doi.org/10.26434/chemrxiv.11969010.v1

Dynamics of Metal Complex Binding in Relation to Catalytic Activity and Selectivity of an Artificial Metalloenzyme

Lara Villarino, Shreyans Chordia, Lur Alonso-Cotchico, Eswar R. Reddem, Andy-Mark Thunnissen, Jean-Didier Maréchal, Gerard Roelfes

Submitted date: 11/03/2020 • Posted date: 12/03/2020

Citation information: Villarino, Lara; Chordia, Shreyans; Alonso-Cotchico, Lur; Reddem, Eswar R.; Thunnissen, Andy-Mark; Maréchal, Jean-Didier; et al. (2020): Dynamics of Metal Complex Binding in Relation to Catalytic Activity and Selectivity of an Artificial Metalloenzyme. ChemRxiv. Preprint. https://doi.org/10.26434/chemrxiv.11969010.v1

We present an artificial metalloenzyme based on the transcriptional regulator LmrR that exhibits a unique form of structural dynamics involving the positioning of its abiological cofactor. The position of the cofactor was found to relate to the preferred catalytic activity, which is either the enantioselective Friedel-Crafts alkylation of indoles with beta-substituted indoles or the tandem Friedel-Crafts alkylation / enantioselective protonation of indoles with alpha-substituted enones. The artificial metalloenzyme could be specialized for one of these reactions by introducing a single mutation in the protein.

Dynamics of metal complex binding in relation to catalytic activity and selectivity of an artificial metalloenzyme

*Lara Villarino,[a] Shreyans Chordia,[a] Lur Alonso-Cotchico,[a] Eswar Reddem,[a] Andy-Mark W.H. Thunnissen,[b] Jean-Didier Maréchal[c] and Gerard Roelfes[a]**

(a) Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

(b) Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

(c) Departament de Química, Universitat Autònoma de Barcelona, Edifici C.n., 08193, Cerdanyola del Vallés, Barcelona, Spain.

email: j.g.roelfes@rug.nl

Abstract

Structural dynamics is important in enzymes to achieve optimal complementarity to the activated complex of the catalyzed reaction and, hence, give rise to rate acceleration and (enantio)selectivity. Here, we present an artificial metalloenzyme based on the transcriptional regulator LmrR that exhibits a unique form of structural dynamics involving the positioning of its abiological metal cofactor. The position of the cofactor, in turn, was found to be related to the preferred catalytic reactivity, which is either the enantioselective Friedel-Crafts alkylation of indoles with β-substituted enones or the tandem Friedel-Crafts alkylation / enantioselective protonation of indoles with α -substituted enones. The artificial metalloenzyme could be specialized for one of these catalytic reactions by introducing a single mutation in the protein. The switching of catalytic activity by dynamic interconversion of the position of a metal cofactor has not been described for natural enzymes and, to date, appears to be unique to supramolecularly assembled artificial metalloenzymes.

Main text

Enzymes are remarkable catalysts, capable of catalyzing chemical transformations with high rates and selectivities. Key to this is their ability to provide structural complementarity of the active site to the activated complex of the catalyzed reaction.[1] The dynamics of the protein involve conformational changes, e.g. domain rearrangement, loop motions, partial folding/unfolding etc., crucial to reach the optimal structure of the active site. Those flexible regions are also frequently involved in the emergence of alternative active site structures and relate to promiscuous catalytic activities. Thus they are important targets for mutagenesis in the natural- or directed evolution of enzymes to improve the activity.[2–6]

A popular approach to achieve enzymatic catalysis of reactions that have no equivalent in nature involves the creation of artificial metalloenzymes, which are rationally designed hybrids of proteins with abiological catalytically active metal cofactors.[7–13] In this approach, the basal catalytic activity is supplied by the metal complex, whereas the second coordination sphere interactions provided by the protein scaffold are envisioned to contribute to rate acceleration and (enantio-)selectivity. Since the protein scaffolds used have not naturally evolved for the reaction of interest, usually the active site structure is far from optimal. Recent examples underscore that structural dynamics in artificial metalloenzymes can be important in this context, similar to natural enzymes.[14,15]

Here, we report that dynamics in the binding position of an abiological metal cofactor in an artificial metalloenzyme leads to alternative active site structures. We also show how different reactions catalyzed by the artificial metalloenzyme, and its substrate selectivity, are dependent on the position of the cofactor in the protein scaffold. Finally, by a single mutation the artificial metalloenzyme can be specialized towards either of these catalytic reactivities.

The design of the artificial metalloenzyme is based on the transcription factor Lactococcal multidrug resistance Regulator (LmrR), which is a homodimeric protein with a size of 13.5 kDa per monomer that contains an unusual large hydrophobic pore at the dimer interface.[16] This hydrophobic pore serves as a promiscuous binding pocket where planar aromatic molecules bind, as shown in X-ray and NMR structures of LmrR with various planar drugs bound.[16–18] Two tryptophan residues, one from each subunit, i.e., W96 and W96', play a key role in binding by sandwiching the guest molecule via π-stacking interactions. Previously, we have shown that this arrangement is attractive for the supramolecular self-assembly of a novel artificial metalloenzyme, by combining the protein LmrR with a Cu(II) complex with a planar aromatic ligand, like 1,10-phenanthroline (phen) (Fig. 1a).[19,20] LmrR showed a moderately strong affinity for Cu(II)-phen, with a dissociation constant (K_d) of 2.6 ± 2 μM. The importance of the central tryptophans for binding Cu(II)-phen is illustrated by the fact that in case of the mutant LmrR_W96A, the K_d was one order of magnitude lower, i.e., 45 μM.

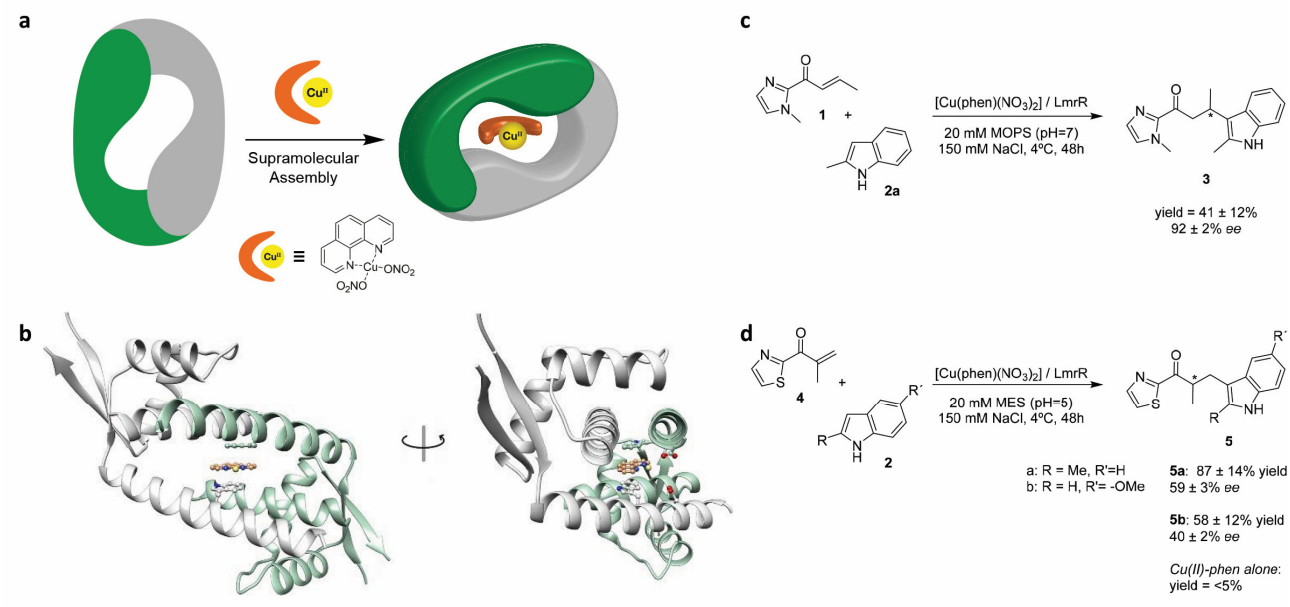


Figure 1. a) Schematic representation of the self-assembly of the artificial metalloenzyme. b) Crystal structure of the LmrR/Cu(II)-phen artificial metalloenzyme (PDB: 6R1L). Considerable disorder is observed in the binding mode of Cu-phen, as evidenced by its relatively weak associated electron density and high atomic B-factors. Disordered ligand binding is a general observation in crystal structures of LmrR and may be an inherent property of this protein. Unfortunately, the weak electron density around the copper, and its special position in the crystal on a crystallographic dyad, prohibited an unambiguous identification of its coordination geometry and ligands other than phenantroline.. c) catalyzed FC reaction (d) catalyzed FC/EP reaction. C and D: Typical reaction conditions: enone **1** or **4** (1mM), indole **2** (1mM) [Cu(II)-phen] (9 mol%; 90 µM) , LmrR (12 mol%; 120 µM) in 20 mM MOPS buffer pH 7.0 (FC reaction) or 20 mM MES buffer pH 5.0 (FC/EP reaction), 150 mM NaCl, at 4°C; Results are the average of at least two independent experiments, both carried out in duplicate. Error margins are listed as standard deviations.

The binding of Cu(II)-phen to LmrR was confirmed by X-ray crystallography, which showed the phenanthroline ligand of the complex sandwiched between W96 and W96', with the indole rings somewhat tilted with respect to each other and the Cu(II) complex (Fig. 1b). The Cu(II) ion is facing the front entrance of the pore. Two carboxylate side chains, from D100 and D100' are oriented towards the Cu(II) ion at a distance of ~5 Å and may interact with bound ligands, e.g. water, at the remaining coordination sites at the copper, albeit that these could not be identified with certainty. Protein residues other than W96 and D100 that surround the copper within a distance of 8 Å are predominantly hydrophobic, i.e., V15, A92, S97, V99 and I103 (and their equivalents from the dimer mate)

In this study, we focused on the application of the artificial enzymes in two catalytic reactions. First, the previously reported enantioselective vinylogous Friedel-Crafts alkylation of indoles with α, β unsaturated 2-acyl-imidazoles (FC reaction), which gives rise to excellent enantioselectivities, to up to 92% when catalyzed by LmrR/Cu(II)-phen (Fig. 1c).[19] The second reaction is the tandem Friedel-Crafts alkylation/enantioselective protonation reaction (FC/EP reaction, Fig. 1d). It involves the conjugate addition of indoles to α-substituted enones. In this case, the chirality is introduced not in the conjugate addition step, but in the protonation step and, hence, this reaction represents an enantioselective protonation in water, which is a highly challenging

reaction.[21] These two reactions share similiraties, but the chiral center is created in different elementary reaction steps.[22]

The conjugate addition of 5-methoxy-1H-indole (**2b**) to 2-methyl-1-(thiazol-2-yl)prop-2-en-1-one (**4**) was used as benchmark FC/EP reaction (Fig. 1d). The artificial metalloenzyme was prepared *in situ* by self-assembly from 9 mol% of [Cu(phen)$(NO_3)_2$] with a slight excess (1.3 equiv) of LmrR in MES buffer at pH 5.0. (Table S3). Under these conditions, the product was obtained in 58 % yield and 40 % ee. Interestingly, in absence of LmrR, so when using Cu(II)-phen alone, very low conversion was observed (Table S4, entry 2). This shows that the reaction is highly protein accelerated, that is, it requires the presence of LmrR to occur. Evaluation of the indole scope showed that the best results were obtained using 2-methyl-1H-indole (**2a**), with 87% yield and 59% *ee* (Table S3, entry 4).

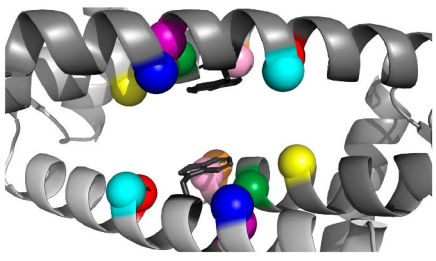
Mutagenesis of residues at various positions in the hydrophobic pocket in spatial proximity to W96 was performed to establish where catalysis of both these reactions occur and which residues are important for activity. This included positions at the front entrance, i.e. D100, F93 and A92, residues in the pocket interior, i.e. M8, Q12 and V99 and residues V15, E7, which are placed in the back entrance of the pocket. Most of these residues were probed by converting them to alanine (alanine scanning), except at position A92 were alanine was already present. In this case mutation to glutamate, A92E was performed.

All mutants were evaluated in both the FC reaction of 2-methyl-indole (**2a**) with *(E)*-1-(1-methyl-*1H*-imidazol-2-yl)but-2-en-1-one (**1**) and the tandem FC/EP reaction of **2a** with 2-methyl-1-(thiazol-2-yl)prop-2-en-1-one (4).

In the FC reaction significant effects on catalysis were observed in case of the front entrance mutants and the mutant M8A, which is located inside the pore (Fig. 2, Table S5). These results confirm that the reaction occurs in the pore, at the front entrance, close to the tryptophan residues where the Cu(II)-phen complex is bound. In most cases, the effect of the mutation was negative on both activity and enantioselectivity. Two mutations gave rise to significantly improved enantioselectivity: the mutation M8A resulted in a strong increase in the *ee* to 99% (Table S5, entry 11) and A92E gave rise to both complete enantioselectivity (>99 % *ee*) and a significantly increased product yield (Table S5, entry 6).

In contrast, almost all of these mutations had only a small effect on the results of the FC/EP reaction (Fig. 2b, Table S6). The only mutation that gave rise to a significant, negative, effect proved to be A92E; a decrease of the *ee* to 21 % was obtained (Table S6, entry 7). This was a surprising result since this same mutation proved to be the most beneficial for the FC reaction (vide supra).

a



b

Mutant:	ee (%) FC	ee (%) FC/EP
Wild type	92 ± 2	58 ± 2
D100A	46 ± 4	63 ± 1
F93A	51 ± 5	66 ± 2
A92E	> 99	21 ± 7
W96A	25 ± 9	64 ± 6
Q12A	98 ± 1	44 ± 6
V99A	95 ± 1	60 ± 4
M8A	99 ± 0	45 ± 1
V15A	94 ± 1	42 ± 1
E7A	96 ± 1	56 ± 1

c

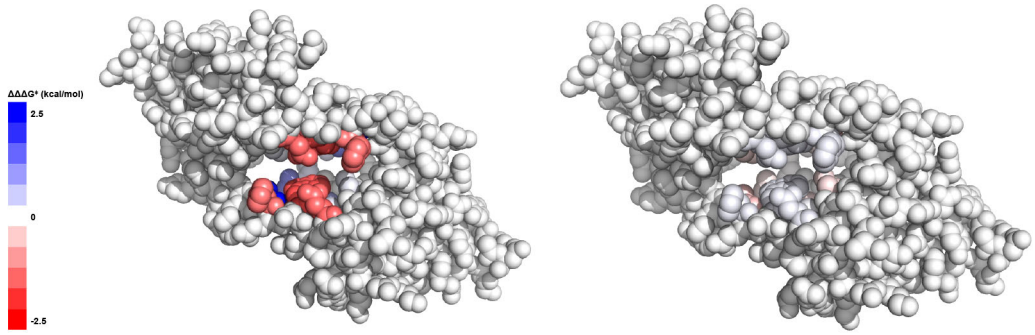


Figure 2. a) Close up of the hydrophobic pore of the LmrR. Residues used in the mutagenesis study are highlighted as spheres (color code as indicated). b) Effect of mutations in LmrR on the *ee* of the catalyzed FC reaction between **1** and **2a** and FC/EP reaction between **4** and **2a**. Colors represent the difference between the $\Delta\Delta G^{\ddagger}$ values calculated from the corresponding ee's as defined in (c); c) Visualization of the mutated residues, and the effects on enantioselectivity, onto the crystal structure of LmrR/Cu(II)-phen (metal complex omitted for clarity). The effect of the mutation on the *ee,* compared to the wild-type LmrR, is visualized as a heatmap where the colors represent the difference between the $\Delta\Delta G^{\ddagger}$ values calculated from the corresponding ee's ($\Delta\Delta\Delta G^{\ddagger} = \Delta\Delta G^{\ddagger}_{mutant} - \Delta\Delta G^{\ddagger}_{wild\ type}$), indicating an increase (blue) or decrease (red) of enantioselectivity of the reaction catalyzed by the mutant compared to the wild type LmrR.

The role of the hydrophobic pocket in catalysis was probed by inhibition studies using Hoechst H33342, which has been shown to bind with nanomolar affinity in the pocket, sandwiched between W96 and W96', analogous to binding of Cu(II)-phen as observed by X-ray crystallography.[16] Hence, H33342 can act as an inhibitor, blocking the binding of Cu(II)-phen, which in view of the fact that both reactions are protein accelerated, should have an effect on the results of catalysis. Indeed, addition of increasing amounts of H33342 up to 4 equivalents with respect to Cu(II)-phen caused a significant decrease in the enantioselectivity in the FC reaction, to 57% *ee*. In contrast, no significant effect was observed in the case of the FC/EP reaction (Table S7).

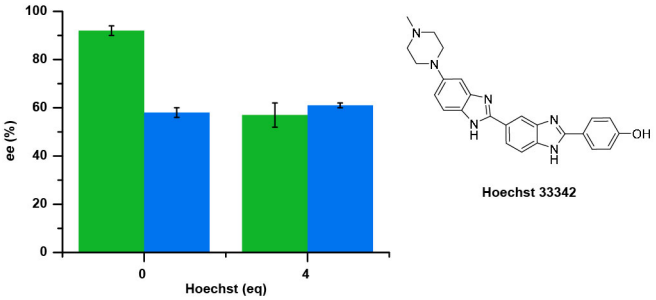


Figure 3. *ee* obtained in catalysis of the FC (green) and FC/EP (blue) reaction in presence of 0 and 4 equivalents of Hoechst 33342, with standard deviations shown.

Next, competition experiments were performed with the substrates **1** and **4** , using wild type LmrR and the mutants LmrR_W96A and LmrR_A92E (Figure 4, Figure S7). When 2-methylindole (**2a**) was combined with equimolar amounts of **1** and **4**, the corresponding products **3** and **5** were obtained in 66% and 17% yield, respectively, in the reaction catalyzed by LmrR/Cu(II)-phen, which corresponds to a selectivity of 80 % for the FC reaction. The enantioselectivity of the products was similar to that obtained in the independent experiments. Notably, when we carried out these competition experiments with the LmrR variant A92E, the selectivity for the FC reaction increased to 96 %; product **3** was obtained in 66% yield and 98% *ee*, while only trace amounts of nearly racemic product **5** were obtained. In contrast, when we used the mutant LmrR_W96A, the FC/EP reaction became the dominant activity, with a selectivity of 76 %; product **5** was obtained with 22 % yield and 69% *ee*. These results show how with one single mutation, either one of these catalyzed reactions can be made the dominant activity of the artificial metalloenzyme.

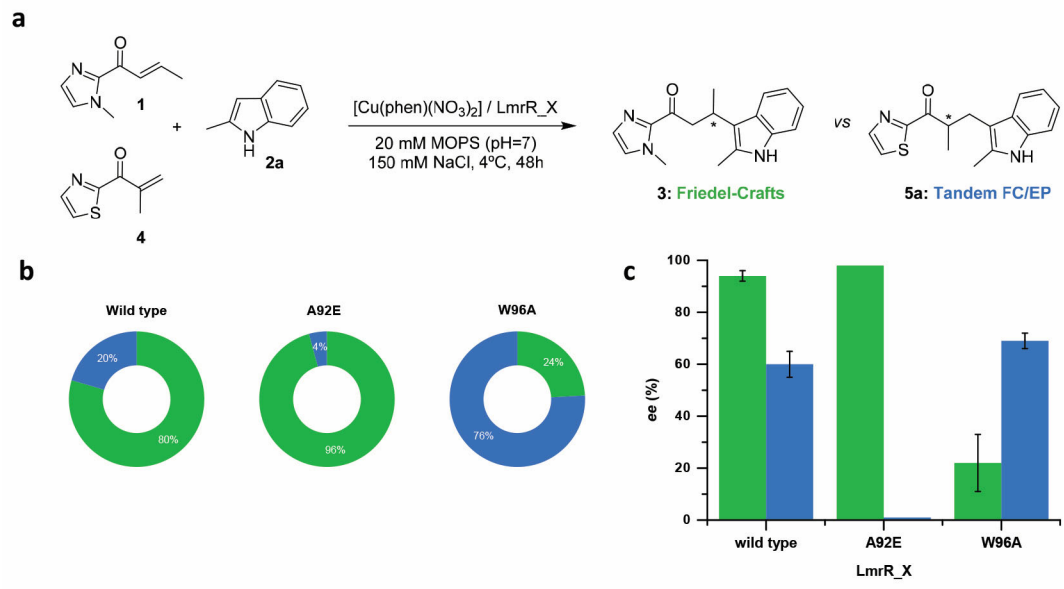


Figure 4. Competition experiment between FC reaction of **1** with **2a** (and FC/EP reaction of **4** with **2a**) catalyzed by LmrR/Cu(II)-phen, LmrR_A92E/Cu(II)-phen and LmrR_W96A/Cu(II)-phen. All substrates were present in equimolar amounts (1 mM). B) Relative product distribution (%) of the competing FC (green) and FC/EP reactions (blue). C) *ee* values for products of the FC and FC/EP reaction in the competition experiment catalyzed by LmrR mutants with standard deviations shown.

While the effect of the W96A mutation is clear, that is, it eliminates a crucial part of the Cu(II)-phen binding site, the role of the glutamate residue in the A92E mutant was less obvious. The binding affinity of the $[Cu(phen)(NO_3)_2]$ was determined (SI, Section VII) and a dissociation constant (K_d) of 65 ± 19 nM and 59 ± 16 nM was found at pH=7 and pH=5, respectively. This represents a two order of magnitude increase in binding affinity compared to WT LmrR (Table S8). Our initial hypothesis was that the carboxylate moieties would contribute to binding of the copper complex by interaction with the Cu(II) ion. For this reason, the corresponding glutamine mutant (i.e., A92Q) was prepared, since glutamine is sterically similar to glutamate, but is not a good ligand for Cu(II). However, the A92Q mutant also showed an increased affinity for Cu(II)-phen (K_d 103 ± 41 nM) (Figure S9). The increased affinity of the A92E mutant for Cu(II)-phen also allowed to determine the apparent catalytic efficiency of this improved mutant for the FC reaction: k_{cat}/K_M = 73.3 M^{-1} min^{-1}(Figure S10). This information could not be obtained for the wild type LmrR-based artificial metalloenzyme since, in case of the wild type protein, the binding affinity of the Cu(II)-phen complex is moderate. At higher concentrationsthis causes the substrates to start displacing the Cu(II)-phen complex, which leads to a significant decrease of the catalytic activity.[19] Unfortunately, due to substrate solubility issues, the individual Michaelis Menten parameters k_{cat} and K_M could not be determined.

Computation was then used to gain a better understanding about the effect of mutation A92E. For this purpose, the bis aqua form of the copper bound phenanthroline cofactor $[Cu(phen)(H_2O)_2]^{2+}$ was optimized via quantum calculations and embedded into the WT, A92E and A92Q variants of LmrR via protein-ligand docking (see details in Supporting Information). The best scored structures, which in all cases showed the copper-phenanthroline moiety at the centre of the cavity packed between tryptophans W96/W96', consistent with the X-ray structure, were submitted to 300 ns MD simulation.

The MD simulations for WT LmrR showed that hydrophobic interactions between A92 and V15 contribute to a somewhat closed arrangement of the active site, in which the indole rings of W96/W96' are slightly tilted with respect to each other, in agreement with the X-ray structure (Figure 5, left). Hence, the π-stacking interactions between the indole rings of residues W96/W96' and the phenanthroline ligand are not optimal. Hydrogen bonding interactions between the water ligands and residues D100/D100' contribute to maintain the positioning of the copper cofactor.

Instead, according to the MD simulations, the mutation of the alanine to glutamate in the A92E mutant disrupts the hydrophobic interaction and generates a hydrogen bonding network at the back of the active site, involving mainly residues N88 of helix α4 and N14, located at the opposite side of the cavity in helix α1 (Figure S12 and S13, left). This contributes to a different structural arrangement of the pore, with a parallel orientation of the W96/W96' indoles and an expanded hydrophobic free volume of the cavity (Figure S18, S19). The result is a better packing of the Cu(II)-phen complex by a dual π -stacking interaction from W96/W96'. Tentatively, this is related to the increased binding affinity, and causes the Cu(II)-phen being

positioned deeper inside the pore (Figure S20). While these structures are found frequently in the A92E mutant (40% of the MD simulation), they are virtually non-existent in case of the WT protein (1% of the MD simulation). Similar structures were observed for the A92Q mutant, albeit that these were less frequent than for A92E mutant (Supporting information). These results suggest that the A92E mutation does not have a direct effect on catalysis, but mainly has a structural effect, which translates a stronger binding of the Cu(II)-phen complex.

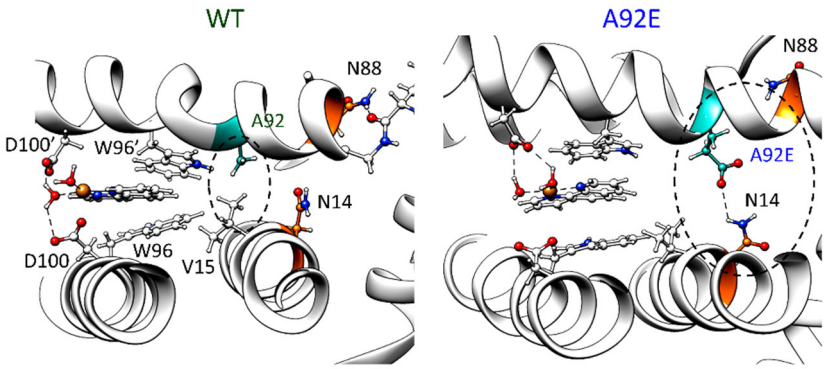


Figure 5. Comparison between representative structures of the pore along 300 ns of MD simulation for WT (left) and A92E (right) variants of LmrR with Cu(II)-phen bound. For WT, hydrophobic interactions between A92 and V15 promote closing of the active site, resulting in a not optimal π-stacking between both W96/W96' and the phenanthroline ligand of the Cu(II)-phen cofactor. In contrast, the A92E mutant enables polar interactions with N14, which contribute to the opening of the active site and a parallel orientation of the W96/W96' residues and, thus, cause a better binding of the phenanthroline ligand via π-stacking.

Combined, the data unambiguously shows that the two reactions, FC and FC/EP, do not occur at the same site in the LmrR protein. In other words, there are two – possibly more – different active sites, which each are more suited for either the FC or FC/EP reaction. The mutagenesis data and inhibition studies show that the FC reaction occurs in the hydrophobic pocket near the front entrance, close to the central tryptophans where the Cu(II)-phen complex preferentially binds. In contrast, the FC/EP reaction does not occur here. Yet, the fact that the reaction does not occur, or only very slowly, in the absence protein, combined with the enantioselectivity of product **5**, shows that the FC/EP reaction does occur at another position in or on the protein.

