Designing artificial enzymes with unnatural amino acids

Drienovská, Ivana

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:
2017

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

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).

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.
CHAPTER 3

Rational design of an enantioselective artificial metallo-hydratase

Here, we present a novel enzyme able to catalyze the conjugate addition of water to alkenes whose design has been based on combining (bio)chemical knowledge and intuition on natural and artificial enzymes with state-of-the-art modelling protocols. After initial hypothesis and the insight gained through the modelling of the LmrR template, in silico analysis allowed to drive the optimization of the second coordination sphere of the metal. Experimental validation confirmed that this approach is capable of providing variants displaying improved efficiency and enantioselectivity.

This chapter will be submitted for publication:
3.1 INTRODUCTION

The design of tailored enzymes for new-to-nature reactions is a long-standing challenge that, when achieved, could bring great reward for organic synthesis. Yet, our limited understanding of the relationship between protein structure and catalytic activity and how to design stable protein folds, makes that de novo enzyme design is still far from routine.

Computational enzyme design methods offer great promise for the creation of novel enzymes. This is demonstrated by the recent progress in methods such as Rosetta has given rise to novel enzymes by matching of a “theozyme”, that is a tentative active site geometry designed for a calculated transition state, to existing protein structures in structural databases. The resulting designs often initially display modest catalytic efficiencies, albeit that these can be improved by posterior computational redesign and laboratory evolution. However, this may result in drastic changes of the structure of the active site when compared to the original computation design. In the case of metalloenzymes, the challenges are compounded by the difficulty of designing a stable metal binding site using canonical amino acids only. This is illustrated by the fact that while computational redesign of metalloenzymes has been achieved, the de novo design of new catalytically active metalloenzymes is scarce.

The amber stop-codon suppression methodology allows the introduction of unnatural bidentate metal binding amino acids into a protein at the genetic level. Recently, we have applied this expanded genetic code method for the preparation of a catalytically active artificial metalloenzyme, created by in vivo incorporation of (2,2'-bipyridin-5yl)alanine (BpyAla) at a predefined position into the transcriptional regulator LmrR. From a computational design perspective, this is an attractive approach, since it reduces some of the complications in metalloenzyme design by providing a stable metal binding site at the onset of the design process.

Here, we show computational design of an artificial metalloenzyme containing an unnatural BpyAla metal binding residue for the catalysis of a chemically challenging reaction: the enantioselective conjugate addition of water to enones. Asymmetric hydration is still an unsolved problem for conventional asymmetric catalysis, due to the challenges associated with using water as a nucleophile. Indeed, the available mechanistic information for natural hydratases, as well as the few reported examples of enantioselective hydration of enones by hybrid catalysts, suggest that these reactions critically depend on 2nd coordination sphere interactions to bind, activate and direct the water
nucleophile, which makes it an attractive target for computational metalloenzyme design.

Figure 1. Outline of the rational metalloenzymes design approach for a target reactivity used in this study.
3.2 RESULTS

3.2.1 Computational design

The approach followed is outlined in Figure 1. The design starts with (bio-)
chemical knowledge of the reaction of interest and the protein scaffold, which are
then used as input for the computational design. This then, in several steps, results
in a few designs that are experimentally validated.

Natural hydrateses have been studied extensively and key to their activity is
a dual activation process, which involves 1) an electrophilic activation of the
enone, e.g. by hydrogen bonding to the carbonyl and 2) the activation of the water
nucleophile by a judiciously placed general base.\textsuperscript{23,25} Quantum chemical
calculations have shown that a similar mechanism is operative in acetylene
hydratase, where the metal center activates the triple bond and an aspartate residue
assists the nucleophilic attack by a water molecule.\textsuperscript{28–30}

Our design is based on the same principle: a Cu\textsuperscript{II}-(2,2’-bipyridine) complex,
involving the BpyAla unnatural amino acid, that acts as Lewis acid to activate the
conjugated double bond and a judiciously positioned carboxylate moiety that acts
as a general base (Scheme 1). Moreover, the positioning of this general base with
respect to one prochiral face of the enone can be used to induce enantioselectivity
in the water addition step.

![Scheme 1. Dual activation strategy for catalysis of the enantioselective hydration reaction. The Cu\textsuperscript{II} ion is bound to 2,2’-bipyridine, the side chain of the unnatural amino acid at position 89 of LmrR, and to the α,β-unsaturated 2-acyl pyridine, the substrate.](image)

DFT calculations were performed on a reduced model of a Cu\textsuperscript{II}-(2,2’-
bipyridine) complex surrounded with water molecules and a carboxylate in the
second coordination sphere of the metal. Those calculations showed that proper
arrangement of the electrophilic and nucleophilic activation may result in a low
activation barrier for the addition of a water molecule to the conjugated double
bond (8.7 kcal/mol). Based on these calculations, the distance between the oxygen
of the side chain carboxylate and the β carbon of the substrate that is attacked by the nucleophilic water molecule needs to be in the range of ca. 3.5 to 5 Å. This corresponds to the distances found in x-ray structures of hydratases with substrate bound.\textsuperscript{31,32}

Using unnatural metal binding amino acids eliminates the need for screening of protein databases for suitable scaffolds. Therefore, we selected the Lactococcal multidrug resistance regulator (LmrR) for our design. LmrR is a homodimeric protein with a size of 13.5 kDa per monomer, that acts as a transcriptional repressor for the production of a multidrug ABC transporter. LmrR contains a unique flat shaped hydrophobic pore at the dimer interface, where normally hydrophobic antibiotic molecules are bound. This makes LmrR attractive for artificial enzyme design, as the hydrophobic nature of the interior of the pore already provides a generic driving force for binding of organic substrates.\textsuperscript{19} Since LmrR is a homodimer, genetic introduction of the unnatural amino acid will result in 2 of these metal binding moieties to be present in the hydrophobic pore. Position 89 was selected for introduction of BpyAla, since it is located at the far end of the hydrophobic pocket, thus avoiding the risk of formation of poorly active 2:1 ligand to metal complexes.

