J|A|C|S ARTICLES

Published on Web 09/29/2009

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^{-1}, 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.[1] 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_2 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.[2,3] Since then, each step of the catalytic cycle in XO has been intensely investigated. The intraenzymatic electron transfer and the Mo(V) intermediates have been studied by rapid-freeze-ESR techniques.[4−13] The oxidative half-reaction of XO has also drawn some attention, since it may be a source of reactive oxygen species, such as

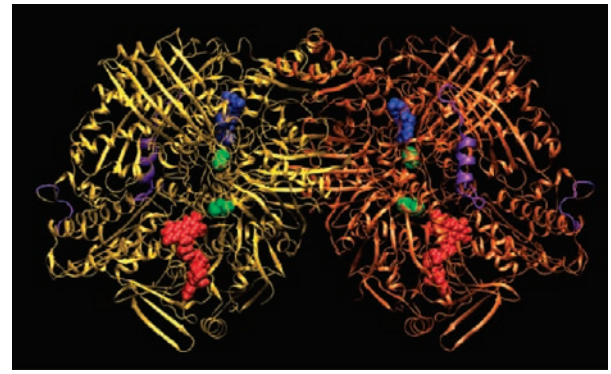
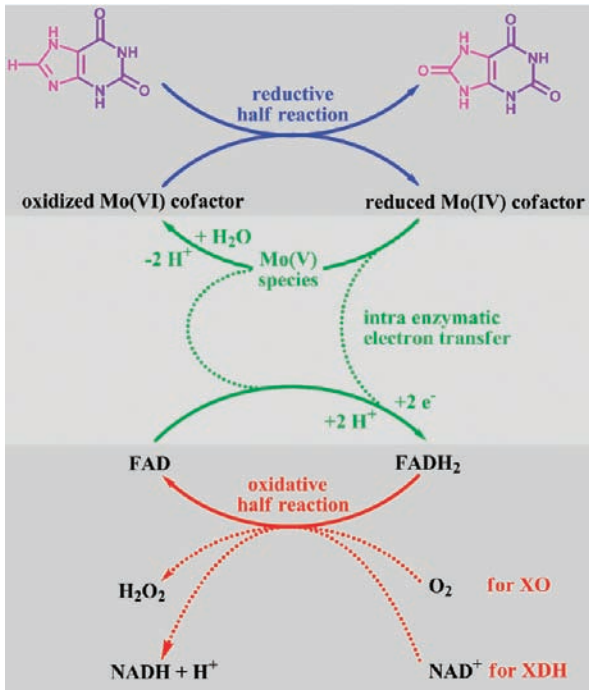


Figure 1. Bovine xanthine oxidase dimer (yellow and orange) with its prosthetic groups: the molybdopterin cofactor (blue), the S_2Fe_2 clusters (green), and FAD (red). Residues that are missing in the crystal structure and have thus been modeled are drawn in purple.

(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.

10.1021/ja9045394 CCC: $40.75

Scheme 1. Sketch of the Mechanism of Xanthine Oxidase and Xanthine Dehydrogenase[a]



[a] 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.[14−18] 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.[19] For many years, allopurinol has been the "gold standard" inhibitor, but there is an increasing number of other potent inhibitors.[20−28]

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 oxidoreductase (AOR),[29,30] crystal structures of XO and XDH,[31,32] kinetic experiments with xanthine and related purine substrates for wild-type xanthine oxidase,[33−39] and kinetic experiments with mutants of XO or XDH.[39−42] 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.[43] In a recent EXAFS analysis, the labile oxygen was experimentally identified to be a hydroxide ligand[44] instead of a bound water molecule as previously assumed. Based on the pH-dependence of the XO activity toward xanthine and lumazine[36] as well as 1-methylxanthine,[45] Lewis-base catalysis by glutamate (Glu1261 in XO and Glu869 in AOR) has been proposed[30−32,38,39,42] and confirmed by QM/MM calculations for the similar AOR system.[46] 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.[45,47]

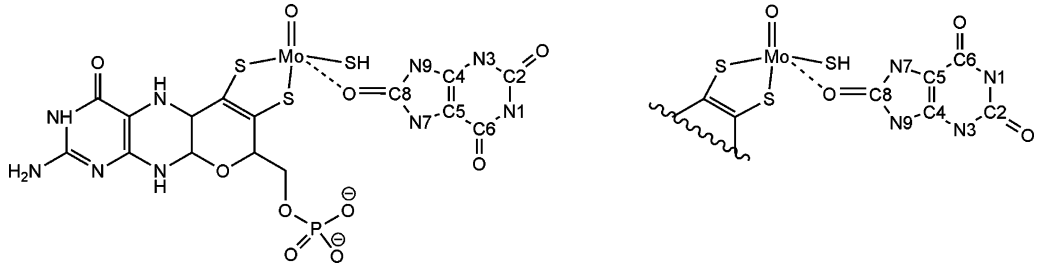
Most experimental results[48−50] have been analyzed using simple Michaelis−Menten assumptions:[51,52]

$$[\mathrm{XO}]_{\mathrm{ox}} + \mathrm{Xant} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\mathrm{XO}]_{\mathrm{ox}} \cdot \mathrm{Xant} \xrightarrow{k_2} [\mathrm{XO}]_{\mathrm{red}} + \mathrm{Uric\ acid}$$

For the reductive half-reaction, an activation barrier of E_A = 14.5 kcal mol^{-1} was obtained from an Arrhenius plot.[50] Using the Eyring equation[53] 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[a]



[a] For details on the protonation state, see section 2.2.