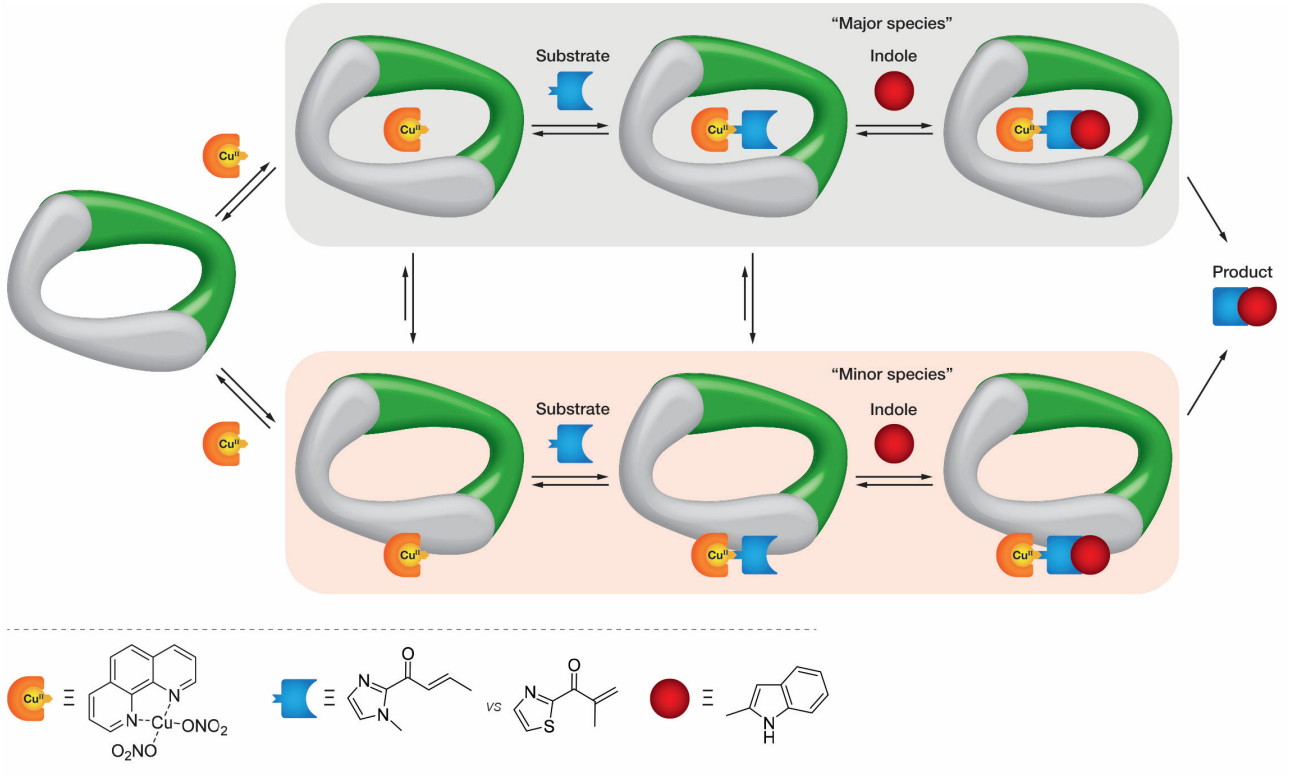


Figure 6. Schematic explanation of the catalytic promiscuity of LmrR-based artificial metalloenzymes. Most of the Cu(II)-phen complex will bind in the cavity between the two Trp residues W96/W96', where it can bind and activate the substrate to undergo conjugate addition by an indole nucleophile, resulting in formation of the product (top pathway, "major species") This pathway is preferred for the FC reaction. However, a small but non-neglible fraction of the Cu(II)-phen can bind at other positions in LmrR where it can also activate a substrate for reaction with the indole (lower pathway, "minor" species). This is the pathway that is favored for the FC/EP reaction.

This is a possibility in view of the moderate binding affinity of Cu(II)-phen for WT LmrR, which is in the micromolar range. This means that under the conditions of catalysis, most of the Cu(II)-phen is bound between the tryptophans (major species), but a non-neglible fraction of the Cu(II) complex does not bind there and, most likely, can interact at other positions of the protein, be it in the pore or on the surface of the WT LmrR (minor species) (Figure 6). This means that the catalyst is actually a mixture of copper complexes in a different environment and, depending on the reaction, it is one of these that is most reactive and may dominate the outcome of the catalyzed reaction. This is highly reminiscent of what was observed before in our work on salmon testes DNA-based catalysis.[23]

This behavior of the present artificial metalloenzymes can be rationalized by considering the Curtin-Hammett principle, which states that in case of competing pathways involving rapidly interconverting intermediates, the outcome of the reaction is solely determined by the relative kinetics of the pathways. In the present case this means that there are multiple intermediate LmrR/Cu(II)-phen_substrate complexes. In the case of the FC reaction, it is the major species, i.e. with the Cu(II)-phen_substrate complex bound between W96/W96', that reacts with high activity and enantioselectivity. However, in case of the FC/EP, this appears to be an

unproductive situation and does not result in reaction. In contrast, it is the minor substrate-bound complex that is bound at a different location, possibly on the protein exterior, that reacts much faster and, hence, is responsible for the observed catalysis.

The competition experiments support this hypothesis. Using LmrR, both the FC and FC/EP reaction are possible, but the former is preferred. Making a single mutation, A92E, results in a much higher binding affinity for the Cu complex, stabilizing the metal cofactor binding between the tryptophans, which is the most optimal conformation for the FC reaction. Hence, the reaction occurs almost exclusively via the major species (top pathway) and the lower pathway (via the minor species) is effectively shut down. This is reflected in the results: the A92E mutant catalyzes only the FC reaction and not the FC/EP reaction. Conversely, by removing the central tryptophans via the W96A mutation only the lower pathway can be followed, since it eliminates the preferred binding site for the Cu(II)-phen complex. The FC reaction is still possible, albeit with lower activity and selectivity. In this case, the FC/EP reaction is now the favored reaction, resulting in an increased yield and enantioselectivity of the FC/EP product.

In conclusion, the LmrR/Cu(II)-phen artificial metalloenzyme shows dynamic behavior in the positioning of its abiological metal cofactor, which , in turn, is related to the preferred catalytic reactivity: the FC or FC/EP reaction. In the WT LmrR/Cu(II)-phen, the FC reaction, which occurs in the hydrophopbic pore of the protein, is the main activity. However, it exhibits lower, but significant, levels of activity for another reaction, the FC/EP reaction, which does not occur at the same location in the LmrR pore, but at another site near or on the protein surface.. By only 1 mutation, A92E, this artificial metalloenzyme became almost fully selective for the FC reaction, while by another mutation, i.e. W96A, the FC/EP reaction becomes the dominant activity. The switching of catalytic activity by dynamic interconversion of the position of a metal cofactor has not been described for natural enzymes and, to date, appears to be unique to supramolecularly assembled artificial metalloenzymes. Thus, this study underlines the importance of structural dynamics as a key element of artificial enzyme design.

Acknowledgments

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities. This project was supported by the European Research Council (ERC starting grant no. 280010), the Netherlands Organisation for Scientific Research (NWO) (Vici grant 724.013.003), a postdoctoral grant from Xunta de Galicia (Plan I2C, to L.V.) and a grant from the Spanish MINECO (CTQ2017- 87889-P). G.R. acknowledges support from the Ministry of Education Culture and Science (Gravitation programme no. 024.001.035).

Keywords

Artificial metalloenzymes, biocatalysis, structural dynamics, enzyme design, copper

Supporting Information

Detailed experimental procedures, characterization data for all new compounds and proteins, additional catalysis and biophysical data, x-ray crystallography data, details for molecular dynamics (MD) simulations, and DFT calculations.

References

(1) Pauling, L. Nature of Forces between Large Molecules of Biological Interest. *Nature* **1948**, *161* (4097), 707–709.

(2) James, L. C.; Tawfik, D. S. Conformational Diversity and Protein Evolution - A 60-Year-Old Hypothesis Revisited. *Trends Biochem. Sci.* **2003**, *28* (7), 361–368.

(3) Tokuriki, N.; Tawfik, D. S. Protein Dynamism and Evolvability. *Science* **2009**, *324* (5924), 203–207.

(4) Hammes-Schiffer, S. Catalytic Efficiency of Enzymes: A Theoretical Analysis. *Biochemistry* **2013**, *52* (12), 2012–2020.

(5) Petrović, D.; Risso, V. A.; Kamerlin, S. C. L.; Sanchez-Ruiz, J. M. Conformational Dynamics and Enzyme Evolution. *J. R. Soc. Interface* **2018**, *15* (144), 20180330.

(6) Leveson-Gower, R. B.; Mayer, C.; Roelfes, G. The Importance of Catalytic Promiscuity for Enzyme Design and Evolution. *Nature Reviews Chemistry* **2019**, *3* (12), 687–705.

(7) 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* (1), 142–231.

(8) Davis, H. J.; Ward, T. R. Artificial Metalloenzymes: Challenges and Opportunities. *ACS Cent. Sci.* **2019**, *5* (7), 1120–1136.

(9) Drienovská, I.; Roelfes, G. Artificial Metalloenzymes for Asymmetric Catalysis by Creation of Novel Active Sites in Protein and DNA Scaffolds. *Isr. J. Chem.* **2015**, *55* (1), 21–31.

(10) Pàmies, O.; Diéguez, M.; Bäckvall, J.-E. Artificial Metalloenzymes in Asymmetric Catalysis: Key Developments and Future Directions. *Adv. Synth. Catal.* **2015**, *357* (8), 1567–1586.

(11) Lewis, J. C. Artificial Metalloenzymes and Metallopeptide Catalysts for Organic Synthesis. *ACS Catal.* **2013**, *3* (12), 2954–2975.

(12) Natoli, S. N.; Hartwig, J. F. Noble-Metal Substitution in Hemoproteins: An Emerging Strategy for Abiological Catalysis. *Acc. Chem. Res.* **2019**, *52* (2), 326-335.

(13) Oohora, K.; Onoda, A.; Hayashi, T. Hemoproteins Reconstituted with Artificial Metal Complexes as Biohybrid Catalysts. *Acc. Chem. Res.* **2019**, *52* (4), 945–954.

(14) Huang, X.; Garcia-Borràs, M.; Miao, K.; Kan, S. B. J.; Zutshi, A.; Houk, K. N.; Arnold, F. H. A Biocatalytic Platform for Synthesis of Chiral α-Trifluoromethylated Organoborons. *ACS Cent. Sci.* **2019**, *5* (2), 270–276.

(15) Villarino, L.; Splan, K. E.; Reddem, E.; Alonso-Cotchico, L.; Gutiérrez de Souza, C.; Lledós, A.; Maréchal, J.-D.; Thunnissen, A.-M. W. H.; Roelfes, G. An Artificial Heme Enzyme for Cyclopropanation Reactions. *Angew. Chemie Int. Ed.* **2018**, *57* (26), 7785–7789.

(16) Madoori, P. K.; Agustiandari, H.; Driessen, A. J. M.; Thunnissen, A.-M. W. H. Structure of the Transcriptional Regulator LmrR and Its Mechanism of Multidrug Recognition. *EMBO J.* **2009**, *28* (2), 156–166.

(17) Takeuchi, K.; Tokunaga, Y.; Imai, M.; Takahashi, H.; Shimada, I. Dynamic Multidrug Recognition by Multidrug Transcriptional Repressor LmrR. *Sci. Rep.* **2015**, *4* (1), 6922.

(18) van der Berg, J. P.; Madoori, P. K.; Komarudin, A. G.; Thunnissen, A.-M.; Driessen, A. J. M. Binding of the Lactococcal Drug Dependent Transcriptional Regulator LmrR to Its Ligands and Responsive

Promoter Regions. *PLoS One* **2015**, *10* (8), e0135467.

(19) Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G. Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold. *J. Am. Chem. Soc.* **2015**, *137* (31), 9796–9799.

(20) Roelfes, G. LmrR: A Privileged Scaffold for Artificial Metalloenzymes. *Acc. Chem. Res.* **2019**, *52* (3), 545–556.

(21) Mohr, J. T.; Hong, A. Y.; Stoltz, B. M. Enantioselective Protonation. *Nat. Chem.* **2009**, *1* (5), 359–369.

(22) García-Fernández, A.; Megens, R. P.; Villarino, L.; Roelfes, G. DNA-Accelerated Copper Catalysis of Friedel-Crafts Conjugate Addition/Enantioselective Protonation Reactions in Water. *J. Am. Chem. Soc.* **2016**, *138* (50), 16308-16314.

(23) Boersma, A. J.; Klijn, J. E.; Feringa, B. L.; Roelfes, G. DNA-Based Asymmetric Catalysis: Sequence-Dependent Rate Acceleration and Enantioselectivity. *J. Am. Chem. Soc.* **2008**, *130* (35), 11783–11790.

Dynamics of metal complex binding in relation to catalytic activity and selectivity of an artificial metalloenzyme

*Lara Villarino, Shreyans Chordia, Lur Alonso-Cotchico, Eswar Reddem, Andy-Mark W.H. Thunnissen, Jean-Didier Maréchal and Gerard Roelfes**

Supporting Information

Contents:

I. Molecular Biology

General remarks

Unless otherwise noted, all chemicals and reagents were commercially available and used without further purification. *E. coli* NEB 5-aplha-T1 Phage Resistant (New England Biolabs) strain was used for cloning and *E. coli* BL21(DE3)-C43-T1 Phage Resistant (New England Biolabs) strain was used for expression of proteins. Primers were synthesized by Eurofins MWG Operon (Ebersberg, Germany). Restriction endonucleases (including DpnI) were procured from New England Biolabs. T4 DNA ligase, DNA Gel Extraction Kit and Plasmid Isolation Kit were purchased from Roche. PCRs were carried out using an Eppendorf Mastercycler Personal apparatus. Pfu Turbo polymerase was purchased from Stratagene. DNA sequencing was carried out by GATC-Biotech (Berlin, Germany). FPLC columns were purchased from GE Healthcare.

Gene Optimization

The synthesized gene of LmrR was ordered from GenScript (USA). The codon usage was adapted to the codon bias of *E. coli* genes. The gene was delivered in the cloning vector pUC57, containing a C-terminal Strep-tag and K55D, K55Q mutations as previously described.[1] The gene was subsequently re-cloned to pET17b expression vector using the restriction sites *NdeI* and *HindIII*.

Optimized gene sequence

ATGGGTGCCGAAATCCCGAAAGAAATGCTGCGTGCTCAAACCAATGTCATCCTGCTGAATGTCCTGAA
ACAAGGCGATAACTATGTGTATGGCATTATCAAACAGGTGAAAGAAGCGAGCAACGGTGAAATGGAA
CTGAATGAAGCCACCCTGTATACGATTTTTGATCGTCTGGAACAGGACGGCATTATCAGCTCTTACTGG
GGTGATGAAAGTCAAGGCGGTCGTCGCAAATATTACCGTCTGACCGAAATCGGCCATGAAAACATGCG
CCTGGCGTTCGAATCCTGGAGTCGTGTGGACAAAATCATTGAAAATCTGGAAGCAAACAAAAAATCTG
AAGCGATCAAATCTAGAGGTGGCAGCGGTGGCTGGAGCCACCCGCAGTTCGAAAAATAA

Site-directed mutagenesis

All the LmrR mutants were prepared by site-directed mutagenesis of pET17b_LmrR _coopt plasmid. The primers required for the mutagenesis are summarized in Table S1. The following PCR cycle was used: initial denaturation at 95 ^{0}C for 1 min, denaturation at 95 ^{0}C for 30 s, annealing at 58-65 ^{0}C for 1 min (with respect to Ta of the particular mutant) and extension at 68 ^{0}C for 5 min. The thermal cycle was repeated 16 times. The resulting PCR product was digested with restriction endonuclease DpnI for 1h at 37 ^{0}C. Chemically competent *E. coli* NEB 5-aplha-T1 Phage Resistant cells were transformed with 5 µL of the DpnI digested sample.

Table S1: Primers used for site-directed mutagenesis

Primer name	**DNA Sequence (5' -> 3')**
LmrR_M8A_F	ATC CCG AAA GAA GCG CTG CGT GCT CAA
LmrR_M8A_R	TTG AGC ACG CAG CGC TTC TTT CGG GAT
LmrR_V99A_F	TCC TGG AGT CGT GCG GAC AAA ATC ATT
LmrR_V99A_R	AAT GAT TTT GTC CGC ACG ACT CCA GGA
LmrR_F93A_F	CGC CTG GCG GCG GAA TCC TGG
LmrR_F93A_R	CCA GGA TTC CGC CGC CAG GCG
LmrR_Q12A_F	CTG CGT GCT GCG ACC AAT GTC A
LmrR_Q12A_R	TGA CAT TGG TCG CAG CAC GCA G
LmrR_Q12E_F	ATG CTG CGT GCT GAA ACC AAT GTC ATC
LmrR_Q12E_R	GAT GAC ATT GGT TTC AGC ACG CAG CAT
LmrR_E7A_F	GAA ATC CCG AAA GCG ATG CTG CGT GCT
LmrR_E7A_R	AGC ACG CAG CAT CGC TTT CGG GAT TTC
LmrR_V15A_F	GCT CAA ACC AAT GCG ATC CTG CTG AAT
LmrR_V15A_R	ATT CAG CAG GAT CGC ATT GGT TTG AGC
LmrR_D100A_F	TGG AGT CGT GTG GCG AAA ATC ATT GAA
LmrR_D100A_R	TTC AAT GAT TTT CGC CAC ACG ACT CCA
LmrR_D100E_F	TGG AGT CGT GTG GAA AAA ATC ATT GAA
LmrR_D100E_R	TTC AAT GAT TTT TTC CAC ACG ACT CCA
LmrR_A92E_F	AAC ATG CGC CTG GAA TTC GAA TCC TGG
LmrR_A92E_R	CCA GGA TTC GAA TTC CAG GCG CAT GTT
LmrR_A92Q_F	AAC ATG CGC CTG CAG TTC GAA TCC TGG
LmrR_A92Q_R	CCA GGA TTC GAA CTG CAG GCG CAT GTT

Expression and purification

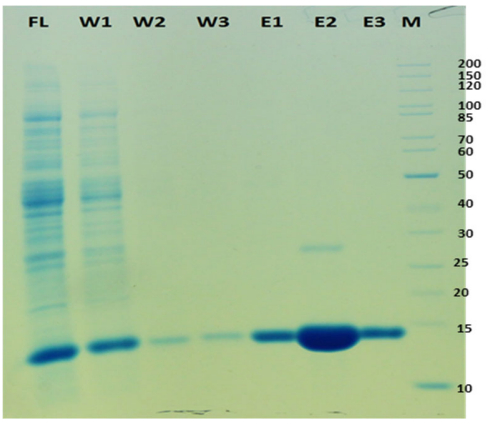
Chemically competent *E. coli* BL21(DE3)-C43-T1 Phage Resistant cells were transformed with a pET17b expression vector carrying various LmrR constructs for production of the wild type and mutant proteins. Single colonies were picked and inoculated into a starter culture of 10 mL of fresh LB medium containing 100 μg/mL of ampicillin. The entire starter culture was used to inoculate 1 L of fresh LB medium containing 100 μg/mL of ampicillin. When the culture reached an optical density at 600 nm between 0.8–0.9, isopropyl β-D-1-thiogalactopyranoside (IPTG) at final concentration of 1 mM was added to induce the expression of the target protein. Expressions were carried out at 30 °C, overnight. Cells were harvested by centrifugation (6000 rpm, JA10, 20 min, 4 °C, Beckman). The pellet was re-suspended in 25mL of 50 mM NaH_2PO_4, pH 8.0, 150 mM NaCl along with half a tablet of mini complete EDTA-free protease inhibitor cocktail (Roche). The re-suspended cells were flash frozen with liquid nitrogen or at -20 °C, overnight. The frozen solution was thawed quickly in a water bath at 37 °C and transferred to ice. Next, DnaseI (final

concentration, 0.1 mg/mL) and $MgCl_2$ (final concentration, 10mM) were added. Sonication was carried out (75% (200W) for 8 min (10 s on, 15 s off). The lysed cells were incubated with for 30 min on ice. Additional sheer forcing with a syringe and a long needle was applied for at least 2 times. After centrifugation (16000 rpm, JA-17, 45 min, 4 °C, Beckman), 12.5 mL of the supernatant was equilibrated with 2.5 mL of pre-equilibrated Strep-tag Tactin column material for 1 h (mixed at 200 rpm on a rotary shaker) at 4 °C. The column was washed with 3 x 1 CV (column volume) of re-suspension buffer (same as buffer used before), and eluted with 0.6 CV, 1.6 CV and 0.8 CV of resuspension buffer containing 5 mM desthiobiotin. Fractions were analyzed on a 12% polyacrylamide SDS-Tris Tricine gel followed by Coomassie Blue staining. Fractions containing protein were re-buffered to 20 mM MOPS, pH 7.0, 150 mM NaCl using an Illustra NAP-10 Desalting Column (GE Healthcare Life Sciences). The concentration of LmrR and the various mutants were determined by using the calculated extinction coefficient obtained from Protparam on the Expasy server (for example, LmrR monomer has $\varepsilon 280 = 25440\ M^{-1}\ cm^{-1}$).[2] Expression yields typically were 20-30 mg L^{-1} media.

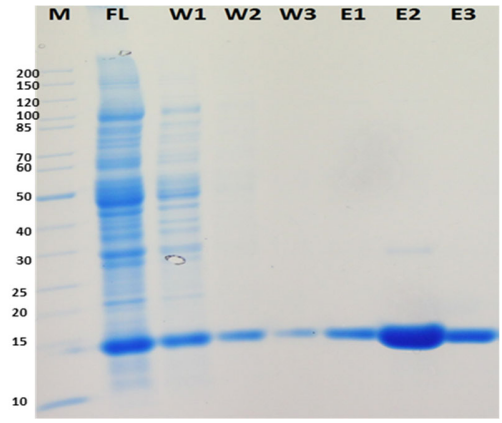
SDS-PAGE

Figure S1: SDS-PAGE gels of the LmrR mutants after Strep-Tag purification

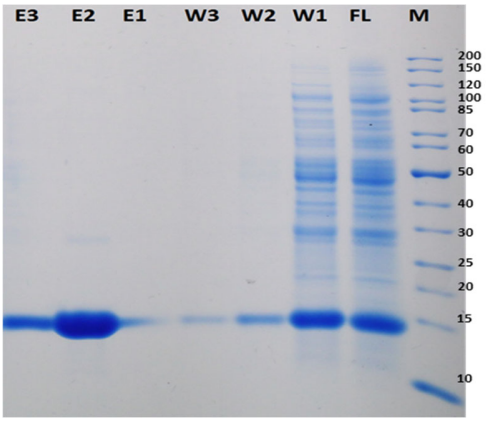
LmrR



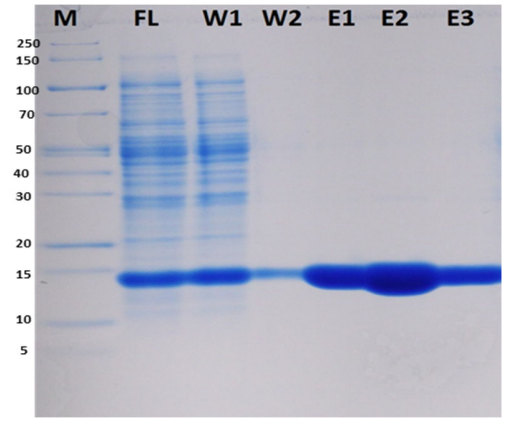
LmrR_ **W96A**



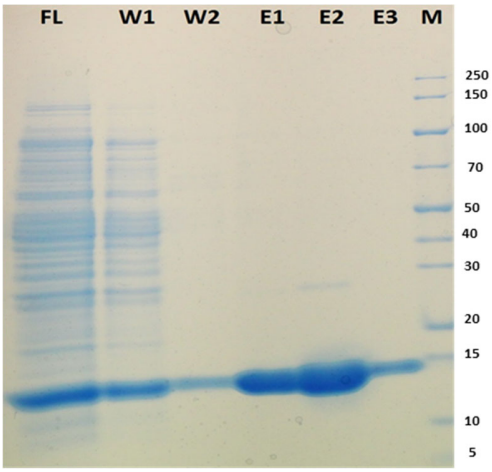
LmrR_ **M8A**



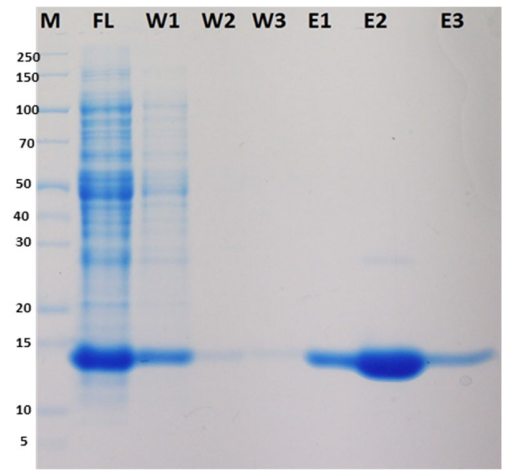
LmrR_ **V99A**



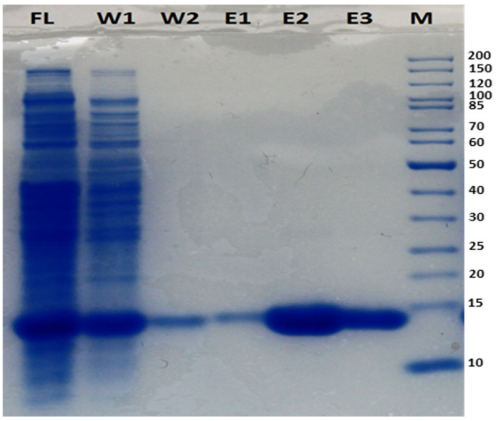
LmrR_ **F93A**



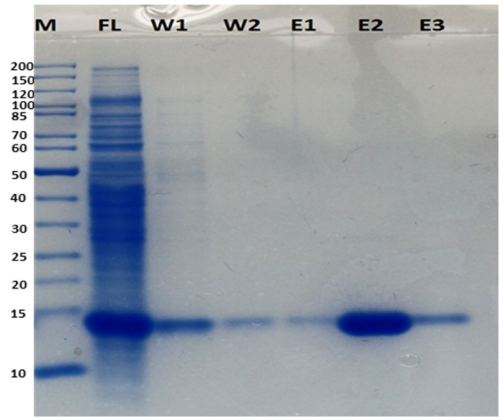
LmrR_ **Q12A**



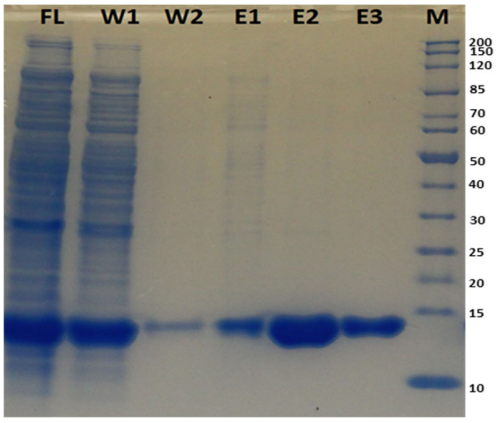
LmrR_ **Q12E**



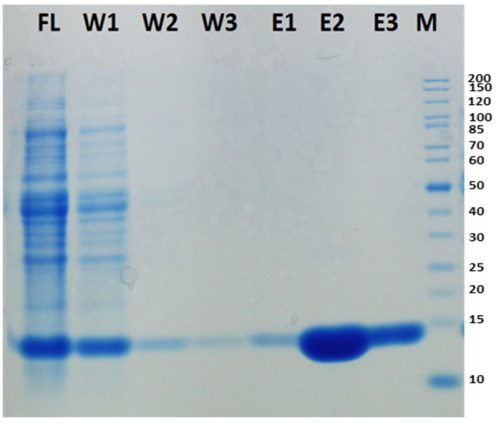
LmrR_ **E7A**



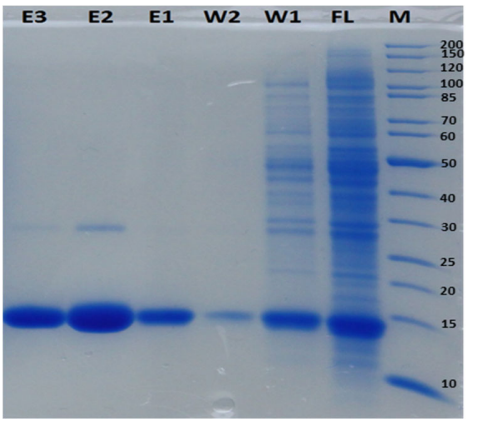
LmrR_ **V15A**



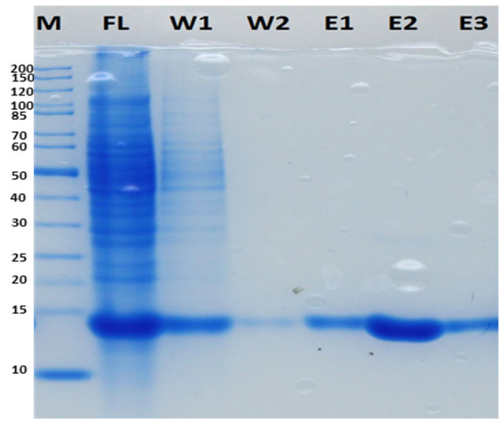
LmrR_ **D100A**



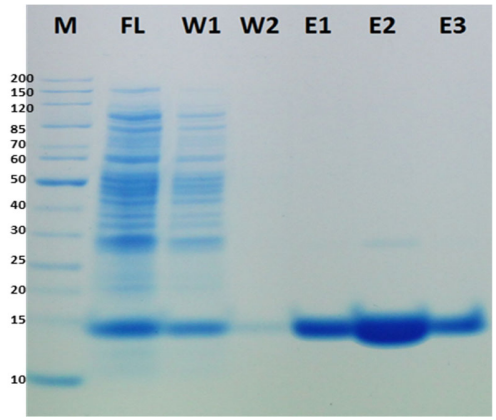
LmrR_ **D100E**



LmrR_ **A92E**



LmrR_ **A92Q**



Legend:

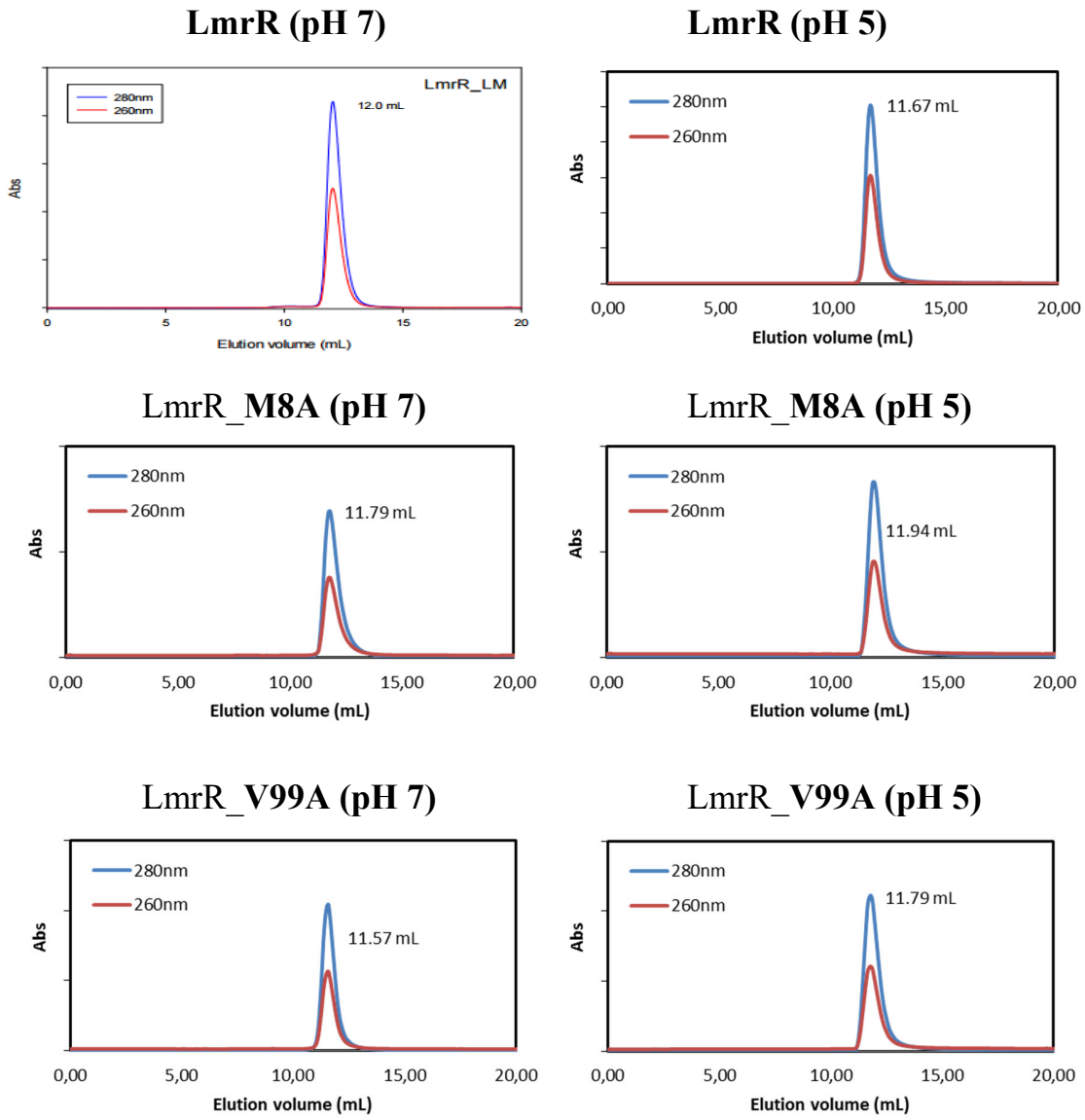
M- Marker (PageRuler™ Unstained Protein Ladder, mol wt 10kDa to 200kDa or PageRuler™ Unstained Broad Range, mol wt 5 to 250kDa). FL – Flow-through column. W - Wash fraction. E - Elution fraction. Gels were stained with InstantBlueTM (Expedeon).

Analytical Size Exclusion Chromatography

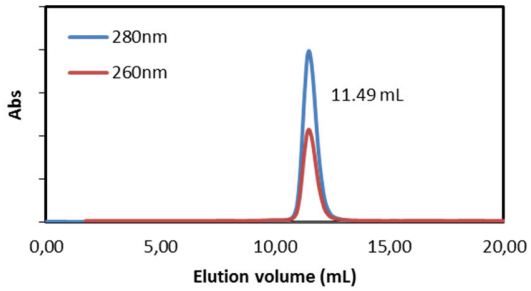
General procedure

Analytical size exclusion chromatography was performed on a Superdex 75 10/300 GL (GE Healthcare). 100 µL of purified protein sample in 20 mM MOPS, pH 7.0 and 150 mM NaCl or 20 mM MES, pH 5.0 and 150 mM NaCl was injected (flow 0.5 mL/min). The column was calibrated using the standard Gel Filtration LMW Calibration Kit (GE Healthcare).

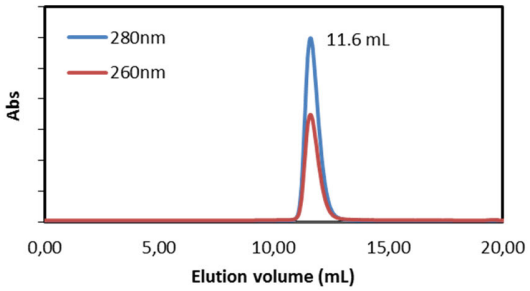
Figure S2: Analytical size-exclusion chromatography (Superdex-75 10/300 GL) of LmrR and mutants in (left) 20 mM MOPS, pH 7.0 and 150 mM NaCl; (right) 20 mM MES, pH 5.0 and 150 mM NaCl



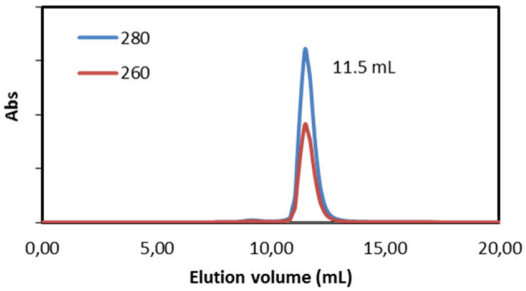
LmrR_**F93A (pH 7)**



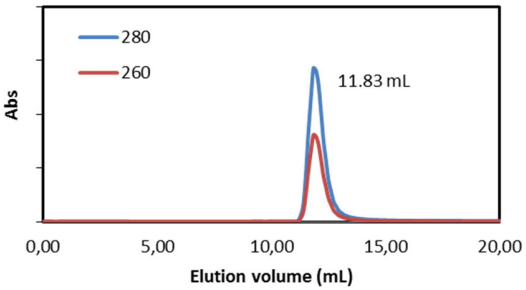
LmrR_**F93A (pH 5)**



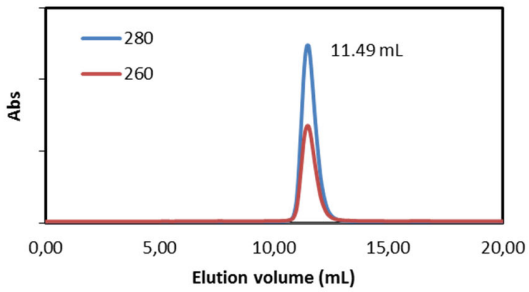
LmrR_**Q12A (pH 7)**



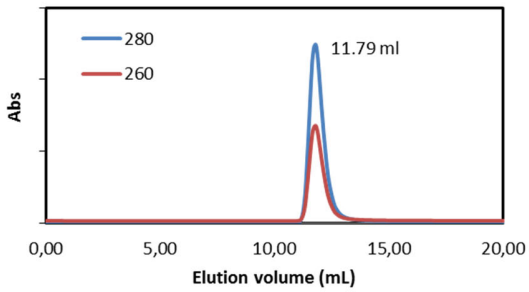
LmrR_**Q12A (pH 5)**



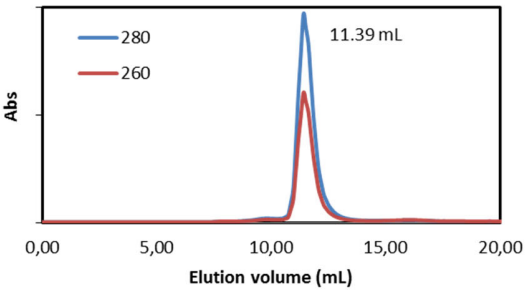
LmrR_**Q12E (pH 7)**



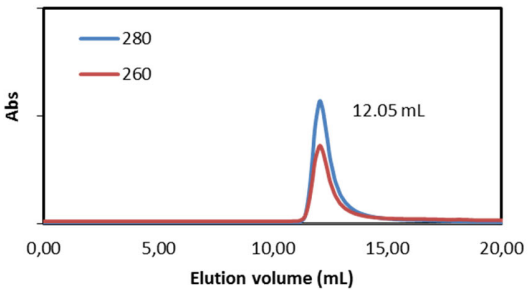
LmrR_**Q12E (pH 5)**



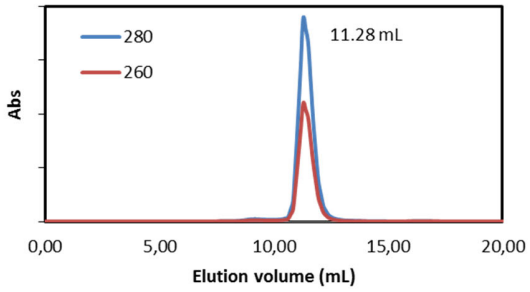
LmrR_**E7A (pH 7)**



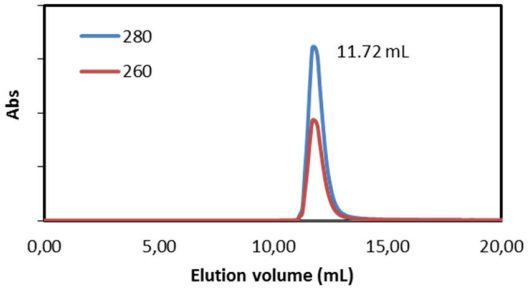
LmrR_**E7A (pH 5)**



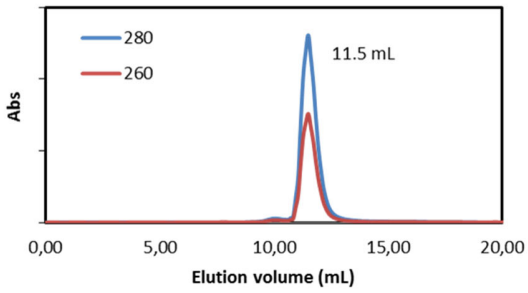
LmrR_**V15A (pH 7)**



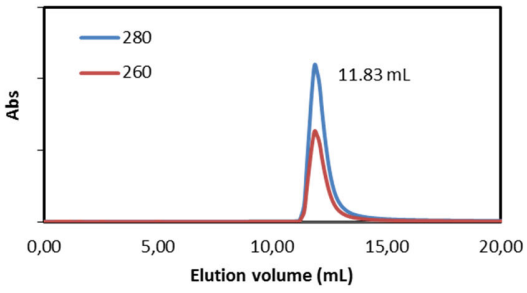
LmrR_**V15A (pH 5)**



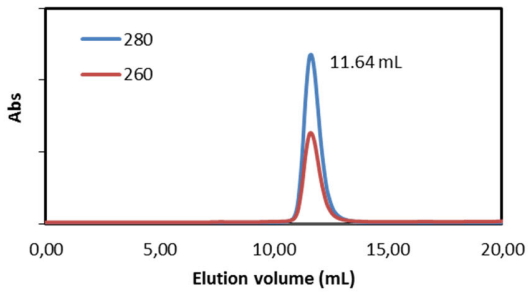
LmrR_**D100A (pH 7)**



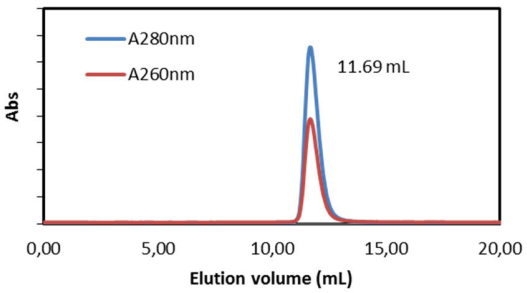
LmrR_**D100A (pH 5)**



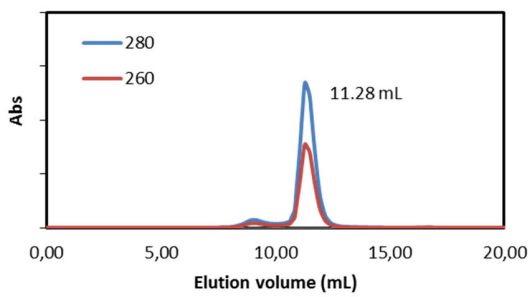
LmrR_**D100E (pH 7)**



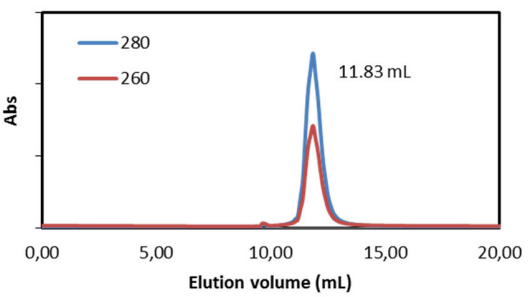
LmrR_**D100E (pH 5)**



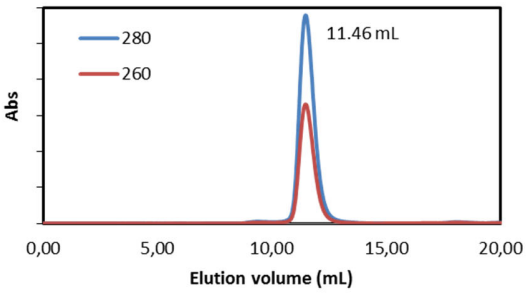
LmrR_**A92E (pH 7)**



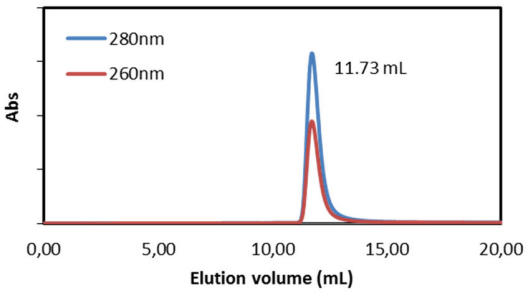
LmrR_**A92E (pH 5)**



LmrR_**A92Q (pH 7)**



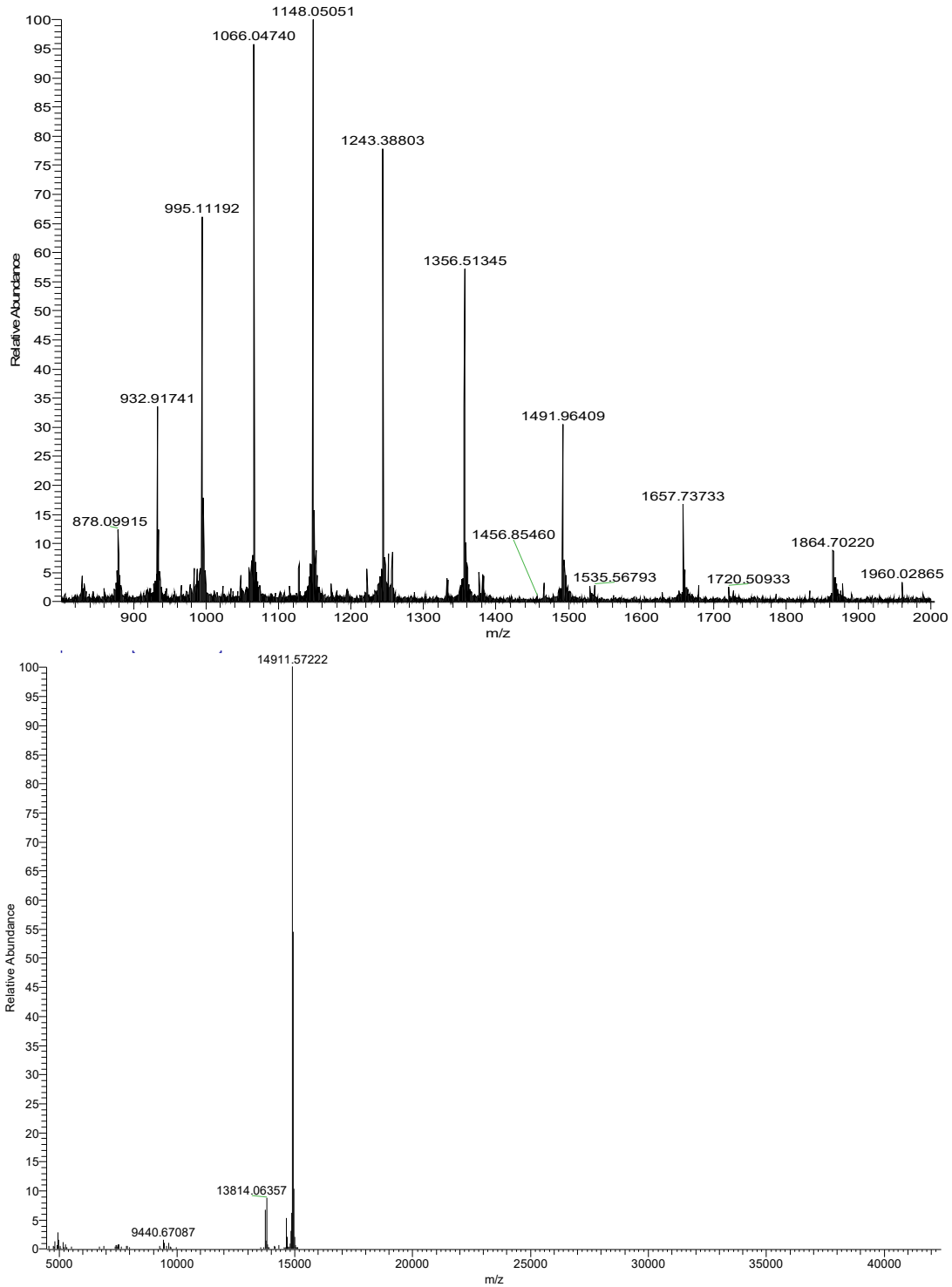
LmrR_**A92Q (pH 5)**



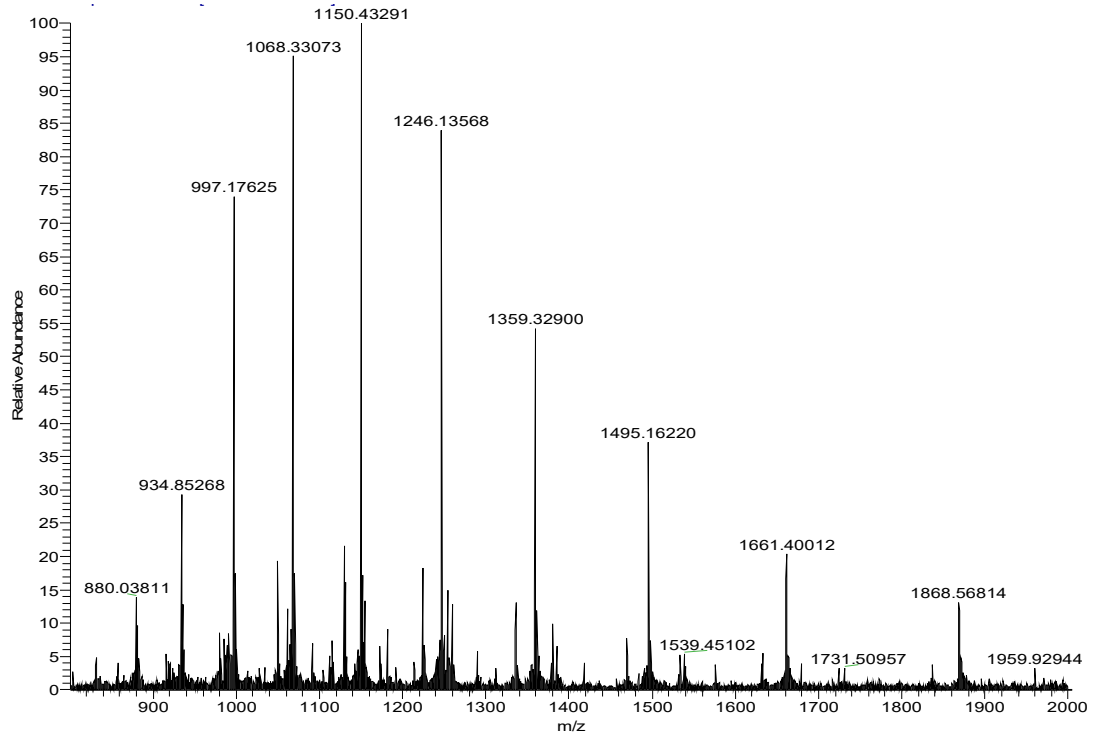
LC-MS ESI spectra

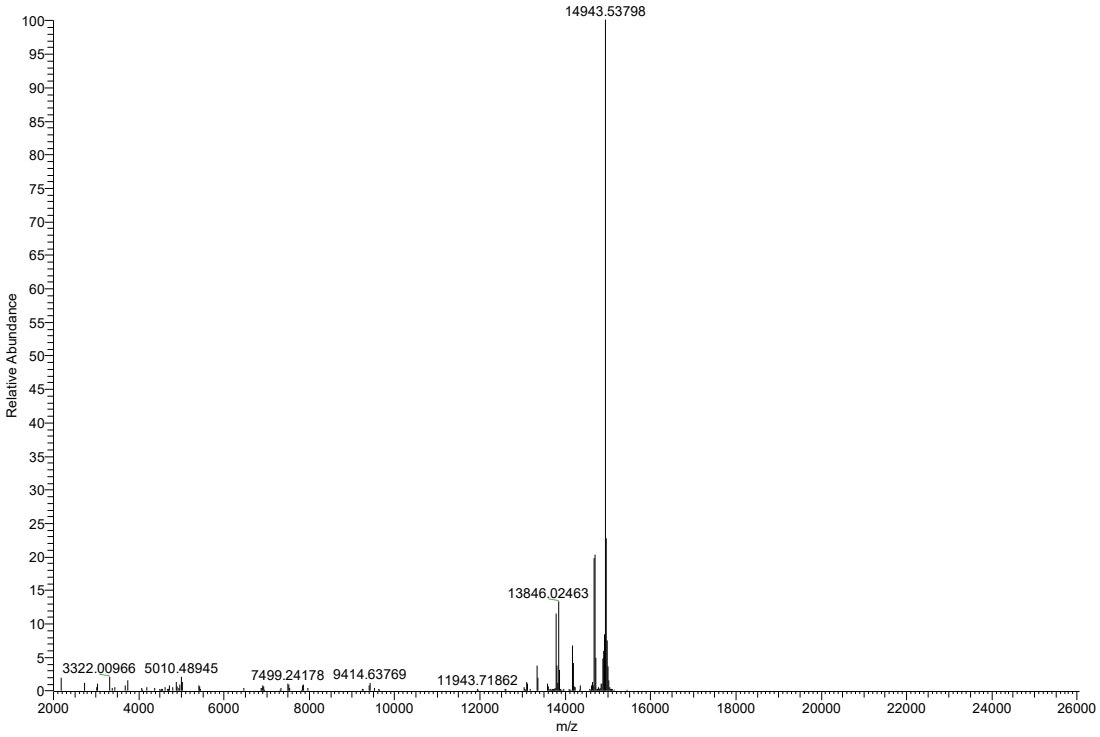
Figure S3: Electronspray ionization (ESI) mass spectra of LmrR and mutants