Protein-ligand docking experiments were performed to generate 3D models of the LmrR_LM_M89BpyAla (further referred to as LmrR_LM_M89X) with the α,β-unsaturated 2-acyl pyridine substrate $1\text{a}$ bound to the Cu$^{II}$ ion as a bidentate ligand in a square planar geometry (BpyAla-Cu$^{II}$-$1\text{a}$). The structure of BpyAla-Cu$^{II}$-$1\text{a}$ was optimized using the same DFT scheme as in our initial cluster models. The docking experiments lead to good predicted interaction energies of 42 ChemScore units. The two cofactor-$1\text{a}$ complexes, one per monomeric unit of LmrR, were found to fit well between the α4 and α4’ helices of the dimer (Figure 2) and, in general, a good hydrophobic complementarity was observed between the aromatic rings of the cofactor and the binding site pocket of the protein with a major edging or stacking interaction with F93. Additionally, the pyridine moiety of each substrate was observed to make hydrophobic interactions with the V15 side chain, while their isopropyl groups are sandwiched between the central tryptophans W96 and W96’. Moreover, the docking calculations also revealed that D100 and D100’ are the only potential general base residues in the vicinity of the substrates.
Figure 2. a) Docking of BpyAla-1a at position 89 of LmrR. b) Positions selected for the introduction of the glutamate residue that serves as catalytic residue. Models including mutations coloured in grey did not lead to reactive arrangements of the protein/cofactor/substrate assembly.

A Molecular Dynamics (MD) simulation was performed to explore the conformational properties of the design. 100 ns simulation were collected starting from the best docking solution. From analysis of the simulation, the following features emerged: i) the two cofactor-substrate complexes remain well stabilized at the dimeric interface of LmrR by hydrophobic interactions with neighbouring residues including F93 and W96. However, some flexibility was observed: one of the substrate bound cofactors was displaced toward the entrance of the LmrR binding pocket and became more exposed to the solvent. This flexibility and the change in the nature of the interaction of the substrate with the artificial metalloenzyme could affect negatively the catalytic profile of the enzyme but also suggest the possibility for further optimization ii) water accessibility is not equivalent on the prochiral faces of the substrate, suggesting that the protein may screen one side of the copper complex and thus induce enantioselectivity iii) D100, a candidate for the role of general base in the hydration process, was found not to approach the double bond of the substrate closely enough; the distance between the oxygen of the carboxylate moiety of the amino acid and the β carbon of the substrate is at the best of 8 Å. However, pre-reactive configurations, that is the configuration in which the cofactor-substrate complex, the general base and surrounding waters are arranged appropriately for the reaction to proceed, in which a water molecule is close to the β carbon and hydrogen bonding the carboxylate may still be achieved in a fraction of configurations (about 10%). The majority of these configurations involve approach from the pro-R face of the substrate. From
this, we predicted that LmrR_LM_M89X would show some catalytic activity, forming one enantiomer preferentially, but likely with a modest yield.

The docking results were analysed further to identify amino acids in the second coordination sphere of the metal that could be mutated to either aspartate or glutamate and fulfil our structural criteria for efficient hydration. Residues V15, W96 and F93 were found to have terminal side chain atoms at 5 Å from the electrophilic carbon of 1a. F93 was discarded as possible mutation site since it is involved in π-π stacking with the bipyridine complex. Taking into account the flexibility observed in the MD and the rearrangement of the receptor, we then extended our search to a 7 Å radius, which resulted in 5 more sites as possible candidates: M8, Q12, A92, S95, D100 (Figure 2). Models building the corresponding D or E variants and subsequent docking of the cofactor-substrate moiety showed however that positions 8, 12, 92 and 95 were not suitable, as the low energy docking solutions were characterized by a too long distance between the side chain oxygen of the D/E residues and the electrophilic carbon. Hence, three mutants were selected: M89X_D100E, M89X_W96E and M89X_V15E; and further investigated by MD simulations (Figure 3).

In case of LmrR_LM_M89X_D100E, the cofactors were not equivalent on the time scale of the simulation. BpyAla-CuII-1a explored three successive conformations, approximately equally populated, the last of which featured the cofactor outside the cavity formed by the monomers. Water accessibility at the prochiral faces was similar, yet pre-reactive conformations involving E100’ and E100 (first conformation) and E107 (second conformation) formed to the pro-R face of the substrate. For BpyAla-CuII-1a two conformations were observed and pre-reactive conformations (involving E100 and the pro-R face of the substrate) were observed only in a small set of conformations. These data suggest that LmrR_LM_M89X_D100E should display reactivity and enantiomeric preference similar to LmrR_LM_M89X.

