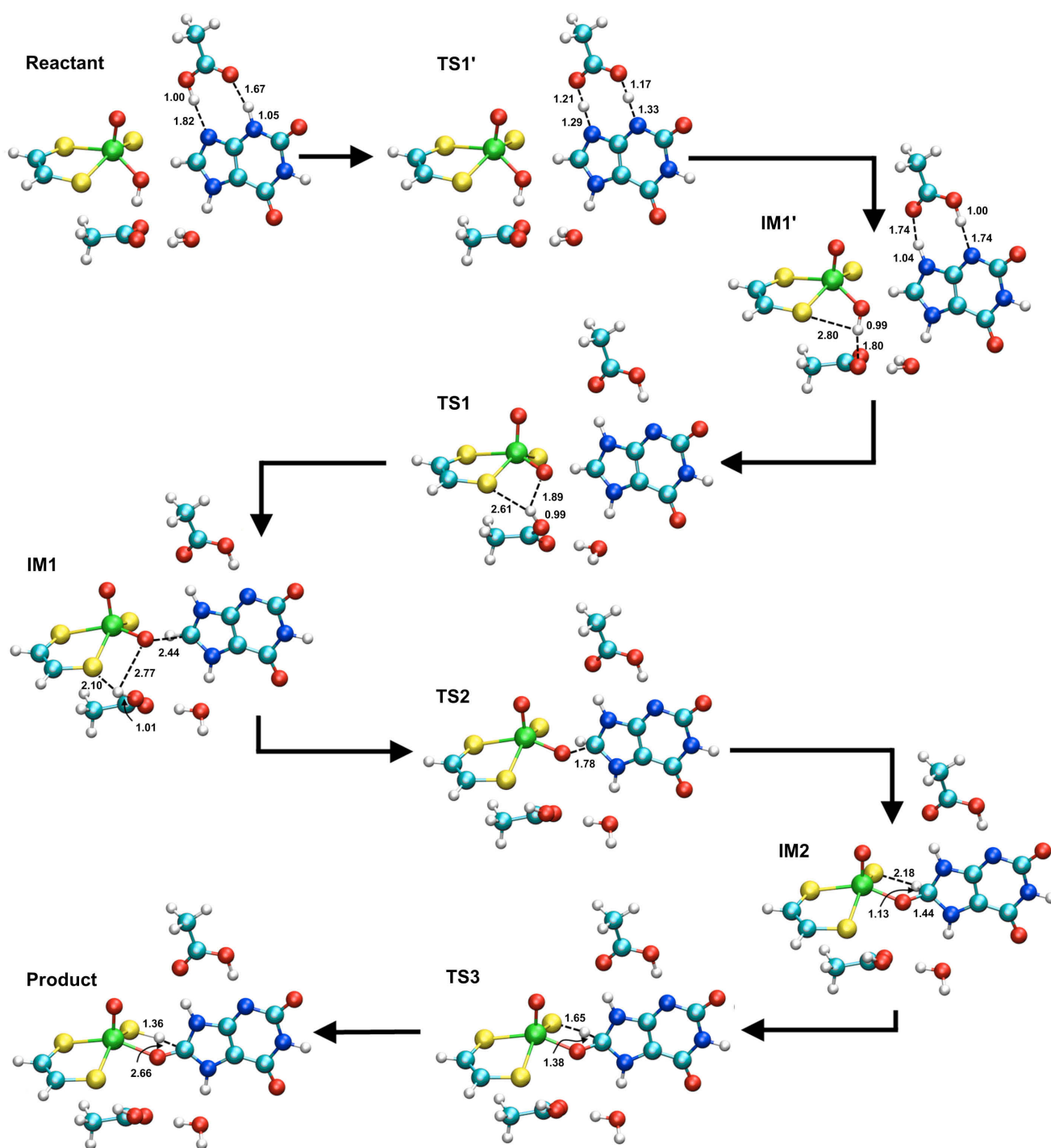
### 3.2 Mechanism of xanthine oxidase

We performed QM/MM calculations to investigate the conversion of xanthine to uric acid in XO. Compared to AOR, more experimental data are available for XO so that a direct comparison of the calculated barriers to the kinetic data is possible.

In contrast to the situation in AOR, where the substrate orientation is well-defined and where only two distinctive protonation states had to be taken into account (protonated and deprotonated Glu869), the setup for XO is quite intricate. In total, seven mechanistic variants with different tautomeric forms of xanthine, different protonation states of the active-site residues, and different substrate orientations were considered. Based on these calculations, Glu802 has to be protonated in the enzymatic environment, in agreement with recently published results for XDH.<sup>94</sup> This protonation state is supported by the deprotonation of the substrate that otherwise occurs in the QM/MM calculations with deprotonated Glu802, which is incompatible with the experimental finding that XO is acting on neutral xanthine.<sup>95</sup> According to the standard protonation state at physiological pH and in agreement with experimental findings,<sup>41,95,96</sup> Arg880 is taken to be protonated in all setups. Finally, Glu1261 is assumed to act as a Lewis base and is therefore deprotonated, in agreement with the results from our AOR study.<sup>85</sup> For full details on the protonation states of the different setups, see paper III.<sup>97</sup>

In addition to the different protonation states taken into account, two essentially different substrate orientations, called "upside" and "upside down", were explored. As the interconversion between them within the binding pocket is prohibited by the enzymatic environment, calculations on large cluster models were performed to establish a common energy scale.

The mechanistic studies started from the product-bound state, which is commonly accepted in the recent literature to be part of the catalytic cycle.<sup>41,43,94,96,98</sup> From the seven setups investigated, four dealt with the "upside" orientation of the substrate and three of them with the "upside down" orientation (see Scheme 2). For each setup, calculated reaction pathways for two or three snapshots obtained by MD simulation were studied. Within this thesis, one example of each orientation is presented, the most reactive setup with "upside" (Figure 3.5) and "upside down" (Figure 3.6) orientation. For the other setups that cover additional substrate tautomers and various proton transfers involving active-site residues, please see paper III.<sup>97</sup>



**Figure 3.5:** Reaction mechanism with the substrate oriented "upside".

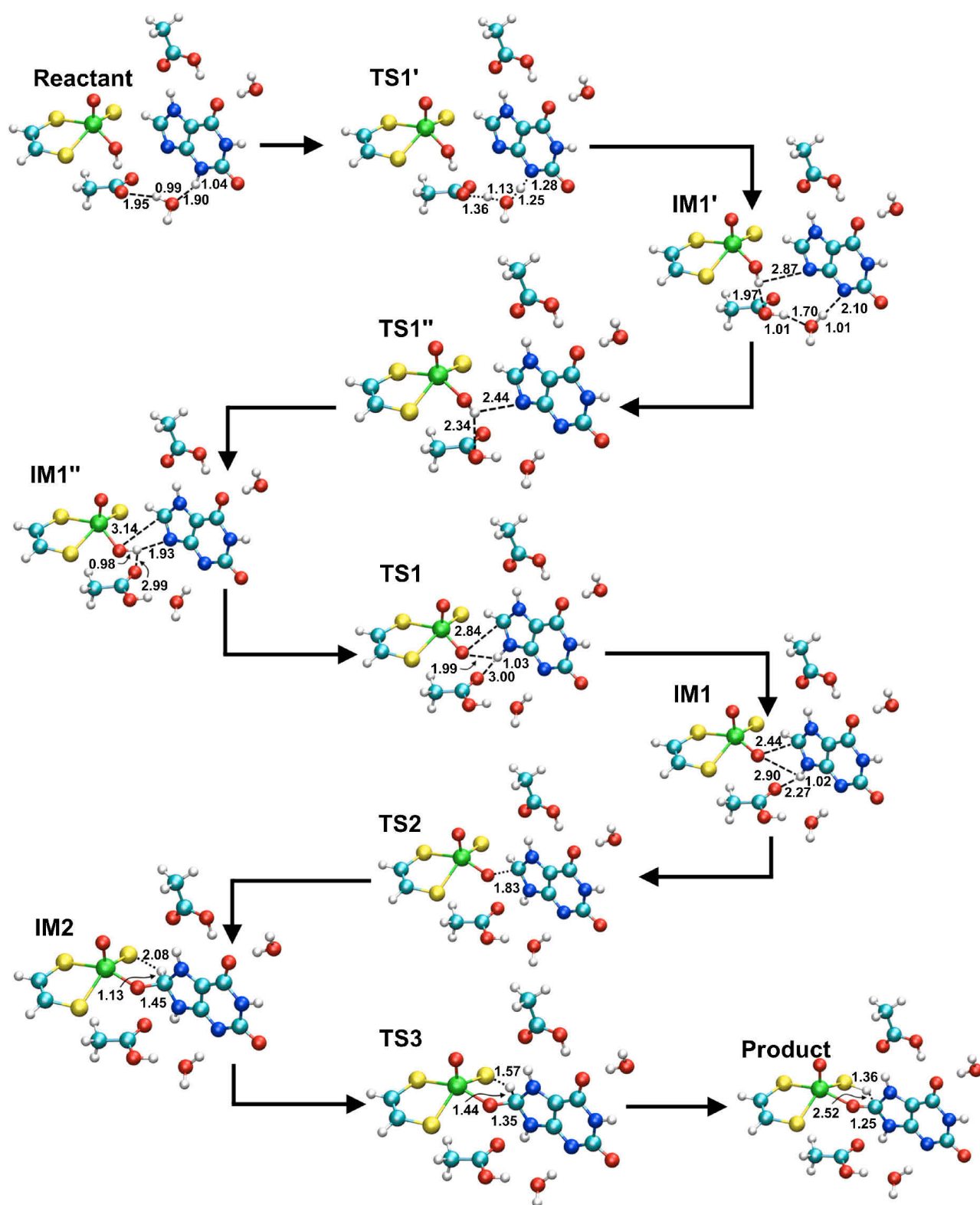
For the "upside" orientation, see Figure 3.5, the reaction starts with a tautomerization of neutral xanthine from its (N1,N3,N7) protonation state (IUPAC numbering convention), which is predominant in aqueous solution,<sup>99</sup> to its (N1,N7,N9) form in IM1'. This step is followed by a second proton transfer (via TS1), this time from the molybdenum-attached OH group to Glu1261. The thus activated oxygen at the cofactor then attacks the substrate to form a tetrahedral intermediate (IM2). In the final step, the H8 atom is transferred to the sulfido-group of the cofactor, and the product becomes planar again.

According to the calculated relative energies, the overall reaction barrier at the B3LYP/MM level is determined by the energy difference between the reactant complex and the transition state TS3 for the hydride transfer. It is calculated to be about 20 kcal mol<sup>-1</sup>, and thus considerably higher than the experimental value of around 15 kcal mol<sup>-1</sup>. As DFT methods tend to underestimate rather than overestimate energy barriers,<sup>100</sup> it is unlikely that this setup represents the enzymatic pathway.

For the "upside down" orientation, see Figure 3.6, the reaction mechanism starts with Glu1261 deprotonating the xanthine at the N3 position (via TS1'), followed by reorientation of the catalytically active OH group via TS1" and a subsequent proton transfer from the cofactor to the N9 atom of xanthine to form IM1. The active tautomer and the activated cofactor are therefore identical in IM1 of both setups. The activated cofactor and substrate then react to form a tetrahedral intermediate (IM2), and a subsequent rate-limiting hydride transfer generates the product. Again, these reaction steps show a great similarity within both presented setups. However, for the favored mechanism with "upside down" orientation of the substrate, a barrier of 13-15 kcal mol<sup>-1</sup> at the B3LYP/MM level was obtained, consistent with the available experimental data.<sup>22,42,101-104</sup>

The barriers presented so far refer to individually prepared setups with QM/MM energies that are not directly comparable among each other. Therefore cluster calculations<sup>105-108</sup> were performed to establish a common energy scale, such that we can compare approximate absolute energy differences for the intermediates of all setups.

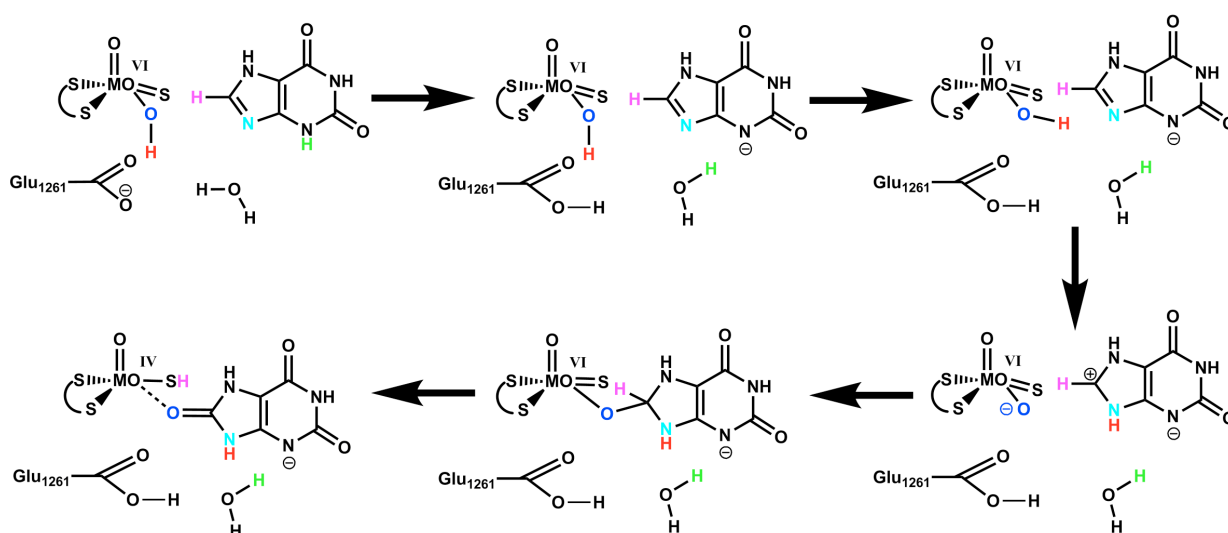
These results confirm the presence of the "upside" conformation of xanthine in the reactant complex, as also observed in a recent X-ray structure.<sup>98</sup> On the other hand, the lowest overall barrier (defined by the TS of the hydride transfer, TS3) is obtained for the "upside down" conformation, which should thus be the catalytically active one. This can be rationalized by the electrostatic influence of the Arg880 residue, whose stabilizing effect on the negatively charged N3 atom of the substrate increases when forming the tetrahedral intermediate and the following rate-limiting transition state.



**Figure 3.6:** Reaction mechanism with the substrate oriented "upside down".

For the favored reaction mechanism, the main function of the Glu1261 residue is to (indirectly) deprotonate the cofactor and to mediate the conversion of the substrate into the reactive tautomeric form. This is achieved by deprotonating the substrate at N3 followed by reorientation of the molybdenum-attached OH group and a second proton transfer that converts cofactor and substrate into their activated forms. The Arg880 residue facilitates substrate binding through stabilizing electrostatic interactions, but its main role during the reaction is to stabilize the substrate in the IM2 intermediate and the preceding (TS2) and following (TS3) transition states, especially by the interaction with the negatively charged N3 atom. According to a recent mutation study,<sup>94</sup> the Glu802 residue seems to have a relatively minor effect on the catalytic activity in XO. This was explored in some more detail by modeling the explicit effect of a Glu802→Gln mutation, see below.

The favored mechanism is quite intricate in its way to activate the substrate by an initial double proton transfer, mediated by Glu1261. In the later stages of the reaction, it shares a number of characteristic features with the most favorable pathway in AOR, see Figures 3.4 and 3.7. However, in contrast to the AOR system, the identification of alternative pathways remains elusive: Due to the much more specific binding of the substrate, its reorientation is hindered and there is no metal center activated or one-step mechanism<sup>85</sup> for XO. Attempts to find such pathways were unsuccessful and led to the mechanisms reported above. Whereas the QM/MM results are broadly compatible with the general mechanistic notions about the reductive half-reaction in AOR and XO, it is obvious that they go significantly beyond these general notions by offering insight into mechanistic details that are hard to unravel by other means, e.g., with regard to substrate orientation and the role of individual active-site residues in XO.



**Figure 3.7:** Scheme of the energetically preferred reaction mechanism.

Additional calculations have been performed to analyze the influence of the recently reported

Glu802→Gln mutation.<sup>94</sup> For both the wild-type and the mutated enzyme, we investigated two competing pathways with the same substrate orientation but different orientations of Glu/Gln802, called setups F and G. The mutant prefers an alternative pathway (setup F) compared to the wild-type enzyme (setup G) for two reasons: First, setup F is disfavored for the wild-type enzyme due to the strong stabilization of IM1, acting as a thermodynamic sink. In the absence of the possible proton transfer, the intermediate IM1 is less stabilized in the mutation, as the stabilizing proton transfer is no longer possible. The overall landscape is therefore flatter and leads to a lower barrier. Secondly, setup G is disfavored in the mutated enzyme as the stabilization of negative charge arising at the O6 position is less strong for Gln than for Glu.

The data in Table 3.4 address the key interactions between the substrate and the active-site residues Arg880 and Glu802, both for the wild-type enzyme (setup G) and the Glu802→Gln mutant with two different orientations of the Gln residue (setups G and F). Arg880 provides significant stabilization when going from the reactant to IM2 in all three systems (see Table 3.4): In setup G, this stabilization of about 40 kcal mol<sup>-1</sup> is caused by shorter distances and higher negative charges of O2 and N3. Additionally, the repulsive interaction of Arg880 with H3, which is part of a water molecule in IM2, is reduced. The stabilization for setup G of the mutant is only slightly diminished (to about 38 kcal mol<sup>-1</sup>), while the individual interactions differ strongly. In contrast, setup F of the mutant gains significantly less stabilization (23 kcal mol<sup>-1</sup>) from its interaction with Arg880. The preference of setup F over setup G on the mutant does therefore not result from a stronger interaction with Arg880. Instead, the dominant factor is the greater stabilization from the interaction of Gln802 with the cofactor/substrate when going from the reactant to IM2, see Table 3.4.

In addition to the effects of mutating a side-chain residue, the reaction of 2-oxo-6-methylpurine as substrate was investigated, taking three possible tautomers into account. The calculated pathways are similar to those found previously with xanthine as substrate, contrary to claims in the literature that these two substrates follow a substantially different reaction path.

Based on the B3LYP/MM energies for the largest QM region, rate-limiting barriers of 25.8, 7.3, and 11.3 kcal mol<sup>-1</sup> were obtained for the three tautomers studied. The first barrier that corresponds to the setup with the most stable gas-phase tautomer is much higher than the others. Therefore, as in the case of xanthine as substrate, the kinetically active setups involve tautomers other than the lowest-energy gas-phase species. The overall barriers for these two setups are of comparable size and range between 7.3 and 11.6 kcal mol<sup>-1</sup> (covering different QM regions). The value of 14.8 kcal mol<sup>-1</sup> derived from experiment is underestimated, but the QM/MM calculations show the right trend in predicting 2-oxo-6-methylpurine to be oxidized faster than xanthine. This finding is supported by

QM model studies for the two substrates, which show the higher reactivity of 2-oxo-6-methylpurine to be an intrinsic feature of the substrate. These results contradict the paradigm of an orientation-dependent reactivity.

**Table 3.4:** Selected ESP charges (e), distances (Å), and differences  $\Delta E$  (kcal mol<sup>-1</sup>) of the corresponding Coulomb interaction energies in IM2 and the reactant (IM2-reactant, see text).

interaction	Reactant			IM2			
	charge 1	charge 2	distance	charge 1	charge 2	distance	$\Delta E$
<b>wild type, Setup G, SN400</b>							
H7-O:Glu802	0.64	-0.67	1.72	0.55	-0.68	1.96	19.4
H:Glu802-O6	0.47	-0.60	1.83	0.50	-0.66	1.66	-14.8
H:Arg880-O2	0.48	-0.70	1.90	0.51	-0.78	1.70	-19.0
H:Arg880-H3	0.48	0.50	2.47	0.51	0.37	2.79	-9.8
H:Arg880-N3	0.48	-0.66	2.90	0.51	-0.73	2.63	-10.7
<b>Glu802→Gln, Setup G, SN400</b>							
H7-O:Gln802	0.53	-0.71	1.69	0.52	-0.76	2.01	8.6
H:Gln802-O6	0.54	-0.86	2.09	0.54	-0.86	1.84	-10.0
H:Arg880-O2	0.60	-0.94	1.87	0.63	-0.98	1.68	-21.8
H:Arg880-H3	0.60	0.66	2.55	0.63	0.29	2.80	-29.9
H:Arg880-N3	0.60	-0.92	2.96	0.63	-0.61	2.66	14.0
<b>Glu802→Gln, Setup F, SN400</b>							
H7-O:Gln802	0.46	-0.84	1.78	0.52	-0.81	1.96	0.7
H:Gln802-O <sub>apikal</sub>	0.51	-0.42	2.17	0.66	-0.57	2.45	-18.2
H:Gln802-O <sub>equatorial</sub>	0.51	-0.80	3.79	0.66	-0.74	2.40	-31.8
H:Arg880-O2	0.59	-0.82	1.77	0.59	-0.86	1.63	-12.6
H:Arg880-H3	0.59	0.13	2.75	0.59	0.12	2.85	-1.0
H:Arg880-N3	0.59	-0.36	3.09	0.59	-0.48	2.92	-9.4

### 3.3 Influence of a modified cofactor

Due to the continuing discussion about the nature of the X<sub>eq</sub> functionality in the molybdopterin cofactor, we performed a series of calculations with X<sub>eq</sub> = O, S, and Se. It is widely believed that the cofactor in XO carries a sulfur ligand and that the oxo form is inactive. The influence of the

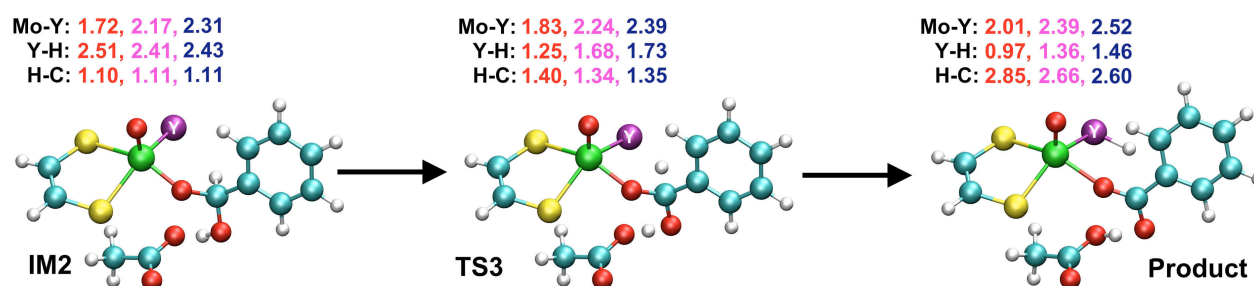
different ligands was analyzed by optimizing the complete reaction path for all three modifications. This provided the changes in the barriers for each individual step within the reductive half-cycle, see paper IV. The main influence is found for the hydride transfer step. Compared with the sulfido form with a barrier of the hydride transfer of 7 kcal mol<sup>-1</sup> and an overall barrier of about 13 kcal mol<sup>-1</sup>, the oxo species is found to be inactive: The barrier of the hydride transfer increases to more than 20 kcal mol<sup>-1</sup>, and the overall barrier to about 30 kcal mol<sup>-1</sup>. By contrast, the barrier for the selenido form drops to about 5 kcal mol<sup>-1</sup> and the overall barrier to about 11 kcal mol<sup>-1</sup>. These data explain the experimentally observed deactivation of the enzyme in its oxo form and confirm the barrier lowering when a selenium atom is introduced.

To evaluate whether the observed trends in XO are valid also for AOR, the same modifications for the molybdopterin cofactor of AOR were performed, using benzaldehyde as substrate. The calculations were restricted to the hydride transfer reaction, as this is the rate-limiting step for the reductive half-reaction of native AOR. The results for XO indicate that this step is the only one that is significantly influenced by the described modifications of the cofactor.

Compared with XO, the same trends for the barrier heights were obtained for AOR: The barrier of the hydride transfer for the oxo form gets prohibitively high (about 27 kcal mol<sup>-1</sup>). For the selenido form the barrier drops to 6.5 kcal mol<sup>-1</sup>, compared to 7.2 kcal mol<sup>-1</sup> for the sulfido form, see Table 3.5.

**Table 3.5:** QM/MM-energies in kcal mol<sup>-1</sup> for oxo, sulfido, and selenido form.

	AOR + benzaldehyde						XO + xanthine					
	Oxo		Sulfido		Selenido		Oxo		Sulfido		Selenido	
IM2	0.0		0.0		0.0		0.0		0.0		0.0	
TS3	26.6	(26.6)	7.2	(7.2)	6.5	(6.5)	20.3	(20.3)	7.0	(7.0)	4.9	(4.9)
Product	-11.9		-19.3		-16.2		-5.3		-6.8		-5.0	



**Figure 3.8:** Hydride transfer reaction for benzaldehyde in AOR. Given are the geometrical key



---

distances at B3LYP/B1 level of theory with Y = O(red), Y = S(pink) and Y = Se(purple).

Contrary to the case of acetaldehyde, no alternative pathway was found to be feasible when using benzaldehyde as substrate, as the bulky phenyl group is less flexible in its arrangement compared to the methyl group in acetaldehyde. Since we find the same increase in the barriers for AOR and XO when modifying the cofactor, it seems extremely unlikely that the cofactor of AOR contains an oxo rather than a sulfido group in its active form. This holds especially true as the origin of the increased barrier can be attributed to geometrical rearrangement effects (see Figure 3.8), which will be involved in any alternative reaction mechanism as well.



## 4 Concluding remarks

For both AOR and XO, several detailed reaction mechanisms were investigated, taking different substrate orientations and/or protonation states into account. From these results a general reaction scheme for both enzymes, namely a Lewis base catalyzed stepwise mechanism, could be established, see Figures 3.4 and 3.7. The mechanistic insights gained go well beyond experimentally detected intermediates. Despite the similarities between AOR and XO, different mechanisms to activate the substrates within these two enzymes were obtained: While for AOR the substrate is active enough to be directly attacked by the deprotonated cofactor, the conversion of xanthine into an activated tautomer is necessary in XO.

Free-energy calculations for AOR provide significant corrections to the computed barriers, but confirm the qualitative conclusions derived from the potential energy profiles. Additional calculations using a different substrate demonstrate the steric influence of the enzymatic environment, which permits only one particular substrate orientation on the preferred Lewis base activated pathway.

The issue of substrate orientation in XO leads to a situation reminiscent of the major-minor paradigm in organometallic chemistry: The "upside" orientation of the substrate represents the thermodynamically most stable reactant complex (in agreement with recent X-ray structures). However, the "upside down" orientation emerges as the kinetically active species with the lowest overall reaction barrier. This can be attributed to the strong stabilization of the tetrahedral intermediate and of the preceding and following transition states by Arg880 for this substrate orientation. In cooperative fashion, Glu1261 and Arg880 convert the protonated cofactor and the substrate in its N1,N3,N7 form into their active forms, i.e., the deprotonated cofactor and the N1,N7,N9 xanthine tautomer. Glu802 plays only a minor mechanistic role in XO, but the Glu802→Gln mutation still causes some changes in the preferred pathway that account for the experimentally observed rate reduction.

Replacement of the equatorial sulfido group in the cofactor by an oxo group leads to a prohibitive increase of the overall reaction barrier for the hydride transfer in both enzymes, AOR and XO. It is therefore extremely unlikely that the active cofactor of AOR bears an oxygen atom as equatorial ligand. Replacement by a selenido ligand, however, is calculated to lower the barrier, consistent with its presence in nicotinate dehydrogenase.

There are a number of topics that are worthwhile to be studied in future theoretical work. These include free-energy calculations in XO and the investigation of the electron-transfer reactions within the enzymes and in the oxidative half-reaction of both enzymes.

## 5 Appendix

## 5.1 List of Abbreviations

AO	Aldehyde Oxidase
AOR	Aldehyde Oxidoreductase
B1, B2	basis set 1, basis set 2, as described in paper I
B3LYP	A hybrid density functional
BP86	A ‘pure’ density functional
CC	Coupled Cluster
CHARMM	Chemistry-At-Harvard-Molecular-Mechanics, a molecular modelling software package
FAD	flavin adenine nucleotide
FADH <sub>2</sub>	reduced flavin adenine nucleotide
FEP	Free-energy perturbation
Glu869, Arg880 etc.	Three-letter amino acid names and primary sequence number
IM1,IM2,...	Intermediate structures of a reaction mechanism
LMP2	Local Møller-Plesset Perturbation theory in second order
LCCSD(T0)	Local Coupled Cluster theory with Single and Double excitations and an approximate perturbative treatment of triple excitations
LMO	localized molecular orbital
M1	MM atom adjacent to boundary
MD	Molecular Dynamics
MM	Molecular Mechanics
PMO	projected molecular orbital
Q1	QM atom adjacent to boundary
QM	Quantum Mechanics
QM/MM	(combined) Quantum Mechanical/Molecular Mechanical approach
RI	Resolution of the Identity approximation
TS	Transition State
vdW	van der Waals
XO	Xanthine Oxidase

## 5.2 List of Figures

<b>Figure 1.1:</b> Substrate specificity of AOR and XO: Aldehydes (red) are naturally catabolized by AOR, whereas purine derivatives (blue) are converted by XO. There are, however, substrates (purple) that can be oxidized by either enzyme. Note that the shaded part of the depicted cofactor is not present in XO. ....	3
<b>Figure 1.2:</b> Shown is the cytidine monophosphate molybdopterin cofactor as present in AOR. The nucleotidic part is drawn in grey and differs between the several cofactor species, see text above. ....	4
<b>Figure 1.3:</b> General sketch of the coordination sphere of the central molybdenum atom. ....	5
<b>Figure 1.4:</b> Active site of AOR (left) and XO with the substrate in an "upside down" orientation (right). In AOR, the interactions of the substrate with enzymatic side chains are dominated by nonpolar interactions, whereas in XO, they are dominated by charge-charge interactions. ....	5
<b>Figure 1.5:</b> Complete catalytic cycle of XO. ....	6
<b>Figure 1.6:</b> Reductive half-cycle of XO with different suggested pathways, see text. One representation of possible protonation states and substrate orientations was selected for this figure. The overall charge of the depicted species is given in parentheses, colored in red. ....	7
<b>Figure 1.7:</b> Reaction sequence reported by Bray et al. <sup>27</sup> The gray colored transition state was not localized. ....	8
<b>Figure 1.8:</b> Reaction sequence reported by Voityuk et al. <sup>29</sup> The gray colored transition states were not localized. ....	8
<b>Figure 1.9:</b> Mechanisms considered in previous theoretical work. (a) Protonated cofactor: Concerted and stepwise pathways reaction sequences were reported. <sup>29-31</sup> (b) Deprotonated cofactor (without proton shown in red): only a concerted and no stepwise mechanism was found. <sup>32</sup> ....	9
<b>Figure 2.1:</b> QM/MM scheme and nomenclature. ....	15
<b>Figure 3.1:</b> Scheme of three competing reaction mechanisms. ....	20
<b>Figure 3.2:</b> Reaction mechanism of AOR with benzaldehyde as substrate. ....	22
<b>Figure 3.3:</b> Small (balls & sticks) and large (including licorice) QM region of AOR with benzaldehyde. ....	24
<b>Figure 3.4:</b> Scheme of the energetically preferred reaction mechanism of AOR. ....	25
<b>Figure 3.5:</b> Reaction mechanism with the substrate oriented "upside" ....	28
<b>Figure 3.6:</b> Reaction mechanism with the substrate oriented "upside down" ....	30
<b>Figure 3.7:</b> Scheme of the energetically preferred reaction mechanism. ....	31
<b>Figure 3.8:</b> Hydride transfer reaction for benzaldehyde in AOR. Given are the geometrical key distances at B3LYP/B1 level of theory with Y = O(red), Y = S(pink) and Y = Se(purple). ....	34

### 5.3 List of Tables

<b>Table 3.1:</b> Energetics of the Lewis base catalyzed mechanism, using acetaldehyde as substrate. Energy values are given in kcal mol <sup>-1</sup> , barriers relative to the preceding minima are given in parentheses. <sup>a</sup> .....	23
<b>Table 3.2:</b> Energetics of the Lewis base catalyzed mechanism, using benzaldehyde as substrate. Energy values are given in kcal mol <sup>-1</sup> , barriers relative to the preceding minima are given in parentheses. <sup>a</sup> .....	23
<b>Table 3.3:</b> Calculated barriers from QM(B3LYP/B1)/MM calculations given in kcal mol <sup>-1</sup> . For details on the nomenclature see paper II.....	26
<b>Table 3.4:</b> Selected ESP charges (e), distances (Å), and differences ΔE (kcal mol <sup>-1</sup> ) of the corresponding Coulomb interaction energies in IM2 and the reactant (IM2-reactant, see text). .....	33
<b>Table 3.5:</b> QM/MM-energies in kcal mol <sup>-1</sup> for oxo, sulfido, and selenido form.....	34



---

## 5.4 Acknowledgements

I thank my supervisor Prof. Thiel for the opportunity to work in a stimulating field of research and for his trust that allowed me to work independently. His accessibility for discussions and his frank and open-minded suggestions helped me a lot to improve my work. I am grateful for the opportunities to extend my knowledge at several scientific meetings and for the continuous financial support.

I am grateful to Prof. Marian and Prof. B. Engels for kindly accepting to act as co-referees for this thesis.

I would like to acknowledge the fruitful collaboration with Johannes M. Dieterich and Prof. Hans-Joachim Werner at the University of Stuttgart as well as Ricardo A. Mata at the University of Göttingen.

I am grateful to Tobias Benighaus, Hans Martin Senn and Mark Waller for proof-reading my thesis.

A great thanks goes to all my colleagues at the MPI for the sociable atmosphere and many interesting scientific discussions. Especially I would like to thank Mario Ramos da Silva, Tobias Benighaus, Anoop Ayyappan, Johannes Kästner, Hans Martin Senn, and Mark Waller for sharing their insights.

I thank the computer department of the MPI Mülheim, especially Horst Lenk, who helped solve many hardware and software problems.

I am obliged to the Fond der Chemischen Industrie for a Kekulé-Stipendium.

I express my deep gratitude to Hannah and my family for their constant support and encouragement.

## 5.5 Bibliography

1. Santos-Silva, T.; Ferroni, F.; Thapper, A.; Marangon, J.; González, P. J.; Rizzi, A.; Moura, I.; Moura, J. J. G.; Romão, M. J.; Brondido, C. D. *J. Am. Chem. Soc.* **2009**, 131, 7990-7998.
2. McCarty, D. J.; Hollander, J. L. *Ann. Intern. Med.* **1961**, 54, 452-460.
3. Baker, B. R.; Hendrickson, J. L. *J. Pharm. Sci.* **1967**, 56, 955-960.
4. Hille, R.; Massey, V. *J. Biol. Chem.* **1981**, 256, 9090-9095.
5. Hawkes, T. R.; George, G. N.; Bray, R. C. *Biochem. J.* **1984**, 218, 961-968.
6. Truglio, J. J.; Theis, K.; Leimkühler, S.; Rappa, R.; Rajagopalan, K. V.; Kisker, C. *Structure* **2002**, 10, 115-125.
7. Okamoto, K.; Eger, B. T.; Nishino, T.; Kondo, S.; Pai, E. F.; Nishino, T. *J. Biol. Chem.* **2003**, 278, 1848-1855.
8. Fukunari, A.; Okamoto, K.; Nishino, T.; Eger, B. T.; Pai, E. F.; Kamezawa, M.; Yamada, I.; Kato, N. *J. Pharmacol. Exp. Ther.* **2004**, 311, 519-528.
9. Pacher, P.; Nivorozhkin, A.; Szabo, C. *Pharmacol. Rev.* **2006**, 58, 87-114.
10. Tamta, H.; Thilagavathi, R.; Chakraborti, A. K.; Mukhopadhyay, A. K. *J. Enzyme Inhib. Med. Chem.* **2005**, 20, 317-324.
11. Nagamatsu, T.; Yamasaki, H.; Fujita, T.; Endo, K.; Machida, H. *J. Chem. Soc., Perkin Trans. I* **1999**, 1, 3117-3125.
12. Clarke, S. E.; Harrell, A. W.; Chemery, R. J. *Drug Metab. Dispos.* **1995**, 23, 251-254.
13. Rashidi, M. R.; Smith, J. A.; Clarke, S. E.; Beedham, C. *Drug Metab. Dispos.* **1997**, 25, 805-813.
14. Kawashima, K.; Hosoi, K.; Naruke, T.; Shiba, T.; Kitamura, M.; Wanatabe, T. *Drug Metab. Dispos.* **1999**, 27, 422-428.
15. Lake, B. G.; Ball, S. E.; Ka, J.; Renwick, A. B.; Price, R. J.; Scatina, J. A. *Xenobiotica* **2002**, 32, 835-847.
16. Obach, R. S. *Drug Metab. Dispos.* **2004**, 32, 89-97.
17. Kisker, C.; Schindelin, H.; Rees, D. C. *Annu. Rev. Biochem.* **1997**, 66, 233-267.
18. Einsle, O.; Tezcan, F. A.; Andrade, S. L. A.; Schmidt, B.; Yoshida, M.; Howard, J. B.; Rees, D. C. *Science* **2002**, 297, 1696-1700.
19. Hille, R. *Chem. Rev.* **1996**, 96, 2757-2816.
20. Hille, R.; Sprecher, H. *J. Biol. Chem.* **1987**, 262, 10914-10917.
21. Doonan, C. J.; Stockert, A.; Hille, R.; George, G. N. *J. Am. Chem. Soc.* **2005**, 127, 4518-4522.
22. Stockert, A. L.; Shinde, S. S.; Anderson, R. F.; Hille, R. *J. Am. Chem. Soc.* **2002**, 124, 14554-14555.
23. Barata, B. A. S.; LeGall, J.; Moura, J. J. G. *Biochemistry* **1993**, 32, 11559-11568.
24. Romão, M. J.; Rösch, N.; Huber, R. *J. Biol. Inorg. Chem.* **1997**, 2, 782-785.
25. Mendel, R. R.; Schwarz, G. *Crit. Rev. Plant Sci.* **1999**, 18, 33-69.
26. Schwarz, G.; Mendel, R. R. *Annu. Rev. Plant Biol.* **2006**, 57, 623-647.
27. Bray, M. R.; Deeth, R. J. *J. Chem. Soc., Dalton Trans.* **1997**, 1267-1268.
28. Ilich, P.; Hille, R. *Inorg. Chim. Acta* **1997**, 263, 87-93.
29. Voityuk, A. A.; Albert, K.; Romão, M. J.; Huber, R.; Rösch, N. *Inorg. Chem.* **1998**, 37, 176-180.
30. Ilich, P.; Hille, R. *J. Phys. Chem. B* **1999**, 103, 5406-5412.
31. Zhang, X. H.; Wu, Y. D. *Inorg. Chem.* **2005**, 44, 1466-1471.
32. Amano, T.; Ochi, N.; Sato, H.; Sakaki, S. *J. Am. Chem. Soc.* **2007**, 129, 8131-8138.
33. Alfaro, J. F.; Jones, J. P. *J. Org. Chem.* **2008**, 73, 9469-9472.
34. Bayse, C. A. *Dalton Trans.* **2009**, 2306-2314.
35. Huber, R.; Hof, P.; Duarte, R. O.; Moura, J. J. G.; Moura, I.; Liu, M. Y.; LeGall, J.; Hille, R.; Archer, M.; Romão, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, 93, 8846-8851.

- 
36. Massey, V.; Edmondson, D. *J. Biol. Chem.* **1970**, 245, 6595-6598.
  37. Wagener, N.; Piersek, A. J.; Ibdah, A.; Hille, R.; Dobbek, H. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 11055-11060.
  38. Howes, B. D.; Bray, R. C.; Richards, R. L.; Turner, N. A.; Bennett, B.; Lowe, D. J. *Biochemistry* **1996**, 35, 1432-1443.
  39. Manikandan, P.; Choi, E. Y.; Hille, R.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, 123, 2658-2663.
  40. Hille, R. *Arch. Biochem. Biophys.* **2005**, 433, 107-116.
  41. Paufl, J. M.; Hemann, C. F.; Jünemann, N.; Leimkühler, S.; Hille, R. *J. Biol. Chem.* **2007**, 282, 12785-12790.
  42. Yamaguchi, Y.; Matsumura, T.; Ichida, K.; Okamoto, K.; Nishino, T. *J. Biochem.* **2007**, 141, 513-524.
  43. Nishino, T.; Okamoto, K.; Eger, B. T.; Pai, E. F.; Nishino, T. *FEBS J.* **2008**, 275, 3278-3289.
  44. Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, 136, B864.
  45. Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, 140, A1133.
  46. Slater, J. C. *Phys. Rev.* **1951**, 81, 385-390.
  47. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, 58, 1200-1211.
  48. Becke, A. D. *Phys. Rev. A* **1988**, 38, 3098-3100.
  49. Perdew, J. P. *Phys. Rev. B* **1986**, 33, 8822-8824.
  50. Perdew, J. P. *Phys. Rev. B* **1986**, 34, 7406.
  51. Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648-5652.
  52. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, 98, 11623-11627.
  53. Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785-789.
  54. Perdew, J. P.; Ernzerhof, M.; Burke, K. *J. Chem. Phys.* **1996**, 105, 9982-9985.
  55. Blomberg, M. R. A.; Siegbahn, P. E. M.; S., S.; Babcock, G. T.; Åkermærk, B.; Åkermærk, B.; Korall, P. *J. Am. Chem. Soc.* **1997**, 119, 8285-8292.
  56. Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, 50, 221-249.
  57. Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, 100, 421-438.
  58. Siegbahn, P. E. M.; Blomberg, M. R. A. *Philos. Trans. R. Soc. London, A* **2005**, 363, 847-860.
  59. Ghosh, A.; Taylor, P. R. *Curr. Opin. Chem. Biol.* **2003**, 7, 113-124.
  60. Himo, F.; Siegbahn, P. E. M. *Chem. Rev.* **2003**, 103, 2421-2456.
  61. Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, 105, 2279-2328.
  62. de Visser, S. P.; Ogliaro, F.; Harris, N.; Shaik, S. *J. Am. Chem. Soc.* **2001**, 123, 3037-3047.
  63. Schöneboom, J. C.; Lin, H.; Reuter, N.; Thiel, W.; Cohen, S.; Ogliaro, F.; Shaik, S. *J. Am. Chem. Soc.* **2002**, 124, 8142-8151.
  64. Claeysens, F.; Harvey, J. N.; Manby, F. R.; Mata, R. A.; Mulholland, A. J.; Ranaghan, K. E.; Schütz, M.; Thiel, S.; Thiel, W.; Werner, H.-J. *Angew. Chem., Int. Ed.* **2006**, 45, 6856-6859.
  65. Mata, R. A.; Werner, H.-J.; Thiel, S.; Thiel, W. *J. Chem. Phys.* **2008**, 128, 025104.
  66. Pulay, P. *Chem. Phys. Lett.* **1983**, 100, 151-154.
  67. Saebo, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, 113, 13-18.
  68. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, 4, 187-217.
  69. MacKerell, A. D., et al. *J. Phys. Chem. B* **1998**, 102, 3586-3616.
  70. Brooks, B. R., et al. *J. Comput. Chem.* **2009**, 30, 1545-1615.
  71. Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, 268, 173-290.
  72. Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, 100, 10580-10594.
  73. Singh, U. C.; Kollmann, P. A. *J. Comput. Chem.* **1986**, 7, 718-730.
  74. de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, 103, 6133-6141.

- 
75. Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, 79-92.
  76. Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, 28, 578-581.
  77. Kirkwood, J. G. *J. Chem. Phys.* **1935**, 3, 300-313.
  78. Kästner, J.; Thiel, W. *J. Chem. Phys.* **2005**, 123.
  79. Kästner, J.; Senn, H. M.; Thiel, W. *J. Chem. Theory Comput.* **2006**, 2, 452-461.
  80. Zhang, Y.; Liu, H.; Yang, W. T. *J. Chem. Phys.* **2000**, 112, 3483-3492.
  81. Zwanzig, R. *J. Chem. Phys.* **1954**, 22, 1420-1426.
  82. Senn, H. M.; Kästner, J.; Breidung, J.; Thiel, W. *Can. J. Chem.* **2009**, 87, 1322-1337.
  83. Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, 59, 573-601.
  84. Henkelman, G.; Uberuaga, B. P.; Jonsson, H. *J. Chem. Phys.* **2000**, 113, 9901-9904.
  85. Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, 131, 4628-4640.
  86. Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, 28, 213-222.
  87. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, 4, 294-301.
  88. Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, 82, 284-298.
  89. Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, 208, 237-240.
  90. Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, 82, 270-283.
  91. Ehlers, A. W.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, 208, 111-114.
  92. Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, 7, 3297-3305.
  93. Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, 8, 1057-1065.
  94. Dietzel, U.; Kuper, J.; Doebbler, J. A.; Schulten, A.; Truglio, J. J.; Leimkühler, S.; Kisker, C. *J. Biol. Chem.* **2009**, 284, 8768-8776.
  95. Kim, J. H.; Ryan, M. G.; Knaut, H.; Hille, R. *J. Biol. Chem.* **1996**, 271, 6771-6780.
  96. Paufl, J. M.; Zhang, J. J.; Bell, C. E.; Hille, R. *J. Biol. Chem.* **2008**, 283, 4818-4824.
  97. Metz, S.; Thiel, W. *J. Am. Chem. Soc.* **2009**, 131, 14885-14902.
  98. Paufl, J. M.; Cao, H.; Hille, R. *J. Biol. Chem.* **2009**, 284, 8760-8767.
  99. Kulikowska, E.; Kierdaszuk, B.; Shugar, D. *Acta Biochim. Pol.* **2004**, 51, 493-531.
  100. Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Science* **2008**, 321, 792-794.
  101. Edmondson, D.; Ballou, D.; Vanheuve, A.; Palmer, G.; Massey, V. *J. Biol. Chem.* **1973**, 248, 6135-6144.
  102. Olson, J. S.; Ballou, D. P.; Palmer, G.; Massey, V. *J. Biol. Chem.* **1974**, 249, 4350-4362.
  103. Mondal, M. S.; Mitra, S. *Biochemistry* **1994**, 33, 10305-10312.
  104. Choi, E. Y.; Stockert, A. L.; Leimkühler, S.; Hille, R. *J. Inorg. Biochem.* **2004**, 98, 841-848.
  105. Sevastik, R.; Himo, F. *Bioorg. Chem.* **2007**, 35, 444-457.
  106. Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput.* **2008**, 4, 1129-1137.
  107. Chen, S.-L.; Fang, W.-H.; Himo, F. *Theor. Chem. Acc.* **2008**, 120, 515-522.
  108. Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, 14, 643-651.
  109. McWhirter, R. B.; Hille, R. *J. Biol. Chem.* **1991**, 266, 23724-23731.

Hiermit versichere ich, die hier vorgelegte Arbeit eigenständig und ohne unerlaubte Hilfe angefertigt zu haben. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner Institution eingereicht. Ich habe keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 17.12.2009

A handwritten signature in blue ink, reading "Sebastian Metz". The signature is written in a cursive, flowing style.

(Sebastian Metz)



## Paper I

*Reductive Half-Reaction of Aldehyde Oxidoreductase  
toward Acetaldehyde: A Combined QM/MM Study*

**Sebastian Metz**, Dongqi Wang,  
and Walter Thiel

*J. Am. Chem. Soc.*, 131 **(2009)** 4628-4640.





## Reductive Half-Reaction of Aldehyde Oxidoreductase toward Acetaldehyde: A Combined QM/MM Study

Sebastian Metz, Dongqi Wang, and Walter Thiel\*

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

Received July 29, 2008; E-mail: thiel@mpi-muelheim.mpg.de

**Abstract:** We report a combined QM/MM study on the mechanism of the reductive half-reaction of aldehyde oxidoreductase. Five possible pathways are explored concerning the binding mode of acetaldehyde and the catalytic effect of the nearby glutamic acid (Glu869), taking both possible protonation states into account. In the most favorable pathway, Glu869 participates and acts as a Lewis base to deprotonate the labile hydroxide group. This proton transfer is essential for the high activity of the enzyme toward substrate because it increases the nucleophilicity of the migrating O atom and strengthens the electrophilicity of the target C atom in the substrate. The subsequent product-forming reactions occur in two discrete steps, first nucleophilic attack and then hydride transfer, which implies that the oxidation of aldehyde is a two-electron process. A variant of this mechanism, with an additional water molecule bridging the Glu869 side chain and the substrate, has similar barriers. Judging from previous gas phase calculations and our present QM/MM data, the catalytic effect of Glu869 mainly lowers the barrier of the nucleophilic attack so that the hydride transfer becomes the rate-determining step in the reductive half-reaction.

### 1. Introduction

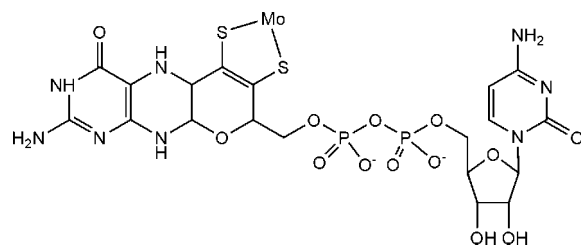
Mononuclear molybdenum enzymes<sup>1–5</sup> constitute a large class of enzymes possessing a pterin cofactor (see Scheme 1) which coordinates to the metal center and may have different forms in different enzymes. They are generally categorized into three families by the structural homology of the active site:<sup>1,6</sup> xanthine oxidase, sulfite oxidase, and DMSO reductase. Important members of the first family are aldehyde oxidoreductase (AOR), xanthine dehydrogenase (XDH), and xanthine oxidase (XO). All these enzymes catalyze the oxidation of their target substrates, aldehydes or xanthine. The mechanism of these enzymatic reactions has been studied both theoretically<sup>7–13</sup> and experimentally.<sup>14–26</sup>

Structure-based catalytic mechanisms for the reductive half-reaction have been proposed<sup>14–17</sup> and adopted for the xanthine oxidase family. The oxygen source consumed in the biological hydroxylation process catalyzed by xanthine oxidase is recognized experimentally to be water rather than atmospheric oxygen, and the labile O should attach to the Mo center at the proximal position prior to the single turnover.<sup>18</sup> It has been suggested that a proton transfer from molybdenum-bound water to the nearby glutamate activates the cofactor. The substrate then interacts with the Mo-cofactor with high stereospecificity to induce the so-called reductive half-reaction. The release of the carboxylic acid product may be assisted by the transient coordination of glutamate to the metal center which will be replaced by a new water.

Due to the participation of glutamate in the metabolism, it is conceivable that the activity of the enzyme could be affected by the pH of the environment. The pH-dependence of the XO activity toward xanthine and lumazine<sup>19</sup> as well as 1-methyl-xanthine<sup>20</sup> has been explored, and a Lewis base-catalyzed scheme<sup>14,17</sup> by the glutamate at the proximal position of the coordinated Mo-center (Glu1261 in XO and Glu869 in AOR)

- (1) Hille, R. *Chem. Rev.* **1996**, 96, 2757–2816.
- (2) Hille, R. *Biochim. Biophys. Acta* **1994**, 1184, 143–169.
- (3) Hille, R. *Arch. Biochem. Biophys.* **2005**, 433, 107–116.
- (4) Sauer, P.; Frébort, I. *Biol. Plant.* **2003**, 46, 481–490.
- (5) Hille, R. *Trends Biochem. Sci.* **2002**, 27, 360–367.
- (6) Mendel, R. R. *Dalton Trans.* **2005**, 3404–3409.
- (7) Bray, M. R.; Deeth, R. J. *J. Chem. Soc., Dalton Trans.* **1997**, 1267–1268.
- (8) Ilich, P.; Hille, R. *Inorg. Chim. Acta* **1997**, 263, 87–93.
- (9) Voityuk, A. A.; Albert, K.; Romão, M. J.; Huber, R.; Rösch, N. *Inorg. Chem.* **1998**, 37, 176–180.
- (10) Ilich, P.; Hille, R. *J. Phys. Chem. B* **1999**, 103, 5406–5412.
- (11) Ilich, P.; Hille, R. *J. Am. Chem. Soc.* **2002**, 124, 6796–6797.
- (12) Zhang, X. H.; Wu, Y. D. *Inorg. Chem.* **2005**, 44, 1466–1471.
- (13) Amano, T.; Ochi, N.; Sato, H.; Sakaki, S. *J. Am. Chem. Soc.* **2007**, 129, 8131–8138.
- (14) Huber, R.; Hof, P.; Duarte, R. O.; Moura, J. J. G.; Moura, I.; Liu, M. Y.; LeGall, J.; Hille, R.; Archer, M.; Romão, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, 93, 8846–8851.
- (15) Romão, M. J.; Archer, M.; Moura, I.; Moura, J. J. G.; Legall, J.; Engh, R.; Schneider, M.; Hof, P.; Huber, R. *Science* **1995**, 270, 1170–1176.
- (16) Romão, M. J.; Rösch, N.; Huber, R. *J. Biol. Inorg. Chem.* **1997**, 2, 782–785.
- (17) Okamoto, K.; Matsumoto, K.; Hille, R.; Eger, B. T.; Pai, E. F.; Nishino, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 7931–7936.

- (18) Hille, R.; Sprecher, H. *J. Biol. Chem.* **1987**, 262, 10914–10917.
- (19) Kim, J. H.; Ryan, M. G.; Knaut, H.; Hille, R. *J. Biol. Chem.* **1996**, 271, 6771–6780.
- (20) Sau, A. K.; Mondal, M. S.; Mitra, S. *J. Chem. Soc., Dalton Trans.* **2000**, 3688–3692.
- (21) Mondal, M. S.; Mitra, S. *Biochemistry* **1994**, 33, 10305–10312.
- (22) Xia, M.; Dempski, R.; Hille, R. *J. Biol. Chem.* **1999**, 274, 3323–3330.
- (23) Greenwood, R. J.; Wilson, G. L.; Pilbrow, J. R.; Wedd, A. G. *J. Am. Chem. Soc.* **1993**, 115, 5385–5392.
- (24) Howes, B. D.; Bray, R. C.; Richards, R. L.; Turner, N. A.; Bennett, B.; Lowe, D. J. *Biochemistry* **1996**, 35, 1432–1443.
- (25) Manikandan, P.; Choi, E. Y.; Hille, R.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, 123, 2658–2663.
- (26) Doonan, C. J.; Stockert, A.; Hille, R.; George, G. N. *J. Am. Chem. Soc.* **2005**, 127, 4518–4522.

**Scheme 1.** Pterin Cytosine Dinucleotide (PCD) Cofactor in AOR

has been proposed. The experiments on temperature-dependent transient kinetics<sup>20,21</sup> suggest that there is at least one intermediate in the course of the turnover of xanthine to the product.

Debates on how Mo and the labile O interact with each other in the resting state have focused on three candidates: Mo–OH<sub>2</sub>, Mo–OH, and Mo=O. The possibility of a Mo=O double bond was ruled out in an investigation that proposed a Mo–O single bond conformation<sup>22</sup> which is consistent with previous suggestions<sup>23</sup> and EPR experiments.<sup>25</sup> In another recent contribution, the labile O was experimentally determined by George and co-workers<sup>26</sup> to be a hydroxide ligand instead of a bound water molecule. In this work, an EXAFS analysis gave detailed information on the pH dependence of the Mo-cofactor conformation: the bond length of Mo–O(labile) shortens from 1.97 Å to 1.75 Å upon the increase of pH from 6 to 10. The value of 1.75 Å is in the range of Mo–O bond lengths (average: 1.77 Å) found in crystallographic databases.<sup>27</sup> Hence, Lewis base-assisted deprotonation is invoked to abstract a proton from Mo–OH to produce a Mo–O (or Mo=O) bond, and the labile O is assigned to be –OH.<sup>26</sup>

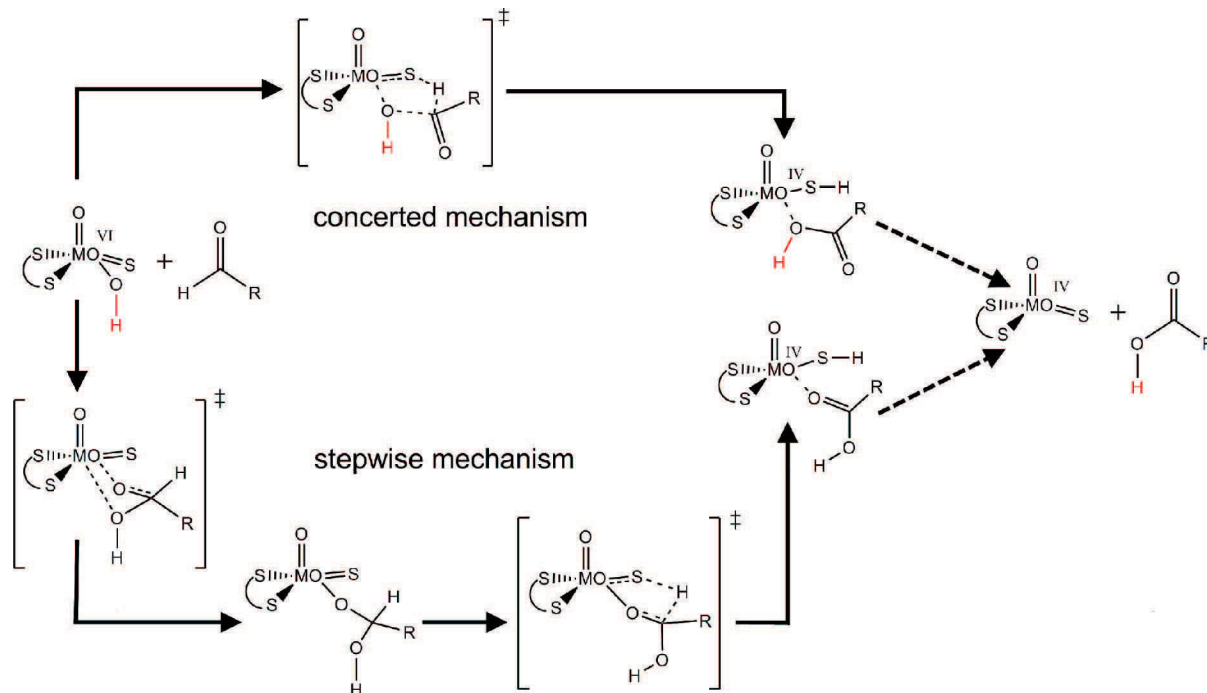
On the theoretical side, there have been a number of QM model studies that mainly focused on the oxidation of formaldehyde<sup>7,9,12</sup> or formamide.<sup>10–13</sup> In one case, acetaldehyde or formamidine were investigated as substrates.<sup>12</sup> The pterin cofactor

(see Scheme 1) was normally modeled by a [S–CR=CR–S]<sup>2–</sup> ligand, R being hydrogen<sup>7,9,10,13</sup> or a methyl group,<sup>11,12</sup> so that the active species was represented by the negatively charged complex [(S–CR=CR–S)Mo(=S)(=O)(–OH)]<sup>–</sup> in most studies.<sup>7–13</sup> Amano et al. also considered the deprotonated and thus dianionic complex [(S–CH=CH–S)Mo(=S)(=O)(–O)]<sup>2–</sup>, as well as a larger cofactor model.<sup>13</sup> None of the studies reported so far included a model for Glu869.

Two essentially different pathways have been identified, a concerted and a stepwise mechanism (see Scheme 2). For the concerted mechanism, a prohibitive barrier of 78 kcal/mol (relative to the infinitely separated reactants) was obtained in a UMP2/Lan12DZ study with formamide as substrate.<sup>10</sup> Ilich and Hille<sup>11</sup> investigated the substituent effect of the Mo-cofactor by comparing the native form of the cofactor (sulfido) to the desulfo-form (oxo or tellurido) at the MP2 level. They found barriers for the reaction between the three congeners and formamide of 91, 78, and 75 kcal/mol for O, S, and Te, respectively, implying that the Mo=O form is inert.

An all-electron DFT(BP86) study<sup>9</sup> with formaldehyde as substrate located a stable intermediate prior to the hydride transfer step, which implies a stepwise mechanism. The barrier for the formation of the intermediate was not computed, while the hydride transfer was calculated to require an activation of 7.7 kcal/mol. The overall reaction was found to be thermoneutral, and hence suggested to be reversible.

The most complete study at the DFT(B3LYP) level of theory was reported by Zhang and Wu.<sup>12</sup> Two possible pathways, concerted and stepwise, were investigated for the reaction with the substrate formaldehyde. For the stepwise pathway, a nucleophilic attack happens first to form the O(labile atom)–C(substrate) bond followed by a hydride transfer step with cleavage of the C–H and formation of the S–H bond. The barriers for the two steps were calculated to be 17.8 and 5.4 kcal/mol, respectively. For the

**Scheme 2.** Mechanisms Considered in Previous Theoretical Work<sup>a</sup>

<sup>a</sup>(i) protonated cofactor: concerted and stepwise pathways, see refs 9, 10 and 12 (ii) deprotonated cofactor (without proton shown in red): concerted pathway, see ref 13 (stepwise pathway not found).

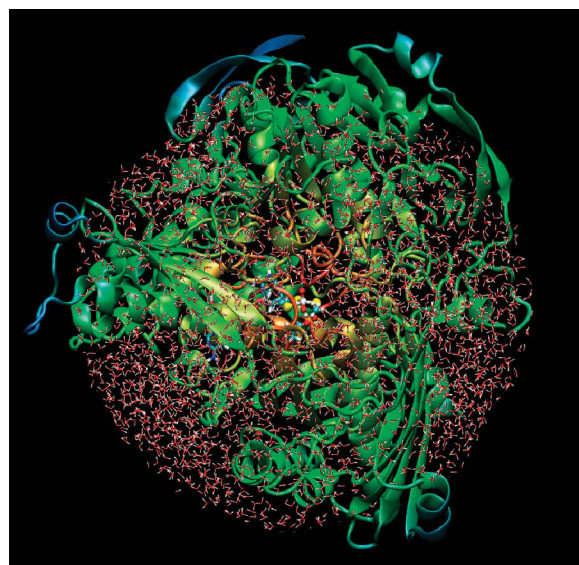
concerted pathway which treats the nucleophilic attack and hydride transfer as a concomitant process, a lower barrier was obtained (11.9 kcal/mol). Solvent effects were taken into account using the PCM model and were found not to change the preference for the concerted path. However, there is experimental evidence against such a concerted mechanism<sup>12</sup> in the case of the oxidation of xanthine by xanthine oxidase. It is known from isotope-labeling experiments that the labeled oxygen from the Mo–<sup>18</sup>OH is transferred to the product<sup>18</sup> and that the product is bound via this labeled oxygen, as has been proven in the case of violapterin oxidized by xanthine oxidase.<sup>28</sup>

In a recent study, Amano et al.<sup>13</sup> applied DFT(B3LYP) and correlated ab initio methods up to CCSD(T) to calculate energy profiles using formamide as a substrate. They considered concerted and stepwise mechanisms both for protonated and deprotonated cofactor models. The reported DFT(B3LYP) and CCSD(T) barriers are in the range of 39–42 kcal/mol (concerted/protonated, stepwise/protonated) and 35–38 kcal/mol (concerted/deprotonated), respectively, while a stepwise mechanism with a deprotonated cofactor could not be found. The authors conclude that a one-step mechanism with a deprotonated active site is most plausible.

QM model studies can provide valuable mechanistic insights at the molecular level, but they normally do not include the effects of the protein/solvent environment. This is achieved by QM/MM methods which have become a popular alternative for exploring enzymatic reactions with reasonable accuracy and affordable computational cost.<sup>29</sup> Here we report our recent QM/MM studies on the reductive half-reaction of aldehyde oxidoreductase with acetaldehyde as substrate. We have investigated five possible pathways to compare the concerted and stepwise protocols and to evaluate the importance of Glu869, checking both possible protonation states. Compared with published QM model studies (see above), pathways A and B (with protonated cofactor) have common features with the previously considered concerted and stepwise mechanisms, respectively. No QM equivalents to pathways C–E have been reported so far; a stepwise mechanism with deprotonated cofactor (as in pathway C) has been searched for, but could not be located at the QM level.<sup>13</sup> Our QM/MM calculations suggest that the reductive half-reaction is a three-step process, and support the Lewis base-assisted scheme with deprotonated Glu869 which activates the nucleophilic attack.

## 2. Methods

Initial coordinates were taken from the X-ray crystal structure<sup>30</sup> (PDB Code: 1VLB, resolution: 1.28 Å) obtained from *Desulfovibrio gigas*. The Mo=OR1 group of the desulfo-form was replaced by Mo=SR1 to prepare the active form. The protonation states of the titratable residues (His, Glu, Asp) were chosen based on the pK<sub>a</sub> values given by the empirical PROPKA procedure<sup>31</sup> and verified through visual inspection. The total charge of the whole system was –12e. In addition, we built a neutral system with zero net charge by selectively protonating titratable residues on the surface of the protein. A partial solvation scheme was used to solvate the region of 35 Å around the Mo center by overlaying a water ball on the enzyme. A potential was imposed on the water sphere to prevent the free water molecules from escaping into the vacuum. The



**Figure 1.** AOR with solvent shell from partial solvation setup.

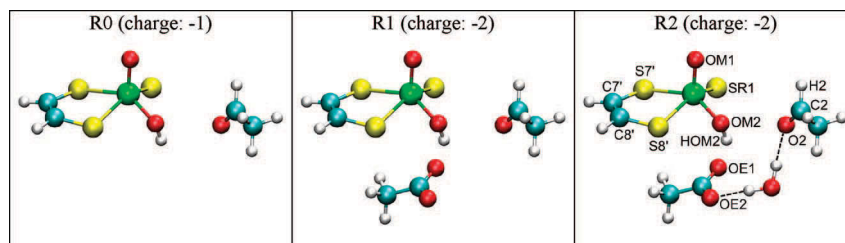
solvated systems were then relaxed by performing energy minimizations and molecular dynamics (MD) simulations at the MM level using the CHARMM27 force field<sup>32</sup> as implemented in the CHARMM program.<sup>33</sup> During the classical energy minimizations and MD simulations, an active region was defined to include the residues within 20 Å, and the water molecules within 35 Å around the Mo atom. The residues and water outside the active region as well as the non-hydrogen atoms of pterin, the pyrano group (dithio-Mo(=O)(=S)(OH)-ene) and the C=O group of the substrate were fixed. The equilibrated systems (see Figure 1) contained 23357 atoms, including 3237 TIP3P water molecules and 5 ions (two Mg<sup>2+</sup> and three Cl<sup>–</sup>).

The chosen QM/MM methodology is analogous to that used in our previous studies. Here we briefly mention some aspects relevant to the present work. Minimized snapshots from the MD trajectories were taken as the initial structures for QM/MM optimization. In the QM/MM calculations, the QM part was treated by the B3LYP<sup>34–39</sup> density functional method, and the MM part was described by the CHARMM27 force field. An electronic embedding scheme<sup>40</sup> was adopted in the QM/MM calculations, the MM charges were incorporated into the one-electron Hamiltonian of the QM calculation, and the QM/MM electrostatic interactions were evaluated from the QM electrostatic potential and the MM partial charges. No cutoffs were introduced for the nonbonding MM and QM/MM interactions. Hydrogen link atoms with the charge shift model<sup>41,42</sup> were employed to treat the QM/MM boundary. The TURBOMOLE program<sup>43</sup> was used for the QM treatment in the QM/MM

- (27) Allen, F.; Kennard, O. *Chem. Des. Autom. News* **1993**, 8, 31–37.  
 (28) Hemann, C.; Ilich, P.; Stockert, A. L.; Choi, E. Y.; Hille, R. *J. Phys. Chem. B* **2005**, 109, 3023–3031.  
 (29) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, 268, 173–290.  
 (30) Rebelo, J. M.; Dias, J. M.; Huber, R.; Moura, J. J. G.; Romão, M. J. *J. Biol. Inorg. Chem.* **2001**, 6, 791–800.

- (31) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, 61, 704–721.  
 (32) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, 102, 3586–3616.  
 (33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, 4, 187–217.  
 (34) Slater, J. C. *Phys. Rev.* **1951**, 81, 385–390.  
 (35) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, 58, 1200–1211.  
 (36) Becke, A. D. *Phys. Rev. A* **1988**, 38, 3098–3100.  
 (37) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648–5652.  
 (38) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, 98, 11623–11627.  
 (39) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785–789.  
 (40) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, 100, 10580–10594.  
 (41) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, 103, 6133–6141.  
 (42) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, 79–92.  
 (43) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, 162, 165–169.





**Figure 2.** Definition of the QM regions.

calculations. The CHARMM27 force field was run through the DL\_POLY program<sup>44</sup> to handle the MM part of the systems. The QM/MM calculations were performed with the ChemShell package<sup>45</sup> that integrates the TURBOMOLE and DL\_POLY programs and also performs geometry optimization with the HDLC optimizer.<sup>46</sup>

Three QM regions were adopted in the QM/MM calculations (Figure 2). The QM region R0 represents the simplest model containing only the molybdenum-cofactor model  $[\text{Mo}(\text{S}_2\text{C}_2\text{H}_2)(=\text{O})(\text{OH})(=\text{S})]^-$  and the substrate acetaldehyde ( $\text{CH}_3\text{CHO}$ ). Since the residue Glu869 is conserved within the XO family and essential for the activity of XHD,<sup>47</sup> it is reasonable to include it in the QM part: QM region R1 is obtained from R0 by adding part of the side chain of Glu869 ( $-\text{CH}_2\text{CO}_2^-$ ). QM region R2 contains an additional water molecule in the active site. For one of the pathways we employ a variant of QM region R1 called R1p with protonated Glu869 ( $-\text{CH}_2\text{CO}_2\text{H}$ ). The total charge of the QM region is  $-1$  for R0 and R1p, and  $-2$  for R1 and R2. The atom labels shown in Figure 2 are taken from the crystal structure 1VLB. Whenever necessary, they will be specified more precisely in the format Resname:AtomName, for example as AALD:O2 for atom O2 in acetaldehyde.

Two basis sets were employed that are defined as follows:

B1: Lanl2DZ<sup>48</sup> + f polarization<sup>49</sup> for Mo, Lanl2DZ<sup>50</sup> + d polarization<sup>51</sup> for S, and 6-31+G\*\*<sup>52,53</sup> for the rest (C, H, O); this is the basis used in ref 12.

B2: Standard def2-TZVP basis set<sup>54</sup> composed of ECP-28-MWB-TZVext<sup>55</sup>+P(f) for Mo, TZV'+P(21/1) for S, TZVPP for C, O, TZP for H.

For QM region R1, B1 and B2 contain 316 and 552 basis functions, respectively. All pathway calculations and optimizations of stationary points were done using basis B1 in combination with the B3LYP hybrid functional. In addition, the stationary points on the energetically favored pathway C were reoptimized using the larger basis B2 in combination with the functionals BP86,<sup>34–36,56,57</sup> BLYP,<sup>34,36,39</sup> B3LYP<sup>34–39</sup> and BHLYP.<sup>34,36,39,58</sup>

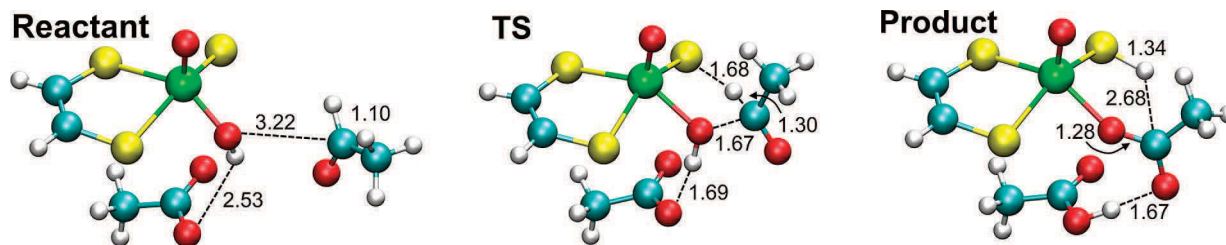
In geometry optimizations at the QM/MM level, the active optimized region included the QM region and all residues and water molecules of the MM region within 10 Å around the Mo-center (see Supporting Information for a detailed list). Reaction paths were scanned along suitably defined reaction coordinates by performing constrained optimizations at each point. This provided starting structures for subsequent full optimizations of all relevant stationary points which employed the low-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm for minima and the partitioned rational function optimizer (P-RFO) for transition states, as implemented in the HDLC code.<sup>46</sup> We ensured by careful reaction path optimization and visual inspection of the optimized structures that the computed stationary points are connected by continuous pathways. Frequency calculations on the QM region confirmed that all reported transition states are characterized by a dominant transition vector that corresponds to the investigated reaction.

### 3. Results

Our mechanistic studies on aldehyde oxidoreductase at the QM/MM level start from the hydroxide-bound resting state. Five pathways (A–E) were investigated which differ with regard to the interaction between the cofactor and substrate, the mechanistic role and protonation state of Glu869, and the involvement of water. The neutral model system was used to perform the calculations on Pathways A, B, and C, while the charged model was used for Pathways C and D. For Pathway E, we set up a second neutral system with protonated Glu869, starting at the product side with bound acetic acid. All energies values given in this section are QM/MM energies (i.e., including QM, MM, and QM/MM interaction terms), without zero-point and thermal corrections. Unless noted otherwise, these energies were obtained using B3LYP/B1 as QM treatment.

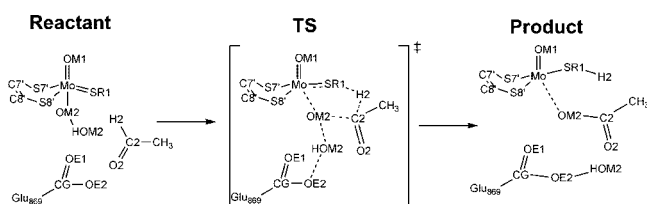
**3.1. Pathway A: Concerted Reaction.** Pathway A is concerted and similar to the one considered in the QM model system.<sup>12</sup> The reaction proceeds by the simultaneous formation of the C2–OM2 and SR1–H2 bonds (see Scheme 3). The reaction coordinate was therefore defined as  $d_{\text{C2-H2}}-d_{\text{OM2-C2}}-d_{\text{SR1-H2}}$ . Transition state (TS) optimization gave a barrier of 20.2 kcal/mol using QM region R0. This value is slightly higher than that found in corresponding QM model calculations (18.5 kcal/mol for acetaldehyde<sup>12</sup>) which may be due to the intrinsic constraints of the protein in the active site. Without direct information on the topology of the active site in the substrate-bound enzyme, we rely on the crystal structure of inhibitor-bound enzyme where there are three water molecules helping to anchor the inhibitor through H-bond interactions with the nearby residue Arg501. In our setup, these water molecules are conserved and may be involved in H-bond interactions with the substrate. In fact, on the energy surface for moving acetaldehyde to the cofactor, there is a very shallow region in the beginning until the OM2–C2 distance is around 3 Å (see Supporting Information). The closer approach of acetaldehyde to the cofactor implies some loss of hydrogen bonding with

- (44) Smith, W.; Forester, T. R. *J. Mol. Graph.* **1996**, *14*, 136–141.
- (45) Sherwood, P.; et al. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1–28.
- (46) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
- (47) Leimkühler, S.; Stockert, A. L.; Igarashi, K.; Nishino, T.; Hille, R. *J. Biol. Chem.* **2004**, *279*, 40437–40444.
- (48) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270–283.
- (49) Ehlers, A. W.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 111–114.
- (50) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284–298.
- (51) Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 237–240.
- (52) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- (53) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (54) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (55) Andrae, D.; Häussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123–141.
- (56) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (57) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
- (58) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.



**Figure 3.** Structures for the concerted pathway from R1/B1 calculations.

**Scheme 3.** Mechanism of the Concerted Reaction (Total Charge  $-2$ )



the water molecules which is expected to make a minor contribution to the barrier.

We have also studied the concerted path by including residue Glu869 in the QM region (i.e., using QM region R1) and allowing it to participate in the reaction (see Scheme 3). The calculations on the neutral model predict a barrier of 20.2 kcal/mol which is same as that from R0 calculations. In spite of this similarity, there are differences concerning HOM2. In the reactant, there is an H-bond interaction between Glu869:OE1 and HOM2, which is conserved throughout the R0 calculation. By contrast, in the R1 calculation, this interaction is lost, and HOM2 forms a new bond with Glu869:OE2. The electrostatic effect from Glu869 was examined by setting the point charges in the side chain of Glu869 to zero in an additional R0 test calculation; the barrier was found to increase by about 8 kcal/mol, and the reaction became endothermic by 4 kcal/mol. This clearly demonstrates the stabilizing electrostatic effect of Glu869 on this pathway.

Figure 3 shows the stationary points found at the R1/B1 level. A five-membered ring forms in this phase, and we observe the partial formation of C2–OM2 (1.667 Å) and SR1–H2 (1.683 Å), as well as the partial cleavage of Mo–OM2 (2.070 Å) and C2–H2 (1.298 Å).<sup>59</sup> The Mo=S bond is weakened and is becoming a single bond (2.243 Å). These geometrical changes reflect the concerted nucleophilic attack and hydride transfer (see also the population analysis in Figure S3 of the Supporting Information).

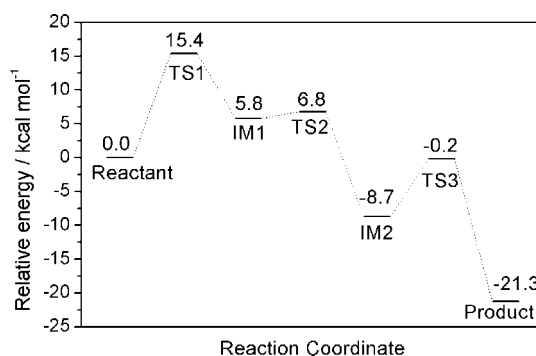
**3.2. Pathway B: Initial Dative Bond between Mo and Substrate.** Since the Mo atom can generally be penta- or hexacoordinated, we have considered an alternative pathway B with initial coordination of the substrate. The first step in pathway B is the formation of a dative C2–OM2 bond coupled with the coordination of AALD:O2 to Mo. An intermediate is obtained with OM2 shared by Mo and AALD:C2. Rotation around the C2–OM2 bond then breaks the interaction between Mo and OM2. The last step is the hydride transfer (H2) from acetaldehyde to SR1 (see Scheme 4).

The possible binding of the carbonyl oxygen of aldehyde to molybdenum was proposed in a previous crystallographic

study<sup>15</sup> but was characterized as being unfavorable in a recent QM study.<sup>12</sup> We carried out calculations on this pathway to find out whether it can become a possible channel for the oxidation of acetaldehyde when the protein environment is included. As the results with QM regions R0 and R1 are very similar (see Supporting Information), we discuss only the R1 results.

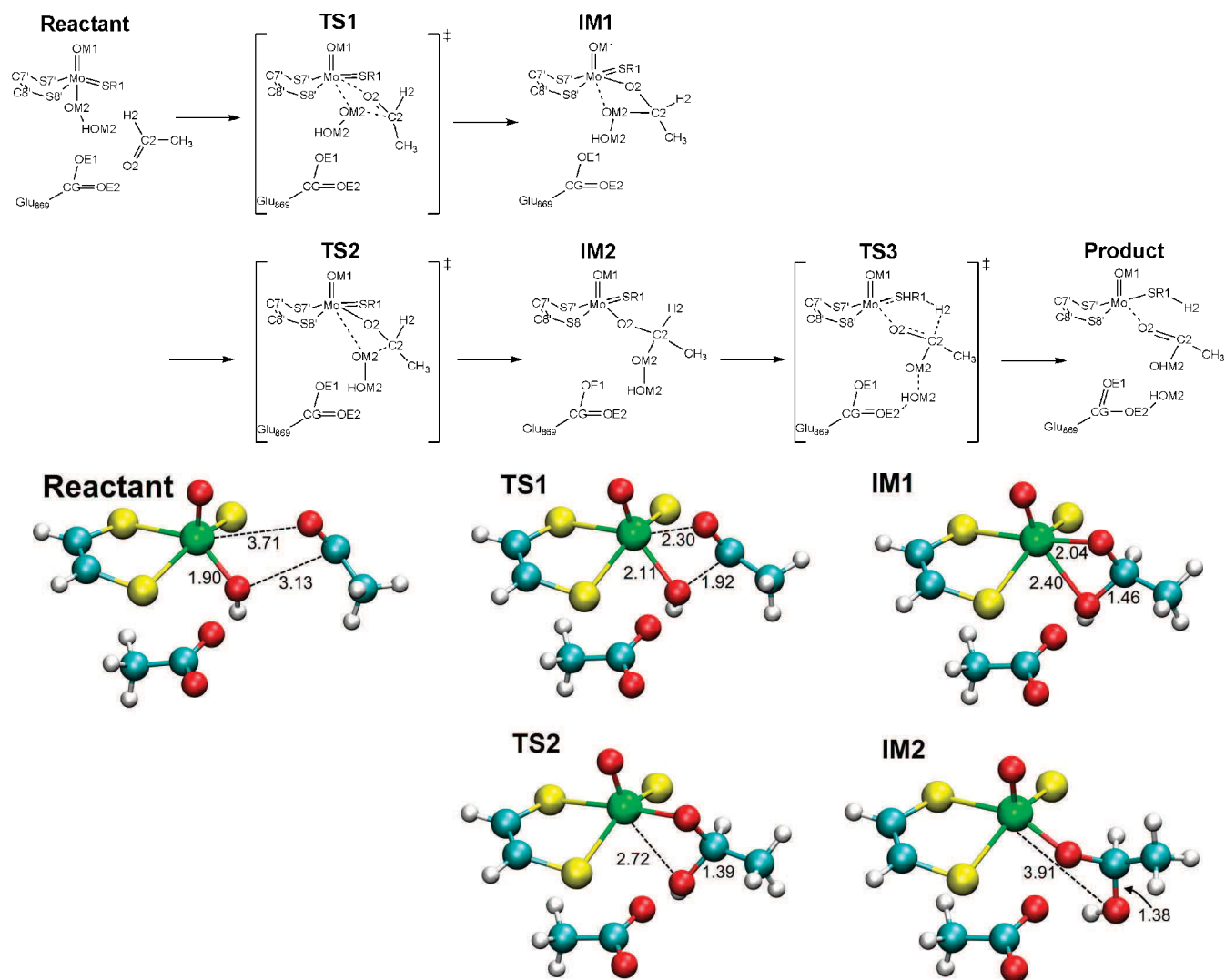
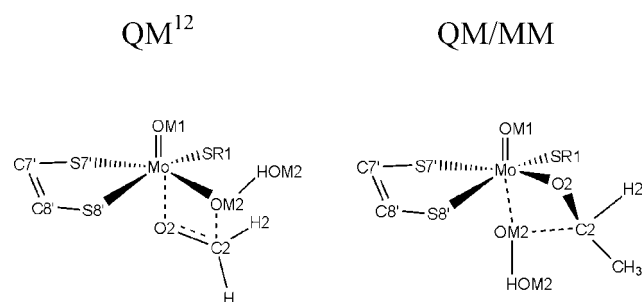
For pathway B, all calculations gave a mechanism in which the nucleophilic attack and hydride transfer happen sequentially. The migration of OM2 to C2 occurs first and an intermediate forms. Hence, we defined the reaction coordinate for formation of the C2–OM2 bond, coupled with the coordination of the carbonyl group (O2) of the substrate to the Mo atom, and the following reorientation of the hydroxide group as  $d_{\text{Mo-OM2}}-d_{\text{OM2-C2}}$ . The final hydride transfer is represented by the reaction coordinate  $d_{\text{C2-H2}}-d_{\text{H2-SR1}}$  and yields the product acetic acid and the cofactor in its reduced form.

In a previous QM study, Wu and Zhang<sup>12</sup> explored the possibility of the formation of a Mo–O dative bond between the cofactor model and formaldehyde, and reported a barrier of 17.8 kcal/mol for the nucleophilic attack. In the QM/MM calculations, the interactions with the surrounding protein environment favor a different orientation of the hydrogen bond network around the hydroxo group, and in addition, the enzyme provides a lipophilic region (Phe425, Tyr535) that further helps to orient the substrate (see Figure S1 in the Supporting Information). This leads to somewhat different transition states: in the QM study, the O2 atom attaches to the Mo atom from the position trans to OM1 (proximal position) to form a distorted octahedral intermediate, while in the QM/MM study, we do not observe any coordination of the substrate from the proximal position where the negatively charged Glu869 is situated and prevents a facile approach of the O2 atom which instead comes in between OM1 and OM2. If we consider the QM/MM transition structure as a distorted octahedron, the O2 atom occupies one equatorial position (in a plane together with S7',



**Figure 4.** QM/MM energy profile for pathway B from R1/B1 calculations (neutral system).

(59) Smith, P. D.; Slizys, D. A.; George, G. N.; Young, C. G. *J. Am. Chem. Soc.* **2000**, *122*, 2946–2947.