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

The simple Michaelis−Menten scheme has been refined[47,49] by considering the formation of the product complex and product release as two separate steps with rate constants $k_{2'}$ and $k_{2''}$, respectively:

$$[XO]_{ox} + Xant \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [XO]_{ox} \cdot Xant \xrightarrow{k_{2'}} [XO]_{red} \cdot Uric\ acid \xrightarrow{k_{2''}} [XO]_{red} + Uric\ acid$$

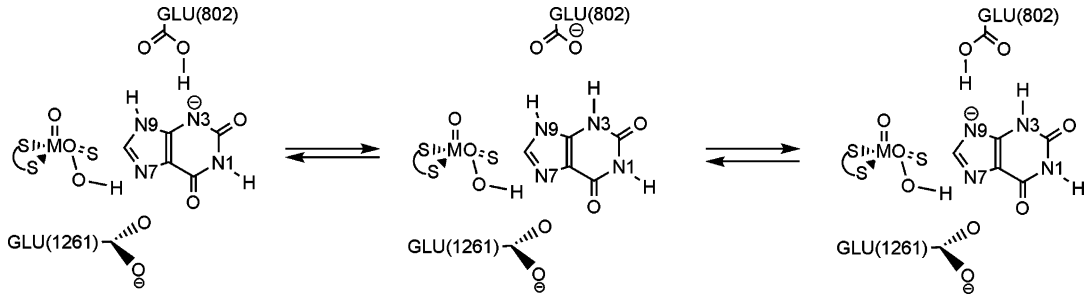
For xanthine[54] and 2-oxo-6-methylpurine,[35] product release has proven to be slower than formation of the product complex, and the barriers of free energy obtained for xanthine (using XDH) are $\Delta G^{\ddagger}(k_{2'}) = 13.4$ kcal mol^{-1} and $\Delta G^{\ddagger}(k_{2''}) = 15.9$ kcal mol^{-1}.[54] From temperature-dependent steady-state and transient kinetics of the reductive half-reaction, Mondal and Mitra[47] obtained similar results using XO with $\Delta G^{\ddagger}(k_{2'}) = 15.1$ kcal mol^{-1} and $\Delta G^{\ddagger}(k_{2''}) = 16.2$ 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[30] 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).[55] On the other hand, there is crystallographic evidence that allopurinol[23,56] and 2-oxo-6-methylpurine,[32] 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.[32,42] By contrast, a recent mechanistic study[39] 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.[35] 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.[57]

On the theoretical side, there have been a number of detailed QM model studies that mainly focused on the oxidation of formaldehyde[58−60] and the less active formamide.[60−63] Our recent QM/MM work has addressed the reaction mechanism in the similar AOR system.[46] Further theoretical investigations have been performed on 6-substituted 4-quinazolinones[64] and imidazole[65] 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.[66] 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.[65] 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.[46]

(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)[a]



[a] 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 molybdopterin 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.[31] 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)[55] and with 2,8-dioxo-6-methylpurin in the binding pocket (PDB: 3B9J)[32] were not yet available, so we used the crystal structure of bovine XO in the XDH form from the Protein Data Bank (PDB: 1FO4).[24] 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.[67] The glycerol molecules and the calcium ions present in the PDB file were kept. We used SWISS-MODEL[68−71] 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*[72] 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*.[73]

Missing hydrogen atoms in the X-ray structure were added using the CHARMM program,[74,75] 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 molybdopterin and Fe_2S_2 cluster from previous calculations.[46] The FAD parameters were taken from the literature.[76] 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 molybdopterin cofactor of chain A, C8:URIC. Within this region and for all following MD simulations, the $[Mo(S_2C_2H_2)(=O)(OR_{unfixed})(-SH)]^{2-}$ moiety and the Fe_2S_2 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[75] as implemented in the CHARMM program package.[74] 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.[77] 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.[36] 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,[66] 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[30−32,38,39,42] that Glu1261 is initially deprotonated. This is consistent with the results of our recent QM/MM study on AOR,[46] which explicitly considered both protonation states of Glu1261 (Glu869 in AOR). Mutation studies of XDH from *Rhodobacter capsulatus*[42] as well as human XO[39] 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.; Bruccoleri, 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.

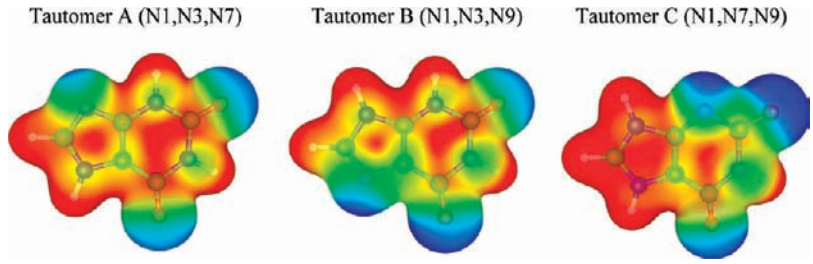


Figure 2. Electrostatic potential for different protonation states of xanthine mapped on the density isosurface (0.01e). Blue (red) color represents a negative (positive) potential. A large positive potential at C8 (leftmost ring atom) will facilitate nucleophilic attack.