MD simulation of LmrR_LM_M89X_V15E showed that the BpyAla-CuII-1a complex is more water exposed than BpyAla-CuII-1a’. BpyAla-CuII-1a mostly maintains a conformation in which the plane of the Cu complex is roughly perpendicular to the axis of the LmrR protein. Interestingly, the substrate displays pre-reactive configurations three times more often than in the simulation of LmrR_LM_M89X. Activation of the H2O nucleophile by E15 preferentially takes place at the pro-R face. Surprisingly, D100’ was found also to engage in pre-reactive conformations from the same prochiral face, even somewhat more frequent than E15. BpyAla-CuII-1a’ maintained a single conformation within the
Figure 3. Pre-reactive conformations from molecular dynamics simulations of LmrR\_LM\_M89X, LmrR\_LM\_M89X\_D100E, LmrR\_LM\_M89X\_V15E, LmrR\_LM\_M89X\_W96E. Pre-reactive conformations are defined on the basis of the closeness of a water molecule to the \(\beta\) carbon of the substrate and the hydrogen-bonding to an aspartate or glutamate residue. Similar pre-reactive conformations are reached in 10% of the simulation by LmrR\_M89X, 1% by LmrR\_M89X\_D100E, 9% by LmrR\_M89X\_V15E and 37% LmrR\_M89X\_W96E.

dimer interface, which is due to the interaction of Cu\textsuperscript{II} ion with E15’. This stabilizing interaction is maintained throughout the simulation and hinders the approach of water to the pro-R face of the substrate. The pro-S face is exposed to the solvent, but only rarely accesses configurations compatible with the pre-reactive geometry of reactants. These data suggest that M89X\_V15E should display higher activity than the template M89X and preference for the formation same hydrated enantiomer as M89X.

As for M89X\_W96E, BpyA\_Cu\textsuperscript{II}\_1a and BpyA\_Cu\textsuperscript{II}\_1a’ display a similar behaviour during the MD simulation. Both stay predominantly at the dimer
interface; only short fluctuations outside the pocket) with the pro-S face of the substrate more accessible to water molecules. For both cofactors, pre-reactive conformations are attained in high number and involve more the pro-S face than the pro-R. Additionally, E97 and E97', which are capable of interacting with both substrates, but also D100 and D100' may be involved in activating the water nucleophile. These data suggest that M89X_W96E should display higher activity and have a preference for the opposite enantioselectivity compared to LmrR_LM_M89X.

3.2.2 Preparation of mutants

The artificial metalloenzymes used in this study were prepared using the amber stop codon suppression methodology to introduce the unnatural metal-binding amino acid BpyAla into the biomolecular scaffold of LmrR at the position M89 (LmrR_M89X).17,19 In addition to the introduction of the amber codon for unnatural amino acid incorporation, LmrR_LM_M89X also contained two additional mutations K55D and K59Q (LM), which have been previously introduced to the sequence in order to prevent the DNA-binding and facilitate the purification and a C-terminal Strep-tag.19,27 The mutants selected from the in silico study were prepared using standard Quick-Change mutagenesis methods (Stratagene), expressed in E. coli BL21(DE3), purified on a Strep-tag affinity column and fully characterized (SDS-PAGE, exact mass measurement, size-exclusion chromatography). Typical purifications yields were 6-11 mg/L. All proteins had observed mass same as calculated confirming successful incorporation of BpyAla and mutagenesis (Experimental section, Table 4). Size-exclusion chromatography showed all LmrR mutants eluting as single peak at an elution volume of 11.6 (± 0.4) mL, which is consistent with a homodimeric structure of molecular weight around 30 kDa. This suggests that the dimeric structure is retained and none of the mutations caused significant perturbation of the structure in the conditions of the catalysis (20 mM MOPS, 250 mM NaCl, pH 7, 4 °C).

3.2.3 Catalytic hydratation

The catalytic activity of the BpyAla-containing artificial metalloenzymes was tested in the Cu"'-catalyzed enantioselective 1,4-addition of water to α,β-unsaturated 2-acyl pyridine 1a to yield in corresponding β-hydroxy ketone product 2a (Table 1). The reaction was carried out for 3 days at 4 °C, using 9 mol% of Cu(NO₃)₂ (90 µM), a small excess of LmrR_LM_M89X (1.25 equivalents, 112.5 µM in LmrR monomers). A conversion of 11 % was obtained in the uncatalyzed
reaction, while catalysis by 9 mol% of Cu(NO$_3$)$_2$, in absence of protein, gave rise to 83% conversion (Table 1, entry 1-2). The uncatalyzed reaction in the presence of protein and the absence of Cu$^{II}$ salts gave similar results as the uncatalyzed reaction: a low conversion of 9% was observed (Table 1, entry 3). Upon binding of Cu$^{II}$ to LmrR_LM_M89X, a moderate ee and conversion was obtained (Table 1, entry 4). All three designed mutants gave rise to an increased conversion of 1a. Using the M89X_D100E mutant, a modest increase of conversion to 52% was obtained, accompanied by slight decrease of ee with compared to LmrR_LM_M89X (Table 1, entry 5). In case of the mutant, M89X_W96E the highest increase in conversion was observed, albeit accompanied by a strong decrease of the ee (Table 1, entry 9). The most interesting case is the mutant M89X_V15E which gives rise to both increased conversion and enantioselectivity (Table 1, entry 7).