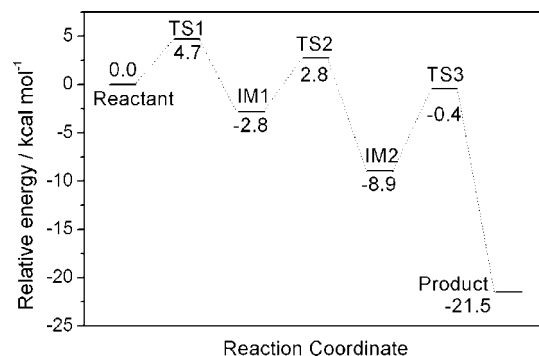
**Scheme 4.** Mechanism of Pathway B (Total Charge −2)**Figure 5.** Selected stationary points along pathway B from R1/B1 calculations.**Scheme 5.** Difference in the Coordination Pattern for TS1 in QM and QM/MM Work

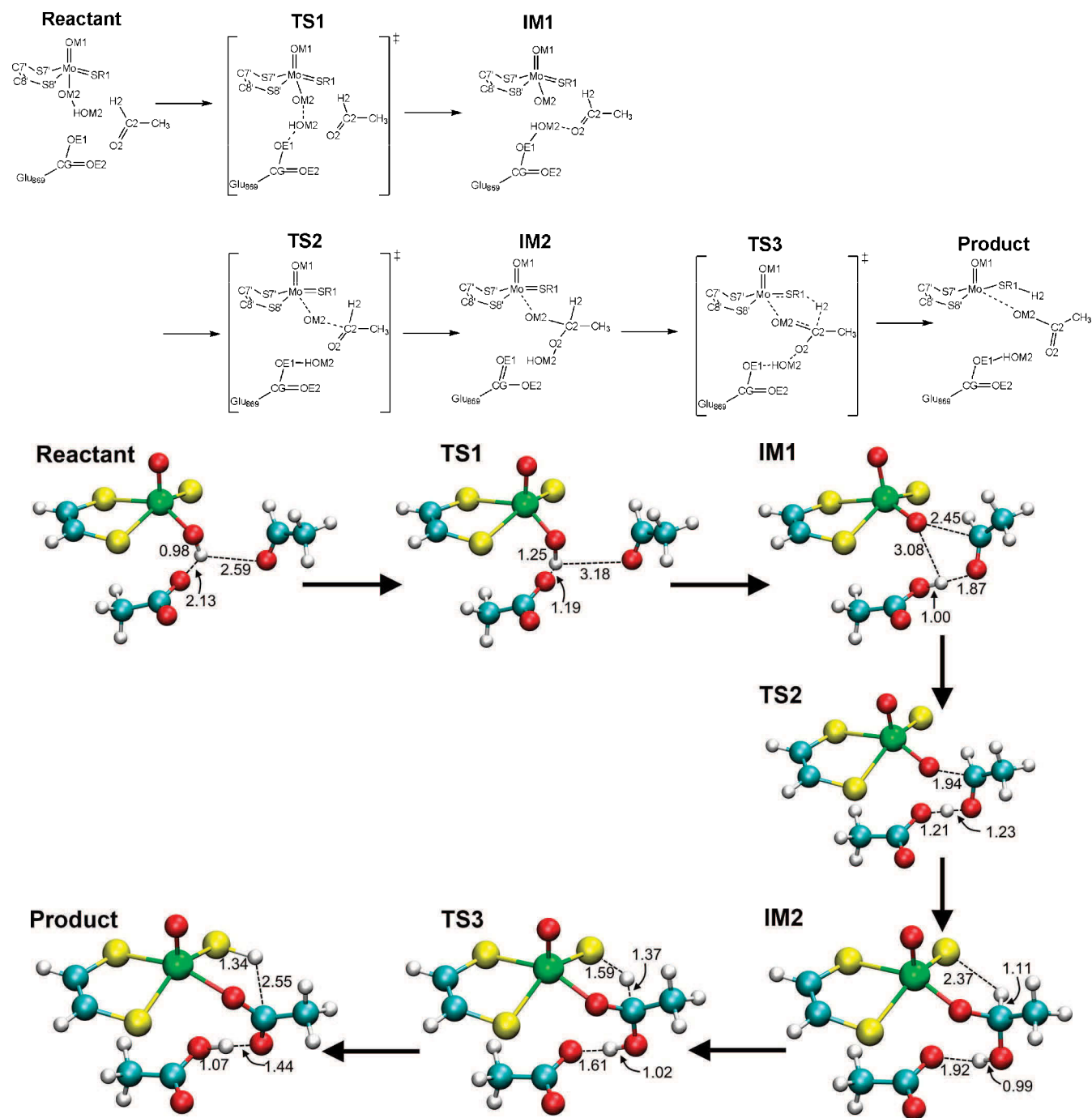
S8', and SR1), and OM2 moves down to stay on the axis containing the OM1 and Mo atoms. The corresponding structures are shown in Scheme 5.

The energy profile for pathway B is depicted in Figure 4. The initial nucleophilic attack of the hydroxide group (OM2–HOM2) via TS1 has to overcome a barrier of 15.4 kcal/mol. The resulting intermediate (IM1) is a shallow minimum, and the dissociation of the Mo–OM2 bond (i.e., the reorientation of the OM2–HOM2 group) via TS2 requires an activation of only 1.0 kcal/mol. The final hydride transfer via

TS3 has a barrier of 8.5 kcal/mol (relative to IM2). The rate-determining barrier of 15.4 kcal/mol for TS1 is lower than the corresponding QM value (21.9 kcal/mol<sup>12</sup>). The overall reaction is exothermic by 21.3 kcal/mol.

Compared to pathway A, the formation of the hexa-coordinated molybdenum complex not only guides the reaction to a stepwise mechanism, but also decreases the activation energy significantly. This suggests that the stepwise pathway

**Figure 6.** QM/MM energy profile for pathway C from R1/B1 calculations on the neutral model.

**Scheme 6.** Mechanism of Pathway C (Total Charge –2)**Figure 7.** Stationary points along pathway C from R1/B1 calculations.

is more favorable for the reductive half-reaction, with the nucleophilic attack of hydroxide to acetaldehyde being the rate-determining step for this pathway.

Figure 5 shows selected stationary points along this pathway. We can see that in the intermediate (IM2), the OM2–HOM2 group migrates to the target C atom (AALD:C2) of the substrate and the AALD:O2 atom bridges Mo and C2 (distances of 1.901 and 1.448 Å, respectively). After the hydride transfer, the interaction between Mo and O2 is weakened due to the recovery of the C2=O2 double bond.

For each of the pathways A and B considered so far, R0 and R1 calculations (without and with Glu869 in the QM region)

gave similar results. This indicates that the role of Glu869 can be equally well described at the MM and QM level in these two mechanistic scenarios where Glu869 serves as an H-bond acceptor. This is no longer true for pathways C and D which involve an initial proton transfer from the Mo-cofactor to Glu869 that acts as a base. The QM/MM calculations for pathways C and D will therefore employ the larger QM regions R1 and R2, respectively.

**3.3. Pathway C: Glu869 Promoted Pathway.** Pathway C consists of three steps and starts with a proton transfer (HOM2) from the Mo-cofactor to Glu869 prior to the reductive half-reaction. The Glu869:OE1 and AALD:O2



atoms are bridged by the proton (HOM2), and thus there is no coordination between Mo and O2 as observed in Pathway B (see Scheme 6).

A recent mutation study on xanthine dehydrogenase (XDH)<sup>47</sup> indicates that the glutamate (Glu730, corresponding to Glu869 in AOR) is essential for catalysis and may contribute ~10 kcal/mol in stabilizing the transition state. A Lewis base-assisted scheme has been proposed to address the crucial role of glutamate.<sup>60</sup> In the following, we explore this scheme in which a proton transfer step from the Mo-cofactor to the deprotonated Glu869 occurs prior to the nucleophilic attack step.

The current QM/MM calculations suggest that the proton transfer from Mo-cofactor to Glu869 does not happen spontaneously during the formation of the OM2–C2 bond. Hence, to take the catalytic effect of residue Glu869 into account explicitly, we moved the proton PCD:HOM2 to Glu869. This step was found to have a barrier of 4.7 kcal/mol for the neutral model (3.5 kcal/mol for the charged model), and the proton could stay on Glu869 only with the assistance of the H-bond acceptor nearby, the substrate AALD. The proton transfer produces an intermediate (IM1) wherein the Mo-cofactor and AALD are bridged by this proton (HOM2) through H-bond interactions (see Figure 6).

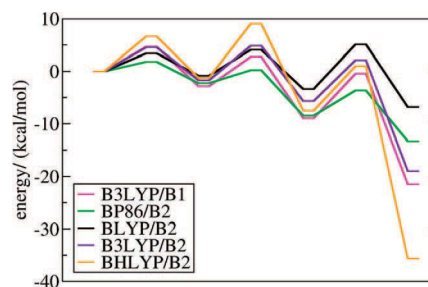
In the intermediate, a new Mo=O bond is formed between Mo and OM2 (1.74 Å, comparable to the axial Mo=OM1 bond) and the oxo-group (OM2) is expected to be more active toward the reaction with acetaldehyde. The calculated barrier for the nucleophilic attack is found to be 5.6 kcal/mol (QM/MM energy relative to IM1, 7.6 kcal/mol for the charged model).

In contrast to the significant activation of the nucleophilic attack step by Glu869, the following hydride transfer is not affected much. The barrier to the hydride transfer in pathway C is calculated to be 8.5 kcal/mol (relative to IM2, 8.6 kcal/mol for the charged model) which is comparable to the values in pathway B. Hence, the main role of Glu869 as a base is to activate the hydroxide addition to the substrate, while it has little effect on the barrier to the hydride transfer. The barrier for the hydride transfer is slightly larger than that for the C2–OM2 bond formation (by 2.9 kcal/mol for the neutral model). Thus, hydride transfer becomes the rate-determining step in pathway C.

Unlike pathways A and B, pathway C involves a two-step proton transfer from OM2 to AALD:O2 in the semiacetal intermediate (IM2). By excluding Glu869 from the QM region, we calculated the direct proton transfer in a single step and obtained a similar exothermicity as for pathway C (–11.2 kcal/mol vs –8.9 kcal/mol) but a much larger barrier of 23.4 kcal/mol. This underlines the crucial mechanistic role of Glu869: the initial proton transfer on pathway C is facile only if it proceeds as a two-step process via an intermediate (IM1) where the proton is “stored” at Glu869.

The semiacetal intermediate (IM2), the precursor for the hydride transfer step, is the same as the one on pathway B. Thus, the QM/MM calculations predict an identical energy profile for the last step in pathways B and C (neutral model).

The optimized geometries along pathway C are similar for the charged and neutral models, and therefore we only show the stationary points obtained from the calculations on the neutral model in Figure 7. As can be seen here, after the proton transfer, the migrated proton HOM2 acts as a bridge to acetaldehyde through an H-bond interaction with OM2 (1.871



**Figure 8.** Influence of basis set and functional on energy profile.

**Table 1.** Calculated QM/MM-Energies in kcal/mol Relative to the Energy of the Reactant for Different Basis Sets and Functionals; Activation Barriers Relative to the Preceding Minima Are Given in Parentheses

	B3LYP/B1	B3LYP/B2	BP86/B2	BLYP/B2	BHLYP/B2
reactant	0.0	0.0	0.0	0.0	0.0
TS1	4.7 (4.7)	4.6 (4.6)	1.8 (1.8)	3.4 (3.4)	6.7 (6.7)
IM1	–2.8	–1.7	–2.3	–0.8	–1.2
TS2	2.8 (5.6)	4.9 (6.6)	0.2 (2.5)	4.1 (4.9)	9.1 (10.3)
IM2	–8.9	–5.6	–8.4	–3.3	–7.5
TS3	–0.5 (8.4)	2.1 (7.7)	–3.6 (4.4)	5.1 (8.4)	1.0 (8.5)
product	–21.5	–19.0	–13.3	–6.8	–35.6

Å, IM1). During the nucleophilic attack step, this proton builds a covalent bond with O2 in the semiacetal intermediate (1.232 Å in TS2 and 0.991 Å in IM2) to facilitate the C2–OM2 bond formation. The third step, hydride transfer, has a transition state (TS3) with a five-membered ring (C2–H2: 1.368 Å, H2–SR1: 1.594 Å, Mo–SR1: 2.265 Å, Mo–OM2: 1.972 Å, OM2–C2: 1.345 Å).

For pathway C we reoptimized all stationary points with basis set B2 and different functionals. Comparing the energy profiles of B3LYP/B1 and B3LYP/B2 (see Figure 8 and Table 1), there is hardly any influence of the enlarged basis set (for details see Supporting Information). Concerning the different functionals, we find the expected behavior for the calculated transition states: the more exact exchange the functional incorporates, the higher are the calculated barriers for TS1 and TS2. TS3 differs from this finding, as the calculated barriers (except the one for BP86) are very similar in energy; in this case, the amount of exact exchange in the functional affects the stability of the resulting Mo(IV) more strongly than the barrier for TS3. For all functionals except BHLYP, hydride transfer via TS3 is the rate-determining step. Overall, the computed reaction profiles are not too sensitive to the choice of basis set or functional, at least in the qualitative sense, which provides justification for using B3LYP/B1 calculations for a consistent comparison of different pathways. We also note in this context, that B3LYP and CCSD(T) results have been found to agree well in a related system.<sup>13</sup>

The results reported here are from calculations on partially solvated models (see section 2). We have also performed calculations on pathway C using a more expensive, fully solvated model which contains 34565 atoms (see Supporting Information for model setup). The computed B3LYP/B1 (B3LYP/B2) barriers are 5.8 (7.3), 2.9 (3.2) and 11.1 (12.4) kcal/mol for proton transfer, nucleophilic attack, and hydride transfer, respectively. They are consistent with the B3LYP results for the partially solvated model (see Table 1).

**3.4. Pathway D: Glu869-Promoted Pathway with One Additional Water Molecule.** Pathway D is a variant of pathway C with an additional water molecule bridging Glu869 and

(60) Hille, R. *Eur. J. Inorg. Chem.* **2006**, 191, 3–1926.



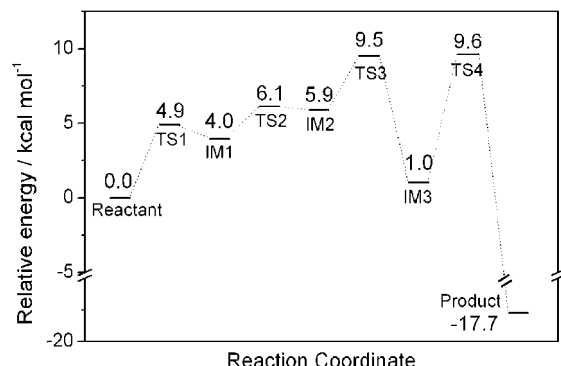


Figure 9. QM/MM energy profile for pathway D from R2/B1 calculations.

AALD, which may facilitate the reaction by a better precoordination of the substrate. The calculations on this pathway explore the possible effect of water present in the active site using the charged model. Classical MD simulations confirm that the additional water molecule remains stable in the active site during a 200 ps MD run (see Figure S11 in the Supporting Information). Due to the similarities in energetics and geometries between the charged and neutral models in pathway C, we did not perform calculations using the neutral model for pathway D.

Following the most favorable protocol found in pathway C, we carefully moved the proton (HOM2) from the cofactor to Glu869. In the calculations on pathway C, an important feature of the first step is the formation of an H-bond between HOM2 and AALD:O2 immediately after the proton transfer which is necessary to assist in the approach of AALD. The presence of one water molecule between Glu869 and AALD does not favor the reorientation of OE1–HOM2 group to form this H-bond. Hence, after the proton HOM2 is delivered to OE1 with a barrier of 4.9 kcal/mol, an additional energy of 2.1 kcal/mol (TS2 in Figure 9) is needed for the corresponding rotation.

Taking advantage of the well-arranged conformation of the active site, the C–O formation becomes a facile process with a barrier of 3.6 kcal/mol. The hydride transfer is the rate-

determining step in the reductive half-reaction with a barrier of 8.7 kcal/mol. The whole reaction is exothermic by  $-17.7$  kcal/mol. Hence, with regard to the energetics, pathway D is comparable to pathway C and can be considered as an alternative.

Figure 10 shows selected stationary points of pathway D. In the educt, a water molecule is found between the substrate acetaldehyde and Glu869 acting as an H-bond bridge (distances of 1.820 and 1.739 Å, respectively). This water molecule helps to anchor the substrate, and the H-bond with AALD:O2 may also make AALD:C2 more electrophilic. After the oxidation of the substrate, the water molecule may move up to bind to the Mo atom and thus replenish the oxygen source. The initial proton transfer proceeds via TS1 and produces intermediate IM1 which, due to the presence of the water molecule, has a large distance between HOM2 and AALD:O2 (3.339 Å). The rotation of the OE1–HOM2 group via TS2 leads to a strong HOM2–O2 interaction: the substrate adjusts its orientation such that the HOM2–O2 distance decreases from 2.694 Å in TS2 to 1.676 Å in IM2. In the resulting IM2 conformation, the target atom for the nucleophilic attack, C2, is already quite close to OM2 (2.364 Å) which facilitates the oxygen transfer from molybdenum to C2 (barrier: 3.6 kcal/mol).

**3.5. Pathway E: Protonated Glu869.** Pathway E is the only one that involves protonated Glu869: Initially, the substrate is rearranged by changes in the H-bond orientation. Afterward, there is a synchronous double proton transfer with concomitant formation of the C2–OM2 bond followed by a hydride transfer (H2) from acetaldehyde to SR1 (see Scheme 7).

In contrast to the previous models, it is not possible to store the HOM2 proton from the cofactor at Glu869:OE1, as the latter residue is already protonated. Therefore, HOE2 has to be transferred at the same time to the substrate, i.e. to AALD:O2. The thus activated substrate concomitantly forms the C2–O2 bond. As there is one additional proton in the QM region, the subsequent hydride transfer, in contrast to all other setups, yields a protonated acetic acid as product. The stationary points are depicted in Figure 11, and the energy profile is shown in Figure 12.

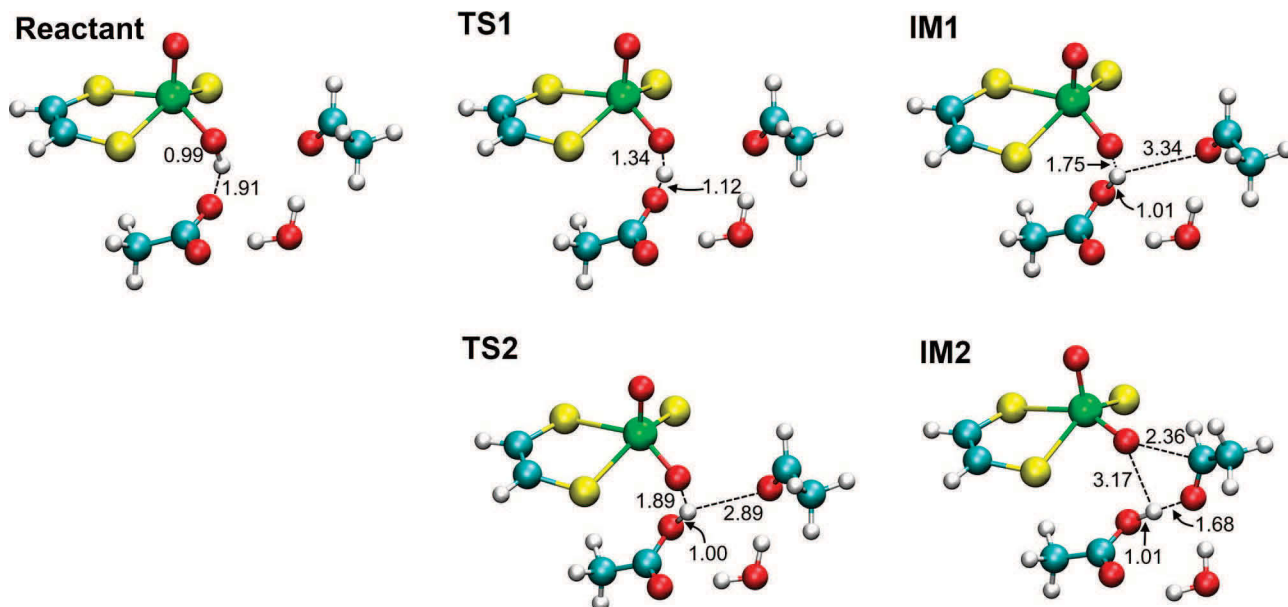
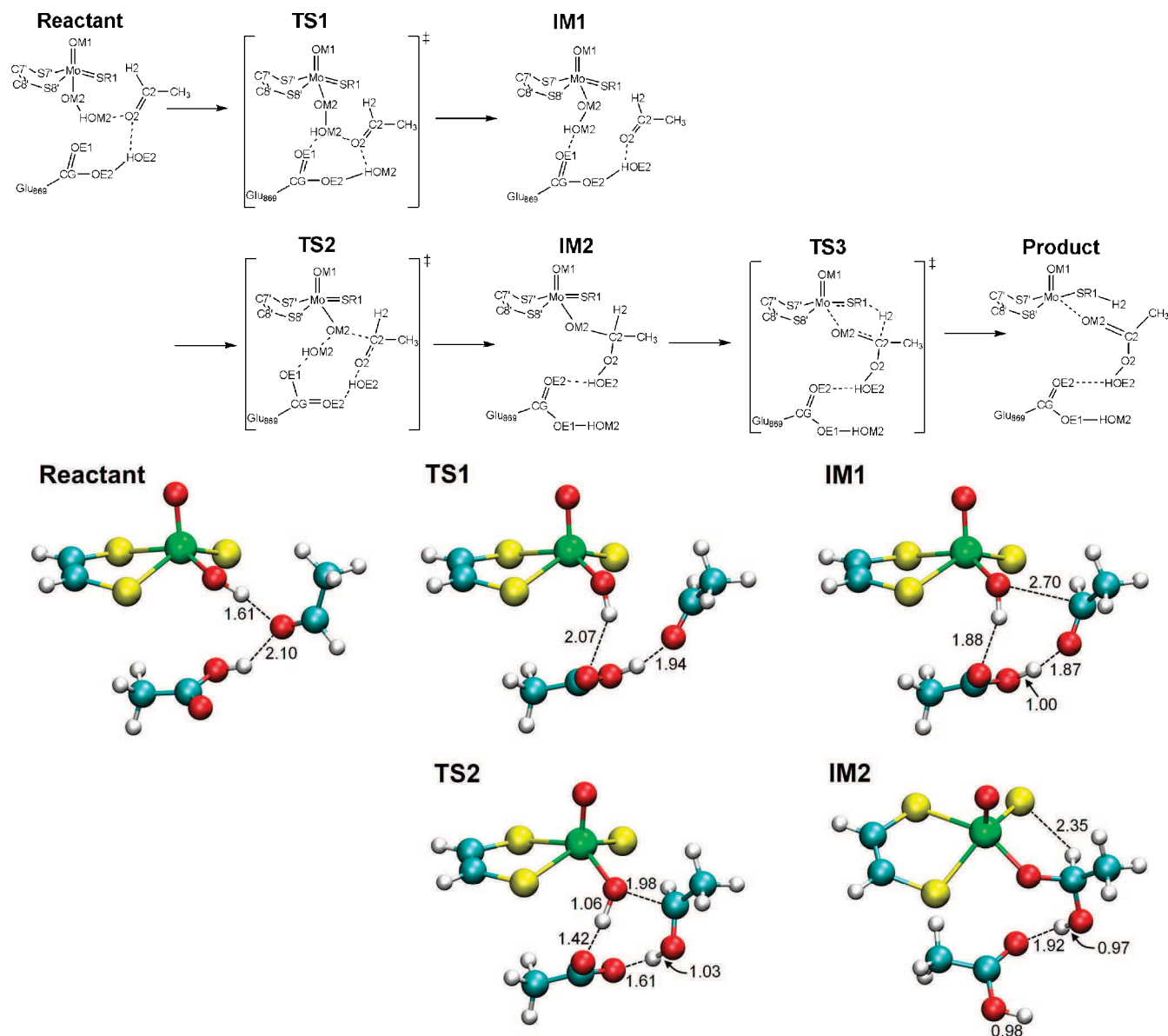
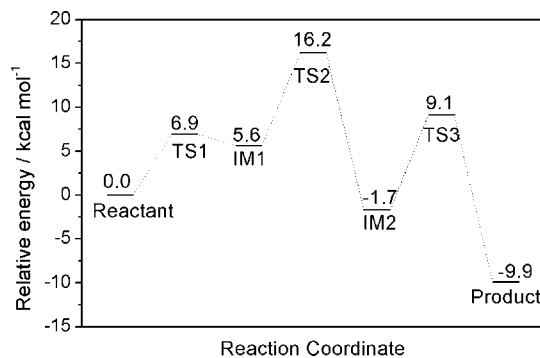


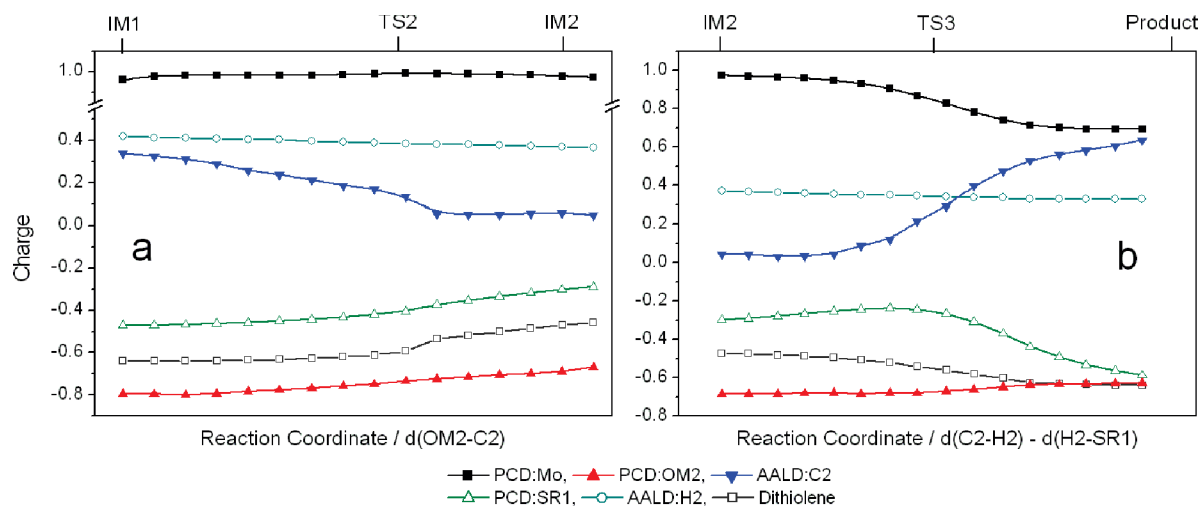
Figure 10. Selected stationary points along pathway D from R2/B1 calculations.

**Scheme 7.** Mechanism of Pathway E (Total Charge  $-1$ )**Figure 11.** Stationary points along pathway E from R1p/B1 calculations.**Figure 12.** QM/MM energy profile for pathway E from R1p/B1 calculations.

The barrier for such a “double proton transfer” in a single step is rather high. Compared with pathway C (R1 region) there is another complicating feature: In the lowest-lying educt, HOM2 is no longer oriented toward Glu869, but forms a

hydrogen bond to the substrate, which leads to a less effective precoordination. The initial phase of the reaction thus involves reorientation of this hydrogen bond (reactant  $\rightarrow$  IM1) as well as C–O bond formation (IM1  $\rightarrow$  IM2). The effective barrier for this initial phase is 16.2 kcal/mol (see Figure 12) as IM1 resides in a very shallow minimum and is thus in a fast equilibrium with the reactant. For the final hydride transfer (IM2  $\rightarrow$  product), the barrier of 10.8 kcal/mol for pathway E is just slightly higher than that for pathway C (8.5 kcal/mol). Comparing the rate-determining barriers from the R1/R1p calculations, pathway C is clearly favored over pathway E (8.5 vs 16.2 kcal/mol).

**3.6. Pathway C: Mulliken Charges and Fold Angles.** In this section we analyze the changes of Mulliken charge distributions and of the puckering of the five-membered ring during the reductive half-reaction of pathway C. The results are similar for the charged and neutral models, and thus we only discuss the former ones to avoid repetition.



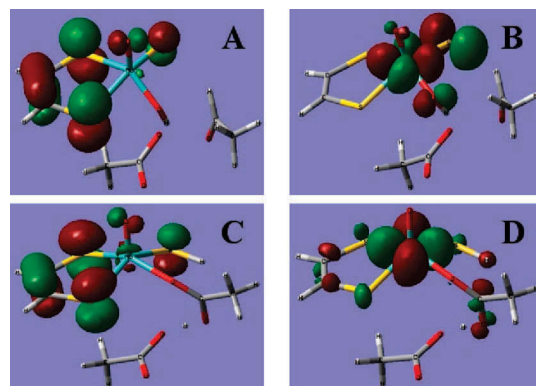
**Figure 13.** Mulliken charge distributions along pathway C. (Left) C–O formation. (Right) Hydride transfer. The position of the corresponding minima and transition states is indicated (see Scheme 6 and Figure 7).

**3.6.1. Partial Charges.** In organometallic catalysis, the consumption of the substrate is often accompanied by charge transfer and by a change of oxidation state of the metal center. Hence, monitoring the charge flow may help to characterize redox reactions. The reaction under study involves the formation of a C–O bond (with concomitant weakening of Mo–O) and the dissociation of a C–H bond (coupled to S–H formation). The Mo atom is reduced formally from oxidation state VI to IV, while the substrate (acetaldehyde, AALD) is oxidized. To characterize this charge flow in more detail, Figure 13 shows the change of the partial charges on selected groups (Mo, C2, H2, OM2, SR1, and dithiolene) along pathway C.

In the case of C–O bond formation, the negative charge flows from the Mo-cofactor to the substrate and the target C2 is reduced (see Figure 13a). In the intermediate (IM2 in Figure 7), AALD:O2 receives the proton HOM2 from OE1 and hence becomes part of a hydroxide group. As a result of the accumulation of negative charge on C2, the character of C2 as a base is enhanced upon formation of the C2–OM2 bond which will facilitate the following hydride transfer. For the hydride transfer step (Figure 13b), C2 becomes more acidic when losing the hydrogen atom H2, while Mo and SR1 are reduced. The group charge on the Mo-cofactor becomes more negative which implies a charge transfer concomitant with the migration of H2 from C2 to SR1.

The reduction of the Mo atom also affects the frontier orbitals for the stationary points along the reaction path (Figure 14). In the reactant (oxidized state), the HOMO mainly contains contributions from dithiolene with  $\pi_{C=C}$  and sulfur  $p_z$ -orbital components (denoted as  $\pi_{C=C} - p_z(S7') - p_z(S8'')$ ), and the LUMO is essentially an antibonding combination of the Mo in-plane d-orbital and the p-orbital of SR1 (denoted as  $d_{xy}(\text{Mo}) - p_x(\text{SR1})$ ). In the product (reduced state), the  $d_{xy}$ -orbital of Mo is occupied due to the reduction of Mo atom. This is consistent with a study on model compounds.<sup>61</sup> The frontier orbitals of the intermediates, IM1 and IM2, have similar composition as those of the reactant, and are thus not shown.

It has been proposed that the reduction of Mo is possibly a Michaelis-like process through two single-electron steps<sup>62</sup> via a radical intermediate (formed by hydrogen rather than hydride transfer). However, a later experimental study on the reduction



**Figure 14.** Frontier orbitals for reactant and product from the charged model: Reactant: (A) HOMO. (B) LUMO. Product: (C) HOMO-1. (D) HOMO.

potential<sup>63</sup> using different substrates did not support such an assumption. We have tried to locate a radical intermediate (open-shell singlet), which could further be reduced from Mo(V) to Mo(IV) by a one-electron transfer along the Mo–O2 dative bond. The calculation produced a closed-shell compound and thus did not provide any evidence for an open-shell singlet. This gives indirect support to the concerted two-electron reduction of Mo atom.

**3.6.2. Fold Angle.** Another issue is the puckering of the Mo-dithiolene ring. The conformation of the dithiolene subunit has been observed to be slightly different in the oxidized and reduced states both in model compounds<sup>64</sup> and in the enzyme.<sup>15</sup> Enemark and co-workers<sup>61,65,66</sup> suggested that the dynamic variation of this unit may modulate the nucleophilic attack of the hydroxide group to the substrate and the following hydride transfer.

The fold angle (see Scheme 8) was monitored along pathway C (Figure 15). It increases upon C–O formation (by  $\sim 3^\circ$ ) and decreases again during the hydride transfer. This is not surprising when considering the trends in the partial charges on the reactive moiety (Figure 13): the changes in the fold angle are such that they help the cofactor to adapt to the changes in the charge distributions. The dithiolene ligand is an electron-rich group and can act as a “buffer” for

electrons. During C–O bond formation, the flow of electron density from Mo to C2 is compensated by additional interactions between  $p_z$ -orbitals of the sulfur atoms of the dithiolene-ligand (S7' and S8') and the empty  $d_{xy}$ -orbital of Mo (see Figure 14) which are strengthened by increasing the fold angle of the dithiolene ring.<sup>61</sup> By contrast, during the hydride transfer, the positive charge on the cofactor is partially neutralized, and the dithiolene now can recover its original conformation. The more planar structure of the dithiolene helps to delocalize the negative charge flowing to Mo from the Mo=SR1 bond during the hydride transfer. As the  $d_{xy}$ -orbital has become the HOMO, the interactions with the  $p_z$ -orbitals of the sulfur atoms of the dithiolene become repulsive, and the fold angle decreases.

#### 4. Discussion

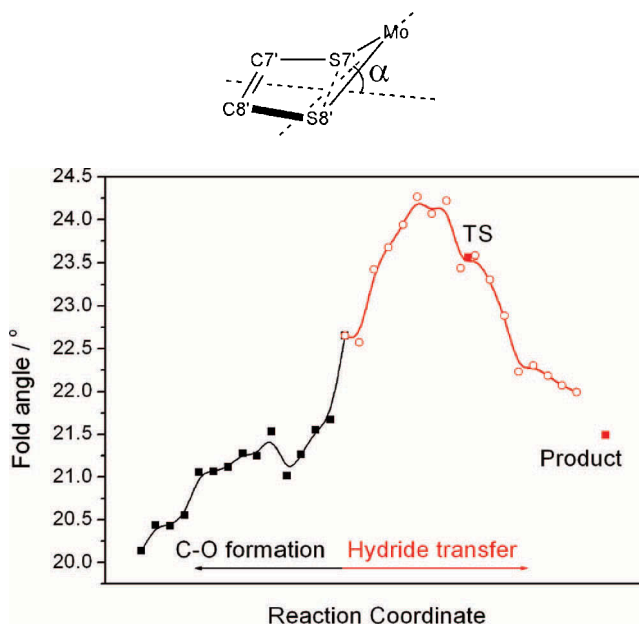
Our QM/MM results for the preferred pathway C suggest that Glu869 acts as Lewis base to deprotonate the labile hydroxide ligand and thus to activate the cofactor for nucleophilic attack at the substrate followed by hydride transfer. The oxidation of the substrate in the enzyme is thus predicted to be a Lewis base-catalyzed stepwise process.

In comparison with recent QM model studies, we note that Amano et al.<sup>13</sup> also favor a scenario where the reaction proceeds via a deprotonated cofactor; however, with formamide as substrate they only find a concerted pathway with a high barrier of ca. 35 kcal/mol, whereas QM/MM calculations with acetaldehyde as substrate yield a much more facile stepwise mechanism via a semiacetal intermediate that is stabilized by the electrostatic interaction with the environment. The structure of the intermediate and some features of our preferred pathway C are reminiscent of the stepwise mechanism with a protonated cofactor reported by Zhang and Wu;<sup>12</sup> however, in their QM model study, the stepwise mechanism is less favorable than the concerted one, contrary to the present QM/MM results. Generally speaking, the published QM model studies have not included the Glu869 residue which plays a crucial mechanistic role according to our QM/MM results and the experimental evidence for XHD.<sup>47</sup> Inclusion of Glu869 may pose problems at the QM level since this will create a dianionic QM model system, with strong repulsion between the two negatively charged entities (cofactor and Glu869), which may necessitate the use of geometric constraints. By contrast, such situations are handled quite naturally at the QM/MM level which properly accounts for the stabilization of charged active-site species by the protein environment.

#### 5. Conclusion

The current QM/MM calculations aim at understanding the reaction mechanism in aldehyde oxidoreductase, in particular the catalytic effect of Glu869. Comparison among pathways A, B, and C suggests that Glu869 plays a crucial role as a Lewis base which promotes the cofactor to a more active species. The QM/MM calculations indicate that the energetics of the oxidation of acetaldehyde change significantly when Glu869 is allowed to act as a Lewis base. By deprotonating the labile hydroxide group, Glu869 facilitates C–O bond formation by nucleophilic attack of the oxy-anion on the substrate, and the subsequent hydride transfer becomes the rate-determining step. The benefits from the proton transfer prior to the C–O bond formation are 2-fold: (a) the OM2

**Scheme 8.** Fold Angle  $\alpha$  Is Defined by Three Points: Mo, the Center of S7' and S8', and the Center of C7'=C8'



**Figure 15.** Fold angle along pathway C.

atom becomes more basic and hence its nucleophilicity increases; (b) the formation of an H-bond between HOM2 and O2 after the proton transfer perturbs the C2=O2 bond and induces an electron flow to O2, which makes C2 more electrophilic.

Two alternative scenarios based on pathway C have been considered for the reductive half-reaction. In pathway D there is an additional water molecule between Glu869 and acetaldehyde which may help to position the substrate and may move to the Mo center after the oxidation of the substrate, in order to facilitate product release and to act as oxygen source for the next turnover. The QM/MM energy profiles are similar for pathways C and D so that both seem feasible, with a slight edge for pathway C. By contrast, pathway E is less favorable. It involves a protonated Glu869 (unlike A–D) which precludes the activation of the cofactor by Glu869 acting as a Lewis base (as in C and D), and the computed effective barrier for pathway E is thus significantly higher than those for pathways C and D.

In summary, our QM/MM calculations describe the oxidation of acetaldehyde by the AOR enzyme as a Lewis base-catalyzed stepwise process. The initial deprotonation of the cofactor by Glu869 initiates a nucleophilic attack at the substrate followed by hydride transfer. We are not aware of any detailed experimental studies of this mechanism in AOR. Our QM/MM results are consistent with experimental evidence in related enzymes, i.e., on the crucial role of a nearby Glu residue in xanthine dehydrogenase<sup>47</sup> and oxygen isotope labeling in xanthine oxidase.<sup>18,28</sup> We hope that our

(61) Joshi, H. K.; Enemark, J. H. *J. Am. Chem. Soc.* **2004**, *126*, 11784–11785.

(62) Page, C. C.; Moser, C. C.; Chen, X. X.; Dutton, P. L. *Nature* **1999**, *402*, 47–52.

(63) Stockert, A. L.; Shinde, S. S.; Anderson, R. F.; Hille, R. *J. Am. Chem. Soc.* **2002**, *124*, 14554–14555.

(64) Stiefel, E. I.; Miller, K. F.; Bruce, A. E.; Corbin, J. L.; Berg, J. M.; Hodgson, K. O. *J. Am. Chem. Soc.* **1980**, *102*, 3624–3626.

current QM/MM predictions will trigger corresponding experimental studies in AOR.

**Acknowledgment.** We thank Prof. Yun-Dong Wu for suggesting this project and for valuable discussions. This work was

- 
- (65) Joshi, H. K.; Cooney, J. J. A.; Inscore, F. E.; Gruhn, N. E.; Lichtenberger, D. L.; Enemark, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3719–3724.
- (66) McNaughton, R. L.; Helton, M. E.; Cosper, M. M.; Enemark, J. H.; Kirk, M. L. *Inorg. Chem.* **2004**, *43*, 1625–1637.

supported by the Max-Planck-Gesellschaft. S.M. thanks the Fonds der Chemischen Industrie for a Kekulé scholarship.

**Supporting Information Available:** Technical details on system preparation and QM/MM setup. Energy scans, charge distributions, MD results, and detailed numerical data. Complete refs 32 and 45. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA805938W



# SUPPORTING INFORMATION

## Reductive Half-reaction of Aldehyde Oxidoreductase toward Acetaldehyde: A Combined QM/MM Study

*Sebastian Metz, Dongqi Wang, and Walter Thiel\**

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

[thiel@mpi-muelheim.mpg.de](mailto:thiel@mpi-muelheim.mpg.de)

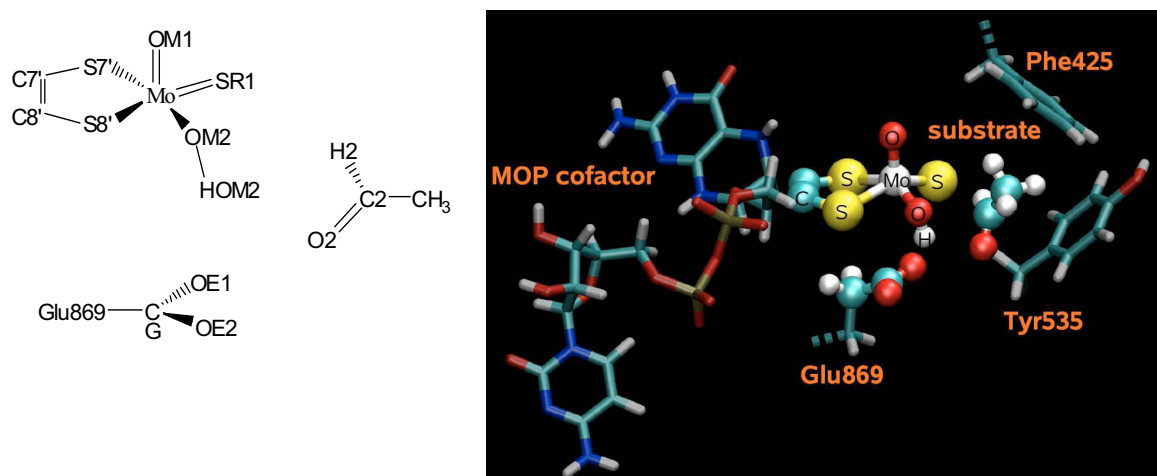


## 1. System preparation using CHARMM C29b2

The active MM region contains the following residues (notation according to PDB: 1VLB):

HSP365 ILE390 GLY391 HSE395 THR420 PHE421 GLY422 TYR423 LYS424 PHE425 SER426  
LYS459 ARG460 PHE494 LEU497 LEU498 ARG501 SER530 ALA531 PHE532 ARG533 GLY534  
TYR535 GLY536 ALA537 PRO538 TYR622 GLY623 SER624 GLY625 LEU626 GLU651 HSE653  
GLN655 PRO694 SER695 GLY696 GLY697 SER698 ARG699 GLN700 MET767 GLN807 GLN814  
VAL867 GLY868 GLU869 LEU870 PRO871 LEU872 PCD AALD

**Figure S1.** Reaction center and lipophilic active-site environment.



The atoms labels are taken from the crystal structure 1VLB. A more precise specification makes use of the format Resname:AtomName, for example, the atom O2 in acetaldehyde is denoted as AALD:O2.



## A. Protonation scheme:

Histidine:

Nδ: 219, 488, 855;

Nη: 86, 395, 526, 611, 653, 830, 876;

Both N: 325, 365, 386, 393, 445, 647, 666, 752

In the neutral model, residues are also protonated if:

1. there is no H-bond donor in the surrounding within a range of 5 Å, and
2. there is no counterion or charge pair in this region.

This leads to the following residues being protonated in the neutral model:

Asp 379, 587; Glu 12, 83, 273, 298, 306, 319, 334, 341, 406, 579;

For the model used in pathway E, we protonated Glu 869 and therefore deprotonated Glu 334.

## B. Partial charges for force field calculations:

The partial charges on the atoms in the cofactor (PCD) and in the  $[2\text{Fe-2S}]^{2+}$  cluster were derived from DFT calculations using the B3LYP functional and the following basis set: Lanl2DZ for Mo, Fe; 6-31G\* for the other atoms.

The following partial charges were adopted:

```
! Mo(=O) (=S) (-OH) (-1,2-dithio-ene)
GROUP ! -1
ATOM  MO      MO      0.947  !
ATOM  OM1     OMO     -0.552  !          OM1
ATOM  OM2     OH1     -0.923  !          ||
ATOM  HOM2    H       0.503  !      S7' --- Mo==OR1
ATOM  OR1     OMO     -0.271  !      /      / \
ATOM  S8'     SS      -0.172  ! ---C7' S8' OM2
ATOM  S7'     SS      -0.055  !      \ \ /      |
ATOM  C8'     CE1     -0.235  !      ---C8' HOM2
ATOM  C7'     CE1     -0.242  !
```

```

!2Fe-2S cluster
RESI FEST 2.00 ! Fe(III)-Fe(III)
GROUP ! [FeS]1+
ATOM FE1 FE 1.064 !
ATOM S1 SF -0.064 !
GROUP ! [FeS]1+
ATOM FE2 FE 1.064 !
ATOM S2 SF -0.064 !

```

The newly defined  $[2\text{Fe}-2\text{S}]^{2+}$  residue is bonded to four cysteine residues through Fe-S bonds. The links are created by a CHARMM patching procedure. The cysteine residues bonded to the cluster are:

```

FES908: FE1, 100, 139; FE2, 103, 137
FES909: FE1, 40, 45; FE2, 48, 60

```

The patch to create the link between Cysteine and FE1 is defined as follows:

```

PRES FES1 -1.00 ! patch for FE1-Cys. Must be 1-CYS and 2-FEST.
GROUP ! HB1
ATOM SG SS -0.80 ! |
ATOM CB CS -0.38 ! --CB--SG
ATOM HB1 HA 0.09 ! | \
ATOM HB2 HA 0.09 ! HB2 (FE1-S1)
DELETE ATOM 1HG1
BOND 1SG 2FE1
ANGLE 1CB 1SG 2FE1 1SG 2FE1 2S1
DIHE 1HB1 1CB 1SG 2FE1 1HB2 1CB 1SG 2FE1
DIHE 1CA 1CB 1SG 2FE1
DIHE 1CB 1SG 2FE1 2S1

```

The patch for the links to FE2 is analogous and will thus not be shown here.

Reminder: The topology files for the Mo-containing group and the  $\text{Fe}_2\text{S}_2$  cluster are used only for the preparation of the QM/MM model systems. They are not recommended for classical MM production calculations without extensive validation.

### C. System setup of totally solvated model

The protonation states are same as in the partially solvated neutral model. A 10 Å water shell was used to solvate the whole enzyme. The outer layer of 5 Å was fixed during the molecular dynamic simulation to prevent the inner free water from escaping.

The solvated systems were then relaxed by performing energy minimizations and molecular dynamic (MD) simulations at the MM level using the CHARMM27 force field as implemented in the CHARMM program.

During the classical energy minimizations and the MD simulations, the non-hydrogen atoms of pterin, the pyrano group (dithio-Mo(=O)(=S)(OH)-ene) and the C=O group of the substrate were fixed. The equilibrated systems contained 34565 atoms, including 6969 TIP3P water molecules and 5 ions (two  $\text{Mg}^{2+}$  and three  $\text{Cl}^-$ ).

## 2. QM/MM setup

Programs: Turbomole + DL\_POLY embedded in ChemShell;

Functional: B3LYP;

Basis set B1: Lanl2DZ + f polarization for Mo; Lanl2DZ + d polarization for S; 6-31+G\*\* for the rest;

Basis set B2: TZVPP for C, O, TZP for H, TZV'+P(21/1), and ECP-28-MWB-TZVext+P(f) for Mo.

Initial coordinates: aldehyde oxidoreductase (PDB Code: 1VLB)

### A. QM region:

R0: 18 atoms:  $\text{Mo}(\text{S}_2\text{C}_2\text{H}_2)(=\text{O})(\text{OH})(=\text{S}) + \text{CH}_3\text{CHO}$  (AALD)

R1(p): 26 (27) atoms: R0 + Glu869 (  $-\text{CH}_2\text{CO}_2^-$  /  $-\text{CH}_2\text{COOH}$  )

R2: 29 atoms: R0 + Glu869 (  $-\text{CH}_2\text{CO}_2^-$  ) +  $\text{H}_2\text{O}$

## **B. MM region:**

(a) Charged model: There are 891 atoms in the active MM region which belong to the following residues (within 10 Å around the Mo-center):

Enzyme and PCD + AALD911 852 atoms:

HSP365 ILE390 GLY391 HSE395 THR420 PHE421 GLY422 TYR423 LYS424 PHE425 SER426  
LYS459 ARG460 PHE494 LEU497 LEU498 ARG501 SER530 ALA531 PHE532 ARG533 GLY534  
TYR535 GLY536 ALA537 PRO538 TYR622 GLY623 SER624 GLY625 LEU626 GLU651 HSE653  
GLN655 PRO694 SER695 GLY696 GLY697 SER698 ARG699 GLN700 MET767 GLN807 GLN814  
VAL867 GLY868 GLU869 LEU870 PRO871 LEU872 PCD AALD

Crystal water: 21 atoms

CRYW197 CRYW220 CRYW369 CRYW370 CRYW371 CRYW375 CRYW376

Solvent: 18 atoms

(b) Neutral model: 843 atoms in the active MM region:

Enzyme and PCD + AALD911 804 atoms:

HSP365 ILE390 HSE395 THR420 PHE421 GLY422 TYR423 LYS424 PHE425 SER426 ARG460  
PHE494 LEU498 ARG501 SER530 ALA531 PHE532 ARG533 GLY534 TYR535 GLY536 ALA537  
PRO538 TYR622 GLY623 SER624 GLY625 LEU626 GLU651 HSE653 GLN655 PRO694 SER695  
GLY696 GLY697 SER698 ARG699 GLN700 MET767 GLN807 GLN814 VAL867 GLY868 GLU869  
LEU870 PRO871 LEU872 PCD AALD

Crystal water: 24 atoms

CRYW197 CRYW209 CRYW220 CRYW369 CRYW370 CRYW371 CRYW375 CRYW376

Solvent: 15 atoms

Table 1 collects selected geometry parameters for the Mo-cofactor from our QM/MM optimized reactant structure and crystal structure 1VLB. Data for two xanthine dehydrogenases, 1VDV and 1V97 in their *sulfo* form, are also given.

**Table S1.** Comparison of selected geometry parameters<sup>(a)</sup> of the Mo-cofactor between the QM/MM optimized resting states and crystal structures

	Mo-OM1	Mo-OM2	Mo-SR1 <sup>(b)</sup>	Mo-S7'	Mo-S8'	S7'-Mo-S8'	Fold angle
<b>Neutral</b>	1.712	1.897	2.157	2.484	2.487	78.4	22.7
<b>Charged</b>	1.732	1.879	2.169	2.490	2.465	78.2	17.5
<b>1VLB</b>	1.740	1.989	1.791	2.409	2.495	79.5	16.6
<b>1VDV<sup>(c)(d)</sup></b>	1.730	1.968	2.174	2.440	2.414	83.6	25.8
<b>1V97<sup>(d)(e)</sup></b>	1.689	2.036	2.406	2.418	2.390	84.4	22.0

(a) Distances in Å, and angles in degree.

(b) In 1VLB, this atom is resolved as OR1 (desulfo-form).

(c) 1VDV: Bovine milk xanthine dehydrogenase (XDH), Y-700 bound form, resolution: 1.98 Å.

(d) The values shown are averaged from the two cofactors. In the XDHs, the atoms have different labels (not shown here).

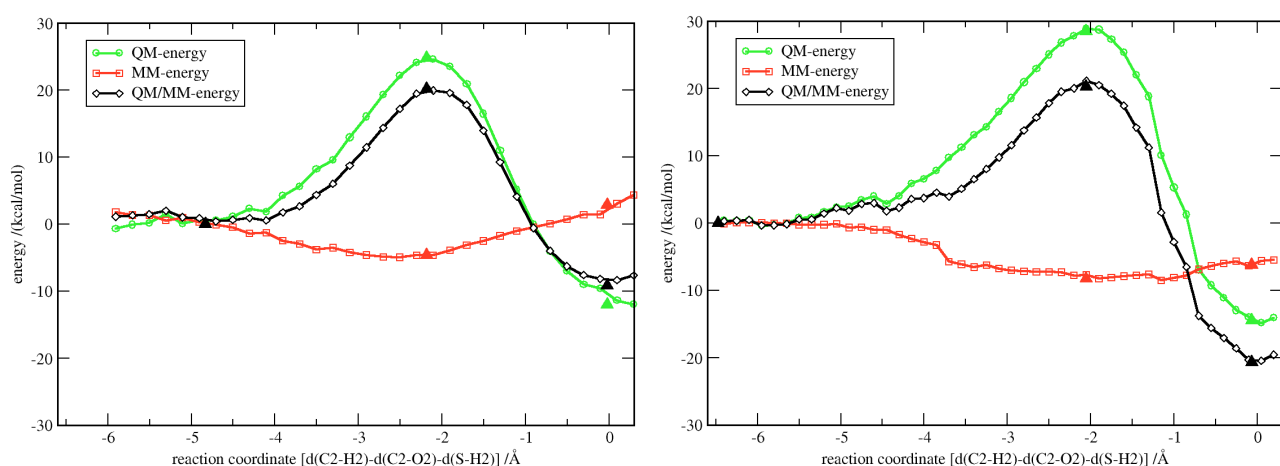
(e) 1V97: Bovine milk xanthine dehydrogenase FYX-051 bound form, resolution: 1.94 Å

### 3 Results and discussions:

Within the energy profiles, all points from constrained optimization are represented by unfilled symbols. The filled triangles represent the fully optimized stationary points. The QM energies include the electrostatic interactions with the MM point charges.

#### 3.1 Pathway A: concerted reaction

a) Neutral model, partial solvation

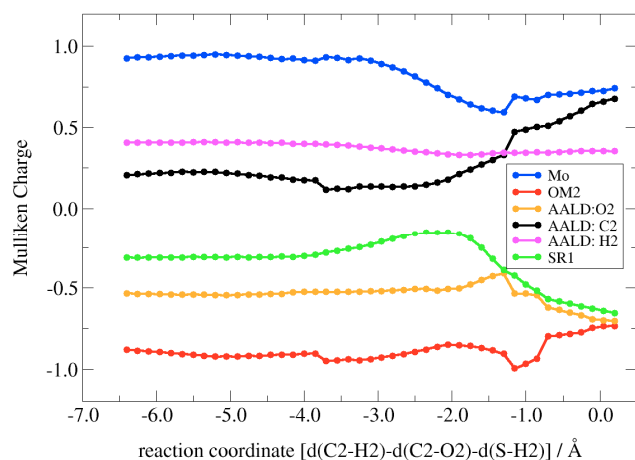


**Figure S2.** Energy profiles of pathway A from R0/B1 (left) and R1/B1 calculations (right).

**Table S2.** Calculated energies for optimized stationary points along pathway A (R0/B1 und R1/B1) (neutral model, in kcal mol<sup>-1</sup>)

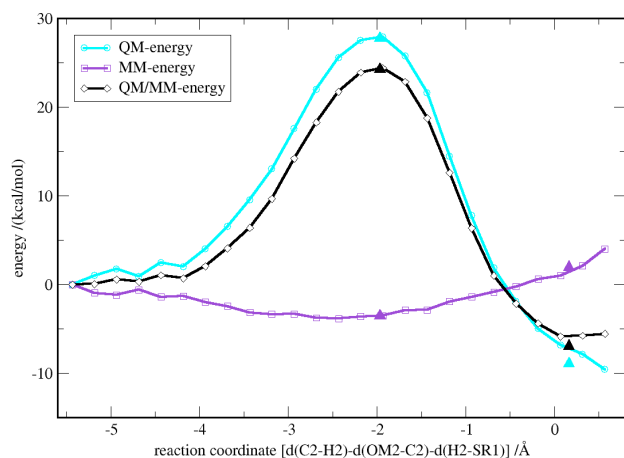
	R0/B1			R1/B1		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactants <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0
TS	24.8	-4.6	20.2	28.5	-8.3	20.2
Product	-12.0	2.8	-9.2	-14.5	-6.2	-20.7

<sup>a</sup> absolute energies (a.u.): R0: -480.23804 -89.65157 -569.88961, R1: -708.87585 -88.85496 -797.73081



**Figure S3.** Mulliken charges for pathway A from R1/B1 calculations.

b) Total solvation scheme



**Figure S4.** Energy profiles of pathway A from R0/B1.

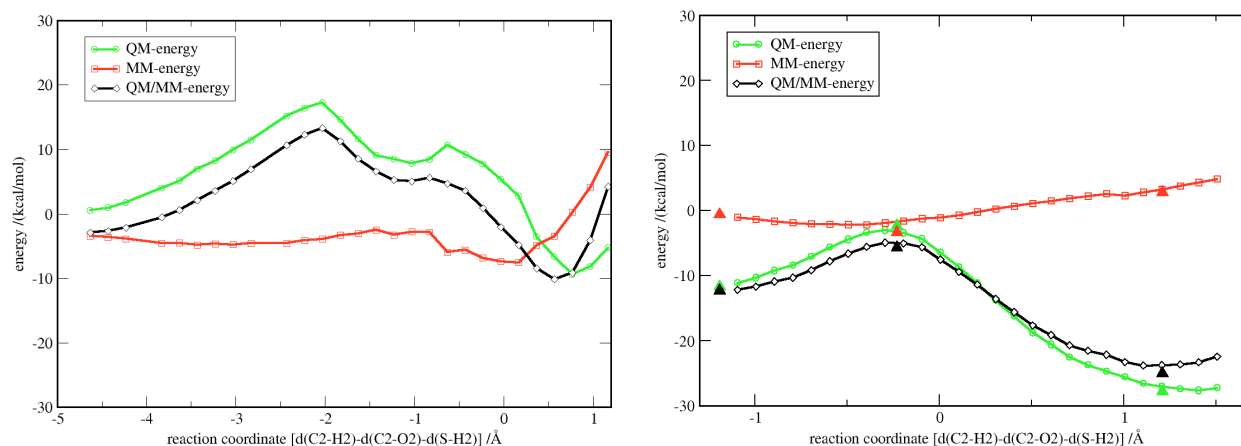
**Table S3.** Calculated energies for optimized stationary points along pathway A (R0/B1) (total solvation scheme, in kcal mol<sup>-1</sup>)

	R0/B1		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactants <sup>a</sup>	0.0	0.0	0.0
TS	27.8	-3.5	24.2
Product	-8.9	2.0	-6.9

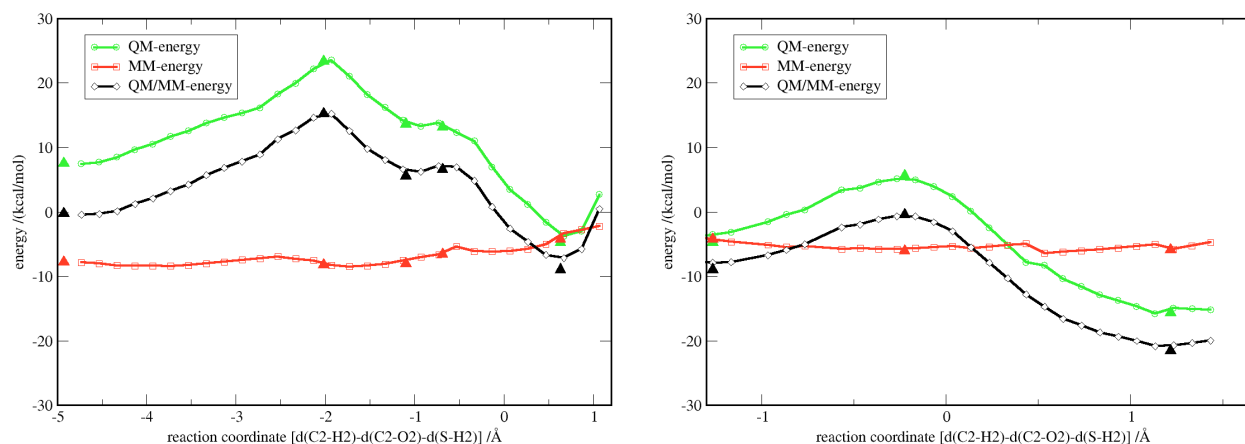
<sup>a</sup> absolute energies (a.u.): -480.19352 -154.40136 -634.59488

### 3.2 Pathway B: initial dative bond between Mo and substrate

a) Neutral model, partial solvation



**Figure S5.** Energy profiles along pathway B for neutral model from R0 calculations.



**Figure S6.** Energy profiles along pathway B for neutral model from R1 calculations.

**Table S4.** Calculated energies for optimized stationary points along pathway B (R0/B1 und R1/B1) (neutral model, in kcal mol<sup>-1</sup>)

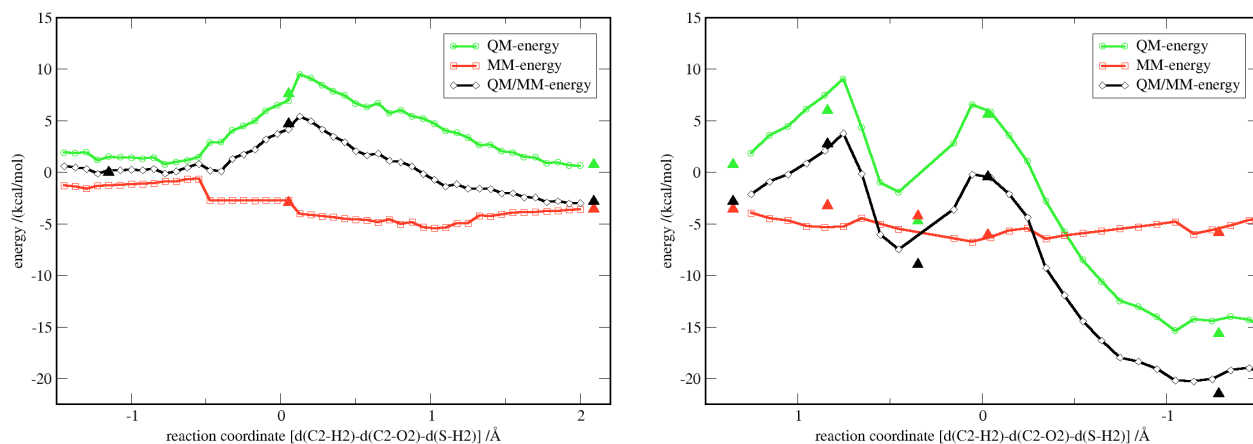
	R0/B1			R1/B1			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactants <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	17.0	-1.5	15.5	15.8	-0.4	15.4	Nucleophilic attack
IM1	9.5	-2.3	7.3	6.1	-0.3	5.8	
TS2	10.4	-2.6	7.8	5.6	1.2	6.8	Reorientation of OM2-HOM2
IM2	-11.9	2.7	-9.2	-12.3	3.6	-8.7	
TS3	-2.6	0.0	-2.6	-1.9	1.7	-0.2	Hydride transfer
Product	-28.0	6.1	-21.9	-23.2	1.9	-21.3	

<sup>a</sup> absolute energies (a.u.): R0:-480.23773 -89.65688 -569.89461, R1:-708.86384 -88.86734 -797.73117



### 3.3 Pathway C: Glu869 promoted pathway

a) Neutral model, partial solvation

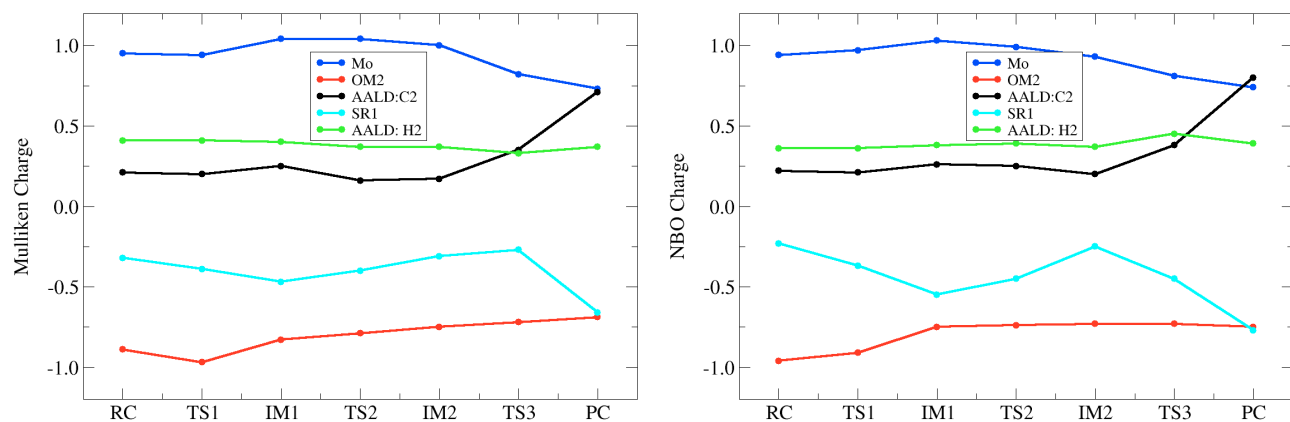


**Figure S7.** Energy profiles for pathway C from R1 calculations for neutral model.

**Table S5.** Calculated energies (R1/B1) for optimized stationary points along pathway C (neutral model, in kcal mol<sup>-1</sup>)

	$E_{QM}$	$E_{MM}$	$E_{QM/MM}$	Remark
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS1	7.6	-2.9	4.7	Proton transfer
IM1	0.7	-3.6	-2.8	
TS2	6.0	-3.2	2.8	C-O formation
IM2	-4.7	-4.2	-8.9	
TS3	5.6	-6.1	-0.5	Hydride transfer
Product	-15.6	-5.9	-21.5	

<sup>a</sup> absolute energies (a.u.): -708.87585 -88.85496 -797.73081



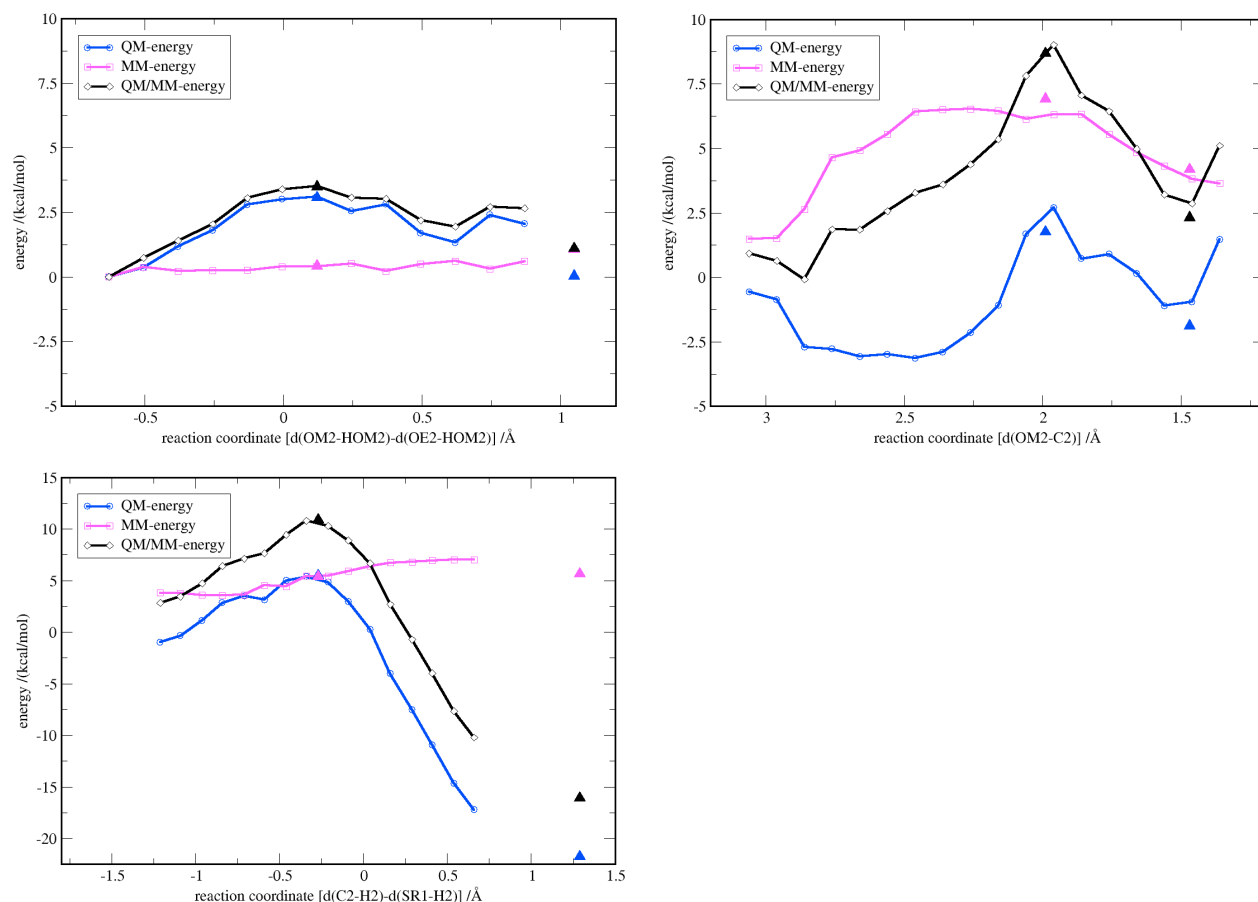
**Figure S8.** Mulliken and NBO charges for stationary points along pathway C from R1 calculations.

**Table S6.** Comparison of different functionals for optimized stationary points (R1/B2) of pathway C (neutral model, in kcal mol<sup>-1</sup>)

	BP86			BLYP			B3LYP			BHLYP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
TS1	3.6	-1.8	1.8	5.3	-1.9	3.4	7.8	-3.2	4.6	9.6	-3.0	6.7
IM1	0.3	-2.5	-2.3	1.8	-2.7	-0.8	2.3	-4.0	-1.7	2.4	-3.6	-1.2
TS2	3.3	-3.0	0.2	7.2	-3.1	4.1	9.1	-4.2	4.9	12.7	-3.6	9.1
IM2	-4.4	-4.0	-8.4	0.6	-3.9	-3.3	-1.1	-4.5	-5.6	-3.9	-3.6	-7.5
TS3	1.9	-5.5	-3.6	10.8	-5.7	5.1	8.8	-6.7	2.1	7.5	-6.5	1.0
Product	-8.1	-5.3	-13.3	-1.2	-5.6	-6.8	-12.5	-6.5	-19.0	-29.7	-5.9	-35.6

<sup>a</sup> absolute energies (a.u.):  
BP86: -1874.82321 -88.85850 -1963.68171  
BLYP: -1874.34554 -88.85651 -1963.20206  
B3LYP: -1874.01443 -88.85512 -1962.86955  
BHLYP: -1873.98457 -88.85559 -1962.84016

b) Charged model, partial solvation



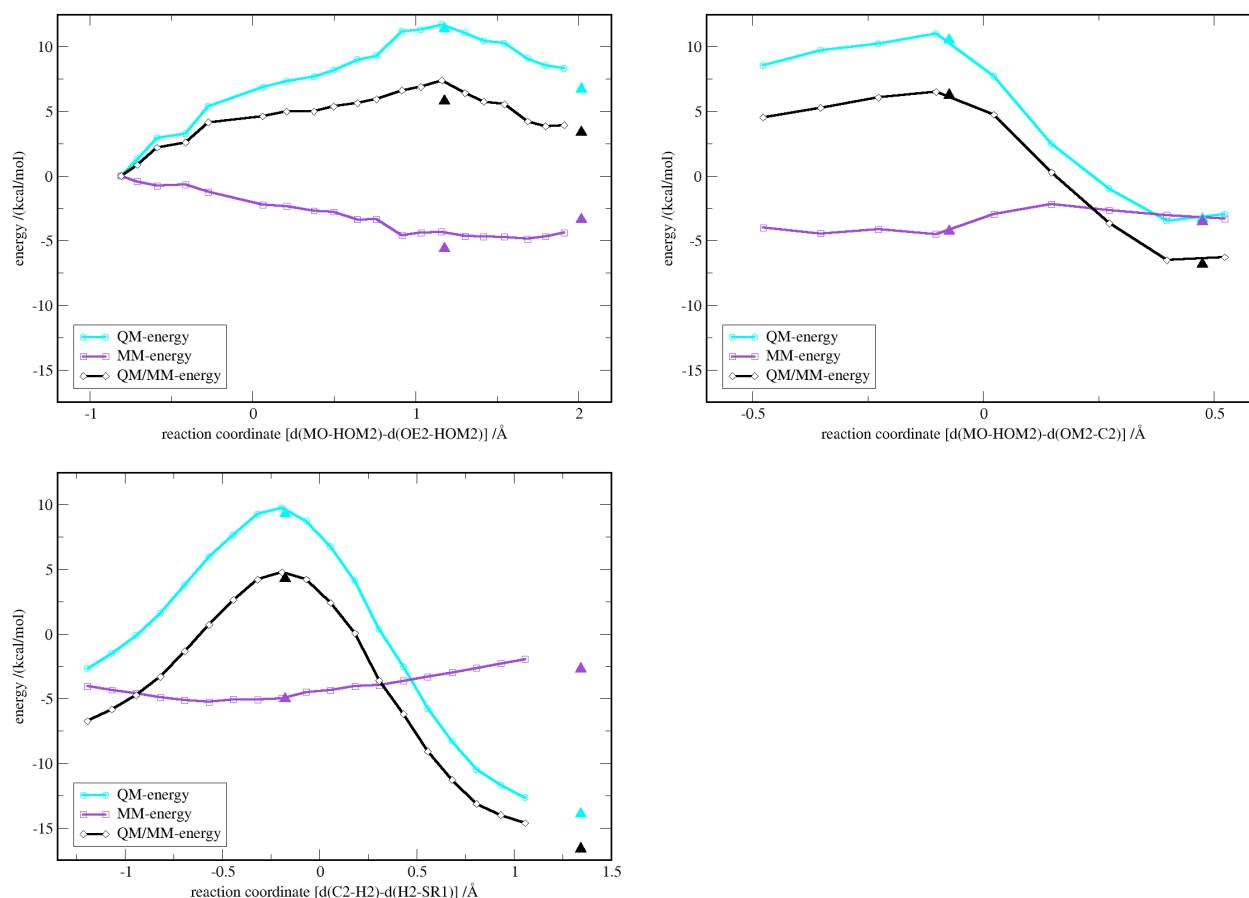
**Figure S9.** Energy profiles (R1/B1) of the charged model (partial solvation scheme) for pathway C.

**Table S7.** Calculated energies (R1/B1) for optimized stationary points along pathway C and QM/MM energies for single point calculations with basis set B2 (charged model, in kcal mol<sup>-1</sup>)

	B3LYP/B1			SP-B3LYP/B2			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	3.1	0.4	3.5	13.2	-5.9	7.3	Proton transfer
IM1	0.0	1.1	1.1	9.4	-3.5	5.8	
TS2	1.8	6.9	8.7	13.7	-4.7	9.0	C-O formation
IM2	-1.9	4.2	2.3	-1.7	-3.5	-5.2	
TS3	5.5	5.4	10.9	12.4	-5.3	7.2	Hydride transfer
Product	-21.8	5.7	-16.1	-8.4	-3.0	-11.3	

<sup>a</sup> absolute energies (a.u.):  
 B1: -708.51878 -89.40269 -797.92147,  
 B2: -1873.65751 -89.40209 -1963.05960

c) Total solvation scheme



**Figure S10.** Energy profiles (R1/B1) of the total solvation scheme for pathway C.

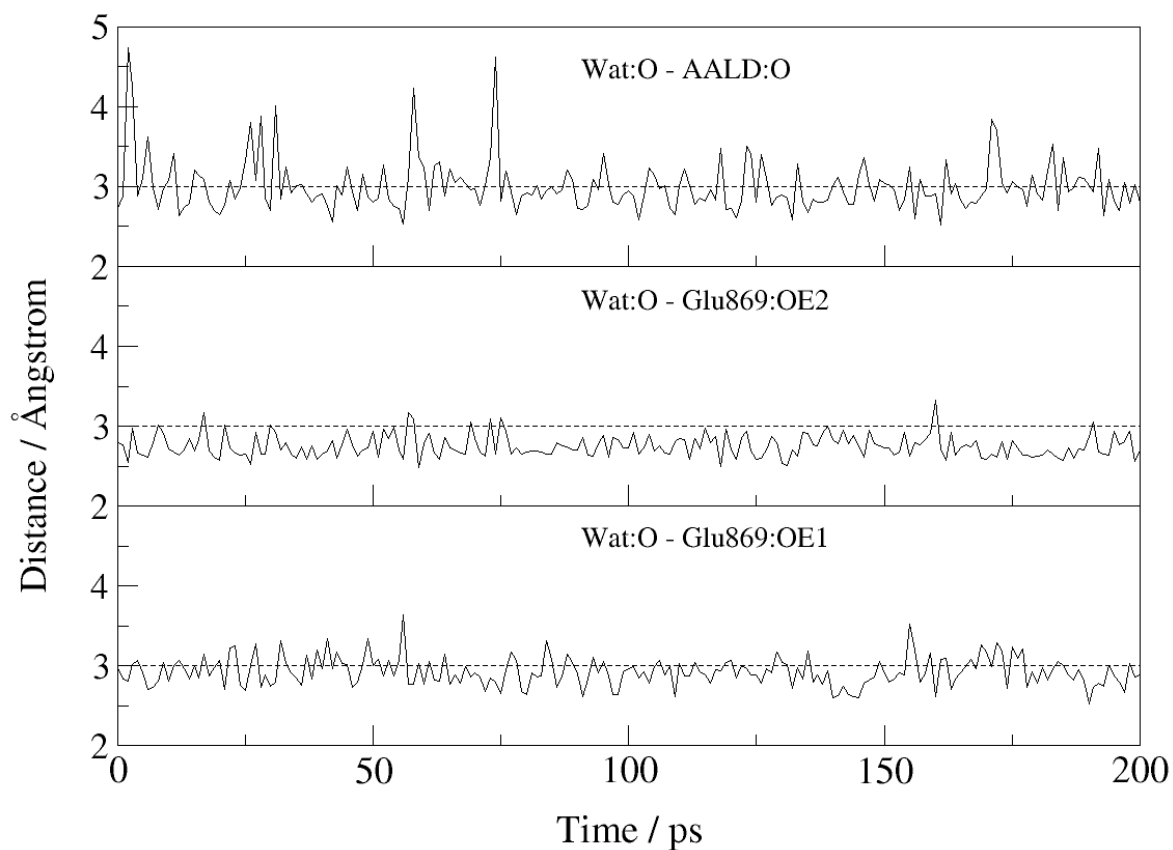
**Table S8.** Optimized stationary points (R1/B1) for pathway C (total solvation scheme, in kcal mol<sup>-1</sup>)

	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	Remark
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS1	11.4	-5.6	5.8	Proton transfer
IM1	6.7	-3.4	3.4	
TS2	10.5	-4.3	6.3	C-O formation
IM2	-3.3	-3.5	-6.8	
TS3	9.3	-5.0	4.3	Hydride transfer
Product	-13.9	-2.7	-16.6	

<sup>a</sup> absolute energies (a.u.): -708.82242 -154.18240 -863.00482

### 3.4 Pathway D: Glu869 promoted pathway with one additional water molecule

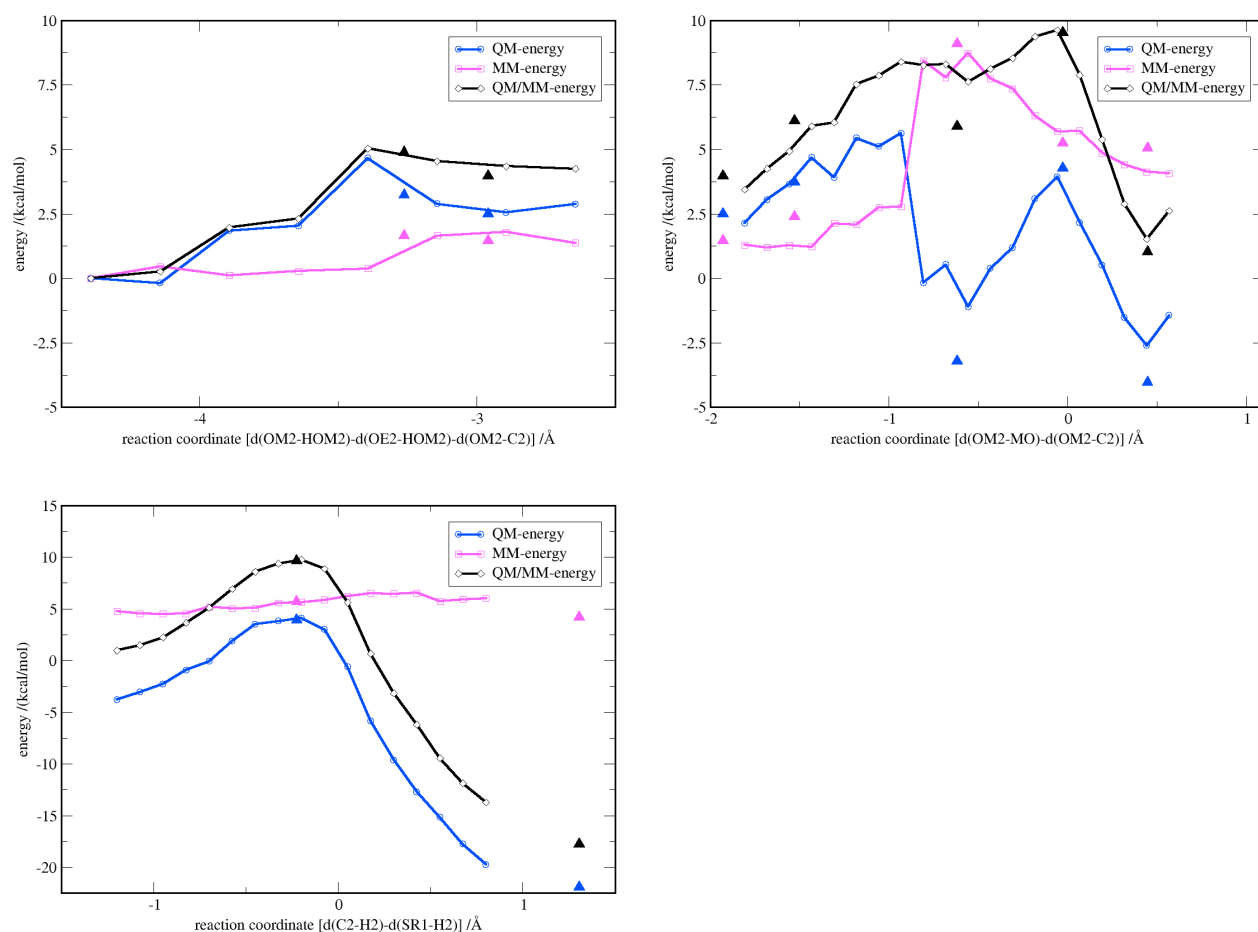
a) Charged model, partial solvation

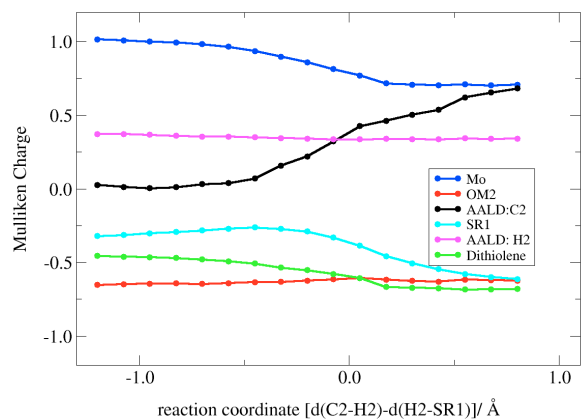


**Figure S11.** Monitoring the position of the additional water molecule relative to the substrate and Glu869 during 200 ps of classical molecular dynamics.

**Table S9.** Optimized stationary points (R2/B1) for pathway D (charged model, in kcal mol<sup>-1</sup>)

	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	Remark
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS1	3.2	1.7	4.9	Proton transfer
IM1	2.5	1.5	4.0	
TS2	3.7	2.4	6.1	H-bond rearrangement
IM2	-3.2	9.1	5.9	
TS3	4.3	5.2	9.5	C-O formation
IM3	-4.0	5.1	1.0	
TS4	3.9	5.7	9.6	Hydride transfer
Product	-21.9	4.2	-17.7	

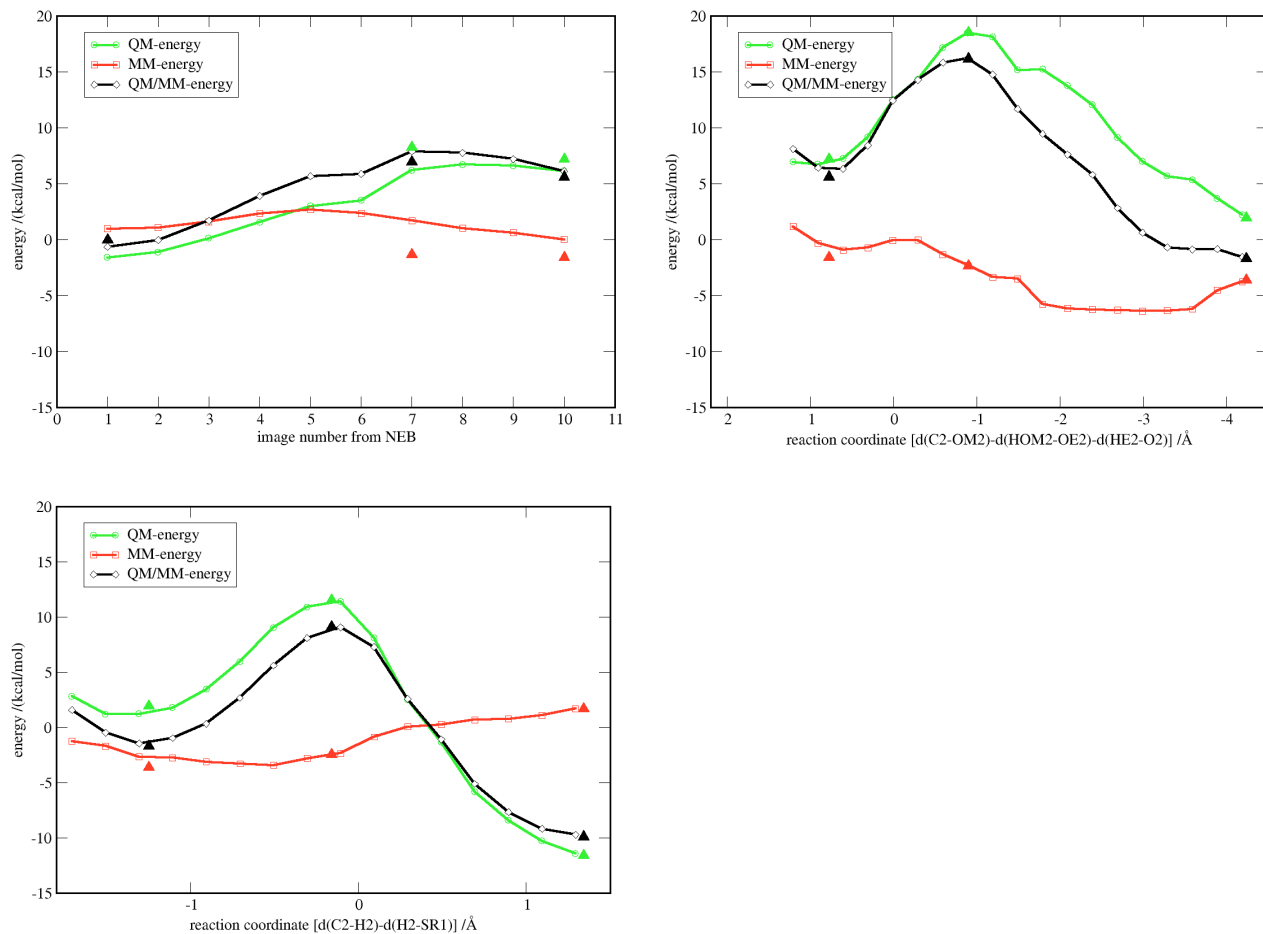
<sup>a</sup> absolute energies (a.u.): -784.92624 -89.36574 -874.29197**Figure S12.** Energy profiles for pathway D, charged model, partial solvation scheme; R2/B1.



