University of Groningen

Integrated Computational Study of the Cu-Catalyzed Hydration of Alkenes in Water Solvent and into the Context of an Artificial Metallohydratase

Alonso-Cotchico, Lur; Sciortino, Giuseppe; Vidossich, Pietro; Rodriguez-Guerra Pedregal, Jaime; Drienovska, Ivana; Roelfes, Gerard; Lledos, Agusti; Marechal, Jean-Didier

Published in:
ACS Catalysis

DOI:
[10.1021/acscatal.8b04919](https://doi.org/10.1021/acscatal.8b04919)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Alonso-Cotchico, L., Sciortino, G., Vidossich, P., Rodriguez-Guerra Pedregal, J., Drienovska, I., Roelfes, G., Lledos, A., & Marechal, J-D. (2019). Integrated Computational Study of the Cu-Catalyzed Hydration of Alkenes in Water Solvent and into the Context of an Artificial Metallohydratase. *ACS Catalysis*, 9(5), 4616-4626. <https://doi.org/10.1021/acscatal.8b04919>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Integrated Computational Study of the Cu-Catalyzed Hydration of Alkenes in Water Solvent and into the Context of an Artificial Metallohydratase

Lur Alonso-Cotchico,^{†,‡} Giuseppe Sciortino,[†] Pietro Vidossich,^{†,§} Jaime Rodríguez-Guerra Pedregal,[†] Ivana Drienovská,[‡] Gerard Roelfes,[‡] Agusti Lledós,^{*,†} and Jean-Didier Maréchal^{*,†}

[†]Departament de Química, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès, Barcelona, Spain

[‡]Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, Netherlands

[§]COBO Computational Bio-Organic Chemistry Bogotá, Department of Chemistry, Universidad de los Andes, Carrera 1 N° 18A 10, Bogotá 111711, Colombia

Supporting Information

ABSTRACT: Despite the increasing efforts in the last few years, the identification of efficient catalysts able to perform the enantioselective addition of water to double bonds has not been achieved yet. Natural hydratases represent an interesting pool of biocatalysts to generate chiral alcohols, but modifying their substrate scope remains an issue. The use of artificial metalloenzymes (ArMs) appears as a promising solution in this field. In the last few years, Roelfes and co-workers have been designing a variety of DNA- and protein-based ArMs able to carry out the copper-mediated addition of water to conjugated alkenes with promising enantioselective levels. Still, from a mechanistic point of view, the copper-mediated hydration reaction remains unclear and a matter of debate.

This lack of information greatly hampers further designs and optimizations of the LmrR-based copper hydratases in terms of substrates and/or enantioselective profiles. In this study, we aim to provide a better understanding of the copper-catalyzed hydration of alkenes occurring both in water solvent and into the context of the LmrR protein as designed by Roelfes and co-workers. For that purpose, we make use of an integrated computational protocol that combines quantum mechanics (QM) (including small and large cluster models as well as ab initio molecular dynamics (AIMD)) and force-field approaches (including protein–ligand docking and classical molecular dynamics (MD) simulation). This integrative study sheds light on the general doubts around the copper-catalyzed hydration mechanism and also paves the way toward more conscious designs of ArMs able to efficiently catalyze the enantioselective addition of water to double bonds.

KEYWORDS: integrated molecular modeling, artificial metallohydratase, enantioselectivity, solvent versus protein environment, alkene hydration

INTRODUCTION

Enantiomerically pure alcohols are key intermediates in chemical industries with applications spreading from fine chemistry to perfume and cosmetics. One of the most interesting routes to synthesize those compounds is the enantioselective direct addition of a water molecule to alkenes, a mechanistic option that provides atom economy and is environmentally benign. However, achieving this transformation with high enantiomeric excesses (ee) represents a major challenge in synthetic organic chemistry.^{1,2} The reaction faces two major issues: on one hand, water is a poor nucleophile that needs to be activated, and, on the other hand, an asymmetric environment is required to enantioselectively add the water substrate, which is also the solvent.³ To reach this goal, biocatalysis is an interesting option. Enzymes can use water as

a substrate and also to provide asymmetric and protective environments. Several naturally occurring enzymes like fumarase and enoyl-CoA hydratase are able to carry out this reaction with excellent catalytic and enantioselective profiles.^{4–7} However, natural hydratases are extremely specific in terms of substrate recognition and biochemical modifications for expanding their scope, and therefore, their industrial uses have not been reached yet.^{1,8,9} Only a few examples of naturally occurring hydratases with a broad substrate scope have been described.^{10,11} Still, they proceed with low ee levels. The design of hydratases able to recognize a wider range of

Received: December 10, 2018

Revised: March 4, 2019

Published: April 9, 2019

chemically relevant reactants appears, therefore, as a particularly interesting goal. A promising strategy to address this challenge is the use of artificial metalloenzymes (ArMs), a rapidly growing family of non-natural biohybrids that combines homogeneous transition metal catalysts with biological receptors^{12–15} to drive the activation of the water nucleophile and induce asymmetry by the chiral second coordination sphere of the protein scaffold.