These results suggest that the placement of a general base at an appropriate position with respect to the active complex is an effective approach to improve catalysis. This was further confirmed by mutagenesis of the introduced glutamates to glutamines, which is sterically similar to glutamate but lacks the negative charge and, therefore, the ability to act as a general base. As expected, the activity of the corresponding glutamine mutants dropped back to levels similar to, or even lower than LmrR_LM_M89X, (Table 1, entries 6,8,10). Moreover, with all three glutamine mutants, the enantioselectivity of the reaction was also significantly lower, confirming that the glutamate residue, in particular in case of LmrR_LM_M89X_V15E, plays a role also in selectively positioning of the water nucleophile with respect to one prochiral face of the enone.

### 3.2.4 Substrate scope

The substrate tolerance of the designed metallo-hydratase was explored by varying the substituent at the β-position of the α,β-unsaturated 2-acyl pyridine (Table 1, substrate 1a-1d). Using LmrR_LM_M89X, good ee’s were obtained for products 2a and 2b (Table 1, entries 4 and 12), while with substrates 1c and 1d low conversions and ee’s were obtained (Table 1, entries 15,18). In the case of LmrR_LM_M89X_V15E, with substrates 1a, 1b good ee’s were obtained (Table 1, entries 6,13) Interestingly, with all 4 substrates conversion was higher than in the reactions catalyzed by LmrR_LM_M89X. In case of substrates 1c-1d, the conversions obtained with LmrR_LM_M89X_V15E were also higher than those obtained when using Cu(NO$_3$)$_2$ alone, further illustrating the importance of the protein scaffold in catalysis.
Table 1. Results of the water-addition reaction catalyzed by LmrR_LM_X_Cu\textsuperscript{II}\textsuperscript{a}.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Catalyst</th>
<th>Substrate</th>
<th>Product</th>
<th>Conv. (%)</th>
<th>ee (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>1a</td>
<td>2a</td>
<td>11 ± 3</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Cu(NO\textsubscript{3})\textsubscript{2}</td>
<td>1a</td>
<td>2a</td>
<td>83 ± 9</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>LmrR_LM_M89X (no Cu\textsuperscript{II})</td>
<td>1a</td>
<td>2a</td>
<td>9 ± 3</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>LmrR_LM_M89X_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>39 ± 7</td>
<td>42 ± 6</td>
</tr>
<tr>
<td>5</td>
<td>LmrR_LM_M89X_D100E_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>52 ± 3</td>
<td>30 ± 2</td>
</tr>
<tr>
<td>6</td>
<td>LmrR_LM_M89X_D100Q_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>35 ± 3</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>7</td>
<td>LmrR_LM_M89X_V15E_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>75 ± 9</td>
<td>64 ± 2</td>
</tr>
<tr>
<td>8</td>
<td>LmrR_LM_M89X_V15Q_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>28 ± 5</td>
<td>15 ± 3</td>
</tr>
<tr>
<td>9</td>
<td>LmrR_LM_M89X_W96E_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>79 ± 3</td>
<td>6 ± 4</td>
</tr>
<tr>
<td>10</td>
<td>LmrR_LM_M89X_W96Q_Cu\textsuperscript{II}</td>
<td>1a</td>
<td>2a</td>
<td>64 ± 1</td>
<td>6 ± 1</td>
</tr>
</tbody>
</table>

**Substrate scope\textsuperscript{b}**

<table>
<thead>
<tr>
<th>Entry</th>
<th>Catalyst</th>
<th>Substrate</th>
<th>Product</th>
<th>Conv. (%)</th>
<th>ee (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>Cu(NO\textsubscript{3})\textsubscript{2}</td>
<td>1b</td>
<td>2b</td>
<td>86 ± 4</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>LmrR_LM_M89X_Cu\textsuperscript{II}</td>
<td>1b</td>
<td>2b</td>
<td>37 ± 2</td>
<td>41 ± 4</td>
</tr>
<tr>
<td>13</td>
<td>LmrR_LM_M89X_V15E_Cu\textsuperscript{II}</td>
<td>1b</td>
<td>2b</td>
<td>64 ± 8</td>
<td>50 ± 2</td>
</tr>
<tr>
<td>14</td>
<td>Cu(NO\textsubscript{3})\textsubscript{2}</td>
<td>1c</td>
<td>2c</td>
<td>26 ± 4</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>LmrR_LM_M89X_Cu\textsuperscript{II}</td>
<td>1c</td>
<td>2c</td>
<td>24 ± 4</td>
<td>19 ± 5</td>
</tr>
<tr>
<td>16</td>
<td>LmrR_LM_M89X_V15E_Cu\textsuperscript{II}</td>
<td>1c</td>
<td>2c</td>
<td>58 ± 5</td>
<td>14 ± 1</td>
</tr>
<tr>
<td>17</td>
<td>Cu(NO\textsubscript{3})\textsubscript{2}</td>
<td>1d</td>
<td>2d</td>
<td>37 ± 12</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>LmrR_LM_M89X_Cu\textsuperscript{II}</td>
<td>1d</td>
<td>2d</td>
<td>17 ± 4</td>
<td>22 ± 2</td>
</tr>
<tr>
<td>19</td>
<td>LmrR_LM_M89X_V15E_Cu\textsuperscript{II}</td>
<td>1d</td>
<td>2d</td>
<td>45 ± 5</td>
<td>57 ± 3</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Standard conditions: 9 mol\% Cu(H\textsubscript{2}O)\textsubscript{6}(NO\textsubscript{3})\textsubscript{2} (90 µM) loading with 1.25 eq LmrR_LM_X in 20 mM MOPS buffer (pH 7.0), 250 mM NaCl, for 3 days at 4 °C. All data are the average of 2 independent experiments, each carried out in duplicate. Errors are reported as standard deviation. \textsuperscript{b} Conditions same as in the experiments with the mutants.
3.2.5 Kinetics