**Figure S13.** Mulliken charges for the hydride transfer step along pathway D for the charged model, partial solvation scheme; R2/B1.

### 3.5 Pathway E: protonated Glu869

#### a) Neutral model, partial solvation



**Figure S14.** Energy profiles for pathway E, the partial solvation scheme; R1p/B1.

**Table S11.** Optimized stationary points (R1p/B1) for pathway E (neutral model, in kcal mol<sup>-1</sup>)

	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	Remark
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS1	8.3	-1.3	6.9	Substrate reorientation
IM1	7.2	-1.6	5.6	
TS2	18.5	-2.3	16.2	Proton transfer and C-O formation
IM2	1.9	-3.6	-1.7	
TS3	11.5	-2.4	9.1	Hydride transfer
Product	-11.6	1.7	-9.9	

<sup>a</sup> absolute energies (a.u.): -709.29668 -88.89489 -798.19158

### Complete references 32 and 45

[32] CHARMM27 force field:

(a) MacKerell, A.D., Jr.; Feig, M.; Brooks, C.L., III J. Comput. Chem. 2004, 25, 1400-1415.

(b) Foloppe, N.; Mackerell, A. D., Jr. J. Comput. Chem. 2000, 21(2), 86-104

(c) MacKerell, Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R.L.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, 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.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J. Phys. Chem. B, 1998, 102, 3586-3616.

[45] (a) 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.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E; Sjovoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. J. Mol. Struc. (Theochem) 2003, 632, 1.

(b) ChemShell is a modular QM/MM program developed in the European QUASI project under the coordination of P. Sherwood (see <http://www.chemshell.org>).





## Paper II

*Reductive Half-Reaction of Aldehyde Oxidoreductase  
toward Acetaldehyde: A Combined QM/MM Study*

Johannes M. Dieterich, Hans Joachim Werner,  
Ricardo A. Mata,  
**Sebastian Metz**, and Walter Thiel

*J. Chem. Phys.*, submitted.



# Reductive Half-Reaction of Aldehyde Oxidoreductase toward Acetaldehyde: Ab Initio and Free Energy QM/MM Calculations

Johannes M. Dieterich and Hans-Joachim Werner  
*Institut für Theoretische Chemie, Universität Stuttgart,  
Pfaffenwaldring 55, D-70569 Stuttgart, Germany*

Ricardo A. Mata  
*Institut für Physikalische Chemie,  
Tammannstrasse 6, D-37077 Göttingen, Germany*

Sebastian Metz and Walter Thiel  
*Max-Planck-Institut für Kohlenforschung,  
Kaiser-Wilhelm-Platz 1,  
D-45470 Mülheim an der Ruhr, Germany*

Energy and free energy barriers for acetaldehyde conversion in aldehyde oxidoreductase are determined for three reaction pathways using quantum mechanical/molecular mechanical (QM/MM) calculations on the solvated enzyme. Ab initio single-point QM/MM energies are obtained at the stationary points optimized at the DFT(B3LYP)/MM level. These ab initio calculations employ local correlation treatments [LMP2 and LCCSD(T0)] in combination with augmented triple- and quadruple-zeta basis sets, and the final coupled cluster results include MP2-based corrections for basis set incompleteness and for the domain approximation. Free energy perturbation (FEP) theory is used to generate free energy profiles at the DFT(B3LYP)/MM level for the most important reaction steps, by sampling along the corresponding reaction paths using molecular dynamics. The ab initio and FEP QM/MM results are combined to derive improved estimates of the free energy barriers which differ from the corresponding DFT(B3LYP)/MM energy barriers by about 3 kcal mol<sup>-1</sup>. The present results confirm the qualitative mechanistic conclusions from a previous DFT(B3LYP)/MM study. Most favorable is a three-step Lewis base catalyzed mechanism with an initial proton transfer from the cofactor to the Glu869 residue, a subsequent nucleophilic attack that yields a tetrahedral intermediate (IM2), and a final rate-limiting hydride transfer. The competing metal center activated pathway has the same final step, but needs to overcome a higher barrier in the initial step on the route to IM2. The concerted mechanism has the highest free energy barrier and can be ruled out. While confirming the qualitative mechanistic scenario proposed previously on the basis of DFT(B3LYP)/MM energy profiles, the present ab initio and FEP QM/MM calculations provide corrections to the barriers that are important when aiming at high accuracy.

## I. INTRODUCTION

Aldehyde oxidoreductase (AOR), aldehyde oxidase (AO), xanthine dehydrogenase (XDH), and xanthine oxidase (XO) are members of the xanthine oxidase family, which have a molybdopterin cofactor as common structural feature.<sup>1,2</sup> It has been proven for XO that the central molybdenum atom in this cofactor carries two thiolate ligands (Mo-S), one terminal oxygen ligand (Mo=O), one terminal sulfido ligand (Mo=S), and one hydroxy ligand (Mo-OH).<sup>3</sup> It is generally assumed that all these enzymes contain the same active center,<sup>2</sup> especially since the replacement of Mo=S against the alternative M=O moiety leads to complete loss of activity in XO.<sup>1,4</sup> However, there are recent indications that AOR may also work with an equatorial oxo ligand.<sup>5</sup> In this study, we adopt the commonly accepted form of the cofactor with a sulfido ligand.

The molybdopterine cofactor within the xanthine oxidase family is known to catalyze oxidative hydroxylation reactions in a variety of aldehyde and purine substrates.<sup>6-12</sup> This function can also be exploited to ac-

tivate or clear drugs and xenobiotics, as has been demonstrated for human liver AO.<sup>13-16</sup> The AOR enzyme investigated presently shows the highest activity when acting on acetaldehyde or benzaldehyde.<sup>7</sup> As all other members of the xanthine oxidase family, AOR uses a water molecule rather than O<sub>2</sub> as source of oxygen<sup>2</sup> and produces reductive equivalents rather than consuming them.

Quantum mechanical/molecular mechanical (QM/MM) approaches have emerged as the method of choice for describing chemical reactions that occur within the active site of an enzyme.<sup>17,18</sup> They provide valuable insight into enzymatic reaction mechanisms and may discriminate between different conceivable pathways. A common procedure for this purpose is to explore the relevant parts of the potential energy surface by QM/MM calculations and to optimize the structures of representative minima and transition states for various pathways. In the case of AOR and XO, this has recently been done in two detailed mechanistic studies at the DFT/CHARMM level,<sup>19,20</sup> using density functional theory (DFT) for the QM part and the CHARMM force field for the MM part.

There are several ways to improve on such standard

QM/MM results. Two crucial issues are the accuracy of the chosen QM component and the role of free energy corrections. In this paper, we address both these issues for AOR. First, we go beyond the DFT(B3LYP) level employed previously,<sup>19</sup> by using local correlated ab initio methods for the QM region, and perform single point ab initio QM/MM calculations at the available optimized DFT/MM structures to determine more accurate energies, up to coupled cluster quality. Secondly, since thermodynamics and kinetics are governed by free energies, we use free energy perturbation (FEP) theory in an approximate QM/MM implementation<sup>21–23</sup> to obtain free energy profiles for the most relevant reaction steps at the DFT(B3LYP)/CHARMM level. These two sets of calculations provide corrections to the previous results which should allow us to determine high-quality free energy QM/MM barriers.

## II. QM/MM SETUP AND GEOMETRY OPTIMIZATIONS

The chosen QM/MM approach has been described in much detail in a recent study on the reductive half-reaction of AOR toward acetaldehyde,<sup>19</sup> so that we only summarize the essential points here. Starting from the experimental crystal structure,<sup>24</sup> the system was prepared following standard protocols that involve a series of classical energy minimizations and molecular dynamics (MD) simulations with the CHARMM27 force field.<sup>25</sup> In the subsequent QM/MM optimizations of selected MD snapshots, the QM part was treated by DFT using the B3LYP<sup>26–31</sup> density functional as implemented in TURBOMOLE,<sup>32</sup> while the CHARMM force field was run through the DL\_POLY program<sup>33</sup> to handle the MM part of the system. Geometries were optimized using a linear-scaling microiterative algorithm working in hybrid delocalized coordinates,<sup>34</sup> as implemented in the ChemShell QM/MM package.<sup>35,36</sup> The QM region consisted of the acetaldehyde substrate, the essential part of the cofactor, and the deprotonated Glu869 residue (see Figure 1). The active region included all residues within a radius of 10 Å around the reaction center (843 atoms in total), whereas all other atoms were kept fixed. The following basis set (denoted as B1) was used during optimization: Lanl2DZ<sup>37</sup> + f-polarization for Mo,<sup>38</sup> Lanl2DZ<sup>39</sup> + d-polarization<sup>40</sup> for S, and 6-31+G\*\*<sup>41,42</sup> for the remaining atoms (H,C,O). Optimized structures were determined for all minima and transition states of the three mechanisms considered,<sup>19</sup> namely the one-step mechanism, the metal-center activated mechanism, and the Lewis base catalyzed mechanism (see Figure 1).

## III. METHODS

### A. Single-point ab initio QM/MM calculations

Several high-level quantum mechanical calculations were performed at the optimized stationary points obtained previously (see above). All calculations employed electrostatic embedding (by including a lattice of point charges in the one-electron Hamiltonian), with the MM energy terms being added *a posteriori*. The goal was to obtain a converged description of the electronic energy for the proposed structures with respect to the one-particle as well as N-particle expansions, or at least reliable error estimates. This strategy disregards other sources of error in the quantum chemical treatment: for example, relativistic effects are only accounted for by the use of a relativistic pseudopotential for molybdenum, and are neglected for the other atoms (including sulfur).

Given the size of the QM region in AOR, correlated ab initio calculations are computationally very demanding. Therefore we applied the local correlation methods available in the Molpro quantum chemistry program.<sup>43</sup> These methods, as first proposed by Saebo and Pulay,<sup>44–46</sup> are useful alternatives to canonical formulations of post-Hartree Fock (HF) theory. They allow a drastic reduction of computational cost relative to molecular size. In the case of single-reference methods, asymptotic linear scaling algorithms have been successfully implemented for local MP2 and CCSD(T0),<sup>47–49</sup> the latter including a non-iterative approximation to the triples (T0).<sup>50</sup> Since dynamical correlation is a short-range effect, it can be treated efficiently by using small local configuration spaces.

In order to reduce the computational cost of the calculations, density-fitting approximations were used throughout, i.e. for HF<sup>51</sup>, MP2 and LMP2<sup>52</sup> as well as for LCCSD(T0).<sup>53</sup> Local approximations in the density fitting were not applied, although this could further improve the computational scaling.<sup>52,54</sup> Two basis sets were used for the correlation calculations. The first set consisted of the Dunning aug-cc-pVTZ basis for C and O, aug-cc-pV(T+d)Z for S and the non-augmented cc-pVTZ basis for hydrogen atoms.<sup>55,56</sup> The Mo atom was described by the ECP28MDF pseudopotential, together with the correlation consistent aug-cc-pVTZ-PP basis set of Peterson *et al.*<sup>57</sup> The second set was built by increasing the basis sets to valence quadruple-zeta quality (aug-cc-pVQZ for C and O, aug-cc-pV(Q+d)Z for S, cc-pVQZ for H and aug-cc-pVQZ-PP for Mo).<sup>55–57</sup> For ease of notation, we will refer to the two sets as TZ and QZ, respectively.

In the HF calculations with the TZ orbital basis, we used the def2-TZVPP<sup>58</sup> Coulomb and exchange fitting basis sets for Mo and cc-pVTZ/JKFIT<sup>59</sup> for all remaining atoms. In connection with the QZ basis we chose the def2-QZVPP<sup>58</sup> and cc-pVQZ/JKFIT<sup>59</sup> fitting basis sets. The MP2/TZ calculations were done with the def2-TZVPP MP2 fitting basis for Mo<sup>60</sup> and the aug-cc-

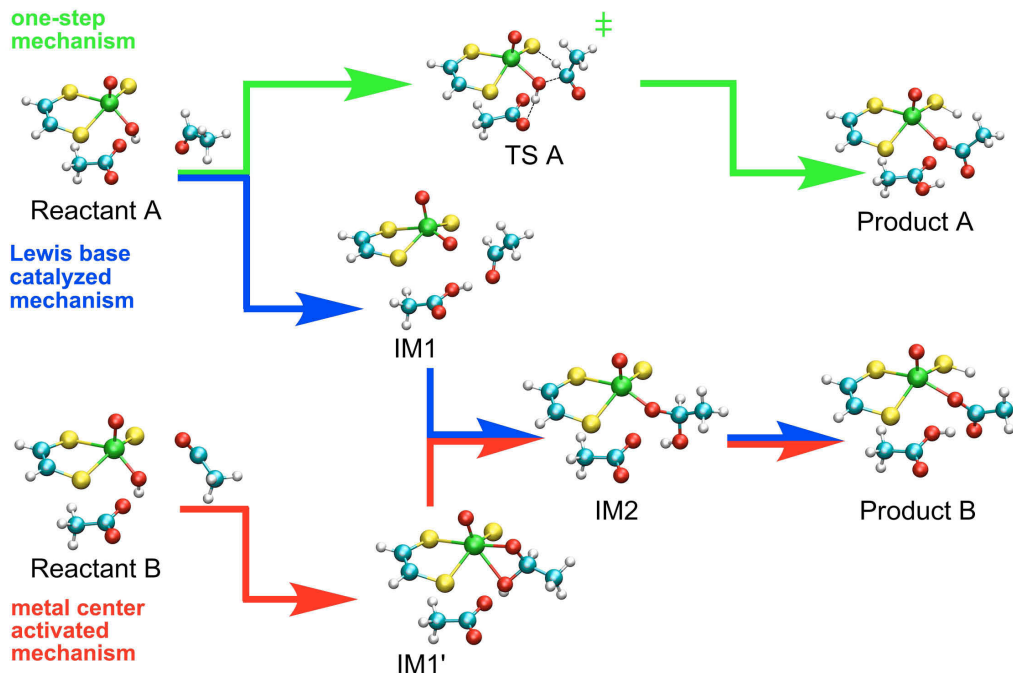


FIG. 1: Proposed mechanisms with key intermediates (IM). A transition state is depicted only for the one-step mechanism. All other transition states are shown in the original publication.<sup>19</sup>

PVTZ/MP2FIT<sup>61</sup> fitting basis for all remaining atoms, while def2-QZVPP<sup>60</sup> and aug-cc-pVQZ/MP2FIT<sup>61</sup> were used for the QZ basis. In the density fitted local coupled cluster calculations with the TZ basis, the fitting for the 4-external integrals was also performed with the latter combination of basis sets (def2-QZVP and aug-cc-pVQZ/MP2FIT).

In the local methods two kinds of approximations are made: first, the excitations from a given pair of localized molecular orbitals (LMOs) are restricted to a subset of non-orthogonal projected atomic orbitals that are spatially close to the LMOs. This is denoted as the *domain approximation*.<sup>44,45,62</sup> Secondly, the orbital pairs are classified according to the distance of the two LMOs involved. The pair correlation energies decrease quickly with increasing distance, and therefore *distant pairs*, which only contribute a negligible amount of correlation energy, can be neglected. Most of the correlation energy (90-95%) is due to the *strong pairs*, in which the two orbital domains overlap. The strong pairs are treated at highest level, e.g. LCCSD(T0). Intermediate are the *close* and *weak pairs*, which contribute only a few percent to the correlation energy. The amplitudes of these pairs can be optimized at lower level (e.g. LMP2, *weak pair approximation*). The close pairs are also optimized by LMP2 and included in the triples calculation. As in previous studies,<sup>63,64</sup> we examined the effects of these approximations separately.

In all local calculations we used natural localized molecular orbitals<sup>65</sup> (NLMOs), and the domains were

determined by natural population analysis (NPA) as recently proposed.<sup>66</sup> We also tested the traditional scheme using Pipek-Mezey localization and the Boughton-Pulay criterium for domain selection,<sup>67</sup> but as in previous studies<sup>66,68</sup> the new NLMO/NPA scheme yielded more consistent domains for the different structures and was found to be more stable with respect to the basis set. The criterium for including atoms into a domain was chosen to be  $T_{\text{NPA}} = 0.03$ .

The effect of the domain approximation can be quantified by comparing LMP2 and MP2 relative energies for a given basis (XZ = TZ or QZ):

$$\Delta E_{\text{domain}} = E_{\text{LMP2/XZ}} - E_{\text{MP2/XZ}} \quad (1)$$

These differences are plotted in Figure 2 for the stationary points of the Lewis base catalyzed mechanism (data for TZ and QZ obtained with identical domains). Evidently, LMP2 overestimates the relative energies of the intermediates and transition states relative to the reactant, by up to 3 kcal mol<sup>-1</sup> (TZ). The domain error drops significantly with increasing basis set size (to a maximum of 1.5 kcal mol<sup>-1</sup> for QZ). This could be partly due to a reduction of the basis set superposition error (BSSE) in the LMP2 method.<sup>69,70</sup> The BSSE lowers the barriers, and therefore the LMP2 method yields barrier heights that are larger than the MP2 ones.

To assess the errors due to basis set incompleteness, we used a two-point  $n^{-3}$  extrapolation of the MP2/TZ

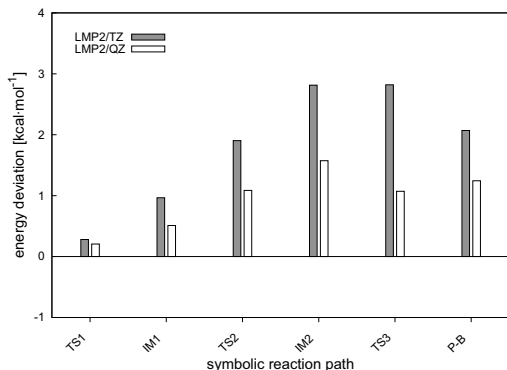


FIG. 2: Differences between LMP2 and MP2 relative energies of the stationary points of the Lewis base catalyzed mechanism, for the aug-cc-pVTZ and aug-cc-pVQZ basis sets

and MP2/QZ correlation energies and added the resulting value to the HF/QZ energy to estimate the MP2 energy at the complete basis set (CBS) limit,  $E_{\text{MP2/CBS}[34]}$ . The basis set incompleteness error for the TZ or QZ basis is then defined as:

$$\Delta E_{\text{basis}} = E_{\text{MP2/CBS}[34]} - E_{\text{MP2/XZ}} \quad (2)$$

These errors are plotted in the top panel of Figure 3, again for the stationary points of the Lewis base catalyzed mechanism. They never exceed 1.0 (0.2) kcal mol<sup>-1</sup> for the TZ (QZ) basis and are thus clearly much smaller than the domain error. The bottom panel of Figure 3 shows the corresponding differences between the LMP2/XZ and MP2/CBS[34] relative energies which measure the combined effect of the errors from the domain approximation and basis set incompleteness. While there is some minor error cancellation, the LMP2/XZ relative energies remain too large, and although they are reduced for the larger QZ basis, they are not negligible even then.

The LCCSD(T0) calculations could only be performed with the TZ basis given the available computational resources. Assuming that the domain error and the basis set incompleteness error are similar at the MP2 and CCSD(T) levels, we approximately account for these errors by adding to the LCCSD(T0)/TZ values the following MP2-based correction:

$$\Delta E_{\text{EC}} = E_{\text{MP2/CBS}[34]} - E_{\text{LMP2/TZ}} \quad (3)$$

In the local coupled cluster methods, the evaluation of the pair correlation energies depends on the distance  $r_p$  between the corresponding LMOs, see Table I. For the classification of the pair types, we used the **Molpro** default values in two cases, i.e.,  $R_c = 1$  Bohr and  $R_{vd} = 15$  Bohr. The remaining parameter  $R_w$  affects the number of pairs taken into account in the perturbative triples calculation. As in the case of chorismate mutase<sup>64</sup> and

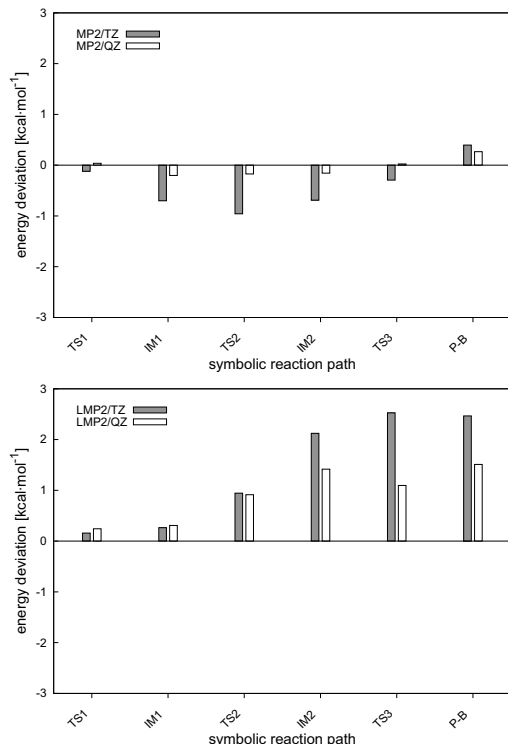


FIG. 3: Differences from the MP2/CBS[34] relative energies of the stationary points of the Lewis base catalyzed mechanism: MP2 (upper panel) and LMP2 (lower panel)

TABLE I: Pair types used in the local coupled cluster calculations.

strong	$r_p < R_c$	treated at the CCSD level
close	$R_c \leq r_p < R_w$	treated at the MP2 level and included in the triples calculation
weak	$R_w \leq r_p < R_{vd}$	treated at the MP2 level
very distant	$r_p \geq R_{vd}$	neglected

PHBH,<sup>63,64</sup> we chose  $R_w = 5$  Bohr (i.e., larger than the **Molpro** default of  $R_w = 3$  Bohr) to properly handle the stretched bond lengths in the transition states. We performed some test calculations for selected stationary points of the Lewis base catalyzed mechanism to validate the chosen values for  $R_c$  and  $R_w$ . The changes in the relative energies were well below 1 kcal mol<sup>-1</sup> when varying these parameters over a reasonable range (data not shown).

In summary, the tests of the local correlation methods indicate that the domain approximation has an effect of about 2 kcal mol<sup>-1</sup> on the energetics in the AOR enzyme, while the errors from basis set incompleteness appear to be smaller than 1 kcal mol<sup>-1</sup>. Adding MP2-based cor-

rections  $\Delta E_{\text{EC}}$  to LCCSD(T0)/TZ energies is expected to remove most of these errors from the coupled cluster results. The effect of the pair approximations is small and will be neglected.

## B. Free Energy Perturbation

Reaction rates are ultimately determined by changes in the free energy of the reacting system. There are several methods to convert energy profiles ( $\Delta E$ ) of an enzymatic reaction to free energy profiles ( $\Delta A$ ) which commonly involve sampling along a predetermined reaction path by means of MD techniques. In the present case of AOR, we apply free energy perturbation (FEP) theory in a QM/MM framework for this purpose. We present a brief summary of the QM/MM FEP method<sup>21–23</sup> to establish notation.<sup>23</sup>

The basic idea of QM/MM FEP is to sample over the MM region while keeping the QM region fixed. The free energy difference between two stationary points can thus be partitioned into a QM part (evaluated from statistical thermodynamics) and an environmental part which accounts for the contributions from the MM region and the QM–MM interactions (obtained from MD sampling over the MM region).

$$\Delta A_{\text{FEP}} = \Delta A_{\text{QM}} + \Delta A_{\text{env}} \quad (4)$$

The QM part  $\Delta A_{\text{QM}}$  consists of four contributions:

$$\Delta A_{\text{QM}} = \Delta E_{\text{QM}} + \Delta E_{\text{QM}}^{\text{ZPE}} + \Delta U_{\text{QM}}^{\text{th}} - T\Delta S_{\text{QM}} \quad (5)$$

We apply the rigid-rotor harmonic-oscillator approximation to obtain the latter three terms, namely the differences in the zero-point vibration energies  $\Delta E_{\text{QM}}^{\text{ZPE}}$ , the thermal contributions to the internal energy  $\Delta U_{\text{QM}}^{\text{th}}$ , and the entropic contributions  $-T\Delta S_{\text{QM}}$ . This requires (normally numerical) QM/MM frequency calculations for the QM region (including all MM atoms linked to QM atoms that are frozen during the FEP-MD sampling).

The first term  $\Delta E_{\text{QM}}$  is computed from the QM wavefunction obtained by electrostatic embedding, as expectation value of the pure QM Hamiltonian without MM point charges.<sup>23</sup> This definition is needed to avoid double counting of the QM–MM electrostatic interactions which are included both in the FEP-MD sampling and in the QM energy  $\Delta E_{\text{SCF}}$  obtained by electrostatic embedding (with the MM point charges included in the Hamiltonian).<sup>23</sup>

Concerning the environmental part in eq. 4, the QM/MM FEP method expresses the free energy difference between two states  $a$  and  $b$  as the exponential average of the corresponding potential energy change (sampled over configurations representative of state  $a$ ). More specifically, for two states along the reaction path, the free energy difference is evaluated by sampling, at the

MM level, the energy difference between the QM part being in configuration  $a$  to being in configuration  $b$ .

$$\Delta A_{\text{env}}^{a \rightarrow b} = -k_B T \ln \left\langle \exp \left\{ -\Delta E_{\text{pert}}^{a \rightarrow b} / (k_B T) \right\} \right\rangle_{\text{MM}}^a \quad (6)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $\Delta E_{\text{pert}}^{a \rightarrow b}$  is given by:

$$\Delta E_{\text{pert}}^{a \rightarrow b} = E_{\text{QM-MM}}(\mathbf{R}_{\text{QM}}^b, \mathbf{R}_{\text{MM}}(t)) - E_{\text{QM-MM}}(\mathbf{R}_{\text{QM}}^a, \mathbf{R}_{\text{MM}}(t)) \quad (7)$$

with  $E_{\text{QM-MM}}$  representing all bonded and non-bonded QM–MM interactions.<sup>23</sup>

In practice, these calculations need to be done for a number of windows along a reaction path, and one has to sum up the increments between the relevant minimum and transition state to determine the environmental contribution  $\Delta A_{\text{env}}$  to the free energy barrier. The FEP procedure can be carried out in the forward and backward direction, and the resulting free energy profiles must coincide within the accuracy of the sampling.

The QM/MM FEP approach outlined so far is computationally feasible only for semiempirical QM components. Two further approximations have been introduced<sup>21</sup> to handle ab initio or DFT QM methods: During FEP-MD sampling, the QM density is frozen and represented by electrostatic-potential derived point charges when evaluating the electrostatic QM–MM interactions. In the present work, we adopt these approximations which have been validated in the previous studies.<sup>22,23</sup>

The current QM/MM FEP calculations were carried out at the same DFT(B3LYP/B1)/CHARMM level that had been used in the original work on AOR.<sup>19</sup> The relative free energies thus obtained can be further improved by including the corrections from the single-point ab initio QM/MM energy calculations with local correlation methods (see above). Our best estimate for free energy differences is:

$$\Delta A_{\text{FEP}}^{\text{best}} = \Delta A_{\text{FEP}} - \Delta E_{\text{QM(B3LYP/B1)/MM}} + \Delta E_{\text{QM(LCCSD(T0))/TZ/MM}} + \Delta E_{\text{EC}} \quad (8)$$

## IV. RESULTS AND DISCUSSION

In this section we present the results for the three mechanisms considered previously<sup>19</sup> for the reductive half-reaction of AOR toward acetaldehyde (see Figure 1). In each case, the reactant contains a complex between the Mo-cofactor and the substrate acetaldehyde, while the product contains a complex between the reduced cofactor and acetic acid. These complexes adopt somewhat different conformations in the different mechanisms (labeled A and B, respectively), with slightly different energies. In the one-step mechanism, reactant A (R-A) is

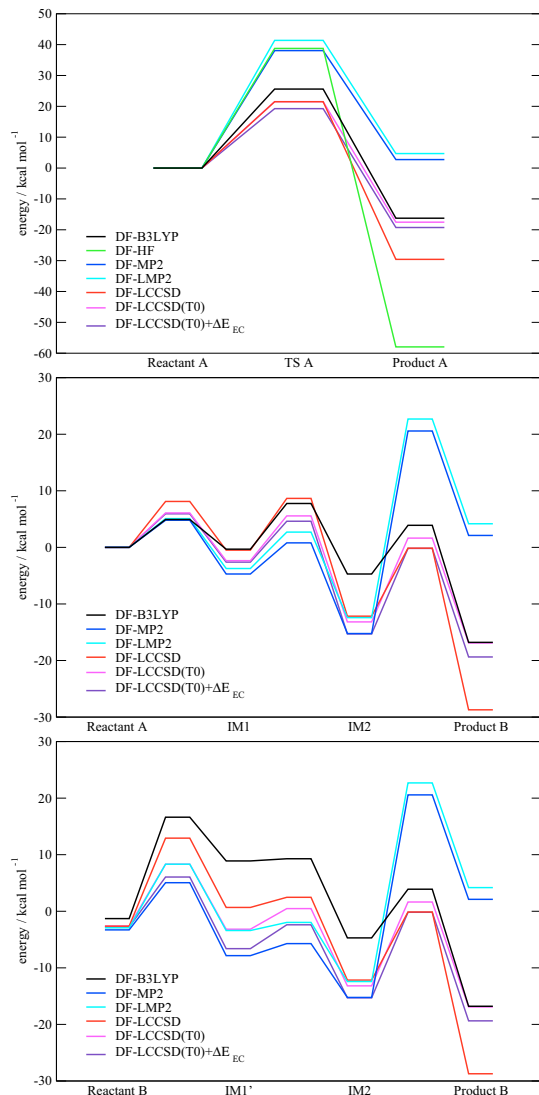


FIG. 4: Relative QM/MM energies for the one-step (top), the Lewis base catalyzed (middle) and the metal center activated mechanism (bottom) at the HF, B3LYP, MP2, LMP2, LCCSD, LCCSD(T0) and LCCSD(T0)+ $\Delta E_{EC}$  levels, using the TZ basis set.

converted to product A (P-A) in a concerted process that involves a nucleophilic attack of the Mo-OH moiety onto the carbonyl carbon atom of acetaldehyde and a concomitant hydride transfer from acetaldehyde to the Mo=S sulfur atom. In the Lewis base catalyzed mechanism, the cofactor in reactant A is first activated by a proton transfer from Mo-OH to the nearby Glu869 residue yielding intermediate IM1; this is followed by a nucleophilic attack that gives the tetrahedral intermediate IM2 which then rearranges to product B (P-B) by a hydride transfer (with simultaneous proton transfer from Glu869 to form acetic acid). In the metal center activated mechanism,

the acetaldehyde substrate first coordinates to the cofactor of reactant B (R-B), with concomitant nucleophilic attack of Mo=O toward acetaldehyde; the resulting intermediate IM1 undergoes a low-barrier internal conversion to another intermediate IM2 which is then transformed into product B by hydride transfer. The latter two mechanisms share the same tetrahedral intermediate IM2 and the same final hydride transfer step,<sup>19</sup> see Figure 1.

Detailed numerical results for all mechanisms are given in the Supporting Information. In the following, we present and discuss the computed energy and free energy profiles. We address the results from the single-point ab initio QM/MM energy calculations first, and thereafter those from the QM/MM FEP treatment.

### A. Single-point ab initio QM/MM results

The energy profiles from single-point QM/MM calculations with the TZ basis at the optimized DFT(B3LYP/B1)/CHARMM geometries are shown in Figure 4. The B3LYP-based relative energies show some changes when replacing the B1 basis employed previously<sup>19</sup> by the larger TZ basis, with deviations up to about 5 kcal mol<sup>-1</sup>, but the qualitative features of the energy profiles are conserved.

Compared with the B3LYP and coupled cluster results, the HF method significantly overestimates the stability of the Mo(IV) product state and thus predicts an extremely large exothermicity for the overall reaction. Furthermore, it gives exceedingly high barriers for some steps, especially those involving nucleophilic attack of the Mo=O moiety onto acetaldehyde. The HF results for AOR thus appear to be unreliable.

The MP2 and LMP2 energy profiles look rather similar: the deviations discussed above (of about 2 kcal mol<sup>-1</sup>) are small on the scale of Figure 4. The MP2 relative energies are rather close to the B3LYP and coupled cluster values in the initial stages (up to IM2) of the Lewis base catalyzed mechanism and the metal center activated mechanism, but they fail for the final step and also for the one-step mechanism. MP2 and LMP2 are thus evidently not able to describe the hydride transfer step that is accompanied by a Mo(VI)  $\rightarrow$  Mo(IV) transition, since they yield much too high energies for the transition state and product: they even make the overall reaction endothermic (the same applies to spin-component-scaled<sup>71</sup> LMP2). This MP2 failure has been noted earlier<sup>72,73</sup> and has been attributed to problems in the description of the Mo=S bond.<sup>74</sup> One may suspect that the Mo(VI)  $\rightarrow$  Mo(IV) transition, with a developing Mo  $d^2$ -configuration, could introduce some multi-reference character that is not captured at the MP2 level. Given this situation, it is not clear *a priori* whether the perturbative triples correction in LCCSD(T0) will behave properly for these species.

Comparing the LCCSD and LCCSD(T0) relative energies, we find that the only major difference indeed



concerns the product state which is destabilized by 12 kcal mol<sup>-1</sup> when including the T0 corrections; it may well be fortuitous that this shifts the overall LCCSD(T0) exothermicity almost on top of the B3LYP value. For the other stationary points, the LCCSD and LCCSD(T0) results are generally quite close to each other (typically within 2 kcal mol<sup>-1</sup>, also for the hydride transfer transition states). Comparing the LCCSD(T0) and B3LYP relative energies, we note good agreement for the one-step mechanism (within 4 kcal mol<sup>-1</sup>) and for the Lewis base catalyzed mechanism (within 2 kcal mol<sup>-1</sup>, except for intermediate IM2 which is more stable in LCCSD(T0) by 8 kcal mol<sup>-1</sup>). The deviations between LCCSD(T0) and B3LYP are larger in the case of the metal center activated mechanism where the intermediates IM1' and IM2 as well as the associated transition states are again appreciably more stable in LCCSD(T0), relative to the reactant.

Concerning barriers, the one-step mechanism requires substantial activation both according to B3LYP and LCCSD(T0) (25.6 vs 21.5 kcal mol<sup>-1</sup>). In the Lewis base catalyzed mechanism, the B3LYP and LCCSD(T0) barriers are similar for the initial proton transfer (4.9 vs 6.1 kcal mol<sup>-1</sup>) and the following nucleophilic attack (8.1 vs 7.9 kcal mol<sup>-1</sup>), but differ for the final hydride transfer step (8.6 vs 14.8 kcal mol<sup>-1</sup>) due to the enhanced stability of IM2 in LCCSD(T0) (see above). In the metal center activated mechanism, the initial coordination and nucleophilic attack step has a higher barrier in B3LYP than in LCCSD(T0) (16.6 vs 8.3 kcal mol<sup>-1</sup>), the following internal rotation is facile at both levels, and the final hydride transfer is the same as in the Lewis base catalyzed mechanism.

In an overall assessment, the B3LYP and LCCSD(T0) results are reasonably close to each other. The energy profiles are qualitatively similar. Both methods identify the one-step mechanism as least favorable, and both prefer the Lewis base catalyzed mechanism (even though only by a small margin in the case of LCCSD(T0)). Analogous conclusions arise in the comparison of B3LYP and LCCSD(T0)+ $\Delta E_{EC}$  results since the corrections for the domain and basis set incompleteness errors are generally rather small.

## B. QM/MM FEP results

In the following we discuss the changes in the reaction profiles when going from energies to free energies. All results in this section were obtained at the QM(B3LYP/B1)/CHARMM level; improvements based on the ab initio single point results will be discussed in the next section. We first consider the QM part and then the environmental part of the free energy differences. For the sake of convenience, we introduce the term

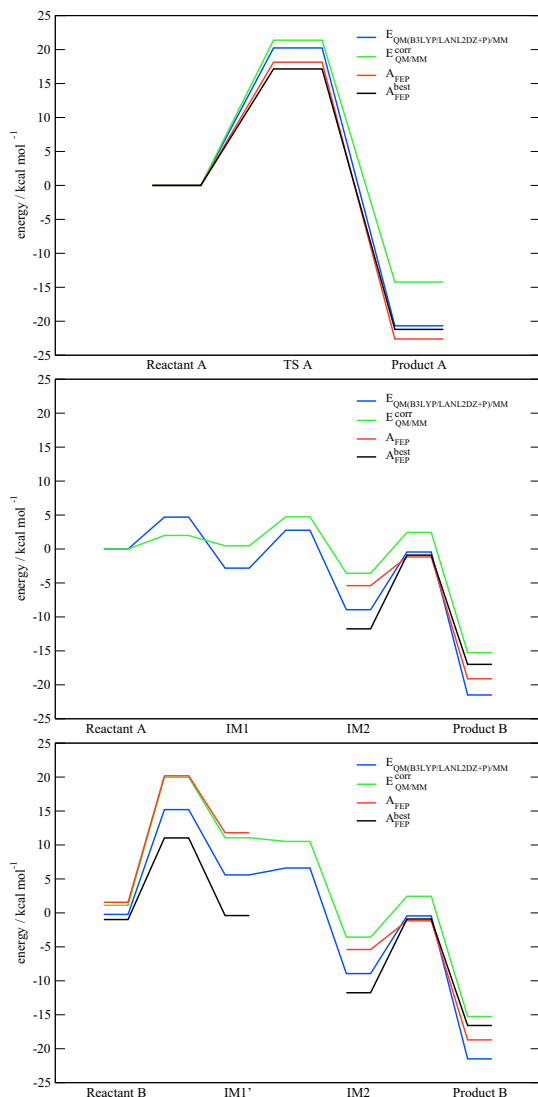


FIG. 5: Relative QM/MM energies ( $E$ ) and free energies ( $A$ ) for the three pathways studied (top: one-step mechanism; middle: Lewis base catalyzed mechanism; bottom: metal center activated mechanism). Blue: QM(B3LYP/B1)/MM energy; green: corrected QM/MM energy (eq. 9); red: QM(B3LYP/B1)/MM free energy (eq. 4); black: best free energy estimate (eq. 8). See text.

$$\Delta E_{QM/MM}^{corr} = \Delta E_{QM/MM} + \Delta E_{QM}^{ZPE} + \Delta U_{QM}^{th} - T\Delta S_{QM} \quad (9)$$

which is the sum of the QM(B3LYP/B1)/MM energy  $\Delta E_{QM/MM}$  (computed with inclusion of the MM point charges) and the ZPE, thermal, and entropic correction terms that also appear in eq. 5. Numerical results for the various terms in eq. 9 are given in the Supporting Information for all minima and transition states. The corresponding results for the barriers of the elementary

TABLE II: Barriers (kcal mol<sup>-1</sup>) from QM(B3LYP/B1)/MM calculations, for the elementary steps in the three mechanisms considered (see text for the definition of the various terms).

	one-step	Lewis base	Lewis base	Lewis base	metal center	metal center
	R-A → P-A	R-A → IM1	IM1 → IM2	IM2 → P-B	R-B → IM1'	IM1' → IM2
$\Delta^\ddagger E_{\text{QM/MM}}$	20.2	4.7	5.6	8.5	15.4	1.0
$\Delta^\ddagger E_{\text{QM}}^{\text{ZPE}}$	-0.8	-2.5	-2.2	-3.1	0.6	-1.5
$\Delta^\ddagger U_{\text{QM}}^{\text{th}}$	-1.3	-0.2	-0.7	-0.5	-1.0	-0.1
$-T\Delta^\ddagger S_{\text{QM}}$	3.2	0.0	1.6	1.1	3.4	0.0
$\Delta^\ddagger E_{\text{QM/MM}}^{\text{corr}}$	21.4	2.0	4.3	6.0	18.4	-0.6
$\Delta^\ddagger E_{\text{QM}}$	9.2	4.8	3.1	7.4	20.4	-0.8
$\Delta^\ddagger A_{\text{QM}}$	10.3	2.1	1.8	4.9	23.4	-2.4
$\Delta^\ddagger A_{\text{env}}$	7.8	—	—	-0.7	-4.8	—
$\Delta^\ddagger A_{\text{FEP}}$	18.1	—	—	4.2	18.7	—
$\Delta^\ddagger A_{\text{FEP}}^{\text{best}}$	17.1	—	—	10.8	12.1	—

steps in the three mechanisms considered are collected in Table II.

For the one-step mechanism, inclusion of the correction terms increases the barrier from 20.2 to 21.4 kcal mol<sup>-1</sup>. This is mainly due to the fact that the transition state for the concerted reaction, with concomitant nucleophilic attack and hydride transfer, is more compact than the reactant and thus carries an entropic penalty of  $-T\Delta^\ddagger S_{\text{QM}} = 3.2$  kcal mol<sup>-1</sup> which dominates over the other two contributions.

For the Lewis base catalyzed process, the small barrier for the initial proton transfer (R-A → IM1) is further reduced by the correction terms, from 4.7 to 2.0 kcal mol<sup>-1</sup>, almost entirely because of the smaller zero-point vibrational energy in the transition state (loss of an O-H vibration). The barrier for the subsequent nucleophilic substitution step (IM1 → IM2) is less affected (5.6 vs 4.3 kcal mol<sup>-1</sup>) since the counteracting changes in the zero-point energy and the entropic term compensate each other to a large extent. The barrier for the final hydride transfer (IM2 → P-B) is lowered again, from 8.5 to 6.0 kcal mol<sup>-1</sup>, mainly due to the reduction in the zero-point energy (loss of a C-H vibration).

In the case of the metal center activated mechanism, the barrier for the first step increases when going from  $\Delta E_{\text{QM/MM}}$  to  $\Delta E_{\text{QM/MM}}^{\text{corr}}$ , from 15.4 to 18.4 kcal mol<sup>-1</sup>. This step (R-B → IM1') involves both coordination of the substrate to the molybdenum atom and nucleophilic attack, and hence the transition state is highly organized and thus entropically disfavored, with  $-T\Delta^\ddagger S_{\text{QM}} = 3.4$  kcal mol<sup>-1</sup>. The following internal rotation (IM1' → IM2) is very facile, with an energy barrier of 1.0 kcal mol<sup>-1</sup>, that even turns slightly negative when including the frequency corrections; this step is mechanistically irrelevant. The final hydride transfer (IM2 → P-B) is the same as in the Lewis base catalyzed process (see above).

The QM(B3LYP/B1)/MM energy  $\Delta E_{\text{QM/MM}}$  is the

sum of the QM energy  $\Delta E_{\text{QM}}$ , the MM energy, and the QM-MM interaction energy. The numerical results for  $\Delta^\ddagger E_{\text{QM}}$  are also included in Table II, as well as those for  $\Delta^\ddagger A_{\text{QM}}$ . We do not discuss these values since the differences  $\Delta A_{\text{QM}} - \Delta E_{\text{QM}}$  and  $\Delta E_{\text{QM/MM}}^{\text{corr}} - \Delta E_{\text{QM/MM}}$  are identical by definition, see eqs. 5 and 9.

We now turn to the environmental contributions to the free energy profiles. FEP calculations at the QM(B3LYP/B1)/MM level were performed for the most important steps in the three mechanisms, i.e., for the concerted one-step reaction, the final hydride transfer that occurs both in the Lewis base catalyzed process and in the metal center activated process, and the initial step in the latter mechanism that involves coordination to the metal center as well as nucleophilic attack. The resulting free energy changes along the reaction path in forward and backward direction are documented in the Supporting Information. The data for the forward and backward direction generally agree well (with only very minor hysteresis effects); average values were used when generating the free energy profiles. These are qualitatively rather similar to the underlying potential energy profiles in all three cases (see Figure 6 for the hydride transfer and the Supporting Information for the other two reactions and for detailed numerical results). The corresponding free energy changes in the barriers due to the environmental contributions are included in Table II.

The  $\Delta^\ddagger A_{\text{env}}$  values measure the total free energy contributions from the environment (including the remaining MM and the QM-MM potential energy terms). For an assessment of the FEP contributions due to the sampling of the environment, one should rather focus on the differences  $\Delta^\ddagger A_{\text{FEP}} - \Delta E_{\text{QM/MM}}^{\text{corr}}$  which amount to -3.3, -1.8, and +0.2 kcal mol<sup>-1</sup> for the three elementary steps considered (see Table II). The one-step reaction is thus facilitated somewhat by the environmental FEP contributions, the hydride transfer is influenced less, and the

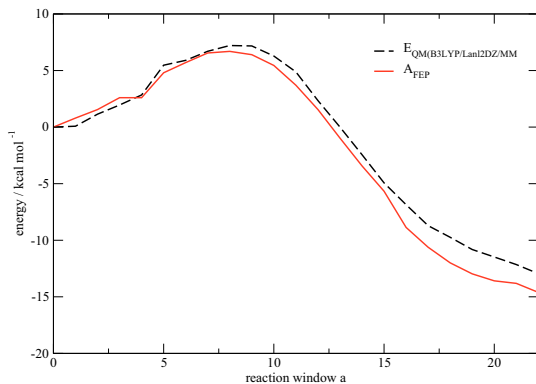


FIG. 6: Reaction profile of the hydride transfer reaction. Black line: Potential energy profile ( $E_{\text{QM(B3LYP/B1)/MM}}$ ). Red line: Free energy profile ( $A_{\text{FEP}}$ ).

initial step of the metal center activated mechanism is almost unaffected. We note, however, that the favorable FEP contributions in the first two cases are almost completely compensated by unfavorable entropic contributions from the QM part (see the  $-T\Delta^\ddagger S_{\text{QM}}$  values in Table II) so that the overall entropic effect on these reactions appears to be rather small. This is plausible since the reactions under investigation are quite local in character, without much re-organization of the active site and its surroundings.<sup>23</sup> This is illustrated for the hydride transfer in Figure 7 which shows an overlay of the optimized structures of the corresponding stationary points.

The three elementary reactions studied by FEP calculations were those with the highest QM(B3LYP/B1)/MM barriers in the three mechanisms considered. These barriers change moderately when going from energies ( $\Delta^\ddagger E_{\text{QM/MM}}$ ) to free energies ( $\Delta^\ddagger A_{\text{FEP}}$ ), namely from 20.2 to 18.1 kcal mol<sup>-1</sup> for the one-step mechanism, from 8.5 to 4.2 kcal mol<sup>-1</sup> for the Lewis base catalyzed process, and from 15.4 to 18.6 kcal mol<sup>-1</sup> for the metal center activated mechanism. The free energy treatment thus reinforces the preference for the Lewis base catalyzed mechanism at the QM(B3LYP/B1)/MM level.

### C. Best estimates

The free energy barriers obtained in the preceding section can be improved by including the corrections from the single-point ab initio QM/MM energy calculations with local correlation methods (see section IV A). Applying eq. 8 we obtain extrapolated free energy barriers  $\Delta^\ddagger A_{\text{FEP}}^{\text{best}}$  of 17.1 kcal mol<sup>-1</sup> for the one-step mechanism, 10.8 kcal mol<sup>-1</sup> for the hydride transfer, and 12.1 kcal mol<sup>-1</sup> for the first step of the metal center activated pathway, see Table II and Figure 5 (black curve).

The concerted one-step mechanism thus remains the

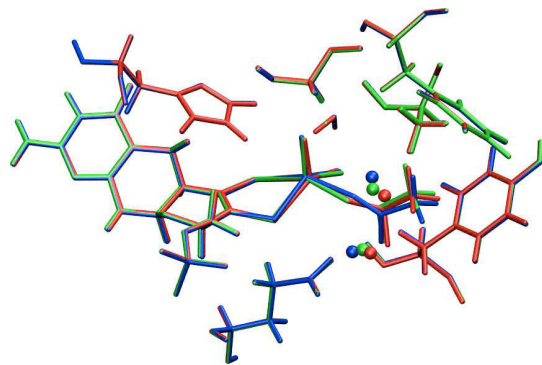


FIG. 7: Overlay of the reaction center of IM2 (red), product B (blue), and the connecting transition state (green). Visible geometrical changes are found only for the transferred hydrogen atom (plotted as ball) and its direct surroundings.

least favorable also after including high-level ab initio energy corrections. Judging from the best estimates for the free energy barriers of the rate-limiting steps in the other two mechanisms, the Lewis base catalyzed process remains favored over the metal center activated process, even though only by a rather small margin. Both mechanisms share the same tetrahedral intermediate (IM2) and the same final hydride transfer step (which is rate-limiting in the Lewis base catalyzed case), but reaching IM2 is much less facile on the metal center activated pathway (where the initial step remains rate-limiting). Judging from the maximum energy spread in the cycle (which determines catalytic turnover) and the energy difference of the barriers preceding IM2, the Lewis base catalyzed mechanism is preferred over both alternatives by a rather large margin.

Finally, there are two further arguments in favor of this mechanism. First, we have not considered the possibility of tunneling up to this point. The hydride transfer occurs in a rather small region of space, without any significant rearrangement of the surroundings, and it thus appears likely that the effective barrier for this step could be lowered to some extent by tunneling, thus increasing the rate of the Lewis base catalyzed process. The rate-limiting steps of the competing mechanisms involve the motion of heavy (non-hydrogen) atoms and, therefore, can not be accelerated in this manner. Secondly, from a chemical point of view, the metal center activated mechanism is sterically more demanding than the Lewis base catalyzed one and would thus seem feasible only for small substrates such as acetaldehyde (if at all). Any large substituent in a substrate (e.g., in benzaldehyde) has to point away from the cofactor which will make this pathway effectively inaccessible.

## V. CONCLUSION

The present high-level ab initio QM/MM calculations and FEP simulations confirm the qualitative mechanistic conclusions of the previous QM/MM study.<sup>19</sup> The three-step Lewis base catalyzed mechanism remains favored, with the final hydride transfer being rate-limiting. Our best estimate for the corresponding free energy barrier is 10.8 kcal mol<sup>-1</sup>. The barriers for the initial proton transfer from Mo-OH to Glu869 and the subsequent nucleophilic attack are significantly lower. The various corrections from the correlated ab initio QM/MM calculations and from the FEP treatment range up to several kcal mol<sup>-1</sup>, but they tend to compensate each other to some extent in the three reactions studied. The best present estimates of the free energy barriers thus differ from the QM(B3LYP/B1)/MM barriers by only about 3

kcal mol<sup>-1</sup>, and the Lewis base catalyzed mechanism remains preferred over the concerted and the metal center activated pathways.

From a methodological point of view, the various corrections considered in the present case study on AOR are too large to be ignored in QM/MM work that aims at accurate predictions. While it is reassuring that standard DFT/MM applications give qualitative mechanistic insights that survive scrutiny at higher levels, it is clearly desirable to strive for enhanced accuracy both for QM/MM energies and free energies, especially in the case of competing mechanisms with similar barriers.

**Acknowledgements.** This work was supported by the Max-Planck-Gesellschaft, by the Deutsche Forschungsgemeinschaft (SFB 706), and by the Fonds der Chemischen Industrie. The support by the latter included a Kekulé scholarship for S.M.

- <sup>1</sup> R. Hille, *Chem. Rev.* **96**, 2757 (1996).
- <sup>2</sup> C. Kisker, H. Schindelin, and D. C. Rees, *Annu. Rev. Biochem.* **66**, 233 (1997).
- <sup>3</sup> C. J. Doonan, A. Stockert, R. Hille, and G. N. George, *J. Am. Chem. Soc.* **127**, 4518 (2005).
- <sup>4</sup> V. Massey and D. Edmondson, *J. Biol. Chem.* **245**, 6595 (1970).
- <sup>5</sup> T. Santos-Silva et al., *J. Am. Chem. Soc.* **131**, 7990 (2009).
- <sup>6</sup> E. B. Skibo, J. G. Gilchrist, and C. H. Lee, *Biochemistry* **26**, 3032 (1987).
- <sup>7</sup> B. A. Barata, J. LeGall, and J. J. G. Moura, *Biochemistry* **32**, 11559 (1993).
- <sup>8</sup> S. L. Andrade, C. Brondino, M. J. Feio, I. Moura, and J. J. G. Moura, *Eur. J. Biochem.* **267**, 2054 (2000).
- <sup>9</sup> R. A. Torres, K. R. Korzekwa, D. R. McMasters, C. M. Fandozzi, and J. P. Jones, *J. Med. Chem.* **50**, 4642 (2007).
- <sup>10</sup> E. Y. Choi, A. L. Stockert, S. Leimkühler, and R. Hille, *J. Inorg. Biochem.* **98**, 841 (2004).
- <sup>11</sup> R. Hille, *Eur. J. Inorg. Chem.*, 1913 (2006).
- <sup>12</sup> Y. Yamaguchi, T. Matsumura, K. Ichida, K. Okamoto, and T. Nishino, *J. Biochem.* **141**, 513 (2007).
- <sup>13</sup> S. E. Clarke, A. W. Harrell, and R. J. Chemery, *Drug Metab. Dispos.* **23**, 251 (1995).
- <sup>14</sup> M. R. Rashidi, J. A. Smith, S. E. Clarke, and C. Beedham, *Drug Metab. Dispos.* **25**, 805 (1997).
- <sup>15</sup> K. Kawashima et al., *Drug Metab. Dispos.* **27**, 422 (1999).
- <sup>16</sup> B. G. Lake et al., *Xenobiotica* **32**, 835 (2002).
- <sup>17</sup> H. M. Senn and W. Thiel, *Top. Curr. Chem.* **268**, 173 (2007).
- <sup>18</sup> H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.* **48**, 1198 (2009).
- <sup>19</sup> S. Metz, D. Wang, and W. Thiel, *J. Am. Chem. Soc.* **131**, 4628 (2009).
- <sup>20</sup> S. Metz and W. Thiel, *J. Am. Chem. Soc.* (doi:10.1021/ja9045394).
- <sup>21</sup> Y. Zhang, H. Liu, and W. Yang, *J. Chem. Phys.* **112**, 3483 (2000).
- <sup>22</sup> J. Kästner, H. M. Senn, S. Thiel, N. Otte, and W. Thiel, *J. Chem. Theory Comput.* **2**, 452 (2006).
- <sup>23</sup> H. M. Senn, J. Kästner, J. Breidung, and W. Thiel, *Can. J. Chem.* **87**, 1322 (2009).
- <sup>24</sup> J. M. Rebelo, J. M. Dias, R. Huber, J. J. G. Moura, and M. J. Romo, *J. Biol. Inorg. Chem.* **6**, 791 (2001).
- <sup>25</sup> MacKerell et al., *J. Phys. Chem. B* **102**, 3586 (1998).
- <sup>26</sup> J. C. Slater, *Phys. Rev.* **81**, 385 (1951).
- <sup>27</sup> S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- <sup>28</sup> A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
- <sup>29</sup> A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
- <sup>30</sup> P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- <sup>31</sup> C. T. Lee, W. T. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- <sup>32</sup> R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Lett.* **162**, 165 (1989).
- <sup>33</sup> W. Smith and T. R. Forester, *J. Mol. Graph.* **14**, 136 (1996).
- <sup>34</sup> S. R. Billeter, A. J. Turner, and W. Thiel, *Phys. Chem. Chem. Phys.* **2**, 2177 (2000).
- <sup>35</sup> Sherwood et al., *J. Mol. Struct. (Theochem)* **632**, 1 (2003), Sp. Iss. SI.
- <sup>36</sup> [www.chemshell.org](http://www.chemshell.org).
- <sup>37</sup> P. J. Hay and W. R. Wadt, *J. Chem. Phys.* **82**, 270 (1985).
- <sup>38</sup> A. W. Ehlers et al., *Chem. Phys. Lett.* **208**, 111 (1993).
- <sup>39</sup> W. R. Wadt and P. J. Hay, *J. Chem. Phys.* **82**, 284 (1985).
- <sup>40</sup> A. Höllwarth et al., *Chem. Phys. Lett.* **208**, 237 (1993).
- <sup>41</sup> P. C. Hariharan and J. A. Pople, *Theor. Chim. Acta* **28**, 213 (1973).
- <sup>42</sup> T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. v. R. Schleyer, *J. Comput. Chem.* **4**, 294 (1983).
- <sup>43</sup> H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, and M. Schütz et al., *Molpro*, version 2008.1, a package of ab initio programs, 2008, see <http://www.molpro.net>.
- <sup>44</sup> P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).
- <sup>45</sup> S. Saebø and P. Pulay, *Chem. Phys. Lett.* **113**, 13 (1985).
- <sup>46</sup> P. Pulay and S. Saebø, *Theor. Chim. Acta* **69**, 357 (1986).
- <sup>47</sup> M. Schütz, G. Hetzer, and H.-J. Werner, *J. Chem. Phys.* **111**, 5691 (1999).
- <sup>48</sup> G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, *J. Chem. Phys.* **113**, 9443 (2000).
- <sup>49</sup> M. Schütz and H.-J. Werner, *J. Chem. Phys.* **114**, 661 (2001).

- <sup>50</sup> M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
- <sup>51</sup> R. Polly, H.-J. Werner, F. R. Manby, and P. J. Knowles, Mol. Phys. **102**, 2311 (2004).
- <sup>52</sup> H.-J. Werner, F. R. Manby, and P. Knowles, J. Chem. Phys. **118**, 8149 (2003).
- <sup>53</sup> H.-J. Werner and M. Schütz, to be published (2009).
- <sup>54</sup> M. Schütz and F. R. Manby, Phys. Chem. Chem. Phys. **5**, 3349 (2003).
- <sup>55</sup> T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
- <sup>56</sup> R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
- <sup>57</sup> K. Peterson, D. Figgen, M. Dolg, and H. Stoll, J. Chem. Phys. **126**, 124101 (2007).
- <sup>58</sup> F. Weigend, J. Chem. Phys. (accepted).
- <sup>59</sup> F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
- <sup>60</sup> A. Hellweg, C. Hättig, S. Höfener, and W. Klopper, Theor. Chim. Acta **117**, 587 (2007).
- <sup>61</sup> F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
- <sup>62</sup> C. Hampel and H.-J. Werner, J. Chem. Phys. **104**, 6286 (1996).
- <sup>63</sup> R. A. Mata, H.-J. Werner, S. Thiel, and W. Thiel, J. Chem. Phys. **128**, 025104 (2008).
- <sup>64</sup> F. Claeysens et al., Angew. Chem., Int. Ed. **45**, 6856 (2006).
- <sup>65</sup> A. E. Reed and F. Weinhold, J. Chem. Phys. **83**, 1736 (1985).
- <sup>66</sup> R. Mata and H.-J. Werner, Mol. Phys. **105**, 2753 (2007).
- <sup>67</sup> J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).
- <sup>68</sup> J. Kaminsky, R. A. Mata, H.-J. Werner, and F. Jensen, Mol. Phys. **106**, 1899 (2008).
- <sup>69</sup> M. Schütz, G. Rauhut, and H.-J. Werner, J. Phys. Chem. A **102**, 5997 (1998).
- <sup>70</sup> R. Mata and H.-J. Werner, J. Chem. Phys. **125**, 184110 (2006).
- <sup>71</sup> S. Grimme, J. Chem. Phys. **118**, 9095 (2003).
- <sup>72</sup> P. Ilich and R. Hille, J. Phys. Chem. B **103**, 5406 (1999).
- <sup>73</sup> P. Ilich and R. Hille, J. Am. Chem. Soc. **124**, 6796 (2002).
- <sup>74</sup> T. Amano, N. Ochi, H. Sato, and S. Sakaki, J. Am. Chem. Soc. **129**, 8131 (2007).



**Supporting Information**  
**Reductive Half-Reaction of Aldehyde Oxidoreductase toward Acetaldehyde: Ab Initio  
and Free Energy QM/MM Calculations**

Johannes M. Dieterich and Hans-Joachim Werner  
*Institut für Theoretische Chemie, Universität Stuttgart,  
Pfaffenwaldring 55, D-70569 Stuttgart, Germany*

Ricardo A. Mata  
*Institut für Physikalische Chemie,  
Tammannstrasse 6, D-37077 Göttingen, Germany*

Sebastian Metz and Walter Thiel  
*Max-Planck-Institut für Kohlenforschung,  
Kaiser-Wilhelm-Platz 1,  
D-45470 Mülheim an der Ruhr, Germany*

TABLE I: QM/MM energies (kcal mol<sup>-1</sup>) of the one-step mechanism for different QM methods (TZ basis; see text).

	HF	B3LYP	LMP2	MP2	MP2/CBS[34]	LCCSD	LSSCD(T0)	LCCSD(T0)+korr
ReactantA*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TS A	38.76	25.59	41.38	38.05	39.10	21.50	21.52	19.24
Product A	-57.95	-16.25	4.70	2.75	2.97	-29.60	-17.54	-19.27
* abs. values (a.u.):	-1956.8960	-1962.6609	-1960.0283	-1960.0579	-1960.4264	-1960.0194	-1960.1998	-1960.59799

TABLE II: QM/MM energies (kcal mol<sup>-1</sup>) of the Lewis base catalyzed mechanism for QM different methods (TZ basis; see text).

	HF	B3LYP	LMP2	MP2	MP2/CBS[34]	LCCSD	LSSCD(T0)	LCCSD(T0)+korr
ReactantA*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TS1	11.49	4.93	5.08	4.80	4.92	8.12	6.09	5.93
IM1	0.83	-0.32	-3.74	-4.70	-4.00	-0.48	-2.35	-2.61
TS2	21.75	7.76	2.70	0.80	1.76	8.66	5.57	4.62
IM2	0.48	-4.71	-12.42	-15.23	-14.54	-12.16	-13.18	-15.3
TS3	5.38	3.90	22.69	20.57	20.86	-0.13	1.64	-0.18
Product B	-56.58	-16.79	4.18	2.11	1.71	-28.72	-16.91	-19.37
* abs. values (a.u.):	-1956.8960	-1962.6609	-1960.0283	-1960.0579	-1960.4264	-1960.0194	-1960.1998	-1960.59799

TABLE III: QM/MM energies (kcal mol<sup>-1</sup>) of the metal center activated mechanism for different QM methods (TZ basis; see text).

	HF	B3LYP	LMP2	MP2	MP2/CBS[34]	LCCSD	LSSCD(T0)	LCCSD(T0)+korr
ReactantB*	-0.28	-1.29	-2.92	-3.29	-3.07	-2.63	-2.62	-2.77
TS1'	31.94	16.63	8.32	5.05	6.05	12.93	8.32	6.05
IM1'	21.79	8.90	-3.40	-7.84	-6.84	0.68	-3.17	-6.61
TS2'	20.17	9.20	-1.97	-5.72	-4.84	2.47	0.47	-2.39
IM2	0.48	-4.71	-12.42	-15.23	-14.54	-12.16	-13.18	-15.30
TS3	5.38	3.90	22.69	20.57	20.86	-0.13	1.64	-0.18
Product B	-56.58	-16.79	4.18	2.11	1.71	-28.72	-16.91	-19.37
* abs. values (a.u.):	-1956.8964	-1962.6630	-1960.0330	-1960.4313	-1960.0330	-1960.0236	-1960.2040	-1960.6023



TABLE IV: QM/MM energies and correction terms (kcal mol<sup>-1</sup>) of the one-step mechanism (B1 basis, except LCCSD(T0)/TZ; see text).

	$E_{QM(B3LYP)/MM}$	$E^{ZPE}$	$U^{th}$	-TΔS	$E_{QM/MM}^{corr}$	$E_{QM(LCCSD(T0))/MM} + \Delta E_{EC}$
ReactantA*	0.00	0.00	0.00	0.00	0.00	0.00
TS A	20.24	-0.79	-1.28	3.21	21.38	19.24
Product A	-20.68	4.46	-1.38	3.37	-14.22	-19.27
* abs. values (a.u.):	-797.7308	0.164267	0.017474	-0.032156	-797.5812	-1960.5979

TABLE V: QM/MM energies and correction terms (kcal/mol) of the Lewis base catalyzed mechanism (B1 basis, except LCCSD(T0)/TZ; see text).

	$E_{QM(B3LYP)/MM}$	$E^{ZPE}$	$U^{th}$	-TΔS	$E_{QM/MM}^{corr}$	$E_{QM(LCCSD(T0))/MM} + \Delta E_{EC}$
ReactantA*	0.00	0.00	0.00	0.00	0.00	0.00
TS	4.70	-2.50	-0.22	0.02	2.00	5.93
IM1	-2.81	2.20	-0.61	1.70	0.47	-2.61
TS	2.77	-0.01	-1.30	3.29	4.74	4.62
IM2	-8.94	3.55	-1.20	3.02	-3.57	-15.30
TS	-0.45	0.48	-1.69	4.10	2.44	-0.18
Product A	-21.49	3.98	-1.56	3.82	-15.26	-19.37
* abs. values (a.u.):	-797.7308	0.164267	0.017474	-0.032156	-797.5812	-1960.5979

TABLE VI: QM/MM energies and correction terms (kcal mol<sup>-1</sup>) of the metal center activated mechanism (B1 basis, except LCCSD(T0)/TZ; see text).

	$E_{QM(B3LYP)/MM}$	$E^{ZPE}$	$U^{th}$	-TΔS	$E_{QM/MM}^{corr}$	$E_{QM(LCCSD(T0))/MM} + \Delta E_{EC}$
ReactantB*	-0.23	1.79	-0.32	-0.12	1.12	-2.77
TS	15.20	2.39	-1.35	3.32	19.56	6.05
IM1'	5.59	3.48	-1.37	3.38	11.07	-6.61
TS	6.61	1.96	-1.42	3.37	10.52	-2.39
IM2	-8.94	3.55	-1.20	3.02	-3.57	-15.30
TS	-0.45	0.48	-1.69	4.10	2.44	-0.18
Product A	-21.49	3.98	-1.56	3.82	-15.26	-19.37
* abs. values (a.u.):	-797.7312	0.167120	0.016962	-0.032350	-797.5794	-1960.6023

TABLE VII: Contributions to the free energies (kcal mol<sup>-1</sup>) of the one-step mechanism.

	$E_{QM}$	$A_{env}^{forward}$	$A_{env}^{backward}$	$A_{env}$	$A_{FEP}$	$A_{FEP}^{best}$
Reactant A	0.00	0.00	0.00	0.00	0.00	0.00
TS A	9.21	6.57	8.93	7.80	18.14	17.14
Product A	-43.67	(-2.82)	14.61	14.61	-22.60	-21.19

The value for  $\Delta A_{forward}$  of product A suffers from an additional proton transfer occurring at window 35. This window was thus left out during the calculation of  $A_{env}$ .

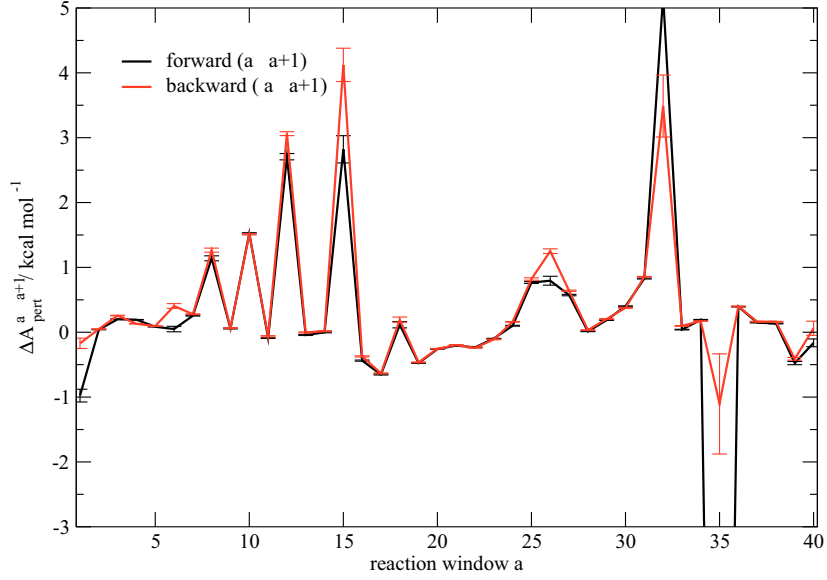
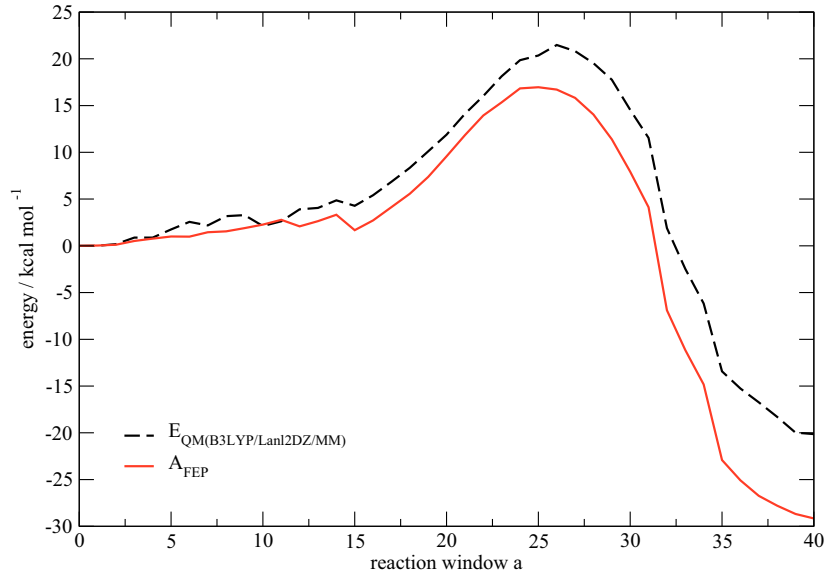
FIG. 1:  $\Delta A_{pert}^{a \rightarrow a+1}$  for the one-step mechanism

FIG. 2: Potential energy profile ( $E_{QM(B3LYP/B1)/MM}$ ) and free energy profile ( $A_{FEP}$ ) for the one-step mechanism. Note that the corrections are not included

TABLE VIII: Contributions to the free energies ( $\text{kcal mol}^{-1}$ ) of the Lewis base catalyzed mechanism.

	$E_{QM}$	$A_{env}^{forward}$	$A_{env}^{backward}$	$A_{env}$	$A_{FEP}$	$A_{FEP}^{best}$
IM2	0.00	0.00	0.00	0.00	0.00	0.00
TS2	7.41	-1.10	-0.34	-0.72	4.21	10.83
Product B	-23.88	8.55	10.00	9.27	-13.74	-5.26

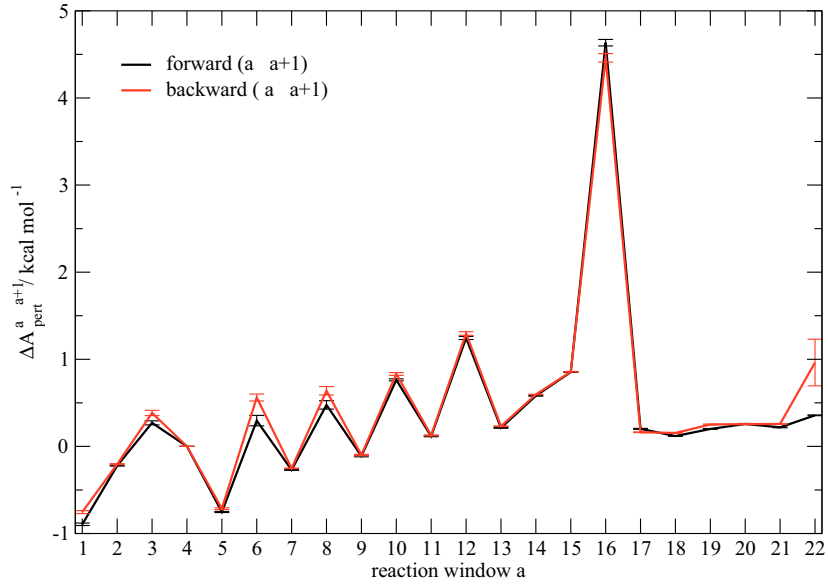
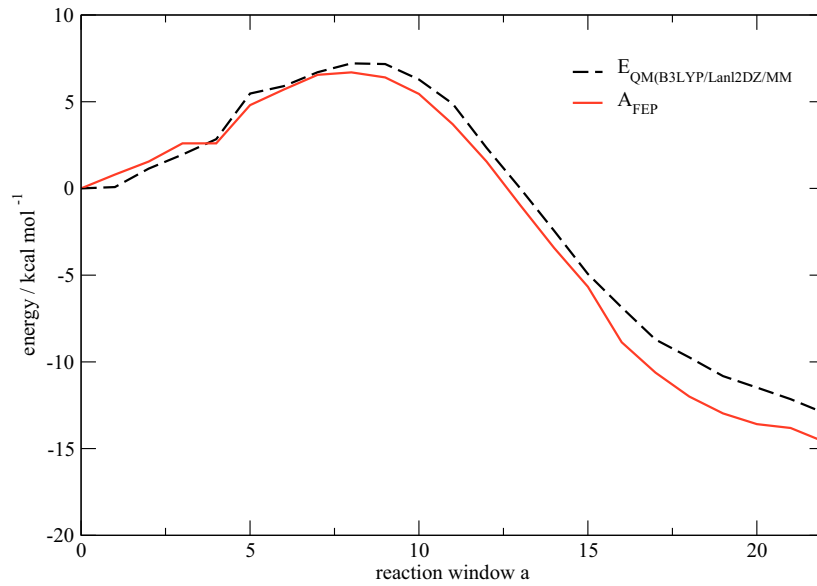
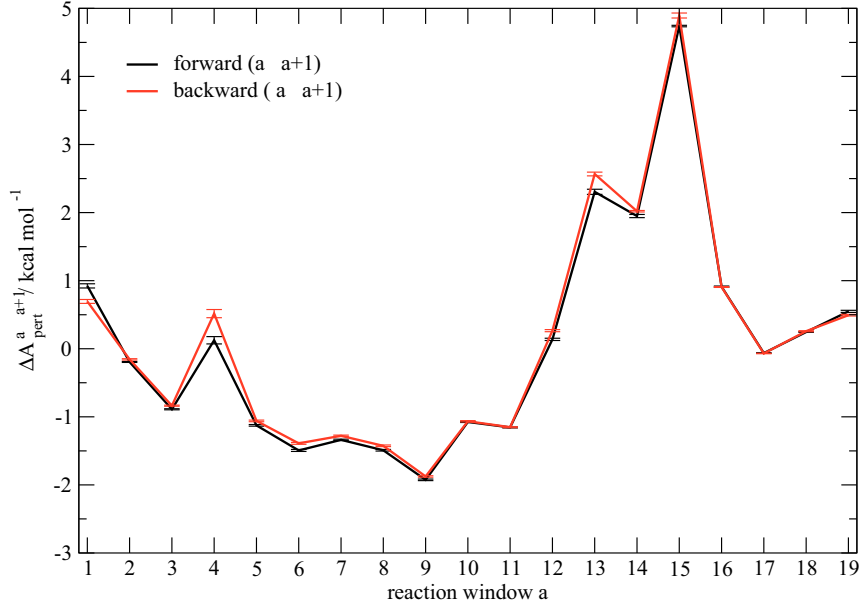
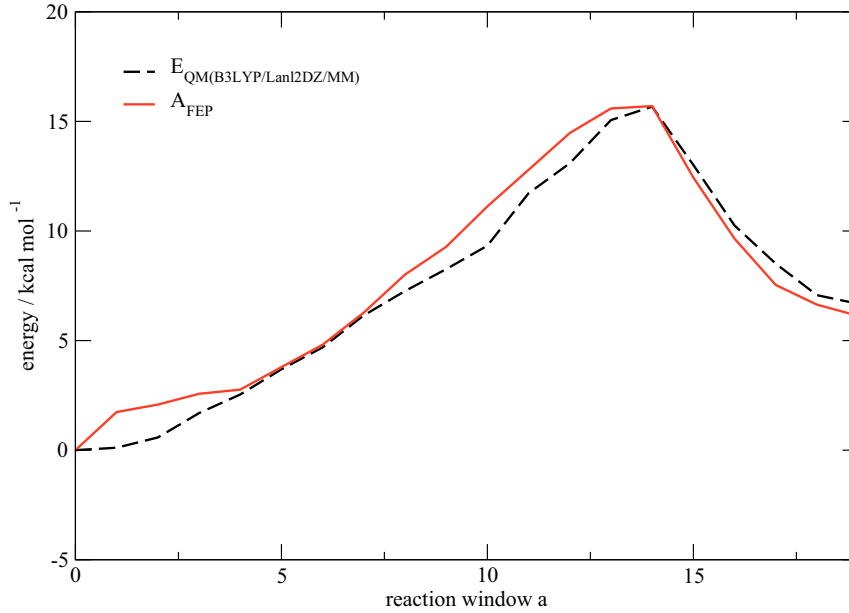
FIG. 3:  $\Delta A_{pert}^{a \rightarrow a+1}$  for the Lewis base catalyzed mechanismFIG. 4: Potential energy profile ( $E_{QM(B3LYP/B1)/MM}$ ) and free energy profile ( $A_{FEP}$ ) for the Lewis base catalyzed mechanism. Note that ZPE corrections are not included.

TABLE IX: Contributions to the free energies ( $\text{kcal mol}^{-1}$ ) of the Lewis base catalyzed mechanism.

	$E_{QM}$	$A_{env}^{forward}$	$A_{env}^{backward}$	$A_{env}$	$A_{FEP}$	$A_{FEP}^{best}$
Reactant B	0.00	0.00	0.00	0.00	0.00	0.00
TS B	20.41	-5.25	-4.18	-4.71	18.71	12.10
IM1'	4.41	1.33	2.31	1.82	10.26	0.71

FIG. 5:  $\Delta A_{pert}^{a \rightarrow a+1}$  for the metal center activated mechanismFIG. 6: Potential energy profile ( $E_{QM(B3LYP/B1)/MM}$ ) and free energy profile ( $A_{FEP}$ ) for the metal center activated mechanism. Note that ZPE corrections are not included.

## **Paper III**

*A Combined QM/MM Study on the  
Reductive Half-reaction of Xanthine Oxidase:  
Substrate Orientation and Mechanism.*

**Sebastian Metz**, and Walter Thiel

*J. Am. Chem. Soc.*, 131 **(2009)** 14885-14902.



## A Combined QM/MM Study on the Reductive Half-Reaction of Xanthine Oxidase: Substrate Orientation and Mechanism

Sebastian Metz and Walter Thiel\*

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

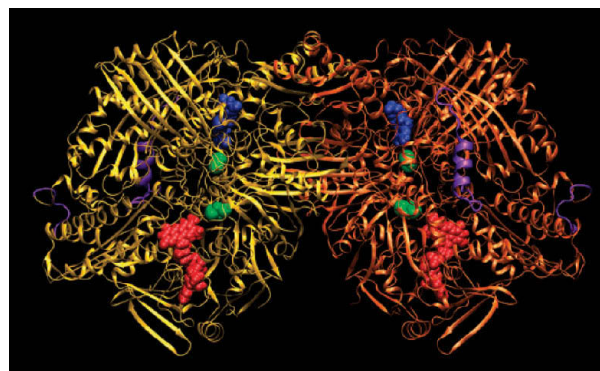
Received June 4, 2009; E-mail: thiel@mpi-muelheim.mpg.de

**Abstract:** Quantum mechanical/molecular mechanical (QM/MM) methods were used to investigate the conversion of xanthine to uric acid in xanthine oxidase. Seven mechanistic variants were considered with different tautomeric forms of xanthine, different protonation states of the active-site residues, and different substrate orientations. The most favorable pathway (setup G) has a B3LYP/MM barrier of about 14 kcal mol<sup>-1</sup>, consistent with the available experimental data. This multistep mechanism starts with Glu1261 deprotonating the xanthine at the N3 position followed by a proton transfer from the cofactor to the N9 atom of xanthine; the thus activated cofactor and substrate then react to form a tetrahedral intermediate, and a subsequent rate-limiting hydride transfer generates the product. The substrate orientation that has commonly been assumed in the literature leads to the most stable reactant complex, but the opposite orientation (“upside down”) is computed to be the most favorable one during the reaction (setup G). In the “upside down” conformation, the Arg880 residue can best stabilize the reactive xanthine species with the negatively charged N3 atom, especially the tetrahedral intermediate and the following transition state for hydride transfer which is generally the highest point on the computed energy profiles. QM-only calculations for a minimal gas-phase model and for larger cluster models are performed for comparison, in particular for establishing intrinsic reactivities and a common energy scale. An analysis of the computational results provides detailed insight into the essential mechanistic role of the active-site residues.

### 1. Introduction

Interconversion of xanthine to uric acid is the last step within the purine nucleotide catabolism in human beings as well as primates, birds, reptiles, and insects. It is performed by two kinds of molybdopterin-containing enzymes, called xanthine oxidase (XO) [EC 1.17.3.2] and xanthine dehydrogenase (XDH) [EC 1.17.1.4], the first one giving the name to the whole structural family.<sup>1</sup> They appear as homodimers (see Figure 1) and have common key components, namely, the molybdopterin cofactor (blue), two iron–sulfur clusters (green), and a flavin adenine dinucleotide (FAD, red). They share the basic mechanism for converting xanthine to uric acid, but differ in how they regenerate the initial state of the enzyme (see Scheme 1). Unlike many oxidative enzymes (e.g., cytochrome P450) that consume elementary O<sub>2</sub> and two reductive equivalents, XO and XDH use water as the ultimate oxygen source and produce reductive equivalents.

Early experiments dealing with xanthine oxidase have been reported more than hundred years ago.<sup>2,3</sup> Since then, each step of the catalytic cycle in XO has been intensely investigated. The intraenzymatic electron transfer and the Mo(V) intermedi-



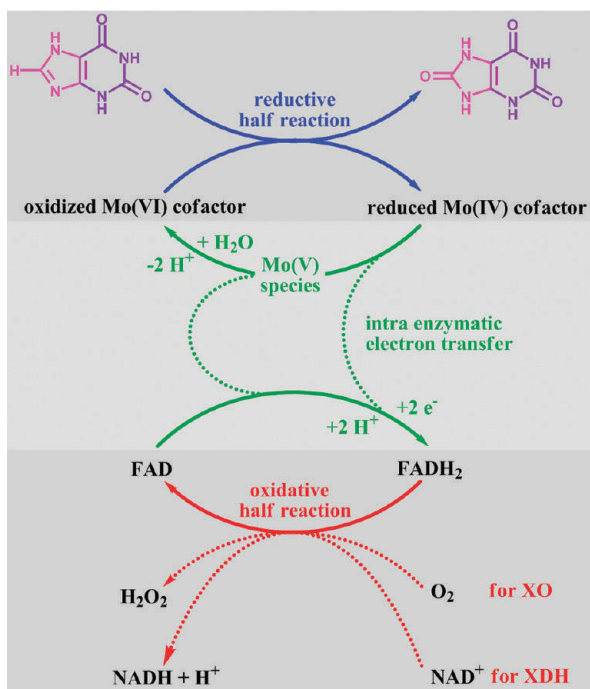
**Figure 1.** Bovine xanthine oxidase dimer (yellow and orange) with its prosthetic groups: the molybdopterin cofactor (blue), the S<sub>2</sub>Fe<sub>2</sub> clusters (green), and FAD (red). Residues that are missing in the crystal structure and have thus been modeled are drawn in purple.

ates have been studied by rapid-freeze-ESR techniques.<sup>4–13</sup> The oxidative half-reaction of XO has also drawn some attention, since it may be a source of reactive oxygen species, such as

- (1) Hille, R. *Chem. Rev.* **1996**, 96, 2757–2816.
- (2) Schardinger, F. Z. *Unters. Nahr. Genussm.* **1902**, 5, 1113–1121.
- (3) Massey, V.; Harris, C. M. *Biochem. Soc. Trans.* **1997**, 25, 750–755.
- (4) Palmer, G.; Bray, R. C.; Beinert, H. *J. Biol. Chem.* **1964**, 239, 2657–2666.
- (5) Bray, R. C.; Palmer, G.; Beinert, H. *J. Biol. Chem.* **1964**, 239, 2667–2676.
- (6) Bray, R. C.; Knowles, P. F.; Pick, F. M.; Vanngard, T. *Biochem. J.* **1968**, 107, 601.
- (7) Bray, R. C.; Vanngard, T. *Biochem. J.* **1969**, 114, 725.

- (8) Swann, J. C.; Bray, R. C. *Eur. J. Biochem.* **1972**, 26, 407.
- (9) Gutteridge, S.; Malthouse, J. P. G.; Bray, R. C. *J. Inorg. Biochem.* **1979**, 11, 355–360.
- (10) Gutteridge, S.; Bray, R. C. *Biochem. J.* **1980**, 189, 615–623.
- (11) Malthouse, J. P. G.; Gutteridge, S.; Bray, R. C. *Biochem. J.* **1980**, 185, 767–770.
- (12) Malthouse, J. P. G.; Williams, J. W.; Bray, R. C. *Biochem. J.* **1981**, 197, 421–425.
- (13) Manikandan, P.; Choi, E. Y.; Hille, R.; Hoffman, B. M. *J. Am. Chem. Soc.* **2001**, 123, 2658–2663.