To date, the most successful ArM design for the hydration of alkenes diastereo- and enantiospecifically has been reported by Roelfes and co-workers.^{16–18} These biohybrids were constructed by embedding a Cu(II) catalyst, with phenanthroline (phen) or bipyridine (bipy) as a ligand, in different biomolecular scaffolds, either DNA or a protein (the transcription factor *Lactococcal multidrug resistance Regulator*, LmrR), using supramolecular¹⁶ or covalent¹⁷ anchoring or the biosynthetic incorporation of unnatural metal-binding amino acids.¹⁸ These artificial metallohydratases have reached enantiomeric excesses up to 84% for the conjugate addition of water to α,β -unsaturated ketones to generate chiral β -hydroxyketones. Interestingly, the reaction also takes place with the isolated copper catalyst in water solvent without the presence of the biomolecule, although with no enantioselectivity.¹⁷ The commonly accepted mechanism of copper-mediated conjugate additions implies the attack of the nucleophile (water) at the C_β position and of the electrophile (proton) at the C_α position. While the enantioselectivity (R/S) is related to the first step (only in the presence of the protein), the stereoselectivity of the reaction (syn/anti) is most likely defined in the second step. Additional mechanistic information has been reported in these experimental studies.^{16–18} Regarding the enantioselective character of the artificial metallohydratase, mutagenesis experiments of the LmrR-phen-Cu(II) artificial metalloenzyme suggested that the critical amino acid for the efficiency of the enzyme was the aspartic acid residue D100, without which both conversion and enantioselectivity significantly decreased.¹⁷ This finding is consistent with evidence of natural tungsten-dependent acetylene hydratases, which are able to use judiciously placed negatively charged amino acids in order to properly activate water molecules for hydration of CC triple bonds.^{19,20} However, it has not been possible to unambiguously establish the role of D100 carboxylate, in particular whether it directly interacts with the nucleophilic water molecule or as a ligand for the Cu(II) ion.¹⁷ In addition, the enantioselectivity of the reaction was found to be substrate-dependent; substrates containing bulky R substituents at C_β , such as R = *t*-butyl, gave rise to the highest enantioselectivities. Finally, from deuteration experiments it was inferred that the hydration occurs preferentially in a syn fashion. The syn addition of water does not result from the presence of the biomolecule in coherence with previous published data reporting syn stereoselectivity with copper(II) complexes.¹⁶

The aim of this work is to increase our understanding of the catalytic mechanism of the copper-mediated hydration of conjugated ketones by means of atomistic simulations, both with the isolated catalyst in aqueous medium and embedded in the LmrR protein. For the latter, this study has been focused on the rationalization of the ArM resulting from the inclusion of the phen-Cu(II) cofactor at position 89 of the LmrR protein (LmrR M89C mutant was produced by substitution of methionine 89 for cysteine), as this is the ArM that best performed the hydration of α,β -unsaturated ketones in terms

of both conversion and enantioselectivity.¹⁷ As illustrated in Figure 1, our strategy involves the simulation of the hydration

Figure 1. Copper-mediated hydration of conjugated ketones was studied with the catalyst embedded (1) in water and (2) into the chiral environment provided by the LmrR homodimeric protein.

reaction of 2-acylpyridine substrates (2a and 2b) mediated by the phen-Cu(II) catalyst, (1) isolated in water solvent and (2) embedded into the LmrR biomolecule. To model reactive events at the atomistic level in such different media, the implemented methodology follows an integrated strategy resulting from both quantum- and molecular mechanics-based approaches as described in the following section.

MODELS AND METHODS

Modeling the copper-mediated hydration reaction in the different environments studied (water solvent and LmrR) requires a complex computational framework that spreads over a large variety of methods. These included density functional theory (DFT) calculations, ab initio molecular dynamics (AIMD) simulations, protein–ligand docking, classical molecular dynamics (MD) simulations, and full-QM cluster models.²¹

Quantum calculations using the DFT formalism with the B3LYP-D3 functional^{22–24} and the Gaussian09 program²⁵ were performed in the first part of the study to elucidate the different steps of the hydration mechanism occurring in water, without the presence of the protein. For this purpose, a model comprising the phen-Cu(II) catalyst bound to the substrates either 2a (R = Me) or 2b (R = *t*-butyl) (Figure 2a) embedded in a solvent-polarizable dielectric continuum model (SMD, water, $\epsilon = 78.35$)²⁶ with six explicit water molecules was constructed (Figure 2b, top). Accordingly with the Cu(II) nature of the catalyst, calculations were performed in the doublet potential energy surface. The spin-density distribution, mainly in place in the copper ion, does not change appreciably along the reaction (Table S3, Supporting Information). To validate the proton-transfer path obtained from this small model, a more extended H-bonding network around the attacking nucleophile was considered and simulated with ab initio molecular dynamics (AIMD) calculations at DFT/BLYP-D3 level^{27–29} using the CP2K program.³⁰ Tests on the consistency between the DFT methods used in static QM and AIMD simulations are good as shown in Table S2 (Supporting Information). In these simulations the intermediate generated after the nucleophilic attack was embedded in a simulation box