LmrR_**M8A** calculated mass (-Met) 14910.6494

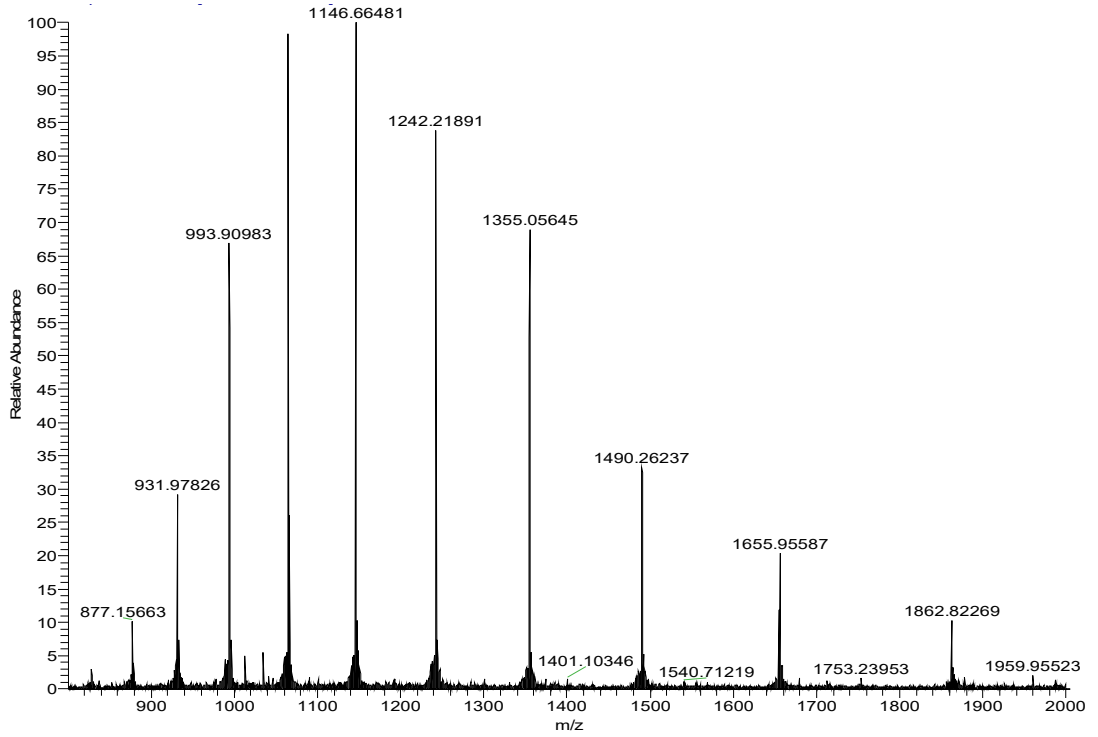


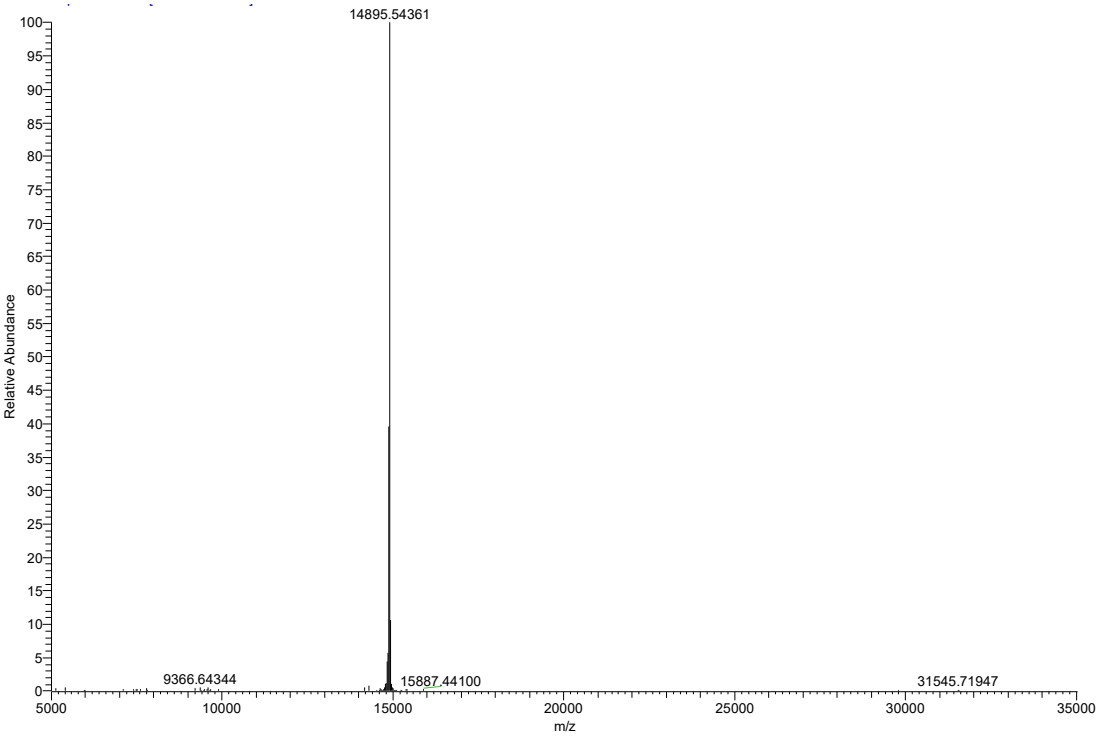
LmrR_**V99A** calculated mass (-Met) 14942.7080



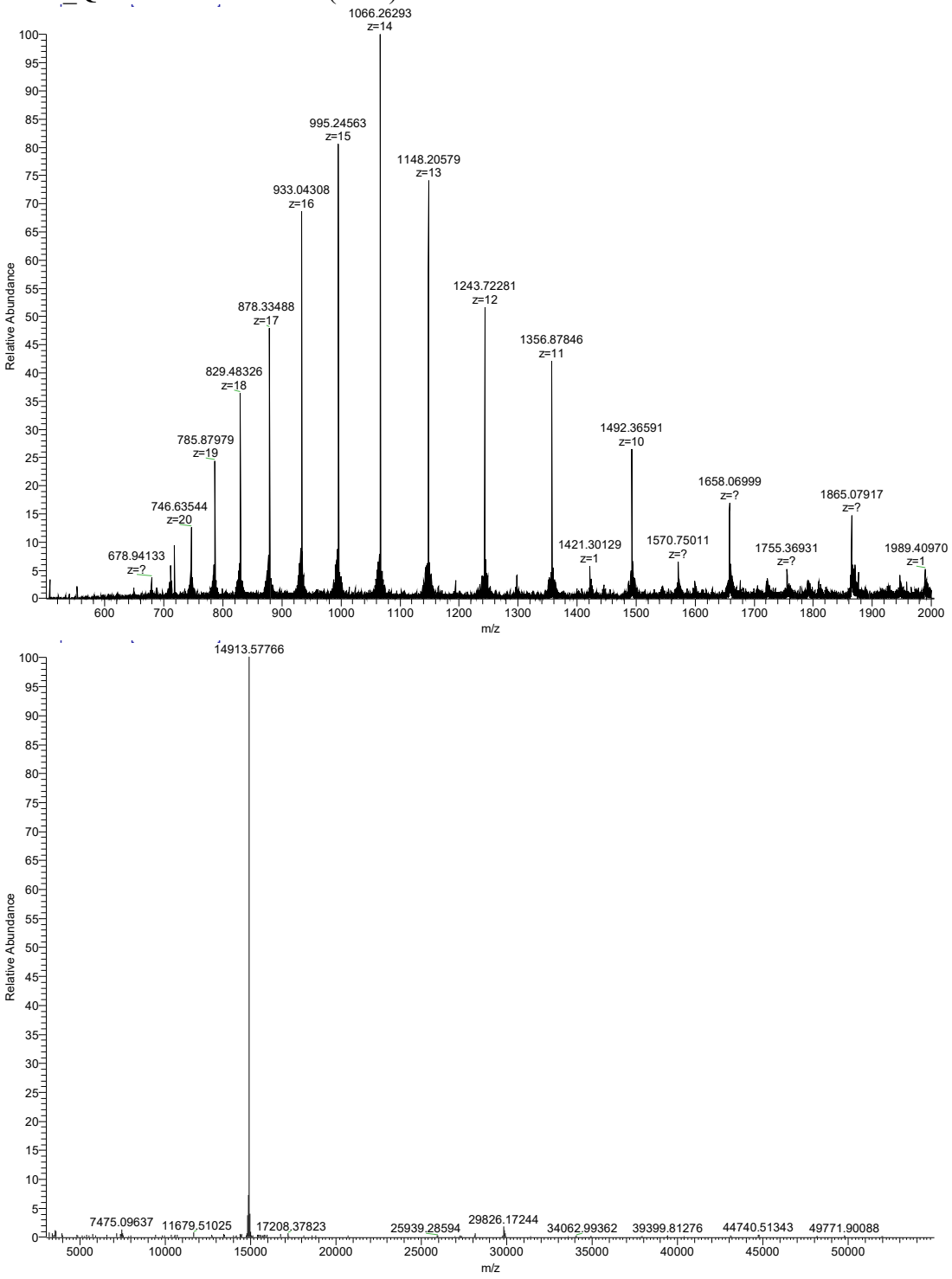


LmrR_**F93A** calculated mass (-Met) 14894.71

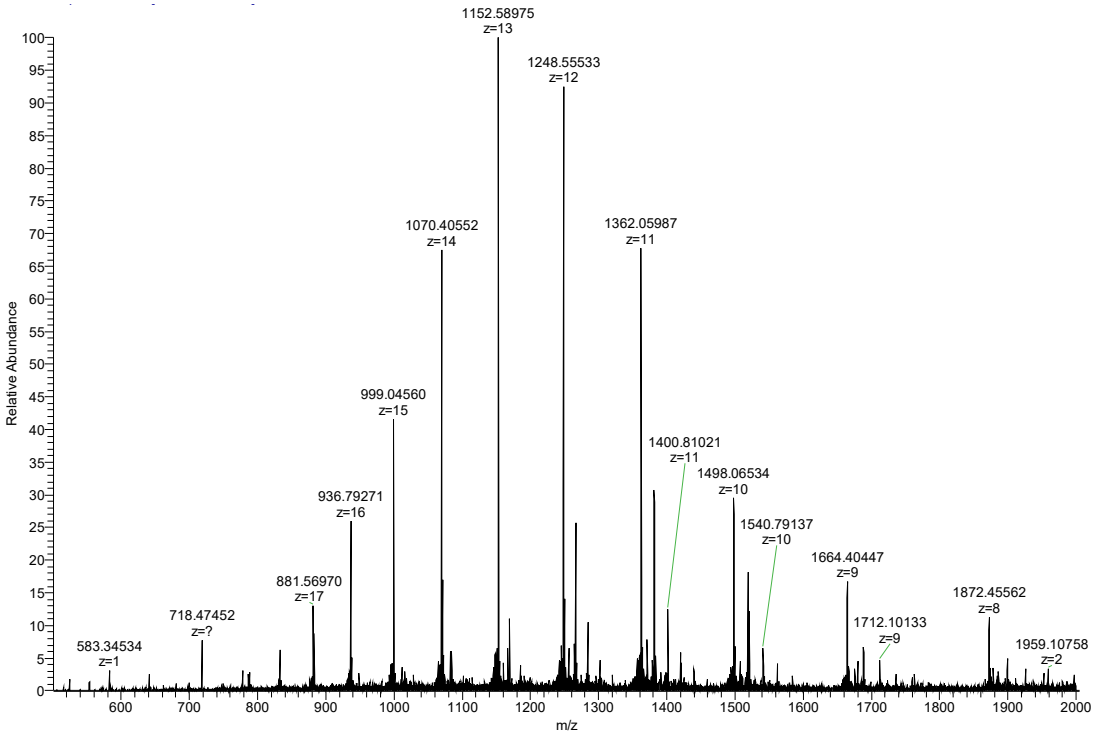


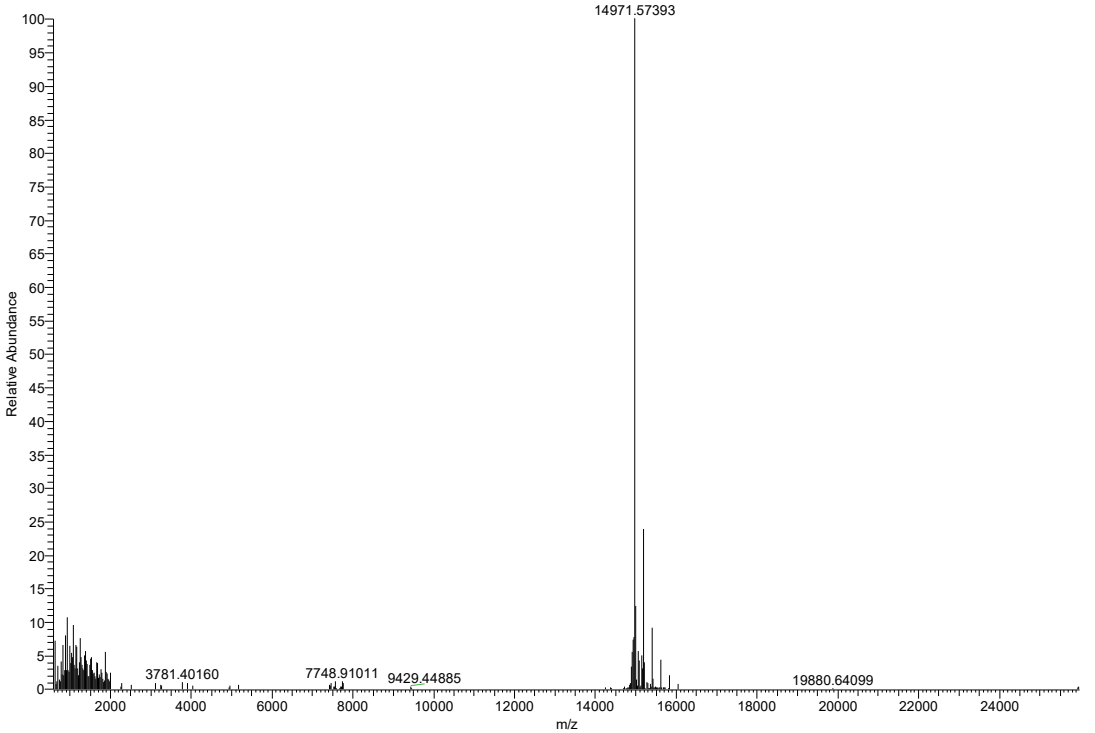


LmrR_**Q12A** calculated mass (-Met) 14913.75

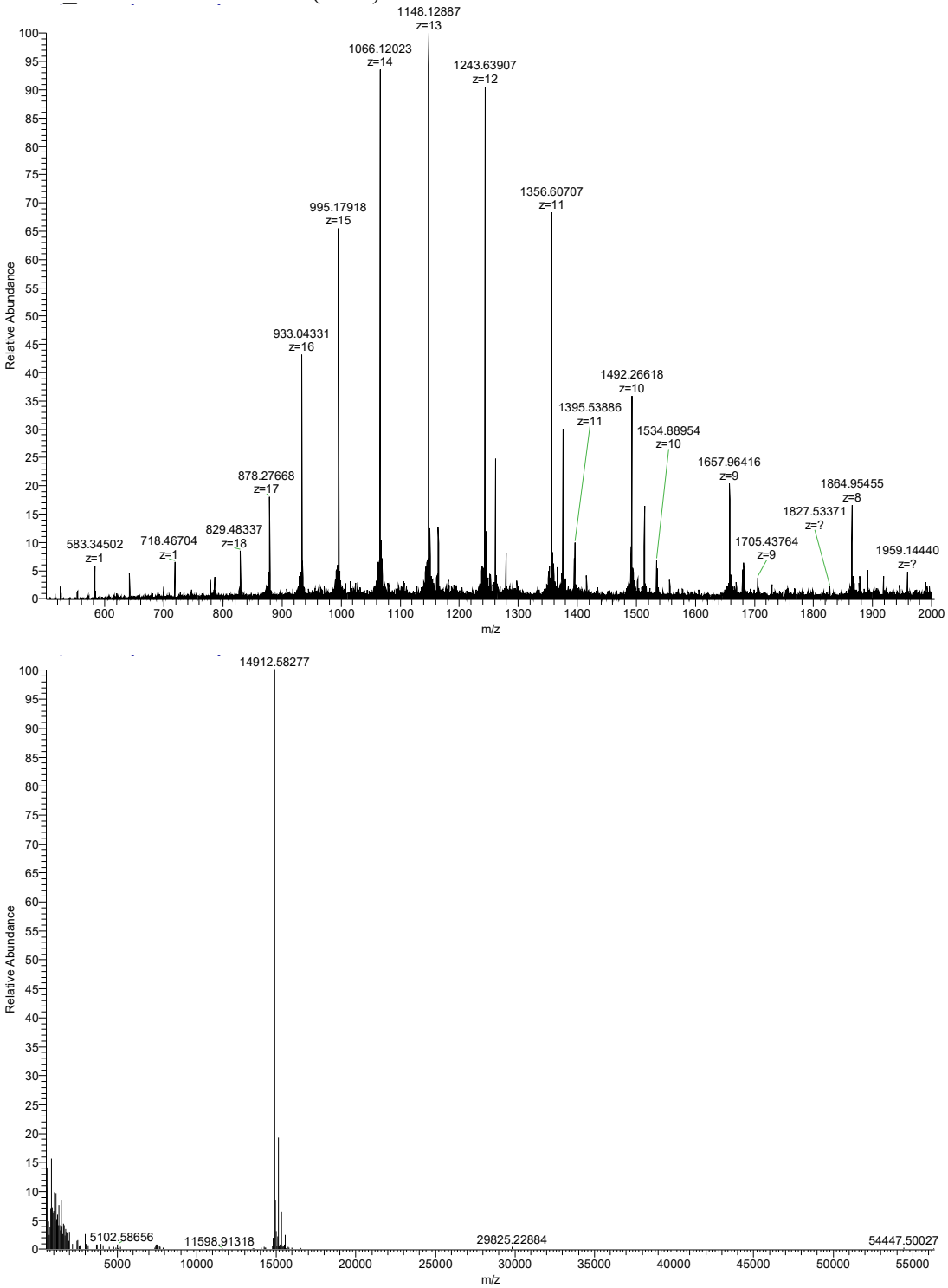


LmrR_**Q12E** calculated mass (-Met) 14971.75

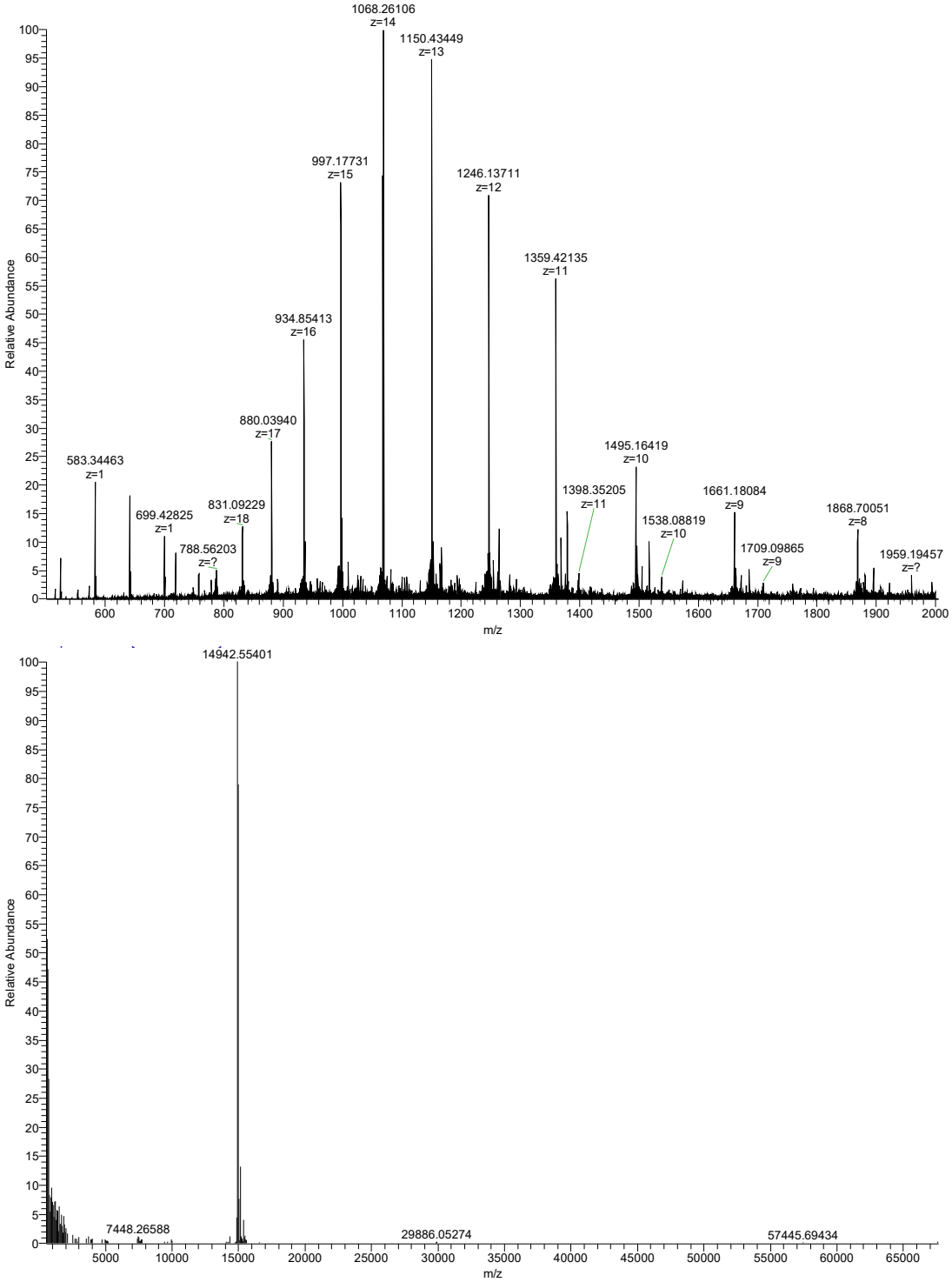




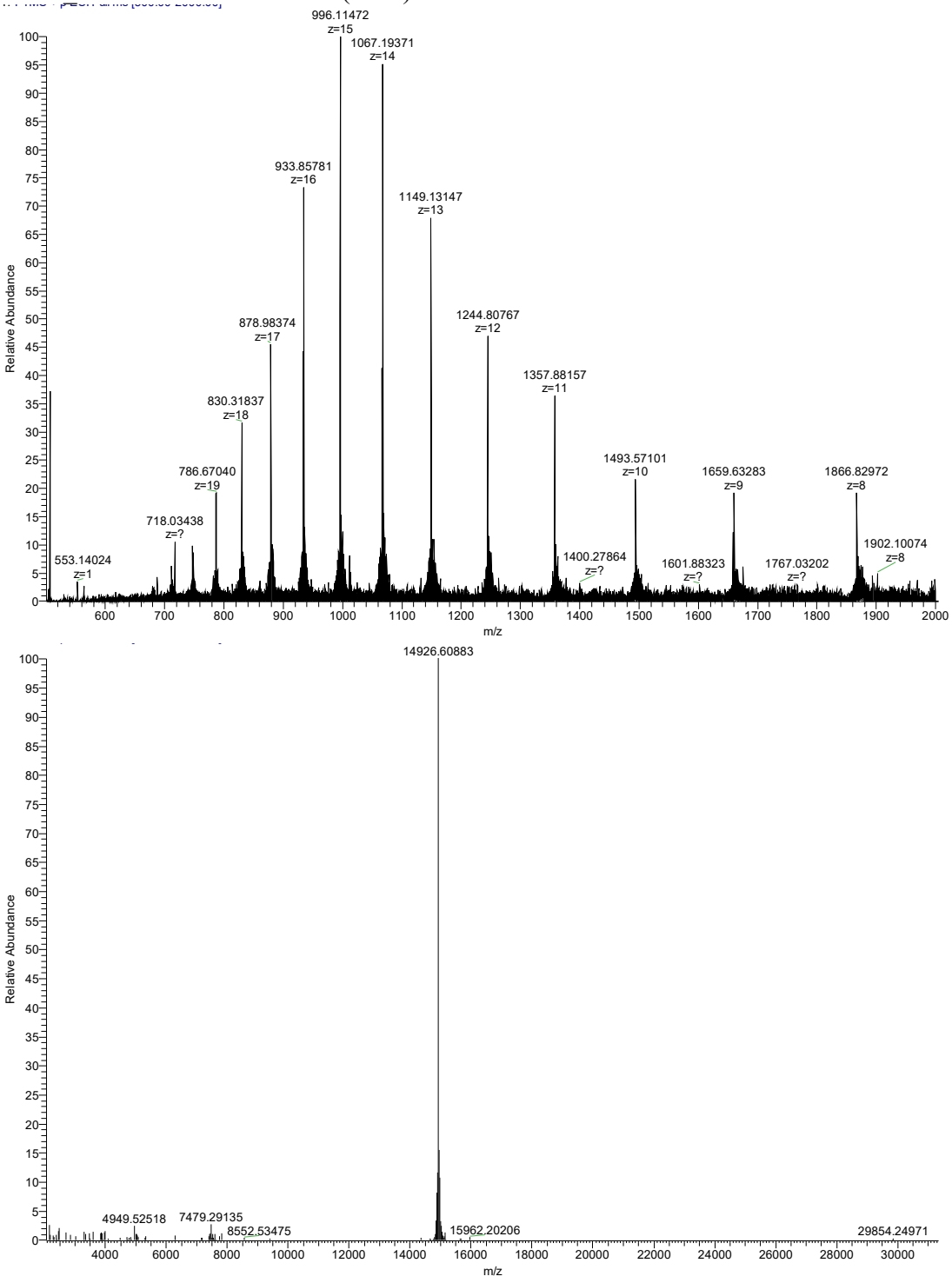
LmrR_**E7A** calculated mass (-Met) 14912.73