**Scheme 1.** Sketch of the Mechanism of Xanthine Oxidase and Xanthine Dehydrogenase<sup>a</sup>



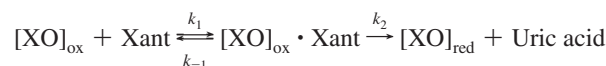
<sup>a</sup> The imidazole part of the substrates is colored in pink, the pyrimidine part in purple. The other color coding is as in Figure 1.

the  $O_2^-$  radical.<sup>14–18</sup> Given the product of its reductive half-reaction, xanthine oxidase was a natural target for drugs against gout, a disease usually caused by having too much uric acid in the body that leads to the deposition of monosodium urate crystals in tissues.<sup>19</sup> For many years, allopurinol has been the “gold standard” inhibitor, but there is an increasing number of other potent inhibitors.<sup>20–28</sup>

Detailed mechanisms of the reductive half-reaction for the xanthine oxidase family have been proposed based on docking studies of xanthine in the structurally similar aldehyde oxi-

doreductase (AOR),<sup>29,30</sup> crystal structures of XO and XDH,<sup>31,32</sup> kinetic experiments with xanthine and related purine substrates for wild-type xanthine oxidase,<sup>33–39</sup> and kinetic experiments with mutants of XO or XDH.<sup>39–42</sup> It has been experimentally shown that water is the source of the oxygen consumed in the biological hydroxylation process and that the labile oxygen should attach to the Mo center at the proximal position prior to the turnover.<sup>43</sup> In a recent EXAFS analysis, the labile oxygen was experimentally identified to be a hydroxide ligand<sup>44</sup> instead of a bound water molecule as previously assumed. Based on the pH-dependence of the XO activity toward xanthine and lumazine<sup>36</sup> as well as 1-methylxanthine,<sup>45</sup> Lewis-base catalysis by glutamate (Glu1261 in XO and Glu869 in AOR) has been proposed<sup>30–32,38,39,42</sup> and confirmed by QM/MM calculations for the similar AOR system.<sup>46</sup> For XO, it is still unclear whether it reacts in a stepwise or a concerted manner or whether there are even further intermediates. The experiments on temperature-dependent transient kinetics suggest that there is at least one intermediate in the course of the turnover of xanthine to the product.<sup>45,47</sup>

Most experimental results<sup>48–50</sup> have been analyzed using simple Michaelis–Menten assumptions:<sup>51,52</sup>

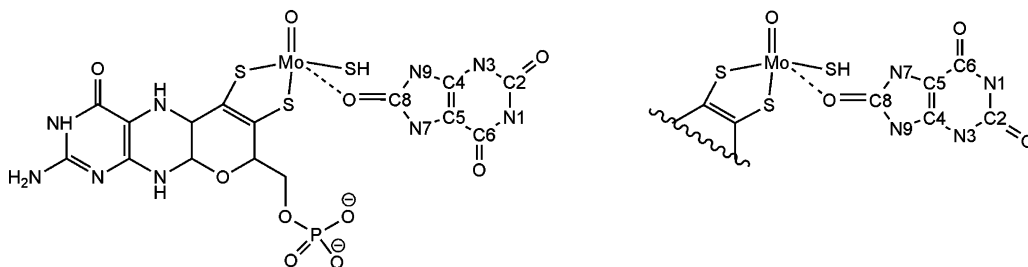


For the reductive half-reaction, an activation barrier of  $E_A = 14.5 \text{ kcal mol}^{-1}$  was obtained from an Arrhenius plot.<sup>50</sup> Using the Eyring equation<sup>53</sup> the reported rate constants of the reductive

- (14) Harrison, R. *Free Radical Biol. Med.* **2002**, *33*, 774–797.
- (15) Olson, J. S.; Ballou, D. P.; Palmer, G.; Massey, V. *J. Biol. Chem.* **1974**, *249*, 4350–4362.
- (16) Hille, R.; Massey, V. *J. Biol. Chem.* **1981**, *256*, 9090–9095.
- (17) Granger, D. N.; Rutili, G.; McCord, J. M. *Gastroenterology*. **1981**, *81*, 22–29.
- (18) Parks, D. A.; Bulkley, G. B.; Granger, D. N. *Surgery* **1983**, *94*, 415–422.
- (19) McCarty, D. J.; Hollander, J. L. *Ann. Intern. Med.* **1961**, *54*, 452–460.
- (20) Baker, B. R.; Hendrickson, J. L. *J. Pharm. Sci.* **1967**, *56*, 955–960.
- (21) Hille, R.; Massey, V. *Pharmacol. Ther.* **1981**, *14*, 249–263.
- (22) Hawkes, T. R.; George, G. N.; Bray, R. C. *Biochem. J.* **1984**, *218*, 961–968.
- (23) Truglio, J. J.; Theis, K.; Leimkühler, S.; Rappa, R.; Rajagopalan, K. V.; Kisker, C. *Structure* **2002**, *10*, 115–125.
- (24) Okamoto, K.; Eger, B. T.; Nishino, T.; Kondo, S.; Pai, E. F.; Nishino, T. *J. Biol. Chem.* **2003**, *278*, 1848–1855.
- (25) Fukunari, A.; Okamoto, K.; Nishino, T.; Eger, B. T.; Pai, E. F.; Kamezawa, M.; Yamada, I.; Kato, N. *J. Pharmacol. Exp. Ther.* **2004**, *311*, 519–528.
- (26) Pacher, P.; Nivorozhkin, A.; Szabo, C. *Pharmacol. Rev.* **2006**, *58*, 87–114.
- (27) Tamta, H.; Thilagavathi, R.; Chakraborti, A. K.; Mukhopadhyay, A. K. *J. Enzyme Inhib. Med. Chem.* **2005**, *20*, 317–324.
- (28) Nagamatsu, T.; Yamasaki, H.; Fujita, T.; Endo, K.; Machida, H. *J. Chem. Soc., Perkin Trans. 1* **1999**, *1*, 3117–3125.
- (29) Romão, M. J.; Archer, M.; Moura, I.; Moura, J. J. G.; Legall, J.; Engh, R.; Schneider, M.; Hof, P.; Huber, R. *Science* **1995**, *270*, 1170–1176.
- (30) Huber, R.; Hof, P.; Duarte, R. O.; Moura, J. J. G.; Moura, I.; Liu, M. Y.; LeGall, J.; Hille, R.; Archer, M.; Romão, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8846–8851.
- (31) Okamoto, K.; Matsumoto, K.; Hille, R.; Eger, B. T.; Pai, E. F.; Nishino, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7931–7936.
- (32) Pauff, J. M.; Zhang, J. J.; Bell, C. E.; Hille, R. *J. Biol. Chem.* **2008**, *283*, 4818–4824.
- (33) Bergmann, F.; Dikstein, S. *J. Biol. Chem.* **1956**, *223*, 765–780.
- (34) Bergmann, F.; Kwietny, H.; Levin, G.; Brown, D. J. *J. Am. Chem. Soc.* **1960**, *82*, 598–605.
- (35) McWhirter, R. B.; Hille, R. *J. Biol. Chem.* **1991**, *266*, 23724–23731.
- (36) Kim, J. H.; Ryan, M. G.; Knaut, H.; Hille, R. *J. Biol. Chem.* **1996**, *271*, 6771–6780.
- (37) Stockert, A. L.; Shinde, S. S.; Anderson, R. F.; Hille, R. *J. Am. Chem. Soc.* **2002**, *124*, 14554–14555.
- (38) Choi, E. Y.; Stockert, A. L.; Leimkühler, S.; Hille, R. *J. Inorg. Biochem.* **2004**, *98*, 841–848.
- (39) Yamaguchi, Y.; Matsumura, T.; Ichida, K.; Okamoto, K.; Nishino, T. *J. Biochem.* **2007**, *141*, 513–524.
- (40) Leimkühler, S.; Stockert, A. L.; Igarashi, K.; Nishino, T.; Hille, R. *J. Biol. Chem.* **2004**, *279*, 40437–40444.
- (41) Hille, R. *Arch. Biochem. Biophys.* **2005**, *433*, 107–116.
- (42) Pauff, J. M.; Hemann, C. F.; Jünemann, N.; Leimkühler, S.; Hille, R. *J. Biol. Chem.* **2007**, *282*, 12785–12790.
- (43) Hille, R.; Sprecher, H. *J. Biol. Chem.* **1987**, *262*, 10914–10917.
- (44) Doonan, C. J.; Stockert, A.; Hille, R.; George, G. N. *J. Am. Chem. Soc.* **2005**, *127*, 4518–4522.
- (45) Sau, A. K.; Mondal, M. S.; Mitra, S. *J. Chem. Soc., Dalton Trans.* **2000**, 3688–3692.
- (46) Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 4628–4640.
- (47) Mondal, M. S.; Mitra, S. *Biochemistry* **1994**, *33*, 10305–10312.
- (48) Edmondson, D.; Ballou, D.; Vanheuve, A.; Palmer, G.; Massey, V. *J. Biol. Chem.* **1973**, *248*, 6135–6144.
- (49) Olson, J. S.; Ballou, D. P.; Palmer, G.; Massey, V. *J. Biol. Chem.* **1974**, *249*, 4363–4382.
- (50) Kim, J. H.; Hille, R. *J. Biol. Chem.* **1993**, *268*, 44–51.
- (51) Michaelis, L.; Menten, M. *Biochem. Z.* **1913**, *49*, 333–369.
- (52) Briggs, G. E.; Haldane, J. B. S. *Biochem. J.* **1925**, *19*, 338–339.
- (53) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107–115.



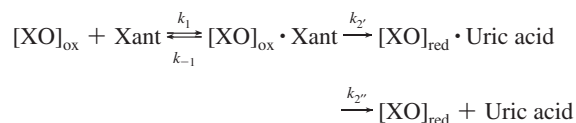
**Scheme 2.** Reduced Molybdopterin Cofactor and Coordinated Product in the “Upside” (left) and “Upside Down” (right) Orientation, with the Numbering Scheme for the Nitrogen and Carbon Atoms<sup>a</sup>



<sup>a</sup> For details on the protonation state, see section 2.2.

half-reaction for XO of  $k_{\text{red}} = 7 \text{ s}^{-1}$  [25 °C, pH = 7.0],<sup>37,38</sup>  $k_2 = 785 \text{ min}^{-1} = 13 \text{ s}^{-1}$  [25 °C, pH = 8.5],<sup>48</sup>  $k_2 = 8 \text{ s}^{-1}$  [25 °C, pH = 8.5],<sup>47</sup>  $k_2 = 1050 \text{ min}^{-1} = 17 \text{ s}^{-1}$  [25 °C, pH = 8.5],<sup>15</sup> and  $k_{\text{cat}} = 1110 \text{ min}^{-1} = 18 \text{ s}^{-1}$  [25 °C, pH = 8.5],<sup>39</sup> translate into free energy barriers of 15.8–16.3 kcal mol<sup>-1</sup>. The published values for the reductive half-reaction of XDH,  $k_{\text{obs}} = 29.3 \text{ s}^{-1}$  [4 °C, pH = 7.8]<sup>42</sup> and  $k_{\text{obs}} = 67.3 \text{ s}^{-1}$  [4 °C, pH = 7.8 ],<sup>40</sup> correspond to free energy barriers of 13.9–14.4 kcal mol<sup>-1</sup>.

The simple Michaelis–Menten scheme has been refined<sup>47,49</sup> by considering the formation of the product complex and product release as two separate steps with rate constants  $k_2'$  and  $k_2''$ , respectively:



For xanthine<sup>54</sup> and 2-oxo-6-methylpurine,<sup>35</sup> product release has proven to be slower than formation of the product complex, and the barriers of free energy obtained for xanthine (using XDH) are  $\Delta G^\ddagger(k_2') = 13.4 \text{ kcal mol}^{-1}$  and  $\Delta G^\ddagger(k_2'') = 15.9 \text{ kcal mol}^{-1}$ .<sup>54</sup> From temperature-dependent steady-state and transient kinetics of the reductive half-reaction, Mondal and Mitra<sup>47</sup> obtained similar results using XO with  $\Delta G^\ddagger(k_2') = 15.1 \text{ kcal mol}^{-1}$  and  $\Delta G^\ddagger(k_2'') = 16.2 \text{ kcal mol}^{-1}$ . In addition, they conclude from the temperature dependency of  $k_2'$ , that even  $\Delta G^\ddagger(k_2')$  is still just an effective barrier, for two consecutive elementary steps of which the first one is endergonic and reversible, whereas the second one is exergonic and irreversible.

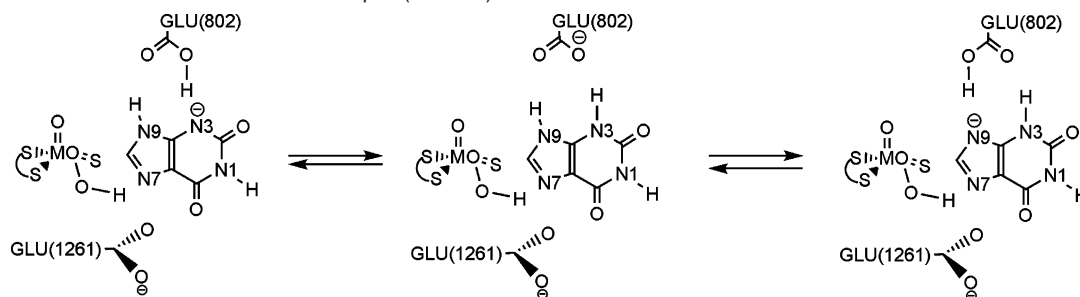
A crucial point for the mechanism of the reductive half-reaction of xanthine oxidase is the substrate orientation within the enzyme. The originally proposed orientation<sup>30</sup> based on docking studies and called “upside” throughout our study (see Scheme 2) has been widely adopted in most subsequent publications. In a recent crystal structure of the inactive desulfo form of XO, xanthine is indeed reported to have this orientation (PDB: 3EUB).<sup>55</sup> On the other hand, there is crystallographic evidence that allopurinol<sup>23,56</sup> and 2-oxo-6-methylpurine,<sup>32</sup> both substrates similar to xanthine, are oriented “upside down”. Hence, different orientations were postulated for “good” substrates like xanthine and “poor” substrates like 2-oxo-6-methylpurine.<sup>32,42</sup> By contrast, a recent mechanistic study<sup>39</sup> proposed

the same “upside down” orientation for xanthine and 2-oxo-6-methylpurine. It should be noted in this context, however, that 2-oxo-6-methylpurine was described as a “poor” substrate because of slow product release.<sup>35</sup> A recent review pointed out that the binding modes of the substrate cannot be clearly distinguished in crystal structures even at a resolution of 1.9 Å and that the electron density may represent a mixture of different substrate orientations.<sup>57</sup>

On the theoretical side, there have been a number of detailed QM model studies that mainly focused on the oxidation of formaldehyde<sup>58–60</sup> and the less active formamide.<sup>60–63</sup> Our recent QM/MM work has addressed the reaction mechanism in the similar AOR system.<sup>46</sup> Further theoretical investigation have been performed on 6-substituted 4-quinazolinones<sup>64</sup> and imidazole<sup>65</sup> as substrates. For xanthine as substrate, there are two QM model studies available. The first one did not model the cofactor explicitly nor give any barrier, but suggested a reasonable change of protonation state in the course of the reaction.<sup>66</sup> The second, very recent study represented the cofactor both as  $[(S-CH=CH-S)Mo(=S)(=O)(-OH)]^-$  anion and as  $[(S-CH=CH-S)Mo(=S)(=O)(-O)]^{2-}$  dianion and treated its reaction with the substrates imidazole and xanthine, considering various protonation states and derivatives of the substrate; xanthine was found to provide much lower barriers than small model substrates due to favorable charge delocalization, and the QM results were interpreted with regard to putative interactions with key residues in the binding pocket.<sup>65</sup> The QM model calculations afford insight into intrinsic reactivities, but they are limited in scope because they do not explicitly account for the steric and electronic influence of nearby active-site residues such as Glu802, Arg880, or Glu1261. Such interactions are captured quite naturally at the QM/MM level, and they have been shown to be of crucial importance in our recent QM/MM work on AOR.<sup>46</sup>

- (54) Schopfer, L. M.; Massey, V.; Nishino, T. *J. Biol. Chem.* **1988**, 263, 13528–13538.
- (55) Pauff, J. M.; Cao, H.; Hille, R. *J. Biol. Chem.* **2009**, 284, 8760–8767.
- (56) Okamoto, K.; Eger, B. T.; Nishino, T.; Pai, E. P.; Nishino, T. *Nucleosides, Nucleotides Nucleic Acids* **2008**, 27, 888–893.

- (57) Nishino, T.; Okamoto, K.; Eger, B. T.; Pai, E. F.; Nishino, T. *FEBS J.* **2008**, 275, 3278–3289.
- (58) Bray, M. R.; Deeth, R. J. *J. Chem. Soc., Dalton Trans.* **1997**, 4005–4009.
- (59) Voityuk, A. A.; Albert, K.; Romão, M. J.; Huber, R.; Rösch, N. *Inorg. Chem.* **1998**, 37, 176–180.
- (60) Zhang, X. H.; Wu, Y. D. *Inorg. Chem.* **2005**, 44, 1466–1471.
- (61) Ilich, P.; Hille, R. *J. Phys. Chem. B* **1999**, 103, 5406–5412.
- (62) Ilich, P.; Hille, R. *J. Am. Chem. Soc.* **2002**, 124, 6796–6797.
- (63) Amano, T.; Ochi, N.; Sato, H.; Sakaki, S. *J. Am. Chem. Soc.* **2007**, 129, 8131–8138.
- (64) Alfaro, J. F.; Jones, J. P. *J. Org. Chem.* **2008**, 73, 9469–9472.
- (65) Bayse, C. A. *Dalton Trans.* **2009**, 2306–2314.
- (66) Ilich, P.; Hille, R. *Inorg. Chim. Acta* **1997**, 263, 87–93.

**Scheme 3.** Different Protonation States within Setup B (see text)<sup>a</sup>

<sup>a</sup> The same interconversion is possible for setup A, if one adds a proton on N7.

## 2. Computational Methods

**2.1. System Preparation and Classical Simulations.** The complex between the substrate and the reduced molybdopterine cofactor (see Scheme 2) is widely accepted as intermediate of the reductive half-cycle, especially as an analogous structure has been found crystallographically for a different substrate.<sup>31</sup> Therefore we modeled this state during system setup. When the project was started, the crystal structures with xanthine in the binding pocket of the inactive desulfo form (PDB: 3EUB)<sup>55</sup> and with 2,8-dioxo-6-methylpurin in the binding pocket (PDB: 3B9J)<sup>32</sup> were not yet available, so we used the crystal structure of bovine XO in the XDH form from the Protein Data Bank (PDB: 1FO4).<sup>24</sup> It contains the dimeric unit, with the residues 1–2, 166–191, and 532(529 in chain B)–536 missing in each chain (see Figure 1), and 2049 crystal water molecules. In this crystal structure, the pocket of the enzyme is occupied by a salicylate molecule, which had to be replaced by uric acid. This was done by performing QM/MM optimizations on docking structures obtained with the program package AutoDock.<sup>67</sup> The glycerol molecules and the calcium ions present in the PDB file were kept. We used SWISS-MODEL<sup>68–71</sup> to add the unresolved residues (except for residues 1–2) by homology modeling. This should cause negligible errors since these residues are far away from the active site. The program *reduce*<sup>72</sup> was used to check and adjust the orientation of Asn, Gln, and His side chains, as well as the protonation states of the His side chains, which were rechecked by visual inspection. After having chosen the protonation state of all titratable residues (see section 2.2), we placed chloride counterions to neutralize the system, using the program *ionize*.<sup>73</sup>

Missing hydrogen atoms in the X-ray structure were added using the CHARMM program,<sup>74,75</sup> which was also employed for classical minimization and molecular dynamics (MD) runs during the hydration and equilibration procedure. We adopted the parameters (all except the charges which were recalculated) for the substrate from guanine and uracil, and took those for the molybdopterine and Fe<sub>2</sub>S<sub>2</sub> cluster from previous calculations.<sup>46</sup> The FAD parameters were taken from the literature.<sup>76</sup> All nonstandard parameters are provided in the Supporting Information.

We defined an active region including all residues within 20 Å of the substrate C8 atom bound to the molybdopterine cofactor of chain A, C8:URIC. Within this region and for all following MD simulations, the [Mo(S<sub>2</sub>C<sub>2</sub>H<sub>2</sub>)(=O)(OR<sub>unfixed</sub>)(–SH)]<sup>2–</sup> moiety and the Fe<sub>2</sub>S<sub>2</sub> cluster were kept fixed. A 35 Å sphere of equilibrated water molecules was superimposed on the enzymatic system, centered at C8:URIC and all water molecules too close to existing atoms were deleted. During the MD simulations, a spherical potential was imposed on the water sphere to prevent water molecules from escaping into the vacuum. Energy minimization and a 100 ps molecular dynamics (MD) simulation at 300 K were performed using the CHARMM force field<sup>75</sup> as implemented in the CHARMM program package.<sup>74</sup> This solvation procedure was iterated until the number of added water molecules was approximately constant. During this procedure, positional restraints were applied to the residues in the active region which were

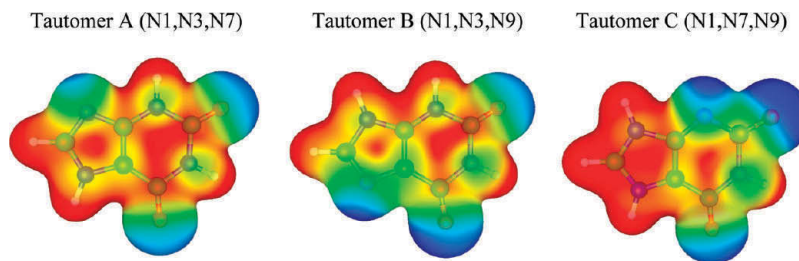
successively lowered in each iteration. The final MD simulation of 500 ps was performed without any such restraints.

For the setup of systems with different protonation states or different substrate orientation, we started from the final structure of this 500 ps equilibration run and modified it to our needs. Thereafter it was subjected to QM/MM minimization followed by a 100 ps pre-equilibration MD run with additional restraints applied to the residues in the active region and a 500 ps MD run without these constraints, from which we took our snapshots. See the Supporting Information for full details on the hydration procedure.

**2.2. Protonation States of the Active Site and Substrate.** The protonation states of the substrate and of the active-site residues are crucial factors during system setup. It is known from experiments that xanthine at physiological pH consists of a 1:1 mixture of neutral and anionic forms.<sup>77</sup> Additionally, within each of these forms, there are several tautomers present. Based on pH-dependent kinetic measurements it was concluded that XO acts on a neutral rather than on an anionic substrate.<sup>36</sup> The protonation state of the substrate itself is again dependent on the surrounding residues and their protonation states (see Scheme 3). It has been postulated that the protonation state of xanthine changes during the course of the reaction,<sup>66</sup> which can be rationalized by electrostatic potential plots of xanthine in its different protonation states, see Figure 2. The relative energies of the tautomers from gas-phase calculations indicate that strong interactions with the surrounding enzyme are necessary to stabilize tautomer C, see Table 1.

It is widely accepted by experimentalists<sup>30–32,38,39,42</sup> that Glu1261 is initially deprotonated. This is consistent with the results of our recent QM/MM study on AOR,<sup>46</sup> which explicitly considered both protonation states of Glu1261 (Glu869 in AOR). Mutation studies of XDH from *Rhodobacter capsulatus*<sup>42</sup> as well as human XO<sup>39</sup> have shown that Arg880 (Arg310 in XDH, Arg881 in human XO) is crucial for enzymatic activity, presumably because it stabilizes the transition state by compensating negative charge

- (67) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.
- (68) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714–2723.
- (69) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, *31*, 3381–3385.
- (70) Kopp, J.; Schwede, T. *Nucleic Acids Res.* **2004**, *32*, D230–D234.
- (71) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, *22*, 195–201.
- (72) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735–1747.
- (73) <http://www.ks.uiuc.edu/Development/MDTools/ionize>.
- (74) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (75) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (76) Luo, G. B.; Andricioaei, I.; Xie, X. S.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 9363–9367.
- (77) Kulikowska, E.; Kierdaszuk, B.; Shugar, D. *Acta Biochim. Pol.* **2004**, *51*, 493–531.
- (78) Kim, J. H.; Odutola, J. A.; Popham, J.; Jones, L.; von Laven, S. *J. Inorg. Biochem.* **2001**, *84*, 145–150.
- (79) Rogstad, K. N.; Jang, Y. H.; Sowers, L. C.; Goddard, W. A. *Chem. Res. Toxicol.* **2003**, *16*, 1455–1462.



**Figure 2.** Electrostatic potential for different protonation states of xanthine mapped on the density isosurface (0.01e). Blue (red) color represents a negative (positive) potential. A large positive potential at C8 (leftmost ring atom) will facilitate nucleophilic attack.

**Table 1.** Relative Energies of Tautomers A–C in kcal mol<sup>−1</sup>

	$\Delta E^a$	$\Delta G^{\text{gas } b}$	$\Delta G^{\text{solution } b}$	$\Delta E^c$	$\Delta E^d$
tautomer A	0.0	0.0	0.0	0.0	0.0
tautomer B	8.9	9.1	0.6	9.3	9.2
tautomer C	not calculated	not calculated	not calculated	23.1	23.0

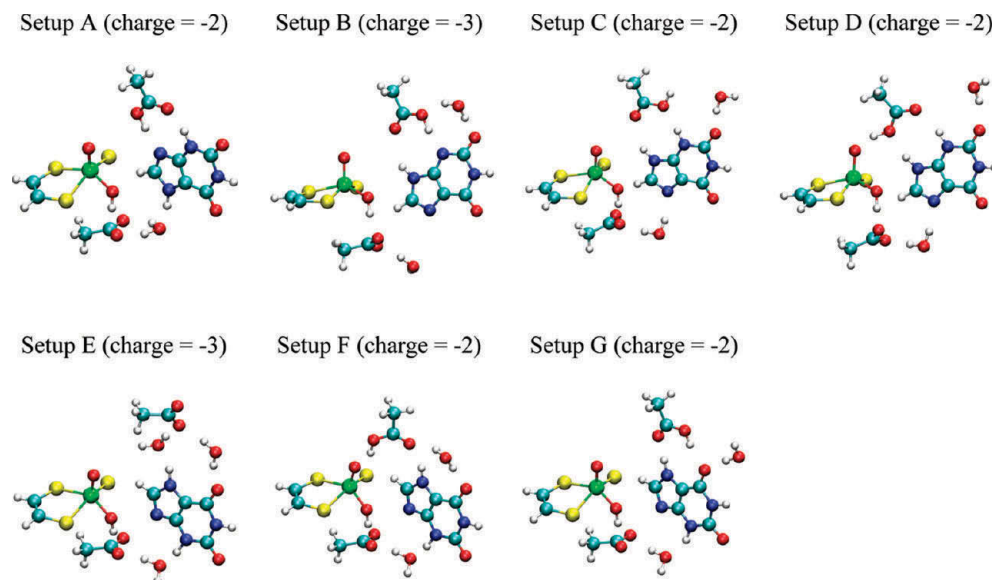
<sup>a</sup> B3LYP/6-31G(d). <sup>b</sup> B3LYP/6-31++G\*\*. <sup>c</sup> B3LYP/6-31+G\*\*. <sup>d</sup> B3LYP/def2-TZVP.

accumulation on XAN:O6 in the course of the reaction with its positive countercharge.<sup>32,42</sup> In view of these experimental results and according to the standard protonation state at physiological pH, we took Arg880 to be protonated. On the basis of the available crystal structures with different substrates/inhibitors,<sup>23,25,39,56</sup> Glu802 is normally assumed to be protonated and to participate in hydrogen bonding of the enzyme to the substrate, but one may expect facile deprotonation.

Taking the various possibilities into account, we set up four systems for the “upside” orientation of xanthine: Neutral xanthine (N1,N3,N9) with protonated Glu802 (setup A), neutral xanthine (N1,N3,N9) with deprotonated Glu802 that converts to a deprotonated substrate and protonated Glu802, see Scheme 3 (setup B), and neutral xanthine (N1,N3,N9) with protonated Glu802, the latter coordinating via a hydrogen bond to a water molecule of the solvent (setup C) or coordinating via a hydrogen bond to the apical oxygen atom of the cofactor (setup D). For the “upside down” orientation we set up three systems: Neutral xanthine (N1,N3,N7) with deprotonated Glu802 (setup E), with protonated Glu802 where the proton is pointing toward the apical oxygen of the cofactor (setup F), and with protonated Glu802 where the proton is pointing toward the O6 atom of the substrate (setup G). The QM regions of all setups are shown in Figure 3.

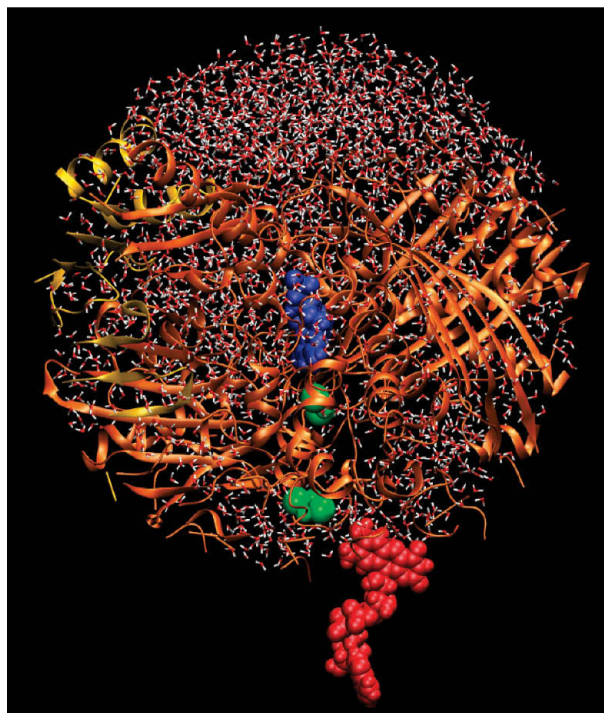
**2.2. QM/MM Calculation.** To reduce computational effort, we cut down the full dimeric system, applying a cutoff radius around C8:URIC of 35–38 Å dependent on the individual setup and snapshot, and making sure that the resulting system is still neutral in charge (see Figure 4). QM/MM geometry optimizations of the stationary points were performed with a linear scaling microiterative algorithm working in hybrid delocalized coordinates.<sup>80</sup> All residues and water molecules within 13 Å of the substrate were included in the optimization; the remaining atoms were kept fixed. Additionally, we reoptimized all stationary points for snapshot 400 of setup G using the complete enzymatic system. All QM/MM calculations were performed with the modular program package ChemShell,<sup>81,82</sup> using the program package TURBOMOLE<sup>83</sup> to obtain the QM-(DFT) wave functions as well as the corresponding energies and gradients. MM energy and gradient were evaluated by DL\_POLY,<sup>84</sup> which is provided in the ChemShell package, using the CHARMM topology and parameter data. We employed electrostatic embedding for the QM region.<sup>85</sup> No cutoff was applied for the MM point charges when calculating the electrostatic interaction within the MM region or the electrostatic QM/MM interaction. To prevent overpolarisation at the QM/MM boundary we applied the *charge-shift* scheme<sup>86,87</sup> for all setups. We also cross-checked the results for setups A and G using the alternative *L2* link atom scheme.<sup>81,88</sup>

The QM regions for the different setups contained a truncated model of the molybdopterin cofactor, part of the side chains of Glu802 and Glu1261, the substrate and one or two water molecules (see Figure 3), one of which proves to be mandatory for the reaction to occur. This partitioning coincides with the CHARMM charge-group boundaries, cutting only C–C single bonds across the QM/MM boundary. For comparison, gas-phase model calculations were performed using the smallest possible setup, namely the truncated model of the cofactor and the substrate. An extended QM region

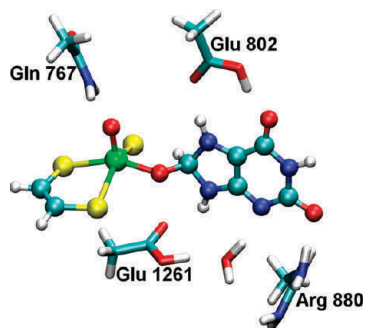


**Figure 3.** QM regions of the seven chosen setups.





**Figure 4.** System used in QM/MM calculations to establish the pathways within setup A–G. Same color coding as in Figure 1.



**Figure 5.** Extended QM region used for single-point energy evaluations (see text and section 3.9). The calculations on the smallest possible gas-phase model for setup G included only the atoms represented in balls and sticks (see section 3.10). Shown is the tetrahedral intermediate (IM2) of setup G.

that includes parts of the side chains of Gln767, Glu802, Arg880, and Glu1261 (see Figure 5) was used for additional QM/MM optimizations and gas-phase as well as COSMO<sup>89</sup> corrected single-point calculations to establish a common energy scale for the different setups.

Starting from the product-bound structure, we first used the BP86 functional<sup>90–94</sup> to determine various minima on the potential energy surface, employing the resolution of the identity (RI) approximation.<sup>95,96</sup> Easily detectable transition structures were localized starting from a preoptimized structure obtained by constrained optimization. In difficult cases we used the nudged elastic band (NEB) method<sup>97</sup> in combination with transition-state optimization to determine the mini-

mum energy pathway and the associate transition state. All stationary points were reoptimized using the B3LYP hybrid functional<sup>90–92,98–100</sup> as implemented in TURBOMOLE.

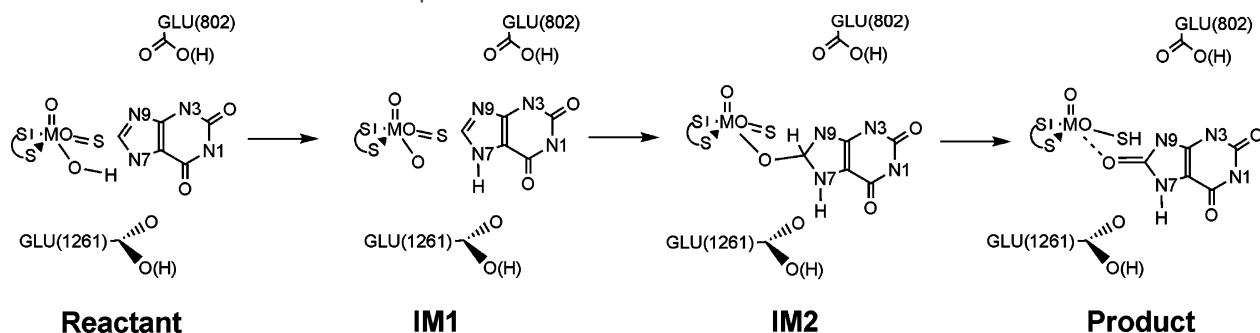
Our standard basis set (B1) was composed as follows: Mo, Lanl2DZ<sup>101</sup> with an additional f polarization function;<sup>102</sup> S, Lanl2DZ<sup>103</sup> with an additional d polarization function;<sup>104</sup> and all other atoms (H, C, O, N) 6-31+G\*\*.<sup>105,106</sup> The BP86 calculations employed the def2-TZVP auxiliary basis set.<sup>107</sup> This combination has already proven to give reasonable results.<sup>46</sup> For the most relevant setup G, we reoptimized all stationary points for snapshot 400 using the def2-TZVP basis set<sup>108</sup> (B2) for all atoms. To check the def2-TZVP results, we replaced the basis of the molybdenum atom by an all-electron basis set TZVPalls2<sup>109</sup> while keeping the def2-TZVP basis for all other atoms (B3). In these latter calculations, we used the corresponding auxiliary basis set<sup>107</sup> with BP86. To check the influence of relativistic effects on these results, we applied the ZORA<sup>110,111</sup> and the Douglas–Kroll–Hess<sup>112–114</sup> methods as implemented in ORCA.<sup>115</sup> All figures showing molecular structures were generated using VMD.<sup>116</sup>

### 3. Results

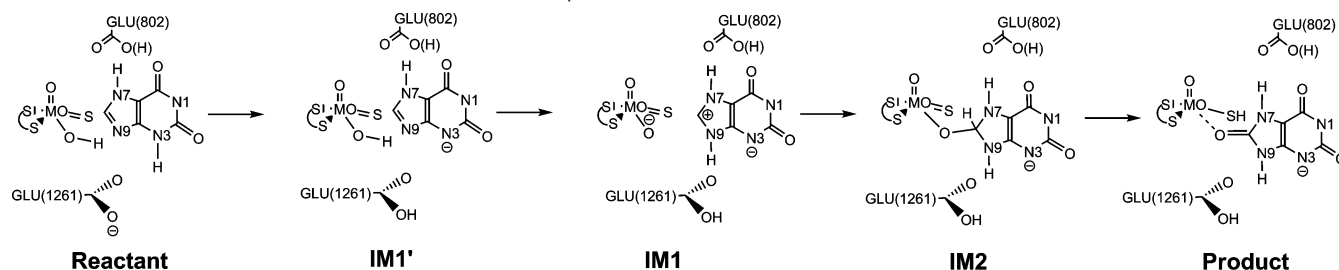
Our mechanistic studies on xanthine oxidase at the QM/MM level start from the product-bound state. Seven setups were

- (86) de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133–6141.
- (87) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss. Chem. Soc.* **1997**, 79–92.
- (88) Antes, I.; Thiel, W. In *Hybrid Quantum Mechanical and Molecular Mechanical Methods*; Gao, J.; Thompson, M. A., Eds.; ACS Symposium Series 712, American Chemical Society: Washington, DC, 1998; pp 50–56.
- (89) Klamt, A.; Schürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 5, 799–805.
- (90) Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.
- (91) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (92) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (93) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (94) Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
- (95) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.
- (96) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
- (97) Henkelman, G.; Uberuaga, B. P.; Jonsson, H. *J. Chem. Phys.* **2000**, *113*, 9901–9904.
- (98) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (99) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (100) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (101) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270–283.
- (102) Ehlers, A. W.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 111–114.
- (103) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284–298.
- (104) Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 237–240.
- (105) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- (106) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (107) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (108) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (109) Ahlrichs, R.; May, K. *Phys. Chem. Chem. Phys.* **2000**, *2*, 943–945.
- (110) Lenthe, E. v.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1993**, *99*, 4597–4610.
- (111) Wüllen, C. V. *J. Chem. Phys.* **1998**, *109*, 392–399.
- (112) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89–155.
- (113) Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756–763.
- (114) Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742–3748.
- (115) www.thch.uni-bonn.de/tc/orca.
- (116) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

- (80) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
- (81) Sherwood, P.; et al. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1–28.
- (82) www.chemshell.org.
- (83) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (84) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136–141.
- (85) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

**Scheme 4.** General Reaction Schemes for Setups A–D<sup>a</sup>

<sup>a</sup> The reactant is converted into IM1 by deprotonation of the initial hydroxo group of the cofactor. The proton at N7 does not in general originate from the cofactor but may have been present already in the reactant. C–O bond formation in IM1 yields the tetrahedral intermediate IM2. Subsequent hydride transfer generates the product. Depending on the setup, there may be additional intermediates; more detailed schemes for each of the setups A–D are shown in the Supporting Information.

**Scheme 5.** General Reaction Mechanism Scheme for Setups E–G<sup>a</sup>

<sup>a</sup> In the reactant structure, the substrate is protonated at N3. It is converted into IM1' by transferring this proton to Glu1261. A second proton transfer, this time from the cofactor to the substrate, yields IM1 with neutral substrate and activated cofactor. The tetrahedral intermediate IM2 is formed by subsequent C–O bond formation. The product is obtained in a final hydride transfer step. Depending on the setup, there may be additional intermediates; more detailed schemes are shown for setup E in the Supporting Information and for setups F and G in Figures 6 and 7.

investigated, four of them (A–D) with the “upside” orientation of the substrate, and three of them (E–G) with the “upside down” orientation (see Scheme 2), taking various protonation states of the substrate and the nearby residues into account. From the 500 ps production MD with (approximately) constant temperature of 300 K (temperature coupling constant = 5 ps) we took a snapshot each 100 ps and performed a QM/MM energy minimization. For each setup, we selected two or three of these snapshots (SN) as being representative (by visual inspection), in some cases including a third one with a slightly different conformational arrangement (e.g., setup G, SN200). Reaction pathways were calculated for these selected snapshots. They share some key features with the most favorable mechanism for the reductive half cycle of acetaldehyde in AOR<sup>46</sup> (see Schemes 4 and 5). However, in contrast to the AOR system, we could not identify alternative pathways: Due to the much more specific binding of the substrate, there is no metal center activated or one-step mechanism<sup>46</sup> for XO. Attempts to find such pathways were unsuccessful and led to the mechanisms reported below. In addition, we carefully checked alternative protonation states for all intermediates and all snapshots. Unless mentioned otherwise (e.g., setup D, SN400) these protonation states are energetically irrelevant for the reaction mechanisms. Attempts to find pathways, where the activated cofactor acts on a previously protonated substrate, did not succeed.

In the following we present the QM/MM results for the individual setups A–G, a QM model study for the most favorable setup G, and a comparison between the seven pathways considered. Detailed mechanistic schemes with optimized geometries are provided in Figures S1–S7 of the Supporting Information (SI) for setups A–E and in Figures 6–7 for setups F–G. All energy values given in this section are QM/

**Table 2.** QM/MM Energies in kcal mol<sup>−1</sup> Calculated for Setup A, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN100-BP86	SN100-B3LYP	SN500-BP86	SN500-B3LYP
reactant	0.0	0.0	0.0	0.0
TS1'	4.2 (4.2)	7.0 (7.0)	3.6 (3.6)	8.9 (8.9)
IM1'	0.3	−0.4	−0.7	2.4
TS1	15.8 (15.5)	14.8 (15.2)	18.3 (19.0)	19.1 (16.7)
IM1	11.8	12.2	13.6	14.9
TS2	17.2 (5.4)	20.2 (8.0)	16.1 (2.5)	20.3 (5.4)
IM2	15.2	17.9	13.1	16.4
TS3	18.8 (3.6)	20.8 (2.9)	17.9 (4.8)	21.7 (5.3)
product	12.1	1.2	9.6	0.5
$\Delta E_{\max}$	18.8	21.2	19.0	21.7

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

MM energies (i.e., including QM, MM, and QM/MM interaction terms), without zero-point and thermal corrections.

**3.1. Setup A (see Figure S1, SI).** The reaction starts with a tautomerization of neutral xanthine from its (N1,N3,N7) protonation state, which is predominant in aqueous solution,<sup>77</sup> to its (N1,N7,N9) form. This step is followed by a second proton transfer, this time from the labile OH group to Glu1261. The thus activated labile oxygen at the cofactor then attacks the substrate to form a tetrahedral intermediate. In the final step, the H8:URIC atom is transferred to the sulfido-group of the cofactor, and the product becomes planar again.

According to the calculated relative energies (see Table 2) the two xanthine tautomers (N1,N3,N7 in the reactant, N1,N7,N9 in IM1') are comparably stable in the enzyme (within 1–2 kcal mol<sup>−1</sup>). As in the case of AOR,<sup>46</sup> the BP86 and B3LYP energetics are generally similar, except for the Mo(IV) product

**Table 3.** QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup B, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN300-BP86	SN300W-BP86	SN300-B3LYP	SN500-BP86	SN500-B3LYP
reactant	0.0	0.0	0.0	0.0	0.0
TS1	19.8 (19.8)	16.2 (16.2)	18.4 (18.4)	18.6 (18.6)	20.2 (20.2)
IM1'	9.8	7.8	11.5	14.3	14.1
TS1'	11.9 (2.1)	9.7 (1.9)	15.1 (3.6)	16.3 (2.0)	20.4 (6.3)
IM1	8.4	8.7	10.8	14.4	16.0
TS2	10.7 (2.3)	14.0 (5.3)	17.9 (7.1)	14.3 (0.1)	18.0 (2.0)
IM2	8.6	12.9	14.4	10.9	14.2
TS3	14.5 (5.9)	18.5 (5.6)	19.9 (5.5)	15.9 (5.0)	19.5 (5.3)
product	8.6	13.0	0.3	12.7	3.5
$\Delta E_{\max}$	19.8	18.5	19.9	18.6	20.4

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses. SN300W describes the pathway with a water molecule positioned between Glu1261 and the substrate.

with its d<sup>2</sup> configuration at the molybdenum center. Referring to high level ab initio results,<sup>63</sup> we consider the B3LYP values to be more realistic. For SN500, the reaction barrier (B3LYP) which is determined by the energy difference between the reactant complex and the transition state TS3 for the hydride transfer (see Table 2) is calculated to be about 20 kcal mol<sup>-1</sup>, and thus considerably higher than the experimental value of around 15 kcal mol<sup>-1</sup>, see section 1. As DFT methods are expected to underestimate rather than to overestimate energy barriers,<sup>117</sup> it is unlikely that setup A represents the enzymatic pathway.

We performed two further checks for SN500 at the B3LYP level. First, the use of the *L2* instead of the *charge shift* link atom scheme leads to only minor changes in the computed overall barrier (of around 1 kcal mol<sup>-1</sup>, see Tables S2 and S3, SI). Second, deprotonation of Glu802 is unfavorable in setup A: deprotonated Glu802 abstracts a proton from the substrate in the reactant state, the tetrahedral intermediate is very high in energy (no effective stabilization of the accumulating charge at O6 by the environment), and the overall barrier increases up to more than 40 kcal mol<sup>-1</sup> (see Table S4 and Figure S2, SI).

**3.2. Setup B (see Figure S3, SI).** The essential difference between setup A and previously reported QM/MM results for AOR<sup>46</sup> is the missing hydrogen bond acceptor functionality of the substrate close to Glu1261. We therefore chose the (N1,N3,N9) protonation state (see Figure 2) in setup B, in combination with a deprotonated Glu802. However, during QM/MM optimization, one of the protons (H3) moves to Glu802, leading to an anionic substrate in the reactant.

In the initial deprotonation step, a proton is transferred from the labile oxygen to Glu1261 via TS1. The substrate is so far still negatively charged, which is favorable for reprotonation, but not for nucleophilic attack. So a second proton transfer occurs, this time from Glu1261 to the N7 atom of the substrate (via TS1'). Having thus regained a neutral substrate with both nitrogen atoms of the imidazole ring being protonated (see Figure 2), a C–O bond is formed in the next step (via TS2). Thereafter, H8:URIC is transferred to the sulfido-group (via TS3), producing the Mo(IV) product species.

Judging from the computed relative energies (Table 3) the highest point on the reaction profile corresponds either to TS1 (snapshots SN300-BP86 and SN500-BP86), TS1' (SN500-B3LYP) or TS3 (SN300W-BP86 and SN300-B3LYP). It should be noted, however, that these transition states are normally quite

**Table 4.** QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup C, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN200-BP86	SN200-B3LYP/BP86	SN200-B3LYP	SN400-BP86	SN400-B3LYP
reactant	0.0	0.0	0.0	0.0	0.0
TS1	9.6 (9.6)	11.1 (11.1)	10.9 (10.9)	13.2 (13.2)	14.9 (14.9)
IM1	5.5	8.0	7.9	11.0	13.4
TS2	10.5 (5.0)	18.1 (10.1)	17.9 (10.0)	18.2 (7.2)	26.9 (13.5)
IM2	8.3	16.5	16.3	13.5	18.5
TS3	12.7 (4.4)	18.4 (1.9)	19.1 (2.8)	19.8 (6.3)	27.9 (5.0)
product	5.6	1.1	0.3	11.0	2.9
$\Delta E_{\max}$	12.7	18.4	19.1	19.8	27.9

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

close in energy (e.g., within 1 kcal mol<sup>-1</sup> for SN500-B3LYP). More importantly, each of these effective barriers is computed to be around 20 kcal mol<sup>-1</sup>, and, thus again, too high compared with experiment.

In snapshot SN300W we checked the influence of a water molecule positioned between Glu1261 and the substrate. Compared with snapshot SN300 the overall mechanism remains unchanged. The two initial proton transfers become slightly more facile (TS1 and TS1' energies lowered), presumably because the more extended hydrogen bonding network in SN300W provides additional stabilization. By contrast, the final two steps become less favorable (TS2 and TS3 energies raised) such that the effective overall barrier does not change much.

In summary, the nucleophilic attack of the cofactor involves a neutral substrate molecule, in both setups A and B, since the initially deprotonated, negatively charged substrate in setup B is reprotonated by Glu1261. The essential difference between setup A and setup B is therefore not the protonation state of the substrate, but the protonation state of Glu1261 during the nucleophilic attack. Nevertheless, both of these setups provide a rather high activation barrier.

**3.3. Setup C (see Figure S4, SI).** In the two previous pathways, the energy of the tetrahedral intermediate (IM2) is in most of the cases already about 15 kcal mol<sup>-1</sup> higher than that of the reactant complexes. In setup C, we used a neutral (N1,N3,N9) substrate in combination with a protonated Glu802 to check whether a presumably more stable all-N protonated tetrahedral intermediate is encountered in this case.

The reaction starts again with a proton transfer from the cofactor (via TS1) to Glu1261 coordinating to a nearby water molecule in IM1. Thereafter the thus activated labile oxygen atom performs a nucleophilic attack on the aromatic carbon atom of the imidazole ring, C8:URIC (via TS2). In the last step, the tetrahedral intermediate is converted to the product complex by transfer of the H8:URIC hydrogen atom to the sulfido-group (via TS3). Remarkably, the proton that originally resides at the cofactor and that is transferred to Glu1261 does not move on to activate the imidazole ring. So the activated cofactor still acts on the neutral substrate.

The overall barrier for setup C is determined by the energy difference between the reactant complex and the transition state TS3 of the hydride transfer, see Table 4. There is one low overall barrier (SN200-BP86) which however increases substantially in a B3LYP-based single-point calculation (SN200-B3LYP//BP86). Since we consider the B3LYP energetics more trustworthy, we conclude that the overall barrier for setup C is still around 20 kcal mol<sup>-1</sup> and therefore too high with respect to experimental values.

(117) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Science* **2008**, *321*, 792–794.



**Table 5.** QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup D, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN200-BP86		SN200-B3LYP		SN400-BP86		SN400-B3LYP	
reactant	0.0		0.0		0.0		0.0	
TS1''	1.0	(1.0)	5.4	(5.4)				
IM1''	-1.0		4.8					
TS1''	1.0	(2.0)	6.6	(1.8)	0.6	(0.6)	3.0	(3.0)
IM1''	-2.2		3.8		-5.4		-3.5	
TS1	0.2	(2.4)	7.6	(3.8)	-2.3	(3.1)	2.3	(5.8)
IM1	0.2		6.7		-4.6		-1.7	
TS2	13.9	(13.7)	27.0	(20.3)	15.7	(20.3)	25.0	(26.7)
IM2	13.7		25.9		14.5		23.9	
TS3	18.0	(4.3)	28.5	(2.6)	18.5	(4.0)	28.5	(4.2)
product	7.5		6.0		4.7		-0.1	
$\Delta E_{\max}$	20.2		28.5		23.9		32.0	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

**3.4. Setup D (see Figure S5, SI).** In setups A–C the intermediates with deprotonated cofactor are about 10–15 kcal mol<sup>-1</sup> higher in energy than the initial reactant complex. This raise in energy could, at least partially, be compensated by directing the proton of Glu802 toward the apical oxygen of the cofactor which becomes significantly more acidic during deprotonation. We tested this pathway using the (N1,N3,N9) tautomer state of the xanthine substrate.

The initial deprotonation step is calculated to occur in two steps for snapshot SN200 and in a single step for snapshot SN400. In the additional minimum for snapshot SN200 (IM1') the labile proton is already transferred to Glu1261, but is still pointing toward the activated oxygen OM2:MOCO. In the next minimum structure (IM1''), the proton is hydrogen bonded to the water molecule residing between Glu1261 and the substrate. By losing the coordination of this proton, the molybdopterine cofactor accumulates negative charge at the apical oxygen atom, which enables a concomitant double proton transfer (from N9: URIC to OE2:Glu802 and from OE1:Glu802 to OM1:MOCO via TS1'') that forms an apical OH group in IM1''. The difference between the two snapshots in the initial phase of the reaction is mechanistically not relevant since it does not affect the general shape of the energy profile (and, in particular, not the rate-limiting part). In the next phase, another proton is transferred, in both snapshots, via the water molecule on the substrate (IM1'' → IM1). So we once again do not obtain a protonated substrate in IM1, but generate a cofactor, which is protonated at the apical position and whose labile oxygen is expected to be a less active nucleophilic agent. At the same time, the active substrate does not bear two protons at the imidazole ring, which would have facilitated the formation of the tetrahedral intermediate (IM2). For SN400 we find two

possible pathways for the interconversion of the tetrahedral intermediate to the product complex. In the first one, the proton is still residing on the apical oxygen (as in SN200), while in the second one, this proton is removed onto Glu802. These two pathways are energetically very close; results are given for the former in Table 5 and for the latter in the SI.

In contrast to setups A–C, most of the structures with the labile oxygen of the molybdopterine cofactor being deprotonated are now lower in energy than the reactant complex, due to the favorable additional proton transfer to the apical oxygen. Thus they serve as reference to determine the overall barrier in setup D, which is calculated to be about 25–30 kcal mol<sup>-1</sup> for the different snapshots.

The pathways examined so far, all with the substrate in “upside” orientation, give energy barriers, which are too high with respect to experiment and too high to expect an effective catalysis and turnover of the enzyme. We therefore studied three additional setups E–G with “upside down” orientation (see Scheme 5).

**3.5. Setup E (see Figure S6, SI).** We first assumed Glu802 to be deprotonated. During the MD simulation the Glu802 side chain moves away from the cofactor and even loses the direct hydrogen bond to the substrate, which is replaced by a hydrogen bond to a bridging water molecule. For the conversion of the reactant to the product complex, we find a four-step mechanism, with an initial proton transfer from the substrate to Glu1261, followed by a second proton transfer from the labile OH group to N9 of the substrate. The thus activated cofactor forms a C–O bond with the substrate leading to the tetrahedral intermediate which subsequently reacts via a hydride transfer to the product.

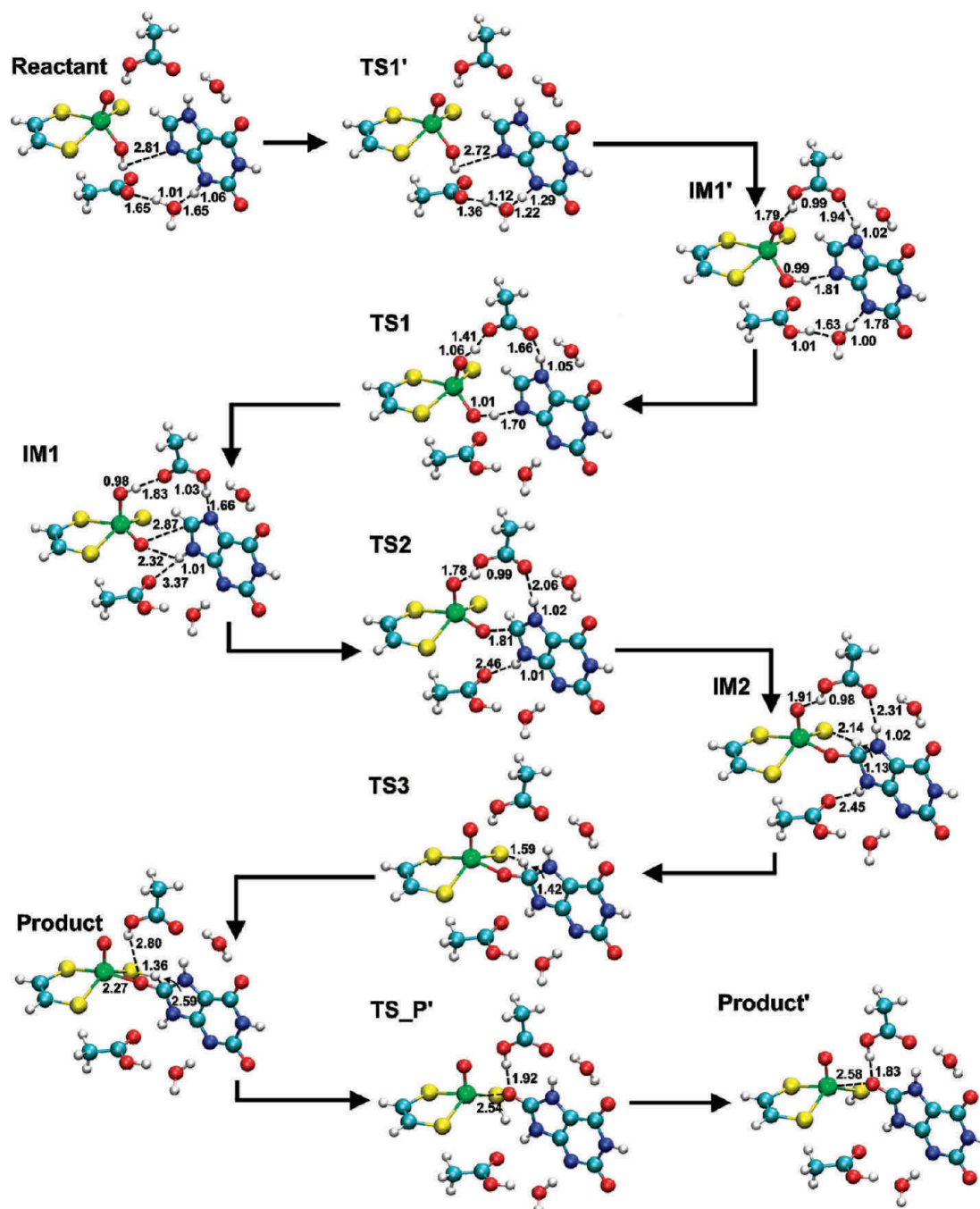
Comparing the calculated results for the three investigated snapshots, the reaction barrier is determined by the energy difference between IM1' and the transition state TS3 for the hydride transfer (see Table 6). For snapshot 100, we find the lowest overall barriers of 13 kcal mol<sup>-1</sup> (with BP86) and 17 kcal mol<sup>-1</sup> (with B3LYP). For the other two snapshots, we obtain higher values of about 16 kcal mol<sup>-1</sup> (for BP86) and about 20 kcal mol<sup>-1</sup> (for B3LYP).

**3.6. Setup F (see Figure 6).** To improve the active-site interactions of Glu802 we protonated this residue in setup F: this allows for a hydrogen bond with the apical oxygen atom of the cofactor (as in setup D). For the conversion of this reactant complex, we essentially find a four-step mechanism to obtain the product which then rearranges to a more stable structure, in which the substrate at least partly dissociated from the cofactor, see Figure 6. In the initial step, the proton at N3 position is transferred via a water molecule to Glu1261 which leads to a deprotonated substrate in IM1'. This step is followed by a proton transfer cascade via TS1. While the proton of the labile OH

**Table 6.** QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup E, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN100-BP86		SN100-B3LYP		SN300-BP86		SN300-B3LYP		SN500-BP86		SN500-B3LYP	
reactant	0.0		0.0		0.0		0.0		0.0		0.0	
TS1'	1.2	(1.2)	2.8	(2.8)	0.5	(0.5)	4.0	(4.0)	2.0	(2.0)	2.8	(2.8)
IM1'	-5.1		-5.9		-5.8		-7.1		-8.1		-9.3	
TS1	2.4	(7.5)	2.5	(8.6)	2.6	(8.4)	4.4	(11.5)	4.0	(12.1)	5.2	(14.5)
IM1	-0.9		-0.4		-0.8		2.3		1.7		2.6	
TS2	6.3	(7.2)	10.1	(10.5)	6.0	(6.8)	12.5	(10.2)	6.7	(5.0)	10.1	(7.5)
IM2	4.8		8.0		5.6		9.8		3.6		7.3	
TS3	8.1	(3.3)	11.0	(3.0)	10.1	(4.5)	14.5	(4.7)	8.2	(4.6)	11.0	(3.7)
product	5.1		-5.8		2.4		-6.3		5.8		-5.0	
$\Delta E_{\max}$	13.2		16.9		15.9		21.9		16.3		20.3	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.



**Figure 6.** Reaction mechanism for setup F, SN400-B3LYP/MM. All bond lengths are given in Å.

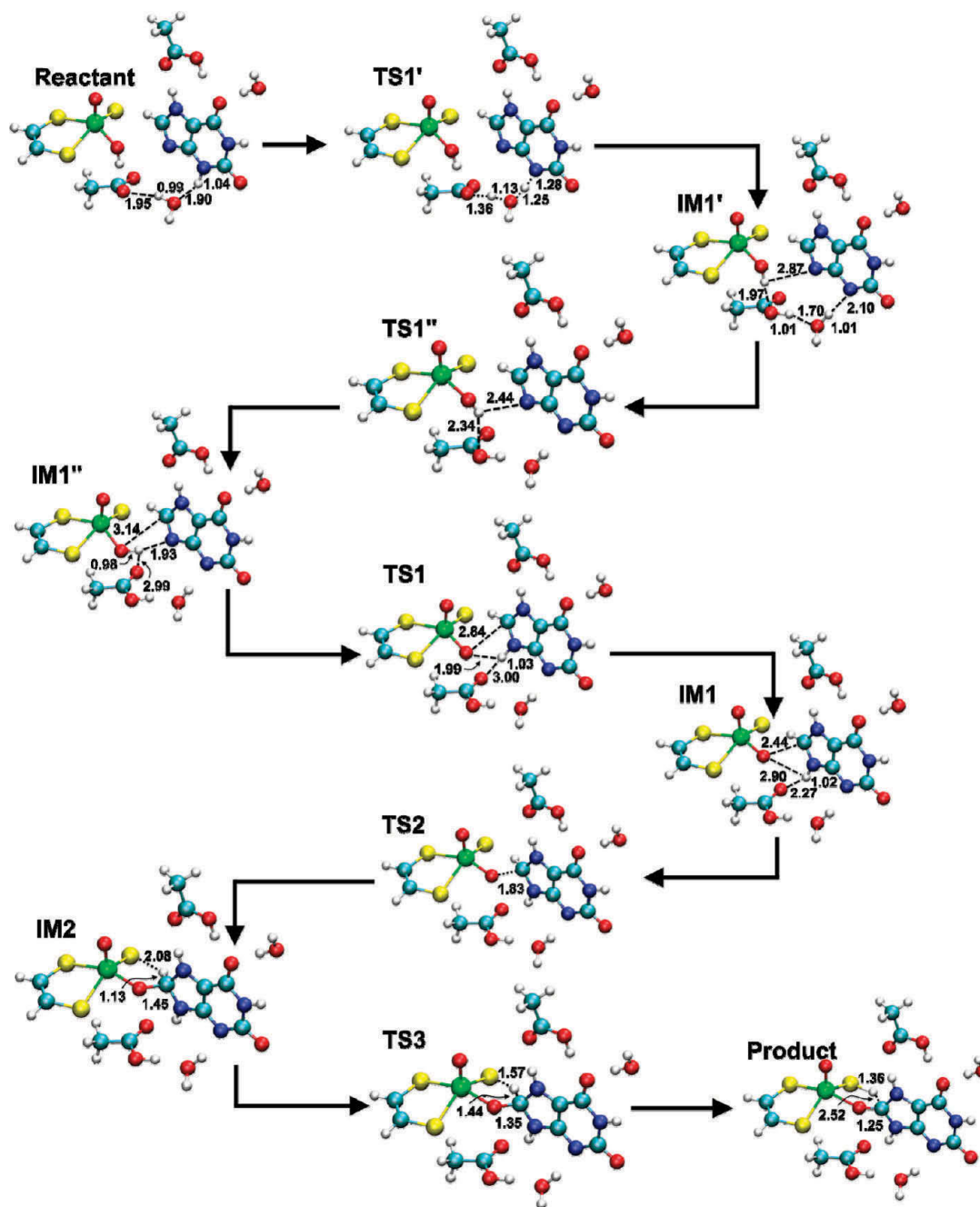
**Table 7.** QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup F, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN200-BP86	SN200-B3LYP	SN400-BP86	SN400-B3LYP
reactant	0.0	0.0	0.0	0.0
TS1'	1.1 (1.1)	2.7 (2.7)	2.0 (2.0)	4.2 (4.2)
IM1'	-7.5	-9.2	-8.6	-9.4
TS1	-5.0 (2.5)	-2.7 (6.5)	-3.3 (5.3)	-0.4 (9.0)
IM1	-17.0	-16.8	-12.9	-13.1
TS2	-2.0 (15.0)	1.3 (15.5)	2.5 (15.4)	6.0 (19.1)
IM2	-4.0	0.3	-1.6	2.0
TS3	3.5 (7.5)	3.7 (4.0)	4.6 (6.2)	7.2 (5.2)
product	1.2	-8.2	0.8	-9.1
TS_P'	6.6 (5.4)	-3.8 (4.5)	4.7 (3.8)	-5.3 (3.8)
product'	-1.2	-12.1	-4.7	-14.6
$\Delta E_{\max}$	23.6	20.5	17.6	20.3

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

group is transferred to N9 of the substrate, the apical oxygen of the cofactor becomes strongly acidic and is protonated by Glu802 whereas Glu802 is reprotonated by H7:URIC. In the thus obtained intermediate IM1, the cofactor still bears one oxo and one hydroxy group, but the position of them has interchanged. Subsequently, the cofactor forms a C–O bond with the substrate (via TS2) leading to the tetrahedral intermediate (IM2) which then reacts via a hydride transfer to the product. The initially formed product complex can undergo a facile rearrangement of the formed thiole group accompanied with a change in coordination of Glu802 toward the uric acid: The proton of Glu802 is no longer pointing toward the apical oxygen atom of the cofactor, but to O8:URIC, and the Mo–O8 bond distance increases to 2.58 Å.





**Figure 7.** Reaction mechanism for setup G, SN400-B3LYP/MM. All bond lengths are given in Å.

Comparing the calculated results for three investigated snapshots, the reaction barrier is determined by the energy difference between IM1 and the transition state TS3 for the hydride transfer (see Table 7). The strong stabilization of IM1 can again be attributed to the favorable proton transfer to the apical oxygen atom as in setup D. For the two snapshots, we obtain overall barriers of about 18–21 kcal mol<sup>−1</sup> (for BP86) and about 20 kcal mol<sup>−1</sup> (for B3LYP). In this setup, IM1 is a rather stable Michaelis complex that acts as a thermodynamic sink and heightens the barrier to be overcome.

**3.7. Setup G (see Figure 7).** For setup G we kept Glu802 protonated, but chose an orientation in which it forms a hydrogen bond with O6:URIC. For the conversion of the reactant to the product complex, we find five elementary steps. Initially the proton at N3 position is transferred to Glu1261 (via

TS1') to obtain IM1'. This is followed by a rotation of the hydroxy group of the cofactor, which establishes a new hydrogen bond to N9 of the substrate in IM1''. In the next step, the proton is transferred to the substrate (via TS1) which then rearranges its orientation such that a new hydrogen bond between the substrate and Glu1261 is formed in IM1. After these three steps, the cofactor has been deprotonated, and the substrate has changed its protonation state from tautomer A to tautomer C (see Figure 2). Subsequently, the cofactor forms a C–O bond with the substrate (via TS2) leading to the tetrahedral intermediate (IM2) which then reacts via a hydride transfer to the product (see Figure 7).

For setup G, we have investigated three snapshots. SN200 differs from SN100 and SN400 in that Glu802 does not coordinate directly to the substrate but via an additional water

**Table 8.** QM/MM Energies in kcal mol<sup>−1</sup> Calculated for Setup G Using Basis Set B1, Relative to the Energy of the Reactant for Different Snapshots<sup>a</sup>

	SN100- BP86	SN100-B3LYP	SN200-BP86	SN200-B3LYP	SN400-BP86	SN400-B3LYP
reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	0.9 (0.9)	4.5 (4.5)			2.9 (2.9)	8.5 (8.5)
IM1'	−5.9	−5.6			−8.0	−7.5
TS1''	−5.4 (0.5)	−5.3 (0.3)	4.2 (4.2)	5.1 (5.1)	−7.4 (0.6)	−7.2 (0.3)
IM1''	−10.5	−9.8	−4.6	−4.8	−9.0	−8.6
TS1	0.6 (11.1)	−0.1 (9.7)	3.1 (7.7)	3.3 (8.1)	1.0 (10.0)	2.5 (11.1)
IM1	−0.3	−0.2	2.6	3.2	−3.4	−1.8
TS2	0.6 (0.9)	3.4 (3.6)	7.8 (5.2)	12.5 (9.3)	−2.1 (1.3)	2.5 (4.3)
IM2	−4.7	−1.2	4.4	9.7	−7.8	−2.5
TS3	3.1 (7.8)	4.8 (6.0)	10.1 (5.7)	13.5 (3.8)	0.7 (8.5)	4.5 (7.0)
product	1.9	−7.2	4.9	−4.1	−0.8	−9.3
$\Delta E_{\max}$	13.6	14.6	14.7	18.3	10.0	13.1

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.**Table 9.** QM(B3LYP)/MM Energies in kcal mol<sup>−1</sup> for Setup G Calculated for Different Basis Sets, Relative to the Energy of the Reactant<sup>a</sup>

	B1	B2	B3/B2	B3	B3-ZORA	B3-DKH2
reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	8.5 (8.5)	7.7 (7.7)	8.0 (8.0)	8.0 (8.0)	8.0 (8.0)	8.0 (8.0)
IM1'	−7.5	−8.9	−8.7	−8.7	−8.6	−8.6
TS1''	−7.2 (0.3)	−8.2 (0.7)	−8.0 (0.7)	−8.0 (0.7)	−8.0 (0.6)	−8.0 (0.6)
IM1''	−8.6	−9.2	−9.5	−9.6	−8.2	−8.2
TS1	2.5 (11.1)	2.4 (11.2)	4.7 (14.2)	4.7 (14.3)	3.8 (12.0)	3.8 (12.0)
IM1	−1.8	−1.3	0.8	0.8	0.5	0.5
TS2	2.5 (4.3)	3.1 (4.4)	3.6 (2.8)	3.8 (3.0)	3.5 (3.0)	3.4 (2.9)
IM2	−2.5	−1.3	−3.1	−3.3	−2.3	−2.5
TS3	4.5 (7.0)	7.7 (9.0)	2.5 (5.6)	2.7 (6.0)	5.7 (8.0)	5.3 (7.8)
product	−9.3	−5.0	−14.7	−14.7	−8.5	−9.0
$\Delta E_{\max}$	13.1	16.9	14.2	14.3	14.3	13.9

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

molecule, which raises the energy barrier compared to the other snapshots. For all snapshots the reaction barrier is determined by the energy difference between IM1'' (i.e., the complex with deprotonated substrate and protonated cofactor) and the transition state TS3 for the hydride transfer (see Table 8, SN400-BP86 differs slightly). For the snapshots SN100 and SN400, we obtain overall barriers of about 10–14 kcal mol<sup>−1</sup> (for BP86) and about 13–15 kcal mol<sup>−1</sup> (for B3LYP). Barriers calculated with the alternative L2 link atom scheme for SN400 differ by 0.4–0.5 kcal mol<sup>−1</sup> (see Table S19 in SI).

For snapshot SN400 with the lowest overall barriers, we performed additional calculations with other basis sets (B2 and B3) and reoptimized all stationary points. For the following discussion, we will restrict ourselves to the basis set dependence of the B3LYP results; the BP86 results are given in the SI.

When going from the small B1 basis to the larger B2 basis (def2-TZVP) there are only minor changes in the computed relative energies (see Table 9), typically of the order of 1 kcal mol<sup>−1</sup>, except for the last two points (TS3, product) which are destabilized by 3–4 kcal mol<sup>−1</sup>. Basis B3 differs from B2 only by using an all-electron description of Mo rather than an effective core potential (ECP) with a valence basis set. This replacement causes negligible geometry changes; the B3 energies from full optimization (column B3) and from single-point calculations at optimized B2 geometries (column B3//B2) are virtually identical (maximum deviation of 0.2 kcal mol<sup>−1</sup>). Scalar relativistic effects associated with the inner shells of Mo are taken into account by the ECP treatment for B1 and B2, but not for the all-electron B3 basis where an explicit relativistic treatment is required for this purpose. The nonrelativistic and relativistic B3-based energies turn out to be quite close to each

other in general (columns B3 vs B3-ZORA and B3-DKH2, typical deviations of about 1 kcal mol<sup>−1</sup>), again except for the last two points (TS3, product) where inclusion of relativistic effects raises the relative energies by 3–6 kcal mol<sup>−1</sup> and brings them closer to the B1 and B2 values (as expected).

Overall the QM(B3LYP)/MM energies for setup G thus appear to be quite stable with regard to basis set extension. The lowest points on the pathway are the initial intermediates IM1' and IM1'' which lie within 1 kcal mol<sup>−1</sup> and are separated by a very small barrier. On the route from these intermediates to the product, the highest point to be overcome is normally the final transition state TS3 (see columns B1, B2, B3-ZORA, B3-DKH2), except in the case of the nonrelativistic B3 treatment (considered less reliable for an all-electron basis). The computed energy differences  $\Delta E_{\max}$  between these lowest and highest points lie between 13 and 17 kcal mol<sup>−1</sup>, and the formally “best” values from the relativistic B3-based calculations are around 14 kcal mol<sup>−1</sup> (see Table 9).

So far, all reported QM/MM results refer to our standard simulation system with a radius of ca. 35 Å around the central C8 atom (see Figure 4). To check the effects of the more remote parts of the initially prepared complete system (see Figure 1), we “glued” the optimized stationary points for SN400 of setup G into the original enzymatic “frame” and reoptimized them. The results differ only slightly from those of the truncated simulation system. The relative energies of the different stationary points typically differ by less than 1 kcal mol<sup>−1</sup> with a maximum deviation of 2.6 kcal mol<sup>−1</sup> (for the B1 basis). The calculated barriers deviate at most by 1.5 kcal mol<sup>−1</sup> (for detailed results see SI).

**Table 10.** Results for the Extended 66-atom QM Region: B3LYP Relative Energies in kcal mol<sup>-1</sup> for Six Snapshots and RMSD Values Averaged over the Six Snapshots (see text)

	reactant rmsd = 0.150 Å				IM1 rmsd = 0.132 Å				IM2 rmsd = 0.101 Å				product rmsd = 0.110 Å			
	$\epsilon = 1$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$	$\epsilon = 1$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$	$\epsilon = 1$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$	$\epsilon = 1$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$
A500	0.0	0.0	0.0	0.0	20.3	10.0	8.3	7.2	0.0	3.3	4.8	5.7	1.4	0.0	0.0	0.0
C200	31.5	18.7	15.2	13.1	40.0	16.7	11.6	8.7	28.0	20.9	19.5	18.5	25.7	13.7	10.9	9.3
D200	7.9	7.1	6.9	6.7	0.0	2.2	4.2	5.3	7.7	11.5	13.4	14.5	0.0	4.2	6.0	7.1
F400	24.0	13.6	10.7	9.1	4.6	0.0	0.0	0.0	2.4	1.6	2.1	2.4	10.7	8.8	8.9	8.9
G100	34.8	16.9	12.3	9.7	33.3	12.8	8.7	6.5	2.9	0.1	0.1	0.1	19.8	9.0	6.6	5.3
G400	34.0	17.1	12.7	10.1	31.2	11.3	7.2	4.9	2.7	0.0	0.0	0.0	17.4	6.9	4.6	3.2

**3.8. Consistency Checks.** One concern in large-scale QM/MM geometry optimizations is whether the calculated stationary points are connected by contiguous reaction paths and how one can exclude artifacts due to distant conformational changes, e.g., flips of amino acid side chains or reorientations of water molecules far away from the active site. We have carefully checked for such artifacts both by visual inspection and by calculating the root-mean-square deviations (rmsd) between the optimized structures of the reactant complex (see Table S30, SI for a detailed definition) and the rate-limiting transition state TS3. The SI documents the overall rmsd values for all setups and snapshots considered (Table S30, SI) and presents a residue-specific analysis for snapshot 500 of setup A (Figure S11, SI) and snapshot 400 of setup G (Figure S12, SI) along with a detailed discussion. These data confirm that the structural changes in the protein environment are rather small and that there are no structural inconsistencies in distant parts of the optimized active region.

**3.9. Setup Comparison.** The QM/MM results reported so far refer to individually prepared setups A–G in which each QM region (Figure 3) is embedded in an individually equilibrated and optimized MM environment. Therefore, these setups generally differ in several aspects, e.g., the number of solvent molecules included, the number of atoms considered (total, QM, MM), the protonation states of certain residues, etc. (see section 2 and SI). As a consequence, the computed total QM/MM energies as well as the corresponding QM and MM contributions are not directly comparable for different setups.

To put the QM/MM energies of different setups on a common scale, one could attempt to study their interconversion at the QM/MM level. This may be feasible for closely related setups (e.g., F and G), but is difficult in general. For example, the orientation of the substrate is “upside” in setups A–D and “upside down” in setups E–G (see Scheme 2) so that a direct interconversion inside the binding pocket would involve a rotation of the substrate which should be associated with a prohibitively high barrier. Such an interconversion would probably occur via a dissociation-association mechanism which can hardly be treated by QM/MM with the required accuracy because of the need for extensive sampling. Furthermore, there are differences in the charges of the chosen QM regions and in the assignment of MM protonation states for different setups (see Figure 3 and SI) which also prevent straightforward QM/MM studies of their interconversion.

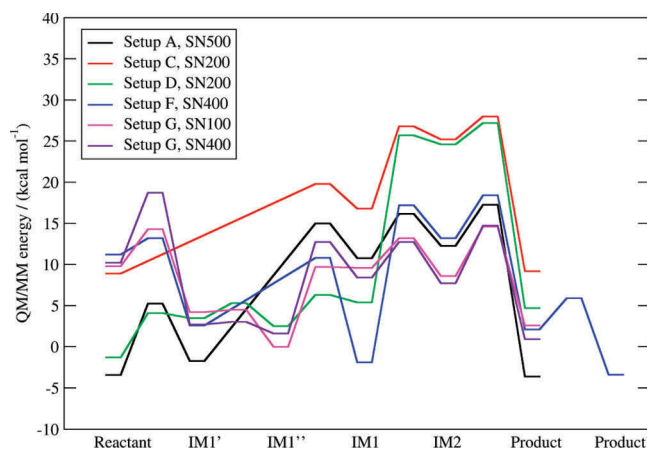
In spite of these caveats, we decided to perform some numerical experiments by manually putting the QM regions of different setups into one given “foreign” MM environment and reoptimizing these systems. We employed the intermediate IM2 for this exercise since its QM region is relatively rigid, with the substrate covalently bound to the cofactor. We chose the most favored setup G as the “foreign” MM environment and replaced the QM region of snapshot 400 by that of setup A, C,

D, and F (labeled G2A, G2C, G2D, and G2F, respectively). As a crosscheck, we also considered the inverse replacement where the QM regions of setups A, C, D, and F were replaced by that of setup G (labeled A2G, C2G, D2G, and F2G, respectively). The results of the corresponding QM/MM reoptimizations are documented in the SI (Table S31, SI). As expected, they are not conclusive: some of these reoptimizations did not converge to an IM2 structure (G2C, G2D) while the resulting energies are unreasonably high for others (G2A, G2F). Thus, the QM regions of setups A, C, D, and F apparently do not “fit well” into the binding pocket of setup G even after local reoptimization, indicating that a more extensive rearrangement of the MM environment would be needed to accommodate them well. This is consistent with the well-known concept of induced fit<sup>118,119</sup> according to which the enzyme adapts its geometry as the substrate enters. Induced-fit effects arising from different substrate orientations and different tautomeric situations are taken into account in our individual setups (see section 2), but are evidently not captured by the local reoptimizations described above.

In another attempt to arrive at a common energy scale for the different setups, we turned to cluster calculations.<sup>120–123</sup> Using an extended QM region with 66 atoms (Figure 5) we performed QM(B3LYP/B1)/MM geometry optimizations for six snapshots. These QM/MM reoptimizations caused only minor changes in the QM/MM geometries previously obtained with standard QM regions, as can be seen from the corresponding root-mean-square deviations (rmsd) in Table 10. Subsequent QM single-point calculations were done both in vacuo and in continuum solvent using COSMO with different dielectric constants. The COSMO calculations include the effect of the MM environment of the different snapshots in an average manner, and it has been argued<sup>120–123</sup> that such a continuum representation of the protein environment is acceptable if the computed relative energies do not depend much on the assumed dielectric constant. Following this strategy we used dielectric constants of 4, 8, and 16 in the COSMO calculations:  $\epsilon = 4$  is a commonly accepted choice for proteins, and values up to 10 and above are often employed<sup>124–126</sup> to reflect the influence of

- (118) Koshland, D. E., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.
- (119) Koshland, D. E., Jr. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 2375–2378.
- (120) Sevastik, R.; Himo, F. *Bioorg. Chem.* **2007**, *35*, 444–457.
- (121) Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput.* **2008**, *4*, 1129–1137.
- (122) Chen, S.-L.; Fang, W.-H.; Himo, F. *Theor. Chem. Acc.* **2008**, *120*, 515–522.
- (123) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651.
- (124) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267–298.
- (125) Simmons, T.; Brooks, C. L., III. *J. Am. Chem. Soc.* **1996**, *118*, 8452–8458.





**Figure 8.** QM/MM reaction profiles for different setups on a common energy scale derived from COSMO ( $\epsilon = 8.0$ ) calculations on the extended QM region (see text and SI).

nearby bulk water (with  $\epsilon = 80^{127}$ ). Since parts of our QM region are close to the bulk solvent, values of  $\epsilon > 4$  seem justified in our case.

Table 10 presents the results of such single-point B3LYP/B1 calculations in the gas phase ( $\epsilon = 1$ ) and in the COSMO-MM environment ( $\epsilon = 4, 8, 16$ ). Relative energies are given for the reactant state, the intermediates IM1 and IM2 as well as for the product. The COSMO results depend on the dielectric constant only slightly for IM2, somewhat more for the reactant and the product, and rather strongly for IM1. The relative energies of the different setups should thus be reflected best<sup>120–123</sup> by the data for IM2. This choice is further supported by the results of Torres et al.<sup>128</sup> who reproduced the observed regioselectivities for different drugs or drug-like molecules quite well using the computed relative energies of the tetrahedral intermediates. The COSMO results for IM2 should thus provide a reasonable starting point to define a common energy scale for all setups. The resulting set of energy profiles (see Figure S15, SI) turns out to be similar to those obtained with the reactant or IM1 or the product as reference species (see Figures S13–S16, SI), and therefore, we show a representative average set in Figure 8 (see page S43 of SI for the conventions adopted to define the common energy scale).

It should be emphasized, of course, that the chosen procedure is very approximate and can only lead to tentative conclusions. It seems that setups A and D with “upside” orientation of the substrate are favored at the reactant stage and would thus be expected to be present in corresponding crystal structures. Other setups are higher in energy initially, but not prohibitively so (keeping in mind our very approximate energy adjustment procedure). Kinetically, setups A and D are no longer favored (see Figure 8). Most energy profiles have their highest point in the final hydride transfer step  $\text{IM2} \rightarrow \text{product}$  (via TS3), and here setup G with “upside down” orientation of the substrate seems preferred, which makes the situation similar to the so-called major-minor problem in asymmetric catalysis.<sup>129</sup> Setup G emerges as the best one also when comparing the largest

single-step barriers or the overall energy span between the lowest and highest point on the profiles for the different pathways (see Tables 2–8).

To rationalize the trends in the relative energies (Table 10) we have analyzed the charge distributions obtained in the single-point B3LYP calculations for the extended QM region. We focus on setup A (SN500) and setup G (SN100 and SN400) as they are representative of the “upside” and “upside down” substrate orientations (with the lowest effective barriers). The computed ESP (electrostatic potential based) charges of a given species (reactant, IM1, IM2, product) are generally quite similar for these three snapshots (see Figure S17, SI). Furthermore, the overall ESP charges of individual residues remain approximately constant in these four species for Arg880 (close to +1) as well as Gln767, Glu802, and Wat224 (close to 0), while there are obvious changes for substrate, cofactor, and Glu1261 that are caused by the proton transfer during the reaction (see Figure S17, SI). The ESP data thus indicate that the basic charge distribution in the QM region is not affected much by substrate orientation.

In an attempt to explain the different relative stability of the reactant and IM2 in setup A (“upside”) and setup G (“upside down”) we have evaluated classical electrostatic interactions between key partners (see Scheme 6) using the computed ESP charges and the relevant distances in the optimized structures. Details are given in the SI (Table S32 and associated discussion). It turns out that the electrostatic interactions involving Arg880 are decisive for the relative stabilities. In the reactant, they favor setup A over setup G by about 13–19 kcal mol<sup>−1</sup> (O6 in setup A is more negatively charged and closer to Arg880 compared with O2 in setup G), while in IM2 they favor setup G over setup A by about 10–15 kcal mol<sup>−1</sup> (mostly because of the interactions with the negatively charged N3 atom which is close to Arg880 only in the “upside down” orientation of setup G, see Scheme 6). Concerning the changes during the reaction, the key electrostatic interactions with Arg880 stabilize IM2 relative to the reactant in both setups, but significantly more so in the case of the favored setup G. Arg880 is thus important both for the overall activity and the stereochemical course of the enzymatic reaction. Further analysis shows that the electrostatic interactions with Glu802 also play a minor role: they become weaker when proceeding from the reactant to IM2, but less so in the favored setup G (see SI for more details).

**3.10. QM Model Study on Setup G.** QM calculations on small active-site models are useful for exploring the intrinsic reactivity of the substrate, and the comparison of such QM results with the full QM/MM results allows for an assessment of the influence of the protein environment. Using the most favorable setup G we have therefore carried out B3LYP/B1 and B3LYP/B2 calculations on a minimal gas-phase model consisting of the cofactor and the substrate only. To be able to compare with the QM/MM results for pathway G (see Figure 7) we started from the analogue of IM1′, i.e., the negatively charged cofactor interacting with the corresponding xanthine anion (N1,N7). The gas-phase reaction proceeds in analogy to the enzymatic one: an initial proton transfer from the cofactor to the N9 atom of xanthine ( $\text{IM1}'' \rightarrow \text{IM1}$ ) is followed by C–O bond formation ( $\text{IM1} \rightarrow \text{IM2}$ ) and hydride transfer ( $\text{IM2} \rightarrow \text{product}$ ), see Figure 9. The computed QM energies are given in Table 11 together with the corresponding QM/MM data adapted from Table 9.

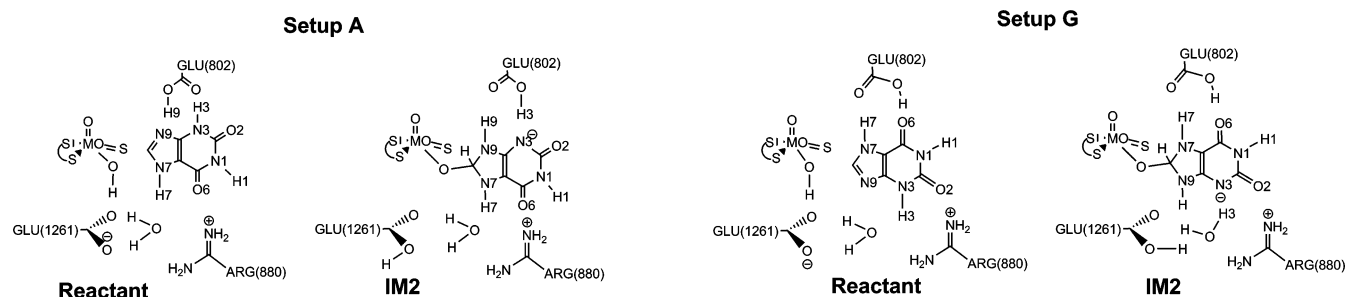
It is obvious that the overall mechanism for setup G is the same in the gas-phase model system and in the enzyme. There is a three-step conversion in each case starting from IM1′, and

(126) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421–438.

(127) Archer, D. G.; Wang, P. J. *Phys. Chem. Ref. Data* **1990**, *19*, 371–411.