Table 1. Relative Energies of Tautomers A−C in kcal mol^{-1}

	ΔE^a	$\Delta G^{gas\ b}$	$\Delta G^{solution\ b}$	ΔE^c	ΔE^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

[a] B3LYP/6-31G(d).[78] [b] B3LYP/6-31++G**.[79] [c] B3LYP/6-31+G**. [d] B3LYP/def2-TZVP.

accumulation on XAN:O6 in the course of the reaction with its positive countercharge.[32,42] 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,[23,25,39,56] 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.[80] 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,[81,82] using the program package TURBOMOLE[83] to obtain the QM-(DFT) wave functions as well as the corresponding energies and gradients. MM energy and gradient were evaluated by DL_POLY,[84] which is provided in the ChemShell package, using the CHARMM topology and parameter data. We employed electrostatic embedding for the QM region.[85] 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[86,87] for all setups. We also cross-checked the results for setups A and G using the alternative *L2* link atom scheme.[81,88]

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

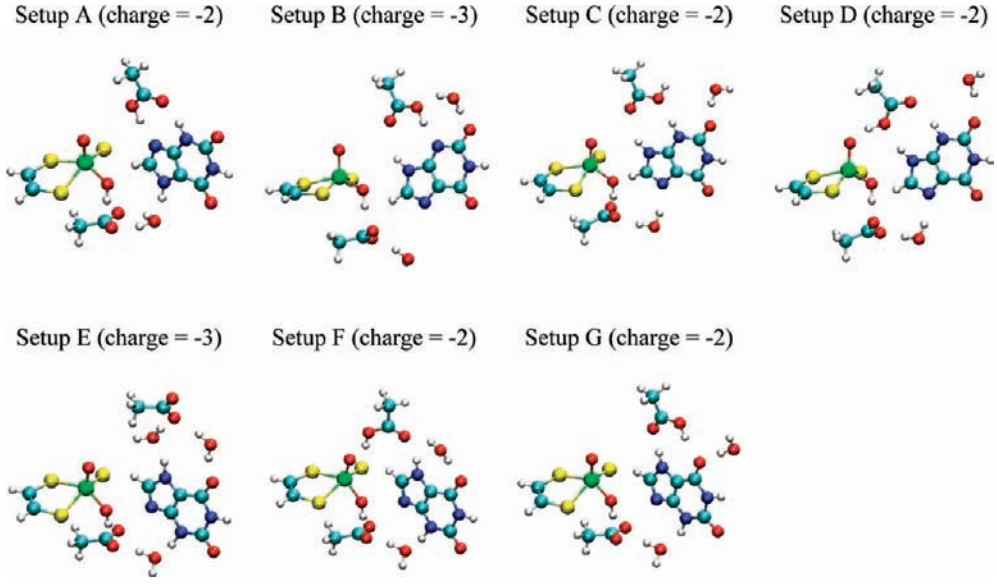


Figure 3. QM regions of the seven chosen setups.

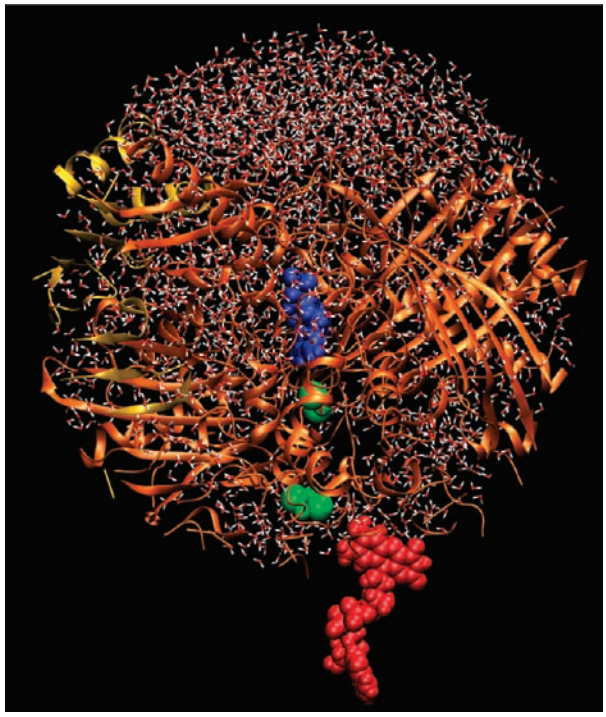


Figure 4. System used in QM/MM calculations to establish the pathways within setup A−G. Same color coding as in Figure 1.

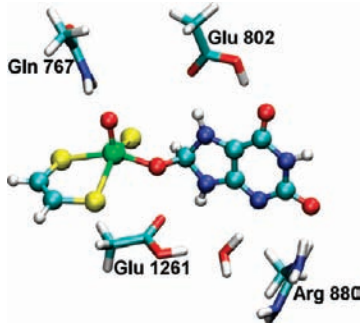


Figure 5. Extended QM region used for single-point energy evaluations (see text and section 3.9). The calculations on the smallest possible gas-phase model for setup G included only the atoms represented in balls and sticks (see section 3.10). Shown is the tetrahedral intermediate (IM2) of setup G.

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[89] 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[90−94] to determine various minima on the potential energy surface, employing the resolution of the identity (RI) approximation.[95,96] 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[97] in combination with transition-state optimization to determine the minimum energy pathway and the associate transition state. All stationary points were reoptimized using the B3LYP hybrid functional[90−92,98−100] as implemented in TURBOMOLE.