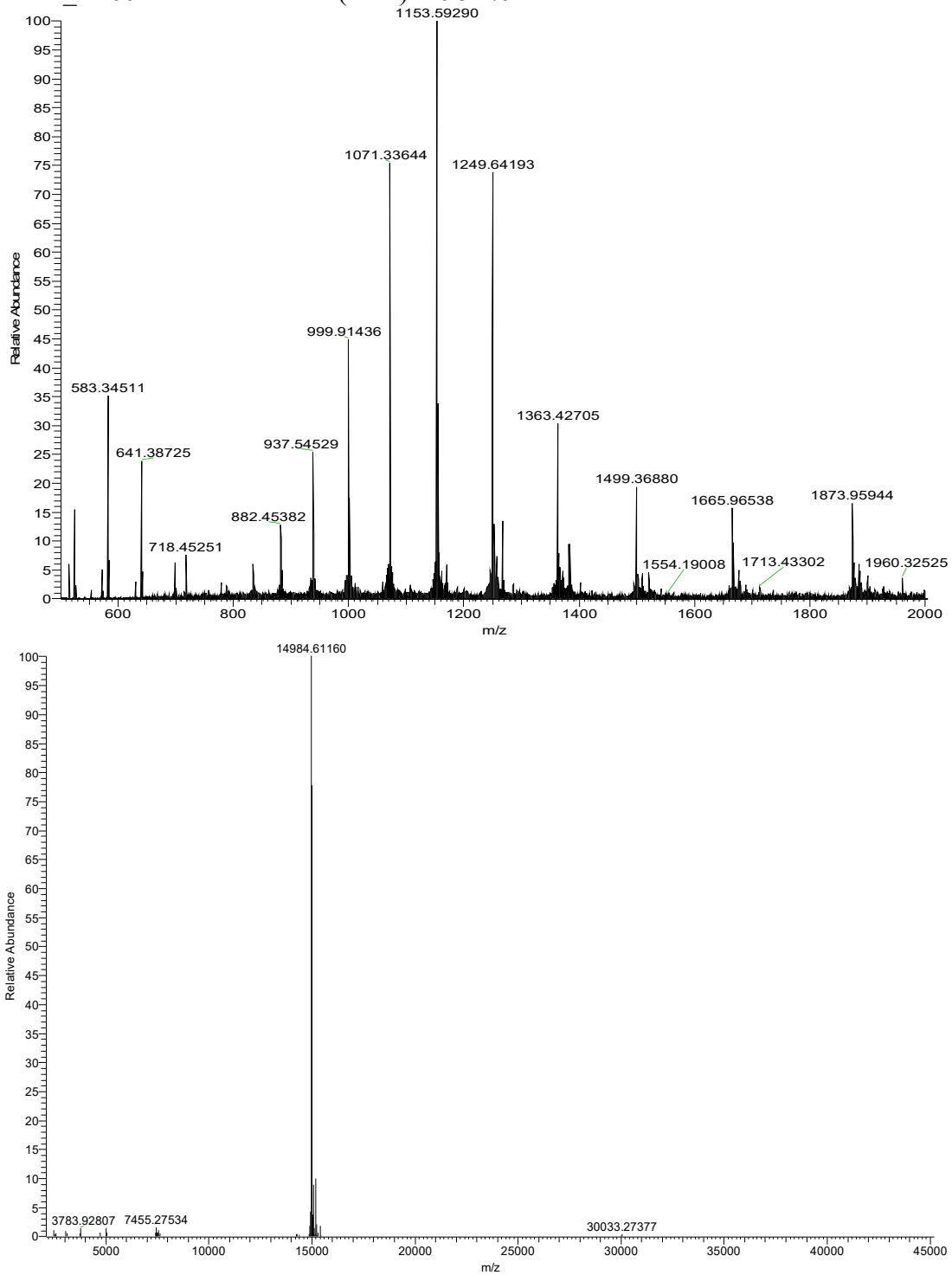
LmrR_**V15A** calculated mass (-Met) 14942.71



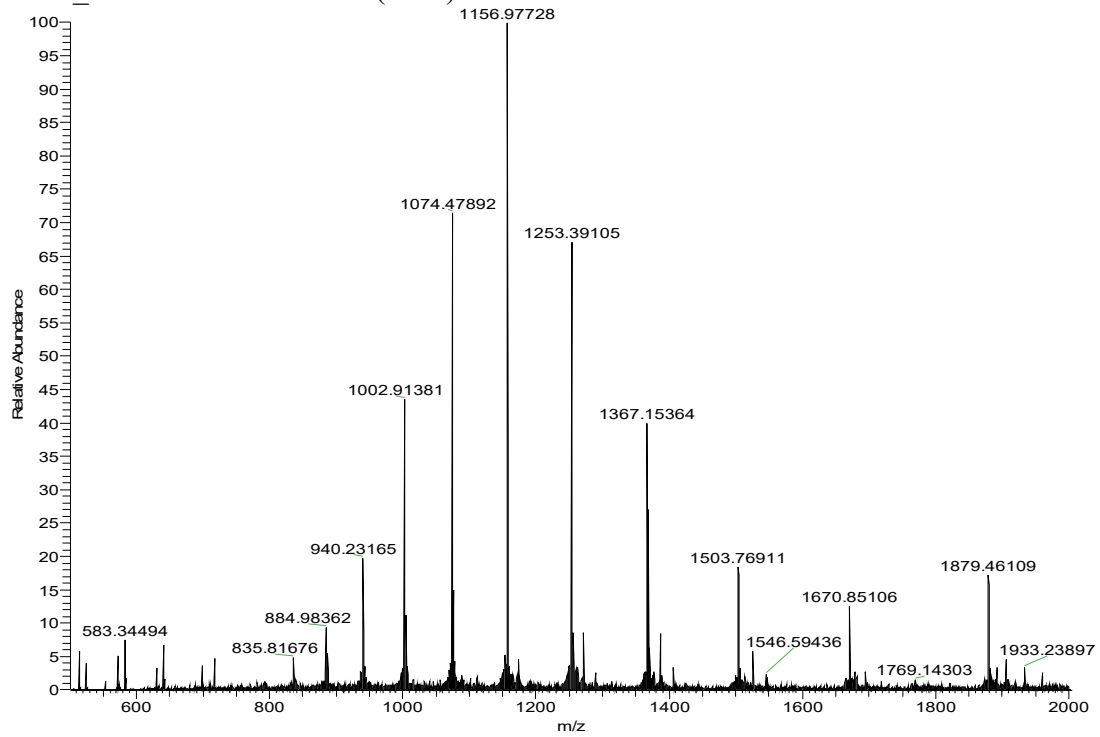
LmrR_**D100A** calculated mass (-Met) 14926.79

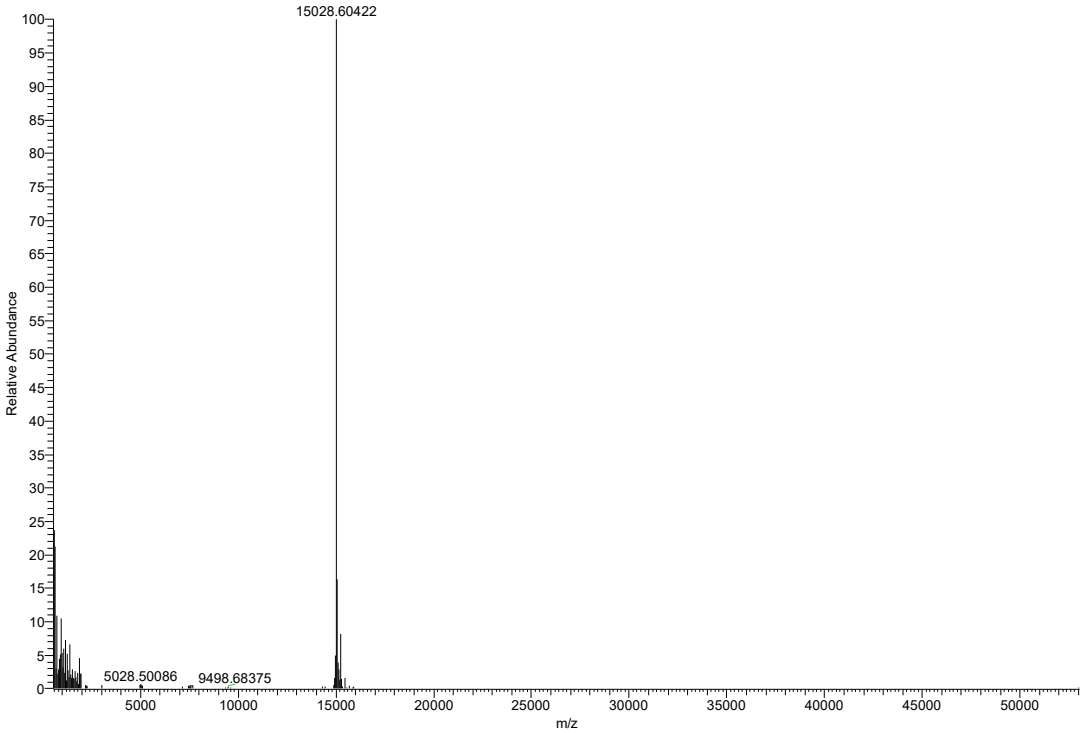


LmrR_**D100E** calculated mass (-Met) 14984.79

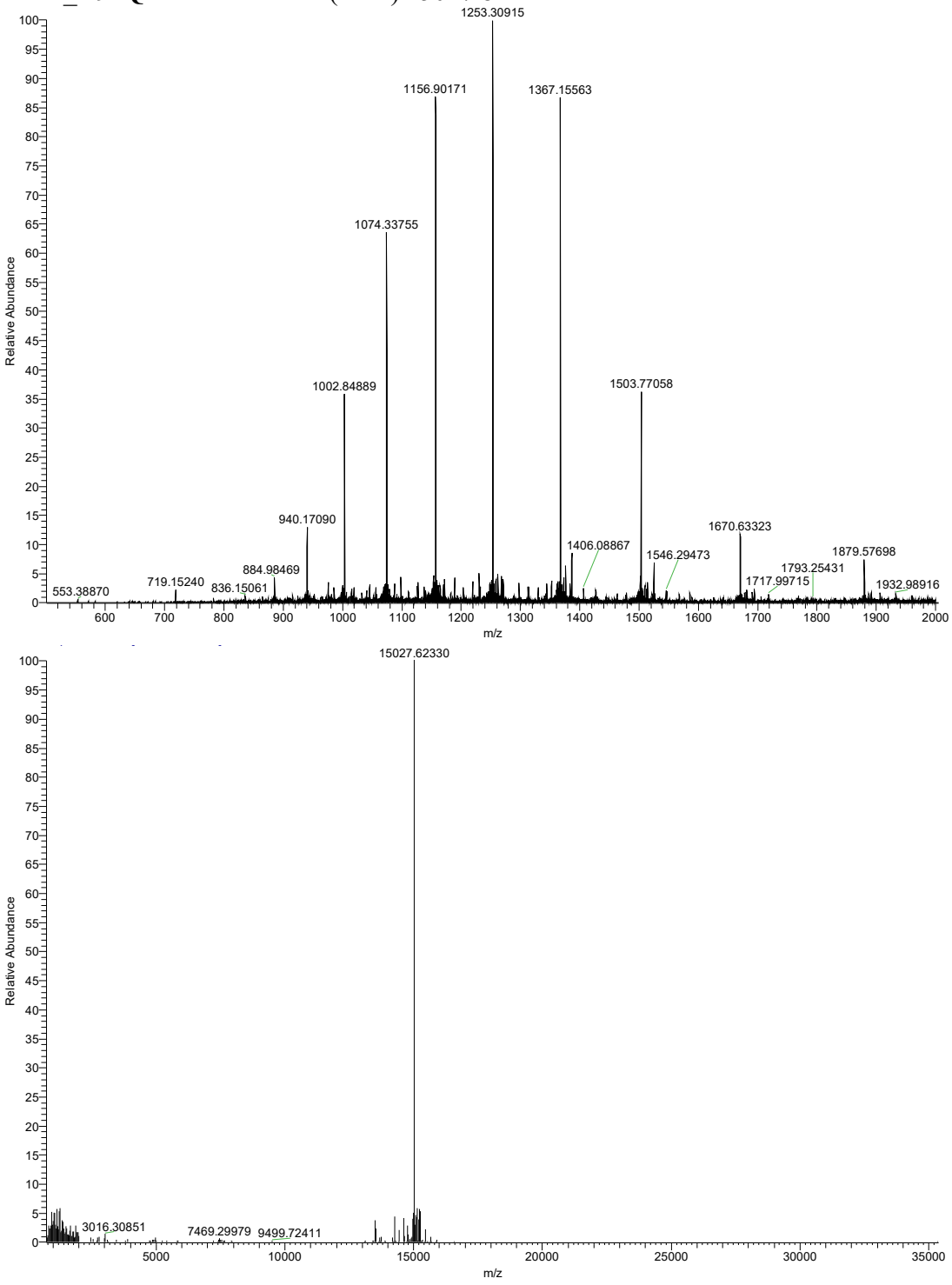


LmrR_**A92E** calculated mass (-Met) 15028.84





LmrR_**A92Q** calculated mass (-Met) 15027.81



II. Crystallization and structure determination

Recombinant wild-type LmrR (with the K55 and K59 residues reintroduced) carrying a C-terminal strep-tag was expressed and purified following previously published protocols.[2] Prior to crystallization the protein was loaded on a size exclusion column (Superdex 200 10/300 GL, GE Healthcare) and eluted in 20 mM MOPS buffer, pH 7.0, 150 mM NaCl. The protein was concentrated to 15 mg/ml in elution buffer and either frozen at 193 K for storage or used immediately for crystallization. Screening for initial crystallization conditions was carried out in 96-well sitting drop plates using the vapour-diffusion method, with the help of a Mosquito crystallization robot (TTP LabTech) and applying various commercially available crystallization screens. Cu-phenanthroline (pre-dissolved in H_2O/DMSO) was added to the reservoir solutions at a final concentration of 5 mM prior to drop mixing. Crystal optimization was performed manually using the hanging-drop vapour diffusion method. Well diffracting crystals grew at 295 K in drops prepared by mixing 1.0 µl protein solution (15 mg/ml) with 1.0 µl well solution containing 0.1 M sodium cacodylate buffer, pH 6.5, 0.3 M sodium acetate, 25% PEG 2000 MME and 5 mM Cu(II)-phen. X-ray diffraction data to 2.09 Å resolution were collected at 100 K from a single flash-cooled crystal on beamline ID30A at the European Synchrotron Radiation Facility (ESRF), Grenoble. The cryo-protecting solution was prepared from the crystallization solution by adding 25% PEG 400. Diffraction data were processed with the program XDS[3] and scaled using AIMLESS[4] from the CCP4 software suite.[5] Molecular replacement was performed with PHASER,[6] using PDB entry 3F8C as a search model. Manual rebuilding of the structure using COOT[7] was alternated with refinement using REFMAC5[8] until convergence was reached. The final rounds of refinement were performed with Phenix.refine.[9] The Cu(II)-phen complex was included in the refinement using appropriate restraints. The Molprobity server[10] was used for structure validation and PyMOL (version 1.8, Schrödinger, LLC) for structure visualization. A summary of the X-ray data collection and refinement statistics are shown in Table S2.

Description and quality of the Cu-phenanthroline bound LmrR crystal structure

LmrR crystallized in a tetragonal crystal form (space group $P4_32_12$) with one polypeptide chain occupying the asymmetric unit; the functional dimer being formed by a crystallographic dyad. In the electron density maps, the polypeptide chain is well defined, except for the tip region of the β-wing (residues 70–73) and the N- and C-termini (residues 1–4 and 109–126, including the C-terminal strep-tag). These regions show a high degree of disorder and were therefore excluded from the final model. Extra density in the central drug-binding pore confirmed the presence of the metal-phenanthroline complex. The overall binding mode of Cu(II)-phen is similar to those of drug-like ligands in previously determined LmrR structures.[2,11] Considerable disorder is observed in the binding mode of Cu(II)-phen, as evidenced by its relatively weak associated electron density and high atomic B-factors. Unfortunately, the weak electron density around the copper, and its special position in the crystal on a crystallographic dyad, prohibited an unambiguous identification of its coordination geometry and ligands other than phenanthroline. Putatively, the copper shows a distorted tetragonal coordination geometry with two bound waters molecules, which in turn are

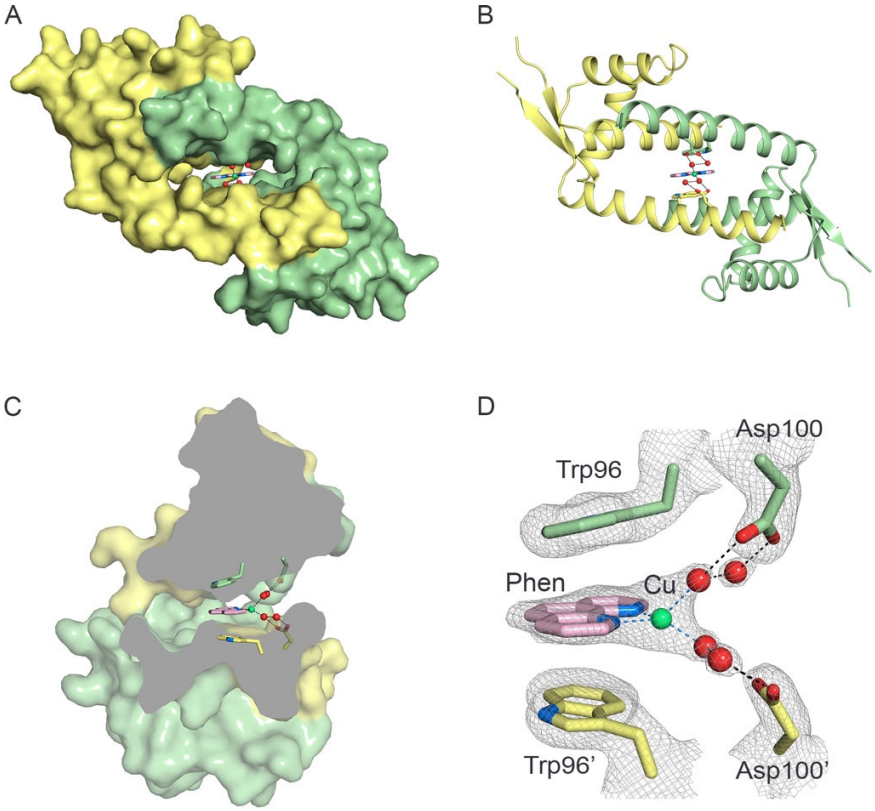
hydrogen bonded to an additional water molecule and to the carboxylate group of the Asp100 side chain (and of its dimer-related equivalent). Protein residues other than Trp96 and Asp100 that surround the copper within a distance of 8 Å are predominantly hydrophobic, i.e., Val15, Ala92, Ser97, Val99 and Ile103 (and their equivalents from the dimer mate).

Table S2: Summary of X-ray crystallographic statistics

	LmrR⊂Cu(II)-Phen
Data collection	
Beamline	ESRF/ID30A
Wavelength (Å)	0.9660
Space group	$P4_32_12$
Unit cell a, b, c (Å)	34.8, 34.8, 177.5
Resolution (Å)	35 – 2.09 (2.16-2.09)
Total observations	65777 (5249)
Unique reflections	7167 (556)
$\langle I/\sigma \rangle$	14.8 (2.5)
$CC_{1/2}$	0.998 (0.811)
Completeness (%)	99.7 (98.3)
R_{merge} (%)	8.8 (89.1)
Refinement	
R_{work}/R_{free} (%)	20.1/24.8
Nr non-H atoms /avg B (Å^2)	
protein (1 chain in AU)	824/45.0
waters	22/43.1
Cu-phenanthroline	15/70.0
Average B (Å^2)	46.2
RMSD	
bond lengths (Å)/angles (°)	0.005/0.7
Ramachandran plot	
% allowed/outliers	100/0.0
Clashscore score	1.8
PDB entry	6R1L

Values in parentheses are for the highest resolution shell.

Figure S4: Crystallographic analysis of Cu(II)-phen binding to LmrR. A) Overall structure of the LmrR dimer ⊂Cu(II)-phen in surface representation showing the bound Cu-phen as sticks. B) Overall structure of the LmrR•Cu-phen dimer in cartoon representation. C) Cut-through side view of the LmrR dimer showing the bound Cu(II)-phen ligand in the central pore. D) 2Fo-Fc electron density for the Cu(II)-phen ligand and interacting residues calculated at 2.09 Å resolution and contoured at 1σ. The putative coordination bonds of the copper ion to two water molecules (red spheres), in addition to the those with the phenanthroline amines, are shown as blue dashed lines. Hydrogen bonds between the waters coordinating the copper and the carboxylate group of Asp100 and Asp100' are shown as black dashed lines.

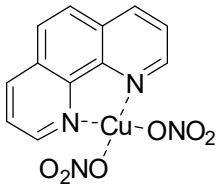


III. Synthetic procedures

Unless otherwise noted, all chemicals and reagents were purchased from Sigma Aldrich or Acros and used without further purification. Flash chromatography was performed using silica gel 60 Å (Merck, 230-400 mesh). Thin-layer chromatography was performed on silica plates. Compounds were visualized by UV light (254 nm), $KMnO_4$ and/or vanillin staining. ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Varian 400 (400 and 100MHz) in $CDCl_3$. Chemical shifts values (δ) are denoted in ppm using residual solvent peaks as the internal standard ($CHCl_3$: δ 7.26 for ^{1}H, δ 77.0 for ^{13}C). ^{1}H NMR data are reported as follows: chemical shifts, multiplicity (br = broad, d = doublet, dd = doublet of doublet, ddd = doublet of doublet of doublets, dq = doublet of quartets, m = multiplet, q = quartet, s = singlet, t = triplet,), coupling constants (Hz), and integration. Mass spectra (HRMS) were recorded on an Orbitrap XL (Thermo Fisher Scientific; ESI pos. mode). Enantiomeric excess determinations were performed by np-HPLC analysis using UV-detection (Shimadzu SCL-10Avp). Fluorescence experiments were recorded using a JASCO FP-6200 spectrometer.

Copper complexes

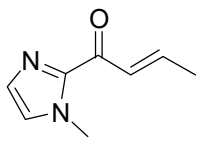
Cu(II)-phen was prepared by following published procedure.[12–14] Elemental analysis data match those reported in the literature.



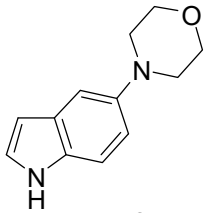
Cu(II)-phen

Substrates

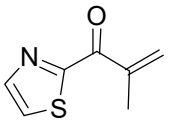
Compounds 1-(1-methyl-1*H*-imidazol-2-yl)but-2-en-1-one (**1**),[15] 5-morpholin-4-yl-1*H*-indole (**2f**)[16] and 2-methyl-1-(thiazol-2-yl)prop-2-en-1-one (**4**)[17] were synthesized according to literature procedures. ^{1}H-NMR and ^{13}C-NMR data matched those reported in the literature.



1



2f

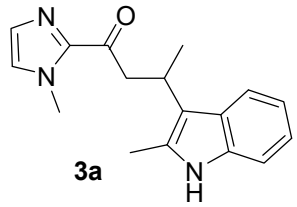


4

Synthesis and characterization of reaction products

Friedel-Crafts product (**3**)[14]

Procedure was adapted from literature. The NMR data match those reported in the literature.

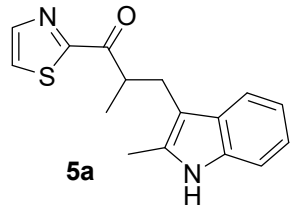


3a

1-(1-methyl-1H-imidazol-2-yl)-3-(2-methyl-1H-indol-3-yl)butan-1-one (3a). ^{1}H-NMR (CD_3OD, 400 MHz) δ = 7.48 (d, J = 7.9 Hz, 1H), 7.20 (s, 1H), 7.14 (d, J = 9.2 Hz, 1H), 7.05 (s, 1H), 6.94 (t, J = 7.8 Hz, 1H), 6.88 (t, J = 7.0 Hz, 1H), 3.70 – 3.77 (m, 1H), 3.73 (s, 3H), 3.41 – 3.53 (m, 2H), 2.29 (s, 3H), 1.44 (d, J = 7.1 Hz, 3H). ^{13}C-NMR ($CDCl_3$, 101 MHz) δ = 192.3, 143.3, 135.3, 130.3, 128.8, 127.4, 126.5, 120.5, 119.2, 118.9, 115.4, 110.1, 46.2, 36.0, 27.0, 21.1, 12.3. HRMS (ESI) calcd for $C_{17}H_{20}N_3O$: 282.1601, found 282.1600. Ee's were determined by HPLC analysis (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 90:10, 1 mL/min). Retention times: 16 and 21 min. (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 88:12, 0.5 mL/min). Retention times: 26 and 36 min.

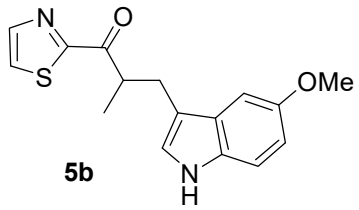
Tandem Friedel-Crafts/Enantioselective protonation products (**5**)[16]

Procedure was adapted from literature. The NMR data match those reported in the literature.



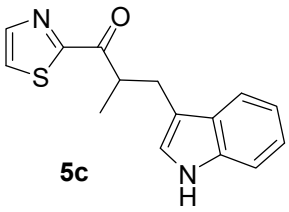
5a

2-methyl-3-(2-methyl-1*H*-indol-3-yl)1-(thiazol-2-yl)propan-1-one (5a). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 8.00 (d, *J* = 3.0 Hz, 1H), 7.78 (br, 1H), 7.70 (dd, *J* = 5.9, 3.1 Hz, 1H), 7.61 (d, *J* = 3.0 Hz, 1H), 7.25 – 7.20 (m, 1H), 7.16 – 7.02 (m, 2H), 4.31 – 4.16 (m, 1H), 3.32 (dd, *J* = 14.2, 5.7 Hz, 1H), 2.82 (dd, *J* = 14.2, 8.8 Hz, 1H), 2.39 (s, 3H), 1.26 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz,) δ = 197.6, 166.9, 144.7, 135.2, 132.1, 128.8, 126.1, 120.9, 119.2, 118.4, 110.1, 109.3, 42.8, 28.0, 16.29, 11.8. HRMS (ESI) calcd for $C_{16}H_{17}N_2OS$ 285.106; found 258.106. Ee's were determined by HPLC analysis (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 95:5, 1 mL/min). Retention times: 16 and 19 min. (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 94:6, 0.5 mL/min). Retention times: 28 and 32 min.



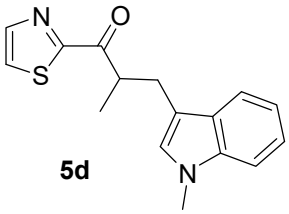
5b

5-(5-methoxy-1*H*-indol-3-yl)-2-methyl-1-(thiazol-2-yl) propan-1-one (5b). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 8.02 (br, 1H), 8.00 (d, *J* = 3.0 Hz, 1H), 7.63 (d, *J* = 3.0 Hz, 1H), 7.27 (d, *J* = 2.3 Hz, 1H), 7.20 (d, *J* = 8.7 Hz, 1H), 6.97 (d, *J* = 2.1 Hz, 1H), 6.85 (dd, *J* = 8.8, 2.4 Hz, 1H), 4.25 (dt, *J* = 7.0, 6.5 Hz, 1H), 3.91 (s, 3H), 3.38 (dd, *J* = 14.4, 5.9 Hz, 1H), 2.85 (dd, *J* = 14.4, 8.1 Hz, 1H), 1.30 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ = 197.4, 167.0, 153.9, 144.66, 131.4, 127.9, 126.3, 123.4, 113.4, 112.2, 111.8, 101.0, 55.9, 42.4, 29.0, 16.5. HRMS (ESI) calcd for $C_{16}H_{17}N_2O_2S$ 301.100; found 301.101. Ee's were determined by HPLC analysis (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 95:5, 1 mL/min). Retention times: 33 and 36 min. (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 93:7, 0.5 mL/min). Retention times: 49 and 54 min.