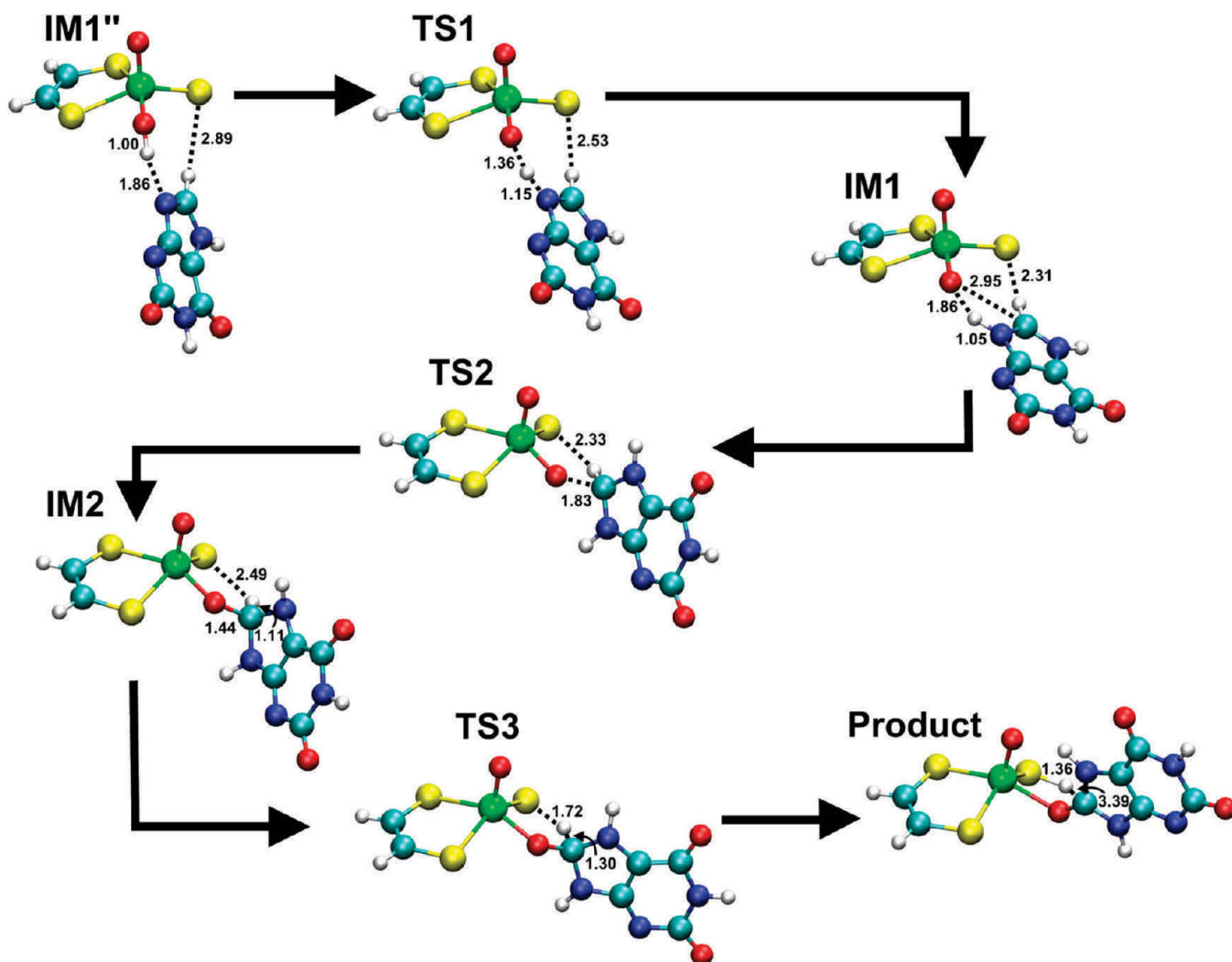
(128) Torres, R. A.; Korzekwa, K. R.; McMasters, D. R.; Fandozzi, C. M.; Jones, J. P. J. *Med. Chem.* **2007**, *50*, 4642–4647.

(129) Halpern, J. *Science* **1982**, *217*, 401–407.

**Scheme 6.** Structures of Reactant and IM2 from Setups A and G

the highest point on the energy profile is associated with the final hydride transfer via TS3. Comparing the optimized structures in the gas phase and the enzyme (see Figures 7 and 9) the relative orientation of the cofactor and the substrate is very similar in the late stages of the reaction, i.e., for TS2, IM2, and TS3; the relative energies of these three species are around 4–5 kcal mol<sup>-1</sup> lower in the enzyme compared with the gas phase (see Table 11) which can be attributed to the effective stabilization of the negative charge at the substrate N3 atom in these species by the neighboring Arg880 residue (which is missing in the gas-phase model). Since TS3 is the highest point on the reaction profile, this specific interaction will accelerate the overall enzymatic reaction for setup G. By contrast, in the

early stage of the reaction, i.e., for IM1'', TS1, and IM2, the relative orientation of the cofactor and the substrate is somewhat different in the gas phase compared with the enzyme: since there are no steric constraints from the protein environment, the gas-phase optimization yields a more “bent” arrangement with a somewhat different hydrogen-bond pattern, and this greater structural freedom may be the reason that the initial proton-transfer barrier is actually lower in the gas phase than in the enzyme. This is less relevant mechanistically since TS1 lies below TS3 in any case (see Table 11). In summary, the comparison between the QM and QM/MM results for setup G emphasizes the catalytic role of the Arg880 residue for the rate-limiting hydride transfer step.

**Figure 9.** Detailed mechanism for the gas-phase reaction analogous to setup G, SN400-B3LYP. All bond lengths are given in Å.

**Table 11.** Gas-Phase QM Energies Compared with QM/MM Energies (QM = B3LYP) from Setup G SN400<sup>a</sup>

	gas phase		complete enzyme	
	B1	B2	B1	B2
IM1''	0.0	0.0	0.0	0.0
TS1	5.5 (5.5)	6.3 (6.3)	11.1 (11.1)	11.2 (11.2)
IM1	5.4	6.0	6.8	7.9
TS2	15.4 (10.0)	17.2 (11.2)	11.1 (4.3)	12.3 (4.4)
IM2	10.8	13.3	6.1	7.9
TS3	16.8 (6.0)	21.0 (7.7)	13.1 (7.0)	16.9 (9.0)
product	-13.3	-8.0	-0.7	4.2
$\Delta E_{\text{max}}$	16.8	21.0	13.1	16.9

<sup>a</sup> All energies are given in kcal mol<sup>-1</sup>, relative to the energy of IM1''. Activation barriers relative to the preceding minima are given in parentheses.

A recently published QM model study<sup>65</sup> at the DFT(mPW1-PW91) level considered the reaction between the deprotonated cofactor and xanthine in three different protonation states. In each case, only a single transition state was reported which corresponds to the hydride transfer step (TS3). Its energy was computed relative to that of the separated reactants (protonated cofactor + xanthine (N1,N3,N7) + formate anion), with HCOO<sup>-</sup>/HCOOH added to balance the stoichiometry and to define a common energy scale.<sup>65</sup> The calculated TS3 energies were 40.5 kcal mol<sup>-1</sup> for xanthine (N1,N3,N7), 33.4 kcal mol<sup>-1</sup> for xanthine (N1,N7,N9), and 6.4 kcal mol<sup>-1</sup> for protonated xanthine (N1,N3,N7,N9). We note that the latter low barrier is not relevant for the situation in the enzyme since protonated xanthine with its pK<sub>a</sub> value of about 0.8<sup>79</sup> will easily transfer a proton to another active-site residue (e.g., Glu1261) and will thus not be present as such in the binding pocket. The remaining two TS3 energies reported for the neutral xanthine tautomers are rather high for an enzymatic reaction which is not surprising since stabilizing active-site residues such as Arg880 (see above) are not included in these QM model calculations.<sup>65</sup> A direct comparison with our results is possible in the case of the (N1,N7,N9) tautomer which appears in our setup G (see Figures 7 and 9). Both the previous<sup>65</sup> and our current QM model study yield rather similar TS3 geometries (e.g., with regard to the optimized bond lengths), and we confirm the prohibitively high barrier in the gas phase relative to the separated reactants<sup>65</sup> (currently 43.4 kcal mol<sup>-1</sup> (48.0 kcal mol<sup>-1</sup>) using B1 (B2) compared to 33.4 kcal mol<sup>-1</sup> reported by Bayse<sup>65</sup>). On the other hand, it is clear that the published QM model study<sup>65</sup> has some intrinsic limitations by simply calculating the energy barrier as energy difference of TS3 and the separated reactants, ignoring the essential influence of the reactant complex on the energy profile (see also Amano et al.<sup>63</sup>). Moreover, pure QM studies by design do not account for the steric constraints that orient the substrate in the binding pocket and for the decisive role of active-site residues such as Arg880 and Glu1261 (all of which is captured in the current QM/MM work).

## 4. Discussion

**4.1. Overview.** Our QM/MM calculations have explored several mechanistic scenarios for the reductive half-reaction in xanthine oxidase. This complexity arises from a number of factors. First, the substrate xanthine can be present in several forms: according to previous B3LYP/6-31++G\*\* calculations with a Poisson–Boltzmann continuum solvation approach,<sup>79</sup> the neutral tautomers (N1,N3,N7) and (N1,N3,N9) are almost isoenergetic in aqueous solution (within 0.6 kcal mol<sup>-1</sup>), and

the same applies to the anionic tautomers (N1,N9), (N1,N7), and (N1,N3) (within 1.2 kcal mol<sup>-1</sup>). Moreover, with a pK<sub>a</sub> value of around 7 (computed: 6.9, experimental: 7.44),<sup>79</sup> both the neutral and anionic forms of xanthine are accessible under physiological conditions. It is thus not surprising that different tautomeric and ionized xanthine species are encountered during the QM/MM pathway calculations, for example, (N1,N3,N7) in setups A and D–G, (N1,N3,N9) in setups B–D, (N1,N7,N9) in setups A, F and G, (N1,N9) in setups B and F, (N1,N7) in setups E–G, and (N1,N3) in setup D. Second, the cofactor is always modeled as [(S–CH=CH–S)Mo(=S)(=O)(–OH)]<sup>-</sup> anion in the initial reactant complex which is normally deprotonated to yield the active [(S–CH=CH–S)Mo(=S)(=O)(–O)]<sup>2-</sup> dianion in the IM1 intermediate (setups A–C, E, G); however, two of the pathways (D, F) also involve at the IM1 stage a different cofactor species with a protonated apical oxygen atom. Third, the orientation of xanthine in the binding pocket (see Scheme 2) can be either “upside” (setups A–D) or “upside down” (setups E–G).

**4.2. Role of the Active-Site Residues.** Given the multitude of energetically competitive arrangements of different forms of the substrate and cofactor, it is clear that the protein environment will play a crucial role in determining the mechanistic preferences. Particularly important are the residues Glu1261 and Glu802 which mediate the various proton transfers in setups A–G, and Arg880 which is essential for stabilizing any developing negative charge in the neighboring region of xanthine (e.g., around N3 in setups E–G). In the following, we first address the influence of the active-site residues on the basis of our QM/MM results and published experimental data.

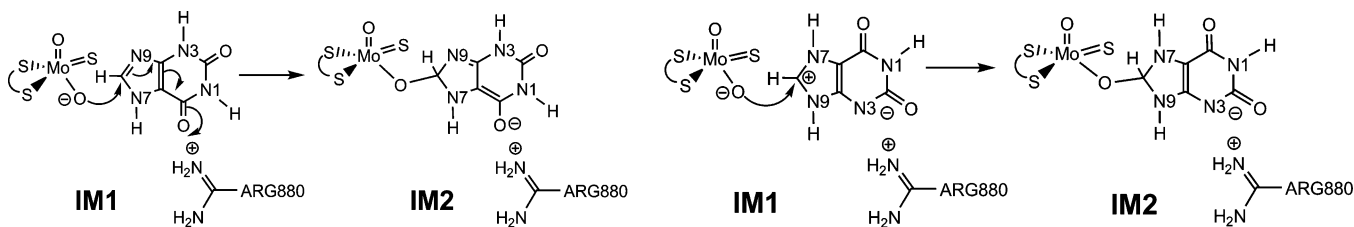
A common experimental approach to judge the effect of specific amino acids is to compare the rate constants of mutants and the wild-type enzyme and to derive the associated change of the activation energy (e.g., via the Eyring equation). This value is normally referred to as transition-state stabilization of the mutated residue. Of course, even a single mutation may affect different steps in the catalytic cycle, and the experimentally observed transition-state stabilization of a given residue thus provides the change in the “effective” barrier which has to be interpreted with some care. For example, it may happen that a mutation influences the balance between the rates for the chemical reaction and for product release (*k*<sub>2</sub>' vs *k*<sub>2</sub>''), implying that a small change in the overall rate is not necessarily associated with a small effect of the mutation on an individual rate constant.

The change in the effective barrier for Glu802 is reported to be small, i.e., 1.6 kcal mol<sup>-1</sup> for a Glu802 → Val802 mutant<sup>39</sup> and 1.4 kcal mol<sup>-1</sup> for a Glu802 → Ala802 mutant.<sup>40</sup> It seems likely in these mutants that the substrate may be coordinated by water from the bulk (rather than Glu802) which can also stabilize the accumulating negative charge at the O6 position of the substrate. The observed values for Glu802 (see above) would then reflect effects beyond pure electrostatic stabilization of the substrate (e.g., steric effects due to Glu802).

A much stronger influence is exerted by Arg880 whose transition-state stabilization is reported to be 4.5 kcal mol<sup>-1</sup> for a Arg880 → Met880 mutant<sup>42</sup> (and calculated to be 5.6 kcal mol<sup>-1</sup> from the published rate constants using the Eyring equation). This is largely, but not entirely an electrostatic effect, since there is still a small amount of transition-state stabilization of 1.6 kcal mol<sup>-1</sup> upon substitution of Arg880 by Lys880<sup>42</sup> which should also be protonated and thus positively charged (as Arg880). We have already shown that Arg880 is important



**Scheme 7.** Lewis Structures for the “Upside” Orientation of the Substrate<sup>32,42</sup> (left) and the “Upside Down” Orientation in the Actual Protonation State of Setup G (right)



for the substrate binding in the reactant state and essential for the stabilization of IM2 and the preceding and following transition states (see section 3.9).

A complete loss of activity is found upon replacement of Glu1261.<sup>40</sup> It is obvious from the calculations (and in agreement within the literature) that Glu1261 acts as a proton acceptor to initiate the reaction. Additionally, the flexibility of Glu1261 and its position relative to the substrate is of importance, as a shortening of its side chain by one CH<sub>2</sub> group in the Glu1261 → Asp880 mutant already leads to complete loss of activity. In the favored setup G, Glu1261 is involved in the tautomerization of xanthine from tautomer A to tautomer C which is achieved by deprotonating the substrate which then deprotonates the cofactor. This mechanism has not been proposed in the literature so far. Instead, Nishino et al.<sup>57</sup> suggested a proton transfer from Glu1261 to the substrate to compensate the evolving negative charge at the N9 position. However, the resulting all-N protonated tetrahedral intermediate is computed to be 7.0 kcal mol<sup>-1</sup> higher in energy than the reported IM2 structure and should thus not be relevant for the reaction mechanism.

**4.3. Substrate Orientation.** According to the QM/MM calculations, the substrate is oriented “upside down” in the kinetically favored pathway (setup G) even though the “upside” orientation of the substrate is thermodynamically favored at the reactant stage (setup A). The published crystal structure<sup>55</sup> of the desulfo-enzyme, with a resolution of 2.6 Å, indeed confirms the theoretically favored “upside” orientation in the reactant complex. An obvious question is whether this is the only binding mode or whether there are different binding modes that may be realized in the enzyme. An earlier report<sup>57</sup> pointed out that even at a resolution of 1.9 Å, the different binding modes of the substrate cannot be clearly distinguished in the X-ray structure. A mixture of various binding modes was assumed to rationalize<sup>57</sup> the observed inhibitory effect of xanthine at high substrate concentrations.<sup>130</sup>

On the theoretical side, the “upside” orientation was originally proposed on the basis of docking studies in a structurally similar AOR enzyme.<sup>30</sup> In addition, the “upside” orientation is often claimed to be favored kinetically because one can draw Lewis structures in which the negative charge accumulating on O6 in the transition state is stabilized.<sup>32,42</sup> Our present calculations of electrostatic interaction energies (see section 3.9 and Table S32, SI) show that such stabilization is indeed found for the “upside” orientation, but is even more pronounced for the “upside down” orientation. This is illustrated by the relevant Lewis structures for setup G in Scheme 7 which indicate a strong electrostatic stabilization of the xanthine tautomer C (IM1) and of the subsequent tetrahedral intermediate (IM2) by Arg880.

There is further experimental evidence that supports the “upside down” orientation. In the metabolic pathway of purine,

the C2 position is hydroxylated prior to the C8 position, presumably because the oxidation of the C2 position introduces the “upside down” recognition pattern that is required in the following step of the metabolic pathway. A mutual relationship between carbon C2 and C8 of the purine skeleton has been known for a long time;<sup>33,34</sup> both 2-oxopurine and 8-oxopurine are oxidized by xanthine oxidase to 2,8-dioxopurine whereas 6-oxopurine (= hypoxanthine) is not oxidized to 6,8-dioxopurine, but to 2,6-dioxopurine (= xanthine).

The observed reactivity toward 1-methyl-6-oxopurine and 1-methyl-2,6-dioxopurine strongly supports the “upside down” orientation: whereas the first substrate shows no reactivity,<sup>34</sup> the latter one does.<sup>34,38</sup> If an “upside” orientation is assumed, both substrates interact with Arg880 in exactly the same manner via O6 (structures A and B in Scheme 8) and should therefore have similar reactivity (contrary to experiment). On the other hand, in the “upside down” orientation, 1-methyl-6-oxopurine lacks the strong interaction of O2 with Arg880 (structure C in Scheme 8) which stabilizes 1-methyl-2,6-dioxopurine (structure D in Scheme 8). For the latter, Arg880 will be able to stabilize the reactant state as well as the relevant intermediate and transition state (IM2, TS3) much better than for 1-methyl-6-oxopurine. The observed reactivity of the two substrates is thus compatible only with an “upside-down” orientation.

**4.4. Product State.** Finally we should point out that our QM/MM calculations model the reactive half-reaction of xanthine oxidase up to a product complex that contains deprotonated uric acid, mostly (setups A–B and E–G) in the form of the (N1,N7,N9) anion which is commonly considered to be the most favorable tautomer.<sup>131–133</sup> Given experimental pK<sub>a</sub> values in the range 5.4–5.8 and a measured ionization enthalpy of –5.2 kcal mol<sup>-1</sup>,<sup>134</sup> protonation of the product complex to neutral uric acid should be feasible, considering the rather large energy difference between the rate-limiting transition state for hydride transfer (TS3) and the product complex (see Tables 8 and 9 for setup G). We have not studied this protonation step.

## 5. Conclusion

We have explored seven different system setups to examine the detailed mechanism for the conversion of xanthine to uric acid catalyzed by xanthine oxidase. For each setup, we have determined the reaction pathways for at least two snapshots using both the BP86 and B3LYP functionals. The favored mechanism (setup G) has a barrier of 13–15 kcal mol<sup>-1</sup> at the B3LYP/MM level, consistent with the available experimental data. For this favored pathway, the reactive xanthine species is

(130) Massey, V.; Brumby, P. E.; Komai, H. *J. Biol. Chem.* **1969**, *244*, 1682–1691.

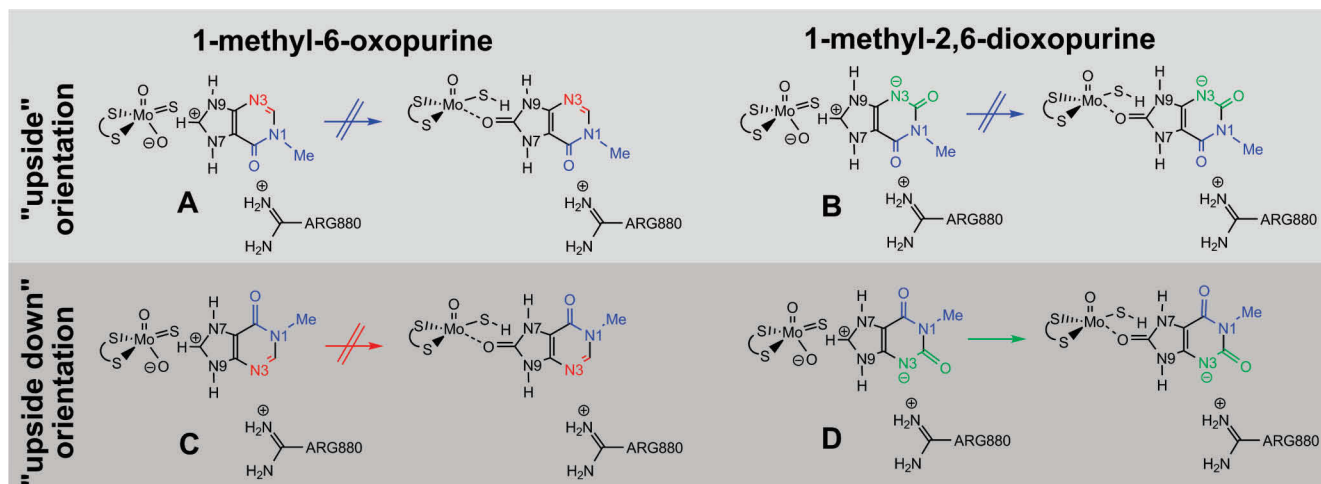
(131) Jimenez, V.; Alderete, J. B. *J. Mol. Struct. (THEOCHEM)* **2005**, *755*, 209–214.

(132) Shukla, M. K.; Mishra, P. C. *J. Mol. Struct.* **1996**, *377*, 247–259.

(133) Allen, R. N.; Shukla, M. K.; Leszczynski, J. *Int. J. Quantum Chem.* **2004**, *100*, 801–809.

(134) Finlayson, B.; Smith, A. *J. Chem. Eng. Data* **1974**, *19*, 94–97.

**Scheme 8.** Reactivity Control by the Substrate–Arg880 Interaction Rationalizes the Different Reactivity of 1-Methyl-6-oxopurine and 1-Methyl-2,6-dioxopurine; the Former Is Unreactive (A, C), while the Latter Is Reactive Only in the “Upside Down” Orientation (D), but Not in the Opposite One (B)



oriented “upside down” in the binding pocket, contrary to most of the previously published suggestions, but analogous to the experimentally derived orientation of 2-oxo-methylpurine. However, in the reactant complex, the alternative “upside” conformation of xanthine seems to be preferred (setup A) which has been observed in X-ray structures with different substrates. This situation is reminiscent of the minor–major paradigm in organometallic catalysis.

Different tautomeric and ionized xanthine species are encountered on the various pathways investigated. The stability order of these species depends on the enzyme environment, and the reactive species in the favored mechanism (setup G) does not correspond to the most stable gas-phase tautomer. In all mechanistic scenarios considered, there are proton transfers that involve active-site residues, and the transition state for the final hydride transfer step is generally the highest point on the computed energy profiles. Product release has not been addressed in the present computational study.

The protein environment is essential for the reductive half-reaction of xanthine oxidase. Focusing on the favored pathway (setup G) the main function of the Glu1261 residue is to (indirectly) deprotonate the cofactor and to mediate the conversion of the substrate from its tautomer A form into the reactive tautomer C. This is achieved by deprotonating the substrate at N3 followed by a second proton transfer that converts cofactor and substrate into their activated forms. The Arg880 residue facilitates substrate binding through stabilizing electrostatic interactions, but its main role during the reaction (setup G) is to stabilize the substrate in the IM2 intermediate and the neighboring transition states (TS2, TS3), especially by the

interaction with the negatively charged N3 atom. Finally, the Glu802 residue seems to have a relatively minor effect on the catalytic activity in xanthine oxidase with xanthine as substrate, but may well contribute to the proper alignment of xanthine in the binding pocket.

The current mechanism is quite intricate in that an initial double proton transfer, mediated by Glu1261, is used to activate the substrate. In the later stages of the reaction, it shares a number of characteristic features with the most favorable pathway in aldehyde oxidoreductase.<sup>46</sup> Our QM/MM results are broadly compatible with the general mechanistic notions about the reductive half-reaction in these enzymes. It is obvious, however, that the QM/MM calculations go significantly beyond these general notions by offering insight into mechanistic details that are hard to unravel by other means, e.g., in the present case with regard to substrate orientation and the role of individual active-site residues in xanthine oxidase.

**Acknowledgment.** This work was supported by the Max-Planck-Gesellschaft. S.M. thanks the Fonds der Chemischen Industrie for a Kekulé scholarship.

**Supporting Information Available:** System setup; detailed QM/MM results including QM, MM, and QM/MM energies of all stationary points and plots of their optimized structures; consistency checks; further setup comparisons; CHARMM parameters; complete refs 75 and 81. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA9045394



# Supporting Information for

## A Combined QM/MM Study on the Reductive Half-reaction of Xanthine Oxidase: Substrate Orientation and Mechanism

*Sebastian Metz and Walter Thiel\**

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

[thiel@mpi-muelheim.mpg.de](mailto:thiel@mpi-muelheim.mpg.de)

1. System preparation using CHARMM C29b2	S2
2. QM/MM results (Tables S1-S29, Figures S1-S10)	S6
3. Consistency checks (Table S30, Figures S11-S12)	S40
4. Setup comparisons (Tables S31-S32, Figures S13-S18)	S43
5. CHARMM parameters	S50
6. Complete references 75 and 81	S61

## 1. System preparation using CHARMM C29b2

### *A. Preparation of the PDB structure.*

The PDB structure of the bovin xanthine oxidase dimer (PDB 1FO4) lacks several residues. Using SWISS-MODEL<sup>1-4</sup> we obtained a complete structure after an automatically performed energy minimization. To add the missing residues to the original crystal structure, we overlayed the complete structure with the original crystal structure and glued residues 165 to 192 and 529 to 538 for chains A and B into the crystal structure, i.e. more residues than missing in order to ensure a smooth change over at the interface of the modeled residues with the original crystal structure. For the missing residues 1 and 2 of each chain, we used the standard N-terminus for Ala 3 (positively charged ammonium) and the standard C-terminus (carboxylate) for Val 1332.

As C, N, and O can often not be distinguished unambiguously from X-ray diffraction data, the orientation of Asn and Gln side-chain amides and of the His imidazole rings was adjusted with the program *Reduce*.<sup>5</sup> The automatic assignments of Asn, Gln, and His side-chain orientation and of the His protonation state were reevaluated manually and adjusted where necessary. Flipped residues are listed below (if the residue has been flipped in just one of the chains, the chain ID is given in brackets)

### **FLIPPED RESIDUES:**

ASN 71, GLN 131, ASN 146(B), ASN 351, GLN 473, ASN 556, GLN 626, ASN 650, HSE 683(A), HSD 747, HSD 875, ASN 1145, HSD 1151, GLN 1284, ASN 1287(B) ASN 1324(A)

All Lys and Arg side chains are in the ammonium or guanidinium forms, respectively, and all Asp side chains are in the carboxylate form. For all setups, Glu267 was protonated and Cys992 directly neighboring the doubly protonated histidine Hsp677 was deprotonated. Glu802 was protonated in both chains (except for setup E where it was protonated only in chain B). The chosen protonation states for all histidines are listed below (if the protonation state has been assigned in just one of the two chains, the chain ID is given in brackets)

### **PROTONATION STATES**

GLU 267p, GLU802p (SETUP E: just chain B), CYS 992d

HSD: 67 187 252 292(A) 552(B) 614 747 821 840 863 875 884 1151 1171(B) 1212 (B)

HSE: 81 82 99 109 292(B) 387 552(A) 579 665 683 741 954 1022 1033 1043 1171(A) 1212 (A)1220

HSP: 677 1285

SETUP A, C, D, E, F, G: GLU 1261 (just chain A)

To keep the whole system neutral in charge, Cl<sup>-</sup> counterions were added. This was done using the program *ionize*,<sup>6</sup> which puts a three dimensional grid on the enzyme and then iteratively calculates the electrostatic potential on all grid points, adding a single ion at the most favorable position until the system is neutral.

The residues in the direct surrounding of the counterions were reoptimized using CHARMM. As a result of this procedure all counterions were placed on the surface of the enzyme and more than 35 Å away from the molybdenum center. Their coordinates are given below for direct use in the original crystal structure.

#### ADDED COUNTERIONS

ATOM	41516	CLA	CLA	8	75.360	86.126	20.686
ATOM	41517	CLA	CLA	9	28.661	-47.776	50.009
in SETUP A,C,D,F & G							
ATOM	41519	CLA	CLA	10	93.374	14.979	26.149

#### B. CHARMM Parameters.

To be able to prepare the system by extended classical MD runs prior to the QM/MM optimizations, we needed CHARMM parameters for the full system, including FAD, FEST, MOC and URIC. The FAD parameters were provided by M. Karplus.<sup>7</sup> We used the FEST parameters published in our prior publication.<sup>8</sup> We kept the molybdenum center fixed during MD simulation and used ESP charges for the interaction with the enzymatic environment. The parameters for URIC (uric acid) were adopted from similar molecules like guanine or uracil; the atomic charges were taken from ESP calculations.

Note that our goal was only to obtain reasonable parameters suitable for the preparatory MM-MD simulations. We therefore did not extensively validate the newly derived parameters, nor do we recommend their use in pure MM calculations without further tests. All parameters not provided with the CHARMM package can be found in Section 4 (see pages S50-S60).

#### C. Hydration Procedure.

The protein structure with added H atoms was hydrated using a sphere of 35 Å radius containing 6840 water molecules cut from an equilibrated simulation of TIP3P water under periodic boundary conditions. The water sphere was superimposed onto the protein structure, centered on MOC(A):C8. All added water molecules whose O atom was within 2.8 Å of any existing non-H atom were deleted. The O atoms of all water molecules (i.e., added and existing ones) within the 35 Å sphere were subjected to a spherical quartic boundary potential with parameters FORCE = 0.2, P1 = 2.25, and DROFF = 33.5. The water molecules were kept rigid using SHAKE constraints.

Keeping all other atoms fixed, the water molecules within the sphere were geometry-optimized, performing first 300 steps of steepest descent followed by 300 steps of adapted-basis Newton–Raphson minimization. In the first solvation run, we performed another two minimizations of 300 steps each (steepest descent and adapted-basis Newton Raphson), optimizing the active region without the water molecules, adding a restraint of  $k=30 \text{ mol } \text{\AA}^{-2}$  on the non-fixed atoms. Finally, the water molecules were subjected to 30 ps molecular dynamics at 300 K with (subsequently decreasing) restraints. Afterward, we resteped to the next solvation cycle. Information on the different solvation cycles is listed below.

## HYDRATION CYCLES

Run	number of water molecules added	$k/(\text{mol } \text{\AA}^{-2})^*$
1	1626	fixed
2	64	50
3	69	45
4	51	40
5	44	35
6	30	30
7	32	25
8	26	20
9	21	15
10	52	10
11	27	5
12	24	3
13	34	1

\*force constant of harmonic constraints on all residues within 20 Å around C8(URIC), all other residues and the cofactor are fixed

The nonbonded interactions in all CHARMM simulations were treated using  
 NBOND ELEC ATOM FSWITCH CDIE VDW VATOM VFSWITCH CUTNB 17. CTONNB  
 11. CTOFNB 15. EPS 1.  
 NBXMOD 5 E14FAC 1. WMIN 1.5.

### Active residues

#### chain A

GLN112, CYS113, CYS148, ARG149, CYS150, THR151, GLY152, HSE579, TYR592, THR646, GLY647, LEU648, PHE649, ASN650, ASP651, GLU652, ILE698, TYR708, LEU712, LYS713, ILE714, ILE736, GLY737, GLY738, GLN739, ASP740, HSE741, PHE742, TYR743, LEU744, GLU745, THR746, HSD747, CYS748, VAL764, SER765, THR766, GLN767, ASN768, ALA769, MET770, LYS771, THR772, GLN773, SER774, PHE775, VAL789, VAL791, LYS792, ARG793, MET794, GLY795, GLY796, GLY797, PHE798, GLY799, GLY800, LYS801, GLU802, THR803, ARG804, SER805, THR806, LEU807, VAL808, SER809, VAL810, LEU827, ARG829, ASP832, MET833, ILE835, THR836, GLY837, GLY838, ARG839, HSD840, PRO841, PHE842, LEU843, ALA844, HSD863, TYR864, SER865, ASN866, ALA867, GLY868, ASN869, SER870, ARG871, ASP872, LEU873, SER874, HSD875, SER876, ILE877, MET878, GLU879, ARG880, ALA881, LEU882, PHE883, HSD884, MET885, ASP886, ASN887, CYS888, THR897, GLY898, ARG899, LEU900, CYS901, SER906, SER907, ASN908, THR909, ALA910, PHE911, ARG912, GLY913, PHE914, GLY915, GLY916, PRO917, GLN918, ALA919, LEU920, PHE921, ILE922, ALA923, GLU924, TYR947, THR953, HSE954, PHE955, GLN957, LEU959, PHE962, VAL964, ILE1001,

PRO1002, THR1003, LYS1004, PHE1005, GLY1006, ILE1007, SER1008, PHE1009, THR1010, VAL1011, PRO1012, PHE1013, LEU1014, ASN1015, GLN1016, ALA1017, GLY1018, ALA1019, HSE1033, GLY1034, GLY1035, THR1036, GLU1037, MET1038, GLY1039, GLN1040, GLY1041, LEU1042, HSE1043, THR1044, LYS1045, MET1046, VAL1049, ILE1063, THR1066, SER1067, THR1068, VAL1071, PRO1072, ASN1073, SER1074, SER1075, PRO1076, THR1077, ALA1078, ALA1079, SER1080, VAL1081, SER1082, THR1083, ASP1084, ILE1085, TYR1086, GLY1087, GLN1088, ALA1089, VAL1090, PHE1132, TYR1133, ARG1134, THR1135, PRO1136, ASN1137, LEU1138, GLY1139, TYR1140, SER1141, PHE1142, GLY1147, ASN1148, ALA1149, PHE1150, HSD1151, TYR1152, PHE1153, THR1154, TYR1155, GLY1156, ILE1178, MET1180, ASP1181, VAL1182, GLY1183, SER1185, LEU1186, ASN1187, ILE1190, ASP1191, ILE1192, GLY1193, GLN1194, VAL1195, GLU1196, GLY1197, ALA1198, PHE1199, VAL1200, GLN1201, GLY1202, LEU1203, GLY1204, LEU1205, PHE1206, GLU1209, LEU1211, TYR1213, TYR1227, ILE1229, PRO1230, LEU1243, ASN1249, LYS1251, ALA1252, ILE1253, SER1256, LYS1257, ALA1258, VAL1259, GLY1260, GLU1261, PRO1262, PRO1263, LEU1264, PHE1265, LEU1266, GLY1267, ALA1268, SER1298, FES2001, MOC2003, URIC2, GOL2, CAL6

chain B

THR1025, TYR1121, GLN1122, ASP1123, ARG1124

crystal water

CRYW3, CRYW4, CRYW5, CRYW7, CRYW9, CRYW11, CRYW12, CRYW13, CRYW15, CRYW19, CRYW23, CRYW24, CRYW25, CRYW40, CRYW42, CRYW45, CRYW48, CRYW49, CRYW52, CRYW56, CRYW58, CRYW61, CRYW63, CRYW65, CRYW71, CRYW74, CRYW77, CRYW85, CRYW89, CRYW91, CRYW93, CRYW94, CRYW96, CRYW97, CRYW101, CRYW108, CRYW109, CRYW114, CRYW122, CRYW125, CRYW126, CRYW128, CRYW143, CRYW157, CRYW160, CRYW168, CRYW174, CRYW177, CRYW189, CRYW191, CRYW196, CRYW199, CRYW200, CRYW203, CRYW211, CRYW213, CRYW218, CRYW222, CRYW224, CRYW225, CRYW227, CRYW230, CRYW240, CRYW243, CRYW244, CRYW248, CRYW274, CRYW277, CRYW283, CRYW287, CRYW300, CRYW308, CRYW317, CRYW343, CRYW369, CRYW373, CRYW382, CRYW405, CRYW408, CRYW428, CRYW429, CRYW447, CRYW498, CRYW536, CRYW558, CRYW563, CRYW565, CRYW570, CRYW571, CRYW575, CRYW596, CRYW615, CRYW632, CRYW636, CRYW706, CRYW712, CRYW751, CRYW753, CRYW798, CRYW847, CRYW947, CRYW1034, CRYW1128, CRYW1205, CRYW1340, CRYW1678

and all solvent water

## References:

1. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, 18, 2714-2723.
2. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, 31, 3381-3385.
3. Kopp, J.; Schwede, T. *Nucleic Acids Res.* **2004**, 32, D230-D234.
4. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, 22, 195-201.
5. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, 285, 1735-1747.
6. <http://www.ks.uiuc.edu/Development/MDTools/ionize>.
7. Luo, G. B.; Andricioaei, I.; Xie, X. S.; Karplus, M. *J. Phys. Chem. B* **2006**, 110, 9363-9367.
8. Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, 131, 4628-4640.

## 2. QM/MM results

### SETUP\_A, SN100

**Table S1.** Calculated QM/MM energies of optimized stationary points for snapshot 100 of SETUP A in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	4.0	0.2	4.2	7.1	-0.2	7.0	N3-N9 Proton transfer
IM1'	-2.0	2.3	0.3	-2.8	2.5	-0.4	
TS1	8.5	7.2	15.8	7.9	6.9	14.8	Proton transfer
IM1	6.1	5.6	11.8	6.9	5.3	12.2	
TS2	16.3	0.9	17.2	19.9	0.3	20.2	C-O formation
IM2	14.6	0.5	15.2	17.7	0.1	17.9	
TS3	18.6	0.2	18.8	22.0	-1.2	20.8	Hydride transfer
Product	6.0	6.1	12.1	-5.4	6.5	1.2	

<sup>a</sup> absolute energies (a.u.): -1423.990924 -87.286581 -1511.277505 (BP86)  
-1422.892054 -87.284641 -1510.176695 (B3LYP)

#### active MM region SETUP A SN100 (1588 active atoms)

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794  
GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805  
GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875 SER876  
ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908 THR909  
ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917 GLN918  
LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010 VAL1011  
PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036 MET1038 GLN1040  
LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078 ALA1079 SER1080  
VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086 THR1135 LEU1138  
ALA1149 PHE1150 HSD1151 THR1154 GLN1194 ALA1198 GLN1201 ALA1258  
VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003  
URIC2 GOL2

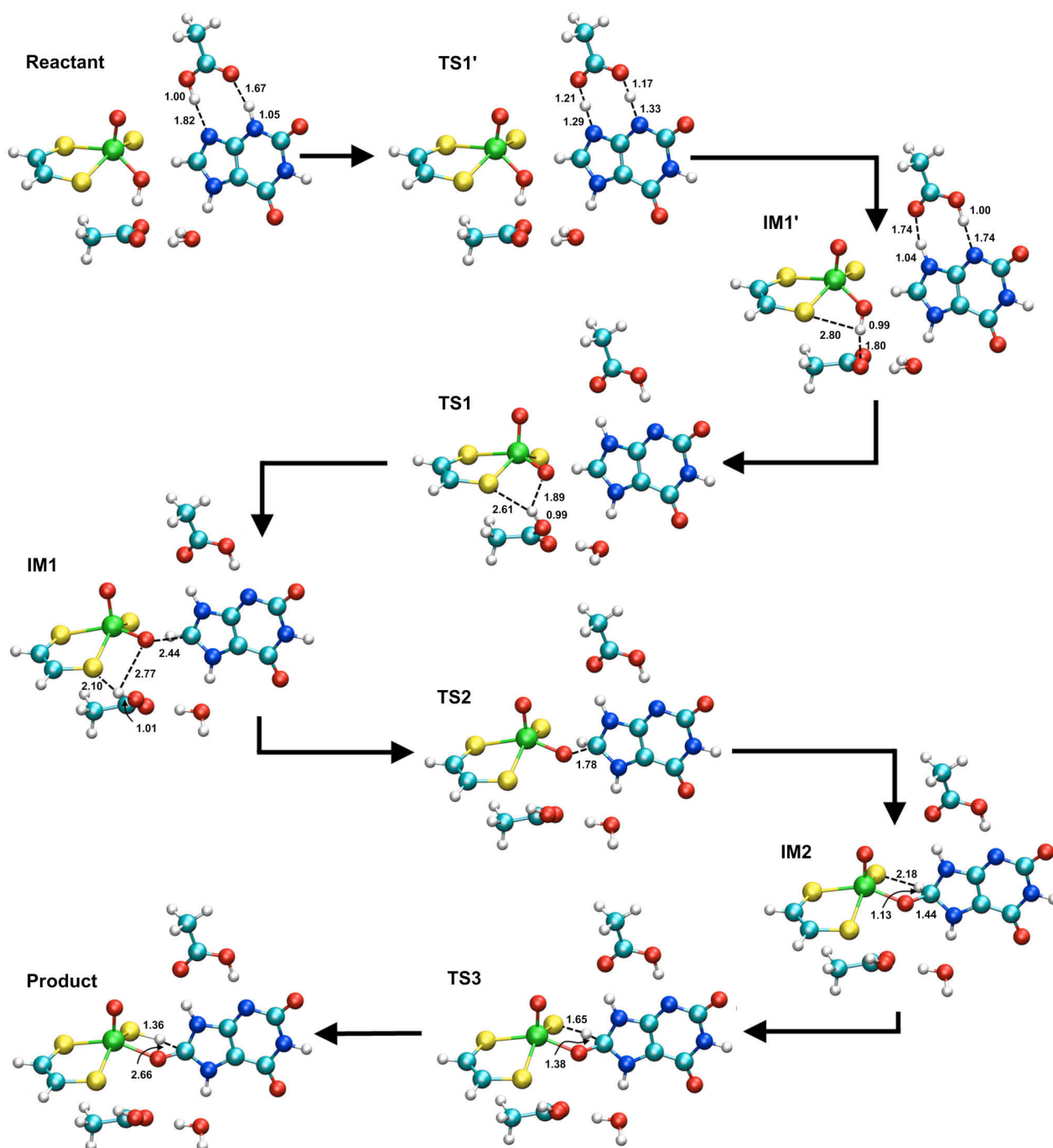
crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW126  
CRYW168 CRYW174 CRYW189 CRYW191 CRYW211 CRYW222 CRYW224 CRYW240  
CRYW244 CRYW248 CRYW287 CRYW288 CRYW343 CRYW369 CRYW373 CRYW405  
CRYW428 CRYW447 CRYW498 CRYW570 CRYW571 CRYW573 CRYW575 CRYW706  
CRYW753 CRYW847 CRYW947 CRYW1034

solvent water

SOLV250 SOLV376 SOLV771 SOLV845 SOLV1087 SOLV1559 SOLV1845  
SOLV1991

# SETUP\_A, SN500



**Figure S1.** Reaction mechanism for setup A, SN500 from B3LYP/MM calculations. The distances are given in Angstrom.

**Table S2.** QM/MM energies of optimized stationary points of SETUP A, SN500 in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	3.6	0.0	3.6	6.8	2.1	8.9	N3-N9 Proton transfer
IM1'	-0.4	-0.3	-0.7	1.6	0.8	2.4	
TS1	16.0	2.3	18.3	15.3	3.8	19.1	Proton transfer
IM1	3.0	10.6	13.6	3.5	11.4	14.9	
TS2	9.0	7.0	16.1	12.5	7.8	20.3	C-O formation
IM2	5.6	7.5	13.1	7.1	9.4	16.4	
TS3	10.0	7.8	17.9	12.7	8.9	21.7	Hydride transfer
Product	-0.7	10.2	9.6	-11.7	12.2	0.5	

<sup>a</sup> absolute energies (a.u.): -1423.935494 -87.129750 -1511.065244 (BP86)  
-1422.837158 -87.128789 -1509.965947 (B3LYP)

**Table S3.** QM/MM energies of the rate limiting stationary points of SETUP A, SN500 in kcal mol<sup>-1</sup> using the L2 link atom scheme, from single point calculations (SP) on the structures obtained with the charge-shift link atom scheme (Figure S1) and on the reoptimized structures (OPT).

	B3LYP/Lanl2DZ (SP)			B3LYP/Lanl2DZ (OPT)			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS3	13.2	7.2	20.4	13.5	6.9	20.4	Hydride transfer

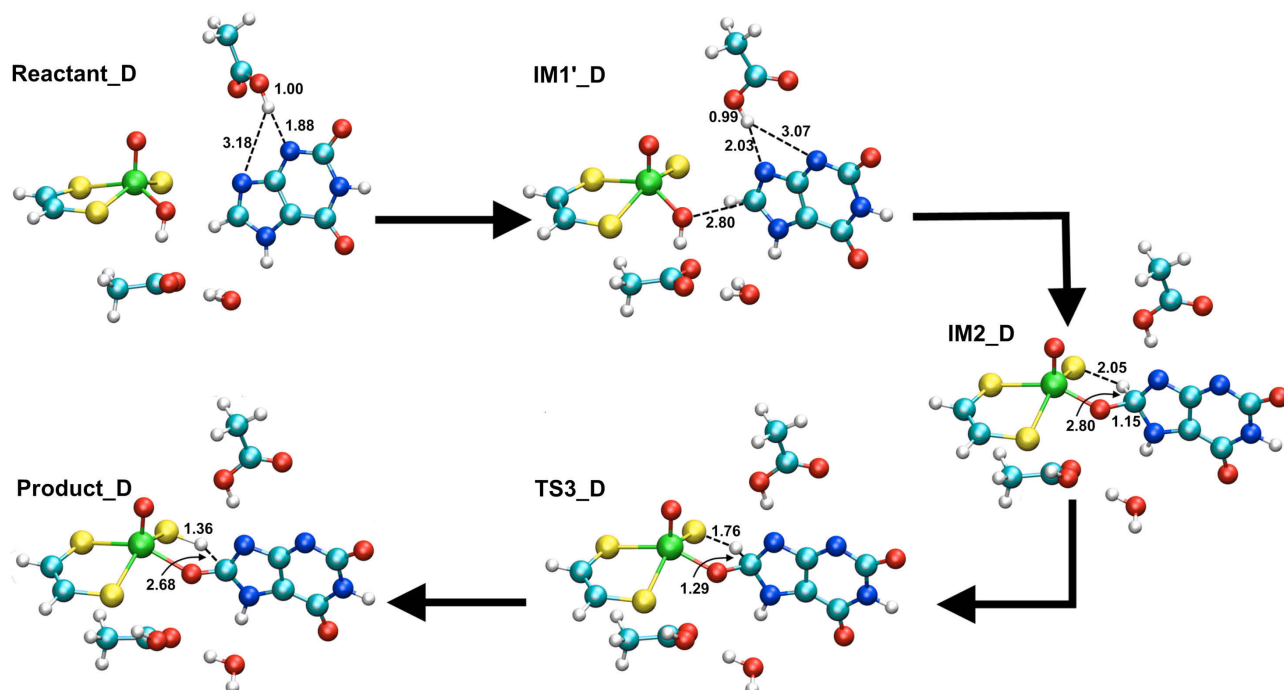
<sup>a</sup> absolute energies (a.u.): -1422.814695 -87.128789 -1509.943484 (single point)  
-1422.815197 -87.129676 -1509.944873 (reoptimized)

**Table S4.** QM/MM energies of the most important stationary points (reoptimized) of SETUP A, SN500 in kcal mol<sup>-1</sup> after deprotonation of Glu802.

	B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant_D <sup>a</sup>	0.0	0.0	0.0	
IM1'_D	7.8	2.7	10.5	
IM2_D	28.2	13.5	41.7	
TS3_D	28.8	13.6	42.3	Hydride transfer
Product_D	-6.5	18.4	12.0	

<sup>a</sup> absolute energies (a.u.): -1422.381161 -87.056611 -1509.437771





**Figure S2.** Intermediates of the reaction mechanism for SETUP A, SN500 from B3LYP/MM calculations after deprotonation of Glu802. The distances are given in Angstrom.

**active MM region SETUP A SN500 (1594 active atoms)**

chain A

LEU648	PHE649	HSE741	TYR743	THR766	GLN767	ASN768	LYS771	MET794
GLY797	PHE798	GLY799	GLY800	LYS801	GLU802	THR803	ARG804	SER805
GLY838	ARG839	HSD840	PHE842	ASP872	LEU873	SER874	HSD875	SER876
ILE877	MET878	GLU879	ARG880	ALA881	PHE883	HSD884	ASN908	THR909
ALA910	PHE911	ARG912	GLY913	PHE914	GLY915	GLY916	PRO917	GLN918
LEU920	PHE1005	GLY1006	ILE1007	SER1008	PHE1009	THR1010	VAL1011	
PRO1012	PHE1013	LEU1014	ASN1015	GLN1016	GLY1035	THR1036	MET1038	
GLN1040	LEU1042	SER1074	SER1075	PRO1076	THR1077	ALA1078	ALA1079	
SER1080	VAL1081	SER1082	THR1083	ASP1084	ILE1085	TYR1086	ALA1149	
PHE1150	HSD1151	THR1154	GLN1194	ALA1198	GLN1201	LEU1205	ALA1258	
VAL1259	GLY1260	GLU1261	PRO1262	PRO1263	LEU1264	PHE1265	MOC2003	

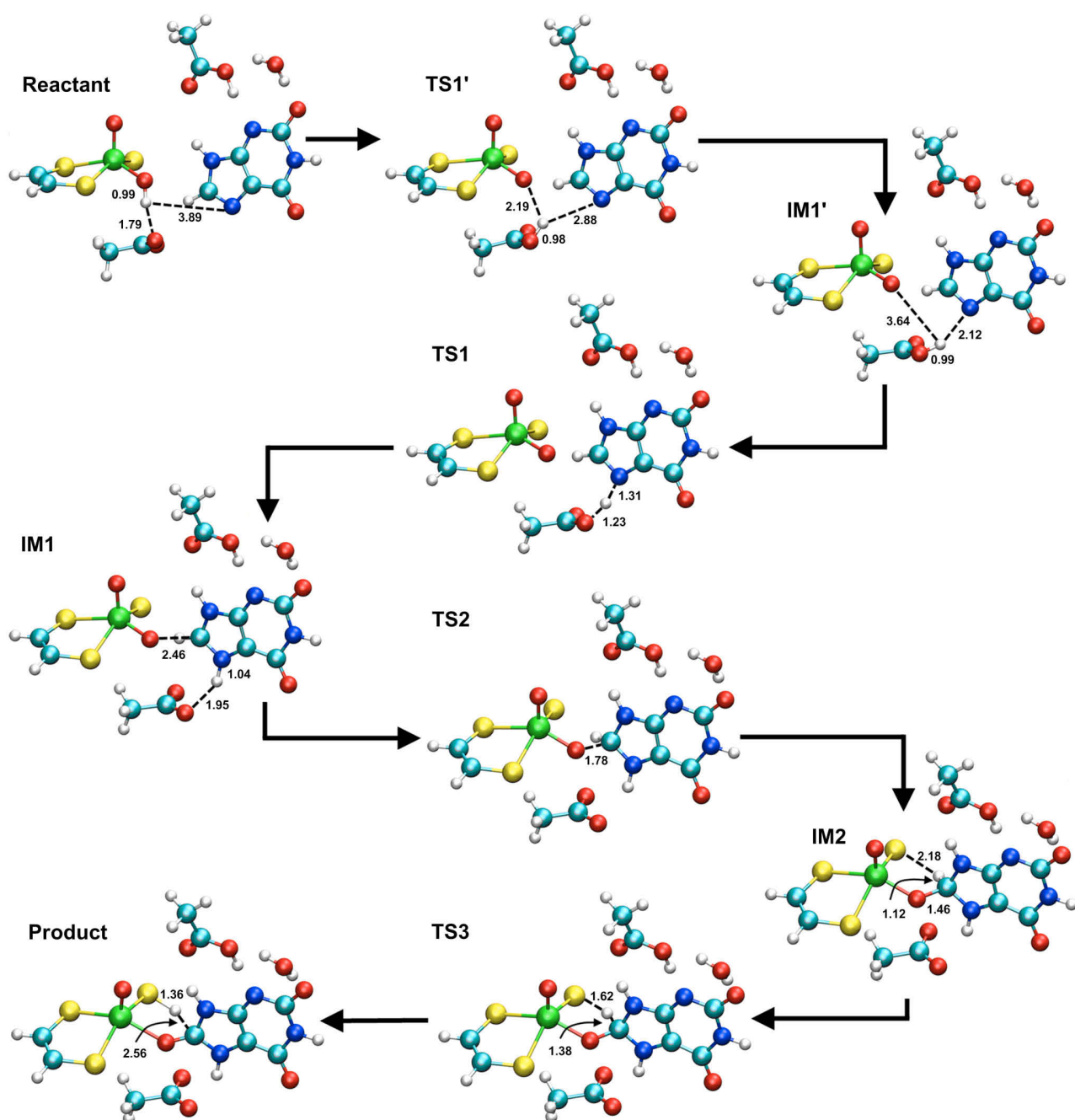
URI2 GOL2

crystal water

CRYW19	CRYW40	CRYW42	CRYW45	CRYW48	CRYW56	CRYW58	CRYW74	CRYW126
CRYW168	CRYW174	CRYW189	CRYW211	CRYW222	CRYW224	CRYW240	CRYW248	
CRYW287	CRYW343	CRYW369	CRYW373	CRYW382	CRYW405	CRYW428	CRYW447	
CRYW498	CRYW536	CRYW571	CRYW575	CRYW706	CRYW753	CRYW847	CRYW1034	

solvent water

SOLV105	SOLV842	SOLV1435	SOLV1559	SOLV1688	SOLV1964	SOLV1991
---------	---------	----------	----------	----------	----------	----------



**Figure S3.** Reaction mechanism for setup B, SN300 from B3LYP/MM calculations. The distances are given in Angstrom. For the sake of clarity, we omitted an additional water molecule close to Glu1261 (cf. Figure 3) that is included in the QM region.

**Table S5.** Calculated QM/MM energies of optimized stationary points for snapshot 300 of SETUP B in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	17.5	2.2	19.8	18.3	0.1	18.4	Proton transfer 1
IM1'	7.7	2.1	9.8	11.2	0.2	11.5	
TS1	8.2	3.6	11.9	12.7	2.4	15.1	Proton transfer 2
IM1	4.9	3.4	8.4	7.4	3.3	10.8	
TS2	10.2	0.5	10.7	19.2	-1.4	17.9	C-O formation
IM2	9.1	-0.5	8.6	16.6	-2.2	14.4	
TS3	15.8	-1.2	14.5	23.3	-3.4	19.9	Hydride transfer
Product	5.0	3.6	8.6	-3.0	-3.3	0.3	

<sup>a</sup> absolute energies (a.u.): -1500.044484 -82.462433 -1582.506917 (BP86)  
-1498.911278 -82.458006 -1581.369284 (B3LYP)

**active MM region SETUP B, SN300 (1635 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794  
GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 GLY837  
GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875 SER876  
ILE877 MET878 GLU879 ARG880 ALA881 LEU882 PHE883 HSD884 ASN908  
THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
GLN918 ALA919 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
THR1010 VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036  
MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
LEU1138 ALA1149 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198  
GLN1201 LEU1205 ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263  
LEU1264 PHE1265 MOC2003 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW168  
CRYW174 CRYW189 CRYW191 CRYW196 CRYW203 CRYW211 CRYW222 CRYW224  
CRYW244 CRYW248 CRYW287 CRYW317 CRYW343 CRYW369 CRYW405 CRYW447  
CRYW498 CRYW571 CRYW573 CRYW575 CRYW847 CRYW947 CRYW1034

solvent water

SOLV60 SOLV95 SOLV338 SOLV506 SOLV589 SOLV737 SOLV1423 SOLV1559  
SOLV1991

## SETUP\_B, SN500

**Table S6.** Calculated QM/MM energies of optimized stationary points for Snapshot 500 of SETUP B in kcal mol<sup>-1</sup>.

	BP86/Lan12DZ			B3LYP/Lan12DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	17.2	1.4	18.6	20.0	0.2	20.2	Proton transfer1
IM1'	11.7	2.7	14.3	13.0	1.0	14.1	
TS1	11.9	4.4	16.3	16.5	4.0	20.4	Proton transfer2
IM1	10.1	4.3	14.4	12.5	3.5	16.0	
TS2	14.1	0.2	14.3	18.9	-0.9	18.0	C-O formation
IM2	12.5	-1.6	10.9	16.6	-2.3	14.2	
TS3	17.5	-1.6	15.9	22.3	-2.8	19.5	Hydride transfer
Product	10.1	2.6	12.7	0.7	2.8	3.5	
<sup>a</sup> absolute energies (a.u.):							
			-1500.045305 -82.390174 -1582.435479 (BP86)				
			-1498.909442 -82.386832 -1581.296274 (B3LYP)				

### active MM region SETUP B SN500 (1613 active atoms)

chain A

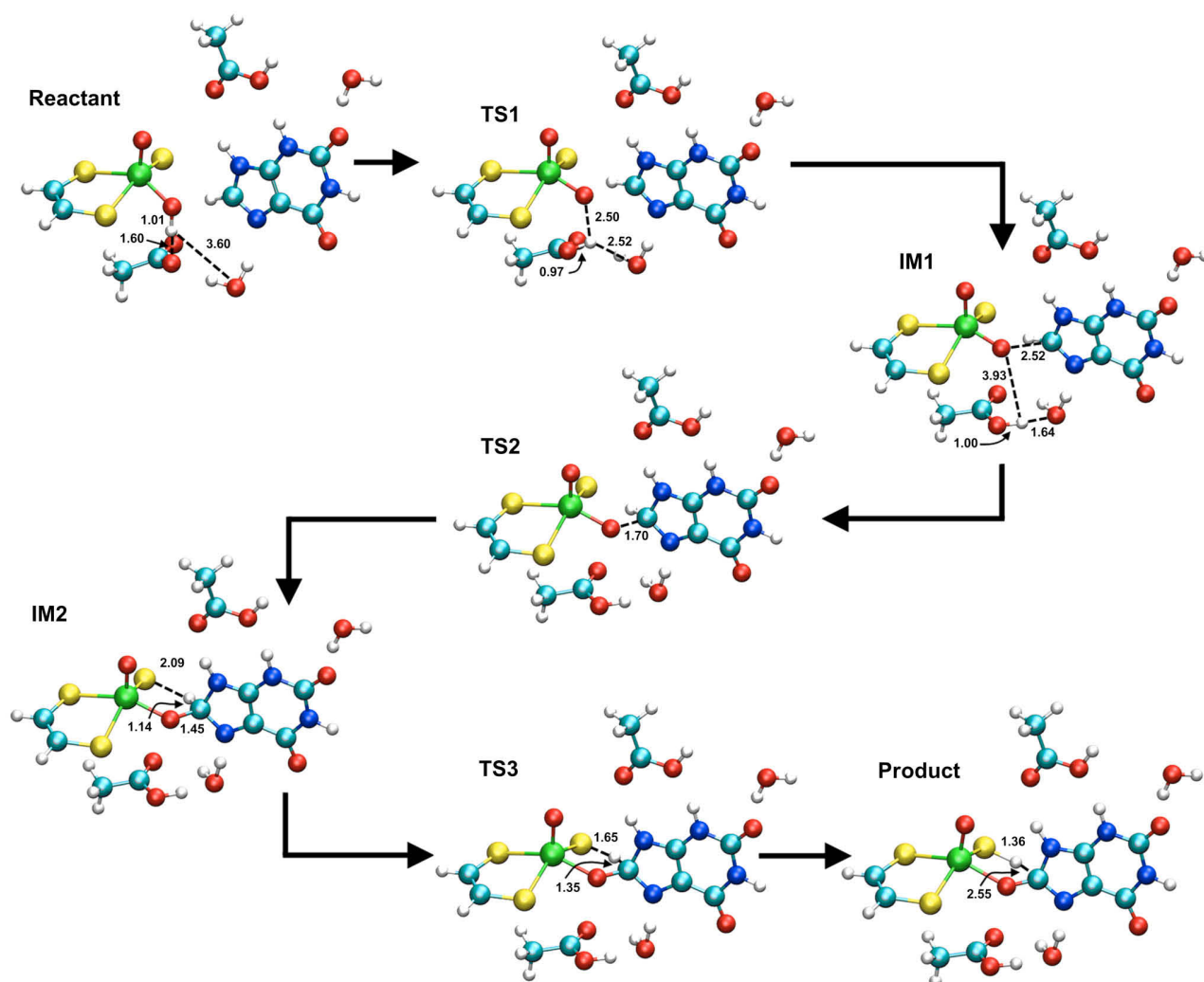
LEU648 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794 GLY797  
PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805 GLY837  
GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875 SER876  
ILE877 MET878 GLU879 ARG880 ALA881 LEU882 PHE883 HSD884 ASN908  
THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036 MET1038  
GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078 ALA1079  
SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086 LEU1138  
ALA1149 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198 GLN1201  
LEU1205 ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264  
PHE1265 MOC2003 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW168 CRYW174  
CRYW189 CRYW191 CRYW196 CRYW203 CRYW211 CRYW222 CRYW224 CRYW244  
CRYW248 CRYW287 CRYW317 CRYW343 CRYW369 CRYW405 CRYW447 CRYW498  
CRYW571 CRYW575 CRYW847 CRYW947 CRYW1034

solvent

SOLV60 SOLV106 SOLV141 SOLV503 SOLV736 SOLV737 SOLV1139 SOLV1559  
SOLV1902 SOLV1991



**Figure S4.** Reaction mechanism for setup C, SN200 from B3LYP/MM calculations. The distances are given in Angstrom.

**Table S7.** Calculated QM/MM energies of optimized stationary points for Snapshot 200 of SETUP C in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	12.5	-2.9	9.6	14.1	-3.2	10.9	Proton transfer
IM1	5.5	0.0	5.5	7.8	0.1	7.9	
TS2	15.3	-4.8	10.5	23.4	-5.4	17.9	C-O formation
IM2	13.7	-5.4	8.3	22.6	-6.3	16.3	
TS3	17.7	-5.0	12.7	24.7	-5.5	19.1	Hydride transfer
Product	6.8	-1.2	5.6	0.8	-0.5	0.3	
<sup>a</sup> absolute energies (a.u.):							
			-1500.366679 -93.136949 -1593.503628 (BP86)				
			-1499.232154 -93.134585 -1592.366739 (B3LYP)				

**active MM region SETUP C, SN200 (1581 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794  
 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805  
 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875  
 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908  
 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
 GLN918 LEU920 LYS1004 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
 THR1010 VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036  
 MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
 ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
 ALA1149 PHE1150 HSD1151 THR1154 GLN1194 ALA1198 GLN1201 ALA1258  
 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003  
 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW126 CRYW168  
 CRYW174 CRYW189 CRYW191 CRYW211 CRYW222 CRYW224 CRYW248 CRYW343  
 CRYW369 CRYW373 CRYW405 CRYW429 CRYW447 CRYW498 CRYW571 CRYW573  
 CRYW575 CRYW706 CRYW753 CRYW847 CRYW947 CRYW1034 CRYW1911

solvent

SOLV154 SOLV506 SOLV735 SOLV845 SOLV1107 SOLV1228 SOLV1559  
 SOLV1566 SOLV1589

**SETUP\_C, SN400****Table S8.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP C in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	15.1	-1.9	13.2	18.0	-3.1	14.9	Proton transfer
IM1	12.0	-1.0	11.0	15.8	-2.4	13.4	
TS2	25.9	-7.6	18.2	35.5	-8.6	26.6	C-O formation
IM2	19.8	-6.2	13.5	25.9	-7.4	18.5	
TS3	28.4	-8.6	19.8	37.1	-9.2	27.9	Hydride transfer
Product	13.9	-2.9	11.0	6.9	-4.0	2.9	

<sup>a</sup> absolute energies (a.u.): -1500.442697 -83.783296 -1584.225993 (BP86)  
-1499.311013 -83.779643 -1583.090656 (B3LYP)

**active MM region SETUP C, SN400 (1612 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804  
SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874  
HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884  
ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916  
PRO917 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
THR1010 VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036  
MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
LEU1138 ALA1149 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198  
GLN1201 ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264  
PHE1265 MOC2003 URIC2 GOL2

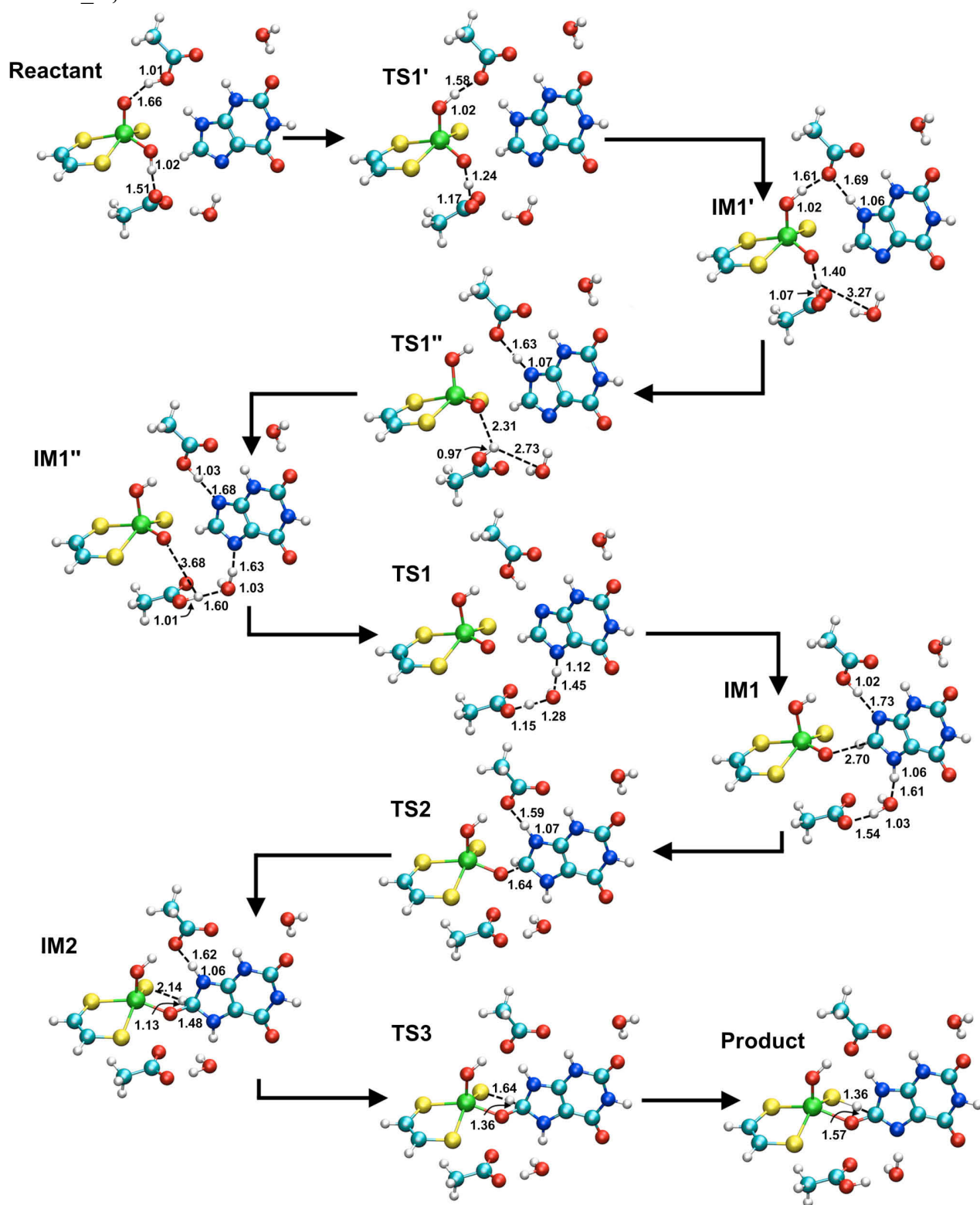
crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW168  
CRYW174 CRYW189 CRYW191 CRYW199 CRYW211 CRYW222 CRYW224 CRYW248  
CRYW287 CRYW343 CRYW369 CRYW373 CRYW405 CRYW428 CRYW447 CRYW498  
CRYW571 CRYW573 CRYW575 CRYW706 CRYW753 CRYW847 CRYW947 CRYW1034

solvent

SOLV506 SOLV686 SOLV799 SOLV1139 SOLV1187 SOLV1559 SOLV1591  
SOLV1976 SOLV1991

SETUP\_D, SN200



**Figure S5.** Reaction mechanism for setup D, SN200 from B3LYP/MM calculations. The distances are given in Angstrom.



**Table S9.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 200 of SETUP D in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1"	4.5	-1.8	1.0	11.0	-5.6	5.4	Proton transfer1
IM1"	2.4	-3.5	-1.0	10.5	-5.8	4.8	
TS1'	4.5	-3.5	1.0	11.8	-5.2	6.6	OH-rotation
IM1'	3.0	-5.2	-2.2	10.2	-6.4	3.8	
TS1	5.4	-5.2	0.2	12.7	-5.1	7.6	Proton transfer2
IM1	5.6	-5.4	0.2	12.8	-6.0	6.7	
TS2	22.3	-8.4	13.9	37.1	-10.1	27.0	C-O formation
IM2	22.4	-8.7	13.7	35.8	-9.9	25.9	
TS3	25.9	-7.9	18.0	37.6	-9.1	28.5	Hydride transfer
Product	13.0	-5.5	7.5	11.9	-5.9	6.0	

<sup>a</sup> absolute energies (a.u.): -1500.418458 -94.061978 -1594.480436 (BP86)  
-1499.290110 -94.058351 -1593.348461 (B3LYP)

**active MM region SETUP D, SN200 (1589 active atoms)**

chain A

LEU648 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771 MET794  
GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805  
GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875  
SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908  
THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
VAL1011 PRO1012 PHE1013 LEU1014 ASN1015 GLN1016 GLY1035 THR1036  
MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
LEU1138 ALA1149 PHE1150 HSD1151 THR1154 ALA1198 GLN1201 ALA1258  
VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003  
URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW126 CRYW168  
CRYW174 CRYW189 CRYW191 CRYW203 CRYW211 CRYW222 CRYW224 CRYW244  
CRYW248 CRYW287 CRYW343 CRYW369 CRYW373 CRYW405 CRYW418 CRYW447  
CRYW498 CRYW571 CRYW575 CRYW706 CRYW753 CRYW847 CRYW947 CRYW1034

solvent

SOLV60 SOLV504 SOLV676 SOLV779 SOLV973 SOLV1559 SOLV1575 SOLV1991

**SETUP\_D, SN400****Table S10.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP C in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	6.8	-6.2	0.6	10.9	-7.8	3.0	Proton transfer1
IM1	6.4	-11.9	-5.4	10.2	-13.4	-3.3	
TS1"	9.1	-11.4	-2.3	14.9	-12.6	2.3	Proton transfer2
IM1"	3.4	-8.0	-4.6	6.1	-7.8	-1.7	
TS2	27.9	-12.2	15.7	37.9	-13.2	24.8	C-O formation
IM2	25.9	-11.4	14.5	36.1	-12.1	23.9	
TS3	30.1	-11.6	18.5	40.9	-12.3	28.5	Hydride transfer
Product	13.7	-9.0	4.7	9.5	-9.7	-0.1	
TS2'	24.4	-9.7	14.7	34.6	-10.5	24.0	Proton transfer3
IM2'	22.7	-8.2	14.5	26.6	-8.3	18.3	
TS3'	28.1	-8.4	19.7	34.8	-8.5	26.3	Hydride transfer
Product'	12.2	-4.6	7.6	4.9	-5.0	-0.1	

<sup>a</sup> absolute energies (a.u.):    -1500.422839 -93.115010 -1593.537849 (BP86)  
                                      -1499.289170 -93.111652 -1592.400822 (B3LYP)

**active MM region SETUP D, SN400 (1575 active atoms)**

chain A

LEU648 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771 MET794  
 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 GLY838  
 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875 SER876 ILE877  
 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908 THR909 ALA910  
 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917 GLN918 LEU920  
 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010 VAL1011 LEU1014  
 ASN1015 GLN1016 GLY1034 GLY1035 THR1036 MET1038 GLN1040 LEU1042  
 SER1074 SER1075 PRO1076 THR1077 ALA1078 ALA1079 SER1080 VAL1081  
 SER1082 THR1083 ASP1084 ILE1085 TYR1086 THR1135 LEU1138 ALA1149  
 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198 GLN1201 ALA1258  
 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003

URIC2 GOL2

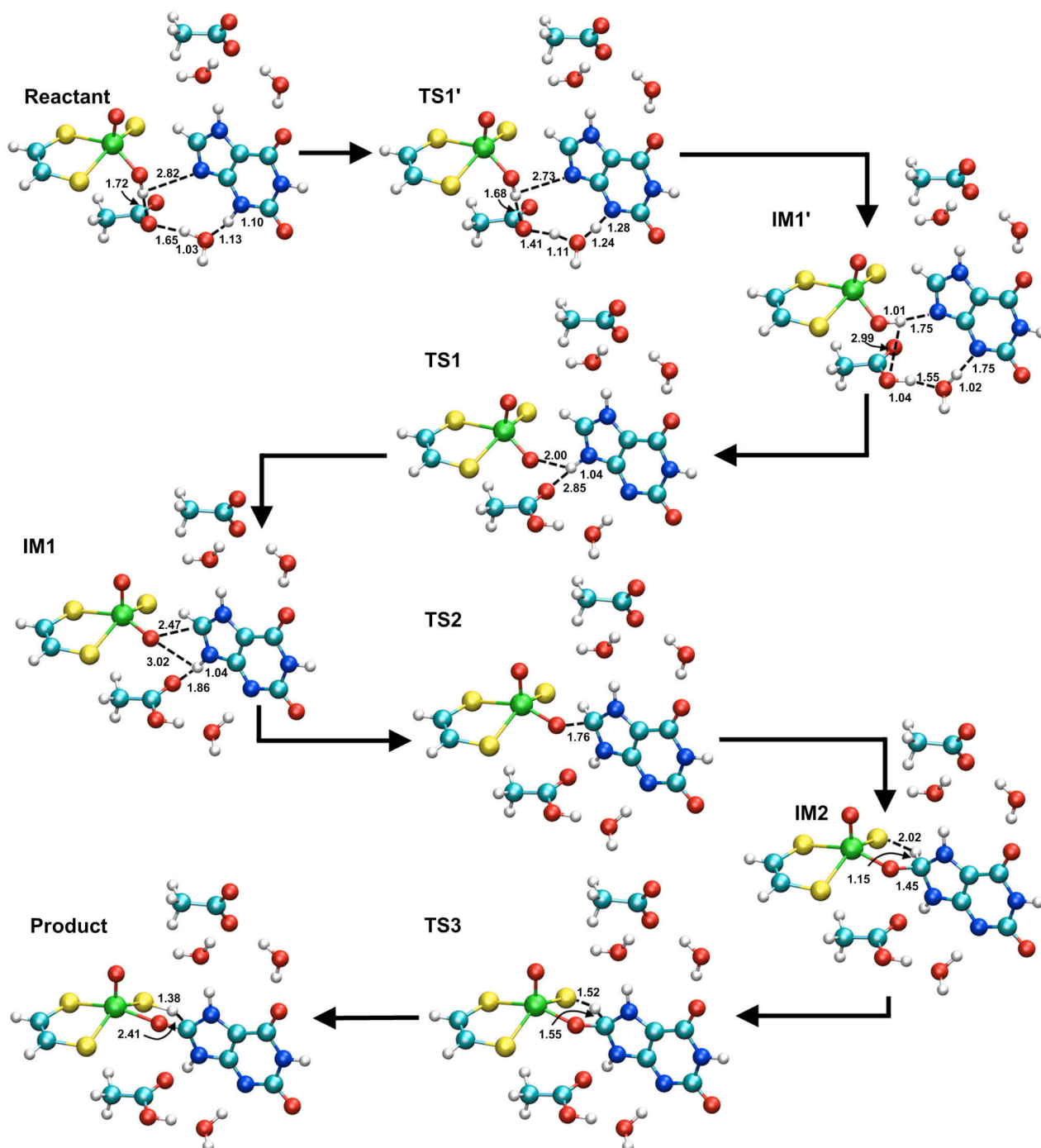
crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW126  
 CRYW168 CRYW174 CRYW189 CRYW211 CRYW222 CRYW224 CRYW248 CRYW287  
 CRYW343 CRYW369 CRYW373 CRYW405 CRYW418 CRYW447 CRYW498 CRYW571  
 CRYW575 CRYW706 CRYW847 CRYW947 CRYW1034

solvent

SOLV60 SOLV97 SOLV245 SOLV359 SOLV660 SOLV814 SOLV1046 SOLV1559  
 SOLV1585 SOLV1991

SETUP\_E, SN100



**Figure S6.** Reaction mechanism for setup E, SN100 from B3LYP/MM calculations. The distances are given in Angstrom.

**Table S11.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 100 of SETUP E in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	1.2	0.0	1.2	2.6	0.2	2.8	Proton transfer 1
IM1'	-4.9	-0.2	-5.1	-6.0	0.1	-5.9	
TS1	0.5	1.8	2.4	0.9	1.6	2.5	Proton transfer 2
IM1	-5.0	4.2	-0.9	-4.6	4.1	-0.4	
TS2	4.7	1.7	6.3	8.7	1.3	10.1	C-O formation
IM2	3.5	1.3	4.8	6.8	1.2	8.0	
TS3	7.4	0.7	8.1	10.5	0.5	11.0	Hydride transfer
Product	0.8	4.3	5.1	-11.4	5.6	-5.8	

<sup>a</sup> absolute energies (a.u.): -1576.550884 -90.891787 -1667.442671 (BP86)  
-1575.372245 -90.889416 -1666.261661 (B3LYP)

**active MM region SETUP E, SN100 (1590 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804  
SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874  
HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884  
ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916  
PRO917 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
THR1010 VAL1011 PRO1012 PHE1013 LEU1014 ASN1015 GLN1016 GLY1034  
GLY1035 THR1036 MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076  
THR1077 ALA1078 ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084  
ILE1085 TYR1086 ALA1149 PHE1150 THR1154 GLN1194 ALA1198 GLN1201  
ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265  
MOC2003 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW126  
CRYW168 CRYW174 CRYW189 CRYW191 CRYW203 CRYW211 CRYW222 CRYW224  
CRYW244 CRYW248 CRYW343 CRYW369 CRYW405 CRYW428 CRYW429 CRYW447  
CRYW498 CRYW570 CRYW571 CRYW573 CRYW575 CRYW706 CRYW753 CRYW847  
CRYW943 CRYW947 CRYW1034

solvent water

SOLV250 SOLV342 SOLV430 SOLV676 SOLV869 SOLV1207 SOLV1423 SOLV1559  
SOLV1991

**SETUP\_E, SN300****Table S12.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 300 of SETUP E in kcal mol<sup>-1</sup>, containing one water molecule less in the QM region than Snapshot 100.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	-1.0	1.5	0.5	2.9	1.1	4.0	Proton transfer 1
IM1'	-7.6	1.8	-5.8	-8.2	1.1	-7.1	
TS1	-1.4	4.1	2.6	1.8	2.5	4.4	Proton transfer 2
IM1	-7.9	7.1	-0.8	-7.3	9.6	2.3	
TS2	-0.8	6.8	6.0	6.3	6.2	12.5	C-O formation
IM2	-1.1	6.8	5.6	3.5	6.3	9.8	
TS3	3.2	6.8	10.1	7.7	6.8	14.5	Hydride transfer
Product	-8.1	10.5	2.4	-17.7	11.4	-6.3	
<sup>a</sup> absolute energies (a.u.):							
				-1500.067181 -91.057880 -1591.125061 (BP86)			
				-1498.932206 -91.055535 -1589.987742 (B3LYP)			

**active MM region SETUP E, SN300 (1585 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794  
 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805  
 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875  
 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908  
 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
 VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1035 THR1036 MET1038  
 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078 ALA1079  
 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086 ALA1149  
 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198 GLN1201 ALA1258  
 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003  
 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW126  
 CRYW168 CRYW174 CRYW189 CRYW203 CRYW211 CRYW222 CRYW224 CRYW244  
 CRYW248 CRYW343 CRYW369 CRYW382 CRYW405 CRYW428 CRYW429 CRYW447  
 CRYW498 CRYW571 CRYW575 CRYW706 CRYW753 CRYW817 CRYW847 CRYW943  
 CRYW947 CRYW1034

solvent water

SOLV260 SOLV498 SOLV676 SOLV869 SOLV1317 SOLV1559 SOLV1927  
 SOLV1991

**SETUP\_E, SN500****Table S13.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 500 of SETUP E in kcal mol<sup>-1</sup>, containing one water molecule less in the QM region than Snapshot 100.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	1.2	0.8	2.0	2.1	0.7	2.8	Proton transfer 1
IM1'	-8.7	0.7	-8.1	-10.6	1.3	-9.3	
TS1	-0.2	4.2	4.0	0.9	4.3	5.2	Proton transfer 2
IM1	-7.4	9.1	1.7	-6.5	9.1	2.6	
TS2	0.2	6.5	6.7	4.0	6.1	10.1	C-O formation
IM2	-2.2	5.8	3.6	1.5	5.9	7.3	
TS3	2.5	5.7	8.2	5.3	5.7	11.0	Hydride transfer
Product	-2.8	8.6	5.8	-14.4	9.3	-5.0	
<sup>a</sup> absolute energies (a.u.):							
			-1500.120054 -91.178763 -1591.298818 (BP86)				
			-1498.982528 -91.176664 -1590.159193 (B3LYP)				

**active MM region SETUP E, SN500 (1624 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 LYS771 MET794  
 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804 SER805  
 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874 HSD875  
 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884 ASN908  
 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916 PRO917  
 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
 VAL1011 PRO1012 PHE1013 LEU1014 ASN1015 GLN1016 GLY1034 GLY1035  
 THR1036 MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077  
 ALA1078 ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085  
 TYR1086 THR1135 LEU1138 ALA1149 PHE1150 HSD1151 THR1154 GLN1194  
 ALA1198 GLN1201 ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263  
 LEU1264 PHE1265 MOC2003 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW168  
 CRYW174 CRYW189 CRYW191 CRYW203 CRYW211 CRYW222 CRYW224 CRYW244  
 CRYW248 CRYW343 CRYW369 CRYW405 CRYW428 CRYW429 CRYW447 CRYW498  
 CRYW571 CRYW575 CRYW706 CRYW847 CRYW947 CRYW1034 CRYW1556 SOLV498

solvent water

SOLV676 SOLV869 SOLV1057 SOLV1168 SOLV1314 SOLV1323 SOLV1559  
 SOLV1662 SOLV1927 SOLV1991

**SETUP\_F, SN200****Table S14.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 200 of SETUP F in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	0.5	0.6	1.1	2.5	0.2	2.7	Proton transfer 1
IM1'	-7.4	-0.1	-7.5	-9.3	0.3	-9.2	
TS1	-1.6	-3.4	-5.0	1.0	-3.7	-2.7	Proton transfer 2
IM1	-14.2	-2.9	-17.1	-14.0	-2.8	-16.8	
TS2	-2.3	0.3	-2.0	1.6	-0.3	1.3	C-O formation
IM2	-5.2	1.3	-4.0	-1.1	1.4	0.3	
TS3	1.8	1.7	3.5	4.0	-0.3	3.7	Hydride transfer
Product	-3.3	4.5	1.2	13.7	5.5	-8.2	
TS_P'	3.3	3.3	6.6	-7.7	3.9	-3.8	Reorientation
Product'	-3.2	2.0	-1.2	-14.3	2.1	-12.1	

<sup>a</sup> absolute energies (a.u.): -1500.381518 -87.771978 -1588.153496 (BP86)  
-1499.245314 -87.769858 -1587.015173 (B3LYP)

**active MM region SETUP F, SN200 (1565 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804  
SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874  
HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884  
ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916  
PRO917 GLN918 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
VAL1011 PRO1012 PHE1013 LEU1014 ASN1015 GLN1016 GLY1034 GLY1035  
THR1036 MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077  
ALA1078 ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085  
TYR1086 LEU1138 ALA1149 PHE1150 THR1154 GLN1194 ALA1198 GLN1201  
ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 MOC2003  
URIC2 GOL2

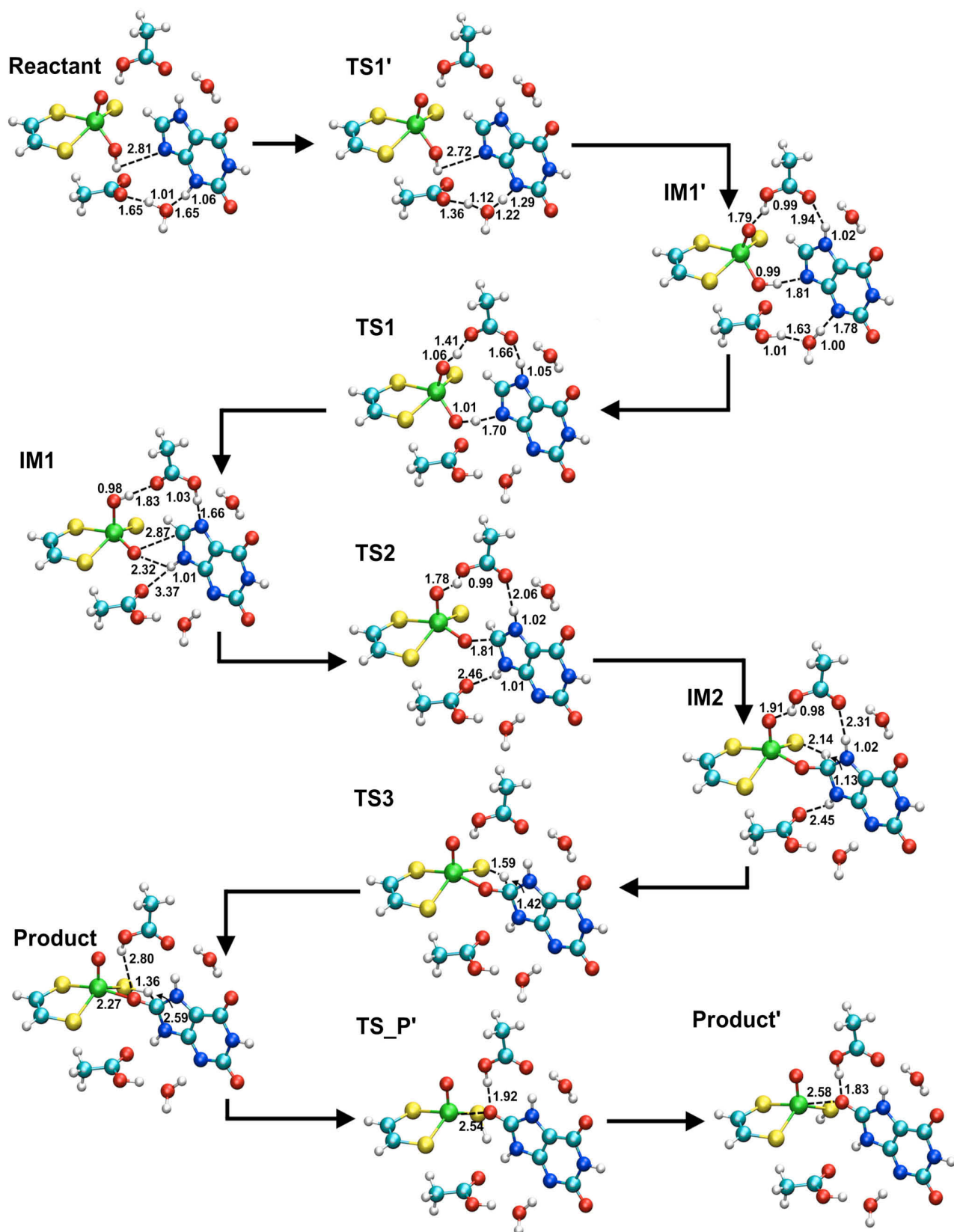
crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW126 CRYW157  
CRYW168 CRYW174 CRYW189 CRYW191 CRYW211 CRYW222 CRYW224 CRYW248  
CRYW343 CRYW369 CRYW373 CRYW405 CRYW447 CRYW498 CRYW570 CRYW571  
CRYW573 CRYW575 CRYW685 CRYW706 CRYW842 CRYW847 CRYW1034

solvent water

SOLV319 SOLV538 SOLV662 SOLV686 SOLV763 SOLV857 SOLV1061 SOLV1301  
SOLV1641 SOLV1753 SOLV1991

SETUP\_F, SN400



**Figure S7.** Reaction mechanism for setup F, SN400 from B3LYP/MM calculations. The distances are given in Angstrom.



**Table S15.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP F in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	1.6	0.5	2.0	3.6	0.7	4.2	Proton transfer 1
IM1'	-8.5	-0.1	-8.6	-9.6	0.2	-9.4	
TS1	0.1	-3.5	-3.3	4.6	-4.9	-0.4	Proton transfer 2
IM1	-15.6	2.7	-12.9	-15.5	2.4	-13.1	
TS2	0.8	1.7	2.5	5.1	0.9	6.0	C-O formation
IM2	-2.1	0.5	-1.6	1.7	0.2	2.0	
TS3	4.3	0.3	4.6	7.1	0.1	7.2	Hydride transfer
Product	-1.4	2.3	0.8	-13.0	3.9	-9.1	
TS_P'	2.9	1.8	4.7	-7.4	2.0	-5.3	Reorientation
Product'	-3.5	-1.2	-4.7	-13.7	-0.9	-14.6	
<sup>a</sup> absolute energies (a.u.):							
			-1500.35366417 -87.22039847 -1587.57406265 (BP86)				
			-1499.21756940 -87.21864060 -1586.43621000 (B3LYP)				

**active MM region SETUP F, SN400 (1523 active atoms)**

chain A

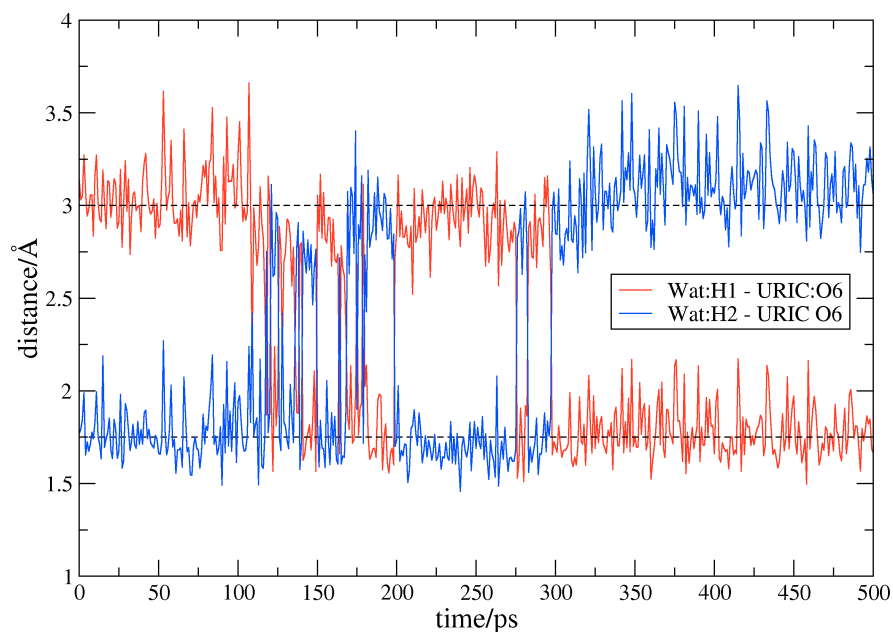
LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804  
SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874  
HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884  
ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916  
PRO917 GLN918 PHE1005 GLY1006 ILE1007 SER1008 PHE1009 THR1010  
VAL1011 PRO1012 LEU1014 ASN1015 GLN1016 GLY1034 GLY1035 THR1036  
MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
THR1135 ALA1149 PHE1150 GLN1194 ALA1198 GLN1201 ALA1258 VAL1259  
GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 MOC2003 URIC2 GOL2

crystal water

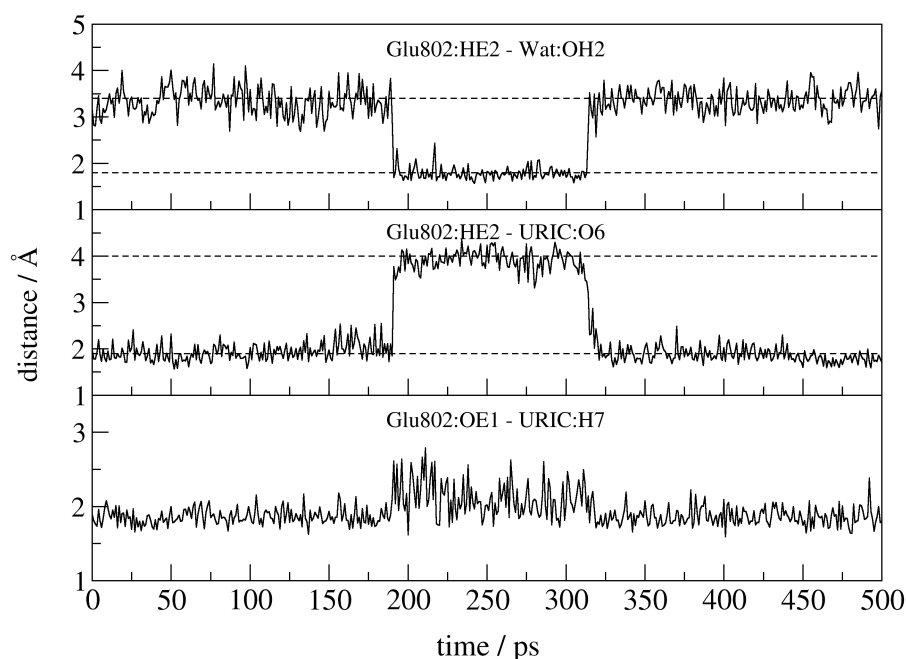
CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW96 CRYW157  
CRYW168 CRYW174 CRYW189 CRYW191 CRYW211 CRYW222 CRYW224 CRYW248  
CRYW343 CRYW369 CRYW373 CRYW405 CRYW447 CRYW498 CRYW571 CRYW573  
CRYW575 CRYW685 CRYW706 CRYW1034

solvent water

SOLV43 SOLV86 SOLV318 SOLV319 SOLV686 SOLV763 SOLV857 SOLV1061  
SOLV1077 SOLV1331 SOLV1641 SOLV1753 SOLV1991



**Figure S8.** Distance plot for the water molecule coordinating towards O6 of xanthine. The water molecule rotates and thereby changes the coordination mode from H2 to H1 coordinating toward URIC:O6.