Figure 2. Models studied. (a) **2a** (R = Me) or **2b** substrate (R = *t*-butyl) bound to the phen-Cu(II) cofactor. (b) Models for the study of the hydration mechanism in water: on top, the phen-Cu(II)-**2a/2b** complex embedded in a discrete-continuum solvent including six explicit water molecules; at the bottom, the phen-Cu(II)-**2b** complex into a cubic box of 228 water molecules. (c) Models for the study of the hydration mechanism into the protein: on top, the phen-Cu(II)-**2b** and phen-Cu(II)-(H₂O)₂ complexes linked to position 89 of each monomer of LmrR protein; at the bottom, a cluster model composed by the phen-Cu(II)-**2b** complex and surrounding residues at the active site.

of 228 water molecules (Figure 2b, bottom), allowing for exploration of the proton's path after dissociation from the nucleophilic water. A similar approach was used to study the acid-catalyzed hydration of ethylene in aqueous solution.³¹

The second part of the study focused on the catalytic mechanism occurring in the chiral environment provided by the LmrR protein. To embed the phen-Cu(II) catalyst into the LmrR dimer interface, protein–ligand dockings were performed throughout a covalent procedure as established in the GOLD 5.2 program, using the ChemScore scoring function.³² For this purpose, and due to the impossibility to fit two substrates at the active site, the substrate-linked catalyst phen-Cu(II)-**2b** (**1**) and its bisqua complex phen-Cu(II)-(H₂O)₂ (**2**) were docked at positions M89C and M89C' of the protein (which is how the experimental ArM was constructed), respectively. This initial model was refined by submitting it to 200 ns of MD simulation using the OpenMM 7.0 program.³³ The information obtained from the DFT calculations allowed us to recognize frames along the MD trajectory consistent with precatalytic configurations. These precatalytic states were extracted from the trajectory and used as starting points to study the copper-mediated hydration mechanism into the protein via full-QM cluster models, which included the phen-Cu(II)-substrate complex and the surrounding residues at the active site (up to 193 atoms). Consistently with the first part of the work, calculations were performed at the DFT/B3LYP-D3 level.^{22–24} For a more detailed description of the implemented methodologies, the reader is referred to the [Supporting Information](#).

RESULTS AND DISCUSSION

Copper-Mediated Hydration of Alkenes in Water Solvent. According to previous experimental works,¹⁷ the copper-mediated hydration of α,β -unsaturated 2-acylpyridines in water is a reversible reaction that proceeds in a syn fashion (both nucleophile and electrophile are added at the same face of the double bond of the substrate). To elucidate the main steps of the copper-mediated hydration, models of phen-Cu(II)-**2a** (R = Me) and phen-Cu(II)-**2b** (R = *t*-butyl) in a water environment were constructed. The homogeneous complexes were embedded in a discrete-continuum water solvent with six explicit water molecules. The number of explicit water molecules was assessed from a series of initial calculations on the nucleophilic attack to the phen-Cu(II)-**2a** with an increasing number of water molecules from 4 to 7 (Figure S5, [Supporting Information](#)). The model with 6 water molecules allows a proper solvation and stabilization of the nucleophilic water (*w_N*), the carbonyl group, and the metal center. Inclusion of an additional water molecule changes the reaction barrier in only 0.5 kcal mol^{−1}, in agreement with a converged model. Additionally, to assess the distribution of water molecules around the cofactor–substrate complex, the hydration of the C–C double bond was investigated by means of a classical MD simulation performed for complex phen-Cu(II)-**2b** in explicit solvent (3685 water molecules and 2 Cl[−] counterions) (see section 3 of the [Supporting Information](#)). The radial distribution function of water molecules around the double bond shows a minimum around 5.25 Å. Within this

Figure 3. Gibbs energy profile for the copper-catalyzed hydration reaction of **2a** and **2b** in water.

Figure 4. Transition state geometries for the copper-catalyzed hydration reaction in water. Selected distances (*d*) are indicated in Å.

distance, about 12 water molecules surround the double bond, six on each side (Figure S6, [Supporting Information](#)).

DFT calculations suggest a stepwise mechanism for the hydration reaction in which, first, the nucleophilic water (w_N) attacks the C_β of the substrate (Figures 3 and 4, TS_{N1}), leading to the formation of a protonated intermediate (**I1**). Next, a proton from **I1** is delivered to the solvent (Figures 3 and 4, TS_{N2}), becoming stabilized at the water chain as a hydronium (**I2**). Although TS_{N2} is found as a real transition state in the potential energy surface, it is displaced slightly below **I1** when thermal and entropic affects are added. This suggests that the delivery of the proton to the solvent after the generation of **I1** occurs spontaneously without a real barrier. Last, the generated hydronium in **I2** acts as an electrophile, leading to the proton

addition to the C_α of the substrate (Figures 3 and 4, TS_E), generating the final β -hydroxyketone product **P**.

Regarding the proton-transfer step from the protonated water to C_α our calculations with the small water cluster agree with an acid–base mechanism involving proton transfer to and from the solvent, instead of a concerted proton-shuttle relay mechanism trough a water chain. The proton-shuttle or acid–base nature of this kind of reaction is a matter of debate.^{34,35} The nature of the proton transfer is certainly dictated by the H-bonding network around the nucleophilic water and the C–C double bond.