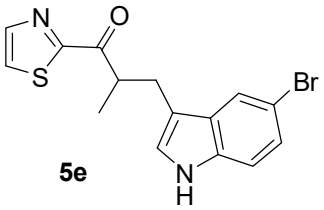
5c

3-(1*H*-indol-3-yl)-2-methyl-1-(thiazol-2-yl)propan-1-one (5c). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 8.05 (br, 1H), 8.01 (d, *J* = 3.0 Hz, 1H), 7.75 (d, *J* = 7.7 Hz, 1H), 7.62 (d, *J* = 3.0 Hz, 1H), 7.32 (d, *J* = 7.8 Hz, 1H), 7.21 – 7.13 (m, 2H), 6.99 (d, *J* = 1.9 Hz, 1H), 4.34 – 4.18 (m, 1H), 3.41 (dd, *J* = 14.4, 6.3 Hz, 1H), 2.91 (dd, *J* = 14.4, 7.9 Hz, 1H), 1.31 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ = 197.4, 166.9, 144.7, 136.2, 127.6, 126.2, 122.6, 121.9, 119.3, 119.1, 113.7, 111.1, 42.5, 28.8, 16.8. HRMS (ESI) calcd for $C_{15}H_{15}N_2OS$ 271.090; found 271.090. Ee's were determined by HPLC analysis (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 95:5, 1 mL/min). Retention times: 23 and 26 min. (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 93:7, 0.5 mL/min). Retention times: 36 and 42 min.



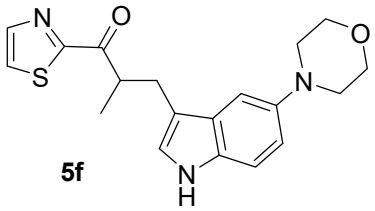
5d

2-methyl-3-(1-methyl-1*H*-indol-3-yl)-1-(thiazol-2-yl)propan-1-one (5d). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 8.02 (d, *J* = 3.0 Hz, 1H), 7.73 (d, *J* = 7.9 Hz, 1H), 7.64 (d, *J* = 3.0 Hz, 1H), 7.28 – 7.20 (m, 2H), 7.17 – 7.08 (m, 1H), 6.88 (s, 1H), 4.30 – 4.15 (m, 1H), 3.72 (s, 3H), 3.39 (dd, *J* = 14.4, 6.2 Hz, 1H), 2.89 (dd, *J* = 14.4, 7.9 Hz, 1H), 1.30 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ = 197.4, 167.0, 144.7, 136.9, 128.0, 127.4, 126.1, 121.4, 119.2, 118.8, 112.2, 109.1, 42.7, 32.6, 28.6, 16.8. HRMS (ESI) calcd for $C_{16}H_{17}N_2OS$ 285.106; found 258.106. Ee's were determined by HPLC analysis (Chiralcel-OD, ^{n}Heptane/^{i}PrOH 98:2, 1 mL/min). Retention times: 17 and 19 min.



5e

3-(5-bromo-1*H*-indol-3-yl)-2-methyl-1-(thiazol-2-yl)propan-1-one (5e). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 7.97 (d, *J* = 3.0 Hz, 1H), 7.91 (br, 1H), 7.80 (d, *J* = 1.7 Hz, 1H), 7.59 (d, *J* = 3.0 Hz, 1H), 7.17 (d, *J* = 1.8 Hz, 1H), 7.12 (d, *J* = 8.6 Hz, 1H), 6.95 (s, 1H), 4.18 – 4.06 (m, 1H), 3.26 (dd, *J* = 14.4, 6.4 Hz, 1H), 2.78 (dd, *J* = 14.5, 7.7 Hz, 1H), 1.22 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ = 197.0, 166.8, 144.7, 134.8, 129.4, 126.2, 124.8, 123.6, 121.9, 113.6, 112.7, 112.4, 42.4, 28.6, 16.7. HRMS (ESI) calcd. for $C_{15}H_{13}BrN_2OSNa^{+}$ 370.982; found 370.982. Ee's were determined by HPLC analysis (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 95:5, 0.5 mL/min). Retention times: 46 and 49 min.



5f

2-methyl-3-(5-morpholino-1*H*-indol-3-yl)-1-(thiazol-2-yl)propan-1-one (5f). ^{1}H NMR ($CDCl_3$, 400 MHz) δ = 7.99 (br, 1H), 7.90 (d, *J* = 3.0 Hz, 1H), 7.55 (d, *J* = 3.0 Hz, 1H), 7.24 (d, *J* = 2.0 Hz, 1H), 7.13 (d, *J* = 8.8 Hz, 1H), 6.91 – 6.80 (m, 2H), 4.20 – 4.10 (m, 1H), 3.90 – 3.80 (m, 4H), 3.28 (dd, *J* = 14.3, 5.8 Hz, 1H), 3.13 – 3.06 (m, 4H), 2.73 (dd, *J* = 14.4, 8.2 Hz, 1H), 1.20 (d, *J* = 6.9 Hz, 3H). ^{13}C NMR ($CDCl_3$, 101 MHz) δ = 197.4, 167.0, 145.5, 144.6, 131.9, 129.0, 126.3, 123.2, 115.0, 113.4, 111.6, 106.1, 67.3, 52.1, 42.3, 29.1, 16.5. HRMS (ESI) calcd for $C_{19}H_{22}N_3O_2S$ 356.143; found 356.140. Ee's were determined by HPLC analysis (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 87:13, 0.5 mL/min). Retention times: 47 and 55 min.

IV. Catalysis

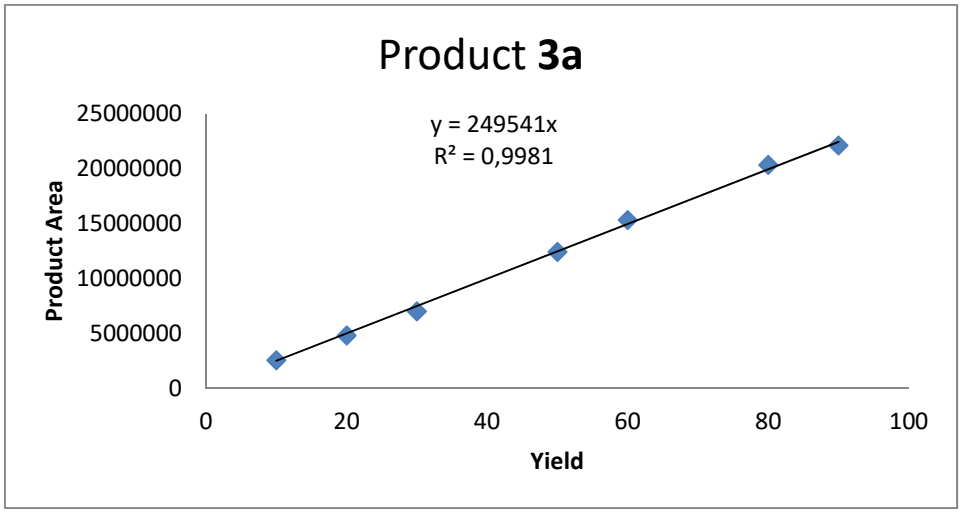
Representative procedure for catalytic FC and FC/EP reactions

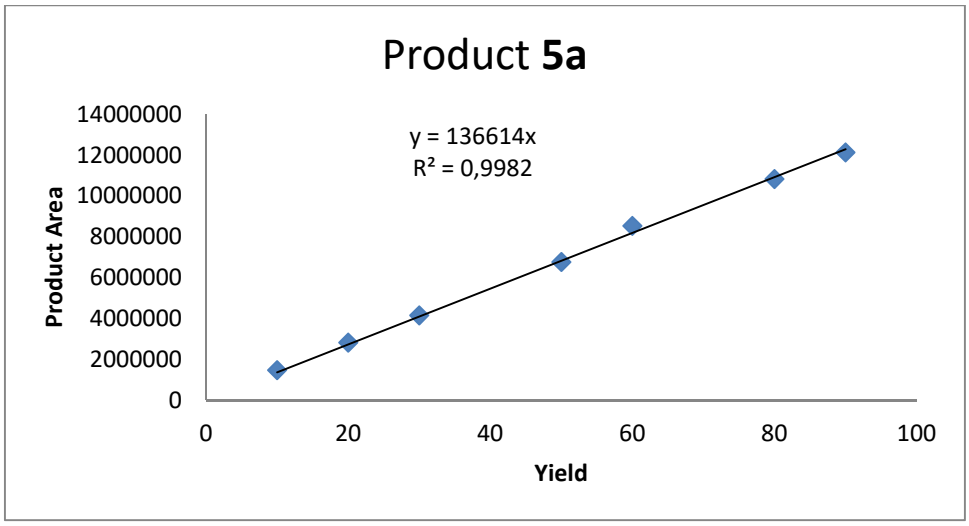
In a typical procedure, a solution containing 90 µM [Cu(phen)(NO_3)$_2$] complex (pre-dissolved in buffer/DMSO, final fraction of DMSO 0.7 v/v%) and 1.3 eq of LmrR (120 µM) in a final volume of 280 µL was mixed by continuous inversion at 4°C for 30 min. Afterwards, a fresh stock solution of enone and **2** in buffer/CH_3CN (final fraction of CH_3CN 1.3 v/v%) were added to the catalyst solution (10 µL each, 1mM final concentration). Reaction was mixed by continuous inversion at 4°C for 48h. Specific conditions are summarized in Table S3, S4, S5 and S6.

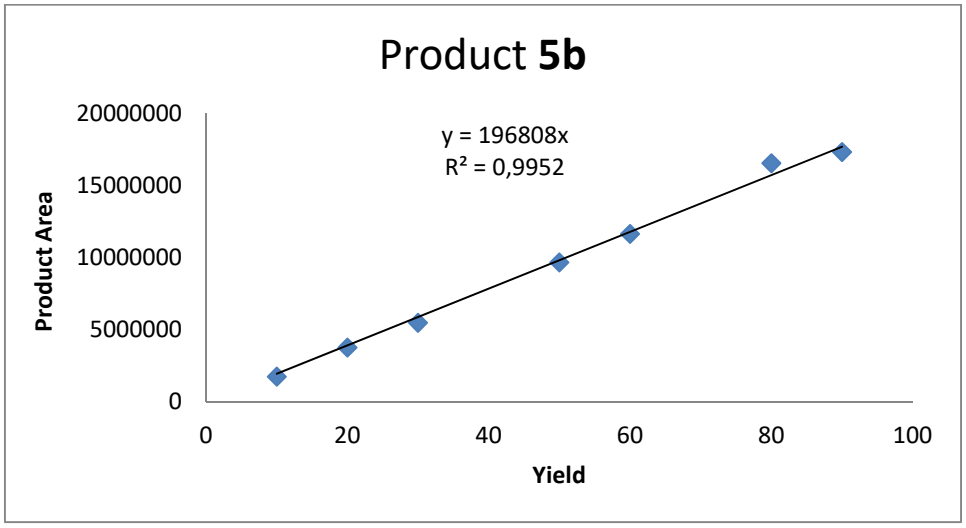
Product analysis

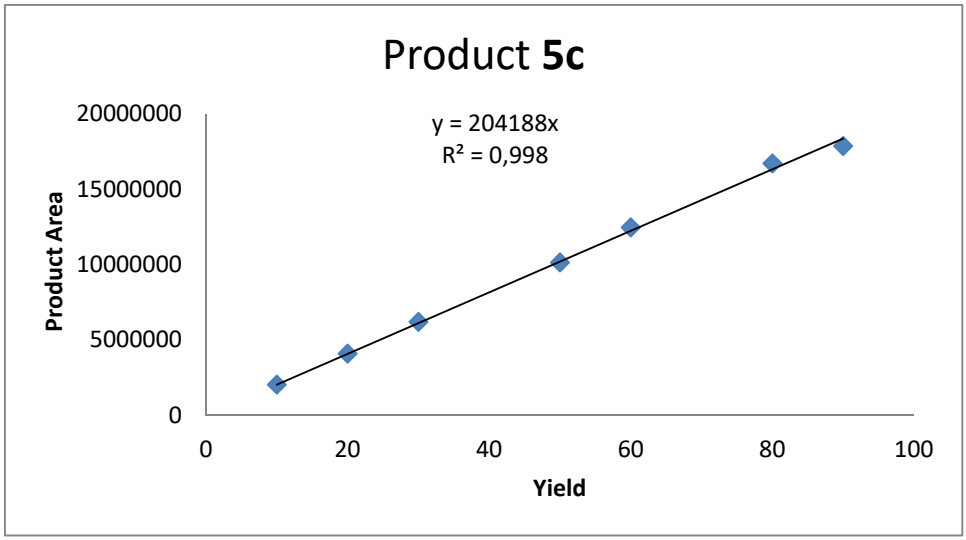
The reactions were analyzed by adding 100 µL of internal standard (2-phenylquinoline, 1mM in CH_3CN) to the reaction mixture, followed by extraction with diethyl ether (3 x 300µL) and separated organic layer were dried over Na_2SO_4 and concentrated under reduced pressure. Subsequently, the samples were redissolved in 180 µL of a heptane:propan-2-ol mixture (5:1) and the yields and enantiomeric excess were determined by HPLC (Figure S5 and S6). Calibration curves for quantification of the different products were constructed using pure reference compounds produced synthetically as described in Synthetic procedures.

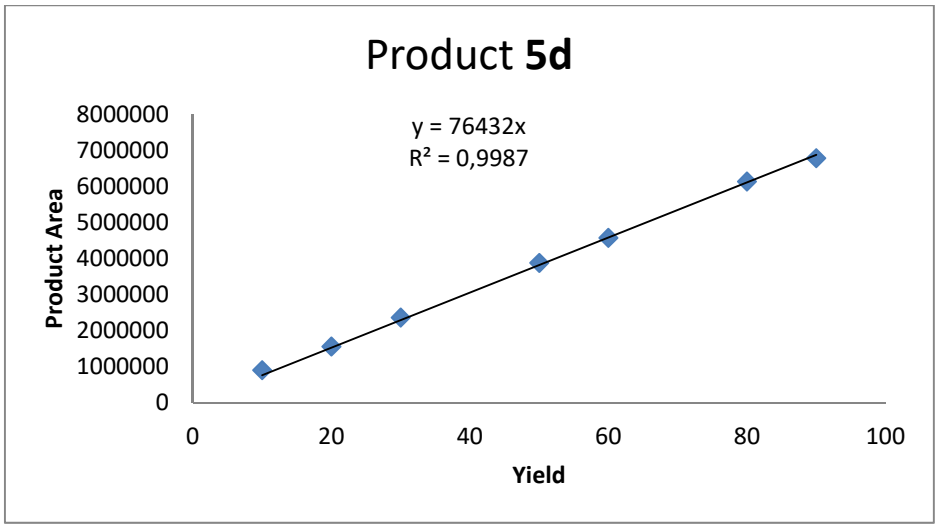
Figure S5. Calibration curves for products 3a, 5a-f.

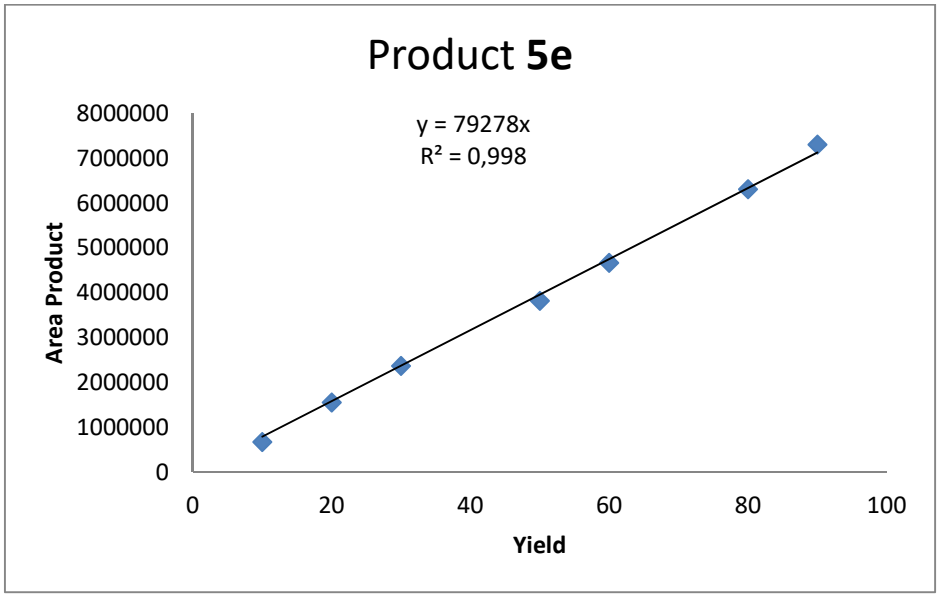



Product 5a
y = 136614x
R² = 0,9982
Product Area
14000000
12000000
10000000
8000000
6000000
4000000
2000000
0
0
20
40
60
80
100
Yield


Product 5b
y = 196808x
R² = 0,9952
Product Area
20000000
15000000
10000000
5000000
0
0
20
40
60
80
100
Yield


Product 5c
y = 204188x
R² = 0,998
Product Area
20000000
15000000
10000000
5000000
0
0
20
40
60
80
100
Yield


Product 5d
y = 76432x
R² = 0,9987
Product Area
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
0
20
40
60
80
100
Yield


Product 5e
y = 79278x
R² = 0,998
Area Product
8000000
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
0
20
40
60
80
100
Yield

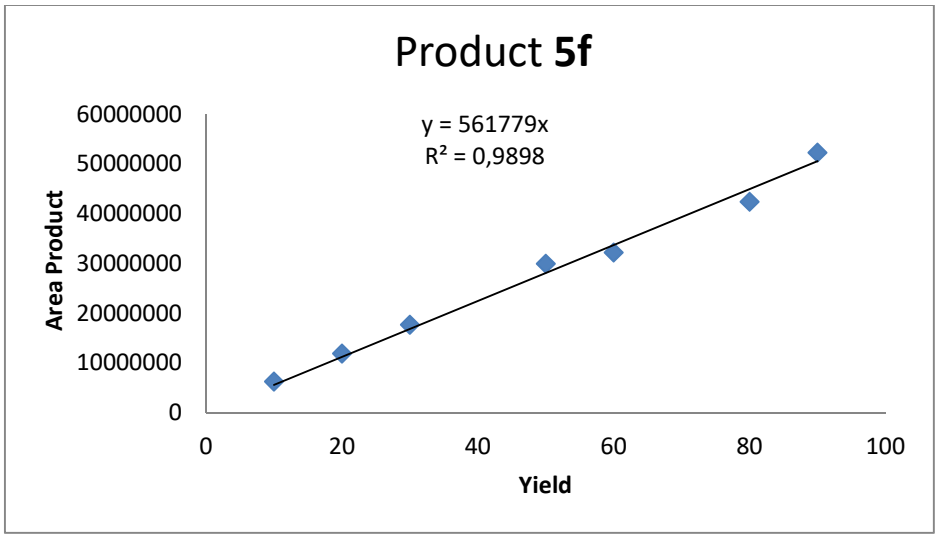
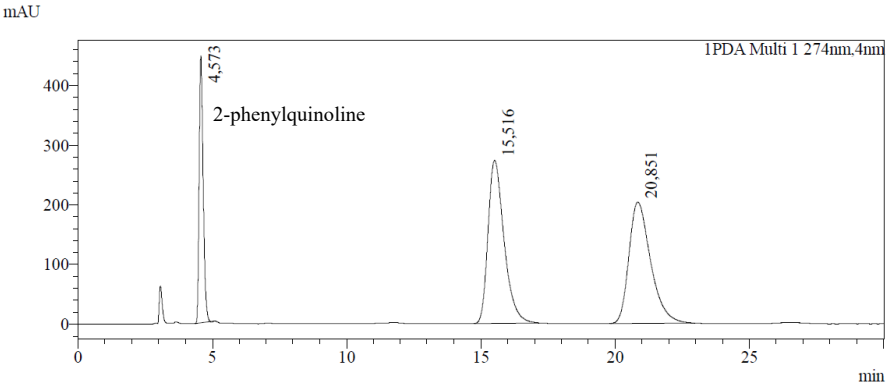

Product 5f
y = 561779x
R² = 0,9898
Area Product
60000000
50000000
40000000
30000000
20000000
10000000
0
0
20
40
60
80
100
Yield

Figure S6. Chiral HPLC traces for products 3a, 5a-f

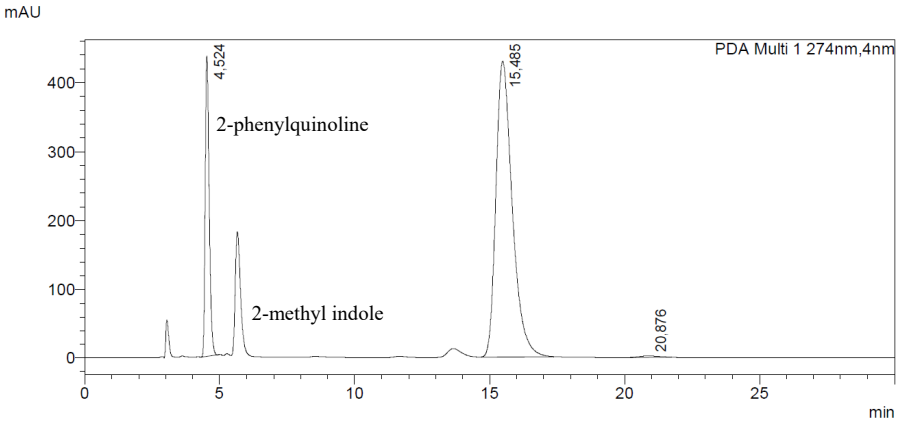
Product **3a** (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 90:10, 1 mL/min)



Peak Table

PDA Ch1 274nm

Peak#	Ret. Time	Area	Area%	Height	Conc.	Unit	Mark
1	4.573	4647917	16.927	448078	16.927		M
2	15.516	11378376	41.439	273553	41.439		M
3	20.851	11431823	41.634	203388	41.634		M
Total		27458116	100.000	925019			



<Peak Table>

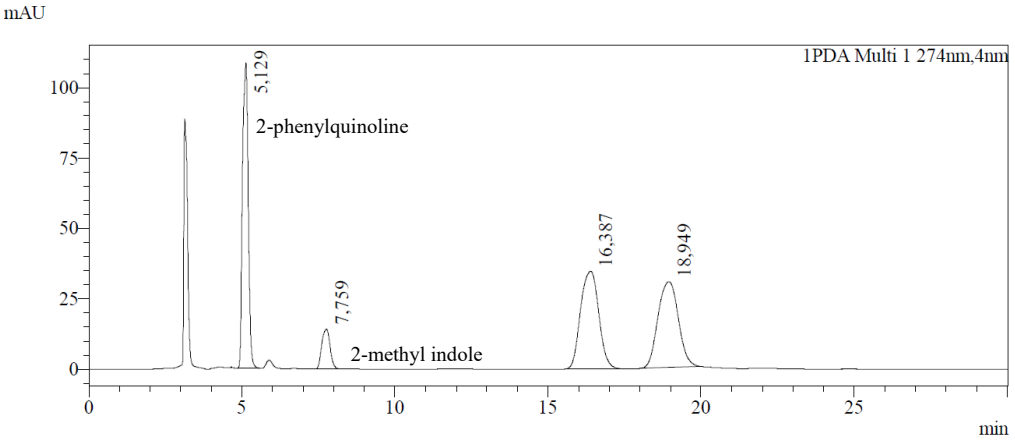
AD2

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
Total							

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	4,524	4466550	435326	19,659		M	
2	15,485	18151795	429468	79,894		M	
3	20,876	101532	2107	0,447		M	
Total		22719878	866901				

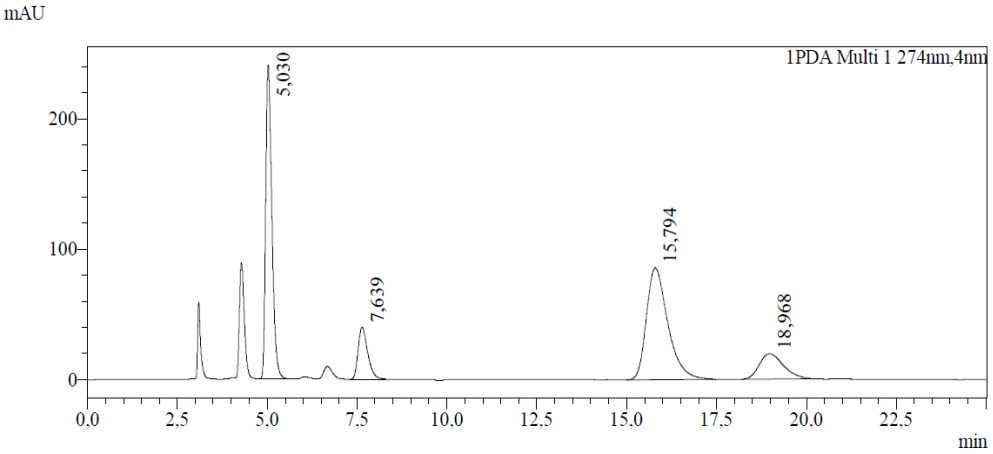
Product **5a** (Chiralcel-AD, ^{n}Heptane/^{i}PrOH 95:5, 1 mL/min)



Peak Table

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	5,129	1404174	108416	31,065		M	
2	7,759	269287	14084	5,958		M	
3	16,387	1426282	34626	31,554		M	
4	18,949	1420382	30499	31,424		M	
Total		4520125	187624				

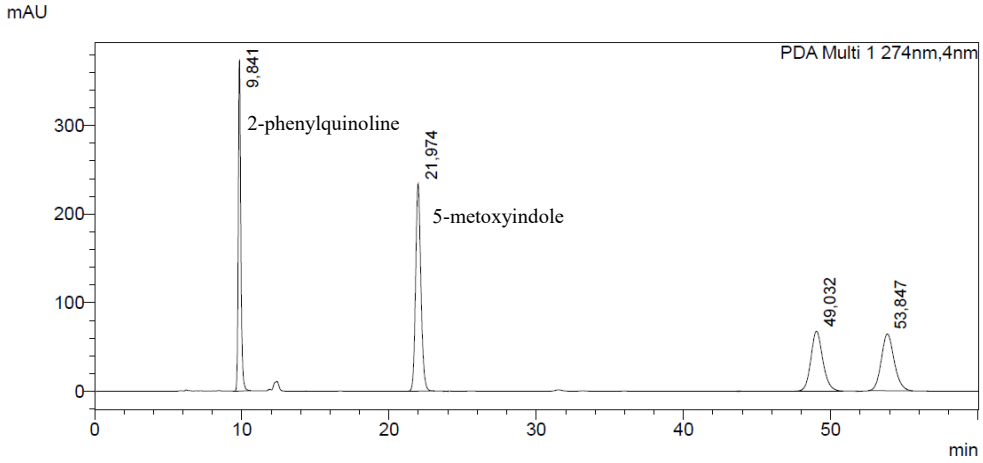


Peak Table

PDA Ch1 274nm

Peak#	Ret. Time	Area	Area%	Height	Conc.	Unit	Mark
1	5,030	2875356	35,681	240474	35,681		M
2	7,639	731958	9,083	39703	9,083		M
3	15,794	3539007	43,916	85432	43,916		M
4	18,968	912185	11,320	19510	11,320		M
Total		8058506	100,000	385118			