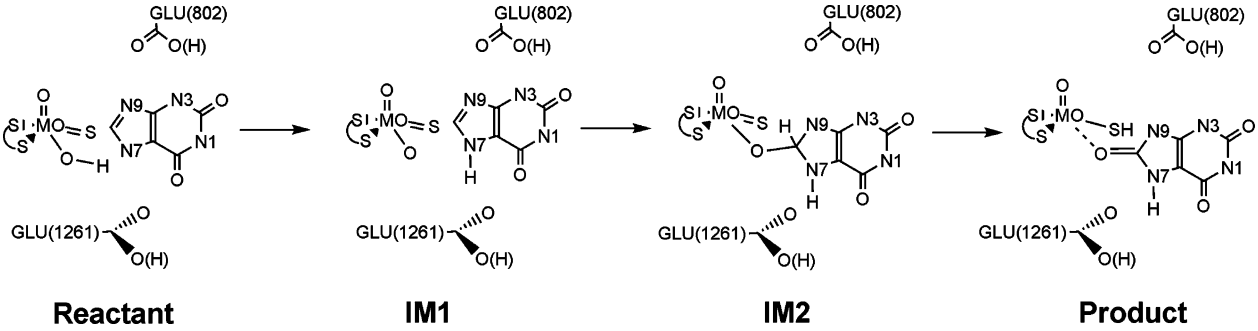
Our standard basis set (B1) was composed as follows: Mo, Lanl2DZ[101] with an additional f polarization function;[102] S, Lanl2DZ[103] with an additional d polarization function;[104] and all other atoms (H, C, O, N) 6-31+G**.[105,106] The BP86 calculations employed the def2-TZVP auxiliary basis set.[107]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

3. Results

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

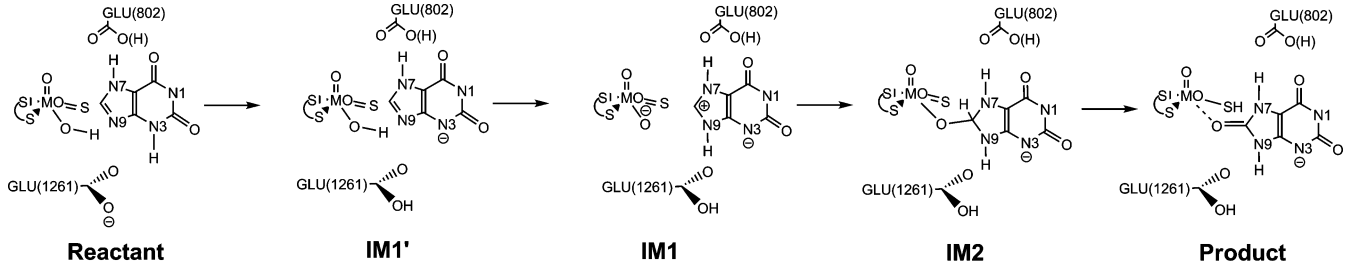
(80) Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2177–2186.
(81) Sherwood, P.; et al. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1–28.
(82) www.chemshell.org.
(83) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
(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.
(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.

Scheme 4. General Reaction Schemes for Setups A−D[a]



[a] 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[a]



[a] 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[46] (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[46] 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/MM energies (i.e., including QM, MM, and QM/MM interaction terms), without zero-point and thermal corrections.

Table 2. QM/MM Energies in kcal mol^{-1} Calculated for Setup A, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	18.8		21.2		19.0		21.7	

[a] Activation barriers relative to the preceding minima are given in parentheses.

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,[77] 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^{-1}). As in the case of AOR,[46] the BP86 and B3LYP energetics are generally similar, except for the Mo(IV) product

Table 3. QM/MM Energies in kcal mol^{-1} Calculated for Setup B, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	19.8		18.5		19.9		18.6		20.4	

[a] 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.

Table 4. QM/MM Energies in kcal mol^{-1} Calculated for Setup C, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	12.7		18.4		19.1		19.8		27.9	

[a] Activation barriers relative to the preceding minima are given in parentheses.

with its d^2 configuration at the molybdenum center. Referring to high level ab initio results,[63] 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^{-1}, and thus considerably higher than the experimental value of around 15 kcal mol^{-1}, see section 1. As DFT methods are expected to underestimate rather than to overestimate energy barriers,[117] 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^{-1}, 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^{-1} (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[46] 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 close in energy (e.g., within 1 kcal mol^{-1} for SN500-B3LYP). More importantly, each of these effective barriers is computed to be around 20 kcal mol^{-1}, 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^{-1} 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^{-1} 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^{-1} Calculated for Setup D, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	20.2		28.5		23.9		32.0	

[a] 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^{-1} 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 molybdopterin 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 molybdopterin 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^{-1} 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^{-1} (with BP86) and 17 kcal mol^{-1} (with B3LYP). For the other two snapshots, we obtain higher values of about 16 kcal mol^{-1} (for BP86) and about 20 kcal mol^{-1} (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^{-1} Calculated for Setup E, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	13.2		16.9		15.9		21.9		16.3		20.3	

[a] Activation barriers relative to the preceding minima are given in parentheses.