To investigate this point, we extended the model solvating intermediate **I1** with 228 water molecules and performed a classical MD simulation to observe the conformational preferences of the H-bond network. Analysis of the MD

Figure 5. Transition-state structures related to the second step of the reaction, using a pyridinium moiety as electrophile to be attacked by the C_α in syn and anti for both *S* and *R* intermediates. Selected distances (*d*) are indicated in Å.

trajectories revealed that H-bonded water chains connecting the proton donor and acceptor atoms are very rare and short-lived. The model was further simulated via *ab initio* MD (AIMD) to improve on the description of molecular interactions and allow proton dissociation. Unfortunately, because of the size of the model (>700 atoms), we were not able to perform a detailed free-energy analysis of competing proton-transfer pathways. Instead, we prepared an ensemble of conformations of state **II** via a restrained AIMD, in which the hydrogens bound to w_N were not allowed to dissociate, and then allowed selected conformations to evolve freely. In three of these simulations, the proton was delivered to the solvent. Delivery to the accepting C_α was never observed. These results support the observations obtained from the discrete-continuum model, showing that the C_α protonation occurs through a two-step acid–base mechanism and not via a water-chain proton-shuttle transfer (see the [Supporting Information](#) for further details).

The effect of the steric bulk of the substrate *R* group was also assessed (Figure 3). The presence of a less-bulky *R* substituent (*R* = Me) eases the attack of the water nucleophile, decreasing the barrier of the first step (TS_{N1}) with respect to a bulkier *R* substituent (*R* = *t*-butyl). On the contrary, *R* substituent does not affect the last step (TS_E).

Deuteration experiments demonstrated that the reaction is stereospecific and occurs in a syn fashion. Moreover, the syn addition of water does not result from the presence of the biomolecule.¹⁶ To assess if there is an intrinsic preference for the syn addition in the electrophilic attack (TS_E), an alternative model was constructed to avoid the artifactual shift of the energy coming from the relocation of the water molecules to approach the different faces of the double bond. Protonation of both *R* and *S* enantiomers of intermediate **I2** was calculated using a pyridinium ion as proton source. To analyze the influence of the *R* substituent, the system with *R* = H was also computed. Consistent with the experimental data,¹⁶ the results favored the syn (Figure 5, TS_E syn) over the anti attack (Figure 5, TS_E anti). This observation was consistent for

both *S* and *R* enantiomers, for which differences of ~3 kcal mol^{−1} (for *R* = methyl) or 4 kcal mol^{−1} (for *R* = *t*-butyl) were found between the syn and anti attacks, respectively (Table 1).

Table 1. Relative ΔG Values of the TS_E Anti Attacks with Respect to the Corresponding TS_E Syn

<i>R</i>	enantiomer	TS_E syn	TS_E anti
<i>t</i> -butyl	<i>R</i>	0.0	4.7
	<i>S</i>	0.7	4.6
Me	<i>R</i>	0.0	3.2
	<i>S</i>	0.2	3.1
H	—	0.0	3.4

The same result was found with *R* = H, precluding substituent effects as the main origin of the stereoselectivity. Deeper structural analyses on these models showed clear differences between syn and anti transition-state structures. The most remarkable were (1) the closer distances of the transferred proton to the attacked carbon in the syn systems (by ca. 0.05 Å, see Figure 5) and (2) a rather short distance between the oxygen of the hydroxyl group of the substrate and the hydrogen of one of the β carbons of the pyridinium group in the syn systems (ca. 2.3 Å). Because this observation pointed to a possible weak interaction between these two atoms, noncovalent interaction analysis (NCI)³⁶ was undertaken. NCI plots clearly highlighted an attractive interaction of hydrogen-bonding nature between these two atoms in the syn geometries that was absent in the anti ones (Figure S7, [Supporting Information](#)). As a consequence, we concluded that the presence of the OH group of the substrate in the proximities of the incoming proton is key in guiding the face of the addition and provides with the syn versus anti preference of the hydration mechanism.

Summarizing, the combination of static QM calculations and AIMD simulations suggests that the copper-mediated hydration of alkenes in water proceeds through a stepwise mechanism. This includes the following: (1) a nucleophilic attack at C_β , (2) proton transfer from the added water to the

Figure 6. Gibbs energy profile of the nucleophilic attack for the hydration reaction in the phen-Cu(II)-2b system in water, in which the water nucleophile (w_N) is activated by an aspartate residue either (a) directly or (b) through a bridging water molecule (w_B). On the right the TS_N for both a and b configurations are represented.

Figure 7. Interactions of the phen-Cu(II)-2b complex at the dimer interface of the LmrR protein after (a) protein–ligand docking and (b) 200 ns MD simulation.

solvent, and (3) a C_α attack from a proton in the solvent, which proceeds in syn fashion. Bulky substituents that are directly connected to the double bond disfavor the nucleophilic attack, increasing the barrier of this step (Figure 3).