The kinetics were determined for the hydration of \( \text{1a} \) catalyzed by LmrR_LM_M89X_Cu\( ^{\text{II}} \) and LmrR_LM_M89X_V15E_Cu\( ^{\text{II}} \). Analysis was performed by reversed phase HPLC, monitoring the formation of product \( \text{2a} \) over the time as a function of the concentration of the substrate \( \text{1a} \) (0.25-3 mM). In both cases, saturation kinetics, typically associated with enzyme catalysis, were observed (Figure 4b). The catalytic parameters are summarized in the Table 2. Fitting the data to Michaelis-Menten kinetics showed a 3-fold increase of the catalytic efficiency of LmrR_LM_M89X_V15E compared to LmrR_LM_M89X. This is due to both a small increase in \( k_{\text{cat}} \) and a decrease of \( K_M \). The observed differences are too small to allow for a detailed discussion about the origin, but it does again confirm the viability of the design.

![Figure 4. a) The formation of the product 2a over time uncatalyzed and catalyzed by LmrR_LM_M89X and LmrR_LM_M89X_V15E. The reaction done at 1 mM concentration of the substrate 1a. Initial rate of uncatalyzed reaction is 0.039 µM.min\(^{-1}\), catalyzed by LmrR_LM_M89X is 0.173 µM.min\(^{-1}\) and LmrR_LM_M89X_V15E is 0.431 µM.min\(^{-1}\). b) Kinetics of the hydration of 1a catalyzed by LmrR_LM_M89X and LmrR_M89X_LM_V15E. The red line represents the fit obtained using Michaelis-Menten equation.](image)

<table>
<thead>
<tr>
<th>Catalyst</th>
<th>( K_M ) (mM)</th>
<th>( k_{\text{cat}} ) (min(^{-1}))</th>
<th>( k_{\text{cat}}/K_M ) (min(^{-1}).M(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>LmrR_LM_M89X_Cu( ^{\text{II}} )</td>
<td>2.55±0.67</td>
<td>0.0078±0.0012</td>
<td>3.06</td>
</tr>
<tr>
<td>LmrR_LM_M89X_V15E_Cu( ^{\text{II}} )</td>
<td>1.34±0.33</td>
<td>0.0104±0.0012</td>
<td>7.76</td>
</tr>
</tbody>
</table>
3.3 DISCUSSION

The observed overall agreement between the predictions from computation and the experimental catalysis results is striking. The MD simulation of LmrR_LM_M89X_Cu\textsuperscript{II}-1a revealed two phenomena: the possibility for the cofactors to move outside the catalytic pocket and the fact that the distance of the only general base in the vicinity of the electrophilic carbon of the enone, D100, is too large for efficient activation of the water nucleophile. Intuitively, mutagenesis of D100 to E would bring a general base closer to the substrate, resulting in higher activity. Yet, the MD simulation of M89X_D100E showed that this is not realized. Indeed, experiments confirmed that this mutant is only marginally better than LmrR_LM_M89X (Table 1, entries 4 and 5).

In contrast, variants LmrR_LM_M89X_V15E and LmrR_LM_M89X_W96E were predicted by computation to be more active than LmrR_LM_M89X because they do show pre-reactive conformations in which a general base is at an optimal distance from the electrophilic carbon from the substrate for significantly more times during the MD simulation. This was reflected in the higher conversions obtained with these mutants, as well as the increase in catalytic efficiency of LmrR_LM_M89X_V15E as observed in the Michaelis-Menten kinetics (Table 1, entries 7 and 9). Interestingly, in the latter mutant the MD simulations suggested a dual role for the E15/E15’ residues: they can act as a general base, in addition to D100, but also assist in stabilizing the cofactor in the interior of the pore by interaction with the Cu\textsuperscript{II} ion.

The MD simulations also allowed qualitative predictions about the enantioselectivity of the reactions by comparison of the prevalence of pro-S and pro-R pre-reactive conformations for the various mutants. In absence of experimental knowledge about the absolute configuration of the products, predicted stereochemical course of the reactions, giving rise to R or S products, cannot be verified to date. However, a relative comparison of the mutants can be made. LmrR_LM_M89X_V15E was predicted to have the same enantiomeric preference as LmrR_LM_M89X, that is, the same enantiomer of the product would be formed in excess, but with higher enantioselectivity. Indeed, the same enantiomeric preference was observed for LmrR_LM_M89X_V15E compared to LmrR_LM_M89X, but with an increase of $ee$ from 42 to 64%.