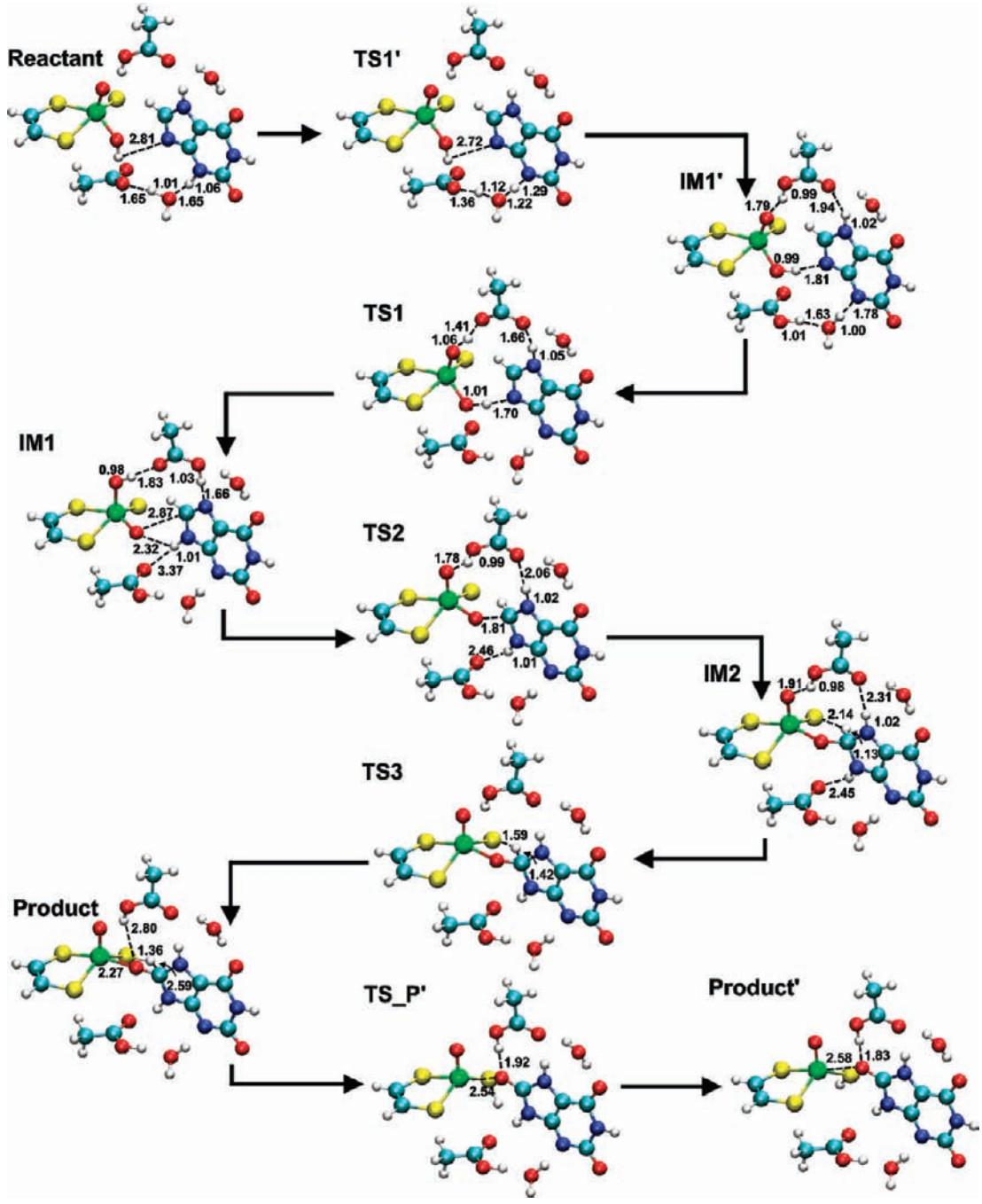


Figure 6. Reaction mechanism for setup F, SN400-B3LYP/MM. All bond lengths are given in Å.

Table 7. QM/MM Energies in kcal mol^{-1} Calculated for Setup F, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	23.6		20.5		17.6		20.3	

[a] 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.