**Figure S9.** Distance plot for the coordination mode of Glu802: While the coordination of OE1 towards H7 stays intact, we find two different features for the Glu802:HE2 - Uric:O6 interaction: one direct interaction (1- 180 ps and 350 - 500 ps) and a water bridged interaction (180 - 350 ps). We investigated the first one in snapshots SN100 and SN400, and the latter one in SN200.

# SETUP\_G, SN100

**Table S16.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 100 of SETUP G in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	-0.7	1.6	0.9	3.1	1.5	4.5	Proton transfer 1
IM1'	-2.8	-3.1	-5.9	-1.9	-3.7	-5.6	
TS1''	-1.9	-3.5	-5.4	-1.3	-4.0	-5.3	OH-Rotation
IM1''	-5.4	-5.1	-10.5	-4.3	-5.6	-9.8	
TS1	1.0	-0.4	0.6	0.8	-0.9	-0.1	Proton transfer 2
IM1	0.1	-0.4	-0.3	0.2	-0.4	-0.2	
TS2	1.0	-0.4	0.6	8.5	-5.1	3.4	C-O formation
IM2	2.3	-7.0	-4.7	6.9	-8.1	-1.2	
TS3	8.6	-5.5	3.1	12.8	-8.0	4.8	Hydride transfer
Product	4.7	-2.8	1.9	-5.2	-2.0	-7.2	

<sup>a</sup> absolute energies (a.u.): -1500.367965 -89.818817 -1590.186781 (BP86)  
-1499.232671 -89.815870 -1589.048541 (B3LYP)

## **active MM region SETUP G, SN100 (1550 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY796 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803  
ARG804 SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873  
SER874 HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883  
HSD884 ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915  
GLY916 PRO917 GLN918 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
THR1010 VAL1011 LEU1014 ASN1015 GLN1016 GLY1035 THR1036 MET1038  
GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078 ALA1079  
SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086 ALA1149  
PHE1150 TYR1152 THR1154 GLN1194 ALA1198 GLN1201 ALA1258 VAL1259  
GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003 URIC2 GOL2

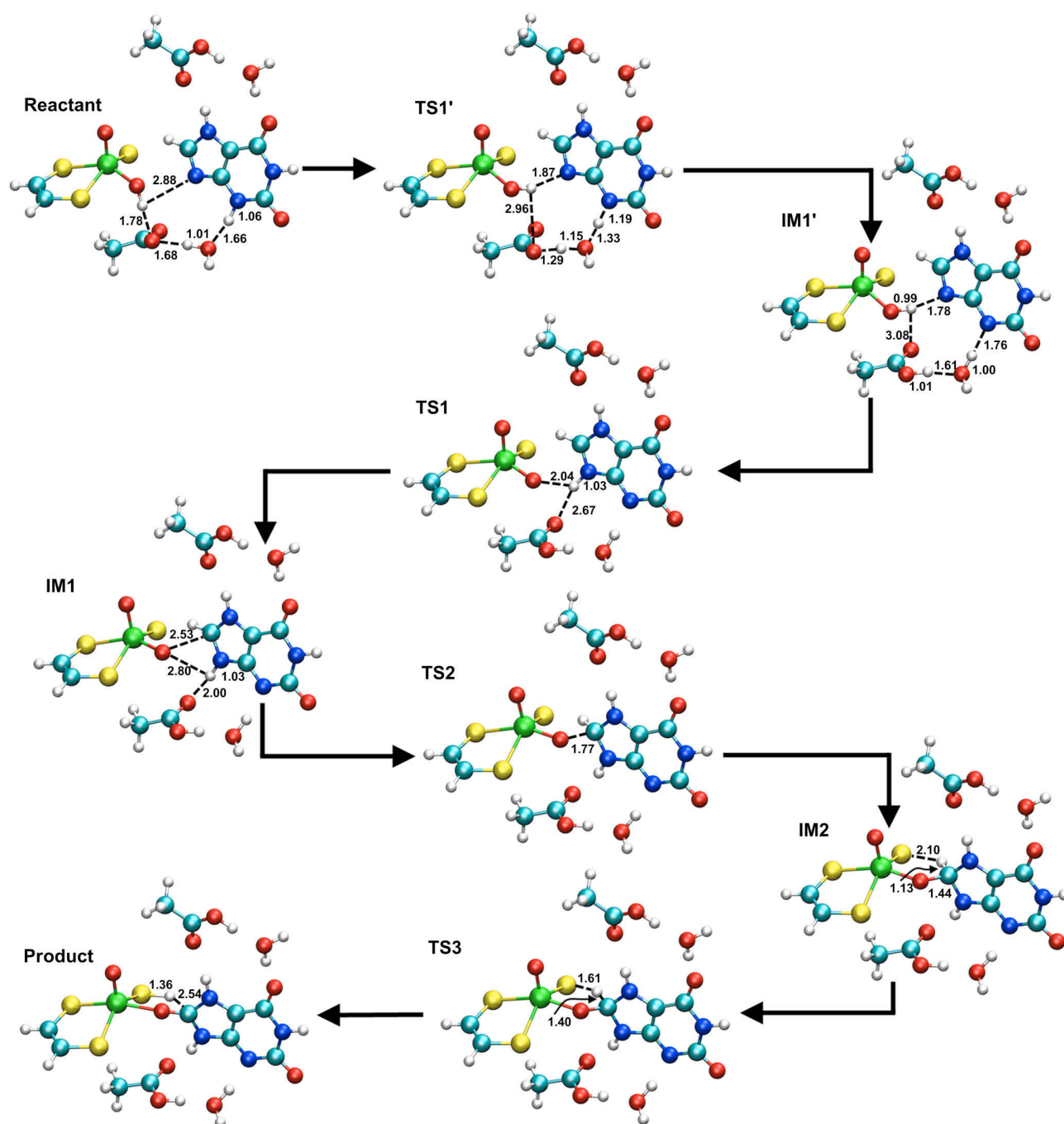
crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW114  
CRYW126 CRYW168 CRYW174 CRYW189 CRYW191 CRYW211 CRYW222 CRYW224  
CRYW248 CRYW343 CRYW369 CRYW382 CRYW447 CRYW498 CRYW571 CRYW575  
CRYW632 CRYW706 CRYW847 CRYW1034 CRYW1205

solvent water

SOLV421 SOLV455 SOLV491 SOLV546 SOLV573 SOLV751 SOLV857 SOLV1321  
SOLV1572 SOLV1641 SOLV1753 SOLV1803

# SETUP\_G, SN200



**Figure S10.** Reaction mechanism for setup G, SN200 from B3LYP/MM calculations. The distances are given in Angstrom.

**Table S17.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 200 of SETUP G in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	3.9	0.3	4.2	5.5	-0.4	5.1	Proton transfer 1
IM1'	-5.2	0.6	-4.6	-5.1	0.3	-4.8	
TS1	-1.3	4.5	3.1	-0.7	3.9	3.3	Proton transfer 2
IM1	-1.3	3.9	2.6	-0.8	4.0	3.2	
TS2	8.5	-0.7	7.8	13.4	-0.9	12.5	C-O formation
IM2	6.8	-2.5	4.4	12.5	-2.8	9.7	
TS3	11.1	-1.1	10.1	14.2	-0.7	13.5	Hydride transfer
Product	1.7	3.2	4.9	-8.0	3.9	-4.1	

<sup>a</sup>absolute energies (a.u.): -1500.354732 -43.210292 -1543.565024 (BP86)  
-1499.218480 -43.208278 -1542.426758 (B3LYP)

**active MM region SETUP G, SN200 (1588 active atoms)**

chain A

LEU648 PHE649 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771  
MET794 GLY796 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803  
ARG804 SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873  
SER874 HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883  
HSD884 ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915  
GLY916 PRO917 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008  
PHE1009 THR1010 VAL1011 PHE1013 LEU1014 ASN1015 GLN1016 GLY1035  
THR1036 MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077  
ALA1078 ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085  
TYR1086 PHE1150 TYR1152 THR1154 GLN1194 ALA1198 GLN1201 ALA1258  
VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265 MOC2003

URIC2 GOL2

crystal water

CRYW189 CRYW191 CRYW211 CRYW222 CRYW224 CRYW244 CRYW248 CRYW343  
CRYW369 CRYW382 CRYW447 CRYW498 CRYW571 CRYW575 CRYW632 CRYW706  
CRYW847 CRYW1034 CRYW1205

solvent water

SOLV138 SOLV421 SOLV455 SOLV491 SOLV546 SOLV573 SOLV857 SOLV893  
SOLV1272 SOLV1321 SOLV1572 SOLV1641 SOLV1752 SOLV1753

# SETUP\_G, SN400

**Table S18.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP G in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	2.8	0.1	2.9	9.4	-0.9	8.5	Proton transfer 1
IM1'	-7.0	1.0	-8.0	-6.0	-1.5	-7.5	
TS1''	-6.7	-0.7	-7.4	-5.3	-1.9	-7.2	OH-Rotation
IM1''	-7.2	-1.7	-9.0	-6.2	-2.4	-8.6	
TS1	0.0	1.0	1.0	3.1	-0.6	2.5	Proton transfer 2
IM1	-13.5	10.1	-3.4	-10.8	9.0	-1.8	
TS2	-6.7	4.5	-2.1	-0.2	2.6	2.5	C-O formation
IM2	-1.5	-6.3	-7.8	-2.2	-0.2	-2.5	
TS3	-1.9	2.6	0.7	3.2	1.3	4.5	Hydride transfer
Product	-8.6	7.9	-0.8	-17.9	8.7	-9.3	

<sup>a</sup> absolute energies (a.u.): -1500.413756 -49.980759 -1550.394515 (BP86)  
-1499.281279 -49.976648 -1549.257927 (B3LYP)

## **active MM region SETUP G, SN400 (1595 active atoms)**

chain A

LEU648 HSE741 TYR743 THR766 GLN767 ASN768 ALA769 LYS771 MET794  
GLY796 GLY797 PHE798 GLY799 GLY800 LYS801 GLU802 THR803 ARG804  
SER805 GLY837 GLY838 ARG839 HSD840 PHE842 ASP872 LEU873 SER874  
HSD875 SER876 ILE877 MET878 GLU879 ARG880 ALA881 PHE883 HSD884  
ASN908 THR909 ALA910 PHE911 ARG912 GLY913 PHE914 GLY915 GLY916  
PRO917 GLN918 LEU920 PHE1005 GLY1006 ILE1007 SER1008 PHE1009  
THR1010 VAL1011 PHE1013 LEU1014 ASN1015 GLN1016 GLY1035 THR1036  
MET1038 GLN1040 LEU1042 SER1074 SER1075 PRO1076 THR1077 ALA1078  
ALA1079 SER1080 VAL1081 SER1082 THR1083 ASP1084 ILE1085 TYR1086  
ALA1149 PHE1150 HSD1151 TYR1152 THR1154 GLN1194 ALA1198 GLN1201  
ALA1258 VAL1259 GLY1260 GLU1261 PRO1262 PRO1263 LEU1264 PHE1265  
MOC2003 URIC2 GOL2

crystal water

CRYW19 CRYW40 CRYW42 CRYW45 CRYW48 CRYW56 CRYW58 CRYW74 CRYW114  
CRYW168 CRYW174 CRYW189 CRYW191 CRYW211 CRYW213 CRYW222 CRYW224  
CRYW244 CRYW248 CRYW343 CRYW369 CRYW382 CRYW429 CRYW447 CRYW498  
CRYW571 CRYW575 CRYW632 CRYW706 CRYW847 CRYW1034

solvent water

SOLV163 SOLV228 SOLV419 SOLV421 SOLV573 SOLV574 SOLV857 SOLV893  
SOLV1276 SOLV1321 SOLV1641 SOLV1752 SOLV1753 SOLV1991

**Table S19.** Calculated QM/MM energies of the rate limiting stationary points for snapshot 400 of SETUP G in kcal mol<sup>-1</sup> using the L2 link atom scheme, from single point calculations (SP) on the structures obtained with the charge-shift link atom scheme and for the reoptimized structures (OPT).

	B3LYP/Lanl2DZ (SP)			B3LYP/Lanl2DZ (OPT)			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
IM1" <sup>a,b</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS3	9.9	3.8	13.6	10.1	3.4	13.5	Hydride transfer

<sup>a</sup> IM1" and TS3 are the lowest and highest point on the energy profile, respectively. The corresponding energy differences for the charge-shift link atom scheme are 9.4, 3.2, and 13.1 kcal mol<sup>-1</sup> (Table S18).

<sup>b</sup> absolute energies (a.u.):      -1499.268497 -49.980544 -1549.249042 (single point)  
    -1499.268638 -49.981706 -1549.250344 (optimized)

**Table S20.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP G in kcal mol<sup>-1</sup> using BP86 as functional, values for B3LYP are given in the paper. The QM and MM contributions are given in Tables S21-S23, see below.

	BP86/Lanl2DZ		BP86/def2-SVP		BP86/def2-TZVP		BP86/TZVP-all electron	
Reactant	0.0		0.0		0.0		0.0	
TS1'	2.9	(2.9)	2.4	(2.4)	3.1	(3.1)	2.5	(2.5)
IM1'	-8.0		-7.2		-7.9		-8.5	
TS1"	-7.4	(0.6)	-7.0	(0.2)	-7.7	(0.2)	-8.0	(0.5)
IM1"	-9.0		-7.9		-8.7		-9.4	
TS1	1.0	(10.0)	1.5	(9.4)	1.3	(9.0)	2.9	(12.3)
IM1	-3.4		-4.4		-2.8		-1.6	
TS2	-2.1	(1.3)	-3.4	(1.0)	-0.9	(1.9)	-0.5	(1.0)
IM2	-7.8		-9.1		-5.7		-7.9	
TS3	0.7	(8.5)	-0.6	(8.5)	4.3	(10.0)	-1.3	(6.6)
Product	-0.8		-1.3		2.4		-5.7	
$\Delta E_{\max}$	10.0		9.4		13.0		12.3	

**Table S21.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP G in kcal mol<sup>-1</sup>. Energies for single point calculation with the SVPalls2 basis set for B3LYP are given in parentheses.

	BP86/def2-SVP			B3LYP/def2-SVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0	0.0	0.0	0.0(0.0)	0.0(0.0)	0.0(0.0)
TS1'	2.8	-0.4	2.4	5.6(4.7)	-0.4(-0.4)	5.3(4.3)
IM1'	-6.4	-0.8	-7.2	-8.9(-10.3)	-0.1(-0.1)	-9.0(-10.4)
TS1''	-6.1	-0.9	-7.0	-7.9(-8.9)	-0.4(-0.4)	-8.3(-9.3)
IM1''	-5.8	-2.1	-7.9	-8.1(-8.4)	-1.6(-1.6)	-9.7(9.9)
TS1	0.6	0.9	1.5	0.2(2.2)	0.6(0.6)	0.8(2.8)
IM1	-14.4	10.0	-4.4	-15.3(-12.2)	10.5(10.5)	-4.8(-1.7)
TS2	-8.3	4.9	-3.4	-5.4(-5.0)	4.2(4.2)	-1.2(0.8)
IM2	-3.2	-5.9	-9.1	-0.6(-3.1)	-6.3(6.3)	-6.8(9.4)
TS3	-3.5	2.9	-0.6	-1.3(-7.6)	2.2(2.2)	0.9(-5.3)
Product	-9.4	8.1	-1.3	-20.5(-30.7)	10.3(10.3)	-10.3(-20.4)



**Table S22.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP G in kcal mol<sup>-1</sup>. Energies for single point calculation with the TZVPalls2 basis set for B3LYP are given in parentheses.

	BP86/def2-TZVP			B3LYP/def2-TZVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0	0.0	0.0	0.0(0.0)	0.0(0.0)	0.0(0.0)
TS1'	4.2	-1.1	3.1	8.9(9.1)	-1.2(-1.2)	7.7(8.0)
IM1'	-6.1	-1.8	-7.9	-7.4(-7.2)	-1.5(-1.5)	-8.9(-8.7)
TS1''	-6.0	-1.7	-7.7	-6.5(-6.2)	-1.8(-1.8)	-8.2(-8.0)
IM1''	-5.9	-2.8	-8.7	-6.7(-7.0)	-2.5(-2.5)	-9.2(-9.5)
TS1	1.6	-0.4	1.3	3.0(5.4)	-0.7(-0.7)	2.4(4.7)
IM1	-12.1	9.3	-2.8	-10.6(-8.6)	9.4(9.4)	-1.3(0.8)
TS2	-4.7	3.8	-0.9	0.3(0.7)	2.8(2.8)	3.1(3.6)
IM2	1.4	-7.0	-5.7	5.9(4.1)	-7.3(-7.3)	-1.3(-3.1)
TS3	2.1	2.2	4.3	6.5(1.2)	1.3(1.3)	7.7(2.5)
Product	-4.8	7.3	2.4	-13.5(-23.2)	8.5(8.5)	-5.0(-14.7)

**Table S23.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of SETUP G in kcal mol<sup>-1</sup>.

	BP86/all-electron-TZVP			B3LYP/all-electron-TZVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	3.2	-0.7	2.5	9.1	-1.1	8.0
IM1'	-7.1	-1.5	-8.5	-7.0	-1.7	-8.7
TS1''	-6.8	-1.2	-8.0	-6.5	-1.5	-8.0
IM1''	-6.6	-2.8	-9.4	-6.7	-2.8	-9.5
TS1	2.8	0.0	2.9	5.3	-0.6	4.7
IM1	-11.0	9.4	-1.6	-8.6	9.3	0.8
TS2	-5.3	4.9	-0.5	0.8	3.0	3.8
IM2	-1.0	-6.9	-7.9	4.2	-7.5	-3.3
TS3	-3.6	2.3	-1.3	1.6	1.2	2.7
Product	-13.7	8.0	-5.7	-23.3	8.7	-14.7

**Table S24.** Results for the complete enzyme (setup G): QM/MM energies in kcal mol<sup>-1</sup> for different basis sets, relative to the energy of the reactant. Activation barriers relative to the preceding minima are given in parentheses. BP86 results.

	ALL- BP86/Lan12 DZ	ALL- BP86/def2- SVP	ALL- BP86/def2- SVP- all_electron (SP)	ALL- BP86/def2- TZVP	ALL- BP86/def2- TZVP- all_electron (SP)	ALL- BP86/TZVP- all_electron
Reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	3.5 (3.5)	1.8 (1.8)	1.6 (1.6)	2.6 (2.6)	2.8 (2.8)	
IM1'	-7.9	-6.9	-7.2	-7.9	-8.1	
TS1''	-7.8 (0.1)	-6.4 (0.5)	-6.3 (0.9)	-7.4 (0.5)	-7.7 (0.4)	
IM1''	-9.0	-8.1	-8.1	-8.9	-9.5	-9.5
TS1	1.0 (10.0)	0.9 (9.0)	2.9 (11.0)	1.1 (10.0)	2.5 (12.0)	2.6 (12.1)
IM1	-2.7	-4.0	-1.2	-3.1	-0.6	-0.6
TS2	-2.1 (0.6)	-3.7 (0.3)	-2.3 (-1.1)	-0.7 (2.4)	-0.5 (0.1)	-0.1 (0.5)
IM2	-8.1	-9.3	-11.2	-5.7	-8.1	-8.2
TS3	0.4 (8.5)	-0.9 (8.6)	-5.4 (5.8)	3.5 (9.2)	-1.7 (6.4)	-1.4 (6.8)
Product	-1.2	-1.1	-7.5	1.3	-5.5	-5.5
$\Delta E_{\max}$	10.0	9.0	11.0	12.4	12.0	12.1

**Table S25.** Results for the complete enzyme (setup G): QM/MM energies in kcal mol<sup>-1</sup> for different basis sets, relative to the energy of the reactant. Activation barriers relative to the preceding minima are given in parentheses. B3LYP results.

	ALL- B3LYP/Lanl 2DZ	ALL- B3LYP/def2 -SVP	ALL- B3LYP/def2- SVP- all_electron (SP)	ALL- B3LYP/def2 -TZVP	ALL- B3LYP/def2- TZVP- all_electron (SP)	ALL- B3LYP/def2- TZVP- all_electron
Reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	7.7 (7.7)	4.7 (4.7)	4.5 (4.5)	8.1 (8.1)	7.9 (7.9)	
IM1'	-8.4	-9.4	-9.2	-8.1	-7.7	
TS1''	-7.6 (0.8)	-8.4 (1.0)	-8.3 (0.9)	-7.4 (0.7)	-7.6 (0.1)	
IM1''	-9.2	-9.5	-9.0	-8.9	-9.2	-9.1
TS1	1.6 (10.8)	0.2 (9.7)	2.9 (11.9)	0.9 (9.8)	3.5 (12.7)	3.6 (12.7)
IM1	-2.2	-3.8	-1.0	-1.4	1.4	1.5
TS2	1.4 (3.6)	-1.3 (2.5)	0.0 (1.0)	3.7 (5.1)	3.6 (2.2)	3.9 (2.4)
IM2	-3.6	-7.0	-8.6	-1.2	-3.2	-3.2
TS3	2.9 (6.5)	0.4 (7.4)	-4.7 (3.9)	7.8 (9.0)	2.6 (5.8)	2.9 (6.2)
Product	-10.8	-10.6	-20.3	-5.8	-14.7	-14.5
$\Delta E_{\text{max}}$	12.1	9.9	11.9	16.7	12.7	13.0

**Table S26.** Results for the complete enzyme (setup G): QM, MM and QM/MM energies in kcal mol<sup>-1</sup> for the Lanl2DZ basis set, relative to the energy of the reactant.

	ALL-BP86/LANL2DZ			ALL-B3LYP/LANL2DZ		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0	0.0	0.0	0.0	0.0	0.0
TS1'	3.7	-0.2	3.5	9.6	-1.9	7.7
IM1'	-6.8	-1.1	-7.9	-5.6	-2.7	-8.4
TS1''	-7.0	-0.8	-7.8	-4.6	-2.9	-7.6
IM1''	-6.8	-2.2	-9.0	-5.4	-3.9	-9.2
TS1	0.0	1.0	1.0	2.6	-1.0	1.6
IM1	-12.4	9.7	-2.7	-10.0	7.8	-2.2
TS2	-6.9	4.8	-2.1	0.0	1.5	1.4
IM2	-1.9	-6.2	-8.1	-1.4	-2.3	-3.6
TS3	-2.1	2.4	0.4	3.2	-0.3	2.9
Product	-8.8	7.5	-1.2	-18.0	7.2	-10.8

**Table S27.** Results for the complete enzyme (setup G): QM, MM and QM/MM energies in kcal mol<sup>-1</sup> for the def2-SVP basis set, relative to the energy of the reactant. The values in parentheses are from single point calculations with the all-electron-SVP basis set.

	ALL-BP86/def2-SVP			ALL-B3LYP/def2-SVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
TS1'	2.3(2.1)	-0.5(-0.5)	1.8(1.6)	5.3(5.1)	-0.7(-0.7)	4.7(4.5)
IM1'	-6.0(-6.3)	-0.9(-0.9)	-6.9(-7.2)	-8.7(-8.6)	-0.6(-0.6)	-9.4(-9.2)
TS1''	-5.4(-5.3)	-1.0(-1.0)	-6.4(-6.3)	-7.6(-7.5)	-0.8(-0.8)	-8.4(-8.3)
IM1''	-4.7(-4.6)	-3.5(-3.5)	-8.1(-8.1)	-7.2(-6.8)	-2.2(-2.2)	-9.5(-9.0)
TS1	-0.3(1.8)	1.1(1.1)	0.9(2.9)	-0.4(2.3)	0.6(0.6)	0.2(2.9)
IM1	-13.9(-11.1)	9.9(9.9)	-4.0(-1.2)	-13.7(10.9)	9.9(9.9)	-3.8(-1.0)
TS2	-9.0(-7.6)	5.3(5.3)	-3.7(-2.3)	-5.0(-3.7)	3.7(3.7)	-1.3(0.0)
IM2	-3.3(-5.2)	-6.0(-6.0)	-9.3(-11.2)	-0.3(-1.9)	-6.7(-6.7)	-7.0(-8.6)
TS3	-3.2(-7.6)	2.3(2.3)	-0.9(-5.4)	-1.2(-6.3)	1.5(1.5)	0.4(-4.7)
Product	-9.0(-15.3)	7.9(7.9)	-1.1(7.5)	-20.2(-29.8)	9.6(9.6)	-10.6(-20.3)

**Table S28.** Results for the complete enzyme (setup G): QM, MM and QM/MM energies in kcal mol<sup>-1</sup> for the def2-TZVP basis set, relative to the energy of the reactant. The values in parentheses are from single point calculations with the all-electron-SVP basis set.

	ALL-BP86/def2-TZVP			ALL-B3LYP/def2-TZVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
TS1'	4.0(4.1)	-1.4(-1.4)	2.6(2.8)	10.2(10.0)	-2.0(-2.0)	8.1(7.9)
IM1'	-6.2(-6.4)	-1.7(-1.7)	-7.9(-8.1)	-5.8(-5.3)	-2.3(-2.3)	-8.1(-7.7)
TS1''	-5.7(-6.1)	-1.6(-1.6)	-7.4(-7.7)	-5.0(-5.1)	-2.5(-2.5)	-7.4(-7.6)
IM1''	-5.8(-6.5)	-3.0(-3.0)	-8.9(-9.5)	-5.4(-5.7)	-3.5(-3.5)	-8.9(-9.2)
TS1	0.5(1.9)	0.6(0.6)	1.1(2.5)	1.5(4.1)	-0.6(-0.6)	0.9(3.5)
IM1	-12.3(-9.7)	9.2(9.2)	-3.1(-0.6)	-10.1(-7.3)	8.7(8.7)	-1.4(1.4)
TS2	-4.3(-4.0)	3.5(3.5)	-0.7(-0.5)	1.7(1.7)	1.9(1.9)	3.7(3.6)
IM2	1.1(-1.2)	1.9(1.9)	-5.7(-8.1)	6.8(4.9)	-8.1(-8.1)	-1.2(-3.2)
TS3	1.6(-3.5)	1.9(1.9)	3.5(-1.7)	7.6(2.4)	0.2(0.2)	7.8(2.6)
Product	-5.7(-12.5)	7.0(7.0)	1.3(-5.5)	-13.7(-22.6)	7.9(7.9)	-5.8(-14.7)

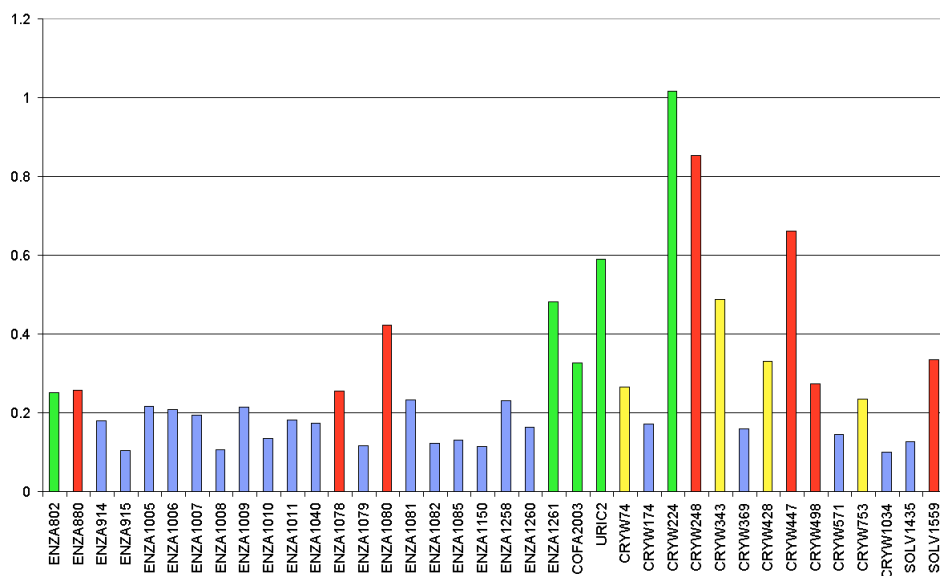
**Table S29.** Results for the complete enzyme (setup G): Calculated relative QM, MM and QM/MM energies of a subset of optimized stationary points for Snapshot 400 in kcal mol<sup>-1</sup>.

	ALL-BP86/all-electron-TZVP			ALL-B3LYP/all-electron-TZVP		
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>
Reactant	0.0	0.0	0.0	0.0	0.0	0.0
IM1''	-6.3	-3.2	-9.5	-4.5	-4.7	-9.1
TS1	2.0	0.6	2.6	5.2	-1.6	3.6
IM1	-9.6	9.0	-0.6	-6.2	7.7	1.5
TS2	-4.5	4.5	-0.1	2.7	1.2	3.9
IM2	-1.1	-7.0	-8.2	6.0	-9.3	-3.2
TS3	-3.0	1.6	-1.4	4.0	-1.2	2.9
Product	-13.1	7.6	-5.5	-21.5	7.0	-14.5

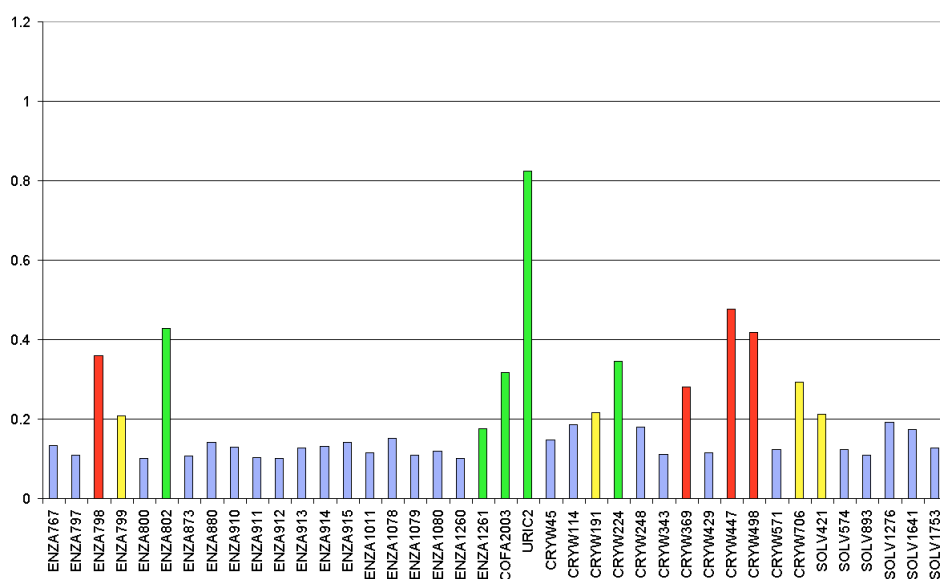
### 3. Consistency checks

Meaningful energy profiles can be obtained only if the calculated stationary points are connected by contiguous reaction paths. For example, there must not be any distant conformational changes such as side chain flips and water reorientations that are unrelated to the reaction under investigation. In our QM/MM work, these requirements are routinely checked by visual inspection. Here we report some additional quantitative checks that involve the RMSD (root-mean square deviation) between the optimized QM/MM structures of the reactant complex and the rate-limiting transition state TS3. For a given setup and snapshot, these structures are easily superimposed because they share the same fixed outer region. Table S30 presents the computed RMSD values for all snapshots considered, in particular those for the active optimized region (typically around 1600 atoms) and for the active region without the QM atoms. The latter RMSD values are most relevant for assessing the structural consistency of the protein environment during the reaction. They are generally found to be about 0.1 Å in all cases which indicates a high degree of overall structural consistency. A residue-specific analysis was performed for setup A (SN500) and setup G (SN400). Figures S11 and S12 show the RMSD values of individual residues that exceed 0.1 Å, respectively. The main structural changes trivially occur in the QM region (colored in green) where the reaction takes place. The side chains of the MM residues next to the QM region (colored in red) also show some appreciable changes because they have to adapt to the reactive events in the QM region. The same holds to a lesser extent for the next shell of residues (colored in yellow). In residues farther away from the QM region (colored in blue), the changes are getting less and less pronounced (RMSD values mostly below 0.2 Å). It should be noted that Figure S11 (S12) shows the RMSD data for only 38 (40) residues out of a total of 135 (141) residues in the active region (see pages S9 and S30) since the RMSD values for the remaining residues are below the threshold of 0.1 Å. The plots in Figures S11 and S12 confirm the absence of structural inconsistencies in distant residues.





**Figure S11.** Residue-specific RMSD values (Å) in Setup A, SN500. Shown are only residues with RMSD values above 0.1 Å. Color code: green, QM region; red, neighboring first-shell residues; yellow, second-shell residues; blue, other more distant residues.



**Figure S12.** Residue-specific RMSD values (Å) in Setup G, SN400. Shown are only residues with RMSD values above 0.1 Å. Color code see Figure S11.

**Table S30.** RMSD values (in Å) between the optimized reactant complex<sup>a</sup> and transition state TS3 for all snapshots considered (see text).

	complete system	active region	active region without QM atoms	QM atoms
Setup A (SN100)	0.042	0.156	0.120	0.639
Setup A (SN500)	0.042	0.157	0.128	0.588
Setup B (SN300)	0.046	0.167	0.098	0.846
Setup B (SN500)	0.041	0.150	0.088	0.757
Setup C (SN200)	0.032	0.123	0.065	0.636
Setup C (SN400)	0.044	0.159	0.097	0.774
Setup D (SN200)	0.036	0.139	0.065	0.751
Setup D (SN400)	0.046	0.178	0.089	0.940
Setup E (SN100)	0.037	0.145	0.094	0.661
Setup E (SN300)	0.040	0.156	0.118	0.635
Setup E (SN500)	0.033	0.125	0.073	0.637
Setup F (SN200)	0.036	0.136	0.084	0.651
Setup F (SN400)	0.039	0.149	0.098	0.673
Setup G (SN100)	0.034	0.131	0.072	0.661
Setup G (SN200)	0.033	0.127	0.073	0.635
Setup G (SN400) <sup>b</sup>	0.035	0.134	0.087	0.621

<sup>a</sup> The reactant complex is defined as the lowest lying minimum preceding TS3, i.e. IM1' for setup E, IM1 for setup F, IM1" for setup G and the starting structure "Reactant" for all other setups.

<sup>b</sup> For Setup G (SN400), we reoptimized the two appropriate structures using the same QM region as in the gas-phase QM study, see Section III.J and Figure 9 of the paper. The RMSD value from the QM/MM study is 0.741 Å (for the QM atoms including the link atoms), much lower than the RMSD value from the gas-phase QM study which is 1.876 Å after structural alignment.

## 4. Setup Comparisons

**Table S31.** Energy differences for snapshots where the QM region has been replaced. All values refer to IM2 structures and are given in kcal mol<sup>-1</sup> using the B3LYP functional.

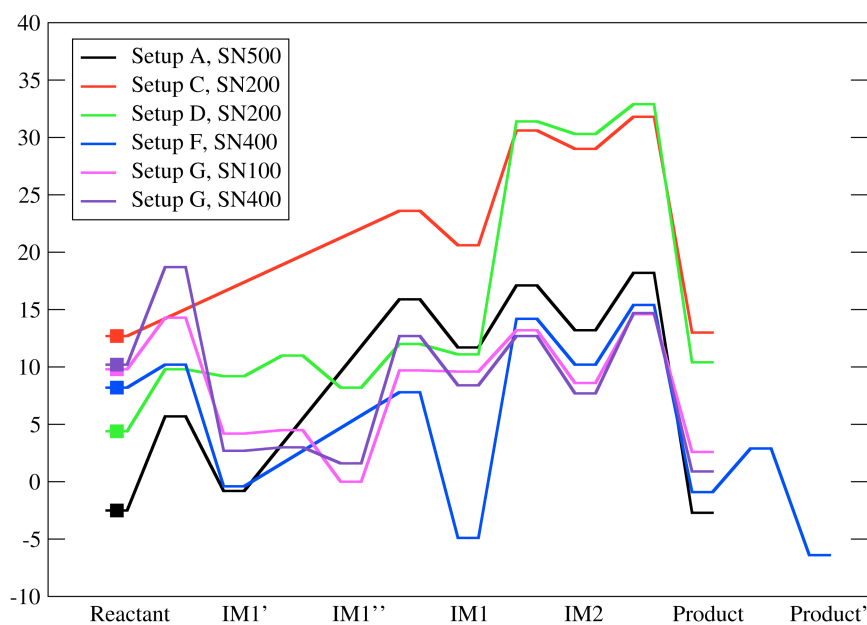
	$\Delta E_{\text{QM}}(\text{LANL})^{\text{a}}$	$\Delta E_{\text{QM/MM}}(\text{LANL})$	$\Delta E_{\text{QM}}(\text{TZVP})^{\text{a}}$	$\Delta E_{\text{QM/MM}}(\text{TZVP})$
G2A	9.6	18.7	9.3	18.4
G2C	-	-	-	-
G2D	-	-	-	-
G2F	9.4	12.6	9.0	12.2
A2G	-5.7	-1.0	-5.8	-1.1
C2G	-5.9	0.7	-5.6	0.8
D2G	4.9	1.6	5.5	2.1
F2G	5.1	6.1	4.5	5.5

<sup>a</sup> including the QM-MM interaction

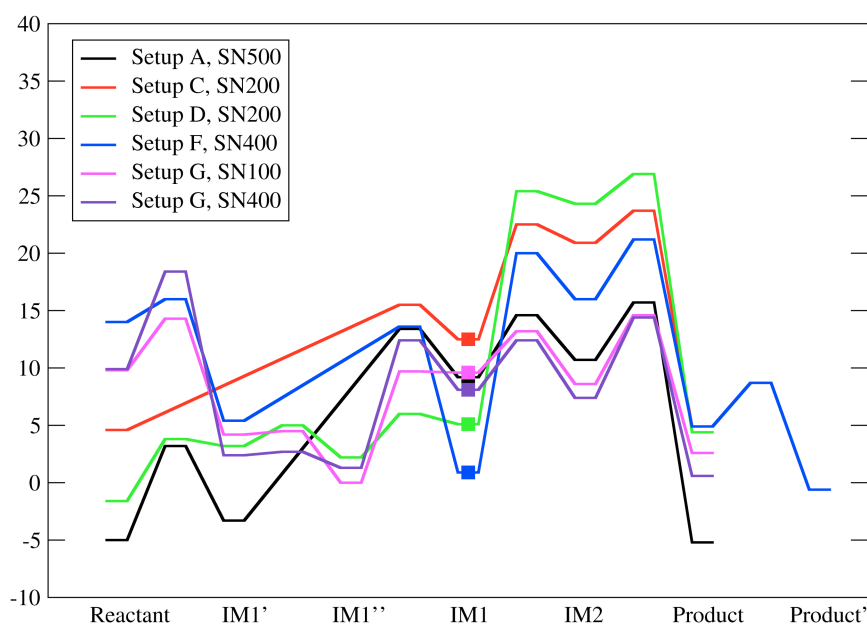
### Conventions used in the following Figures S13-S16:

The lowest point on the energy profile of setup G, SN100 is assigned an energy of zero. The other energy profiles are shifted such that the relative energy of one selected species is equal to the value obtained from COSMO ( $\epsilon=8$ ) single-point calculations on the extended 66-atom QM region (see Table 10 of the paper for the corresponding numerical data). Selected species are the reactant in Figure S13, the intermediate complex IM1 in Figure S14, the tetrahedral intermediate in Figure S15 and the product in Figure S16. Their relative energies are represented by squares in Figures S13-S16.

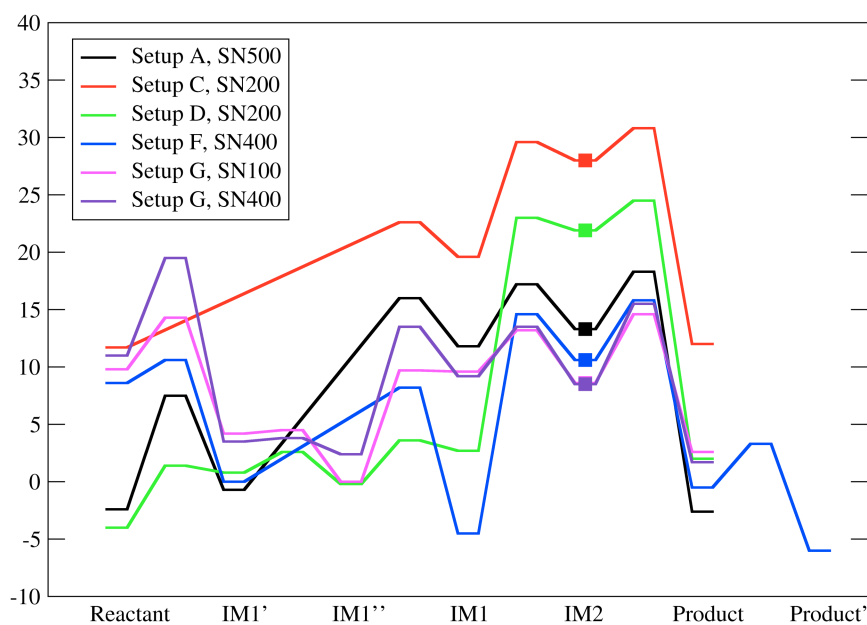
The resulting four sets of energy profiles are quite similar. A representative overall set can be obtained by averaging each of the energy values in the four individual sets, which is feasible since they all refer to the same energy scale (having the same zero energy, see above). The resulting set of energy profiles is shown in Figure 8 of the paper.



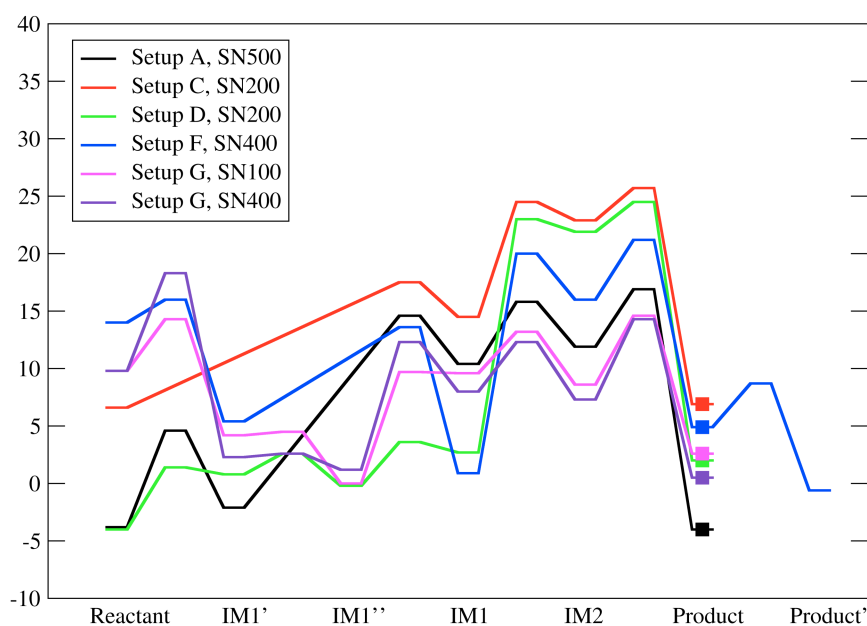
**Figure S13.** Reaction profiles scaled on the relative energies of the **reactant**, represented by the squares (B3LYP-COSMO,  $\epsilon=8$ , see preceding text and Section III.I of the paper).



**Figure S14.** Reaction profiles scaled on the relative energies of **IM1**, represented by the squares (B3LYP-COSMO,  $\epsilon=8$ , see preceding text and Section III.I of the paper).



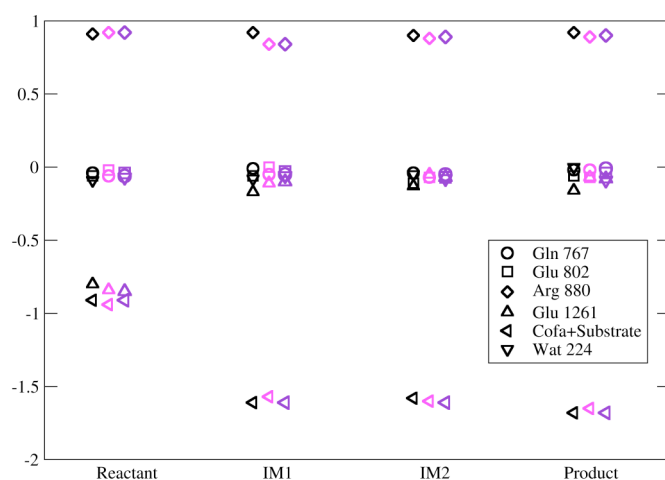
**Figure S15.** Reaction profiles scaled on the relative energies of **IM2**, represented by the squares (B3LYP-COSMO,  $\epsilon=8$ , see preceding text and Section III.I of the paper).



**Figure S16.** Reaction profiles scaled on the relative energies of the **product**, represented by the squares (B3LYP-COSMO,  $\epsilon=8$ , see preceding text and Section III.I of the paper).

## ESP charges and key electrostatic interactions

As discussed in the paper (Section III.I), we have analyzed the charge distributions obtained in the single-point B3LYP calculations for the extended QM region (Figure 5 of the paper). ESP charges were determined for four species (reactant, IM1, IM2, product) in three snapshots of setup A (SN500) and setup G (SN100 and SN400). Figure S17 shows the overall ESP charges of the residues included in the QM region. It is obvious that these charges are quite similar for the different snapshots and often also for the different species, except for the change between the reactant and IM1 that is due to the proton transfer from the cofactor (setup A) or the substrate (setup G) to Glu1261.



**Figure S17:** ESP charges ( $e$ ) for the residues included in the extended QM region (see Figure 5 in the paper). Each triple represents the data for snapshots A-SN500, G-SN100, and G-SN-400, from left to right.

Using the computed ESP charges and the relevant distances from the QM/MM optimized structures, individual electrostatic interactions were calculated for key atom pairs in the reactant and in IM2. Table S32 lists the corresponding ESP charges and distances as well as the difference (IM2-reactant) of the computed Coulomb interaction energies between the two point charges (a positive value indicates that the interaction is stronger in the reactant than in IM2).

**Table S32.** Selected ESP charges ( $e$ ), distances ( $\text{\AA}$ ), and differences  $\Delta E$  ( $\text{kcal mol}^{-1}$ ) of the corresponding Coulomb interaction energies in IM2 and the reactant (IM2-reactant, see text). The notation conforms to Scheme 6 of the paper.

interaction	Reactant			IM2			
	charge 1	charge 2	distance	charge 1	charge 2	distance	$\Delta E$
<b>Setup A, SN500</b>							
H <sub>Glu</sub> -N <sub>acc</sub> <sup>a)</sup>	0.45	-0.67	1.80	0.43	-0.61	1.69	4.1
H <sub>Subs</sub> -O <sub>acc</sub> <sup>b)</sup>	0.51	-0.71	1.65	0.43	-0.68	1.85	20.4
H:Arg880-O6	0.51	-0.74	1.74	0.52	-0.84	1.60	-18.6
H:Arg880-H7	0.51	0.67	3.14	0.52	0.50	3.46	-11.2
<b>Setup G, SN100</b>							
H7-O:Glu802	0.58	-0.63	1.72	0.55	-0.67	2.06	11.1
H:Glu802-O6	0.50	-0.58	1.80	0.52	-0.47	1.61	3.1
H:Arg880-O2	0.46	-0.68	1.90	0.50	-0.77	1.70	-20.5
H:Arg880-H3	0.46	0.40	2.59	0.50	0.36	2.78	2.1
H:Arg880-N3	0.46	-0.40	3.01	0.50	-0.74	2.64	-26.2
<b>Setup G, SN400</b>							
H7-O:Glu802	0.64	-0.67	1.72	0.55	-0.68	1.96	19.4
H:Glu802-O6	0.47	-0.60	1.83	0.50	-0.66	1.66	-14.8
H:Arg880-O2	0.48	-0.70	1.90	0.51	-0.78	1.70	-19.0
H:Arg880-H3	0.48	0.50	2.47	0.51	0.37	2.79	-9.8
H:Arg880-N3	0.48	-0.66	2.90	0.51	-0.73	2.63	-10.7

<sup>a)</sup> Reactant: H9-N9, IM2:H3-N3

<sup>b)</sup> Reactant: H7-O:Glu802, IM2: H3-O:Glu802

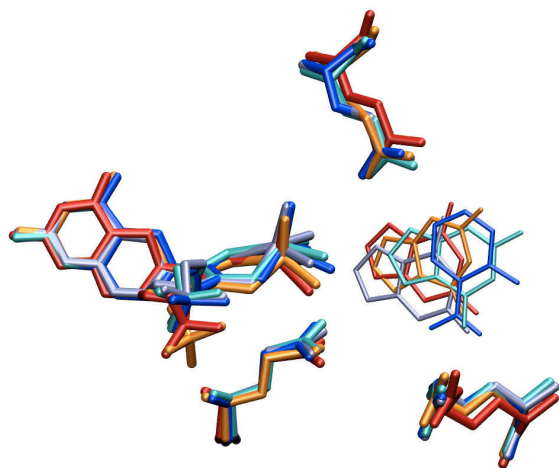
The data in Table S32 address the key interactions between the substrate and the active-site residues Arg880 and Glu802, both for the "upside" (setup A) and "upside down" (setup G) orientation of the substrate. For setup A, the binding mode of Glu802 changes from the reactant to IM2, whereas the interaction of Arg880 with the substrate remains similar. By contrast, for setup G, the binding mode of Glu802 remains the same, whereas the interaction of Arg880 changes due to the proton transfer that leads to a deprotonated N3 center in IM2 (see Scheme 6 in the paper).

We find that Arg880 provides significant stabilization when going from the reactant to IM2 in all three snapshots (see Table S32): In setup A, this stabilization of about 30 kcal mol<sup>-1</sup> is caused by a higher negative charge at and a shorter distance to O6 as well as less repulsion with H7. In setup G, this stabilization is even larger (ca. 45 kcal mol<sup>-1</sup> for SN100 and 40 kcal mol<sup>-1</sup> for SN400) since the substrate-Arg880 interactions in IM2 benefit from shorter distances and higher negative charges of O2 and N3. Additionally, the repulsive interaction of Arg880 with H3, which is part of a water molecule in IM2, is reduced in SN400.

For Glu802, there is again a common trend for all three snapshots: When going from the reactant state to IM2, the substrate-Glu802 interactions become less stabilizing. This can be attributed to lower charges on the interacting atoms and/or increased distances. There is one exception to this trend, namely setup G SN400, where the interaction H:Glu802-O6 becomes stronger which is however overcompensated by the weakening of H7-O:Glu802.

The data in Table S32 also offer an explanation why the "upside" orientation is better stabilized in the reactant state than the "upside down" orientation: In setup A we find the interaction of Arg880 with O6 to be about 13-19 kcal mol<sup>-1</sup> stronger than the interaction of Arg880 with O2 in setup G. Hence, the substrate-Arg880 interaction (with O6) favors the "upside" orientation in the reactant state. On the other hand, the substrate-Arg880 interaction (primarily with O2 and N3) is also the origin of the stronger stabilization of the "upside down" orientation in IM2, which is the catalytically active conformation.





**Figure S18.** Aligned structures for setup A (orange), setup G (red), and the crystal structures 1FO4 (dark blue), 3EUB (ice blue), and 3B9J (cyan).

Intuitively, it seems strange that a binding pocket containing residues with quite polar side chains should bind two different orientations of the same substrate equally or at least comparably well. On the other hand one should keep in mind that XO is not uniquely designed to transform xanthine to uric acid, but also hypoxanthine to xanthine. The binding pocket may thus be a compromise for performing both tasks. Therefore, the induced fit with different substrates may result in rather different orientations in the binding pocket, as can be seen in Figure S18 which shows crystal structures with three different substrates (salicylic acid - dark blue, 2,8-dioxo-6-methylpurin - ice blue, and xanthine - cyan). Obviously, the orientation of the substrates varies and the enzymatic side chains adopt specific orientations. The optimized QM/MM structures (orange and red) differ slightly more, particularly with respect to the position of the molybdopterin phosphate (which is however known from experiment to be very floppy, with different orientations found in different crystal structures, see e.g. Schindelin, H.; Kisker, C.; Rees, D. C. *J. Biol. Inorg. Chem.* **1997**, 2, 773-781); also, the geometry of the molybdenum center is less tilted than in the experimental structures. However, all substrates conserve the important interaction with Arg880 in the binding pocket.

## 5. CHARMM parameters

```
!=====
! CHARMM PARAMETERS for FAD provided by M. Karplus,
!                               cf. J. Phys. Chem. B, 110, 9363-9367 (2006)
!=====
```

### !!!!!!!!!!!!!! FAD TOPOLOGY !!!!!!!!!!!!!!!

MASS	304	NX1	14.00700	N	! FAD-specific
MASS	305	HX1	1.00800	H	! FAD-specific
MASS	306	CX1	12.011000	C	! FAD-specific
MASS	307	OX1	15.999400	O	! FAD-specific
MASS	308	CX2	12.011000	C	! FAD-specific
MASS	309	CX3	12.011000	C	! FAD-specific
MASS	310	NX2	14.00700	N	! FAD-specific
MASS	311	CX4	12.011000	C	! FAD-specific
MASS	312	OX2	15.999400	O	! FAD-specific
MASS	313	NY1	14.00700	N	! FAD-specific
MASS	314	CY1	12.011000	C	! FAD-specific
MASS	315	CY2	12.011000	C	! FAD-specific
MASS	316	NY2	14.00700	N	! FAD-specific
MASS	317	CZ1	12.011000	C	! FAD-specific
MASS	318	HZ1	1.00800	H	! FAD-specific
MASS	319	CZ2	12.011000	C	! FAD-specific
MASS	320	CZ3	12.011000	C	! FAD-specific
MASS	321	CZ4	12.011000	C	! FAD-specific
MASS	322	HZ4	1.00800	H	! FAD-specific

```
RESI FAD -2.000      ! cf. J. Phys. Chem. B, 110, 9363-9367 (2006)
GROUP
```

ATOM	C4'	CN7	0.16
ATOM	H4'	HN7	0.09
ATOM	O4'	ON6B	-0.50
ATOM	C1'	CN7B	0.16
ATOM	H1'	HN7	0.09

GROUP

ATOM	C5	CN5	0.28
ATOM	N7	NN4	-0.71
ATOM	C8	CN4	0.34
ATOM	H8	HN3	0.12
ATOM	N9	NN2	-0.05
ATOM	N1	NN3A	-0.74
ATOM	C2	CN4	0.50
ATOM	H2	HN3	0.13
ATOM	N3	NN3A	-0.75
ATOM	C4	CN5	0.43
ATOM	C6	CN2	0.46
ATOM	N6	NN1	-0.77
ATOM	H61	HN1	0.38
ATOM	H62	HN1	0.38

GROUP					
ATOM C2'	CN7	0.14 !		H61	H62
ATOM H2''	HN7	0.09 !		\	/
ATOM O2'	ON5	-0.66 !		N6	
ATOM H2'	HN5	0.43 !			
GROUP				C6	
ATOM C3'	CN7	!		//	\
ATOM H3'	HN7	0.09 !		N1	C5--N7\\
ATOM O3'	ON5	-0.66 !			C8-H8
ATOM H3T	HN5	0.43 !		C2	C4--N9/
GROUP				/ \\ / \	
ATOM C5'	CN8	-0.08 !		H2	N3
ATOM H5'	HN8	0.09 !			
ATOM H5''	HN8	0.09 !			
ATOM O5'	ON2	-0.62 !			
ATOM PA	P	1.50 !	O1B	O1A	H5' H4' O4' \
ATOM O1A	ON3	-0.82 !			\ / \ \
ATOM O2A	ON3	-0.82 !	O3B-PB-O3A--PA-O5'-C5'---C4'		C1'
ATOM O3A	ON2	-0.74 !			\ / \
ATOM PB	P	1.50 !	O2B	O2A	H5'' C3'--C2' H1'
ATOM O1B	ON3	-0.90 !	HS51-CS5-HS52		/ \ / \
ATOM O2B	ON3	-0.90 !			O3' H3' O2' H2''
ATOM O3B	ON2	-0.74 !	HS41-CS4-OS4-HS42		
GROUP			!	H3'	H2'
ATOM CS1	CT2	0.21 !	HS31-CS3-OS3-HS32		
ATOM HS11	HA	0.09 !			
ATOM HS12	HA	0.09 !	HS21-CS2-OS2-HS22		
ATOM CS2	CT1	0.17 !			
ATOM HS21	HA	0.09 !	HS11-CS1-HS12		
ATOM OS2	OH1	-0.66	\		
ATOM HS22	H	0.43 !	\ \		
ATOM CS3	CT1	0.17 !	\ \		
ATOM HS31	HA	0.09 !	\ \		
ATOM OS3	OH1	-0.66 !	\ \		
ATOM HS32	H	0.43 !	\ \		
ATOM CS4	CT1	0.17 !	\ \	HZ4 HZ61 HZ62	
ATOM HS41	HA	0.09 !	\	/	
ATOM OS4	OH1	-0.66 !	OX2 NX2 NY2 CZ4 CZ6--HZ63		
ATOM HS42	H	0.43 !	\\ / \ / \ / \\ /		
ATOM CS5	CT2	0.22 !	CX4 CX3 CY2 CZ3 HZ53		
ATOM HS51	HA	0.09 !			
ATOM HS52	HA	0.09 !	NX1 CX2 CY1 CZ2---CZ5--HZ52		
GROUP			! / \ / \ // \ //		
ATOM NX1	NX1	-0.89 !	HX1 CX1 NY1 CZ1 HZ51		
ATOM HX1	HX1	0.43 !			
ATOM CX1	CX1	0.82 !	OX1	HZ1	
ATOM OX1	OX1	-0.61			
ATOM CX2	CX2	0.24			
ATOM CX3	CX3	0.37			
ATOM NX2	NX2	-0.81			

ATOM	CX4	CX4	1.09						
ATOM	OX2	OX2	-0.66						
ATOM	NY1	NY1	-0.60						
ATOM	CY1	CY1	0.60						
ATOM	CY2	CY2	-0.30						
ATOM	NY2	NY2	0.20						
ATOM	CZ1	CZ1	-0.46						
ATOM	HZ1	HZ1	0.24						
ATOM	CZ2	CZ2	0.19						
ATOM	CZ5	CT3	-0.44						
ATOM	HZ51	HA	0.09						
ATOM	HZ52	HA	0.09						
ATOM	HZ53	HA	0.09						
ATOM	CZ3	CZ3	0.18						
ATOM	CZ6	CT3	-0.48						
ATOM	HZ61	HA	0.09						
ATOM	HZ62	HA	0.09						
ATOM	HZ63	HA	0.09						
ATOM	CZ4	CZ4	-0.28						
ATOM	HZ4	HZ4	0.19						
BOND	PB	O3A	PB	O1B	PB	O2B	PB	O3B	O3A PA
BOND	PA	O1A	PA	O2A	PA	O5'	O3'	H3T	
BOND	O5'	C5'	C5'	C4'	C4'	O4'	C4'	C3'	O4' C1'
BOND	C1'	N9	C1'	C2'	N9	C4	N9	C8	C4 N3
BOND	C2	N1	C6	N6					
BOND	N6	H61	N6	H62	C6	C5	C5	N7	
BOND	C2'	C3'	C2'	O2'	O2'	H2'	C3'	O3'	
BOND	C1'	H1'	C2'	H2''	C3'	H3'	C4'	H4'	C5' H5'
BOND	C5'	H5''	C8	H8	C2	H2			
BOND	CS1	HS11	CS1	HS12	CS1	CS2	CS2	HS21	CS2 OS2
BOND	OS2	HS22	CS2	CS3	CS3	HS31	CS3	OS3	OS3 HS32
BOND	CS3	CS4	CS4	HS41	CS4	OS4	OS4	HS42	CS4 CS5
BOND	CS5	HS51	CS5	HS52	CS5	O3B	CS1	NY2	
BOND	NX1	HX1	NX1	CX1	CX1	OX1	CX1	CX2	CX2 CX3
BOND	CX3	NX2	NX2	CX4	CX4	OX2	CX2	NY1	NY1 CY1
BOND	CY1	CY2	CY2	NY2	NY2	CX3	NX1	CX4	
BOND	CY1	CZ1	CZ1	HZ1	CZ1	CZ2	CZ2	CZ5	CZ5 HZ51
BOND	CZ5	HZ52	CZ5	HZ53	CZ2	CZ3	CZ3	CZ6	CZ6 HZ61
BOND	CZ6	HZ62	CZ6	HZ63	CZ3	CZ4	CZ4	HZ4	CZ4 CY2
DOUBLE			N1	C6	N3	C2	C4	C5	N7 C8
IMPR	N6	C6	H61	H62					
IMPR	C6	N1	C5	N6					
IMPR	CZ2	CZ1	CZ3	CZ5					
IMPR	CZ3	CZ2	CZ4	CZ6					
IMPR	CY1	CZ1	NY1	CY2					
IMPR	NY2	CY2	CX3	CS1					
IMPR	CX2	NY1	CX1	CX3					
PATC	FIRS	NONE	LAST	NONE					

!!!!!!!!!!!!!! FAD PARAMETER !!!!!!!!!!!!!!!

**BONDS**

CX4	NX2	350.000	1.3700	!	FAD
CX4	NX1	350.000	1.4000	!	FAD
CX4	OX2	640.000	1.1900	!	FAD
NX2	CX3	350.000	1.2900	!	FAD
CX3	CX2	282.000	1.4700	!	FAD
CX2	CX1	282.000	1.5000	!	FAD
CX3	NY2	350.000	1.3600	!	FAD
CX2	NY1	350.000	1.2700	!	FAD
CX1	OX1	640.000	1.1900	!	FAD
CX1	NX1	350.000	1.3600	!	FAD
NX1	HX1	465.600	1.0000	!	FAD
NY2	CT2	310.000	1.4600	!	FAD
NY2	CY2	390.000	1.3900	!	FAD
CY2	CZ4	440.000	1.4000	!	FAD
CY2	CY1	440.000	1.4000	!	FAD
CY1	NY1	390.000	1.3700	!	FAD
CY1	CZ1	440.000	1.4000	!	FAD
CZ4	HZ4	370.000	1.0700	!	FAD
CZ4	CZ3	440.000	1.3800	!	FAD
CZ3	CZ2	440.000	1.4100	!	FAD
CZ2	CZ1	440.000	1.3700	!	FAD
CZ1	HZ1	370.000	1.0700	!	FAD
CZ2	CT3	345.000	1.5100	!	FAD
CZ3	CT3	345.000	1.5100	!	FAD
CT2	ON2	340.000	1.4400	!	FAD

**ANGLES**

OX2	CX4	NX1	98.00	118.6	!	FAD
OX2	CX4	NX2	98.00	123.1	!	FAD
NX1	CX4	NX2	70.00	118.3	!	FAD
CX3	NX2	CX4	60.00	120.5	!	FAD
NX2	CX3	CX2	65.00	124.1	!	FAD
NX2	CX3	NY2	90.00	120.1	!	FAD
CX2	CX3	NY2	65.00	115.8	!	FAD
CX1	CX2	CX3	65.00	116.8	!	FAD
CX1	CX2	NY1	65.00	119.1	!	FAD
CX3	CX2	NY1	65.00	124.1	!	FAD
CX2	CX1	OX1	80.00	124.0	!	FAD
CX2	CX1	NX1	65.00	112.9	!	FAD
NX1	CX1	OX1	98.00	123.2	!	FAD
CX1	NX1	HX1	18.00	117.0	!	FAD
CX4	NX1	CX1	60.00	127.4	!	FAD
CX4	NX1	HX1	18.00	117.0	!	FAD
CX3	NY2	CT2	44.00	119.5	!	FAD
CT2	NY2	CY2	80.00	119.6	!	FAD
CX3	NY2	CY2	60.00	120.9	!	FAD
NY2	CY2	CZ4	65.00	122.7	!	FAD
NY2	CY2	CY1	65.00	118.9	!	FAD
CZ4	CY2	CY1	70.00	118.4	!	FAD
CY2	CY1	NY1	65.00	121.2	!	FAD
CY2	CY1	CZ1	70.00	120.1	!	FAD
NY1	CY1	CZ1	65.00	118.7	!	FAD
CY1	NY1	CX2	60.00	119.2	!	FAD

CY2	CZ4	HZ4	31.00	120.2	!	FAD		
CZ3	CZ4	HZ4	31.00	118.6	!	FAD		
CY2	CZ4	CZ3	70.00	121.1	!	FAD		
CZ4	CZ3	CZ2	70.00	120.5	!	FAD		
CZ4	CZ3	CT3	70.00	119.3	!	FAD		
CZ2	CZ3	CT3	70.00	120.2	!	FAD		
CZ1	CZ2	CZ3	70.00	118.1	!	FAD		
CZ3	CZ2	CT3	70.00	121.2	!	FAD		
CZ1	CZ2	CT3	70.00	120.7	!	FAD		
CZ2	CZ1	CY1	70.00	121.8	!	FAD		
CZ2	CZ1	HZ1	31.00	121.2	!	FAD		
CY1	CZ1	HZ1	31.00	117.1	!	FAD		
CZ2	CT3	HA	35.00	115.0	!	FAD		
CZ3	CT3	HA	35.00	115.0	!	FAD		
NY2	CT2	HA	46.50	106.2	!	FAD		
NY2	CT2	CT1	65.00	109.5	!	FAD		
CN8	CN7	ON6B	110.00	109.0	!	FAD		
ON6B	CN7B	CN7	110.00	105.0	!	FAD		
NN2	CN7B	CN7	70.00	113.7	!	FAD		
CN7	CN7	CN7	53.35	111.0	8.0	2.561	!	FAD
CN7	CN8	ON2	75.70	110.1	!	FAD		
P	ON2	CT2	20.0	120.0	35.	2.33	!	FAD
CT2	CT1	OH1	75.700	110.1000	!	FAD		
ON2	CT2	CT1	75.700	110.1000	!	FAD		
ON2	CT2	HA	45.900	108.8900	!	FAD		

# DIHEDRAL

ON3	P	ON2	CT2	0.10	3	0.0	!	FAD
X	CX4	NX2	X	2.05	2	180.0	!	FAD
X	NX2	CX3	X	2.05	2	180.0	!	FAD
X	CX2	CX3	X	0.50	2	180.0	!	FAD
X	CX1	CX2	X	0.50	2	180.0	!	FAD
X	CX1	NX1	X	2.05	2	180.0	!	FAD
X	CX4	NX1	X	2.05	2	180.0	!	FAD
X	CX3	NY2	X	2.05	2	180.0	!	FAD
X	CY2	NY2	X	2.05	2	180.0	!	FAD
X	CY1	CY2	X	3.10	2	180.0	!	FAD
X	NY1	CY1	X	2.05	2	180.0	!	FAD
X	NY1	CX2	X	2.05	2	180.0	!	FAD
X	CY2	CZ4	X	3.10	2	180.0	!	FAD
X	CZ4	CZ3	X	3.10	2	180.0	!	FAD
X	CZ2	CZ3	X	3.10	2	180.0	!	FAD
X	CZ1	CZ2	X	3.10	2	180.0	!	FAD
X	CY1	CZ1	X	3.10	2	180.0	!	FAD
X	CZ2	CT3	X	0.10	3	0.0	!	FAD
X	CZ3	CT3	X	0.10	3	0.0	!	FAD
X	CT2	NY2	X	0.01	3	0.0	!	FAD
X	CN7	CN7	X	0.15	3	0.0	!	FAD
X	CT2	ON2	X	0.25	3	0.0	!	FAD
X	ON6B	CN7B	X	0.10	3	0.0	!	FAD
X	CN7	CN8	X	0.200	3	0.0	!	FAD
X	CN7	CN7B	X	0.15	3	0.0	!	FAD
X	CN7	ON6B	X	0.10	3	0.0	!	FAD
X	CN8	ON2	X	-0.10	3	0.0	!	FAD
ON2	P	ON2	CT2	0.95	2	0.0	!	FAD

**IMPROPER**

CZ1	X	X	CZ4	15.0000	0	0.0000	!	FAD
CZ2	X	X	CY2	15.0000	0	0.0000	!	FAD
CZ3	X	X	CY1	15.0000	0	0.0000	!	FAD
NY1	X	X	NY2	65.0000	0	0.0000	!	FAD
CY1	X	X	CX3	100.0000	0	0.0000	!	FAD
CX2	X	X	CY2	100.0000	0	0.0000	!	FAD
CX2	X	X	CX4	100.0000	0	0.0000	!	FAD
NY2	X	X	CX4	60.0000	0	180.0000	!	FAD
NY2	X	X	CX1	60.0000	0	180.0000	!	FAD
CX3	X	X	NX1	60.0000	0	0.0000	!	FAD
CX1	X	X	NX2	60.0000	0	0.0000	!	FAD
CZ3	X	X	CT3	130.0000	0	0.0000	!	FAD
CY1	X	X	CY2	15.0000	0	0.0000	!	FAD
CZ2	X	X	CT3	130.0000	0	0.0000	!	FAD
NY2	X	X	CT2	45.0000	0	0.0000	!	FAD
CX2	X	X	CX3	100.0000	0	0.0000	!	FAD
CX4	X	X	OX2	147.0000	0	0.0000	!	FAD
NX1	X	X	HX1	18.0000	0	0.0000	!	FAD
CX1	X	X	OX1	147.0000	0	0.0000	!	FAD

**NONBONDED**

CX1	0.000000	-0.141000	1.870000	!	FAD
CX2	0.000000	-0.141000	1.870000	!	FAD
CX3	0.000000	-0.141000	1.870000	!	FAD
CX4	0.000000	-0.141000	1.870000	!	FAD
CY1	0.000000	-0.141000	1.870000	!	FAD
CY2	0.000000	-0.141000	1.870000	!	FAD
CZ1	0.000000	-0.141000	1.870000	!	FAD
CZ2	0.000000	-0.141000	1.870000	!	FAD
CZ3	0.000000	-0.141000	1.870000	!	FAD
CZ4	0.000000	-0.141000	1.870000	!	FAD
HX1	0.000000	-0.049800	0.800000	!	FAD
HZ1	0.000000	-0.042000	1.330000	!	FAD
HZ4	0.000000	-0.042000	1.330000	!	FAD
NX1	0.000000	-0.090000	1.830000	!	FAD
NX2	0.000000	-0.090000	1.830000	!	FAD
NY1	0.000000	-0.090000	1.830000	!	FAD
NY2	0.000000	-0.090000	1.830000	!	FAD
OX1	0.000000	-0.120000	1.700000	!	FAD
OX2	0.000000	-0.120000	1.700000	!	FAD

```

!=====
! CHARMM topology for molybdenum-bond uric acid
!=====

```

!!!!!!! **TOPOLOGY** !!!!!!!!

```

MASS 353    UN2U    14.00700  N ! URI-specific angles assigned
MASS 354    UHN2    1.00800  H ! URI-specific angles assigned
MASS 355    UC1T    12.011000 C ! URI-specific angles assigned
MASS 356    UON1    15.999400 O ! URI-specific angles assigned
MASS 357    UN2B    14.00700  N ! URI-specific angles assigned
MASS 358    UCN5    12.011000 C ! URI-specific angles assigned
MASS 359    UC5G    12.011000 C ! URI-specific angles assigned
MASS 360    UCN1    12.011000 C ! URI-specific angles assigned
MASS 361    UC8T    12.011000 C ! URI-specific angles assigned
MASS 362    UN7B    14.00700  N ! URI-specific angles assigned
MASS 363    UN9B    14.00700  N ! URI-specific angles assigned

```

```

RESI URIC      -1.00  ! CHARGES are UNREVISED ESP CHARGES!!!!
GROUP

```

```

ATOM N1    UN2U    -0.94  !
ATOM H1    UHN2     0.41  !
ATOM C2    UC1T     1.10  !
ATOM O2    UON1    -0.76  !
ATOM N3    UN2B    -0.94  !
ATOM C4    UCN5     0.68  !
ATOM C5    UC5G    -0.32  !
ATOM C6    UCN1     0.82  !
ATOM O6    UON1    -0.70  !
ATOM N7    UN7B    -0.61  !
ATOM H7    UHN2     0.40  !
ATOM C8    UC8T     0.81  !
ATOM O8    UON1    -0.71  !
ATOM N9    UN9B    -0.63  !
ATOM H9    UHN2     0.39  !