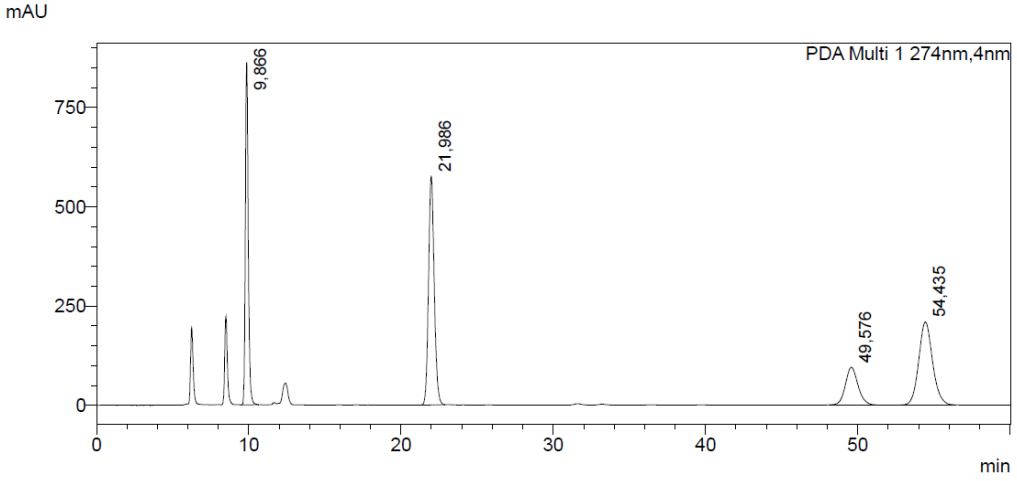
Product **5b** (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 93:7, 0.5 mL/min)



<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Area%
1	9,841	4475110	372034	24,833
2	21,974	5820211	233709	32,297
3	49,032	3821344	67595	21,205
4	53,847	3904101	64127	21,664
Total		18020767	737465	100,000

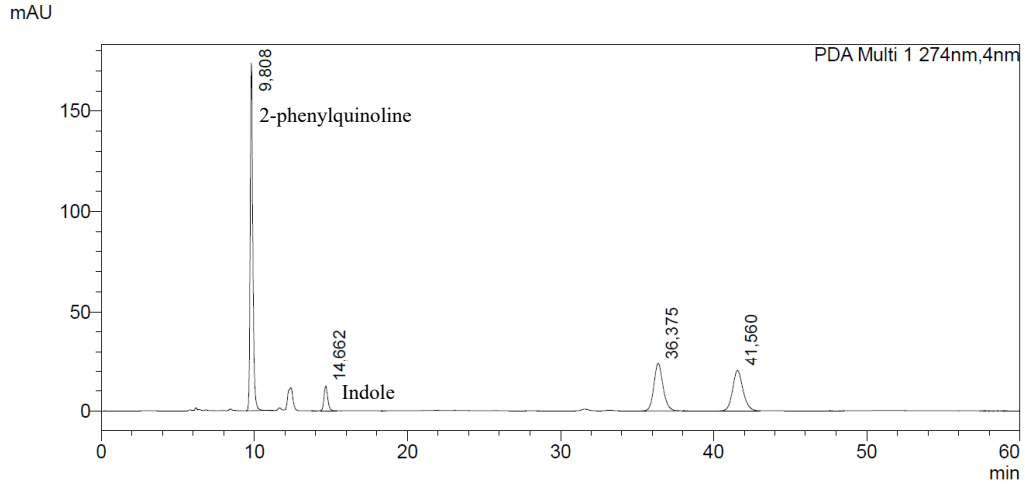


<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Area%
1	9,866	11421784	862732	25,357
2	21,986	14964804	576220	33,222
3	49,576	5445785	94808	12,090
4	54,435	13212063	209230	29,331
Total		45044436	1742990	100,000

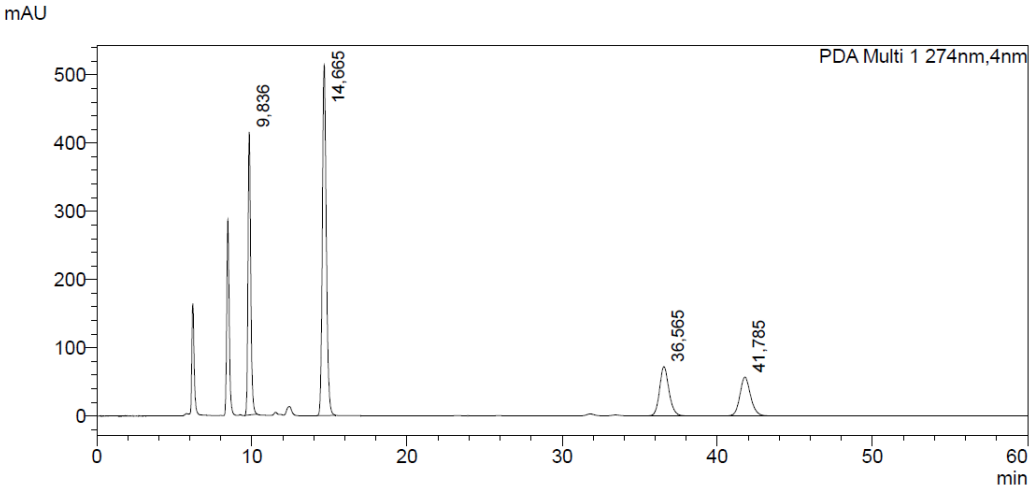
Product **5c** (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 93:7, 0.5 mL/min)



<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Area%
1	9,808	2077017	173386	49,232
2	14,662	208925	12423	4,952
3	36,375	985023	23732	23,348
4	41,560	947850	20126	22,467
Total		4218814	229667	100,000

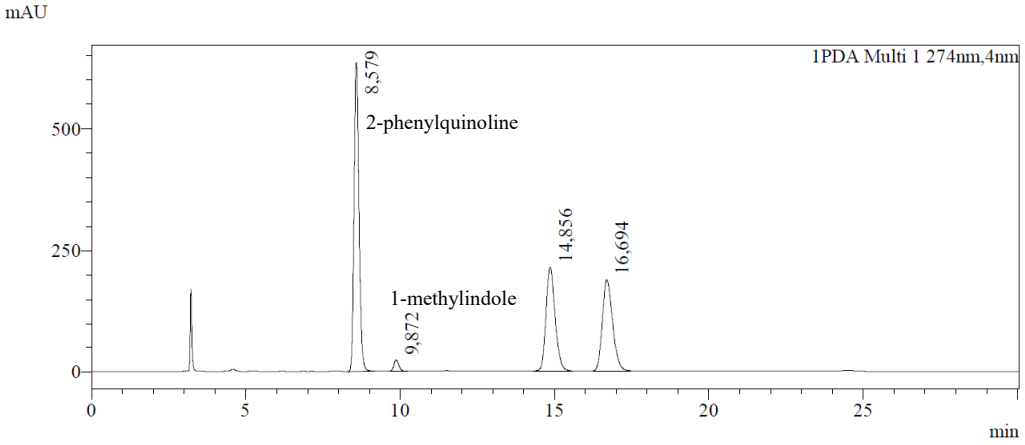


<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,836	5134440	414420	25,644		M	
2	14,665	9132807	513162	45,614		M	
3	36,565	3043356	72252	15,200		M	
4	41,785	2711385	56800	13,542		M	
Total		20021988	1056634				

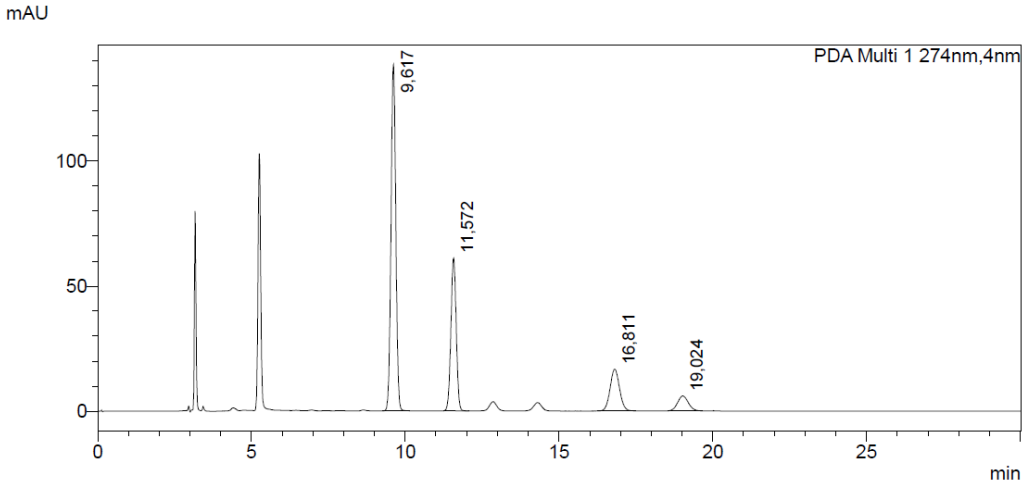
Product **5d** (Chiralcel-OD, ^{n}Heptane/^{i}PrOH 98:2, 1 mL/min)



Peak Table

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	8,579	7089371	633733	43,932		M	
2	9,872	258234	22953	1,600		M	
3	14,856	4417209	214944	27,373		M	
4	16,694	4372363	188533	27,095		M	
Total		16137177	1060163				

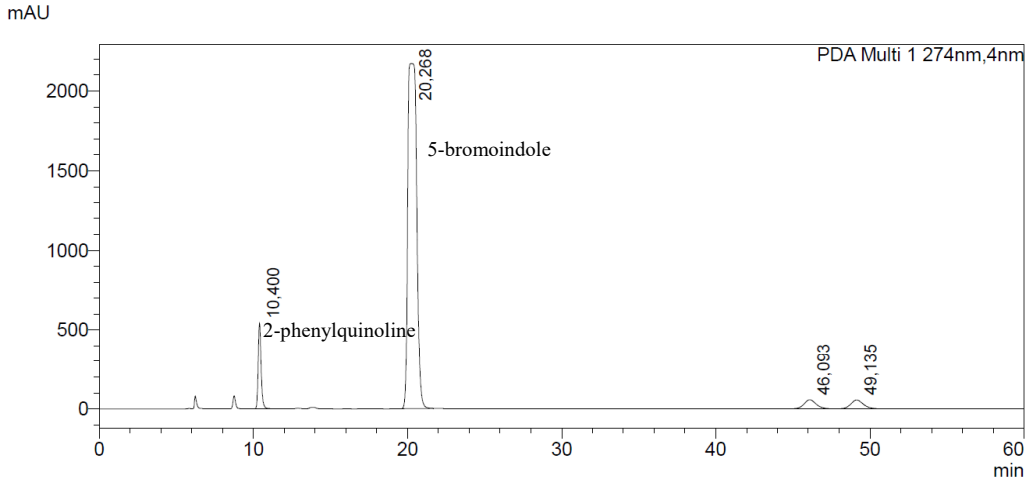


<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,617	1503881	138471	54,917		M	
2	11,572	732365	61178	26,744		M	
3	16,811	356761	16738	13,028		M	
4	19,024	145460	5964	5,312		M	
Total		2738467	222351				

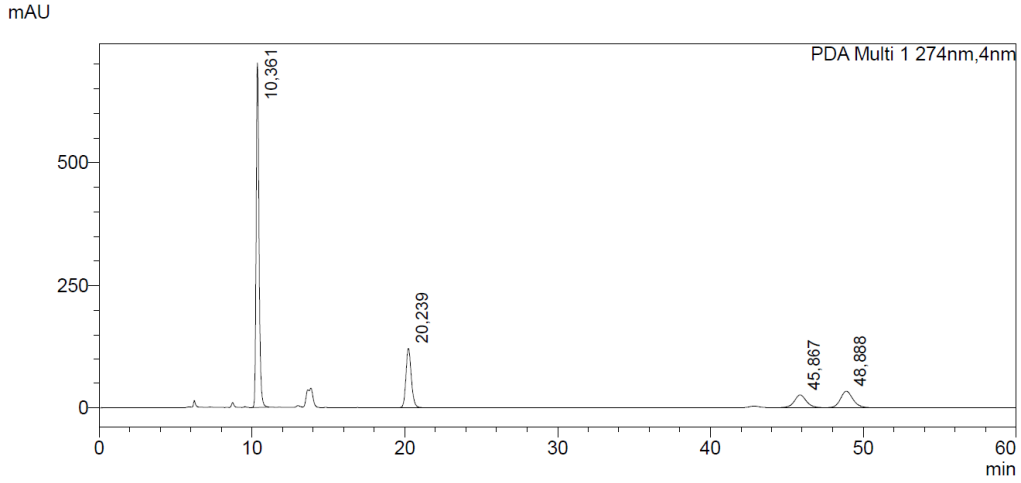
Product **5e** (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 95:5, 0.5 mL/min)



<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	10,400	7064353	539921	7,071		M	
2	20,268	86824508	2171872	86,906		M	
3	46,093	3002951	56210	3,006		M	
4	49,135	3014985	54273	3,018		M	
Total		99906798	2822276				

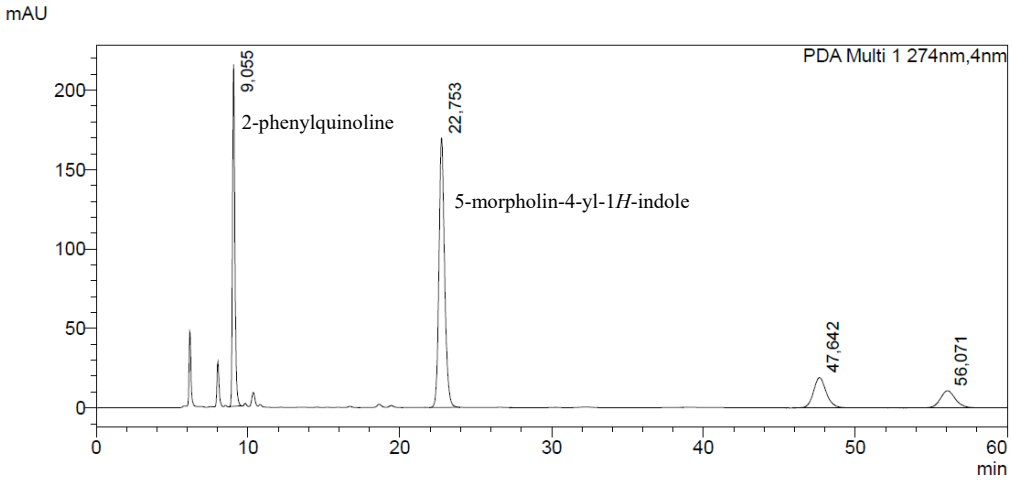


<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Area%
1	10,361	9098171	701781	59,236
2	20,239	2936753	122469	19,121
3	45,867	1419290	25925	9,241
4	48,888	1904865	33609	12,402
Total		15359079	883784	100,000

Product **5f** (Chiralcel-AD-H, ^{n}Heptane/^{i}PrOH 87:13, 0.5 mL/min).



<Peak Table>

PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	9,055	2329391	215163	27,179		M	
2	22,753	4404846	169395	51,395		M	
3	47,642	1108001	18991	12,928		M	
4	56,071	728320	10725	8,498		M	
Total		8570557	414275				

Table S3. Optimization and scope of the tandem Friedel-Crafts/Enantioselective Protonation (FC/EP reaction)

Entry	Indole	Product	4 : 2	yield (%)	*ee* (%)
1	**2b**	**5b**	1	58 ± 12	40 ± 2
2	**2b**	**5b**	1.5	56 ± 21	38 ± 4
3	**2b**	**5b**	0.5	28 ± 12	54 ± 3
4	**2a**	**5a**	1	87 ± 14	59 ± 3
5	**2c**	**5c**	1	30 ± 11	8 ± 6
6	**2d**	**5d**	1	17 ± 6	35 ± 6
7	**2e**	**5e**	1	24 ± 10	14 ± 7
8	**2f**	**5f**	1	17 ± 5	23 ± 5

Typical reaction conditions: **4** (1mM), [Cu(II)-phen] (9 mol%; 90 μM) , LmrR (12 mol%; 120 μM) in 20 mM MES buffer (pH 5.0), 150 mM NaCl, at 4°C for 72h; Results are the average of at least two independent experiments, both carried out in duplicate.

Table S4. Control experiments of the tandem Friedel-Crafts/Enantioselective Protonation (FC/EP reaction)

Entry	Catalyst (μM)	yield of 5 (%)	*ee* (%)
1	$Cu(NO_3)_2.xH_2O$ (90 μM)	< 5	-
2	Cu(II)-phen (90 μM)	< 5	-
3	$Cu(NO_3)_2.xH_2O$ (90 μM) + LmrR (120 μM)	86 ± 10	racemic

Typical reaction conditions: **4** (1mM), **2b** (1mM) in 20 mM MES buffer (pH 5.0), 150 mM NaCl, at 4°C for 72h; Results are the average of at least two independent experiments, both carried out in duplicate.

Table S5. Results of the vinylogous Friedel-Crafts alkylation (FC reaction) catalyzed by LmrR and LmrR mutants

Entry	LmrR_X	yield (%)	*ee* (%)
1	wild type	41 ± 12	92 ± 2
2	W96A	21 ± 11	25 ± 9
3	D100A	15 ± 1	46 ± 4
4	D100E	48 ± 21	28 ± 6
5	F93A	30 ± 5	51 ± 5
6	A92E	61 ± 13	> 99
7	A92Q	55 ± 15	94 ± 1
8	Q12A	45 ± 18	98 ± 1
9	Q12E	68 ± 14	92 ± 1
10	V99A	55 ± 11	95 ± 1
11	M8A	43 ± 13	99 ± 0
12	V15A	59 ± 9	94 ± 1
13	E7A	72 ± 15	96 ± 1

Conditions: **1** (1mM), **2a** (1mM), [Cu(II)-phen] (9 mol%; 90 μM) , LmrR_X (12 mol%; 120 μM) in 20 mM MOPS buffer (pH 7.0), 150 mM NaCl, at 4°C for 48h; Results are the average of at least two independent experiments, both carried out in duplicate.

Table S6. Results of the tandem Friedel-Crafts/Enantioselective Protonation (FC/EP) reaction catalyzed by LmrR/Cu(II)-phen and effect of mutations on catalysis.

Entry	LmrR_X	yield (%)	*ee* (%)
1	*wild type*	70 ± 16	58 ± 2
2[b]	*wild type*	53 ± 16	54 ± 5
3[b]	W96A	32 ± 8	64 ± 6
4	*D100A*	67 ± 9	63 ± 1
5	*D100E*	78 ± 10	54 ± 1
6	*F93A*	48 ± 21	66 ± 2
7	*A92E*	61 ± 6	21 ± 7
8	*A92Q*	75 ± 10	51 ± 1
9	*Q12A*	49 ± 8	44 ± 6
10	*Q12E*	64 ± 7	21 ± 8
11	*V99A*	58 ± 7	60 ± 4
12	*M8A*	49 ± 7	45 ± 1
13	*V15A*	70 ± 7	42 ± 1
14	*E7A*	60 ± 7	56 ± 1

(a) Conditions: **4** (1mM), **2a** (1mM), [Cu(II)-phen] (9 mol%; 90 μM) , LmrR_X (12 mol%; 120 μM) in 20 mM MES buffer (pH 5.0), 150 mM NaCl, at 4°C for 48h; Results are the average of at least two independent experiments, both carried out in duplicate. (b) 20 mM MOPS buffer (pH 7.0), 150 mM NaCl.

V. Competition experiments with Hoechst 33342

Representative procedure for catalytic FC and FC/EP reactions competition experiments with Hoechst 33342

In a typical procedure, a solution containing 120 μM of LmrR (1.3 eq) and 0, 1, 2 or 4 eq of Hoechst 33342 (0 μM, 90 μM, 180 μM and 360 μM , respectively) in a final volume of 270 μL was mixed by continuous inversion at 4°C for 30 min. Afterwards, a solution containing 90 μM [Cu(phen)(NO_3)$_2$] complex (pre-dissolved in buffer/DMSO, final fraction of DMSO 0.7 v/v%) was added to the solution (10 μL) and mixed by continuous inversion at 4°C for 30 min. Finally, a fresh stock solution of enone and **2a** in buffer/CH_3CN (final fraction of CH_3CN 1.3 v/v%) were added to the catalyst solution (10 μL each, 1mM final concentration). Reaction was mixed by continuous inversion at 4°C for 48h.

Table S7. Results of the Friedel-Crafts and tandem Friedel-Crafts/Enantioselective protonation reaction catalyzed by LmrR/Cu(II)-phen in the presence of increasing amounts of Hoechst 33342.[a]

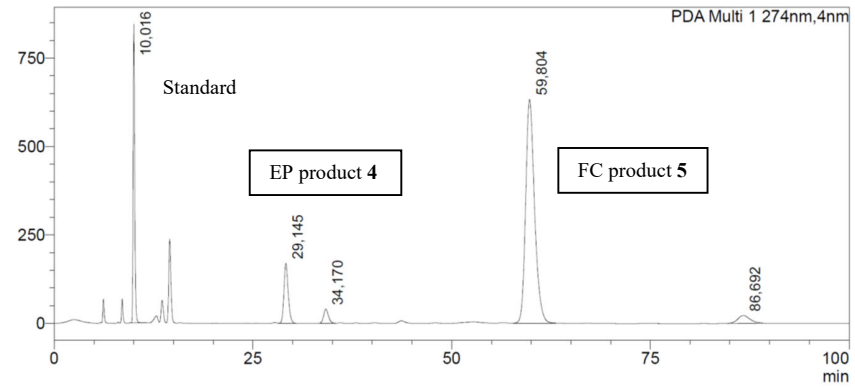
Entry	Hoechst 33342 (eq)	FC		FC/EP [b]	
		yield(%)	*ee* (%)	yield(%)	*ee* (%)
1	0	41 ± 12	92 ± 2	70 ± 16	58 ± 2
2	1	57 ± 6	94 ± 1	68 ± 22	61 ± 2
3	2	30 ± 8	72 ± 5	49 ± 13	53 ± 2
4	4	29 ±18	57 ± 5	60 ± 6	61 ± 1

(a) Conditions: **1** or **4** (1mM), **2a** (1mM), [Cu(II)-phen] (9 mol%; 90 μM) , LmrR (12 mol%; 120 μM), H33342 (0, 1, 2 or 4 equivalents with respect to the concentration of Cu(II)-phen) in 20 mM MOPS buffer (pH 7.0), 150 mM NaCl, at 4°C for 48h; Results are the average of at least two independent experiments, both carried out in duplicate. (b) 20 mM MES buffer (pH 5.0), 150 mM NaCl.

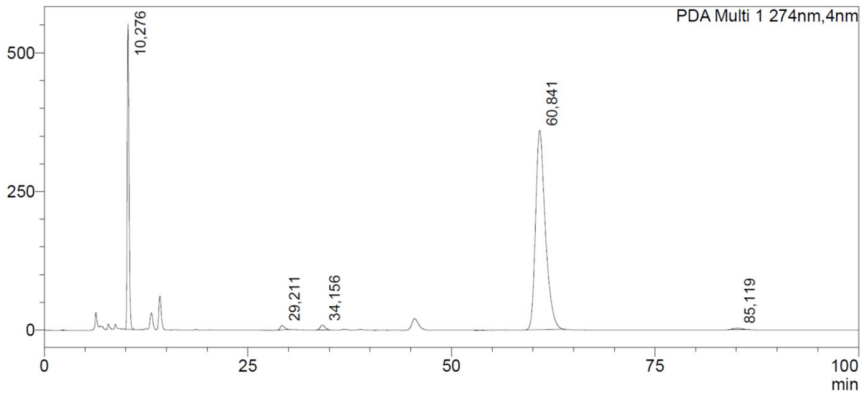
VI. Results of competition catalysis experiments

Figure S7. Chiral HPLC traces for the competition between substrates 1 and 4 in catalysis

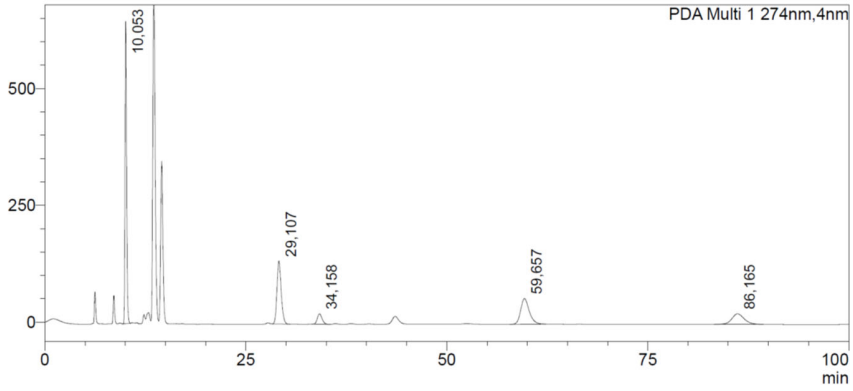
(a) LmrR



(b) LmrR_A92E



(c) LmrR_W96A



VII. Fluorescence titration experiments LmrR_A92E and LmrR_A92Q

Representative procedure

Titrations were performed with 2 µM solutions of protein dimer (2 mL) titrated with a 40 µM solution of [Cu(phen)$(NO_3)_2$] complex in the appropriate buffer solution. 13.5 mM stock solution of [Cu(phen)$(NO_3)_2$] complex **(**[Cu(II)-phen]) was prepared by dissolving the solid in DMSO. By subsequent dilutions with buffer solution, a 40 µM solution of [Cu(II)-phen] was prepared. 20 µM protein conjugates solutions (dimer) in buffer solution (absorbance checked by Nanodrop 2000, Fisher Scientific) was prepared and then diluted x10. Emission spectra were recorded at room temperature after 2 min incubation. The sample was excited at 292 nm and emission spectra were obtained from 310 nm to 450 nm. The excitation and emission slit widths were 5 nm. All the experiments were performed in triplicate with independent batches of the corresponding [Cu(phen)$(NO_3)_2$] complex and protein dimer. The intensity of fluorescence at the maximum of the emission spectra of tryptophan (345nm) was plotted against the concentration ratio between Cu(II) complex and protein dimer.