Before we embarked on characterizing the reaction into the protein, we wanted to assess the role of the aspartate D100 in the activation of the water nucleophile, as was suggested in the experimental study.¹⁷ Thus, a model including the same elements as before (the phen-Cu(II)-2b complex and six water molecules) plus an aspartate moiety was constructed (Figure 6). Two positions of the aspartate moiety with respect to the substrate double bond were assessed: one in which the aspartate was directly interacting with the nucleophilic water w_N (Figure 6a) and a second one including a bridging water molecule (w_B) between the aspartate and w_N (Figure 6b). From now on all models involve only the 2b substrate, as it is the one that provides the highest ee levels in the reaction catalyzed by the artificial metalloenzyme.¹⁷

The results point out that the aspartate boosts the nucleophilic attack (TS_N) by decreasing the barrier by 6 kcal mol⁻¹ with respect to the isolated phen-Cu(II) system (Figure 2) and, more importantly, that the optimum O–C _{β} distances

(d_1) between the carboxylate group of the aspartate and the double bond of the alkene to approach the transition-state geometries are ~ 3.6 – $5.0 \text{ \AA}$. This finding was used to define precatlytic configurations in the following procedure.

Copper-Mediated Hydration of Alkenes into the LmrR Protein. The artificial metalloenzyme model was constructed by linking the copper catalyst at positions M89C and M89C' of the LmrR dimer via covalent protein–ligand docking simulations. Because it was not possible to fit two substrates at the LmrR active site, the docked complexes consisted of the substrate-linked catalyst phen-Cu(II)-2b (1) and the aqua phen-Cu(II)-(H₂O)₂ (2). After the first docking run, which involved docking of 1, the results suggested very good interactions between the cofactor–substrate complex and the dimer interface (55.78 ChemScore units) (see Supporting Information, Table S4 and section 6). The complex appears sandwiched in the hydrophobic cavity between residues W96/W96' (Figure 7a). Additional hydrophobic interactions with residues I103', V15, and F93 also stabilize the complex inside the active site. As expected, residues D100/D100' are the only negatively charged amino acids that appear close to the substrate double bond. The second docking run involved the inclusion of complex 2 into the LmrR containing complex 1. In

Figure 8. (a) Distances in Å between the substrate double bond (C_{β}) and the oxygens of D100 (black)/D100' (red) residues along 200 ns MD simulation, which are also illustrated in (b).

this case, the fitting between the aqueous cofactor and the dimer interface is less favorable (37.19 Chemscore units, Table S4) because it appears slightly displaced toward the solvent due to the lack of space at the dimer interface after the inclusion of complex 1.

The best scored solutions of LmrR containing both complexes 1 and 2 were submitted to 200 ns of MD simulation. Consistent with the observations resulting from docking, the phen-Cu(II)-2b complex appeared well-stabilized at the active site. However, MD refinement suggested more intimate interactions between the cofactor–substrate complex and surrounding amino acids (Figure 7b). On one hand, F93/F93' appears to interact firmly with the aromatic rings of the phenanthroline ligand via π -stacking (see Table S6, Supporting Information). The important role of this residue was already discussed in the experimental work of Bos et al. that found that the point mutations performed on F93/F93' residues led to a drastic decrease in the conversion and enantioselective levels for the hydration reaction.¹⁷ Such observation surely results from the disruption of the stabilizing π -stacking interactions between F93/F93' and the phenanthroline aromatic rings that, if present, helps to maintain the phen-Cu(II)-2b complex inside the LmrR interdimeric region. By breaking this stabilizing interaction, higher propensity of the cofactor to lay outside the binding site may occur, hence having impact on both conversion and ee. On the other hand, polar interactions that were not found during docking were also identified: the phen-Cu(II) complex appeared stabilized by (1) a hydrogen bond with N19 and (2) the interaction between D100' and the metal center either directly or through a bridging water molecule located at the axial position of the copper. This suggests a double role for the D100' residue: on one hand, the stabilization of the cofactor–substrate complex at the active site and, on the other, the activation of the water nucleophile. These interactions, as well as the general structure of the protein, appeared quite stable during the 200 ns MD simulation.

Several convergence analyses were used to assess the stability/flexibility of the system as well as the convergence (if the conformational space of the system is properly visited) of the MD trajectory (see section 7 in the Supporting Information). Around 100 ns, the strong motion of the $\alpha 4'$ N-terminus promotes a cascade effect in the active site, resulting in the displacement of the substrate toward the solvent (see the structure at 100 ns in Figure S12). After this time, the loop comes back to the original position and the cofactor–substrate complex recovers the main interactions with the hydrophobic residues of the pocket (F93, V15, W96'/W96, and I103') as well as hydrogen-bonding interaction with N19 residue.

Altogether, these analyses agree with a converged trajectory after 200 ns.