```

```

BOND      H1 N1    N1 C2    C2 N3    N3 C4    C4 N9
BOND      N9 H9    C8 N7    N7 C5    C5 C6    C6 N1
BOND      N7 H7    N9 C8
DOUBLE   C2 O2    C4 C5    C6 O6    C8 O8

```

```

!Improper and DONO/ACC from URA (& patch PYRU) and GUA (& patch
PURG)

```

```

IMPR C2 N3 N1 O2    C6 N1 C5 O6    N9 C4 H9 C8    N7 C5 C8 H7
DONO H1  N1
DONO H7  N7
DONO H9  N9
ACCE O2  C2
ACCE O6  C6
ACCE O8  C8
AUTO ANGLES DIHE
PATC FIRS NONE LAST NONE

```



# !!!!!!!!!!!!!! URIC PARAMETER !!!!!!!!!!!!!!!

## BONDS

UHN2	UN2U	474.0	1.00	!U, JWK, adm jr. 7/24/91
UN2U	UC1T	340.0	1.44	!U,T adm jr. 11/97
UC1T	UON1	860.0	1.19	!nad/ppi, jjpl/adm jr. 7/95
UC1T	UN2B	302.0	1.40	!U,T adm jr. 11/97
UHN2	UN2B	474.0	1.00	!G, adm jr. 11/97 AS IN PYRO & PURG patch
UCN5	UN2B	320.0	1.37	!U,T adm jr. 11/97 from URA: CN3 NN2B
UCN5	UC5G	320.0	1.30	!G, adm jr. 11/97
UC5G	UCN1	302.0	1.41	!G, adm jr. 11/97
UCN1	UN2U	340.0	1.44	!U,T adm jr. 11/97 from URA: CN1 NN2U
UC5G	UN7B	302.0	1.40	!G, adm jr. 11/97 from GUA: CN5 NN2B
UC8T	UN7B	302.0	1.38	!U,T adm jr. 11/97
UCN5	UN9B	302.0	1.39	!G, adm jr. 11/97 from GUA: CN5 NN2B
UHN2	UN7B	474.0	1.00	!G, adm jr. 11/97 AS IN PYRO & PURG patch
UC8T	UON1	860.0	1.28	!nad/ppi, jjpl/adm jr. 7/95
UHN2	UN9B	474.0	1.00	!G, adm jr. 11/97 AS IN PYRO & PURG patch
UC8T	UN9B	302.0	1.33	!U,T adm jr. 11/97
UCN1	UON1	660.0	1.20	!U,A,G par_a4 adm jr. 10/2

## ANGLES

UN2B	UCN5	UN9B	140.0	126.6	! from NN2B CN5 NN3G
UC5G	UCN5	UN9B	100.0	111.9	! G
UC5G	UN7B	UC8T	100.0	106.6	! from CN4 NN2B CN5
UHN2	UN7B	UC8T	40.5	128.8	! U, h1
UN7B	UC8T	UON1	100.0	123.2	! U, o2
UON1	UC8T	UN9B	100.0	124.2	! U, o2
UCN5	UN9B	UC8T	70.0	103.9	! from CN1T NN2B CN3
UCN5	UN9B	UHN2	40.5	121.0	! U, h1
UC8T	UN9B	UHN2	40.5	135.1	! U, h1
UC1T	UN2U	UHN2	40.5	114.4	! U, h3
UCN1	UN2U	UHN2	40.5	116.2	! U
UC1T	UN2U	UCN1	50.0	129.4	! U
UN2U	UC1T	UON1	100.0	122.1	! U
UN2B	UC1T	UON1	100.0	123.1	! U, o2
UN2B	UC1T	UN2U	50.0	114.9	! U
UC1T	UN2B	UCN5	70.0	121.0	! from URA CN1T NN2B CN3
UC1T	UN2B	UHN2	40.5	119.5	! U, h1
UCN5	UN2B	UHN2	32.0	120.5	! U !!! from URA CN3 NN2B HN2
UC5G	UCN5	UN2B	100.0	121.4	! G
UHN2	UN7B	UC5G	30.0	125.1	! from CN5 NN2B HN2
UCN5	UC5G	UN7B	100.0	104.6	! from CN5G CN5 NN2B
UCN1	UC5G	UN7B	125.0	127.4	! from CN1 CN5G NN4
UCN1	UC5G	UCN5	70.0	122.9	!
UC5G	UCN1	UN2U	100.0	110.3	! from : CN5G CN5 NN2B
UC5G	UCN1	UON1	50.0	125.3	!
UN2U	UCN1	UON1	50.0	122.1	! from NN2G CN1 ON1
UN7B	UC8T	UN9B	50.0	113.5	! U

## DIHEDRALS

UON1	UCN1	UC5G	UN7B	0.0	2	180.0	! from GUA ON1 CN1 CN5G NN4
UN2U	UCN1	UC5G	UN2B	2.0	2	180.0	! from GUA NN2G CN1 CN5G NN4
UN2U	UCN1	UC5G	UN7B	2.0	2	180.0	! from GUA NN2G CN1 CN5G NN4
UC1T	UN2B	UCN5	UCN1	2.0	2	180.0	! from GUA CN2 NN3G CN5 NN2B
UCN5	UN2B	UC1T	UN2U	1.5	2	180.0	! from URA CN3 NN2B CN1T NN2U

UON1	UCN1	UC5G	UCN5	14.0	2	180.0	!	from GUA	ON1	CN1	CN5G	CN5
UCN5	UC5G	UCN1	UN2U	0.2	2	180.0	!	from GUA	CN5	CN5G	CN1	NN2G
UC5G	UCN1	UN2U	UC1T	1.5	2	180.0	!	from URA	CN1T	NN2U	CN1	CN3
UN2B	UCN5	UC5G	UCN1	2.0	2	180.0	!	from GUA	NN3G	CN5	CN5G	CN1
UON1	UCN1	UN2U	UHN2	0.0	2	180.0	!	from URA	ON1	CN1T	NN2U	HN2
UN2B	UC1T	UN2U	UCN1	1.5	2	180.0	!	from URA	NN2B	CN1T	NN2U	CN1
UC5G	UCN5	UN2B	UHN2	1.6	2	180.0	!	from URA	CN3	CN3	NN2B	HN2
UN2U	UC1T	UN2B	UHN2	1.6	2	180.0	!	from URA	NN2U	CN1T	NN2B	HN2
UON1	UC1T	UN2B	UHN2	0.0	2	180.0	!	from URA	ON1	CN1T	NN2B	HN2
UON1	UC1T	UN2U	UHN2	0.0	2	180.0	!	from URA	ON1	CN1T	NN2U	HN2
UN2B	UC1T	UN2U	UHN2	3.8	2	180.0	!	from URA	NN2B	CN1T	NN2U	HN2
UC5G	UCN1	UN2U	UHN2	4.8	2	180.0	!	from URA	CN3T	CN1	NN2U	HN2
UC5G	UCN5	UN2B	UC1T	1.5	2	180.0	!	from URA	CN3	CN3	NN2B	CN1T
UON1	UCN1	UN2U	UC1T	14.0	2	180.0	!	from GUA	ON1	CN1	NN2G	CN2
UHN2	UN2B	UC5G	UCN5	1.2	2	180.0	!	from Gua	HN2	NN2B	CN5	CN5G
UHN2	UN7B	UC5G	UCN5	1.2	2	180.0	!	from Gua	HN2	NN2B	CN5	CN5G
UCN1	UC5G	UN7B	UHN2	1.2	2	180.0	!	from XHN2	XN2B	XCN4	XNN4	
UC1T	UN2B	UCN5	UN9B	2.0	2	180.0	!	from GUA	CN2	NN3G	CN5	NN2B
UN2B	UCN5	UN9B	UC8T	2.0	2	180.0	!	from GUA	NN3G	CN5	NN2B	CN4
UN2B	UCN5	UN9B	UHN2	1.2	2	180.0	!	from GUA	HN2	NN2B	CN4	NN4
UHN2	UN2B	UCN5	UN9B	1.2	2	180.0	!	from XHN2	XN2B	XCN4	XNN4	
UCN5	UN9B	UC8T	UN2B	16.0	2	180.0	!	from GUA	CN5G	NN4	CN4	NN2B
UCN5	UC5G	UN2B	UC8T	6.0	2	180.0	!	from GUA	CN4	NN2B	CN5	CN5G
UCN5	UC5G	UN7B	UC8T	6.0	2	180.0	!	from GUA	CN4	NN2B	CN5	CN5G
UC5G	UCN5	UN9B	UC8T	6.0	2	180.0	!	from GUA	CN5	CN5G	NN4	CN4
UC5G	UCN5	UN9B	UHN2	1.6	2	180.0	!	from URA	CN3	CN3	NN2B	HN2
UCN1	UC5G	UN7B	UC8T	2.0	2	180.0	!	from GUA	CN1	CN5G	NN4	CN4
X	UC1T	UN2U	X	0.9	2	180.0	!					
X	UC1T	UN2B	X	0.9	2	180.0	!	From X	CN1	NN2	X,	for thymine
X	UC5G	UCN5	X	0.0	2	180.0	!					
X	UC8T	UN2U	X	0.9	2	180.0	!					
X	UC8T	UN7B	X	0.9	2	180.0	!	From X	CN1	NN2	X,	for thymine
X	UC8T	UN9B	X	0.9	2	180.0	!	From X	CN1	NN2	X,	for thymine

#### IMPROPER

UC1T	X	X	UON1	90.0	0	0.0	!U					
UC8T	X	X	UON1	90.0	0	0.0	!U					
UCN1	X	X	UON1	90.0	0	0.0	!U					
UN9B	X	X	UC8T	7.0	0	0.0	!G,	adm jr.	11/97			
UN2B	X	X	UHN2	7.0	0	0.0	!G,	adm jr.	11/97	NN2B	CN4	CN5
UN7B	X	X	UHN2	7.0	0	0.0	!G,	adm jr.	11/97	NN2B	CN4	CN5

#### NONBONDED

UN2U	0.0	-0.20	1.85									
UHN2	0.0	-0.0460	0.2245									
UC1T	0.0	-0.10	1.9000									
UON1	0.0	-0.1200	1.70									
UN2B	0.0	-0.20	1.85	!	From NN2,	for N9	in guanines					
UCN5	0.0	-0.075	1.9000									
UC5G	0.0	-0.075	1.9000									
UCN1	0.0	-0.10	1.9000									
UC8T	0.0	-0.10	1.9000									
UN7B	0.0	-0.20	1.85	!	From NN2,	for N9	in guanines					
UN9B	0.0	-0.20	1.85	!	From NN2,	for N9	in guanines					

```

=====
! CHARMM Parameter Molybdopterin, NOT RECOMMENDED FOR CLASSICAL MD
=====
!!!!!!!!!!!!!! TOPOLOGY !!!!!!!!!!!!!!!
MASS    210 MO      95.94000 MO ! FOR MOLYBDENUM
MASS    211 OMO     15.99900 O  ! MO==O
MASS    212 SMO     32.06000 S  ! MO==S

RESI MOC          -3.00    ! Mo-cofactor: -1 (Mo) + -2 (phosphate)
GROUP          ! charge -1 Part 1 (on which PATCH MOPS is applied)
ATOM  Mo        MO      0.947  !
ATOM  OM1       OMO     -0.552  ! OM1
ATOM  OM2       OH1     -0.923  ! ||
ATOM  HOM2      H       0.503  ! S7'---Mo==SR1
ATOM  SR1       SMO     -0.271  ! / / \
ATOM  S8'       SS      -0.172  ! (C6)---C7' S8' OM2
ATOM  S7'       SS      -0.055  ! \ \ / |
ATOM  C8'       CE1     -0.235  ! (C3')---C8' HOM2
ATOM  C7'       CE1     -0.242  !
GROUP          ! charge: 0.00 Part 2
ATOM  N1        NN3G    -0.74
ATOM  C2        CN2      0.75  !
ATOM  N2        NN1     -0.68  !
ATOM  H21'      HN1      0.35  !
ATOM  H22'      HN1      0.32  ! H6'
ATOM  N3        NN2G    -0.34  ! O4 H5' |
ATOM  H3        HN2      0.26  ! || | | (C7')
ATOM  C4        CN1      0.54  ! C4 N5 | /
ATOM  O4        ON1     -0.51  ! / \ / \ | / (C8')
ATOM  C9        CN5G     0.00  ! H3-N3 C9 C6 |
ATOM  C10       CN5      0.26  ! | || | |
ATOM  N8        N2      -0.61  ! C2 C10 C7 C3'---(C4')
ATOM  H8'       HC       0.40  ! / \ \ / \ / | \ / \
GROUP          ! H21'-N2 N1 N8 | O3' H3'
ATOM  N5        N2      -0.61  ! | | |
ATOM  H5'       HC       0.40  ! H22' H8' |
ATOM  C6        CNT2     0.12  ! H7
ATOM  H6'       HA       0.09
ATOM  C7        CN7B     0.16
ATOM  H7        HN7      0.09
ATOM  O3'       ON6B    -0.50
ATOM  C3'       CN7      0.16
ATOM  H3'       HN7      0.09
GROUP          ! charge: -2.0 Part 3, from RESI MP_2
ATOM  C4'       CN9     -0.18  ! (C3')
ATOM  H101      HN9      0.09  ! |
ATOM  H102      HN9      0.09  ! H102---C4'---H101
GROUP          ! |
ATOM  O4'       ON2     -0.40  ! O4'
ATOM  P         P       1.10  ! |
ATOM  O1P       ON3     -0.90  ! O3P(-)---P(+)--O1P(-)
ATOM  O2P       ON3     -0.90  ! |
ATOM  O3P       ON3     -0.90  ! O2P(-)

```

!!!!!!!!!!!!!! TOPOLOGY PATCH !!!!!!!!!!!!!!!

```
PRES MOPS -1.000      ! patch to MO(IV) species
GROUP                !                OM1
ATOM   Mo      MO      0.775  !                ||
ATOM   OM1     OMO     -0.498  !                S7'---Mo---SR1---HSR
ATOM   SR1     SMO     -0.576  !                /      /
ATOM   HSR     HS      0.133  !                /      /
ATOM   S8'     SS      -0.514  !      ---C7'  S8'
ATOM   S7'     SS      -0.261  !                \\  /
ATOM   C8'     CE1      0.072  !      ---C8'
ATOM   C7'     CE1     -0.131  !
```

DELETE ATOM OM2

DELETE ATOM HOM2

BOND SR1 HSR

ANGLE HSR SR1 MO

IC SR1 MO OM1 S8' 2.1700 107.51 -100.79 90.00 0.0000

IC HSR SR1 MO OM1 1.3900 101.66 84.64 000.00 0.0000

!!!!!!!!!!!!!! MOCO PARAMETER !!!!!!!!!!!!!!!

**BONDS** !=====dummy parameters=====

```
SMO HS 275.000 1.385 ! from S HS
MO OMO 00.00 1.73 ! from lvdv
OC MO 00.00 2.15
MO SMO 00.00 2.17 ! from lvdv
MO OH1 250.0 1.97 ! from lvdv K from QM !!!!!!!
MO SS 0.0 0.0
SS CE1 0.0 0.0
```

**ANGLES** !=====dummy parameters=====

```
OMO MO OH1 50.000 105.495 ! From QM
OMO MO OMO 00.000 107.507 ! x-ray
OMO MO SMO 00.000 107.827 ! From QM
OMO MO SS 00.000 101.542 ! From QM
OH1 MO SMO 70.000 100.000 !
OH1 MO SS 70.000 075.000 !
SMO MO SS 00.000 077.239 !
SS MO SS 00.000 079.321 ! From QM
MO OH1 H 35.000 119.659 ! From QM
MO SS CE1 00.000 107.736 ! From QM
SS CE1 CE1 00.000 122.122 ! From QM
HS SMO MO 30.000 100.000 ! FROM QM
CC OC MO 35.000 120.000 !
```

**DIHEDRAL** !=====dummy parameters=====

```
MO SS CE1 CE1 00.00 1 3.908 ! from lvlb
MO SS CE1 CN7 00.00 1 -165.552 ! from lvlb
MO SS CE1 CNT2 00.00 1 178.353 ! from lvlb
OH1 MO SS CE1 20.0000 4 180.0
SMO MO SS CE1 0.0000 4 0.00
H OH1 MO SS 4.5000 2 0.00
OMO MO OH1 H 4.5000 2 0.00
SMO MO OH1 H 4.5000 2 0.00
OMO MO SS CE1 00.00 1 89.770 ! from lvlb
SS MO SS CE1 0.0000 1 0.00
```

**NONBONDED** !=====dummy parameters=====

```
MO 0.0100 0.0000 0.6500 !
OMO 0.0000 -0.1200 1.7000 0.000000 -0.120000 1.400000 !
SMO 0.0000 -0.4700 2.200000
```

## 6. Complete references 75 and 81

[75] CHARMM27 force field:

(a) MacKerell, A.D., Jr.; Feig, M.; Brooks, C.L., III *J. Comput. Chem.* **2004**, 25, 1400-1415.

(b) Foloppe, N.; Mackerell, A. D., Jr. *J. Comput. Chem.* **2000**, 21, 86-104.

(c) MacKerell, Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R.L.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, 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.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, 102, 3586-3616.

[81] (a) 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.; Terstegen, 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. Struc. (Theochem)* **2003**, 632, 1-28.

(b) ChemShell is a modular QM/MM program developed in the European QUASI project under the coordination of P. Sherwood (see <http://www.chemshell.org>).



## **Paper IV**

*QM/MM Studies of Xanthine Oxidase:  
Variations of Cofactor, Substrate and active site Glu802*

**Sebastian Metz**, and Walter Thiel

*J. Phys. Chem. B*, submitted.





# QM/MM Studies of Xanthine Oxidase: Variations of Cofactor, Substrate, and Active-Site Glu802

Sebastian Metz and Walter Thiel\*

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

Received: October 19, 2009

In continuation of our previous QM/MM study on the reductive half-reaction of wild-type xanthine oxidase, we consider the effects of variations in the cofactor, the substrate, and the active-site Glu802 residue on the reaction mechanism. Replacement of the sulfido ligand in the natural cofactor by an oxo ligand leads to a substantial increase in the computed barriers, consistent with the experimentally observed inactivity of this modified cofactor, whereas the selenido form is predicted to have lower barriers and hence higher activity. For the substrate 2-oxo-6-methylpurine, the calculated pathways for three different tautomers show great similarity to those found previously for xanthine, contrary to claims in the literature that the mechanisms for these two substrates are different. Compared with the wild-type enzyme, the conversion of xanthine to uric acid follows a somewhat different pathway in the Glu802 → Gln mutant which exhibits a lower overall activity, in agreement with recently published kinetic data. The present results confirm the basic stepwise reaction mechanism and the orientation of the substrate that has been proposed in our previous QM/MM work on aldehyde oxidoreductase and xanthine oxidase.

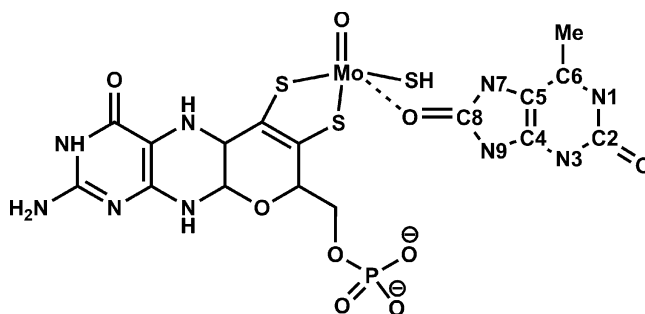
## 1. Introduction

The conversion of xanthine to uric acid is catalyzed by two molybdopterin containing enzymes, xanthine oxidase (XO, giving the name to the whole structural family<sup>1</sup>) and xanthine dehydrogenase (XDH). These enzymes have been studied for more than 100 years.<sup>2,3</sup> Contrary to many other oxygenases, they use water instead of elementary O<sub>2</sub> as their ultimate oxygen source<sup>4</sup> and they generate two reductive equivalents rather than consuming them.

Mechanistic information on the reductive half-reaction for the xanthine oxidase family was inferred from docking studies of xanthine in the structurally similar aldehyde oxidoreductase (AOR),<sup>5,6</sup> from the crystal structures of XO and XDH,<sup>7–9</sup> and from kinetic experiments with xanthine and other substrates of XO.<sup>10–13</sup> A Lewis-base-catalyzed mechanism involving the active-site glutamate residue (Glu1261 in XO and Glu869 in AOR) was proposed on the basis of these data.<sup>6–9,11,12</sup> Recent quantum mechanical/molecular mechanical (QM/MM) calculations<sup>14,15</sup> confirmed the general features of this mechanism and provided detailed insight into several competing pathways and the nature of the postulated intermediates.<sup>16</sup>

Substrate orientation within the enzyme was controversial for a long time. An “upside” conformation of xanthine relative to the molybdopterin cofactor was found in early modeling studies in 1996<sup>6</sup> and widely accepted thereafter. For the inactive oxo form of XO, this orientation of xanthine was recently proven by X-ray crystallography.<sup>17</sup> On the other hand, allopurinol<sup>18</sup> and 2-oxo-6-methylpurine,<sup>9</sup> both substrates similar to xanthine, are oriented “upside down” in their crystal structures (see Scheme 1). Therefore, different orientations for “good” substrates such as xanthine and “poor” substrates such as 2-oxo-6-methylpurine were postulated to explain the different reactivities.<sup>8,9</sup> By contrast, Yamaguchi et al.<sup>12</sup> proposed a reaction mechanism in which xanthine adopts the same “upside down” orientation as

**SCHEME 1: Reduced Molybdopterin Cofactor with Coordinated Product in the “Upside Down” Orientation, with the Numbering Scheme for the Nitrogen and Carbon Atoms in the Product**



found for 2-oxo-6-methylpurine in the crystal structure. The QM/MM results<sup>15</sup> suggest that the “upside” conformation represents the thermodynamic minimum, whereas xanthine is oriented “upside down” in the kinetically active pathway.

Our recent QM/MM work addressed the reaction mechanism in XO<sup>15</sup> and in the similar AOR system.<sup>14</sup> There are several other theoretical investigations on model systems, dealing with formaldehyde,<sup>19–21</sup> formamide,<sup>21–24</sup> as well as imidazole, xanthine, and several methylxanthines<sup>25</sup> as substrates. Since these have been summarized before, we shall not review them again here, except for noting that most of these studies only deal with a concerted one-step mechanism, which is less favorable than a stepwise mechanism in the enzymatic environment.<sup>14,15</sup> In this work, we will focus on this stepwise mechanism for the conversion of xanthine to uric acid in XO and explore the question whether it remains valid in variants of this reaction, i.e., upon changes in the cofactor, the substrate, and the Glu802 residue. Motivated by the need to understand recent experimental data (see below), we will concentrate on the following three issues.

\* To whom correspondence should be addressed. E-mail: thiel@mpi-muelheim.mpg.de.

**Cofactor.** The catalytic activity of XO is lost when the sulfido ligand in the cofactor is replaced by an oxo ligand.<sup>26</sup> We will study the reaction of xanthine in XO with the cofactor in its oxo form and also monitor the influence of a selenido ligand which is present in the nicotinate dehydrogenase system.<sup>27</sup> The lack of activity in the oxo form has been attributed to orbital effects in theoretical investigations of small model compounds.<sup>27</sup> We will provide a more detailed analysis, since we calculate the complete reaction pathway with all individual elementary steps and intermediates.

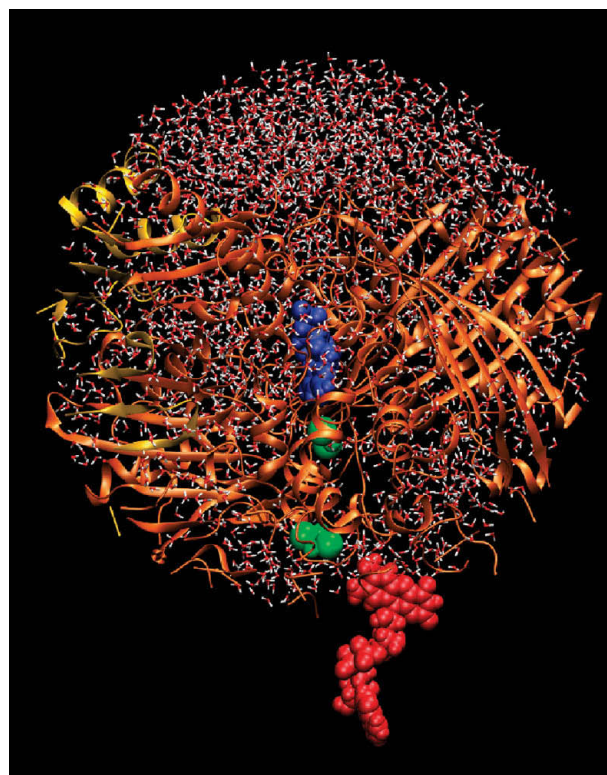
**Substrate.** The formation and decay of ESR active species following the reductive half-cycle of XO with xanthine as a substrate is rather fast. Therefore, 2-oxo-6-methylpurine was often used as a slower substrate to investigate this reaction experimentally.<sup>10,11,13,28,29</sup> However, it is not the conversion of the coordinated substrate to the product complex but the product release that makes 2-oxo-6-methylpurine a “slow” substrate.<sup>13</sup> Nevertheless, its lower overall reactivity led to the suggestion that xanthine and 2-oxo-6-methylpurine adopt different orientations within the binding pocket.<sup>8,17</sup> We will therefore study the reaction pathways of 2-oxo-6-methylpurine and compare them with those computed previously for xanthine in XO.

**Glu802.** Mutagenesis of active-site residues is a helpful tool to analyze their individual influence on the reactivity. Mutation of Glu1261 leads to a complete loss of activity<sup>30</sup> because it shuts down the initial proton transfer that is essential for the activation of the cofactor in XO. Arg880 influences both substrate binding and reaction rate,<sup>8,12</sup> by tuning the amount of electrostatic stabilization provided to the various forms of the substrate during the reaction.<sup>15</sup> Glu802 affects the reactivity of xanthine in XO only to a minor extent but plays a more pronounced role for hypoxanthine.<sup>12,30,31</sup> Recent experimental work with xanthine as a substrate has shown that the conservative mutation Glu802 → Gln in XO hardly changes the Michaelis constant ( $K_{\text{m(xanthine)}}$  =  $64.4 \pm 0.9 \mu\text{M}$  for the wild-type enzyme vs  $K_{\text{m(xanthine)}}$  =  $55.21 \pm 3.1 \mu\text{M}$  for the mutant), whereas the effective rate constant drops from  $k_{\text{cat}}$  =  $108 \pm 1.5 \text{ s}^{-1}$  for the wild-type enzyme to  $1.16 \pm 0.08 \text{ s}^{-1}$  for the mutant.<sup>31</sup> We will analyze this change in reactivity due to the mutation of this “spectator residue” and demonstrate its impact on the mechanism.

This paper is organized as follows: In the next section, we describe the applied computational methods. In the following main section, we present and discuss the QM/MM results for the three topics outlined above. We conclude with a brief summary in the final section.

## 2. Computational Methods

**2.1. System Preparation and Classical Simulations.** To generate starting structures for the oxo and selenido forms of the cofactor and for the Glu802 → Gln mutant, we manually made the corresponding changes in the optimized structures from our previous study<sup>15</sup> and reoptimized them. For 2-oxo-6-methylpurine, we started from the previous solvated setup before the final 500 ps MD run,<sup>15</sup> manually replaced the substrate, and then performed a QM/MM optimization (with the QM region containing part of the cofactor, Glu1261, and the substrate). The resulting starting structures were relaxed by classical 15 ps MD simulations with some constraints: We defined an active region including all residues within 20 Å of the C8 atom of the substrate bound to the molybdopterin cofactor of chain A. Within this region, we kept the  $[\text{Mo}(\text{S}_2\text{C}_2\text{H}_2)(=\text{O})(\text{OR}_{\text{unfixed}})(-\text{SH})]^{2-}$  moiety and the  $\text{Fe}_2\text{S}_2$  cluster fixed during these simulations, while on the outside part of the water sphere (35 Å around the substrate C8 atom) we imposed a spherical



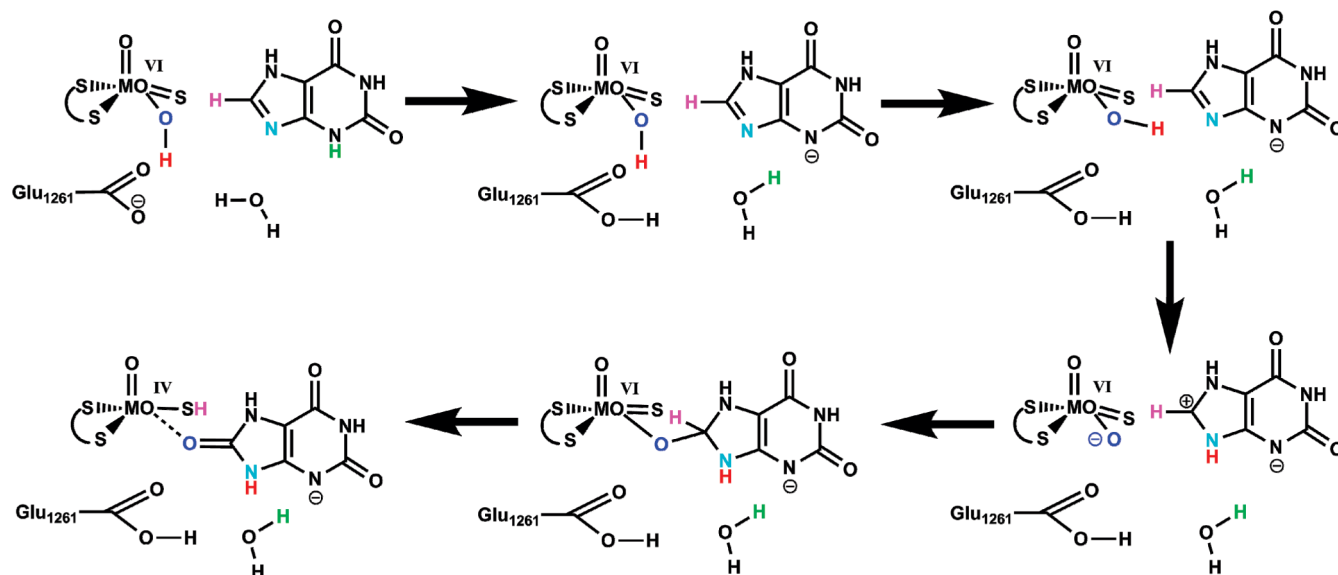
**Figure 1.** System used in the QM/MM calculations.

potential to prevent water molecules from escaping into the vacuum. A final 500 ps MD simulation at a temperature of 300 K using the CHARMM force field<sup>32</sup> as implemented in the CHARMM program package,<sup>33</sup> without imposing additional restraints, concluded the preparation of the enzyme–substrate complex.

As commonly accepted<sup>6–9,11,12</sup> and in analogy to our previous QM/MM work on XO<sup>15</sup> and AOR,<sup>14</sup> the substrate and Glu802 (Gln802) were neutral, Arg880 was protonated, and Glu1261 was deprotonated in all calculations. We adopted the same MM parameters as in our previous work<sup>14,15,34</sup> for the substrate, the molybdopterin cofactor, and the  $\text{Fe}_2\text{S}_2$  cluster. All nonstandard parameters are documented in the Supporting Information.

**2.2. QM/MM Setup.** We used the same neutral QM/MM system as in our previous XO study<sup>15</sup> (see Figure 1). QM/MM geometry optimizations of the stationary points were performed with a linear scaling microiterative algorithm working in hybrid delocalized coordinates.<sup>35</sup> All residues and water molecules within 13 Å of the substrate were included in the optimization; the remaining atoms were kept fixed. All QM/MM calculations were done with the modular program package ChemShell,<sup>36</sup> using TURBOMOLE<sup>37</sup> to obtain the QM(DFT) wave functions as well as the corresponding energies and gradients. MM energies and gradients were evaluated by DL\_POLY,<sup>38</sup> which is provided by the ChemShell package, using the CHARMM topology and parameter data. We used electrostatic embedding for the QM region.<sup>39</sup> No cutoff was applied on the rigid MM point charges, neither to calculate the electrostatic interaction within the MM region nor to calculate the electrostatic QM/MM interaction. To prevent overpolarization at the QM/MM boundary, we applied the charge-shift scheme.<sup>40,41</sup>

Different QM regions were employed in the QM/MM calculations which will be defined below on a case-by-case basis. The largest QM region included part of the cofactor, the substrate, and parts of the side chains of the Gln767, Glu802, Arg880, and Glu1261 residues and was used to perform

**SCHEME 2: Favored Mechanism for the Reaction of Xanthine within XO as Obtained in Our Previous QM/MM Study<sup>15</sup>**

COSMO<sup>42</sup> corrected single-point calculations to establish a common energy scale for the different setups.

Optimizations were generally first done using the BP86<sup>43–47</sup> functional with the resolution-of-the identity (RI) approximation<sup>48,49</sup> and then refined using the B3LYP hybrid functional<sup>43–45,50–52</sup> as implemented in TURBOMOLE. In the case of 2-oxo-6-methylpurine, we started from the product-bound structure and first determined the various minima along the reaction path. Easily detectable transition structures were localized starting from a preoptimized structure obtained by constrained minimization. In difficult cases, the nudged elastic band method<sup>53</sup> as implemented in DL-Find<sup>54</sup> was applied in combination with transition state optimization to find the minimum energy pathway as well as the transition state. The QM model calculations were performed using Gaussian 03.<sup>55</sup> All figures showing molecular structures were generated using VMD.<sup>56</sup>

Molybdenum was described by the Lan12DZ basis set<sup>57</sup> augmented by an f polarization function,<sup>58</sup> sulfur and selenium by Lan12DZ<sup>59</sup> with an additional d polarization function,<sup>60</sup> and all other atoms (C, H, N, O) by the 6-31+G\*\* basis set.<sup>61,62</sup> The RI-BP86 calculations employed the def2-TZV(P) auxiliary basis set.<sup>63</sup> This combination has already proven to give reasonable results.<sup>14,15,21</sup>

### 3. Results and Discussion

In a previous QM/MM study,<sup>15</sup> we investigated the conversion of xanthine to uric acid in XO using seven different setups that covered different tautomers and different orientations of xanthine as well as different protonation states of active-site residues. For the kinetically active “upside down” orientation (setup G),<sup>15</sup> the reaction starts with Glu1261 deprotonating xanthine at the N3 position followed by a proton transfer from the cofactor to the N9 atom of xanthine. The thus activated cofactor and substrate then react to form a tetrahedral intermediate, and a subsequent rate-limiting hydride transfer generates the product. The intermediates of this mechanism are sketched in Scheme 2, while the complete pathway is given in the Supporting Information (Figure S1). This most favorable pathway for xanthine in the wild-type enzyme (setup G) will serve as a reference mechanism in the present study of variants of the

previously studied system. In addition, we shall also consider a closely related pathway (setup F) and one conformation with the opposite “upside” orientation of the substrate (setup A) which was previously found to represent the thermodynamically most stable reactant complex.<sup>15</sup>

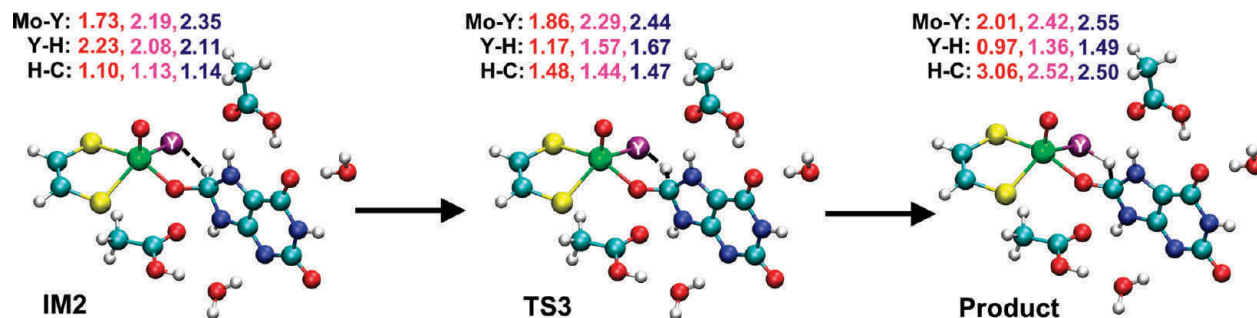
#### 3.1. Influence of the Hydride Acceptor in the Cofactor.

We modified the cofactor in XO and reoptimized the stationary points of the most favorable setup G (SN400) from our previous study to analyze the origin of the experimentally observed deactivation of XO upon replacement of the sulfido ligand by an oxo ligand<sup>26</sup> and to monitor the influence of a selenido ligand, that is present in the related nicotinate dehydrogenase.<sup>27</sup> We used the same QM region as before; see Figure 2. It is obvious from Table 1 that there is hardly any effect of the changed hydride acceptor on the relative energies of the various stationary points and the computed barriers, except for the hydride transfer reaction IM2 → TS3 → product, which shows a dramatically increased barrier for the oxido ligand and a slightly lower barrier when sulfur is replaced by selenium.

The lowering of the barrier for Se has been attributed to the smaller HOMO–LUMO gap between the Mo=Se  $\pi$  and  $\pi^*$  orbitals which should facilitate the rate-limiting hydride transfer step.<sup>27</sup> Indeed, we calculate the HOMO–LUMO gap to decrease from 2.35 eV (O) to 2.02 eV (S) and 1.95 eV (Se) in the intermediate IM2. We note, however, that the computed orbitals are quite delocalized in all relevant stationary points (IM2, TS3, product). Hence, there is no clear-cut  $\pi$  or  $\pi^*$  character, and neither the HOMO nor the LUMO orbital have significant electron density at the transferred H-atom (see Figure S2 in the Supporting Information).

An alternative explanation for the enhanced reactivity when going from the oxo to the sulfido and selenido form of the cofactor can be based on geometrical and energetic arguments. It is obvious that larger structural distortions of IM2 are needed to reach TS3 in the case of the oxo form (see Figure 2). For example, elongating the strong Mo=O bond by 0.13 Å is expected to cost more energy than stretching the weaker Mo=S and Mo=Se bonds by about 0.10 Å. Similar remarks apply to the changes in the Y–H and C–H distances. Looking from the product side, the H–Y bonds are longer by about 0.2 Å in TS3 which is again more difficult to achieve for H–O than for





**Figure 2.** Hydride transfer step with key distances (Å) at the B3LYP/B1 level, for different forms of the cofactor: Y = O (red), Y = S (pink), and Y = Se (purple).

**TABLE 1: QM/MM Energies (in kcal mol<sup>-1</sup>) for the Conversion of Xanthine to Uric Acid, with the Cofactor of XO in the Oxo, Sulfido, and Selenido Form<sup>a</sup>**

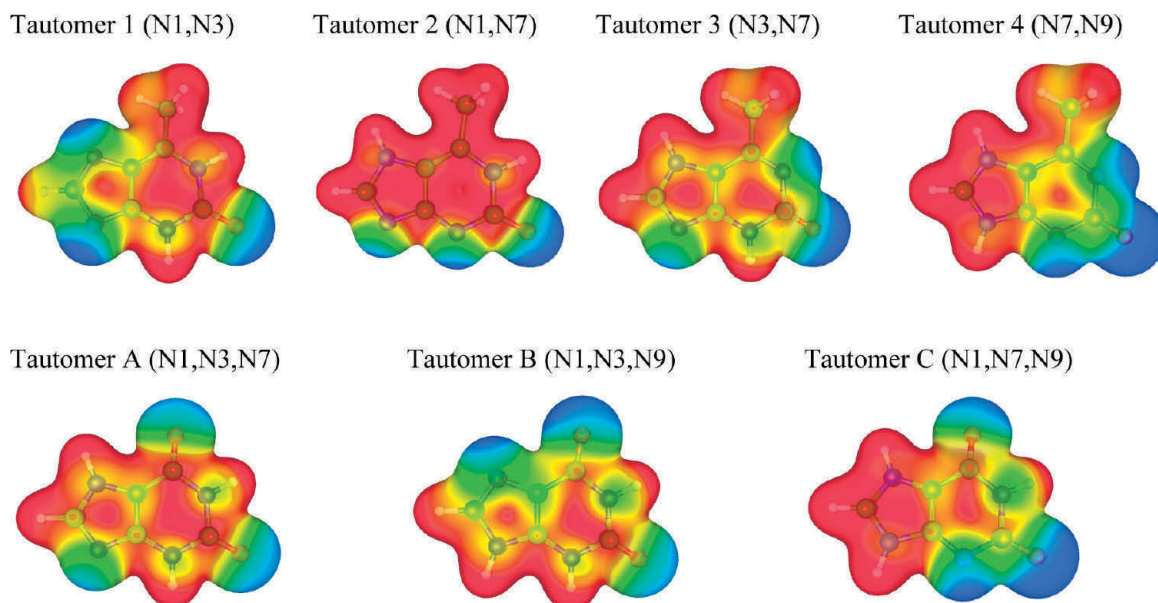
	oxo				sulfido				selenido			
	BP86		B3LYP		BP86		B3LYP		BP86		B3LYP	
reactant	0.0		0.0		0.0		0.0		0.0		0.0	
TS1'	2.5	(2.5)	5.6	(5.6)	2.9	(2.9)	8.5	(8.5)	3.1	(3.1)	5.2	(5.2)
IM1'	-7.4		-7.2		-8.0		-7.5		-8.2		-9.6	
TS1''	-6.9	(0.5)	-7.1	(0.1)	-7.4	(0.6)	-7.2	(0.3)	-7.1	(1.1)	-9.4	(0.2)
IM1''	-8.9		-9.5		-9.0		-8.6		-8.7		-10.7	
TS1	1.8	(10.7)	1.6	(11.1)	1.0	(10.0)	2.5	(11.1)	0.4	(9.1)	-1.3	(9.4)
IM1	-1.1		0.7		-3.4		-1.8		-3.1		-3.8	
TS2	-0.5	(0.6)	2.6	(1.9)	-2.1	(1.3)	2.5	(4.3)	-1.0	(2.1)	0.2	(4.0)
IM2	4.7		0.9		-7.8		-2.5		-6.6		-5.7	
TS3	19.6	(14.9)	21.2	(20.3)	0.7	(8.5)	4.5	(7.0)	0.6	(7.2)	-0.8	(4.9)
product	6.2		-4.4		-0.8		-9.3		-0.8		-10.7	
$\Delta E_{\max}$	28.5		30.6		10.0		13.1		9.3		10.9	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

H-S or H-Se. Overall, TS3 resembles the energetically favored product state more for Y = Se than for Y = O.

**3.2. Influence of the Substrate: 2-Oxo-6-methylpurine.** In the case of xanthine, the reactivity is strongly dependent on the tautomeric form of the substrate.<sup>15</sup> For 2-oxo-6-methylpurine, we investigated four different tautomers that differ in the protonation pattern of the nitrogen atoms of the purine ring; see Figure 3. The electrostatic potential plots for xanthine and

2-oxo-6-methylpurine are quite similar, in spite of the fact that the nitrogen atoms in neutral xanthine carry three protons, instead of two in 2-oxo-6-methylpurine, and regardless of the replacement of the negatively charged O6 atom in the former (acting as a hydrogen bond acceptor) by an essentially neutral methyl group in the latter. For example, tautomer 3 of 2-oxo-6-methylpurine has an electrostatic potential map analogous to tautomer A of xanthine, while tautomer 4 resembles the active



**Figure 3.** Electrostatic potential plots for different tautomers of 2-oxo-6-methylpurine (top panel, tautomers 1-4) and xanthine (bottom panel, tautomers A-C). The electrostatic potential is mapped on a density isosurface (0.01e). Red (blue) indicates a positive (negative) potential.

**TABLE 2: Relative Energies of Tautomers 1–4 of 2-Oxo-6-methylpurine in kcal mol<sup>-1</sup>**

	$\Delta E^a$	$\Delta E^b$	$\Delta E^c$	$\Delta E(\text{COSMO})^d$
tautomer 1	0.0	0.0	0.0	0.0
tautomer 2	11.4	11.2	10.6	0.7
tautomer 3	4.7	4.5	4.1	0.2
tautomer 4	not calculated	33.7	33.0	12.2

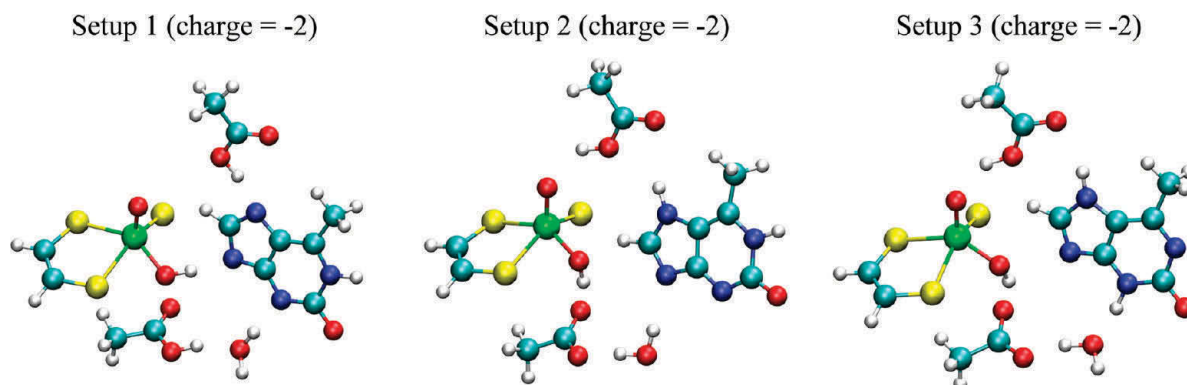
<sup>a</sup> B3LYP/6-31G(d). <sup>b</sup> B3LYP/6-31+G\*\*. <sup>c</sup> B3LYP/def2-TZVP. <sup>d</sup> B3LYP/def2-TZVP,  $\epsilon = 80$ .

xanthine species with a doubly protonated imidazole substructure (tautomer C). Again, as in the case of xanthine,<sup>15</sup> the energies of the four tautomers of 2-oxo-6-methylpurine show a large spread in the gas phase but lie much closer to each other in a continuum (COSMO) solvent where tautomers 1–3 are within 1 kcal mol<sup>-1</sup>; see Table 2. We therefore computed reaction paths for all of these tautomers. The standard QM regions of the starting structures are depicted in Figure 4; the larger QM regions also include Gln767 and Arg880 in the QM part.

**Setup 1 (See Figure S3 in the Supporting Information).** Tautomer 1 is the most stable tautomer in gas-phase calculations. In the enzyme, the lack of a proton at N7 causes coordination by the protonated Glu802 residue. Having two protons at the pyrimidine ring, tautomer 1 is not stable as a neutral species in the binding pocket but immediately transfers its N3 proton to Glu1261 which thus becomes less acidic. Consequently, the hydroxo group of the cofactor rearranges and establishes a hydrogen bond to N9 of the substrate; see Figure 4. Starting from this reactant state, we could not locate an intermediate in which the proton of the cofactor hydroxyl group is transferred to the substrate without the C–O bond being formed. Instead, these two changes happen at the same time so that we directly obtain a tetrahedral structure as the first intermediate (IM2), which bears a proton at N9 (not at N7), establishing a hydrogen bond with OE1 of Glu1261. The second step is a formal hydride transfer of H8 to the sulfido group of the cofactor.

For both snapshots considered, we generally find good agreement in the computed geometries and relative energies: Both the BP86 and B3LYP functionals assign TS2 as the rate-limiting barrier, although TS3 lies only slightly below TS2. The barrier is somewhat lower for BP86 (about 18 kcal mol<sup>-1</sup>) than for B3LYP (about 23 kcal mol<sup>-1</sup>, rising to 25.8 kcal mol<sup>-1</sup> for the large QM region); see Table 3. These values are higher than the barrier derived from experimental data: Using the Eyring equation<sup>65</sup> to convert the reported values for the rate constant of  $k_{\text{red}} = 57 \text{ s}^{-1}$ ,<sup>13</sup> the resulting experimental barrier is 14.8 kcal mol<sup>-1</sup>.

**Setup 2 (See Figure S4 in the Supporting Information).** Using the standard QM region, the reactant state of the (N1, N7) tautomer 2 can be optimized in the enzyme only for snapshot SN400, whereas it rearranges spontaneously to the intermediate IM1 in the case of snapshot SN200. This rearrangement involves a double proton transfer: one proton is shifted from Glu802 to the oxo ligand of the cofactor to form an apical hydroxyl group, and another proton from the equatorial hydroxyl group is transferred to Glu1261. This supports the suggestion<sup>24</sup> that deprotonation of the cofactor is induced by substrate binding. Note that the initial rearrangement differs slightly in SN400 with the large QM region where Gln767 is a QM residue and can thus act as a hydrogen bond acceptor. Considering the reaction pathways from IM1 onward, there are again slight differences between the two snapshots. In the course of C–O bond formation, the proton is transferred back in SN200 from the apical hydroxyl group of the cofactor directly to Glu802, whereas this happens in two distinct steps in SN400. However, the overall transformation is the same in both snapshots. The last step is a combined proton-hydride transfer, in which the hydrogen atom H8 of the substrate migrates to the sulfido group of the cofactor while the proton at N9 is concomitantly transferred back to Glu1261. The extent of proton transfer in the corresponding transition state depends on the functional used: the proton is already transferred to Glu1261 with BP86 (OE2–H bond length: 1.03 Å), whereas it still resides

**Figure 4.** Standard QM regions of the three chosen setups for 2-oxo-6-methylpurine: reactant states of tautomers 1–3.**TABLE 3: QM/MM Energies of Different Snapshots in kcal mol<sup>-1</sup> Calculated for Setup 1 of 2-Oxo-6-methylpurine, Relative to the Energy of the Reactant<sup>a</sup>**

	SN200-BP86	SN200-B3LYP	SN400-BP86	SN400-B3LYP	SN400-B3LYP, large QM
reactant	0.0	0.0	0.0	0.0	0.0
TS2	18.1 (18.1)	22.7 (22.7)	17.6 (17.6)	23.4 (23.4)	25.8 (25.8)
IM2	12.2	17.7	10.4	16.0	20.0
TS3	17.6 (5.4)	22.6 (4.9)	15.6 (5.2)	21.7 (5.7)	24.3 (4.3)
product	13.7	7.5	11.1	6.0	8.1
$\Delta E_{\text{max}}$	18.1	22.7	17.6	23.4	25.8

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

**TABLE 4: QM/MM Energies of Different Snapshots in kcal mol<sup>-1</sup> Calculated for Setup 2 of 2-Oxo-6-methylpurine, Relative to the Energy of IM1 (Relative to the Reactant for the Large QM Region)<sup>a</sup>**

	SN200-BP86		SN200-B3LYP		SN400-BP86		SN400-B3LYP		SN400-B3LYP, large QM region	
reactant					4.4		3.9		0.0	
TS1					5.8	(1.4)	7.3	(3.4)	5.8	(5.8)
IM1	0.0		0.0		0.0		0.0		1.4	
TS2	8.7	(8.7)	9.8	(9.8)	8.5	(8.5)	11.6	(11.6)	7.3	(5.9)
IM2	-0.5		-1.1		0.2		1.0		1.9	
TS2'					2.2	(2.0)	4.6	(3.6)	6.6	(4.7)
IM2'					0.7		3.3		1.8	
TS3	5.4	(5.9)	9.3	(10.4)	5.1	(4.4)	11.2	(7.9)	7.0	(5.2)
product	4.3		-5.7		1.3		-7.7		-7.4	
$\Delta E_{\max}$	8.7		10.4		8.5		11.6		7.3	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.**TABLE 5: QM/MM Energies of Different Snapshots in kcal mol<sup>-1</sup> Calculated for Setup 3 of 2-Oxo-6-methylpurine, Relative to the Energy of IM1 (IM1' for the Large QM Region)<sup>a</sup>**

	SN200-BP86		SN200-B3LYP		SN400-BP86		SN400-B3LYP		SN400-B3LYP, large QM region	
reactant	9.4		8.3		12.1		9.1		8.9	
TS1'	12.2	(2.7)	13.3	(5.0)	12.7	(0.6)	9.5	(0.4)	11.3	(2.4)
IM1'	3.9		1.8		3.9		1.0		0.0	
TS1	4.6	(0.7)	5.5	(3.7)	5.8	(1.9)	6.5	(5.5)	6.5	(6.5)
IM1	0.0		0.0		0.0		0.0		0.9	
TS2	6.7	(0.3)	9.6	(2.5)	5.9	(0.4)	7.4	(1.5)	9.8	(8.9)
IM2	1.1		3.9		0.6		2.0		5.9	
TS3	8.1	(7.0)	8.9	(5.0)	6.7	(6.1)	7.0	(5.0)	11.3	(5.4)
product	7.9		-1.6		4.4		-6.6		-2.8	
$\Delta E_{\max}$	8.1		8.9		6.7		7.4		11.3	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

at the substrate with B3LYP (N9–H bond length: 1.05 Å), which gives a higher barrier for this step. Looking at the overall reaction profile, the barriers for C–O bond formation and hydride transfer are rather similar in the case of B3LYP, the former one being slightly higher and thus rate-determining. Compared with setup 1, the computed barriers for the individual reaction steps are significantly lower in setup 2. This holds especially for the B3LYP results with the large QM region where the highest point on the reaction profile drops to 7.3 kcal mol<sup>-1</sup> (see Table 4), i.e., even below the experimental barrier.

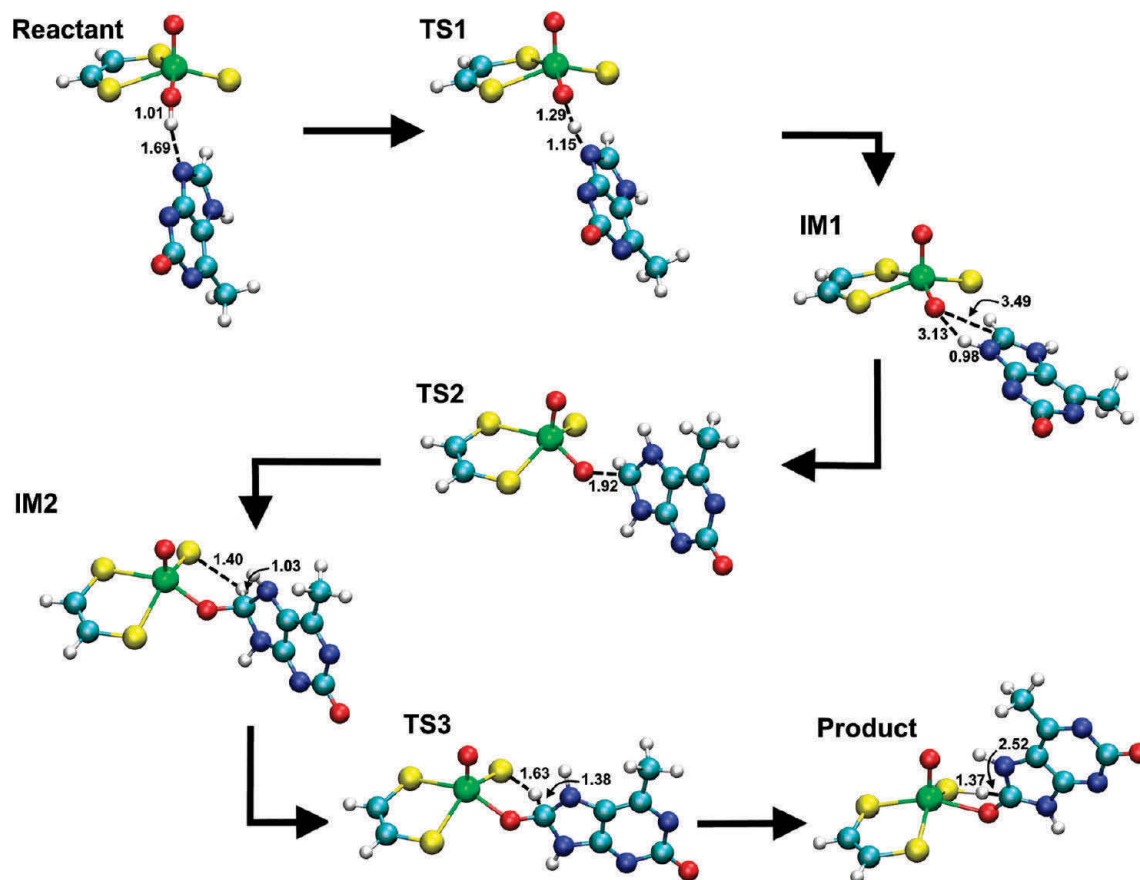
**Setup 3 (See Figure S5 in the Supporting Information).** Initially, the (N3, N7) tautomer 3 is converted into the (N7, N9) tautomer 4 as follows. The proton at N3 is transferred via the crystal water molecule CRYW224 toward Glu1261, and at the same time, the hydroxyl group of the cofactor reorients and establishes a hydrogen bond with the N9 atom of 2-oxo-6-methylpurine. Thereafter, the proton of this hydroxyl group is transferred to the N3 atom of the substrate, and concomitantly, the apical oxygen atom of the cofactor becomes strongly acidic and captures the proton from Glu802 to form an apical hydroxyl group. The thus formed activated tautomer 4 reacts with the cofactor under C–O bond formation, and the resulting tetrahedral intermediate then undergoes a hydride transfer that generates the product.

Both functionals assign IM1 as the Michaelis complex, but they show minor differences in the reaction profiles for C–O bond formation and hydride transfer: For the standard QM region, the latter (former) is assigned to be rate-limiting with BP86 (B3LYP). This can be rationalized by the stronger stabilization of the Mo(IV) product in B3LYP compared with BP86 (by 7–8 kcal mol<sup>-1</sup>) which leads to an “earlier” transition state for the hydride transfer (with a slightly lower energy). The key C8–H8 distance in the transition state is 1.60–1.65 Å in

the two BP86 snapshots and 1.42 Å in both B3LYP snapshots. However, one should emphasize that the energetic differences are small (less than 2 kcal mol<sup>-1</sup>, see Table 5) and thus not very relevant. According to B3LYP, the highest point on the reaction profile is at 7.4–9.6 kcal mol<sup>-1</sup> for the two chosen snapshots (rising to 11.3 kcal mol<sup>-1</sup> for SN400 with the larger QM region, see Table 5).

**QM Model Study (Figure 5).** To examine the intrinsic reactivity of 2-oxo-6-methylpurine compared to xanthine, we performed a QM study on a minimal gas-phase model consisting of parts of the cofactor and the substrate; see Table 6. For both substrates, we find a three-step mechanism, involving a proton transfer (via TS1), C–O bond formation (via TS2), and a final hydride transfer (via TS3). Considering the rate-limiting final step, the barrier relative to the preceding minimum (IM2) is lower for xanthine, but the overall barrier  $\Delta E_{\max}$  is lower for 2-oxo-6-methylpurine due to the stronger stabilization of IM2 by 4.6 kcal mol<sup>-1</sup>. It should, however, be noted that the QM and QM/MM mechanisms differ, since the QM model system does not include Glu802, so that it is impossible to form an apical OH group.

**Summary.** The overall barrier for setup 1 is considerably larger than those for setups 2 and 3, and also much too high compared with experiment.<sup>13</sup> Hence, tautomer 1 of 2-oxo-6-methylpurine (which is the most stable one in the gas phase) can be excluded as the reactive species in the enzyme, in analogy to the case of xanthine. The overall barriers for setups 2 and 3 are of comparable size, i.e., slightly lower for 2 (3) when using the large (standard) QM region, so that both pathways appear feasible. The overall B3LYP/MM barriers for the various snapshots in these two setups lie between 7.3 and 11.6 kcal mol<sup>-1</sup> and are thus somewhat lower than the value of 14.8 kcal mol<sup>-1</sup> derived from experiment (see above). Compared with



**Figure 5.** Reaction mechanism obtained from the QM model study of 2-oxo-6-methylpurine.

**TABLE 6: QM Energies for the Gas-Phase Calculations on Setup 3 of 2-Oxo-6-methylpurine in kcal mol<sup>-1</sup>, Relative to the Energy of the Reactant<sup>a</sup>**

	2-oxo-6-methylpurine		xanthine <sup>15</sup>	
reactant (IM1' <sup>15</sup> )	0.0		0.0	
TS1	4.6	(4.6)	5.5	(5.5)
IM1	3.0		5.4	
TS2	11.2	(8.2)	15.4	(10.0)
IM2	6.2		10.8	
TS3	13.4	(7.2)	16.8	(6.0)
product	-17.4		-13.3	
$\Delta E_{\text{max}}$	13.4		16.8	

<sup>a</sup> Activation barriers relative to the preceding minima are given in parentheses.

xanthine (B3LYP/MM barriers of 13–15 kcal mol<sup>-1</sup>), 2-oxo-6-methylpurine should thus be oxidized more easily by XO, consistent with the calculated intrinsic gas-phase reactivity of these two substrates (see above). Experimentally, the catalytic conversion of 2-oxo-6-methylpurine in XO is only slightly faster than that of xanthine, by a factor of about 10. We note again in this context that 2-oxo-6-methylpurine is considered a “slow” substrate only because of the sluggish product release.<sup>13</sup>

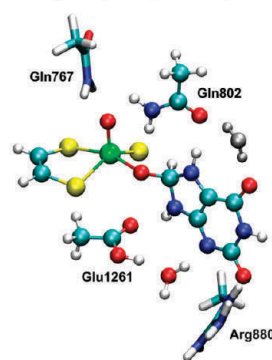
**3.3. Influence of the Glu802 → Gln Mutation.** Our recent QM/MM study<sup>15</sup> on the conversion of xanthine to uric acid in xanthine oxidase showed the essential mechanistic role of Glu1261 and Arg880, for activating the cofactor and for stabilizing the crucial intermediates and transition states, respectively, whereas Glu802 mainly acted as a spectator on the favored pathway (setup G). On the other hand, Glu802 strongly affects the reactivity when hypoxanthine is used as a substrate,<sup>12</sup> and it may have some influence on the different

orientations observed for xanthine and 2-oxo-6-methylpurine in the binding pocket.<sup>8,9,17</sup>

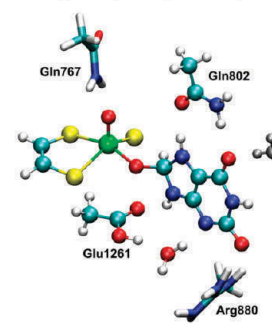
In the case of xanthine, the pathway with an “upside” orientation of the substrate and deprotonated Glu802 was found to have a barrier of at least 40 kcal mol<sup>-1</sup> which is lowered to about 20 kcal mol<sup>-1</sup> in the presence of protonated Glu802.<sup>15</sup> In the latter case, the substrate is able to change from the (N1, N3, N7) to the (N1, N7, N9) tautomer by using Glu802 as a proton shuttle.<sup>15</sup> This functionality should vanish for a Glu802 → Gln mutation which we investigate here.

Recently published kinetic data for the corresponding mutant of XDH show the overall rate constant to drop from  $k_{\text{cat}} = 108$

**Setup F (charge = -1)**

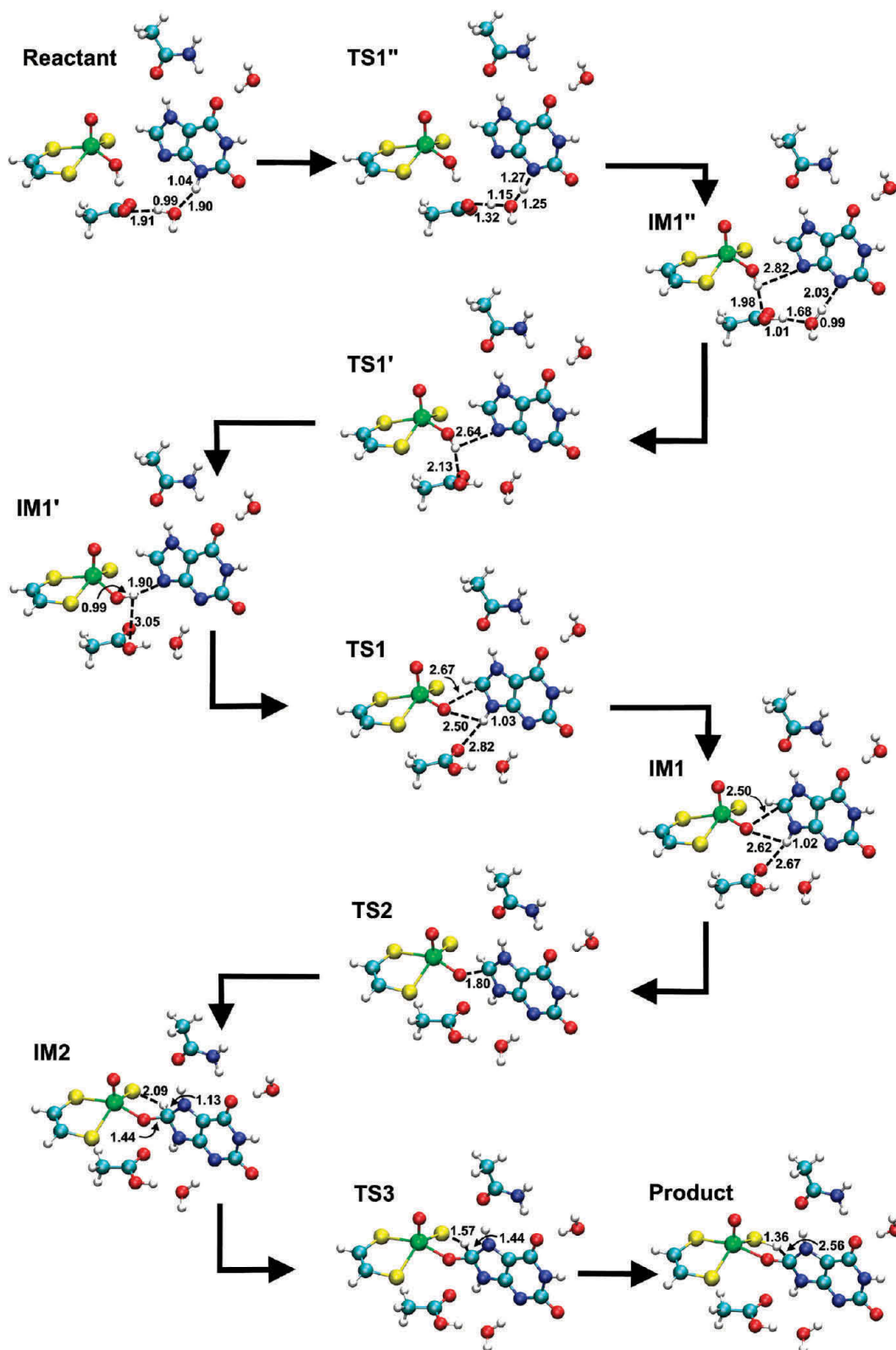


**Setup G (charge = -1)**



**Figure 6.** QM regions of the two chosen setups F and G for the mutant. For consistency reasons, we use the same notation as in our previous publication.<sup>15</sup> The standard QM region is shown in ball-and-stick representation, whereas the large QM region also contains Gln767 and Arg880 but excludes the water molecule colored in gray.





**Figure 7.** Reaction mechanism for the oxidation of xanthine in the mutant: setup G (SN400).

$\pm 1.5 \text{ s}^{-1}$  for the wild-type enzyme to  $1.16 \pm 0.08 \text{ s}^{-1}$  for the mutant.<sup>31</sup> This decrease by a factor of 90 implies an increase in the Eyring activation energy of  $2.5 \text{ kcal mol}^{-1}$ . We examine whether we find a similar effect of this mutation in XO, using two different setups F and G (see Figure 6, same notation as in our previous publication<sup>15</sup>).

Since setup G had been preferred in the wild-type enzyme, we used it as an obvious starting point for the mutant and indeed

obtained essentially the same reaction pathway; see Figure 7. The barriers for the individual steps are similar in both cases, but the intermediates IM1 and IM2 are destabilized in the mutant, by 2.7 and 5.3  $\text{kcal mol}^{-1}$  relative to the wild-type enzyme (see Table 7), which leads to an increase of the overall barrier to  $17.5 \text{ kcal mol}^{-1}$ . For the large QM region, IM2 is further destabilized by another  $3.6 \text{ kcal mol}^{-1}$ , and the overall barrier becomes  $22.0 \text{ kcal mol}^{-1}$ . We crosschecked and con-



**TABLE 7: QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup G (SN 400), Relative to the Energy of Intermediate IM1'<sup>a</sup>**

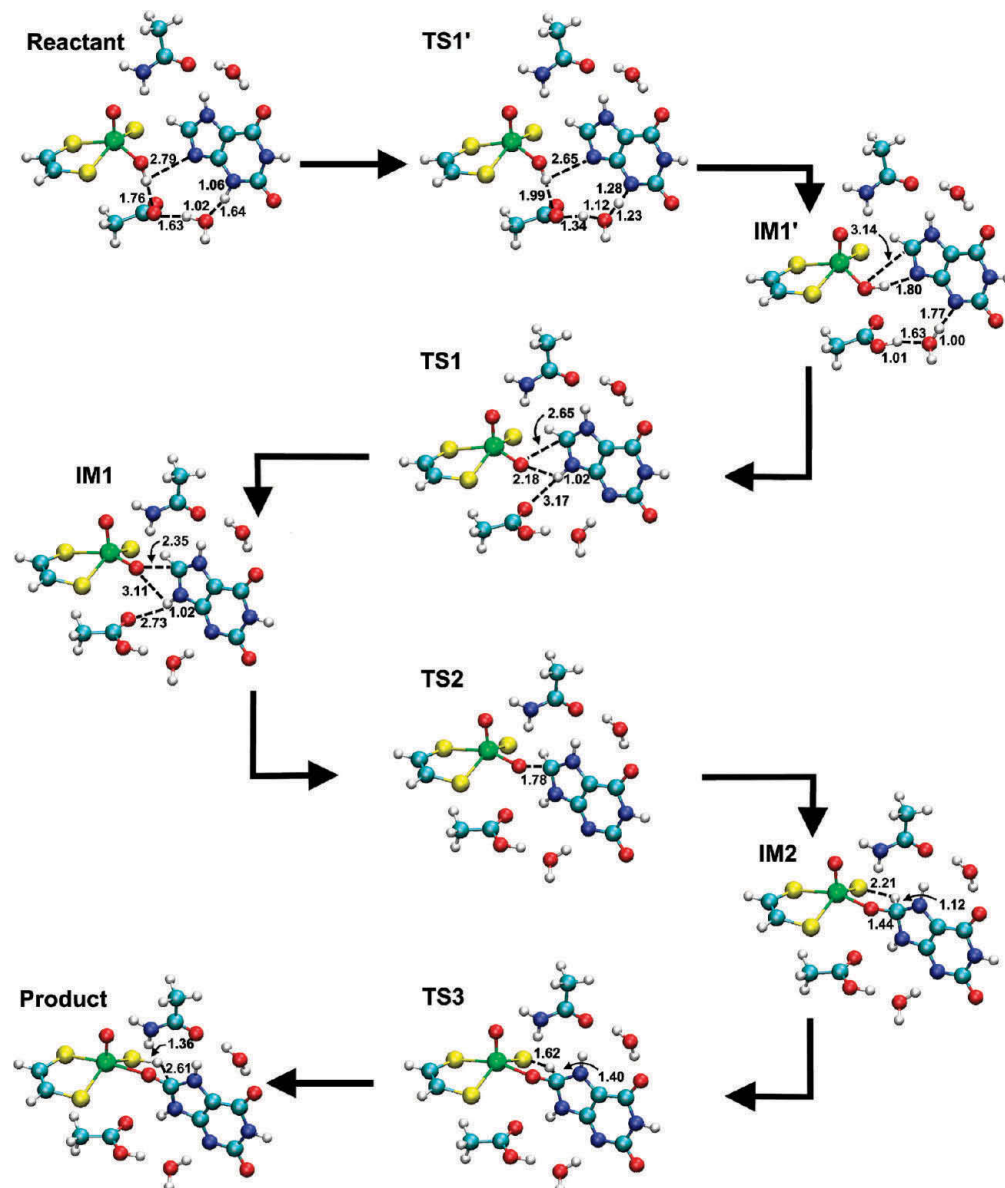
	SN400		SN400, Gln802		SN400, Gln802, large QM	
reactant	8.7		9.1		9.0	
TS1''	17.2	(8.5)	16.0	(6.9)	16.7	(16.7)
IM1''	1.2		2.5		5.5	
TS1'	1.5	(0.3)	2.5	(0.0)	5.7	(0.2)
IM1'	0.0		0.0		0.0	
TS1	11.1	(11.1)	11.3	(11.3)	10.6	(10.6)
IM1	6.8		9.5		9.6	
TS2	11.1	(4.3)	15.1	(5.6)	18.3	(8.7)
IM2	6.2		11.5		15.1	
TS3	13.2	(7.0)	17.5	(6.0)	22.0	(6.9)
product	-0.6		2.7		7.6	
$\Delta E_{\max}$	13.1		17.5		22.0	

<sup>a</sup> The three columns give the values for the wild-type enzyme, the Glu802  $\rightarrow$  Gln mutant using the standard QM region, and the mutant using the large QM region, respectively. Activation barriers relative to the preceding minima are given in parentheses.

firmed these results for snapshot SN100 where we obtained overall barriers of 18.2 (24.2) kcal mol<sup>-1</sup> for the standard (large) QM region.

In the mutant, the computed barrier for setup G is thus significantly higher than that in the wild-type enzyme, and it also considerably exceeds the experimental value. The high barrier is mainly due to the destabilization of IM2, caused by an accumulation of charge on O6. This could be alleviated by a rotation of Gln802 (as in setup F) that would allow a water molecule to coordinate to O6 and thus provide some stabilization.

Therefore, we investigated setup F as an alternative. We again find very similar reaction pathways for the wild-type enzyme and the mutant; see Figure 8. There is, however, one essential difference: The intermediate IM1 is much less stable in the mutant compared with the wild-type enzyme, by more than 11 kcal mol<sup>-1</sup>, since the stabilizing proton transfer (from Glu802) is no longer possible in the Glu802  $\rightarrow$  Gln802 mutant. The potential surface is thus generally flatter in the mutant, with an overall barrier of 15.8 (17.1) kcal mol<sup>-1</sup> for the standard (large) QM region in snapshot SN400. The corresponding values for

**Figure 8.** Reaction mechanism for the oxidation of xanthine in the mutant: setup F (SN400).

**TABLE 8: QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup F (SN400), Relative to the Energy of the Intermediate IM1<sup>a</sup>**

	SN400		SN400, Gln802		SN400, Gln802, large QM	
reactant	9.4		9.1		7.7	
TS1'	13.6	(4.2)	12.0	(2.9)	11.8	(4.1)
IM1'	0.0		0.0		0.0	
TS1	9.0	(9.0)	9.8	(9.8)	9.6	(9.6)
IM1	-3.7		7.7		7.6	
TS2	15.4	(19.1)	13.1	(5.3)	14.9	(7.3)
IM2	11.4		10.3		10.4	
TS3	16.6	(5.2)	15.8	(5.5)	17.1	(6.7)
product	0.3		-3.7		-2.9	
$\Delta E_{\max}$	20.3		15.8		17.1	

<sup>a</sup> The three columns give the values for the wild-type enzyme, the Glu802 → Gln mutant with the standard QM region, and the mutant with the large QM region, respectively. Activation barriers relative to the preceding minima are given in parentheses.

snapshot SN100 are somewhat higher, i.e., 19.6 (21.3) kcal mol<sup>-1</sup> (see Table 8).

Following the same procedure as in our previous QM/MM study,<sup>15</sup> we performed cluster QM calculations<sup>66–69</sup> in an attempt to establish to a common energy scale for the different setups and snapshots (including all residues shown in Figure 6 in the QM cluster and representing the environment of the cluster by a polarizable continuum with a dielectricity constant of  $\epsilon = 8$ ). According to these cluster calculations, the reactants of the closely related setups F and G are essentially isoenergetic and should thus both be populated at room temperature. Hence, the results for these two setups are directly comparable with each other. For the standard QM region, the preferred pathway in the Glu802 → Gln mutant is setup F with an overall B3LYP/MM barrier of 15.8 kcal mol<sup>-1</sup> for the conversion of xanthine to uric acid, while the corresponding best B3LYP/MM value (analogous QM region) in the wild-type enzyme is 13.1 kcal mol<sup>-1</sup> (setup G). In the case of XO, the calculations thus predict an increase in the barrier of 2.7 kcal mol<sup>-1</sup> upon mutation, in

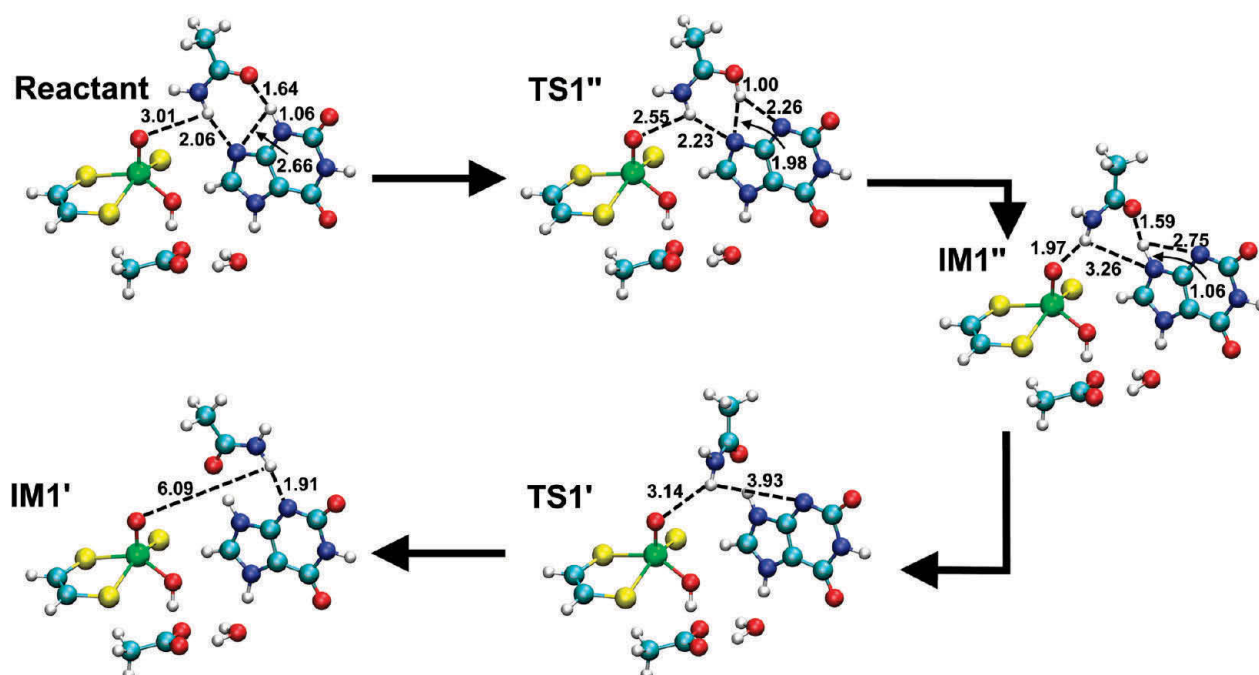
**TABLE 9: QM/MM Energies in kcal mol<sup>-1</sup> Calculated for Setup A, Relative to the Energy of the Reactant<sup>a</sup>**

	SN500		SN500, Gln802		SN500, Gln802, large QM	
reactant	0.0		0.0		0.0	
TS1''			20.0	(20.0)	19.9	(19.9)
IM1''			7.1		10.0	
TS1'	8.9	(8.9)	15.9	(8.8)	17.6	(7.6)
IM1'	2.4		3.7		4.9	
TS1	19.1	(16.7)				
IM1	14.9		18.1		18.1	
TS2	20.3	(5.4)				
IM2	16.4		19.4		21.0	
TS3	21.7	(5.3)	23.9	(4.5)	25.7	(4.7)
product	0.5		3.7		3.7	
$\Delta E_{\max}$	21.7		23.9		25.7	

<sup>a</sup> The three columns give the values for the wild-type enzyme, the Glu802 → Gln mutant with the standard QM region, and the mutant with the large QM region. Activation barriers relative to the preceding minima are given in parentheses.

good accord with the experimental value of 2.5 kcal mol<sup>-1</sup> for XDH.<sup>31</sup> More important than this partly fortuitous numerical agreement is the qualitative conclusion that the mutation induces a change in the favored reaction channel from setup G (for the wild-type enzyme) to setup F (for the Glu802 → Gln mutant).

For setups F and G with an “upside down” orientation of the substrate, the initial proton transfer from the N3 to the N9 position is not affected by the mutation of Glu802. This is quite different in the previously proposed alternative “upside” orientation<sup>8,9,17</sup> where the proton shuttle is strongly affected by the Glu802 → Gln mutation. In the corresponding setup A, the proton transfer now occurs in two separate steps (see Figure 9). The proton at N3 first moves to N9, being bound to Gln802 in the transition state. In the resulting intermediate (IM1''), the amide group of Gln802 is coordinated to the cofactor, and a simple rotation of the Gln802 side chain is needed to affect coordination of the amide group to the substrate (thus generating the same bonding situation as in the wild-type enzyme). The barrier for this two-step proton transfer in the mutant (setup A)

**Figure 9.** Part of the reaction mechanism for the oxidation of xanthine in the mutant: setup A (SN100).

is quite high (20 kcal mol<sup>-1</sup> compared to 8.9 kcal mol<sup>-1</sup> in the wild-type enzyme). However, the rate-limiting step is still the hydride transfer reaction, and the influence on the overall barrier is relatively small (+2.2 kcal mol<sup>-1</sup>); see Table 9.

To summarize, the results for the Glu802 → Gln mutant confirm the “upside down” orientation of the substrate in the kinetically active species, since the computed overall barriers for setup F (“upside down”) are much lower than those for setup A (“upside”), typically by 8–9 kcal mol<sup>-1</sup>.

#### 4. Conclusions

We have extended our previous QM/MM study of the reductive half-reaction of xanthine oxidase by considering variations of the cofactor, the substrate, and the active-site Glu802 residue. The present calculations confirm that the enzyme becomes inactive when the sulfido ligand in the cofactor is replaced by an oxo ligand, while selenido substitution is predicted to increase the activity by lowering the overall barrier. For the alternative substrate 2-oxo-6-methylpurine, we determined the reaction mechanisms for three different tautomers (two snapshots each) and obtained overall barriers (B3LYP/CHARMM, large QM region) of 25.8, 7.3, and 11.3 kcal mol<sup>-1</sup>, respectively. Therefore, only the latter two pathways are viable which involve tautomers 2–4 rather than the lowest-energy gas-phase tautomer 1. Our results agree with the experimental finding that the conversion of the Michaelis complex to the product complex is more facile when using 2-oxo-6-methylpurine instead of xanthine as a substrate. QM model studies also indicate an intrinsically higher reactivity of 2-oxo-6-methylpurine compared with xanthine. In each case, the most favorable pathway involves the “upside down” orientation of the substrate in the binding pocket, contrary to previous claims that the differences in reactivity may be caused by different substrate orientations.

For the Glu802 → Gln mutant, we investigated two competing pathways with “upside down” substrate orientation that had also been studied previously for the wild-type enzyme (setups F and G). These pathways differ mainly in the orientation of the Glu802/Gln802 residue. The mutant and the wild-type enzyme prefer different mechanisms (F and G, respectively) for two reasons. First, setup F is disfavored in the wild-type enzyme, since the intermediate IM1 acts as a thermodynamic sink, being stabilized by a proton transfer from Glu802 which is not possible in the mutant (containing Gln802), and therefore, the overall energy profile is flatter in the mutant and has a lower overall barrier. Second, setup G is disfavored in the mutant, because the negative charge developing at the O6 position during the reaction is less well stabilized by Gln802 compared with Glu802. The computed differences in the overall barriers for wild-type XO (setup G) and the Glu802 → Gln mutant (setup F) are in accord with corresponding experimental data for XDH. For the “upside” orientation of the substrate in the mutant (setup A), the barriers are calculated to be much higher (by 8–9 kcal mol<sup>-1</sup>), again in analogy to wild-type XO.<sup>15</sup>

In summary, the present results are consistent with the stepwise mechanism and the substrate orientation proposed in our previous QM/MM studies and with the available experimental results for the examined variations of cofactor, substrate, and active-site residue.

**Acknowledgment.** This work was supported by the Max-Planck-Gesellschaft. S.M. thanks the Fonds der Chemischen Industrie for a Kekulé scholarship.

**Supporting Information Available:** System setup; CHARMM parameters; QM, MM, and QM/MM energies of all stationary points; and complete refs 32, 36, and 55. This material is available free of charge via the Internet at <http://pubs.acs.org>.

#### References and Notes

- Hille, R. *Chem. Rev.* **1996**, *96*, 2757–2816.
- Schardinger, F. *Z. Unters. Nahr. Genussm.* **1902**, *5*, 1113–1121.
- Massey, V.; Harris, C. M. *Biochem. Soc. Trans.* **1997**, *25*, 750–755.
- Hille, R.; Sprecher, H. *J. Biol. Chem.* **1987**, *262*, 10914–10917.
- Romão, M. J.; Archer, M.; Moura, I.; Moura, J. J. G.; Legall, J.; Engh, R.; Schneider, M.; Hof, P.; Huber, R. *Science* **1995**, *270*, 1170–1176.
- Huber, R.; Hof, P.; Duarte, R. O.; Moura, J. J. G.; Moura, I.; Liu, M. Y.; LeGall, J.; Hille, R.; Archer, M.; Romão, M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8846–8851.
- Okamoto, K.; Matsumoto, K.; Hille, R.; Eger, B. T.; Pai, E. F.; Nishino, T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7931–7936.
- Pauff, J. M.; Hemann, C. F.; Jünemann, N.; Leimkühler, S.; Hille, R. *J. Biol. Chem.* **2007**, *282*, 12785–12790.
- Pauff, J. M.; Zhang, J. J.; Bell, C. E.; Hille, R. *J. Biol. Chem.* **2008**, *283*, 4818–4824.
- Kim, J. H.; Ryan, M. G.; Knaut, H.; Hille, R. *J. Biol. Chem.* **1996**, *271*, 6771–6780.
- Choi, E. Y.; Stockert, A. L.; Leimkühler, S.; Hille, R. *J. Inorg. Biochem.* **2004**, *98*, 841–848.
- Yamaguchi, Y.; Matsumura, T.; Ichida, K.; Okamoto, K.; Nishino, T. *J. Biochem.* **2007**, *141*, 513–524.
- McWhirter, R. B.; Hille, R. *J. Biol. Chem.* **1991**, *266*, 23724–23731.
- Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 4628–4640.
- Metz, S.; Thiel, W. *J. Am. Chem. Soc.* **2009**, *131*, 14885–14902.
- Mondal, M. S.; Mitra, S. *Biochemistry* **1994**, *33*, 10305–10312.
- Pauff, J. M.; Cao, H.; Hille, R. *J. Biol. Chem.* **2009**, *284*, 8760–8767.
- Truglio, J. J.; Theis, K.; Leimkühler, S.; Rappa, R.; Rajagopalan, K. V.; Kisker, C. *Structure* **2002**, *10*, 115–125.
- Bray, M. R.; Deeth, R. J. *J. Chem. Soc., Dalton Trans.* **1997**, 4005–4009.
- Voityuk, A. A.; Albert, K.; Romão, M. J.; Huber, R.; Rösch, N. *Inorg. Chem.* **1998**, *37*, 176–180.
- Zhang, X. H.; Wu, Y. D. *Inorg. Chem.* **2005**, *44*, 1466–1471.
- Ilich, P.; Hille, R. *J. Phys. Chem. B* **1999**, *103*, 5406–5412.
- Ilich, P.; Hille, R. *J. Am. Chem. Soc.* **2002**, *124*, 6796–6797.
- Amano, T.; Ochi, N.; Sato, H.; Sakaki, S. *J. Am. Chem. Soc.* **2007**, *129*, 8131–8138.
- Bayse, C. A. *Dalton Trans.* **2009**, 2306–2314.
- Massey, V.; Edmondson, D. *J. Biol. Chem.* **1970**, *245*, 6595–6598.
- Wagener, N.; Pierek, A. J.; Ibdah, A.; Hille, R.; Dobbek, H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 11055–11060.
- Bray, R. C.; George, G. N. *Biochem. Soc. Trans.* **1985**, *13*, 560–567.
- Xia, M.; Ilich, P.; Dempski, R.; Hille, R. *Biochem. Soc. Trans.* **1997**, *25*, 768–773.
- Leimkühler, S.; Stockert, A. L.; Igarashi, K.; Nishino, T.; Hille, R. *J. Biol. Chem.* **2004**, *279*, 40437–40444.
- Dietzel, U.; Kuper, J.; Doebbler, J. A.; Schulten, A.; Truglio, J. J.; Leimkühler, S.; Kisker, C. *J. Biol. Chem.* **2009**, *284*, 8768–8776.
- MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- Luo, G. B.; Andricioaei, I.; Xie, X. S.; Karplus, M. *J. Phys. Chem. B* **2006**, *110*, 9363–9367.
- Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
- Sherwood, P.; et al. *THEOCHEM* **2003**, *632*, 1–28.
- Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136–141.
- Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- de Vries, A. H.; Sherwood, P.; Collins, S. J.; Rigby, A. M.; Rigutto, M.; Kramer, G. J. *J. Phys. Chem. B* **1999**, *103*, 6133–6141.
- Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, 79–92.
- Klamt, A.; Schürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, *5*, 799–805.
- Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.
- Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

- (45) Becke, A. D. *Phys. Rev. A* **1988**, 38, 3098–3100.
- (46) Perdew, J. P. *Phys. Rev. B* **1986**, 33, 8822–8824.
- (47) Perdew, J. P. *Phys. Rev. B* **1986**, 34, 7406.
- (48) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, 240, 283–289.
- (49) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, 97, 119–124.
- (50) Becke, A. D. *J. Chem. Phys.* **1993**, 98, 5648–5652.
- (51) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, 98, 11623–11627.
- (52) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785–789.
- (53) Henkelman, G.; Uberuaga, B. P.; Jonsson, H. *J. Chem. Phys.* **2000**, 113, 9901–9904.
- (54) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. J. *Phys. Chem. A* **2009**, 113, 11856–11865.
- (55) Frisch, M. J.; et al. *Gaussian 03*, revision D.01; Gaussian Inc.: Wallingford, CT, 2004.
- (56) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, 14, 33–38.
- (57) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, 82, 270–283.
- (58) Ehlers, A. W.; Böhme, M.; Dapprich, S.; Gobbi, A.; Höllwarth, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, 208, 111–114.
- (59) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, 82, 284–298.
- (60) Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, 208, 237–240.
- (61) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, 28, 213–222.
- (62) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, 4, 294–301.
- (63) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, 8, 1057–1065.
- (64) Kim, J. H.; Odutola, J. A.; Popham, J.; Jones, L.; von Laven, S. *J. Inorg. Biochem.* **2001**, 84, 145–150.
- (65) Eyring, H. *J. Chem. Phys.* **1935**, 3, 107–115.
- (66) Sevastik, R.; Himo, F. *Bioorg. Chem.* **2007**, 35, 444–457.
- (67) Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput.* **2008**, 4, 1129–1137.
- (68) Chen, S.-L.; Fang, W.-H.; Himo, F. *Theor. Chem. Acc.* **2008**, 120, 515–522.
- (69) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, 14, 643–651.

JP909999S

# Supporting Information for

## QM/MM Studies of Xanthine Oxidase: Variations of Cofactor, Substrate, and Active-Site Glu802

*Sebastian Metz and Walter Thiel\**

Max-Planck-Institut für Kohlenforschung, D-45470 Mülheim an der Ruhr, Germany

[thiel@mpi-muelheim.mpg.de](mailto:thiel@mpi-muelheim.mpg.de)

1. System preparation using CHARMM	S2
2. Influence of the hydride acceptor (O vs. S vs. Se)	S6
3. Influence of the substrate (2-oxo-6-methylpurine)	S10
4. Influence of the mutation Glu802→Gln	S18
5. CHARMM parameters	S20
6. Complete references 32, 36, and 55	S32

## 1. System preparation using CHARMM

We followed the same procedures as in our previous publication<sup>1</sup> and therefore only give a short summary here.

We used SWISS-MODEL<sup>2-5</sup> to add missing residues and the program *Reduce*<sup>6</sup> to adjust the orientation of Asn and Gln side-chain amides and of the His imidazole rings. Flipped residues are listed below (if the residue has been flipped in just one of the chains, the chain ID is given in brackets).

### Flipped residues:

ASN 71, GLN 131, ASN 146(B), ASN 351, GLN 473, ASN 556, GLN 626, ASN 650, HSE 683(A), HSD 747, HSD 875, ASN 1145, HSD 1151, GLN 1284, ASN 1287(B), ASN 1324(A)

All Lys and Arg side chains are in the ammonium or guanidinium forms, respectively, and all Asp side chains are in the carboxylate form. For all setups, Glu267 was protonated, and Cys992 directly neighboring the doubly protonated histidine Hsp677 was deprotonated. Glu802 was protonated in both chains. The chosen protonation states for all histidines are listed below (if the protonation state has been assigned in just one of the two chains, the chain ID is given in brackets).