For LmrR_LM_M89X_W96E the enantiomeric preference was predicted to be inverse to the other mutants. Unfortunately, this mutant gave rise to nearly racemic product. This might be related to the fact that for this mutant, multiple glutamate and aspartate residues could contribute to the catalysis, which arguably
makes prediction of the enantiomeric outcome more difficult. However, it should be noted that near racemic product is obtained means that this variant does show a stronger preference for formation of the opposite enantiomer than the other mutants.

3.4 CONCLUSIONS

Here, we have presented a rationally designed metallo-hydratase enzyme, comprising an unnatural metal binding amino acid, for catalysis of a chemically challenging reaction: the enantioselective conjugate addition of water. The presence of a stable metal binding site in the protein, provided by the unnatural amino acid, BpyAla, combined with chemical knowledge of the mechanism of the reaction of interest provided an excellent starting point for the computational design of the artificial metalloenzymes. It is gratifying that the computational results provide us with suitable models of our system, allowing us to predict and better understand the positions of a general base in the designed active site. The results illustrate that the combination of QM, docking and MD simulations is a powerful approach to the design of metalloenzymes for novel and challenging reactions.

3.5 EXPERIMENTAL SECTION

3.5.1 General remarks

_E. coli_ strains NEB5-alpha and BL21(DE3) (_New England Biolabs_) were used for cloning and expression. DNA sequencing was carried out by _GATC-Biotech_ (Berlin, Germany). Primers were synthesized by _Eurofins MWG Operon_ (Ebersberg, Germany). Restriction endonucleases were purchased from _New England Biolabs_. Plasmid Purification Kits were purchased from _QIAGEN_. _Pfu_ Turbo polymerase was purchased from _Agilent_. Strep-tactin columns were purchased from _Iba-lifesciences_. Chemicals were purchased from Sigma Aldrich and used without further purification. Concentrations of DNA and protein solutions were estimated based on the absorption at 260 nm or 280 nm on Thermo Scientific Nanodrop 2000 UV-Vis spectrophotometer. The plasmid pEVOL-BpyAla was kindly provided by prof. P.G. Schultz.
3.5.2 Computational study

Docking

Protein-ligand docking calculations were performed to assess the complementarity between the cofactor-substrate complex and the protein frame. The crystal structure of the LmrR bound to the drug daunomycin at the dimer interface was used (PDB code: 3F8F)\(^{18}\).

Crystallographic water molecules and daunomycin were removed from the model. The organometallic complex bipyridine-Cu(II)-substrate was optimized with Gaussian09 program\(^{33}\) at density functional theory level using B3LYP functional\(^{34,35}\) and the 6-31g(d,p) basis set\(^{36,37}\). A bi-coordinated geometry of the substrate to the copper cofactor was considered, as this is the most suitable to fit the binding site. For the inclusion of complex at position 89 of LmrR, M89 of each monomer were mutated to alanine using the Dunbrack rotamer library\(^{38}\) as implemented in UCSF Chimera package\(^{39}\). The docking was performed imposing a covalent link between the beta carbon of alanine and the terminal carbon of the bipyridine ligand. Two successive docking runs were performed, the first at position 89 and the resulting structure used for docking at position M89'. Structures of the LmrR variants M89X_LM_D100E, M89X_LM_V15E and M89X_LM_W96E were generated, first introducing the second mutation using the Drunback rotamer library and then performing the docking of the bipyridine-Cu(II) substrate complex as for the wild type protein. All docking runs were performed with GOLD 5.2 (available through the Cambridge Crystallographic Data Center (CCDC)), and evaluated with ChemScore scoring function\(^{40}\).

Molecular Dynamics

The same crystal structure (PDB code: 3F8F) used for the docking was used to set up models for the all-atoms molecular dynamics (MD) simulations. Side chain conformation for residues 71 and 72 of chain A, not determined in the X-ray experiment, were fixed by superposition to chain B. Terminal residues 117-126 of chain A, 1-4 of chain B and 116-126 of chain B, not determined in the X-ray experiment, were discarded and uncharged terminal motifs were used to end the chain terminals. His86, solvent exposed, was considered protonated at ε. Model systems were set up with the xleap program.\(^{41}\) Each system was embedded into a cubic box including about 37000 water molecules and a number of chloride counterions (4 or 6) as required to neutralize the simulation cell. The AMBER\(^{42}\) and TIP3P\(^{43}\) force fields were used for protein and water, respectively. For chloride anions, parameters from ions94.lib library were used. Parameters for the bipyridine-Cu(II)-substrate complex were developed according to standard approaches. Point charges were calculated with antechamber\(^{41}\) according to the RESP procedure\(^{44}\). Bonded terms at the Cu center were calculated according to Seminario’s method based on second-derivatives\(^{45}\). The GAFF force field\(^{46}\) was adopted for the remaining atoms. A cutoff of 10 Å was used for short range electrostatics and Van der Waals interactions. Long range electrostatic interactions were calculated with the Ewald Particle Mesh method\(^{47}\). Bonds involving hydrogen atoms were constrained using the SHAKE algorithm\(^{48}\). A time step of 1 fs was used to integrate the equation of motion with a Langevin integrator\(^{49,50}\).
Constant temperature and pressure were achieved by coupling the systems to a Monte Carlo barostat\textsuperscript{51} at 1.01325 bar. Simulations were performed with OpenMM 7.0\textsuperscript{52}. Model systems were initially energy minimized (3000 steps) progressively, allowing water molecules, side-chain and backbone atoms to move; then, thermalization of water molecules and side chains was achieved by increasing the temperature from 100 K up to 300 K; finally, 100 ns MD simulations were performed and further analyzed.