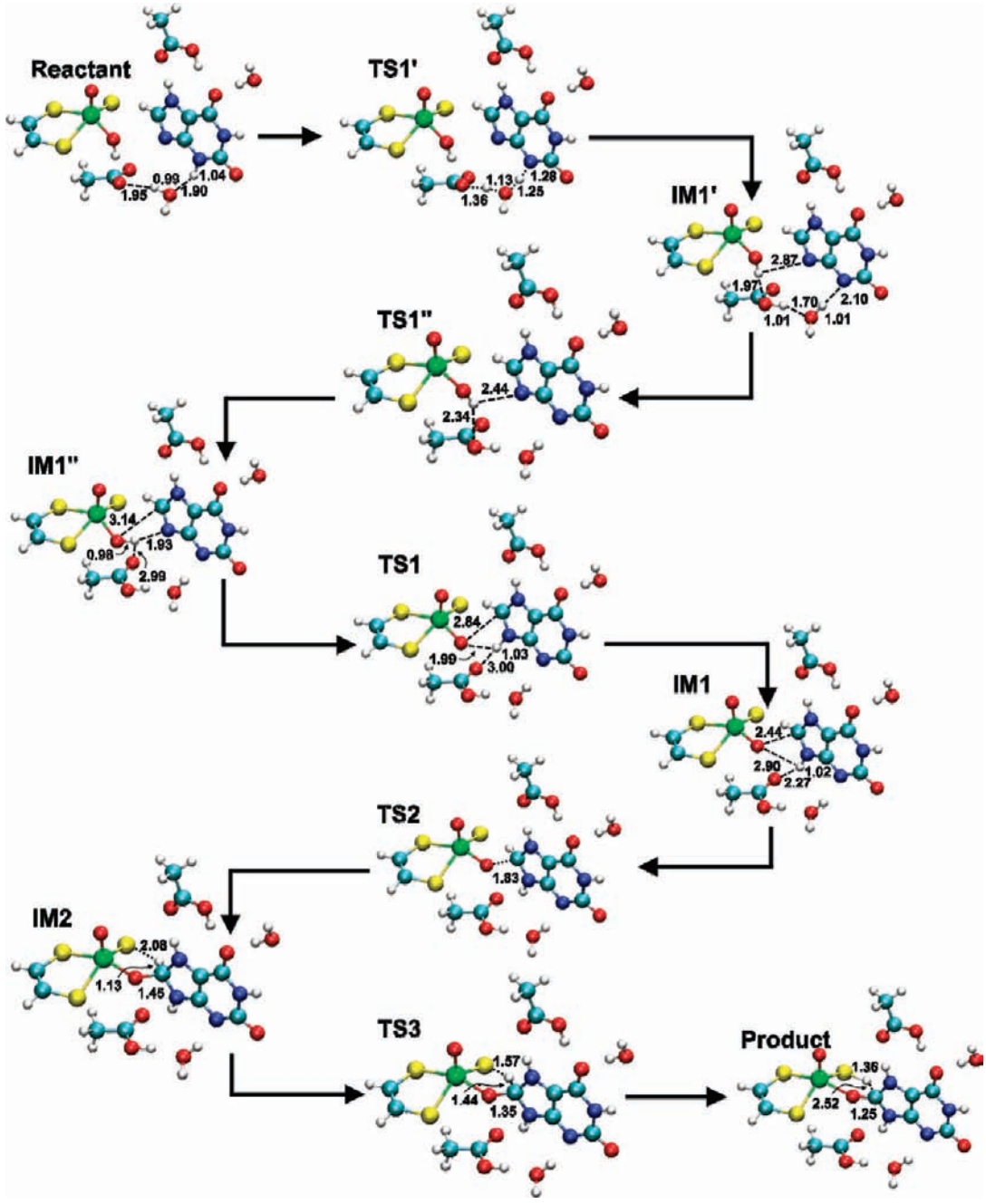


Figure 7. Reaction mechanism for setup G, SN400-B3LYP/MM. All bond lengths are given in Å.

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^{-1} (for BP86) and about 20 kcal mol^{-1} (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^{-1} Calculated for Setup G Using Basis Set B1, Relative to the Energy of the Reactant for Different Snapshots[a]

	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	
ΔE_{max}	13.6		14.6		14.7		18.3		10.0		13.1	

[a] Activation barriers relative to the preceding minima are given in parentheses.

Table 9. QM(B3LYP)/MM Energies in kcal mol^{-1} for Setup G Calculated for Different Basis Sets, Relative to the Energy of the Reactant[a]

	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	
ΔE_{max}	13.1		16.9		14.2		14.3		14.3		13.9	

[a] 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^{-1} (for BP86) and about 13−15 kcal mol^{-1} (for B3LYP). Barriers calculated with the alternative *L2* link atom scheme for SN400 differ by 0.4−0.5 kcal mol^{-1} (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^{-1}, except for the last two points (TS3, product) which are destabilized by 3−4 kcal mol^{-1}. 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^{-1}). 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^{-1}), again except for the last two points (TS3, product) where inclusion of relativistic effects raises the relative energies by 3−6 kcal mol^{-1} 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^{-1} 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 ΔE_{max} between these lowest and highest points lie between 13 and 17 kcal mol^{-1}, and the formally "best" values from the relativistic B3-based calculations are around 14 kcal mol^{-1} (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^{-1} with a maximum deviation of 2.6 kcal mol^{-1} (for the B1 basis). The calculated barriers deviate at most by 1.5 kcal mol^{-1} (for detailed results see SI).

Table 10. Results for the Extended 66-atom QM Region: B3LYP Relative Energies in kcal mol^{-1} 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			
	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 8$	$\varepsilon = 16$	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 8$	$\varepsilon = 16$	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 8$	$\varepsilon = 16$	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 8$	$\varepsilon = 16$
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[118,119] 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.[120−123] 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[120−123] 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: $\varepsilon = 4$ is a commonly accepted choice for proteins, and values up to 10 and above are often employed[124−126] 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.; Åquist, 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.

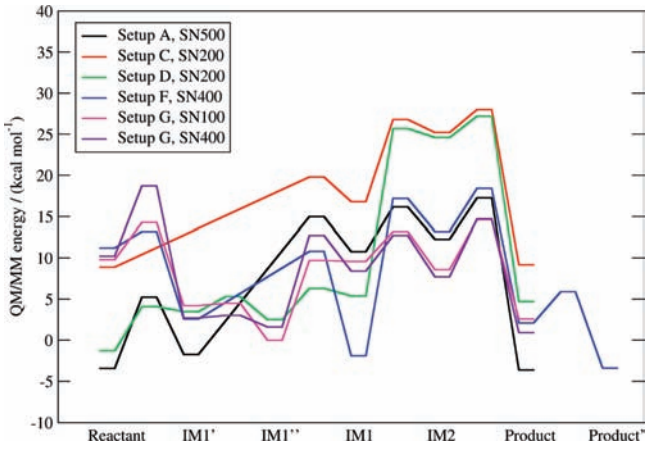


Figure 8. QM/MM reaction profiles for different setups on a common energy scale derived from COSMO ($\varepsilon = 8.0$) calculations on the extended QM region (see text and SI).

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

Table 10 presents the results of such single-point B3LYP/B1 calculations in the gas phase ($\varepsilon = 1$) and in the COSMO-MM environment ($\varepsilon = 4, 8, 16$). 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[120−123] by the data for IM2. This choice is further supported by the results of Torres et al.[128] 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 IM2 → 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.[129] 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.

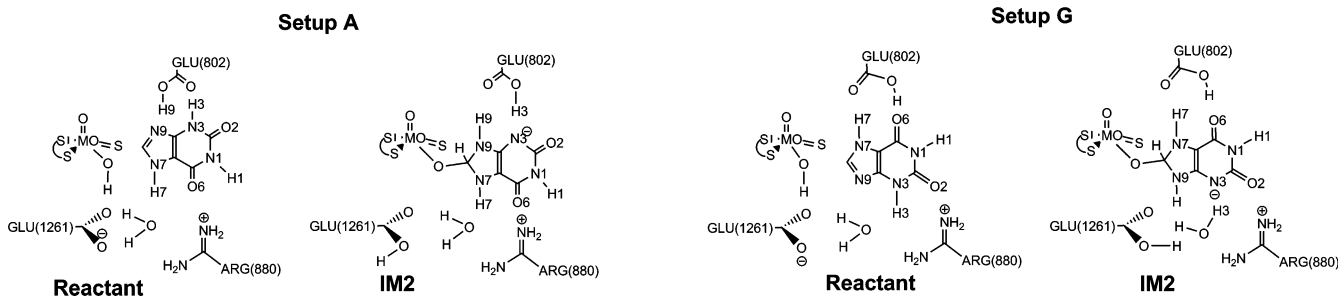
In an attempt to explain the different relative stability of the reactant and IM2 in setup A ("upside") and setup G ("upside down") we have evaluated classical electrostatic interactions between key partners (see Scheme 6) using the computed ESP charges and the relevant distances in the optimized structures. 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^{-1} (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^{-1} (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 (IM1″ → IM1) is followed by C−O bond formation (IM1 → IM2) and hydride transfer (IM2 → 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^{-1} 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.