Because of the dimeric nature of the system, the D100 and D100' residues appear positioned at the central region and at opposite sides of the dimer interface. It appears likely that they should be able to approach opposite faces of the substrate double bond, which could disfavor the enantioselectivity of the reaction. To assess the positioning of the D100/D100' residues with respect to the substrate double bond along the MD trajectory, the distances between the carboxylate groups and the C_{β} of the substrate were analyzed (Figure 8b and c). Interestingly, these were consistent with the catalytic distances found in the isolated models containing the aspartate moiety (from 3.61 to 5.03 Å). Although this analysis indicated close-to-catalysis distances between both D100/D100' and the substrate double bond, visual inspection of these frames suggested that actually the substrate double bond is accessible to D100' but not to D100. The latter residue appears blocked by the bulky substituent R at the C_{β} position of the substrate (Figure 7b). Indeed, experimental evidence highlighted the impact of bulky substituents on the ee.¹⁷ To further assess if these frames correspond with real precatalytic configurations, i.e., structures containing water molecules at proper distances and orientations with respect to the double bond and any negatively charged residue around, the water distribution around the substrate double bond was analyzed (Table 2).

Table 2. Number of Precatalytic Structures Found among 25 000 Frames along the 100 ns MD Simulation

Pro-S	Pro-R
–	D100'–2721
	D100–1
	E104–26

This allowed us to determine the following: (1) the number of times a precatalytic structure is found along the trajectory; (2) the negative residue conforming such configuration; and (3) the side of the double bond exposed to the catalytic water, which allowed us to perform an estimation about the enantioselective tendency of the system. Surprisingly, all the precatalytic structures found involved the pro-R face of the double bond (Table 2) and mainly the D100' residue.

These findings suggest that, when the cofactor–substrate complex is placed at the interdimeric region of LmrR, the R enantiomer is formed as preferred product. However, the fact that there is a second cofactor could be detrimental for the selectivity of the enzyme. The results suggest that the protein is not able to stabilize both cofactors at the inner part of the

dimer interface at the same time, which could mean that the second cofactor is reoriented toward the solvent lacking the chiral environment. It should be noted that this work is limited to describing the catalytic events occurring at the active site of the protein.

Among the precatalytic states identified, two different conformations were found: one in which D100' was directly coordinating the copper (**D100-Cu**) (Figure 9a) and a second one with an additional bridging water molecule located at the axial position of the copper ion (**D100-w-Cu**) (Figure 9b).

Figure 9. Precatalytic configurations found along 100 ns MD simulation, involving (a) direct (**D100-Cu**) or (b) indirect (**D100-w-Cu**) interaction between D100' and the metal center.

To discern the relevance of both types of configurations in the catalytic event, we have studied the two steps of the hydration reaction for both configurations using a quantum chemical cluster approach.³⁷ For this purpose, two precatalytic structures involving both types of interactions between D100' and copper were extracted to assess catalysis into the protein scaffold. The two systems (**D100-Cu** and **D100-w-Cu**) were

reduced to models composed of <200 atoms, which accounted for the phen-Cu(II)-**2b** complex, the residues at the active site, the nucleophilic water (w_N), and the axial water coordinating the copper, if any (Figure 10 and Supporting Information). DFT calculations were performed on these cluster models.

Between the two studied configurations, **D100-Cu** and **D100-w-Cu**, the latter appears clearly favored (Figure 10). In **D100-w-Cu** the aspartate residue efficiently activates the nucleophilic water w_N that attacks the double bond of the substrate at the same time that D100' extracts the remaining proton (TS_N), leading to the intermediate **I** (Figure 11, TS_N and TS_E). The energy barrier of this step ($7.1 \text{ kcal mol}^{-1}$) is consistent with the barrier found for the model including the isolated catalyst–substrate complex, six water molecules, and an aspartate moiety (Figure 6). In contrast, it is $\sim 8 \text{ kcal mol}^{-1}$ lower than the barrier found in the model lacking the aspartate moiety (Figure 3, TS_{N1}). In the second step, the protonated D100' gives back the proton to the substrate in a syn fashion (TS_E), leading to the final product (**P**). Regarding the **D100-Cu** model (direct coordination of the aspartate to the metal), it involves much higher barriers for both steps (Figure 10). This is due to the hindered rearrangement of the cofactor–substrate complex along the different steps of the reaction imposed by the coordination of the aspartate side chain to the copper(II). In the first transition state (Figure 11, TS_N), the restrictions force the ketone group of the substrate to switch from the equatorial to the axial position of the metal; regarding the second transition state (Figure 11, TS_E), the residue D100' is not able to make a hydrogen bond with the hydroxyl group of the substrate, as it is found in the **D100-w-Cu** cluster model.

Summarizing, these results suggest that the preferred configuration for the copper-mediated hydration of conjugated alkenes into the context of the LmrR protein involves a water molecule at the axial position of the copper ion. This

Figure 10. Catalytic pathways for the copper-mediated hydration of alkenes into LmrR assessed via full-QM cluster models for both types of configuration, **D100-Cu** and **D100-w-Cu**.

Figure 11. Transition-state geometries calculated via full-QM cluster models for the hydration reaction into LmrR. These correspond to the nucleophilic (TS_N) and electrophilic (TS_E) attacks for both types of configurations found, D100-Cu and D100-w-Cu.

configuration stabilizes the cofactor–substrate complex through the interaction with D100', which then appears properly placed next to the substrate double bond. This configuration of the second coordination sphere (1) promotes a decrease of ~ 8 kcal mol^{−1} of the barrier related to the nucleophilic attack, (2) leads to the generation of the *R* enantiomer, and (3) drives the electrophilic attack to evolve in syn.