### Protonation states:

GLU 267p, GLU802p (SETUP E: just chain B), CYS 992d

HSD: 67 187 252 292(A) 552(B) 614 747 821 840 863 875 884 1151 1171(B) 1212 (B)

HSE: 81 82 99 109 292(B) 387 552(A) 579 665 683 741 954 1022 1033 1043 1171(A) 1212 (A) 1220

HSP: 677 1285

SETUP A, C, D, E, F, G: GLU 1261 (just chain A)

To keep the whole system neutral, Cl<sup>-</sup> counterions were added. This was done using the program *ionize*,<sup>7</sup> which puts a three dimensional grid on the enzyme and then iteratively calculates the electrostatic potential on all grid points, adding a single ion at the most favorable position until the system is neutral.

The residues in the direct surrounding of the counterions were reoptimized using CHARMM. As a result of this procedure, all counterions were placed on the surface of the enzyme and more than 35Å away from the molybdenum center. Their coordinates are given below for direct use in the original crystal structure.

### Added counterions:

ATOM	41516	CLA	CLA	8	75.360	86.126	20.686
ATOM	41517	CLA	CLA	9	28.661	-47.776	50.009
ATOM	41519	CLA	CLA	10	93.374	14.979	26.149



## CHARMM parameters

To be able to prepare the system by extended classical MD runs prior to the QM/MM optimizations, we needed CHARMM parameters for the full system, including FAD, FEST, MOC and URIC. The FAD parameters were provided by M. Karplus.<sup>8</sup> We used the FEST parameters published in our prior publications.<sup>1,9</sup> We kept the molybdenum center fixed during MD simulation and used ESP charges for the interaction with the enzymatic environment. The parameters for 2-oxo-6-methylpurine were adopted from similar molecules like guanine or uracil; the atomic charges were taken from ESP calculations.

Note that our goal was only to obtain reasonable parameters suitable for the preparatory MM-MD simulations. We therefore did not extensively validate the newly derived parameters, nor do we recommend their use in pure MM calculations without further tests. All parameters not provided with the CHARMM package can be found in Section 5 (see pages S20-S31).

## Hydration procedure

The protein structure with added H atoms was hydrated using a sphere of 35 Å radius containing 6840 water molecules cut from an equilibrated simulation of TIP3P water under periodic boundary conditions. The water sphere was superimposed onto the protein structure, centered on MOC(A):C8. All added water molecules whose O atom was within 2.8 Å of any existing non-H atom were deleted. The O atoms of all water molecules (i.e., added and existing ones) within the 35Å sphere were subjected to a spherical quartic boundary potential with parameters  $FORCE = 0.2$ ,  $P1 = 2.25$ , and  $DROFF = 33.5$ . The water molecules were kept rigid using SHAKE constraints. Keeping all other atoms fixed, the water molecules within the sphere were geometry-optimized, performing first 300 steps of steepest descent followed by 300 steps of adapted-basis Newton–Raphson minimization. In the first solvation run, we performed another two minimizations of 300 steps each (steepest descent and adapted-basis Newton Raphson), optimizing the active region without the water molecules, adding a restraint of  $k=30 \text{ mol } \text{\AA}^{-2}$  on the non-fixed atoms. Finally, the water molecules were subjected to 30 ps molecular dynamics at 300 K with (subsequently decreasing) restraints. Afterward, we resteped to the next solvation cycle. Information on the different solvation cycles is listed below.

The nonbonded interactions in all CHARMM simulations were treated using

```
NBOND ELEC ATOM FSWITCH CDIE VDW VATOM VFSWITCH CUTNB 17. CTONNB 11.  
CTOFNB 15. EPS 1. NBXMOD 5 E14FAC 1. WMIN 1.5.
```

## Active residues

### chain A

GLN112, CYS113, CYS148, ARG149, CYS150, THR151, GLY152, HSE579, TYR592, THR646, GLY647, LEU648, PHE649, ASN650, ASP651, GLU652, ILE698, TYR708, LEU712, LYS713, ILE714, ILE736, GLY737, GLY738, GLN739, ASP740, HSE741, PHE742, TYR743, LEU744, GLU745, THR746, HSD747, CYS748, VAL764, SER765, THR766, GLN767, ASN768, ALA769, MET770, LYS771, THR772, GLN773, SER774, PHE775, VAL789, VAL791, LYS792, ARG793, MET794, GLY795, GLY796, GLY797, PHE798, GLY799, GLY800, LYS801, GLU802, THR803, ARG804, SER805, THR806, LEU807, VAL808, SER809, VAL810, LEU827, ARG829, ASP832, MET833, ILE835, THR836, GLY837, GLY838, ARG839, HSD840, PRO841, PHE842, LEU843, ALA844, HSD863, TYR864, SER865, ASN866, ALA867, GLY868, ASN869, SER870, ARG871, ASP872, LEU873, SER874, HSD875, SER876, ILE877, MET878, GLU879, ARG880, ALA881, LEU882, PHE883, HSD884, MET885, ASP886, ASN887, CYS888, THR897, GLY898, ARG899, LEU900, CYS901, SER906, SER907, ASN908, THR909, ALA910, PHE911, ARG912, GLY913, PHE914, GLY915, GLY916, PRO917, GLN918, ALA919, LEU920, PHE921, ILE922, ALA923, GLU924, TYR947, THR953, HSE954, PHE955, GLN957, LEU959, PHE962, VAL964, ILE1001, PRO1002, THR1003, LYS1004, PHE1005, GLY1006, ILE1007, SER1008, PHE1009, THR1010, VAL1011, PRO1012, PHE1013, LEU1014, ASN1015, GLN1016, ALA1017, GLY1018, ALA1019, HSE1033, GLY1034, GLY1035, THR1036, GLU1037, MET1038, GLY1039, GLN1040, GLY1041, LEU1042, HSE1043, THR1044, LYS1045, MET1046, VAL1049, ILE1063, THR1066, SER1067, THR1068, VAL1071, PRO1072, ASN1073, SER1074, SER1075, PRO1076, THR1077, ALA1078, ALA1079, SER1080, VAL1081, SER1082, THR1083, ASP1084, ILE1085, TYR1086, GLY1087, GLN1088, ALA1089, VAL1090, PHE1132, TYR1133, ARG1134, THR1135, PRO1136, ASN1137, LEU1138, GLY1139, TYR1140, SER1141, PHE1142, GLY1147, ASN1148, ALA1149, PHE1150, HSD1151, TYR1152, PHE1153, THR1154, TYR1155, GLY1156, ILE1178, MET1180, ASP1181, VAL1182, GLY1183, SER1185, LEU1186, ASN1187, ILE1190, ASP1191, ILE1192, GLY1193, GLN1194, VAL1195, GLU1196, GLY1197, ALA1198, PHE1199, VAL1200, GLN1201, GLY1202, LEU1203, GLY1204, LEU1205, PHE1206, GLU1209, LEU1211, TYR1213, TYR1227, ILE1229, PRO1230, LEU1243, ASN1249, LYS1251, ALA1252, ILE1253, SER1256, LYS1257, ALA1258, VAL1259, GLY1260, GLU1261, PRO1262, PRO1263, LEU1264, PHE1265, LEU1266, GLY1267, ALA1268, SER1298, FES2001, MOC2003, URIC2, GOL2, CAL6

### chain B

THR1025, TYR1121, GLN1122, ASP1123, ARG1124

### crystal water

CRYW3, CRYW4, CRYW5, CRYW7, CRYW9, CRYW11, CRYW12, CRYW13, CRYW15, CRYW19, CRYW23, CRYW24, CRYW25, CRYW40, CRYW42, CRYW45, CRYW48, CRYW49, CRYW52, CRYW56, CRYW58, CRYW61, CRYW63, CRYW65, CRYW71, CRYW74, CRYW77, CRYW85, CRYW89, CRYW91, CRYW93, CRYW94, CRYW96, CRYW97, CRYW101, CRYW108, CRYW109, CRYW114, CRYW122, CRYW125, CRYW126, CRYW128, CRYW143, CRYW157, CRYW160, CRYW168, CRYW174, CRYW177, CRYW189, CRYW191, CRYW196, CRYW199, CRYW200, CRYW203, CRYW211, CRYW213, CRYW218, CRYW222, CRYW224, CRYW225, CRYW227, CRYW230, CRYW240, CRYW243, CRYW244, CRYW248, CRYW274, CRYW277, CRYW283, CRYW287, CRYW300, CRYW308, CRYW317, CRYW343, CRYW369, CRYW373, CRYW382, CRYW405, CRYW408, CRYW428, CRYW429, CRYW447, CRYW498, CRYW536, CRYW558, CRYW563, CRYW565, CRYW570, CRYW571, CRYW575, CRYW596, CRYW615, CRYW632, CRYW636, CRYW706, CRYW712, CRYW751, CRYW753, CRYW798, CRYW847, CRYW947, CRYW1034, CRYW1128, CRYW1205, CRYW1340, CRYW1678

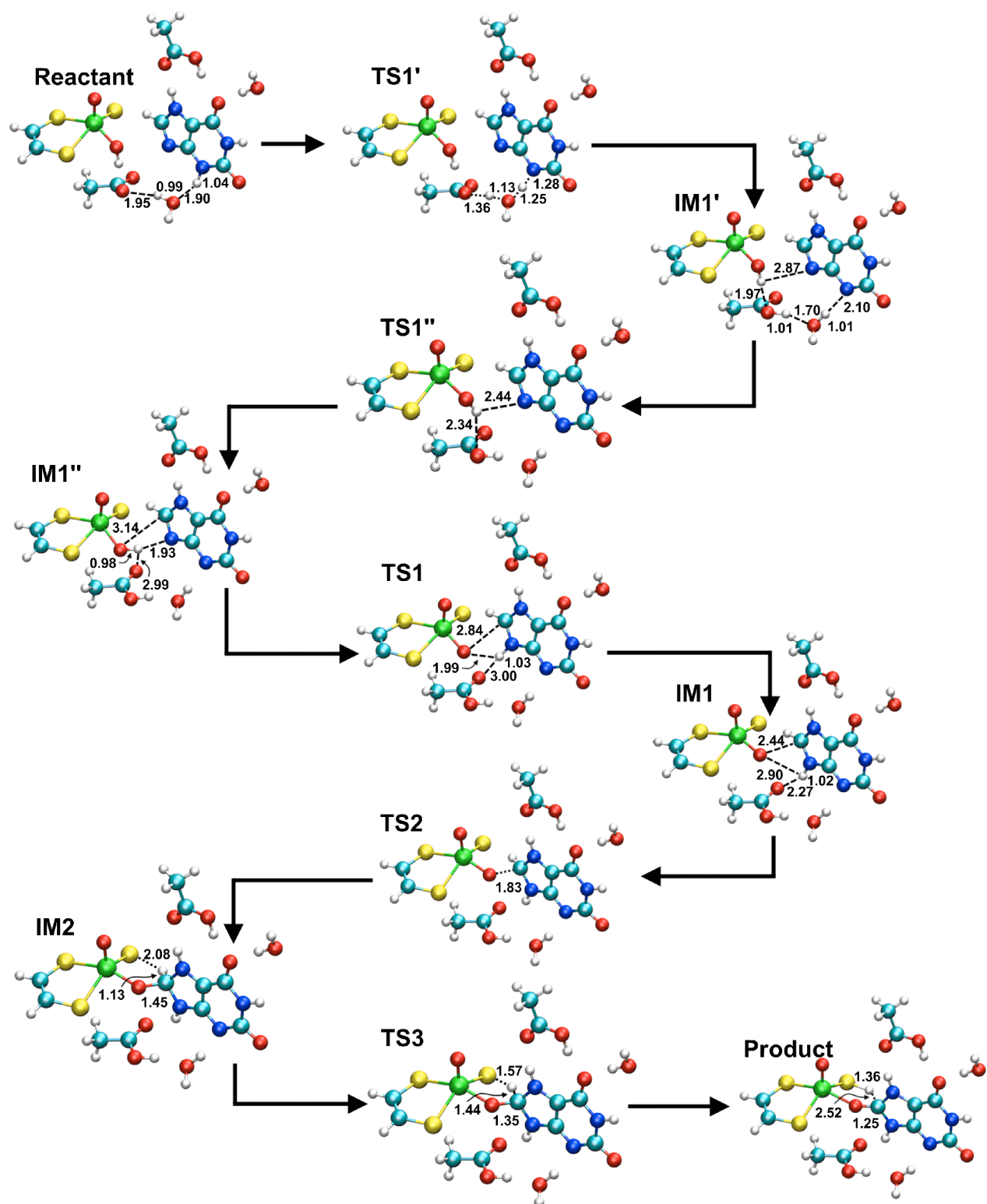
and all solvent water molecules



## References:

1. Metz, S.; Thiel, W. *J. Am. Chem. Soc.* **2009**, ASAP article, DOI:10.1021/ja9045394.
2. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, 18, 2714-2723.
3. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. *Nucleic Acids Res.* **2003**, 31, 3381-3385.
4. Kopp, J.; Schwede, T. *Nucleic Acids Res.* **2004**, 32, D230-D234.
5. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *Bioinformatics* **2006**, 22, 195-201.
6. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, 285, 1735-1747.
7. <http://www.ks.uiuc.edu/Development/MDTools/ionize>.
8. Luo, G. B.; Andricioaei, I.; Xie, X. S.; Karplus, M. *J. Phys. Chem. B* **2006**, 110, 9363-9367.
9. Metz, S.; Wang, D.; Thiel, W. *J. Am. Chem. Soc.* **2009**, 131, 4628-4640.

## 2. Influence of the hydride acceptor (O vs. S vs. Se)



**Figure S1.** Reaction mechanism for Y=S. The structures for Y=O and Y=Se are analogous and therefore not presented in separate figures.

**Table S1.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of the oxo form of SETUP G in kcal mol<sup>-1</sup>.

	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	2.7	-0.1	2.5	6.3	-0.7	5.6	Proton transfer 1
IM1'	-4.9	-2.5	-7.4	-3.5	-3.6	-7.2	
TS1''	-4.9	-2.0	-6.9	-4.5	-2.5	-7.1	OH-Rotation
IM1''	-5.4	-3.5	-8.9	-5.0	-4.5	-9.5	
TS1	-9.0	10.8	1.8	-7.6	9.1	1.6	Proton transfer 2
IM1	-10.1	9.0	-1.1	-7.6	8.4	0.7	
TS2	-6.8	6.3	-0.5	-0.6	3.2	2.6	C-O formation
IM2	-4.9	0.2	-4.7	2.2	-1.3	0.9	
TS3	10.4	2.9	13.3	19.6	1.6	21.2	Hydride transfer
Product	-2.2	8.4	6.2	-12.6	8.1	-4.4	
<sup>a</sup> absolute energies (a.u.):							
			-1565.507584 -49.975596 -1615.483179 (BP86)				
			-1564.386683 -49.971791 -1614.358474 (B3LYP)				

**Table S2.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of the sulfido form of SETUP G in kcal mol<sup>-1</sup>.

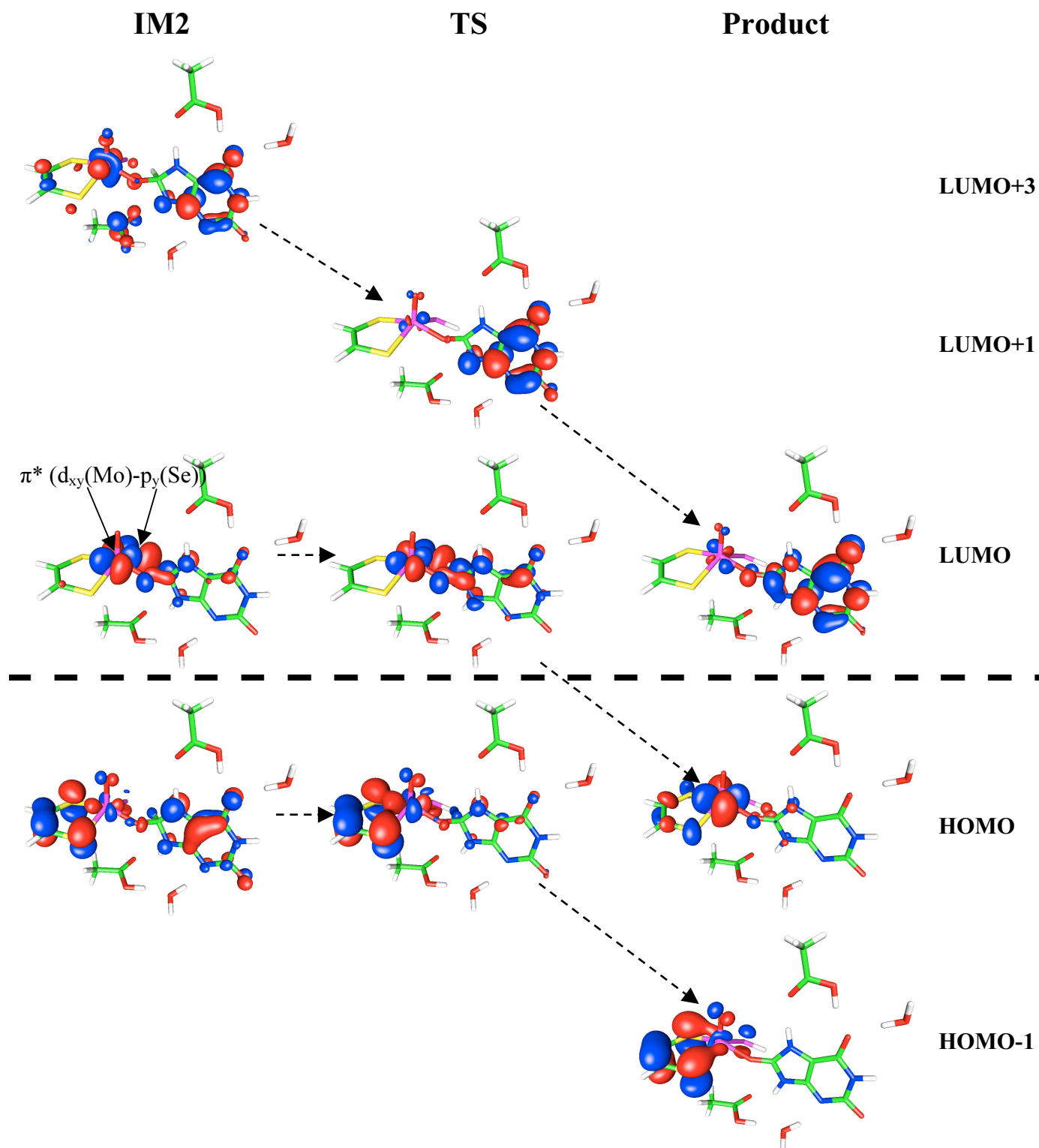
	BP86/Lanl2DZ			B3LYP/Lanl2DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	2.8	0.1	2.9	9.4	-0.9	8.5	Proton transfer 1
IM1'	-7.0	1.0	-8.0	-6.0	-1.5	-7.5	
TS1''	-6.7	-0.7	-7.4	-5.3	-1.9	-7.2	OH-Rotation
IM1''	-7.2	-1.7	-9.0	-6.2	-2.4	-8.6	
TS1	0.0	1.0	1.0	3.1	-0.6	2.5	Proton transfer 2
IM1	-13.5	10.1	-3.4	-10.8	9.0	-1.8	
TS2	-6.7	4.5	-2.1	-0.2	2.6	2.5	C-O formation
IM2	-1.5	-6.3	-7.8	-2.2	-0.2	-2.5	
TS3	-1.9	2.6	0.7	3.2	1.3	4.5	Hydride transfer
Product	-8.6	7.9	-0.8	-17.9	8.7	-9.3	
<sup>a</sup> absolute energies (a.u.):							
			-1500.413756 -49.980759 -1550.394515 (BP86)				
			-1499.281279 -49.976648 -1549.257927 (B3LYP)				

**Table S3.** Calculated relative QM/MM energies of optimized stationary points for Snapshot 400 of the sulfido form of SETUP G in kcal mol<sup>-1</sup>.

	BP86/Lan12DZ			B3LYP/Lan12DZ			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1'	3.4	-0.4	3.1	5.6	-0.4	5.1	Proton transfer 1
IM1'	-7.5	-0.7	-8.2	-8.7	-0.9	-9.6	
TS1''	-6.1	-1.0	-7.1	-8.9	-0.5	-9.4	OH-Rotation
IM1''	-6.2	-2.5	-8.7	-8.5	-2.2	-10.7	
TS1	-0.7	1.0	0.4	-2.0	0.7	-1.3	Proton transfer 2
IM1	-13.4	10.3	-3.1	-14.5	10.6	-3.8	
TS2	-5.0	3.9	-1.0	-2.9	3.1	0.2	C-O formation
IM2	-7.8	1.3	-6.6	0.1	-5.9	-5.7	
TS3	5.0	-4.4	0.6	4.4	-5.2	-0.8	Hydride transfer
Product	-0.3	-0.5	-0.8	-20.9	10.3	-10.7	

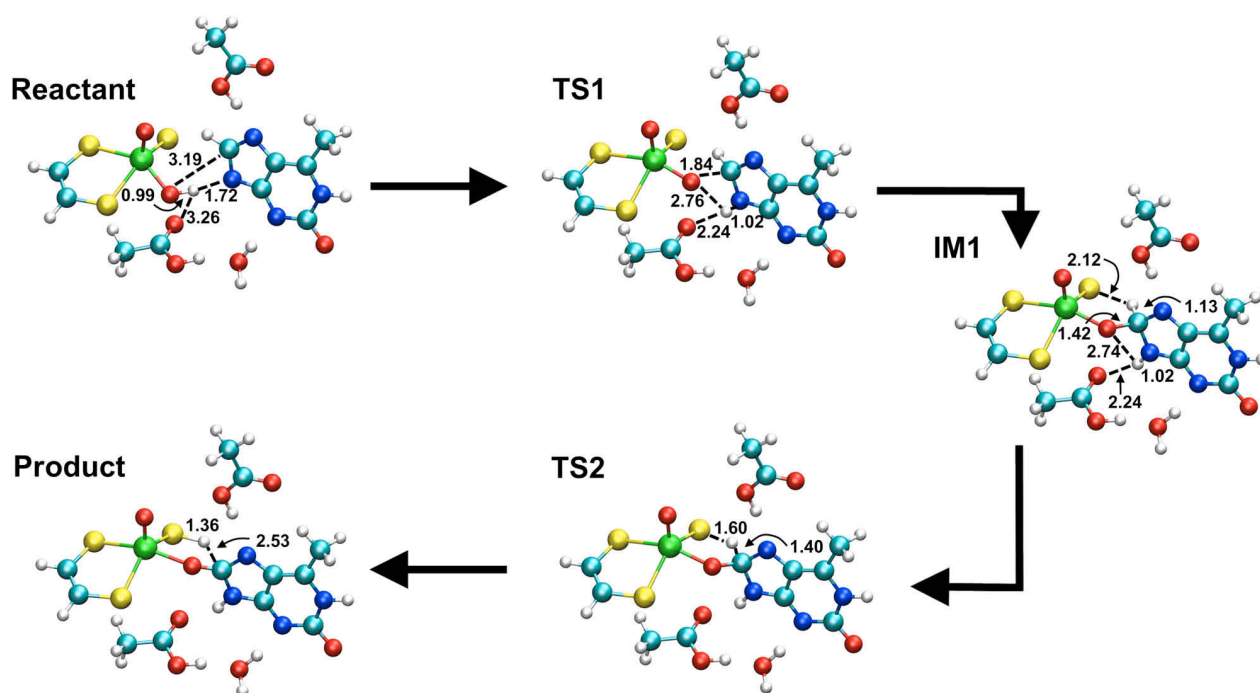
<sup>a</sup> absolute energies (a.u.):  
-1499.532106 -49.979549 -1549.511655 (BP86)  
-1498.389771 -49.977770 -1548.367541 (B3LYP)

**Figure S2.** Molecular orbitals for the stationary points of the hydride transfer step at the B3LYP/B1 level, for the selenido form of the cofactor. The antibonding interaction of  $d_{xy}(\text{Mo})$  and  $p_y(\text{Se})$  is highlighted for the LUMO of IM2.



### 3 Influence of the substrate (2-oxo-6-methylpurine)

#### a) SETUP 1



**Figure S3.** Reaction mechanism for Setup 1, SN400 using the B3LYP functional.

**Table S4.** Calculated relative QM/MM energies of optimized stationary points for SN 200 of SETUP 1 in kcal mol<sup>-1</sup>.

	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	15.3	2.9	18.1	20.5	2.1	22.7	C-O formation
IM2	10.9	1.3	12.2	17.1	0.6	17.7	
TS3	16.5	1.1	17.6	21.5	1.0	22.6	Hydride transfer
Product	9.4	4.3	13.7	1.1	6.5	7.5	

<sup>a</sup> absolute energies (a.u.):  
 -1387.975039 -90.746599 -1478.721637 (BP86)  
 -1386.890254 -90.744099 -1477.634353 (B3LYP)

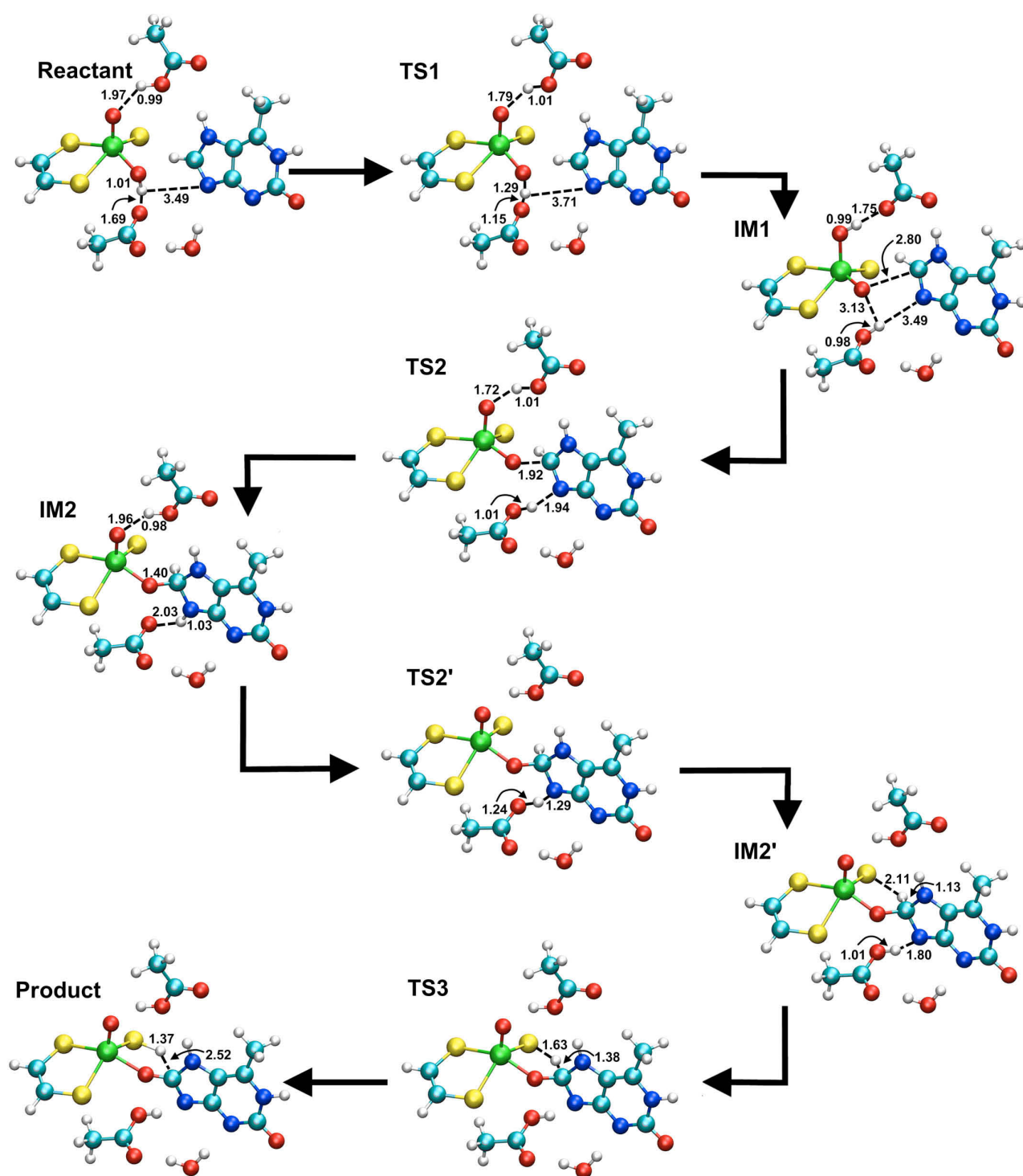
**Table S5.** Calculated relative QM/MM energies of optimized stationary points for SN 400 of SETUP 1 in kcal mol<sup>-1</sup>.

	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	15.5	2.1	17.6	20.7	2.6	23.4	C-O formation
IM2	10.0	0.4	10.4	15.0	0.9	16.0	
TS3	15.0	0.6	15.6	20.3	1.3	21.7	Hydride transfer
Product	9.3	1.8	11.1	2.1	3.9	6.0	
<sup>a</sup> absolute energies (a.u.): -1388.017623 -90.549135 -1478.566758 (BP86) -1386.931320 -90.548575 -1477.479895 (B3LYP)							

**Table S6.** Calculated relative QM/MM energies of optimized stationary points for SN 400 of SETUP 1 in kcal mol<sup>-1</sup>.

	larger QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS2	20.9	4.9	25.8	C-O formation
IM2	17.8	2.2	20.0	
TS3	22.1	2.2	24.3	Hydride transfer
Product	1.5	6.6	8.1	
<sup>a</sup> absolute energies (a.u.): -1841.002797 -90.160203 -1931.163001				

**b) SETUP 2**



**Figure S4.** Reaction mechanism for Setup 2, SN400 using the B3LYP functional.



**Table S7.** Calculated relative QM/MM energies of optimized stationary points for SN 200 of SETUP 2 in kcal mol<sup>-1</sup>.

	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
IM1 <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	5.9	2.9	8.7	5.3	4.5	9.8	C-O formation
IM2	-4.5	4.1	-0.5	-6.9	5.8	-1.1	
TS3	2.9	2.5	5.4	3.3	5.9	9.3	Hydride transfer
Product	-1.4	5.7	4.3	-14.8	9.0	-5.7	
<sup>a</sup> absolute energies (a.u.): -1387.986171 -90.722565 -1478.708736 (BP86) -1386.894598 -90.724067 -1477.618664 (B3LYP)							

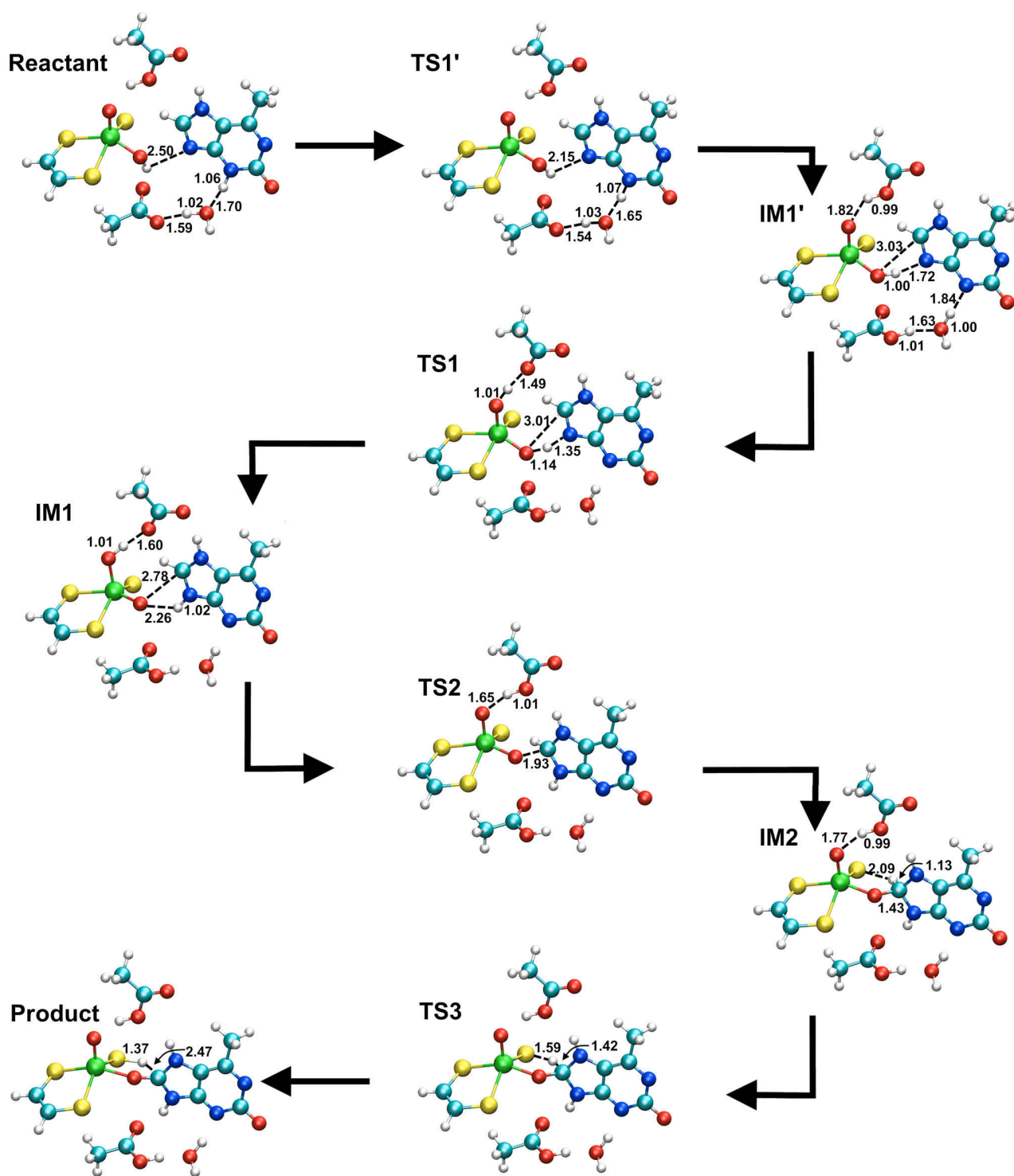
**Table S8.** Calculated relative QM/MM energies of optimized stationary points for SN 400 of SETUP 2 in kcal mol<sup>-1</sup>.

	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	-10.3	14.8	4.4	-11.0	14.8	3.9	
TS1	-8.1	13.9	5.8	-6.9	14.1	7.3	Proton transfer 1
IM1 <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	0.7	7.8	8.5	4.5	7.1	11.6	C-O formation
IM2	-9.0	9.2	0.2	-8.0	9.0	1.0	
TS2'	-6.3	8.5	2.2	-4.2	8.8	4.6	Proton transfer 2
IM2'	-13.7	14.5	0.7	-10.2	13.5	3.3	
TS3	-0.4	5.4	5.1	2.3	8.9	11.2	Hydride transfer
Product	-5.3	6.6	1.3	-14.7	7.0	-7.7	
<sup>a</sup> absolute energies (a.u.): -1388.021422 -90.534915 -1478.556337 (BP86) -1386.935438 -90.533438 -1477.468876 (B3LYP)							

**Table S9.** Calculated relative QM/MM energies of optimized stationary points for SN 400 of SETUP 2 in kcal mol<sup>-1</sup>.

	larger QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	
TS1	-1.6	7.4	5.8	Proton transfer 1
IM1	6.8	-5.4	1.4	
TS2	8.6	-1.2	7.3	C-O formation
IM2	9.7	-7.7	1.9	
TS2'	15.9	-9.2	6.6	Proton transfer 2
IM2'	8.3	-6.5	1.8	
TS3	14.0	-6.9	7.0	Hydride transfer
Product	-5.8	-1.6	-7.4	
<sup>a</sup> absolute energies (a.u.): -1841.010545 -90.141475 -1931.152020				

c) Setup 3



**Figure S5.** Reaction mechanism for Setup 3, SN400 using the B3LYP functional.

**Table S10.** Calculated relative QM/MM energies of optimized stationary points for SN 200 of SETUP 3 in kcal mol<sup>-1</sup>.

	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	5.9	3.5	9.4	4.6	3.7	8.3	
TS1'	8.1	4.1	12.2	9.4	3.8	13.3	Proton transfer 1
IM1'	0.4	3.4	3.9	-2.5	4.2	1.8	
TS1	3.0	1.5	4.6	4.4	1.1	5.5	Proton transfer 2
IM1 <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	0.0	6.7	6.7	4.3	5.4	9.6	CO formation
IM2	-3.7	4.8	1.1	-0.9	4.8	3.9	
TS3	3.2	4.9	8.1	4.8	4.1	8.9	Hydride transfer
Product	0.0	7.9	7.9	-11.4	9.8	-1.6	

<sup>a</sup> absolute energies (a.u.): -1387.996046 -90.722766 -1478.718811 (BP86)  
-1386.907647 -90.720286 -1477.627933 (B3LYP)

**Table S11.** Calculated relative QM/MM energies of optimized stationary points for SN 400 of SETUP 3 in kcal mol<sup>-1</sup>.

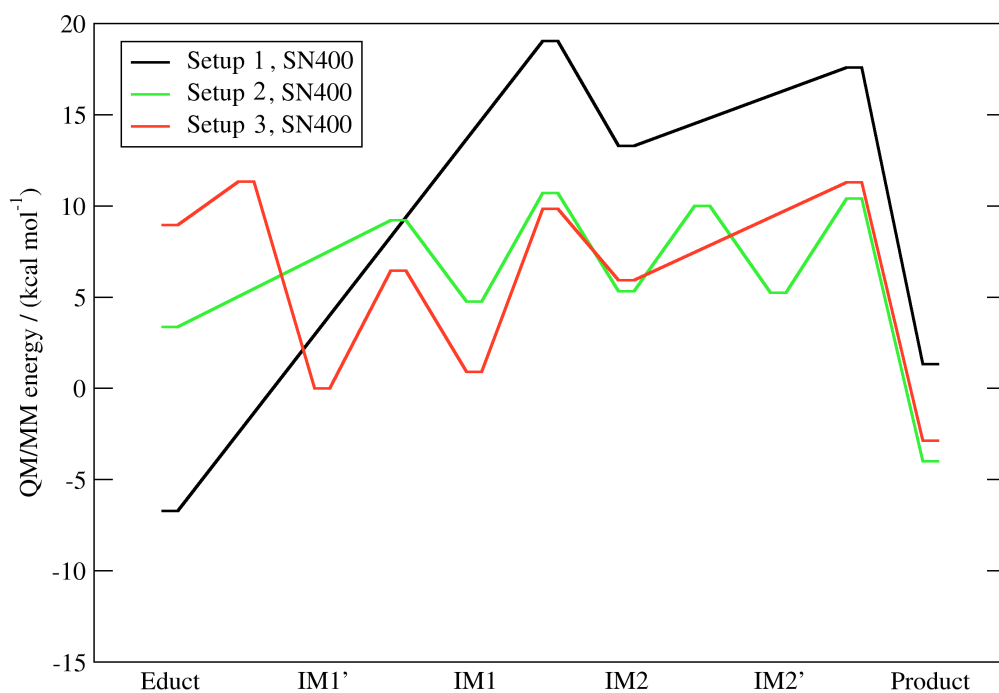
	BP86			B3LYP			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	11.9	0.2	12.1	8.0	1.0	9.1	
TS1'	11.7	1.1	12.7	8.0	1.5	9.5	Proton transfer 1
IM1'	2.6	1.3	3.9	-1.2	2.2	1.0	
TS1	6.7	-0.9	5.8	9.0	-2.5	6.5	Proton transfer 2
IM1 <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS2	3.1	2.8	5.9	4.7	2.7	7.4	C-O formation
IM2	-0.5	1.1	0.6	0.6	1.4	2.0	
TS3	6.3	0.4	6.7	6.7	0.4	7.0	Hydride transfer
Product	1.7	2.6	4.4	-10.6	4.0	-6.6	

<sup>a</sup> absolute energies (a.u.): -1388.037653 -90.545232 -1478.582886 (BP86)  
-1386.949040 -90.543172 -1477.492212 (B3LYP)

**Table S12.** Calculated relative QM/MM energies of optimized stationary points for SN 200 of SETUP 1 in kcal mol<sup>-1</sup>.

	Large QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	6.7	2.2	9.0	
TS1'	9.3	2.0	11.3	Proton transfer 1
IM1' <sup>a</sup>	0.0	0.0	0.0	
TS1	8.5	-2.1	6.5	C-O formation
IM1	-0.3	1.2	0.9	
TS2	5.8	4.0	9.8	Proton transfer 2
IM2	7.9	-1.9	5.9	
TS3	12.9	-1.6	11.3	Hydride transfer
Product	-6.9	4.0	-2.8	

<sup>a</sup> absolute energies (a.u.):      -1841.010545 -90.141475 -1931.152020



**Figure S6.** QM/MM reaction profiles for all three setups on a common energy scale derived from COSMO ( $\epsilon=8.0$ ) calculations on the extended QM region.

#### 4 Influence of the mutation Glu802→Gln

**Table S13.** Calculated relative QM/MM energies of optimized stationary points for the Glu802→Gln mutant, SN 400 of SETUP G in kcal mol<sup>-1</sup>.

	standard QM region			large QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	5.4	3.7	9.1	1.4	7.5	9.0	
TS1"	14.9	1.0	16.0	13.3	3.4	16.7	Proton transfer 1
IM1"	2.3	0.3	2.5	6.0	-0.2	5.8	
TS1'	1.8	0.8	2.5	5.8	0.0	5.7	OH-Rotation
IM1 <sup>ia</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	3.8	7.5	11.3	6.2	4.4	10.6	Proton transfer 2
IM1	-2.8	12.2	9.5	-0.5	10.1	9.6	
TS2	10.9	4.2	15.1	21.3	-2.9	18.3	C-O formation
IM2	12.4	-0.9	11.5	21.2	-6.2	15.1	
TS3	14.7	2.7	17.5	27.2	-5.2	22.0	Hydride transfer
Product	-7.8	10.5	2.7	0.7	6.9	7.6	

<sup>a</sup> absolute energies (a.u.): -1479.430248 -49.978512 -1529.408760 (standard QM region)  
-1857.111604 -49.585650 -1906.697254 (enlarged QM region)

**Table S14.** Calculated relative QM/MM energies of optimized stationary points for the Glu802→Gln mutant, SN 400 of SETUP G in kcal mol<sup>-1</sup>.

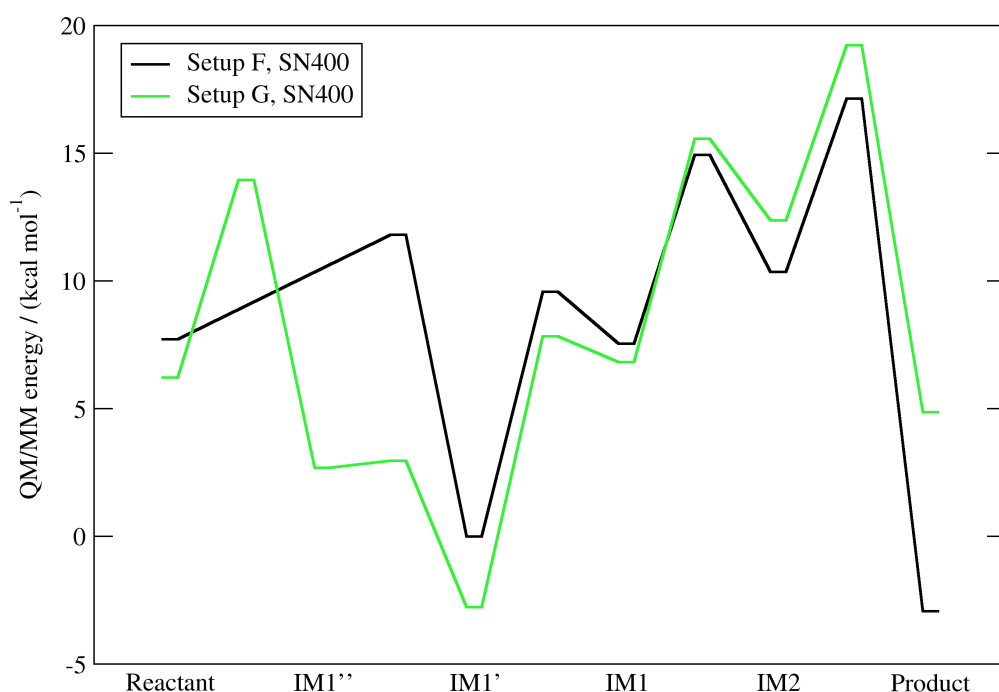
	standard QM region			large QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant	9.4	-0.3	9.1	4.8	2.9	7.7	
TS1'	11.6	0.5	12.0	10.4	1.4	11.8	Proton transfer 1
IM1 <sup>ia</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1	7.2	2.6	9.8	5.0	4.6	9.6	Proton transfer 2
IM1	4.8	3.0	7.7	3.0	4.5	7.6	
TS2	14.7	-1.6	13.1	14.8	0.1	14.9	C-O formation
IM2	12.3	-2.0	10.3	11.3	-0.9	10.4	
TS3	18.1	-2.3	15.8	17.0	0.2	17.1	Hydride transfer
Product	-5.8	2.2	-3.7	-6.7	3.7	-2.9	

<sup>a</sup> absolute energies (a.u.): -1479.372402 -87.219719 -1566.592121 (standard QM region)  
-1857.092836 -86.790014 -1943.882850 (enlarged QM region)

**Table S15.** Calculated relative QM/MM energies of optimized stationary points for the Glu802→Gln mutant, SN 500 of SETUP A in kcal mol<sup>-1</sup>.

	standard QM region			large QM region			Remark
	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	E <sub>QM</sub>	E <sub>MM</sub>	E <sub>QM/MM</sub>	
Reactant <sup>a</sup>	0.0	0.0	0.0	0.0	0.0	0.0	
TS1''	4.0	11.8	15.9	11.4	6.2	17.6	Gln802 Rotation
IM1''	-0.6	7.7	7.1	9.7	0.3	10.0	
TS1'	16.0	4.0	20.0	19.1	0.8	19.9	Proton transfer 1
IM1'	1.2	2.4	3.7	1.8	3.1	4.9	
TS1							Proton transfer 2
IM1	3.1	15.0	18.1	10.5	7.6	18.1	
TS2							C-O formation
IM2	8.9	10.5	19.4	18.6	2.4	21.0	
TS3	13.7	10.2	23.9	21.8	3.8	25.7	Hydride transfer
Product	-9.9	13.6	3.7	-6.6	10.3	3.7	

<sup>a</sup> absolute energies (a.u.): -1402.975969 -87.126232 -1490.102201 (standard QM region)  
-1857.048442 -86.745385 -1943.793827 (enlarged QM region)



**Figure S7.** QM/MM reaction profiles for both setups on a common energy scale derived from COSMO ( $\epsilon=8.0$ ) calculations on the extended QM region.

## 5. CHARMM parameters

```
!=====
! CHARMM PARAMETERS for FAD provided by M. Karplus,
!                               cf. J. Phys. Chem. B, 110, 9363-9367 (2006)
!=====
```

### !!!!!!!!!!!!!! FAD TOPOLOGY !!!!!!!!!!!!!!!

```
MASS 304    NX1      14.00700  N ! FAD-specific
MASS 305    HX1       1.00800  H ! FAD-specific
MASS 306    CX1     12.011000  C ! FAD-specific
MASS 307    OX1     15.999400  O ! FAD-specific
MASS 308    CX2     12.011000  C ! FAD-specific
MASS 309    CX3     12.011000  C ! FAD-specific
MASS 310    NX2      14.00700  N ! FAD-specific
MASS 311    CX4     12.011000  C ! FAD-specific
MASS 312    OX2     15.999400  O ! FAD-specific
MASS 313    NY1      14.00700  N ! FAD-specific
MASS 314    CY1     12.011000  C ! FAD-specific
MASS 315    CY2     12.011000  C ! FAD-specific
MASS 316    NY2      14.00700  N ! FAD-specific
MASS 317    CZ1     12.011000  C ! FAD-specific
MASS 318    HZ1       1.00800  H ! FAD-specific
MASS 319    CZ2     12.011000  C ! FAD-specific
MASS 320    CZ3     12.011000  C ! FAD-specific
MASS 321    CZ4     12.011000  C ! FAD-specific
MASS 322    HZ4       1.00800  H ! FAD-specific
```

```
RESI FAD    -2.000          ! cf. J. Phys. Chem. B, 110, 9363-9367 (2006)
GROUP
```

```
ATOM C4'    CN7      0.16
ATOM H4'    HN7      0.09
ATOM O4'    ON6B    -0.50
ATOM C1'    CN7B     0.16
ATOM H1'    HN7      0.09
```

GROUP

```
ATOM C5     CN5      0.28
ATOM N7     NN4     -0.71
ATOM C8     CN4      0.34
ATOM H8     HN3      0.12
ATOM N9     NN2     -0.05
ATOM N1     NN3A    -0.74
ATOM C2     CN4      0.50
ATOM H2     HN3      0.13
ATOM N3     NN3A    -0.75
ATOM C4     CN5      0.43
ATOM C6     CN2      0.46
ATOM N6     NN1     -0.77
ATOM H61    HN1      0.38
ATOM H62    HN1      0.38
```



GROUP			
ATOM C2'	CN7	0.14	!
ATOM H2''	HN7	0.09	!
ATOM O2'	ON5	-0.66	!
ATOM H2'	HN5	0.43	!
GROUP			
ATOM C3'	CN7		!
ATOM H3'	HN7	0.09	!
ATOM O3'	ON5	-0.66	!
ATOM H3T	HN5	0.43	!
GROUP			
ATOM C5'	CN8	-0.08	!
ATOM H5'	HN8	0.09	!
ATOM H5''	HN8	0.09	!
ATOM O5'	ON2	-0.62	!
ATOM PA	P	1.50	!
ATOM O1A	ON3	-0.82	!
ATOM O2A	ON3	-0.82	!
ATOM O3A	ON2	-0.74	!
ATOM PB	P	1.50	!
ATOM O1B	ON3	-0.90	!
ATOM O2B	ON3	-0.90	!
ATOM O3B	ON2	-0.74	!
GROUP			
ATOM CS1	CT2	0.21	!
ATOM HS11	HA	0.09	!
ATOM HS12	HA	0.09	!
ATOM CS2	CT1	0.17	!
ATOM HS21	HA	0.09	!
ATOM OS2	OH1	-0.66	!
ATOM HS22	H	0.43	!
ATOM CS3	CT1	0.17	!
ATOM HS31	HA	0.09	!
ATOM OS3	OH1	-0.66	!
ATOM HS32	H	0.43	!
ATOM CS4	CT1	0.17	!
ATOM HS41	HA	0.09	!
ATOM OS4	OH1	-0.66	!
ATOM HS42	H	0.43	!
ATOM CS5	CT2	0.22	!
ATOM HS51	HA	0.09	!
ATOM HS52	HA	0.09	!
GROUP			
ATOM NX1	NX1	-0.89	!
ATOM HX1	HX1	0.43	!
ATOM CX1	CX1	0.82	!
ATOM OX1	OX1	-0.61	
ATOM CX2	CX2	0.24	
ATOM CX3	CX3	0.37	
ATOM NX2	NX2	-0.81	
ATOM CX4	CX4	1.09	
ATOM OX2	OX2	-0.66	
ATOM NY1	NY1	-0.60	
ATOM CY1	CY1	0.60	
ATOM CY2	CY2	-0.30	
ATOM NY2	NY2	0.20	

H61 H62

\ /

N6

|

C6

// \

N1 C5--N7\\

| || C8-H8

C2 C4--N9/

/ \\ / \

H2 N3

O1B O1A H5' H4' O4' \

| | | / \

O3B-PB-O3A--PA-O5'-C5'---C4' C1'

| | | \ / \

O2B O2A H5'' C3'--C2' H1'

HS51-CS5-HS52 / \ / \

| O3' H3' O2' H2''

HS41-CS4-OS4-HS42 | |

| H3' H2'

HS31-CS3-OS3-HS32

|

HS21-CS2-OS2-HS22

|

HS11-CS1-HS12

\

HZ4 HZ61 HZ62

| | /

OX2 NX2 NY2 CZ4 CZ6--HZ63

\\ / \ / \ /

CX4 CX3 CY2 CZ3 HZ53

| | | |

NX1 CX2 CY1 CZ2---CZ5--HZ52

/ \ / \ // \ //

HX1 CX1 NY1 CZ1 HZ51

|

OX1 HZ1

ATOM	CZ1	CZ1	-0.46						
ATOM	HZ1	HZ1	0.24						
ATOM	CZ2	CZ2	0.19						
ATOM	CZ5	CT3	-0.44						
ATOM	HZ51	HA	0.09						
ATOM	HZ52	HA	0.09						
ATOM	HZ53	HA	0.09						
ATOM	CZ3	CZ3	0.18						
ATOM	CZ6	CT3	-0.48						
ATOM	HZ61	HA	0.09						
ATOM	HZ62	HA	0.09						
ATOM	HZ63	HA	0.09						
ATOM	CZ4	CZ4	-0.28						
ATOM	HZ4	HZ4	0.19						
BOND	PB	O3A	PB	O1B	PB	O2B	PB	O3B	O3A PA
BOND	PA	O1A	PA	O2A	PA	O5'	O3'	H3T	
BOND	O5'	C5'	C5'	C4'	C4'	O4'	C4'	C3'	O4' C1'
BOND	C1'	N9	C1'	C2'	N9	C4	N9	C8	C4 N3
BOND	C2	N1	C6	N6					
BOND	N6	H61	N6	H62	C6	C5	C5	N7	
BOND	C2'	C3'	C2'	O2'	O2'	H2'	C3'	O3'	
BOND	C1'	H1'	C2'	H2''	C3'	H3'	C4'	H4'	C5' H5'
BOND	C5'	H5''	C8	H8	C2	H2			
BOND	CS1	HS11	CS1	HS12	CS1	CS2	CS2	HS21	CS2 OS2
BOND	OS2	HS22	CS2	CS3	CS3	HS31	CS3	OS3	OS3 HS32
BOND	CS3	CS4	CS4	HS41	CS4	OS4	OS4	HS42	CS4 CS5
BOND	CS5	HS51	CS5	HS52	CS5	O3B	CS1	NY2	
BOND	NX1	HX1	NX1	CX1	CX1	OX1	CX1	CX2	CX2 CX3
BOND	CX3	NX2	NX2	CX4	CX4	OX2	CX2	NY1	NY1 CY1
BOND	CY1	CY2	CY2	NY2	NY2	CX3	NX1	CX4	
BOND	CY1	CZ1	CZ1	HZ1	CZ1	CZ2	CZ2	CZ5	CZ5 HZ51
BOND	CZ5	HZ52	CZ5	HZ53	CZ2	CZ3	CZ3	CZ6	CZ6 HZ61
BOND	CZ6	HZ62	CZ6	HZ63	CZ3	CZ4	CZ4	HZ4	CZ4 CY2
DOUBLE			N1	C6	N3	C2	C4	C5	N7 C8
IMPR	N6	C6	H61	H62					
IMPR	C6	N1	C5	N6					
IMPR	CZ2	CZ1	CZ3	CZ5					
IMPR	CZ3	CZ2	CZ4	CZ6					
IMPR	CY1	CZ1	NY1	CY2					
IMPR	NY2	CY2	CX3	CS1					
IMPR	CX2	NY1	CX1	CX3					
PATC	FIRS	NONE	LAST	NONE					

!!!!!!!!!!!!!! FAD PARAMETER !!!!!!!!!!!!!!!

**BONDS**

CX4	NX2	350.000	1.3700	!	FAD
CX4	NX1	350.000	1.4000	!	FAD
CX4	OX2	640.000	1.1900	!	FAD
NX2	CX3	350.000	1.2900	!	FAD
CX3	CX2	282.000	1.4700	!	FAD
CX2	CX1	282.000	1.5000	!	FAD
CX3	NY2	350.000	1.3600	!	FAD
CX2	NY1	350.000	1.2700	!	FAD
CX1	OX1	640.000	1.1900	!	FAD
CX1	NX1	350.000	1.3600	!	FAD
NX1	HX1	465.600	1.0000	!	FAD
NY2	CT2	310.000	1.4600	!	FAD
NY2	CY2	390.000	1.3900	!	FAD
CY2	CZ4	440.000	1.4000	!	FAD
CY2	CY1	440.000	1.4000	!	FAD
CY1	NY1	390.000	1.3700	!	FAD
CY1	CZ1	440.000	1.4000	!	FAD
CZ4	HZ4	370.000	1.0700	!	FAD
CZ4	CZ3	440.000	1.3800	!	FAD
CZ3	CZ2	440.000	1.4100	!	FAD
CZ2	CZ1	440.000	1.3700	!	FAD
CZ1	HZ1	370.000	1.0700	!	FAD
CZ2	CT3	345.000	1.5100	!	FAD
CZ3	CT3	345.000	1.5100	!	FAD
CT2	ON2	340.000	1.4400	!	FAD

**ANGLES**

OX2	CX4	NX1	98.00	118.6	!	FAD
OX2	CX4	NX2	98.00	123.1	!	FAD
NX1	CX4	NX2	70.00	118.3	!	FAD
CX3	NX2	CX4	60.00	120.5	!	FAD
NX2	CX3	CX2	65.00	124.1	!	FAD
NX2	CX3	NY2	90.00	120.1	!	FAD
CX2	CX3	NY2	65.00	115.8	!	FAD
CX1	CX2	CX3	65.00	116.8	!	FAD
CX1	CX2	NY1	65.00	119.1	!	FAD
CX3	CX2	NY1	65.00	124.1	!	FAD
CX2	CX1	OX1	80.00	124.0	!	FAD
CX2	CX1	NX1	65.00	112.9	!	FAD
NX1	CX1	OX1	98.00	123.2	!	FAD
CX1	NX1	HX1	18.00	117.0	!	FAD
CX4	NX1	CX1	60.00	127.4	!	FAD
CX4	NX1	HX1	18.00	117.0	!	FAD
CX3	NY2	CT2	44.00	119.5	!	FAD
CT2	NY2	CY2	80.00	119.6	!	FAD
CX3	NY2	CY2	60.00	120.9	!	FAD
NY2	CY2	CZ4	65.00	122.7	!	FAD
NY2	CY2	CY1	65.00	118.9	!	FAD
CZ4	CY2	CY1	70.00	118.4	!	FAD
CY2	CY1	NY1	65.00	121.2	!	FAD
CY2	CY1	CZ1	70.00	120.1	!	FAD
NY1	CY1	CZ1	65.00	118.7	!	FAD
CY1	NY1	CX2	60.00	119.2	!	FAD
CY2	CZ4	HZ4	31.00	120.2	!	FAD
CZ3	CZ4	HZ4	31.00	118.6	!	FAD

CY2	CZ4	CZ3	70.00	121.1	!	FAD		
CZ4	CZ3	CZ2	70.00	120.5	!	FAD		
CZ4	CZ3	CT3	70.00	119.3	!	FAD		
CZ2	CZ3	CT3	70.00	120.2	!	FAD		
CZ1	CZ2	CZ3	70.00	118.1	!	FAD		
CZ3	CZ2	CT3	70.00	121.2	!	FAD		
CZ1	CZ2	CT3	70.00	120.7	!	FAD		
CZ2	CZ1	CY1	70.00	121.8	!	FAD		
CZ2	CZ1	HZ1	31.00	121.2	!	FAD		
CY1	CZ1	HZ1	31.00	117.1	!	FAD		
CZ2	CT3	HA	35.00	115.0	!	FAD		
CZ3	CT3	HA	35.00	115.0	!	FAD		
NY2	CT2	HA	46.50	106.2	!	FAD		
NY2	CT2	CT1	65.00	109.5	!	FAD		
CN8	CN7	ON6B	110.00	109.0	!	FAD		
ON6B	CN7B	CN7	110.00	105.0	!	FAD		
NN2	CN7B	CN7	70.00	113.7	!	FAD		
CN7	CN7	CN7	53.35	111.0	8.0	2.561	!	FAD
CN7	CN8	ON2	75.70	110.1	!	FAD		
P	ON2	CT2	20.0	120.0	35.	2.33	!	FAD
CT2	CT1	OH1	75.700	110.1000	!	FAD		
ON2	CT2	CT1	75.700	110.1000	!	FAD		
ON2	CT2	HA	45.900	108.8900	!	FAD		

# DIHEDRAL

ON3	P	ON2	CT2	0.10	3	0.0	!	FAD
X	CX4	NX2	X	2.05	2	180.0	!	FAD
X	NX2	CX3	X	2.05	2	180.0	!	FAD
X	CX2	CX3	X	0.50	2	180.0	!	FAD
X	CX1	CX2	X	0.50	2	180.0	!	FAD
X	CX1	NX1	X	2.05	2	180.0	!	FAD
X	CX4	NX1	X	2.05	2	180.0	!	FAD
X	CX3	NY2	X	2.05	2	180.0	!	FAD
X	CY2	NY2	X	2.05	2	180.0	!	FAD
X	CY1	CY2	X	3.10	2	180.0	!	FAD
X	NY1	CY1	X	2.05	2	180.0	!	FAD
X	NY1	CX2	X	2.05	2	180.0	!	FAD
X	CY2	CZ4	X	3.10	2	180.0	!	FAD
X	CZ4	CZ3	X	3.10	2	180.0	!	FAD
X	CZ2	CZ3	X	3.10	2	180.0	!	FAD
X	CZ1	CZ2	X	3.10	2	180.0	!	FAD
X	CY1	CZ1	X	3.10	2	180.0	!	FAD
X	CZ2	CT3	X	0.10	3	0.0	!	FAD
X	CZ3	CT3	X	0.10	3	0.0	!	FAD
X	CT2	NY2	X	0.01	3	0.0	!	FAD
X	CN7	CN7	X	0.15	3	0.0	!	FAD
X	CT2	ON2	X	0.25	3	0.0	!	FAD
X	ON6B	CN7B	X	0.10	3	0.0	!	FAD
X	CN7	CN8	X	0.200	3	0.0	!	FAD
X	CN7	CN7B	X	0.15	3	0.0	!	FAD
X	CN7	ON6B	X	0.10	3	0.0	!	FAD
X	CN8	ON2	X	-0.10	3	0.0	!	FAD
ON2	P	ON2	CT2	0.95	2	0.0	!	FAD

**IMPROPER**

CZ1	X	X	CZ4	15.0000	0	0.0000	!	FAD
CZ2	X	X	CY2	15.0000	0	0.0000	!	FAD
CZ3	X	X	CY1	15.0000	0	0.0000	!	FAD
NY1	X	X	NY2	65.0000	0	0.0000	!	FAD
CY1	X	X	CX3	100.0000	0	0.0000	!	FAD
CX2	X	X	CY2	100.0000	0	0.0000	!	FAD
CX2	X	X	CX4	100.0000	0	0.0000	!	FAD
NY2	X	X	CX4	60.0000	0	180.0000	!	FAD
NY2	X	X	CX1	60.0000	0	180.0000	!	FAD
CX3	X	X	NX1	60.0000	0	0.0000	!	FAD
CX1	X	X	NX2	60.0000	0	0.0000	!	FAD
CZ3	X	X	CT3	130.0000	0	0.0000	!	FAD
CY1	X	X	CY2	15.0000	0	0.0000	!	FAD
CZ2	X	X	CT3	130.0000	0	0.0000	!	FAD
NY2	X	X	CT2	45.0000	0	0.0000	!	FAD
CX2	X	X	CX3	100.0000	0	0.0000	!	FAD
CX4	X	X	OX2	147.0000	0	0.0000	!	FAD
NX1	X	X	HX1	18.0000	0	0.0000	!	FAD
CX1	X	X	OX1	147.0000	0	0.0000	!	FAD

**NONBONDED**

CX1	0.000000	-0.141000	1.870000	!	FAD
CX2	0.000000	-0.141000	1.870000	!	FAD
CX3	0.000000	-0.141000	1.870000	!	FAD
CX4	0.000000	-0.141000	1.870000	!	FAD
CY1	0.000000	-0.141000	1.870000	!	FAD
CY2	0.000000	-0.141000	1.870000	!	FAD
CZ1	0.000000	-0.141000	1.870000	!	FAD
CZ2	0.000000	-0.141000	1.870000	!	FAD
CZ3	0.000000	-0.141000	1.870000	!	FAD
CZ4	0.000000	-0.141000	1.870000	!	FAD
HX1	0.000000	-0.049800	0.800000	!	FAD
HZ1	0.000000	-0.042000	1.330000	!	FAD
HZ4	0.000000	-0.042000	1.330000	!	FAD
NX1	0.000000	-0.090000	1.830000	!	FAD
NX2	0.000000	-0.090000	1.830000	!	FAD
NY1	0.000000	-0.090000	1.830000	!	FAD
NY2	0.000000	-0.090000	1.830000	!	FAD
OX1	0.000000	-0.120000	1.700000	!	FAD
OX2	0.000000	-0.120000	1.700000	!	FAD

```

=====
! CHARMM Parameter Molypdopterin, NOT RECOMMENDED FOR CLASSICAL MD
=====

```

!!!!!!!!!!!!!! **TOPOLOGY** !!!!!!!!!!!!!!!

```

MASS 210 MO 95.94000 MO ! FOR MOLYBDENUM
MASS 211 OMO 15.99900 O ! MO==O
MASS 212 SMO 32.06000 S ! MO==S
MASS 213 SEL 78.96 SE ! MO==SE

```

```

RESI MOC -3.00 ! Mo-cofactor: -1 (Mo) + -2 (phosphate)
GROUP ! charge -1 Part 1 (on which PATCH MOPS is applied)

```

```

ATOM Mo MO 0.947 !
ATOM OM1 OMO -0.552 ! OM1
ATOM OM2 OH1 -0.923 ! ||
ATOM HOM2 H 0.503 ! S7'---Mo==SR1
ATOM SR1 SMO -0.271 ! / / \
ATOM S8' SS -0.172 ! (C6)---C7' S8' OM2
ATOM S7' SS -0.055 ! \ \ / |
ATOM C8' CE1 -0.235 ! (C3')---C8' HOM2
ATOM C7' CE1 -0.242 !

```

```

GROUP ! charge: 0.00 Part 2

```

```

ATOM N1 NN3G -0.74
ATOM C2 CN2 0.75 !
ATOM N2 NN1 -0.68 !
ATOM H21' HN1 0.35 !
ATOM H22' HN1 0.32 ! H6'
ATOM N3 NN2G -0.34 ! O4 H5' |
ATOM H3 HN2 0.26 ! || | | (C7')
ATOM C4 CN1 0.54 ! C4 N5 | /
ATOM O4 ON1 -0.51 ! / \ / \ | / (C8')
ATOM C9 CN5G 0.00 ! H3-N3 C9 C6 |
ATOM C10 CN5 0.26 ! | || | |
ATOM N8 N2 -0.61 ! C2 C10 C7 C3'---(C4')
ATOM H8' HC 0.40 ! / \ \ / \ / | \ / \
GROUP ! H21'-N2 N1 N8 | O3' H3'
ATOM N5 N2 -0.61 ! | | |
ATOM H5' HC 0.40 ! H22' H8' |
ATOM C6 CNT2 0.12 ! H7
ATOM H6' HA 0.09
ATOM C7 CN7B 0.16
ATOM H7 HN7 0.09
ATOM O3' ON6B -0.50
ATOM C3' CN7 0.16
ATOM H3' HN7 0.09

```

```

GROUP ! charge: -2.0 Part 3, from RESI MP_2

```

```

ATOM C4' CN9 -0.18 ! (C3')
ATOM H101 HN9 0.09 ! |
ATOM H102 HN9 0.09 ! H102--C4'--H101
GROUP ! |
ATOM O4' ON2 -0.40 ! O4'
ATOM P P 1.10 ! |
ATOM O1P ON3 -0.90 ! O3P(-)--P(+)--O1P(-)
ATOM O2P ON3 -0.90 ! |
ATOM O3P ON3 -0.90 ! O2P(-)

```

!!!!!!!!!!!!!! TOPOLOGY PATCH !!!!!!!!!!!!!!!

PRES MOPS -1.000 ! patch to obtain MO(IV) species

```

GROUP
ATOM Mo MO 0.775 ! OM1
ATOM OM1 OMO -0.498 ! ||
ATOM SR1 SMO -0.576 ! S7'---Mo---SR1---HSR
ATOM HSR HS 0.133 ! / /
ATOM S8' SS -0.514 ! ---C7' S8'
ATOM S7' SS -0.261 ! \ /
ATOM C8' CE1 0.072 ! ---C8'
ATOM C7' CE1 -0.131 !

```

DELETE ATOM OM2

DELETE ATOM HOM2

BOND SR1 HSR

ANGLE HSR SR1 MO

IC SR1 MO OM1 S8' 2.1700 107.51 -100.79 90.00 0.0000

IC HSR SR1 MO OM1 1.3900 101.66 84.64 000.00 0.0000

!!!!!!!!!!!!!! MOCO PARAMETER !!!!!!!!!!!!!!!

**BONDS** !=====dummy parameters=====

```

SMO HS 275.000 1.385 ! from S HS
MO OMO 00.00 1.73 ! from lvdv
OC MO 00.00 2.15
MO SMO 00.00 2.17 ! from lvdv
MO OH1 250.0 1.97 ! from lvdv K from QM !!!!!!!
MO SS 0.0 0.0
SS CE1 0.0 0.0

```

**ANGLES** !=====dummy parameters=====

```

OMO MO OH1 50.000 105.495 ! From QM
OMO MO OMO 00.000 107.507 ! x-ray
OMO MO SMO 00.000 107.827 ! From QM
OMO MO SS 00.000 101.542 ! From QM
OH1 MO SMO 70.000 100.000 !
OH1 MO SS 70.000 075.000 !
SMO MO SS 00.000 077.239 !
SS MO SS 00.000 079.321 ! From QM
MO OH1 H 35.000 119.659 ! From QM
MO SS CE1 00.000 107.736 ! From QM
SS CE1 CE1 00.000 122.122 ! From QM
HS SMO MO 30.000 100.000 ! FROM QM
CC OC MO 35.000 120.000 !

```

**DIHEDRAL** !=====dummy parameters=====

```

MO SS CE1 CE1 00.00 1 3.908 ! from lvlb
MO SS CE1 CN7 00.00 1 -165.552 ! from lvlb
MO SS CE1 CNT2 00.00 1 178.353 ! from lvlb
OH1 MO SS CE1 20.0000 4 180.0
SMO MO SS CE1 0.0000 4 0.00
H OH1 MO SS 4.5000 2 0.00
OMO MO OH1 H 4.5000 2 0.00
SMO MO OH1 H 4.5000 2 0.00
OMO MO SS CE1 00.00 1 89.770 ! from lvlb
SS MO SS CE1 0.0000 1 0.00

```

**NONBONDED** !=====dummy parameters=====

```

MO 0.0100 0.0000 0.6500 !
OMO 0.0000 -0.1200 1.7000 0.000000 -0.120000 1.400000 !
SMO 0.0000 -0.4700 2.200000
SEL 0.0000 -0.4700 2.200000 ! PURE DUMMIE FROM SS

```

```

=====
! CHARMM topology for molybdenum-bond 2,8-dioxo-6-methylpurine
=====

```

!!!!!!!!!!!!!!!!!!!! TOPOLOGY !!!!!!!!!!!!!!!!!!!!!

```

MASS 129    CN3T 12.011000 C ! Nucleic acid aromatic carbon, Thy C5
MASS 139    CN9 12.011000 C ! Nucleic acid carbon (equivalent to CT3)
MASS 113    HN9  1.008000 H ! Bound to CN9
MASS 323    7N2U 14.00700  N ! XAN-specific angles assigned
MASS 324    7HN2  1.00800  H ! XAN-specific angles assigned
MASS 325    7C1T 12.011000 C ! XAN-specific angles assigned
MASS 326    7ON1 15.999400 O ! XAN-specific angles assigned
MASS 327    7N2B 14.00700  N ! XAN-specific angles assigned
MASS 328    7CN5 12.011000 C ! XAN-specific angles assigned
MASS 329    7C5G 12.011000 C ! XAN-specific angles assigned
MASS 330    7CN1 12.011000 C ! XAN-specific angles assigned
MASS 331    7CN4 12.011000 C ! XAN-specific angles assigned
MASS 332    7HN3  1.00800  H ! XAN-specific angles assigned
MASS 333    7NN4 14.00700  N ! XAN-specific angles assigned

```

```

RESI ME6X      -1.00  !  !!!! ESP charges  !!!!!
GROUP          !
!
ATOM N1      7N2U   -0.99  !
ATOM C2      7C1T    1.07  !
ATOM O2      7ON1   -0.75  !
ATOM N3      7N2B   -0.91  !
ATOM H3      7HN2    0.44  !
ATOM C4      7CN5    0.95  !
ATOM C5      7C5G   -0.48  !
ATOM C6      CN3T    0.78  !
ATOM C6'     CN9   -0.64  !
ATOM H61     HN9    0.16  !
ATOM H62     HN9    0.16  !
ATOM H63     HN9    0.16  !
ATOM N7      7N2B   -0.60  !
ATOM H7      7HN2    0.37  !
ATOM C8      7CN4    0.95  !
ATOM O8      7ON1   -0.73  !
ATOM N9      7NN4   -0.94  !

BOND      N1 C2    C2 N3    C2 O2    N3 C4    C4 N9    C6 C6'
BOND      C8 O8    C8 N7    N7 H7    N7 C5    C5 C6    C6 N1
BOND      C4 C5    N9 C8    C6' H61  C6' H62  C6' H63    H3 N3
!Improper and DONO/ACC from URA (& patch PYRU) and GUA (& patch PURG)
IMPR C2      N3      N1      O2          C6      N1      C5      C6'          N7      C5      H7      C8
DONO H7      N7
DONO H3      N3
ACCE O2      C2
ACCE N1
ACCE N9
PATCH FIRS NONE LAST NONE

```



!!!!!!!!!!!!!!!!!!!!!! ME6X PARAMETER !!!!!!!!!!!!!!!!!!!!!!!!!!!!!

**BONDS**

7N2U	7C1T	340.0	1.383	!U,T adm jr. 11/97
7C1T	7N2B	302.0	1.348	!U,T adm jr. 11/97
7C1T	7ON1	860.0	1.230	!nad/ppi, jjp1/adm jr. 7/95
7CN5	7N2B	320.0	1.343	!from URA: CN3 NN2B
7CN5	7NN4	310.0	1.365	!from GUA CN5G NN4
CN3T	CN9	230.0	1.478	!T, adm jr. 11/97
7CN4	7ON1	660.0	1.20	!U,A,G par_a4 adm jr. 10/2
7CN4	7N2B	300.0	1.378	!G, adm jr. 11/97
7HN2	7N2B	474.0	1.01	!SAME AS IN PYRO and PURG patch
7C5G	7N2B	302.0	1.375	!from GUA CN5 NN2B
7C5G	CN3T	560.0	1.320	! from THY CN3 CN3T
CN3T	7N2U	400.0	1.342	!from A CN2 NN3A
7CN5	7C5G	320.0	1.350	!G, adm jr. 11/97
7CN4	7NN4	400.0	1.305	!G,A, adm jr. 11/97
CN9	HN9	322.0	1.111	!alkanes

**ANGLES**

7C1T	7N2U	7HN2	40.5	114.2	!U, h3
7CN1	7N2U	7HN2	40.5	115.8	!U
7C1T	7N2U	7CN1	50.0	130.0	!U
7N2U	7C1T	7ON1	100.0	122.7	!U
7N2B	7C1T	7ON1	100.0	123.2	!U, o2
7N2B	7C1T	7N2U	50.0	114.1	!U
7C1T	7N2B	7CN5	70.0	121.6	! from URA CN1T NN2B CN3
7C1T	7N2B	7HN2	40.5	122.0	!U, h1
7CN5	7N2B	7HN2	32.0	117.2	!U !!! from URA CN3 NN2B HN2
7N2B	7CN5	7NN4	140.0	126.9	! !!!! from GUA NN2B CN5 NN3G
7CN5	7C5G	7N2B	100.0	104.5	!G
7NN4	7CN5	7C5G	100.0	112.2	! from GUanin CN5 CN5G NN4
7CN4	7NN4	7CN5	120.0	103.8	! from GUAnini CN4 NN4 CN5G
7NN4	7CN4	7HN3	39.0	124.7	!h8, G,A
7N2B	7CN4	7NN4	100.0	113.0	! 5R)
7HN3	7CN4	7N2B	40.0	122.3	! h8 (NN4 CN4 HN3 124.8)
7CN4	7N2B	7C5G	100.0	106.6	! from GUA CN4 NN2B CN5
7HN2	7N2B	7C5G	30.0	125.4	! from GUA CN5 NN2B HN2
7HN2	7N2B	7CN4	30.0	128.0	! from GUA CN4 NN2B HN2
7C5G	7CN5	7N2B	70.0	121.1	! from CN5G CN5 NN3G
7CN1	7C5G	7N2B	125.0	131.8	! from CN1 CN5G NN4
7CN1	7C5G	7CN5	70.0	123.6	!
7C5G	7CN1	7N2U	100.0	109.8	! from : CN5G CN5 NN2B
7C5G	7CN1	7ON1	50.0	127.8	!
7N2U	7CN1	7ON1	50.0	122.4	! from NN2G CN1 ON1
7C1T	7N2U	CN3T	50.0	130.0	!U
7CN5	7C5G	CN3T	70.0	123.6	!
CN3T	7C5G	7N2B	125.0	131.8	!
7N2U	CN3T	7C5G	100.0	109.8	! from : CN5G CN5 NN2B
7N2U	CN3T	CN9	70.0	113.7	!9-E-G, adm jr
7C5G	CN3T	CN9	38.0	118.7	!T, c5 methyl
CN3T	CN9	HN9	33.43	110.10	22.53 2.179 ! Alkanes, sacred
HN9	CN9	HN9	35.500	108.40	5.40 1.802 !alkane update
7N2U	CN3T	CN9	70.0	113.7	!9-E-G, adm jr
7N2B	7CN4	7ON1	100.0	127.5	!U, o2
7ON1	7CN4	7NN4	100.0	127.5	!U, o2

# DIHEDRALS

7HN3	7CN4	7NN4	7CN5	5.6	2	180.0	!	from	GUA	HN3	CN4	NN4	CN5G
7HN3	7CN4	7N2B	7HN2	0.0	2	180.0	!	from	GUA	HN3	CN4	NN2B	HN2
7HN3	7CN4	7N2B	7C5G	5.6	2	180.0	!	from	GUA	HN3	CN4	NN2B	CN5
7CN1	7C5G	7N2B	7CN4	2.0	2	180.0	!	from	GUA	CN1	CN5G	NN4	CN4
7CN4	7N2B	7C5G	7CN5	6.0	2	180.0	!	from	GUA	CN4	NN2B	CN5	CN5G
7C5G	7CN5	7NN4	7CN4	6.0	2	180.0	!	from	GUA	CN5	CN5G	NN4	CN4
7N2B	7CN5	7NN4	7CN4	2.0	2	180.0	!	from	GUA	NN3G	CN5	NN2B	CN4
7N2B	7CN5	7C5G	7N2B	10.0	2	180.0	!	from	GUA	NN3G	CN5	CN5G	NN4
7NN4	7CN5	7C5G	7N2B	10.0	2	180.0	!	from	GUA	NN2B	CN5	CN5G	NN4
7ON1	7CN1	7C5G	7N2B	0.0	2	180.0	!	from	GUA	ON1	CN1	CN5G	NN4
7N2U	7CN1	7C5G	7N2B	2.0	2	180.0	!	from	GUA	NN2G	CN1	CN5G	NN4
7HN2	7N2B	7CN4	7NN4	1.2	2	180.0	!	from	GUA	HN2	NN2B	CN4	NN4
7CN5	7NN4	7CN4	7N2B	16.0	2	180.0	!	from	GUA	CN5G	NN4	CN4	NN2B
7C1T	7N2B	7CN5	7CN1	2.0	2	180.0	!	from	GUA	CN2	NN3G	CN5	NN2B
7CN5	7N2B	7C1T	7N2U	1.5	2	180.0	!	from	URA	CN3	NN2B	CN1T	NN2U
7ON1	7CN1	7C5G	7CN5	14.0	2	180.0	!	from	GUA	ON1	CN1	CN5G	CN5
7CN5	7C5G	7CN1	7N2U	0.2	2	180.0	!	from	GUA	CN5	CN5G	CN1	NN2G
7C5G	7CN1	7N2U	7C1T	1.5	2	180.0	!	from	URA	CN1T	NN2U	CN1	CN3
7N2B	7CN5	7C5G	7CN1	2.0	2	180.0	!	from	GUA	NN3G	CN5	CN5G	CN1
7ON1	7CN1	7N2U	7HN2	0.0	2	180.0	!	from	URA	ON1	CN1T	NN2U	HN2
7N2B	7C1T	7N2U	7CN1	1.5	2	180.0	!	from	URA	NN2B	CN1T	NN2U	CN1
7C5G	7CN5	7N2B	7HN2	1.6	2	180.0	!	from	URA	CN3	CN3	NN2B	HN2
7N2U	7C1T	7N2B	7HN2	1.6	2	180.0	!	from	URA	NN2U	CN1T	NN2B	HN2
7ON1	7C1T	7N2B	7HN2	0.0	2	180.0	!	from	URA	ON1	CN1T	NN2B	HN2
7ON1	7C1T	7N2U	7HN2	0.0	2	180.0	!	from	URA	ON1	CN1T	NN2U	HN2
7N2B	7C1T	7N2U	7HN2	3.8	2	180.0	!	from	URA	NN2B	CN1T	NN2U	HN2
7C5G	7CN1	7N2U	7HN2	4.8	2	180.0	!	from	URA	CN3T	CN1	NN2U	HN2
7C5G	7CN5	7N2B	7C1T	1.5	2	180.0	!	from	URA	CN3	CN3	NN2B	CN1T
7ON1	7CN1	7N2U	7C1T	14.0	2	180.0	!	from	GUA	ON1	CN1	NN2G	CN2
7C1T	7N2B	7CN5	7NN4	2.0	2	180.0	!	from	GUA	CN2	NN3G	CN5	NN2B
7HN2	7N2B	7C5G	7CN5	1.2	2	180.0	!	from	Gua	HN2	NN2B	CN5	CN5G
7C5G	7N2B	7CN4	7NN4	16.0	2	180.0	!	from	GUA	CN5G	NN4	CN4	NN2B
7HN2	7N2B	7CN5	7NN4	1.2	2	180.0	!	from	XHN2	XN2B	XCN4	XNN4	
7CN1	7C5G	7N2B	7HN2	1.2	2	180.0	!	from	XHN2	XN2B	XCN4	XNN4	
X	7C1T	7N2U	X	0.9	2	180.0	!						
X	7C1T	7N2B	X	0.9	2	180.0	!	from	X	CN1	NN2	X	(thymines)
X	7C5G	7CN5	X	0.0	2	180.0	!						
7N2U	CN3T	7C5G	7CN5	1.5	2	180.0	!	from	URA	CN3	NN2B	CN1T	NN2U
7N2U	CN3T	7C5G	7N2B	2.0	2	180.0	!	from	GUA	NN2G	CN1	CN5G	NN4
7C1T	7N2U	CN3T	7C5G	1.5	2	180.0	!	from	URA	CN1T	NN2U	CN1	CN3
CN3T	7C5G	7N2B	7HN2	1.2	2	180.0	!	from	Gua	HN2	NN2B	CN5	CN5G
CN3T	7C5G	7N2B	7CN4	2.0	2	180.0	!	from	GUA	CN1	CN5G	NN4	CN4
7N2U	CN3T	CN9	HN9	0.46	3	0.0	!	from	CN1	CN3T	CN9	HN9	
7C1T	7N2U	CN3T	CN9	5.6	2	180.0	!	from	NN2B	CN1	CN3T	CN9	
7CN5	7C5G	CN3T	CN9	5.6	2	180.0	!	from	NN2B	CN1	CN3T	CN9	
7C5G	CN3T	CN9	HN9	0.46	3	0.0	!	from	CN1	CN3T	CN9	HN9	
CN9	CN3T	7C5G	7N2B	5.6	2	180.0	!	from	NN2B	CN1	CN3T	CN9	
X	7NN4	7CN4	X	0.9	2	180.0	!	From	X	CN1	NN2	X	(thymines)
X	7N2B	7CN4	X	0.9	2	180.0	!	From	X	CN1	NN2	X	(thymines)

# IMPROPER

7C1T	X	X	7ON1	90.0	0	0.0	!	U
7N2B	X	X	7CN4	7.0	0	0.0	!	G, adm jr. 11/97
CN9	X	X	CN3T	14.0	0	0.0	!	T, adm jr. 11/97

**NONBONDED**

7N2U	0.0	-0.20	1.85	
7HN2	0.0	-0.0460	0.2245	
7C1T	0.0	-0.10	1.9000	
7ON1	0.0	-0.1200	1.70	
7N2B	0.0	-0.20	1.85	! From NN2, for N9 in guanines
7CN5	0.0	-0.075	1.9000	
7C5G	0.0	-0.075	1.9000	
7CN1	0.0	-0.10	1.9000	
7CN4	0.0	-0.075	1.9000	
7NN4	0.0	-0.20	1.85	
7HN3	0.0	-0.046	1.1000	!adm jr. aromatic Hvdw
CN3T	0.0	-0.09	1.9000	! T, adm jr.
CN9	0.0	-0.0780	2.040	0.0 -0.01 1.90 !
HN9	0.0	-0.0240	1.3400	! Hydrogen bound to CN9

## 6. Complete references 32, 36, and 54

[32] CHARMM force field:

(a) MacKerell, A.D., Jr.; Feig, M.; Brooks, C.L., III *J. Comput. Chem.* **2004**, 25, 1400-1415.

(b) Foloppe, N.; Mackerell, A. D., Jr. *J. Comput. Chem.* **2000**, 21, 86-104.

(c) MacKerell, Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R.L.; Evanseck, J.D.; Field, M.J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, 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.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, 102, 3586-3616.

[36] ChemShell package:

(a) 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.; Terstegen, 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.

(b) ChemShell is a modular QM/MM program developed in the European QUASI project under the coordination of P. Sherwood (see <http://www.chemshell.org>).

[55] Gaussian 03:

Gaussian 03, Revision D.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A.

Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