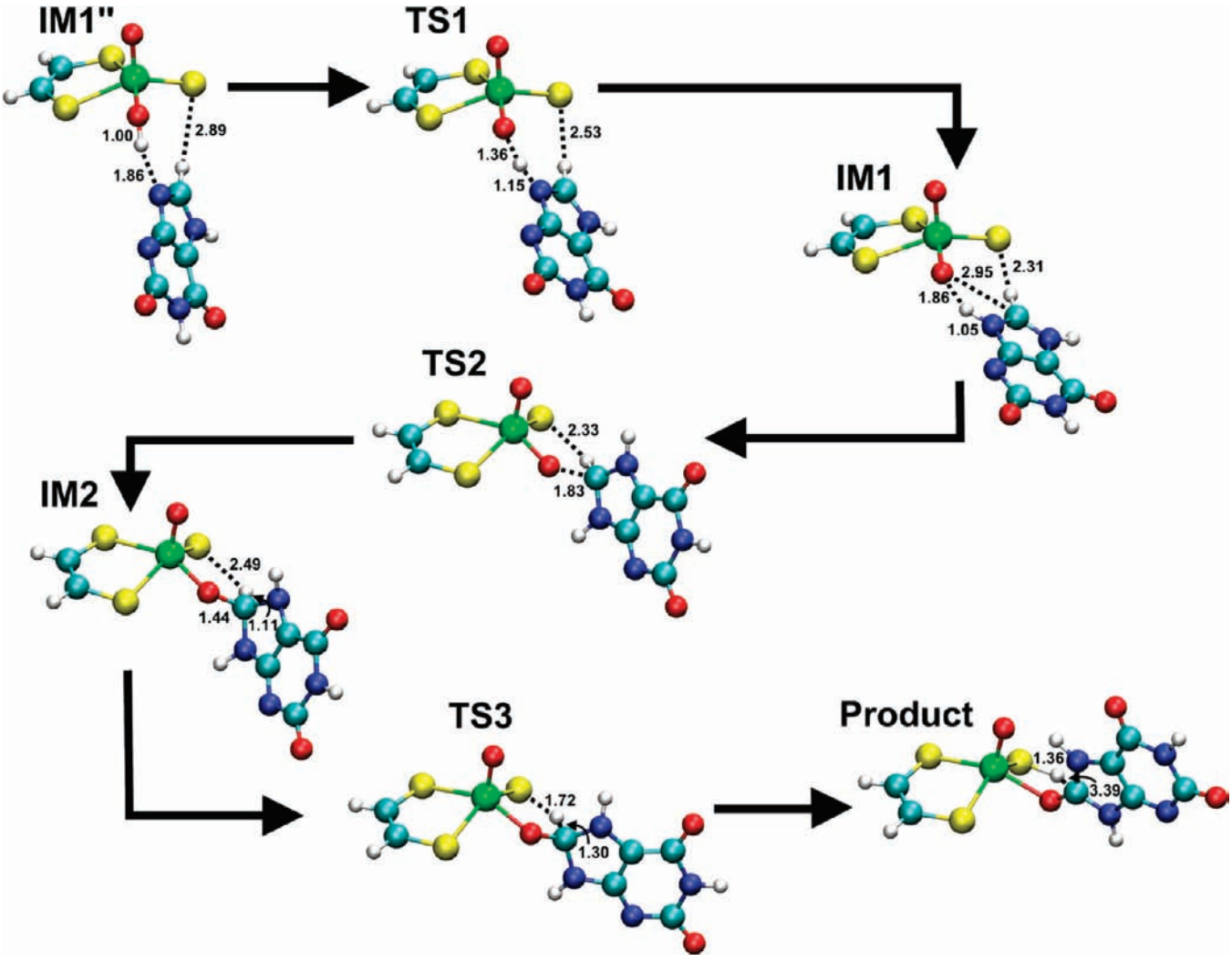


Figure 9. Detailed mechanism for the gas-phase reaction analogous to setup G, SN400-B3LYP. All bond lengths are given in Å.

Table 11. Gas-Phase QM Energies Compared with QM/MM Energies (QM = B3LYP) from Setup G SN400[a]

	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	
ΔE_{max}	16.8		21.0		13.1		16.9	

[a] All energies are given in kcal mol^{-1}, relative to the energy of IM1″. Activation barriers relative to the preceding minima are given in parentheses.