CONCLUSIONS

The copper-mediated conjugated addition of water to ketones has been described based on an integrated computational strategy, which includes both QM-based strategies (AIMD and QM) and force-field-based approaches (protein–ligand docking and MD simulations). Their proper combination has allowed us to elucidate the nature of the copper-mediated hydration reaction both in water solvent and in the context of an artificial metallohydratase.

Quantum-based calculations show that the reaction in water courses through a stepwise mechanism which involves (1) a nucleophilic attack at C_β and (2) an electrophilic attack from a proton in the solvent to C_α. The proton transfer from the oxygen atom of the added water to the C_α proceeds to and from the solvent instead of a direct proton-transfer shuttle through a water chain. The position of the added −OH group guides the face of the addition, favoring syn addition in this step.

Regarding the reaction occurring into the artificial metalloenzyme, computational results are consistent with experimental data, pointing at D100' as the main residue driving the activation of the water nucleophile.¹⁷ Because of the arrangement of the second coordination sphere around the phen-Cu(II)-2b complex, D100' appears properly located to

approach the substrate double bond. Additionally, the majority of precatalytic configurations found involve the D100' residue and are only related to the pro-*R* face of the substrate. Furthermore, results suggest that the lack of precatalytic configurations involving the pro-*S* face may be substrate-dependent: the *R* substituent seems to play an important role by blocking the accessibility of D100 to the substrate double bond. This observation is consistent with experimental data, which shows higher conversion but decreased ee levels for substrates with less bulky substituents.¹⁷

The identification of precatalytic configurations along the MD simulations has allowed the construction of full-QM cluster models to elucidate the potential energy profile of the hydration mechanism into the LmrR protein. The results suggest a double role for the residue D100': (1) to stabilize the phen-Cu(II) by interacting with the metal center through a bridging water molecule located at the axial position of the metal and (2) to drive both the nucleophilic and electrophilic additions. The arrangement of the active site around the cofactor–substrate complex leads to a positioning of D100' with respect to the substrate double bond that, on one hand, promotes the generation of the *R* enantiomer and, on the other hand, makes the reaction evolve in a syn fashion.

The implemented integrative approach overcomes the real challenge of dealing with a small nucleophile such as a water molecule placed in a solvated environment. Our results shed light on the copper-catalyzed addition of water to conjugated alkenes as well as the effect coming from the second coordination sphere of the protein into the LmrR-phen-Cu(II) artificial metalloenzyme. We believe this work contributes to the understanding of the hydration reaction in both organometallic and biocatalytic reactions. Moreover, it expands the computational toolbox for optimizing the efficiency, regarding

both conversion and enantioselectivity, of artificial metal-lohydratases.

■ ASSOCIATED CONTENT

■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b04919.

Detailed description of the computational methodology; spin-density changes along the reaction pathway; selection of the number of water molecules in the smaller water cluster model; NCI analysis of the preference for the syn addition; best docking solutions for the inclusion of **1** and **2** into LmrR; assessment of substrate orientation into LmrR; LmrR conformational flexibility; and Cartesian coordinates and absolute E and G energies of the QM optimized structures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: agusti@klington.uab.es.

*E-mail: jeandidier.marechal@uab.cat.

ORCID

Giuseppe Sciortino: 0000-0001-9657-1788

Jaime Rodríguez-Guerra Pedregal: 0000-0001-8974-1566

Gerard Roelfes: 0000-0002-0364-9564

Agusti Lledós: 0000-0001-7909-422X

Jean-Didier Maréchal: 0000-0002-8344-9043

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Financial support from the Spanish MINECO (CTQ2017-87889-P) is gratefully acknowledged. L.A.-C. and J.R.-G.P. thank the Generalitat de Catalunya for their Ph.D. FI grant. G.R. acknowledges support from The Netherlands Organisation for Scientific Research (NWO, Vici Grant 724.013.003) and the Ministry of Education Culture and Science (Gravitation program no. 024.001.035).

■ REFERENCES

- (1) Liese, A.; Seelbach, K.; Wandrey, C. *Industrial Biotransformations*, 2nd Completely Revised and Extended ed.; Wiley-VCH: Weinheim, Germany, 2006.
- (2) Guo, J.; Teo, P. Anti-Markovnikov Oxidation and Hydration of Terminal Olefins. *Dalton Trans.* **2014**, 43, 6952–6964.
- (3) Resch, V.; Hanefeld, U. The Selective Addition of Water. *Catal. Sci. Technol.* **2015**, 5, 1385–1399.
- (4) Gawron, O.; Fondy, T. P. Stereochemistry of the Fumarase and Aspartase Catalyzed Reactions and of the Krebs Cycle from Fumaric Acid to D-Isocitric Acid. *J. Am. Chem. Soc.* **1959**, 81, 6333–6334.
- (5) Willadsen, P.; Eggerer, H. Substrate Stereochemistry of the Enoyl-CoA Hydratase Reaction. *Eur. J. Biochem.* **1975**, 54, 247–252.
- (6) Agnihotri, G.; Liu, H. Enoyl-CoA Hydratase: Reaction, Mechanism, and Inhibition. *Bioorg. Med. Chem.* **2003**, 11, 9–20.
- (7) Mohrig, J. R.; Moerke, K. A.; Cloutier, D. L.; Lane, B. D.; Person, E. C.; Onasch, T. B. Importance of Historical Contingency in the Stereochemistry of Hydratase-Dehydratase Enzymes. *Science* **1995**, 269, 527–529.
- (8) Jin, J.; Hanefeld, U. The Selective Addition of Water to C=C Bonds; Enzymes Are the Best Chemists. *Chem. Commun.* **2011**, 47, 2502–2510.