### 3.5.3 Molecular Biology

**Site-directed mutagenesis**

Site-directed mutagenesis was used for preparation of all LmrR mutants. It was performed on the previously reported plasmid (according to the desired mutation), pET17b_LmrR_LM_M89X\textsuperscript{19} (LM–referred to lysine mutants exchanged to K55D, K59Q; M89X – TAG codon at positions 89). The primers required for the mutagenesis are summarized in Table 3. The following PCR cycles were used: initial denaturation at 95 °C for 1 min, denaturation at 98 °C for 30 s, annealing at 54-63 °C for 30 s (depending on the \(T_M\) of the particular mutant) and extension at 72 °C for 4 min 30 s. The thermal cycle was repeated 16 times and a final extension at 72 °C for 10 min was used. The resulting PCR product was digested with restriction endonuclease \textit{DpnI} for 2 h at 37 °C and transformed into chemically competent \textit{E. coli} NEB5α cells. A single colony was cultured in 5 mL of LB medium, the plasmid was isolated and mutagenesis was confirmed by sequencing.

<table>
<thead>
<tr>
<th>Primer\textsuperscript{*}</th>
<th>5’ (\rightarrow) 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>LmrR_LM_M89X_V15E_fw</td>
<td>GCT CAA ACC AAT GAA ATC CTG CTG AAT</td>
</tr>
<tr>
<td>LmrR_LM_M89X_V15E_rv</td>
<td>ATT CAG CAG GAT TTC ATT GGT TTG AGC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_V15Q_fw</td>
<td>GCT CAA ACC AAT CAG ATC CTG CTG AAT</td>
</tr>
<tr>
<td>LmrR_LM_M89X_V15Q_rv</td>
<td>ATT CAG CAG GAT CTG ATT GGT TTG AGC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96E_fw</td>
<td>GCG TTC GAA TCC GAA AGT CGT GTG GAC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96E_rv</td>
<td>GTC CAC ACG ACT TTC GGA TTC GAA CGC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96Q_fw</td>
<td>GCG TTC GAA TCC CAG CTG GTG GAC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96Q_rv</td>
<td>GTC CAC ACG ACT CTG GGA TTC GAA CGC</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100E_fw</td>
<td>TGG AGT CGT GTG GAA AAA ATC ATT GAA</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100E_rv</td>
<td>TTC AAT GAT TTT TTC CAC ACG ACT CCA</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100Q_fw</td>
<td>TGG AGT CGT GTG CAG AAA ATC ATT GAA</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100Q_rv</td>
<td>TTC AAT GAT TTT CTG CAC ACG ACT CCA</td>
</tr>
</tbody>
</table>

\textsuperscript{*}In bold is indicated which mutation was introduced by the particular primer set.
3.5.4 Expression and purification

The plasmids pEVOL-BpyAla and pET17b_LmrR_LM_X were cotransformed into *E. coli* BL21(DE3) and a single colony was used to inoculate an overnight culture of 10 mL of fresh LB medium containing 100 μg/mL of ampicillin and 34 μg/mL of chloramphenicol at 37 °C. 2 mL (500x dilutions) of overnight culture was used to inoculate at 37 °C 500 mL of fresh LB medium containing 100 μg/mL of ampicillin 34 μg/mL of chloramphenicol. When the culture reached an optical density at 600 nm of 0.8–0.9, the expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) (final concentration 1 mM) and L-Arabinose (final concentration 0.02%) and 200 mg/L of BpyAla (racemic mixture, synthesis previously reported) was added. Expression was done overnight at 30 °C. Cells were harvested by centrifugation (6000 rpm, JA10, 20 min, 4 °C, Beckman), resuspended in washing buffer (50 mM NaH$_2$PO$_4$, 150 mM NaCl, 50 mM EDTA, pH 8.0) and sonicated (70% (200W) for 7 min, 10 sec on, 15 sec off). After centrifugation (15000 rpm, JA-17, 1h, 4 °C, Beckman), the supernatant was loaded on a Strep-Tactin column (Strep-Tactin®Superflow® high capacity) and incubated for 1 h at 4°C. The column was washed with 3 x 1 CV washing buffer, and eluted with 6 x 0.5 CV of resuspension buffer (same as washing buffer plus 5 mM desthiobiotin). The fractions were analysed by SDS-PAGE electrophoresis on 12% polyacrylamide SDS-TrisTricine gel followed by Coomassie staining (InstantBlue™, Expediton). The concentration of the proteins was determined by using the calculated extinction coefficient for LmrR corrected for the absorbance of the BpyAla. Expression yields were 8-15 mg/L. For the purposes of characterization and catalysis, protein solutions were dialyzed against MOPS buffer (20 mM MOPS, 250 mM NaCl, pH 7.0) overnight at 4 °C with 2 exchanges of buffer. Expression in minimal media was also performed for the mutant LmrR_LM_M89X_V15E and LmrR_LM_M89X_W96E to minimize the iron binding. Protocol as mentioned above was followed, with only exception, that is expression at 30 °C for two days, instead of one.