A recently published QM model study[65] 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 $HCOO^-$/HCOOH added to balance the stoichiometry and to define a common energy scale.[65] The calculated TS3 energies were 40.5 kcal mol^{-1} for xanthine (N1,N3,N7), 33.4 kcal mol^{-1} for xanthine (N1,N7,N9), and 6.4 kcal mol^{-1} 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_a value of about 0.8[79] 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.[65] 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[65] 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[65] (currently 43.4 kcal mol^{-1} (48.0 kcal mol^{-1}) using B1 (B2) compared to 33.4 kcal mol^{-1} reported by Bayse[65]). On the other hand, it is clear that the published QM model study[65] 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.[63]). 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,[79] the neutral tautomers (N1,N3,N7) and (N1,N3,N9) are almost isoenergetic in aqueous solution (within 0.6 kcal mol^{-1}), and the same applies to the anionic tautomers (N1,N9), (N1,N7), and (N1,N3) (within 1.2 kcal mol^{-1}). Moreover, with a pK_a value of around 7 (computed: 6.9, experimental: 7.44),[79] 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)]$^-$ anion in the initial reactant complex which is normally deprotonated to yield the active [(S−CH═CH−S)Mo(═S)(═O)(−O)]$^{2-}$ 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).

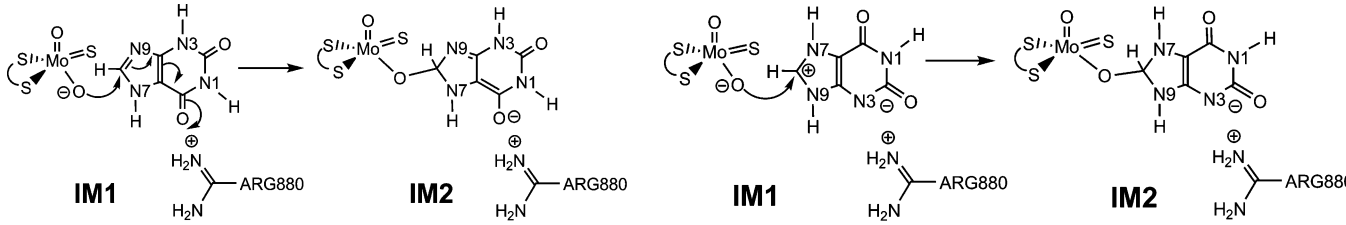
4.2. Role of the Active-Site Residues. Given the multitude of energetically competitive arrangements of different forms of the substrate and cofactor, it is clear that the protein environment will play a crucial role in determining the mechanistic preferences. Particularly important are the residues 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 ($k_{2'}$ vs $k_{2''}$), 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^{-1} for a Glu802 → Val802 mutant[39] and 1.4 kcal mol^{-1} for a Glu802 → Ala802 mutant.[40] 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^{-1} for a Arg880 → Met880 mutant[42] (and calculated to be 5.6 kcal mol^{-1} 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^{-1} upon substitution of Arg880 by Lys880[42] 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[32,42] (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.[40] 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_2 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.[57] 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^{-1} 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[55] 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[57] 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[57] the observed inhibitory effect of xanthine at high substrate concentrations.[130]

On the theoretical side, the "upside" orientation was originally proposed on the basis of docking studies in a structurally similar AOR enzyme.[30] 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.[32,42] 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;[33,34] both 2-oxopurine and 8-oxopurine are oxidized by xanthine oxidase to 2,8-dioxopurine whereas 6-oxopurin (= hypoxanthine) is not oxidized to 6,8-dioxopurine, but to 2,6-dioxopurin (= 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,[34] the latter one does.[34,38] 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.

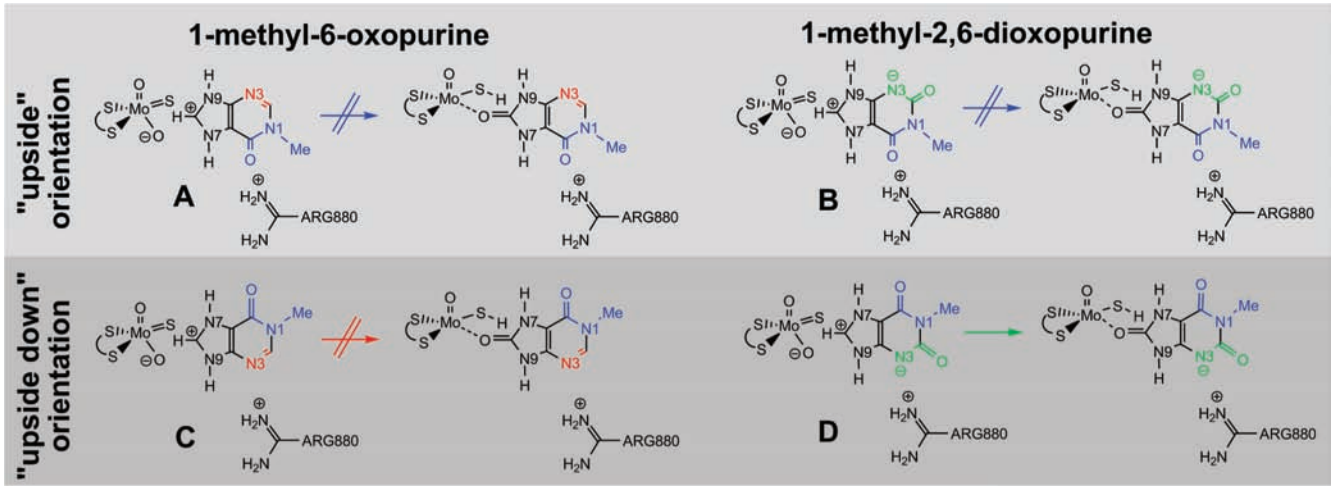
4.4. Product State. Finally we should point out that our QM/MM calculations model the reactive half-reaction of xanthine oxidase up to a product complex that contains deprotonated uric acid, mostly (setups A−B and E−G) in the form of the (N1,N7,N9) anion which is commonly considered to be the most favorable tautomer.[131−133] Given experimental pK_a values in the range 5.4−5.8 and a measured ionization enthalpy of −5.2 kcal mol^{-1},[134] 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^{-1} 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.[46] 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