(9) Sheng, X.; Himo, F. Theoretical Study of Enzyme Promiscuity: Mechanisms of Hydration and Carboxylation Activities of Phenolic Acid Decarboxylase. *ACS Catal.* **2017**, 7, 1733–1741.

(10) Wuensch, C.; Gross, J.; Steinkellner, G.; Gruber, K.; Glueck, S. M.; Faber, K. Asymmetric Enzymatic Hydration of Hydroxystyrene Derivatives. *Angew. Chem., Int. Ed.* **2013**, 52, 2293–2297.

(11) Jin, J.; Oskam, P. C.; Karmee, S. K.; Straathof, A. J. J.; Hanefeld, U. MhyADH Catalysed Michael Addition of Water and in Situ Oxidation. *Chem. Commun.* **2010**, 46, 8588–8590.

(12) Diéguez, M.; Bäckvall, J.-E.; Pàmies, O. *Artificial Metalloenzymes and MetalloDNAzymes in Catalysis: From Design to Applications*; John Wiley & Sons: Weinheim, Germany, 2018.

(13) Rosati, F.; Roelfes, G. Artificial Metalloenzymes. *ChemCatChem* **2010**, 2, 916–927.

(14) Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R. Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **2018**, 118, 142–231.

(15) Jeschek, M.; Panke, S.; Ward, T. R. Artificial Metalloenzymes on the Verge of New-to-Nature Metabolism. *Trends Biotechnol.* **2018**, 36, 60–72.

(16) Boersma, A. J.; Coquière, D.; Geerdink, D.; Rosati, F.; Feringa, B. L.; Roelfes, G. Catalytic Enantioselective Syn Hydration of Enones in Water Using a DNA-Based Catalyst. *Nat. Chem.* **2010**, 2, 991–995.

(17) Bos, J.; García-Herrera, A.; Roelfes, G. An Enantioselective Artificial Metallo-Hydratase. *Chem. Sci.* **2013**, 4, 3578–3582.

(18) Drienovská, I.; Alonso-Cotichico, L.; Vidossich, P.; Lledós, A.; Maréchal, J.-D.; Roelfes, G. Design of an Enantioselective Artificial Metallo-Hydratase Enzyme Containing an Unnatural Metal-Binding Amino Acid. *Chem. Sci.* **2017**, 8, 7228–7235.

(19) Liao, R.-Z.; Yu, J.-G.; Himo, F. Mechanism of Tungsten-Dependent Acetylene Hydratase from Quantum Chemical Calculations. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 22523–22527.

(20) Liao, R. Z.; Himo, F. Theoretical Study of the Chemoselectivity of Tungsten-Dependent Acetylene Hydratase. *ACS Catal.* **2011**, 1, 937–944.

(21) Muñoz Robles, V.; Ortega-Carrasco, E.; Alonso-Cotichico, L.; Rodríguez-Guerra, J.; Lledós, A.; Maréchal, J.-D. Toward the Computational Design of Artificial Metalloenzymes: From Protein-Ligand Docking to Multiscale Approaches. *ACS Catal.* **2015**, 5, 2469–2480.

(22) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, 98, 5648–5652.

(23) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, 98, 11623–11627.

(24) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, 132, 154104.

(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision D. 01*; Gaussian, Inc.: Wallingford, CT, 2009.

(26) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a

Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(27) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.

(28) Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti Correlation Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(29) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157*, 200–206.

(30) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. Cp2k: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.

(31) Van Erp, T. S.; Meijer, E. J. Proton-Assisted Ethylene Hydration in Aqueous Solution. *Angew. Chem., Int. Ed.* **2004**, *43*, 1660–1662.

(32) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 609–623.

(33) Eastman, P.; Pande, V. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Comput. Sci. Eng.* **2010**, *12*, 34–39.

(34) Plata, R. E.; Singleton, D. A. A Case Study of the Mechanism of Alcohol-Mediated Morita Baylis-Hillman Reactions. the Importance of Experimental Observations. *J. Am. Chem. Soc.* **2015**, *137*, 3811–3826.

(35) Liu, Z.; Patel, C.; Harvey, J. N.; Sunoj, R. B. Mechanism and Reactivity in the Morita-Baylis-Hillman Reaction: The Challenge of Accurate Computations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30647–30657.

(36) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLLOT: A Program for Plotting Noncovalent Interaction Regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.

(37) Himo, F. Recent Trends in Quantum Chemical Modeling of Enzymatic Reactions. *J. Am. Chem. Soc.* **2017**, *139*, 6780–6786.