3.5.5 Characterization

**Exact mass measurement**

The high-resolution mass spectrometry (HR-MS) was used to determine exact mass of the prepared proteins. It was performed using LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) with electrospray ionization as a source of ions. The mass of proteins was calculated with expasy peptide mass calculator (http://web.expasy.org/peptide_mass/).

**Analytical size-exclusion chromatography**

Analytical size exclusion chromatography was performed on a Superdex 75 10/300 GL (GE Healthcare). 100 μL of the sample was injected using 20 mM MOPS, 250 mM NaCl pH 7.0, as buffer (flow 0.5 mL/min). The column was calibrated using the standard Gel Filtration LMW Calibration Kit of GE Healthcare.
### Table 4. Mass of LmrR_LM_M89X mutants prepared in this study.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Mass\text{calculated}</th>
<th>Mass\text{found}</th>
</tr>
</thead>
<tbody>
<tr>
<td>LmrR_LM_M89X_V15E</td>
<td>15094.59</td>
<td>15094.62</td>
</tr>
<tr>
<td>LmrR_LM_M89X_V15Q</td>
<td>15093.61</td>
<td>15093.61</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96E</td>
<td>15007.51</td>
<td>15001.61</td>
</tr>
<tr>
<td>LmrR_LM_M89X_W96Q</td>
<td>15006.53</td>
<td>15006.60</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100E</td>
<td>15078.64</td>
<td>15078.63</td>
</tr>
<tr>
<td>LmrR_LM_M89X_D100Q</td>
<td>15077.66</td>
<td>15077.71</td>
</tr>
</tbody>
</table>

### 3.5.6 Catalysis

**Representative procedure for water-addition reaction catalyzed by LmrR_LM_X_Cu\text{II}**

The catalytic solution was prepared by combing Cu(H$_2$O)$_6$(NO$_3$)$_2$ (90 μM, 9 % catalyst loading) with 1.25 equivalents of LmrR_LM_X (112.5 μM) in a final volume of 290 μL MOPS buffer (20 mM MOPS, 250 mM NaCl, pH 7.0) and incubating together at 4 °C for one hour. To this 10 μL of a fresh stock solution of substrate 1 in CH$_3$CN/MOPS (50:50, 30 mM, final concentration in reaction mixture 1 mM). The reaction was mixed for 3 days by continuous inversion at 4 °C. The product was extracted with 3 x 1 mL of diethyl ether, the organic layers were dried on Na$_2$SO$_4$ and evaporated under reduced pressure. The product was redissolved in 150 μL of a heptane:propan-2-ol mixture (10:1) and the conversion and enantiomeric excess were determined using np-HPLC.

2a: column: Chiralpak-ADH n heptane:iPrOH 99:1, 0.5 mL/min, retention times: 76.8 and 88.2 min, 2b: column: Chiralpak-ADH n heptane:iPrOH 99:1, 0.5 mL/min, retention times: 58.5 and 67.9 min, 2c: column: Chiralpak-ASH n heptane:iPrOH 99.5:0.5, 0.5 mL/min, retention times: 23.5 and 26.0 min, 2d: column: Chiralpak-ASH n heptane:iPrOH 98:2, 0.5 mL/min, retention times: 17.6 and 18.5 min. Substrates and products of tested reactions have been synthetized according to previously published procedures.

### 3.5.7 Saturation Kinetics

The catalytic parameters were determined for LmrR_LM_M89X and LmrR_LM_M89X_V15E using reversed-phase HPLC (rp-HPLC), C18 column equipped with pre-column (Phenomenex, 4.6 mm internal diameter) with substrate 1a, using caffeine as a standard. (Acetonitrile/ water gradient, 60 min 0.5 mL/min). The catalytic solution was prepared as in standard catalysis, by combing Cu(H$_2$O)$_6$(NO$_3$)$_2$ (90 μM, 9% catalyst loading) with 1.25 equivalents of LmrR_LM_M89X/M89X_V15E (112.5 μM) in a final volume of 435 μL MOPS buffer (20 mM MOPS, 250 mM NaCl, pH 7.0) and incubating together at 4 °C for one hour. The reaction was started by adding 15 μL of a fresh stock solution of substrate 1a in CH$_3$CN (final concentration in reaction mixture varied from 0.25 to 3 mM), the sample was thoroughly mixed and immediately run on rp-HPLC. Over time
Rational design of an enantioselective artificial metallo-hydration enzyme (typically every 65 min) a sample was taken and injected on the column and analyzed. The same experiment was also performed without any addition of protein and Cu(H₂O)₆(NO₃)₂. For each substrate concentration, the concentrations of the product were plotted against the time and the initial rate of the reaction (v₀) was determined from the linear part of the curve. The kinetic parameters were obtained by fitting the data to Equation 1 using Origin software 8.5.

\[
\frac{v_0}{[E]} = \frac{k_{cat}[S]}{K_M + [S]}
\]

(Equation 1)

3.6 REFERENCES


Rational design of an enantioselective artificial metallo-hydration.


(39) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng,