The data obtained from the titration of the corresponding Cu(II) complex with LmrR were fitted using non-linear regression analysis (program OriginPro 8.5), employing equation **1**.[18]

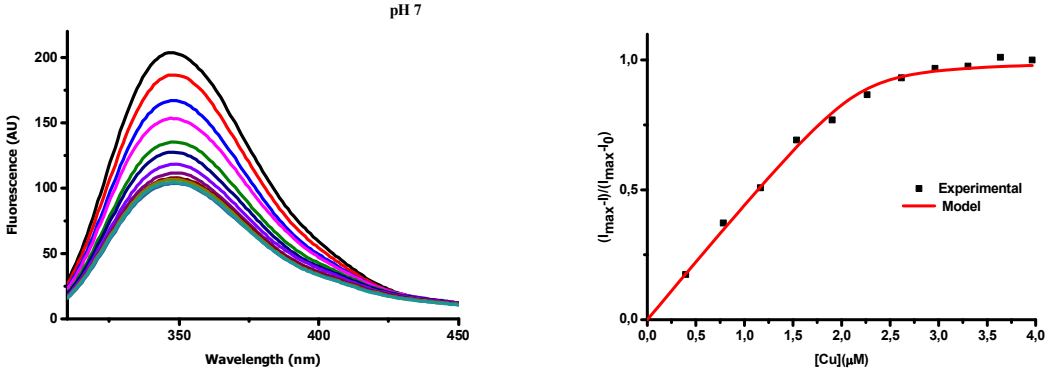
$$FQ = \frac{k}{2}\left(C_P + C_L + K_d - \sqrt{(C_P + C_L + K_d)^2 - 4C_P C_L} \right) \qquad \textbf{(1)}$$

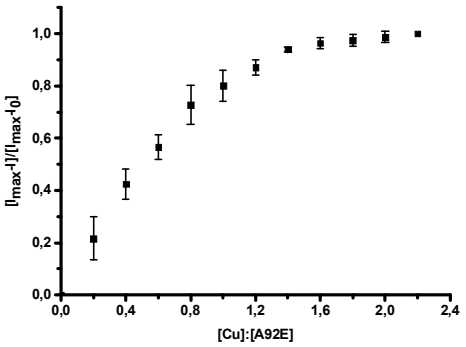
where FQ is the tryptophan fluorescence quenching (ΔQ_F) by [Cu(II)-phen] complex (expressed as (I_{max}-I)/ I_{max}-I_0), with I_{max} the intensity of fluorescence in the absence of quencher, I the intensity of fluorescence upon addition of quencher and I_0 the intensity of fluorescence at saturation). C_p and C_L are the total concentrations of protein and Cu(II) complex, respectively, and *k* is a constant.

Figure S8. Fluorescence titration experiments with LmrR _A92E

(a) Titration of [Cu(II)-phen] with LmrR_ A92E in 20 mM MOPS buffer pH 7.0, containing 150 mM NaCl

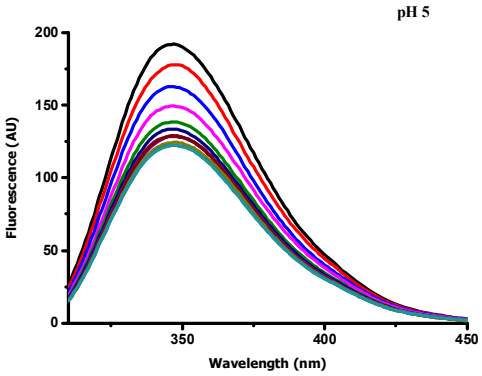
	Value	Standard Error
n	2,19866	0,06619
k	0,03846	0,01129

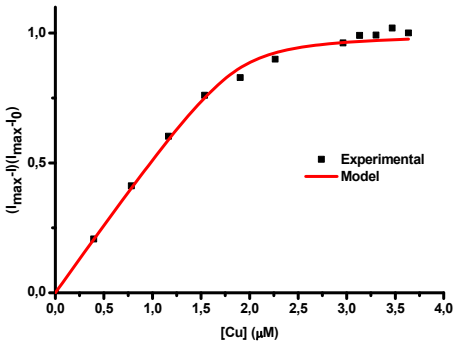




(b) Titration of [Cu(II)-phen] with LmrR_ A92E in 20 mM MES buffer pH 5.0, containing 150 mM NaCl

	Value	Standard Error
n	1,86987	0,07322
k	0,0442	0,01349





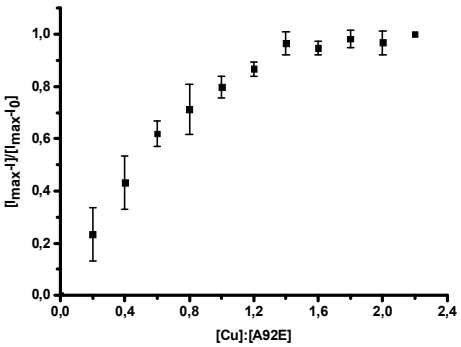


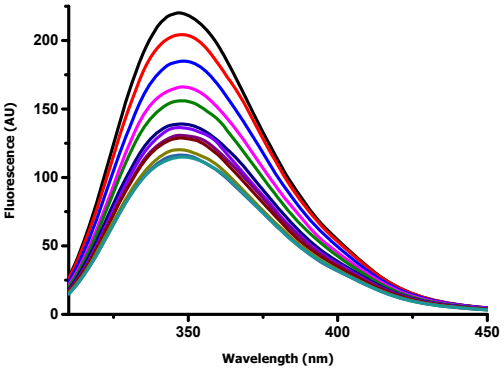
Table S8. Binding constant of Cu(II)-phen with LmrR _A92E.

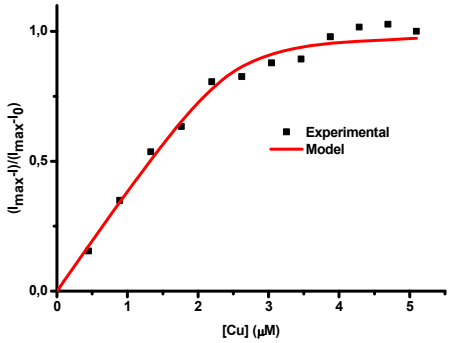
pH	Binding constant K_d (nM)
7	65 ± 19
5	59 ± 16

Figure S9. Fluorescence titration experiments with LmrR _A92Q

Titration of [Cu(II)-phen] with LmrR_ A92Q in 20 mM MOPS buffer pH 7.0, containing 150 mM NaCl

	Value	Standard Error
n	2,48856	0,14628
k	0,07407	0,02843





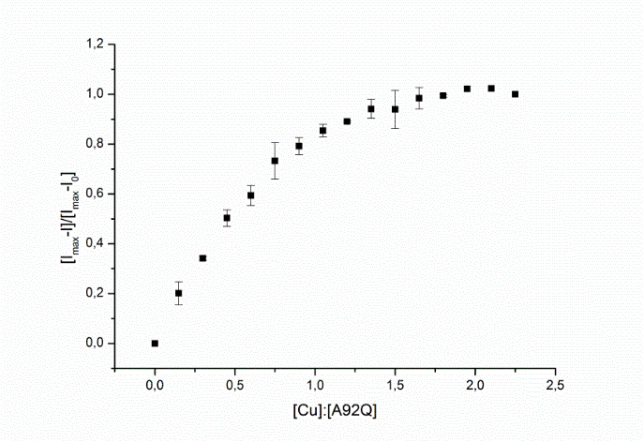


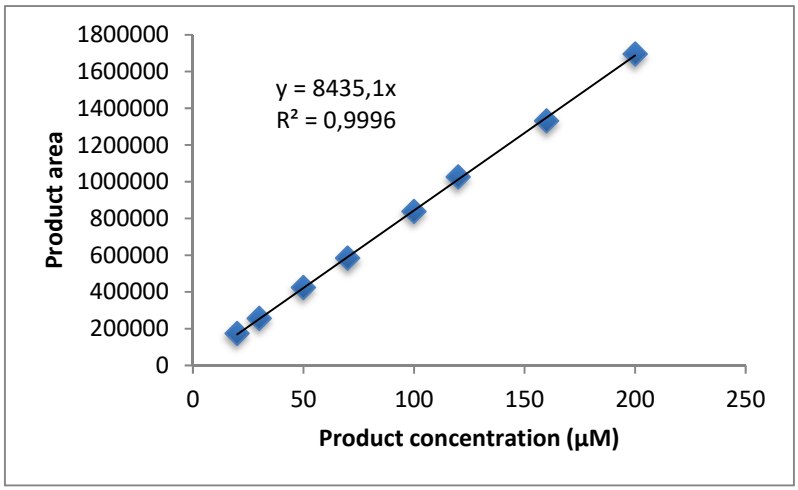
Table S9. Binding constant of Cu(II)-phen with LmrR _A92Q

Binding constant K_d (nM)
103 ± 41

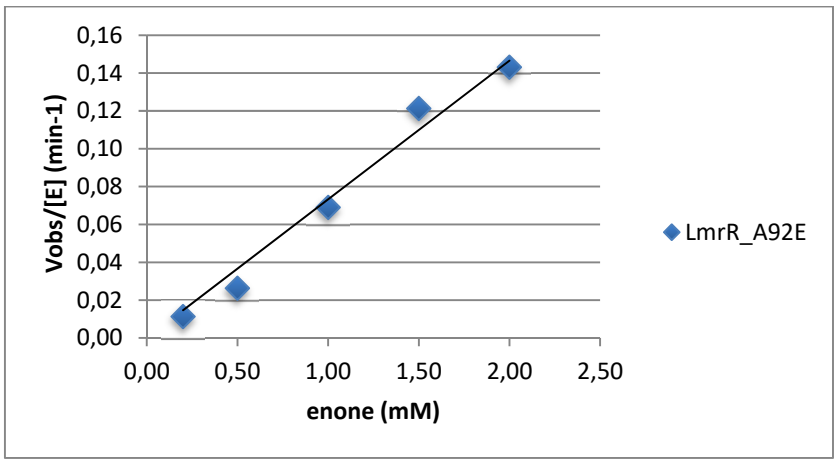
VIII. Kinetic studies with LmrR_A92E mutant

A solution of purified protein LmrR_A92E (6.7 µM) was mixed with Cu(II)-Phen (5 µM) in a 2 mL HPLC vial. After 30 min the substrates – 1 (5 mM) and 2a (0.2 – 2 mM) were added and the vial loaded into the sample tray of a rp-HPLC, maintained at 4 °C. A sample was injected every 30 min onto a Chiralpak AD column, protected by a guard column from possible precipitation in samples. The product separated from the reaction mixture on the HPLC column and the yield was calculated. For the kinetics assay we could not go higher than 5 mM final substrate concentration due to solubility issues. Since, we did not approach substrate saturation and the V_{obs} was still in the linear part of the Michaelis-Menton curve ($S << K_M$), we cannot determine the individual k_{cat} and K_M values of the enzyme.

Figure S10. A) Calibration curve for F-C product



B) Kinetics of the reaction of 1 with 2a catalyzed by LmrR_A92E/Cu(II)-phen



IX. Computational studies

Details of computation

Model systems. The X-ray structure of the LmrR bound to copper-phenanthroline (PDB: 6R1L) was used as protein host. The copper-phenanthroline complex and crystallographic water molecules were manually removed from the model. Lacking protein residues were built by superimposing both monomers as well as the X-ray data of the LmrR bound to Daunomycin (PDB: 3F8F) and to the heme group (PDB: 6FUU) from previous studies.[11,19] Mutants A92E and A92Q were constructed with the Dunbrack backbone-dependent rotamer library[20,21] of the Chimera program.[22]

Ligand-protein docking. The built *apo* LmrR was then used as host to dock the Phen-Cu(II)-2H_2O form of the cofactor (Figure S12). This was first optimized at DFT level using the B3LYP-D3 functional.[23–25] The 6-31g(d,p) basis set[26–28] was used for non-metallic atoms and SDD[29] pseudopotential including *f* polarization functions for the 19 outer electrons of copper. All dockings were performed with GOLD 5.2[30] and evaluated with the ChemScore[31] scoring function.

Molecular Dynamics simulations. The best scored docking solutions were used as starting points for the all-atoms Molecular Dynamics (MD) simulations. Models were set up with the xleap[32] program. Each system was embedded into a cubic box including about 37000 water molecules (10 Å from the protein to the edge of the box) and the corresponding sodium ions to neutralize the simulation cell. The AMBER[32] and TIP3P[33] force fields were used for protein and water, respectively. For sodium ions, parameters from ions94.lib library[32] were used. Parameters for the Phen-Cu(II)-2H_2O complex were developed. RESP[34] charges were calculated with antechamber.[35] Bonded terms involving the metal center were calculated based on the Seminario's approach.[36] The remaining atoms were parameterized according to GAFF force field.[37] A cut off of 10 Å was applied for short range electrostatics and Van der Waals interactions. Long range electrostatic interactions were calculated with the Particle Mesh Ewald method.[38] Bonds involving hydrogen atoms were constrained using the SHAKE algorithm.[39] To integrate the equation of motion a Langevin integrator[40,41] and a time step of 1 fs were used. Constant temperature and pressure were achieved by coupling the systems to a Monte Carlo barostat at 1.01325 bar.[42] Simulations were performed with OpenMM 7.1[43] following a standard protocol specified into OMMProtocol[44]: initially, model systems were progressively minimized (3000 steps) allowing water molecules, side-chain and backbone atoms to accommodate; 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. A92E and A92Q mutants were calculated per triplicate. Visualization of the simulations and the molecular graphics were performed with the UCSF Chimera program.[22]

Figures and Tables

Figure S11. Optimized structure of the biaqua form of the copper bound phenanthroline cofactor (Phen-Cu(II)-2H_2O) used as ligand in this study.

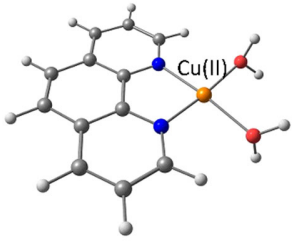


Figure S12. Polar residues at the back of LmrR selected for assessing interactions with A92E and A92Q. Notice that N14/N14' are the only residues located at the opposite side of the active site with respect to A92/A92'. Thus, it is reasonable that is the only interaction with A92E/A92Q that contributes to the opening/closing of the dimer interface.

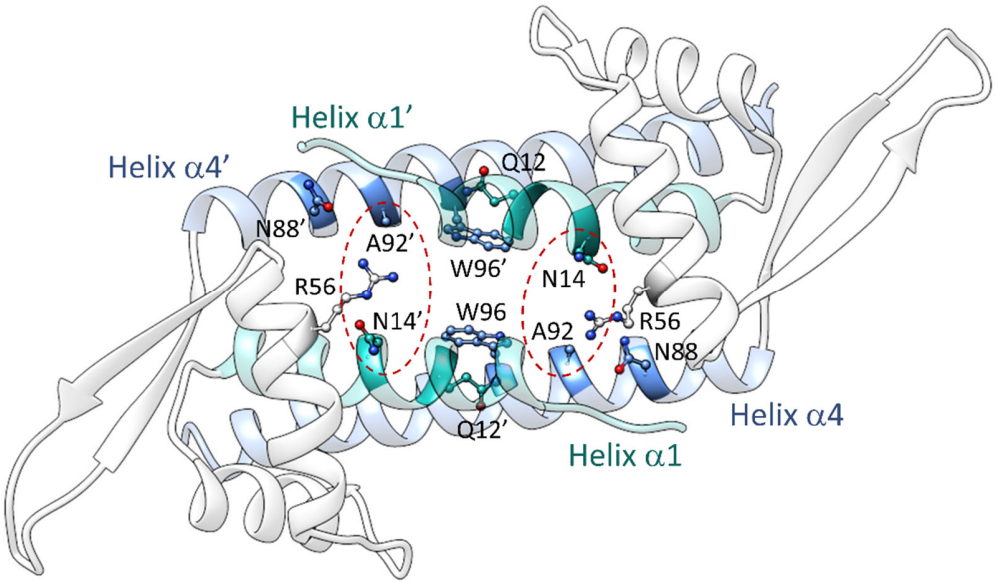


Figure S13. Hydrogen bonding interactions between A92E (blue) or A92Q (red) and all the polar residues at the back of LmrR, as indicated in Figure S13. Direct (left) or indirect interactions through bridging water molecules (right) were assessed. A92E variant interacts with N14 through bridging water molecules significantly more frequently than A92Q.

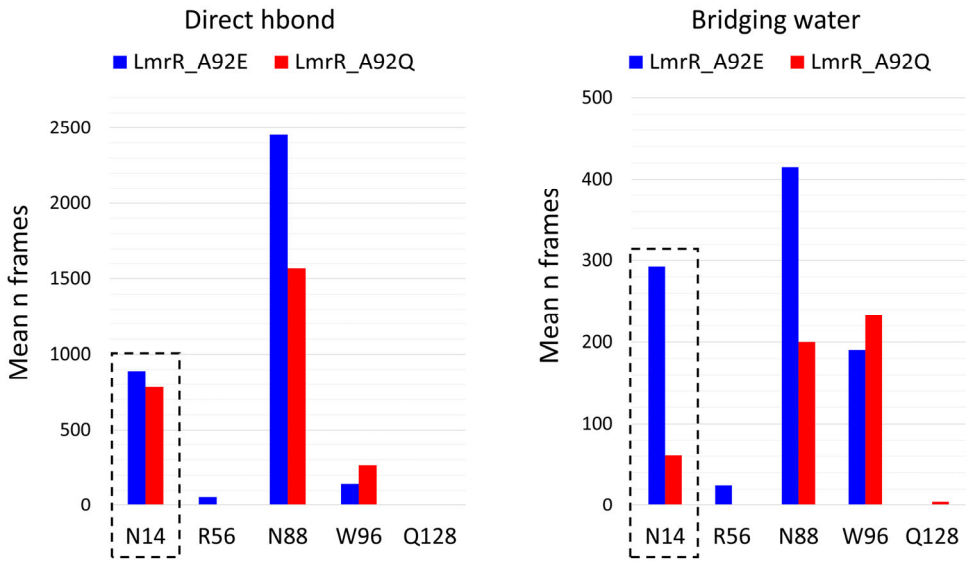


Figure S14. Representation of the methodology implemented to assess the relative positioning of W96/W96' residues in the systems under study. (**a**) Atoms of the tryptophan rings selected for the generation of a plane (blue and orange plates) passing through them. (**b**) Frame by frame and for 10000 frames, axes (blue and orange sticks) passing throughout the centroid of the selected atoms in **a** and parallel to the generated planes are calculated. Then, the angle between these axes is considered for analysis, which corresponds to the relative angle between the rings of tryptophans W96/W96'. Each time an angle < 20° is found, the frame is counted.

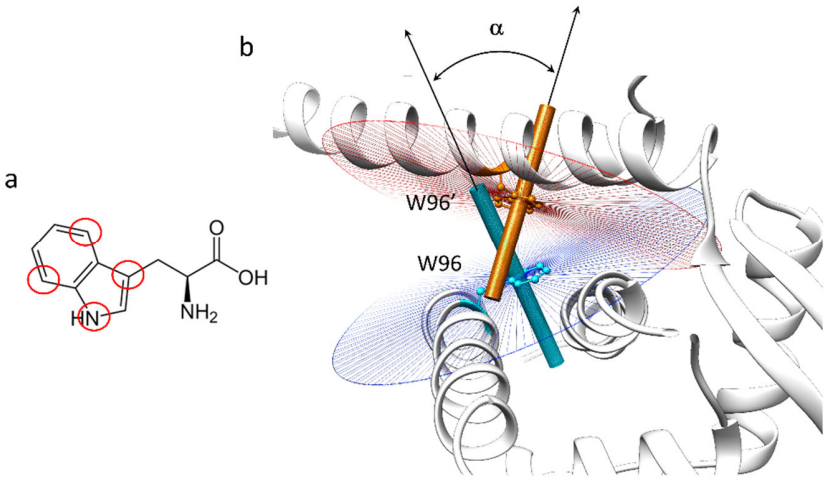
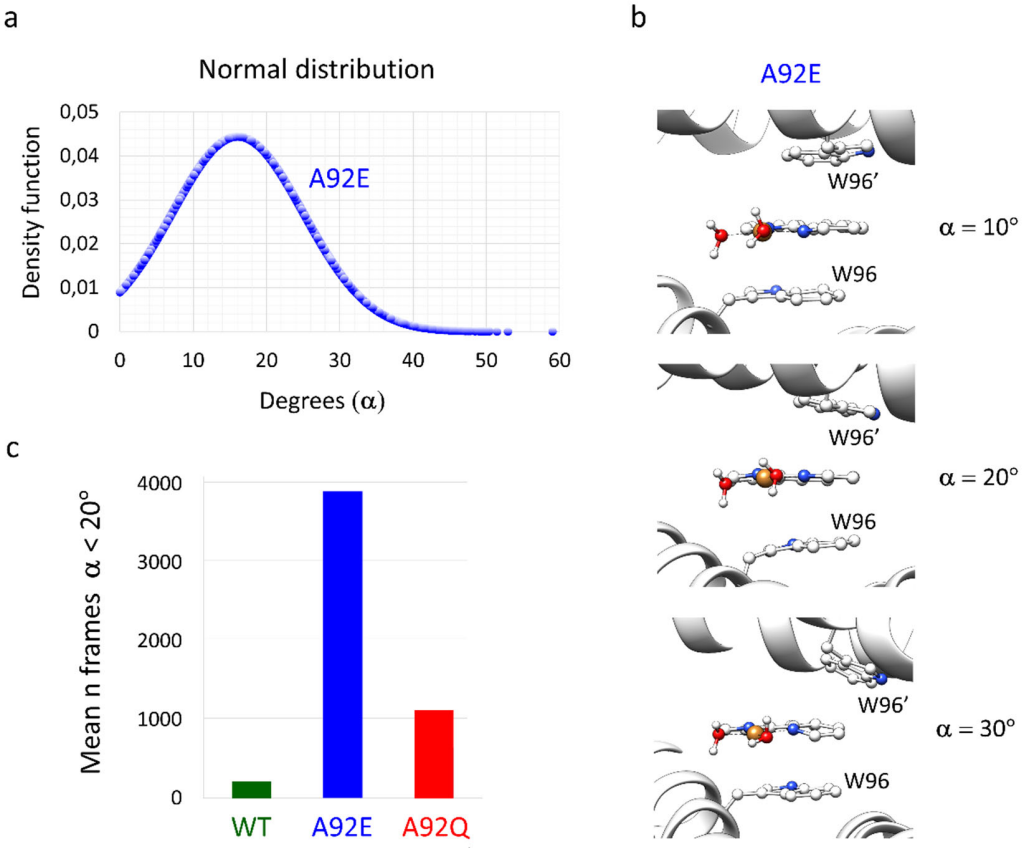
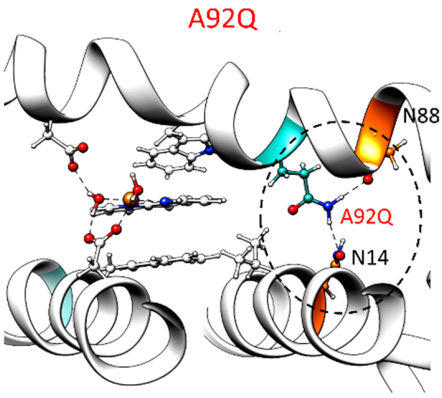


Figure S15. Frames from the three replicas of MD simulations for A92E variant associated to a proper packing of the Phen-Cu(II)-2H_2O cofactor were extracted to calculate the normal distribution of the relative W96/W96' angle associated to this configuration (**a**), being the major density of frames associated to angles between 10° and 20°. Since, close to 30°, a good π-stacking between both W96/W96' and phenanthroline was not observed (**b**), 20° was selected as cut off to compare the structural arrangement of the LmrR variants assessed (**c**). The results show that A92E packs better the Phen-Cu(II)-2H_2O cofactor than A92Q*, and this better than WT.



**Data for A92Q variant corresponds to 100 ns of MD simulation analyzed per duplicate, since in the third replica the phenanthroline cofactor abandoned the active site of the protein from the beginning of the simulation. Thus, this was not considered as part of the protein-cofactor interaction analysis.*

Figure S16. Representative structure along 100 ns of MD simulation of the A92Q variant of LmrR analyzed per duplicate*. As found for the A92E mutant, LmrR_A92Q is also able to perform a hydrogen bonding network involving A92Q, N88 and N14 residues, being A92Q and N14 the ones contributing to the opening of the dimer interface. However, this arrangement is by far less frequent than as it is found for A92E mutant, probably due to the lack of the negative charge in the glutamine amino acid, which makes these interactions less frequent.



**As mentioned in Figure S16, in the third replica of the MD simulation for this variant, the phenanthroline cofactor abandoned the active site of the protein from the beginning of the simulation. Thus, this was not considered as part of the protein-cofactor interaction analysis.*

Table S10. Best scored docking solutions for the Phen-Cu(II)-2H_2O ligand bound to WT, A92E and A92Q forms of LmrR. The ChemScore scoring function estimates the total free energy (**ΔG**) resulting from ligand binding by adding to it the magnitude of each particular physical contribution: lipophilicity (**S_{lipo}**), protein-ligand hydrogen bonding (**S_{hbond}**), metal binding (**S_{metal}**) and the loss of conformational entropy of the ligand after binding (**H_{rot}**). The final ChemScore value (**ChemScore** column) adds to the total free energy: clash penalty resulting from too close contacts between the ligand and the protein (**Clash**), internal ligand torsional strain penalty **(Internal**) and an internal ligand energy correction term (**Intcor**).

Protein host	**ChemScore**	**ΔG**	**S_{hbond}**	**S_{metal}**	**S_{lipo}**	**H_{rot}**	**Clash**	**Internal**	**Intcor**
LmrR_WT	30.05	-30.05	0.00	0.00	210.04	0.00	0.00	0.00	0.00
LmrR_A92E	30.22	-30.22	0.00	0.00	211.49	0.00	0.00	0.00	0.00
LmrR_A92Q	30.27	-30.27	0.00	0.00	211.92	0.00	0.00	0.00	0.00

Figure S17. Clustering of structures along 300 ns MD simulation for WT and A92E forms of LmrR bound to the Phen-Cu(II) cofactor. Notice that cluster c0 for A92E variant is highly populated (47% of frames) in contrast to the most populated cluster for WT (20% of frames). In addition, for A92E, cluster c0 together with c1 and c3 correspond with configurations consistent with parallel tryptophans and dual π-stacking of the Phen-Cu(II) cofactor by tryptophans W96/W96' (see **Figure S19**).

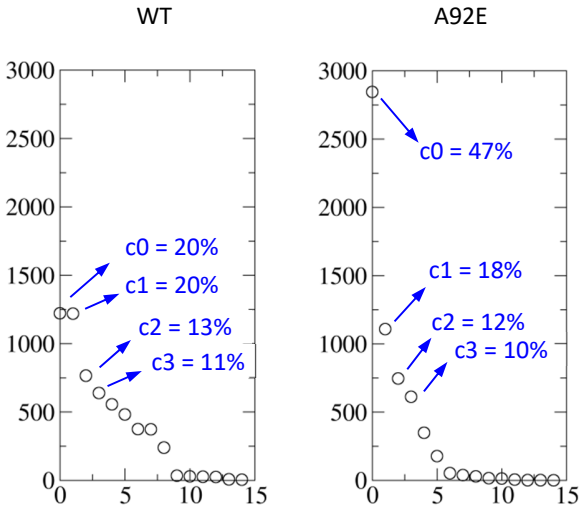


Figure S18. Illustration of the measurement of the hydrophobic surface volume in the representative structures of the most populated cluster c0 for WT and A92E variants of LmrR. For detailed information see **Figure S20**.

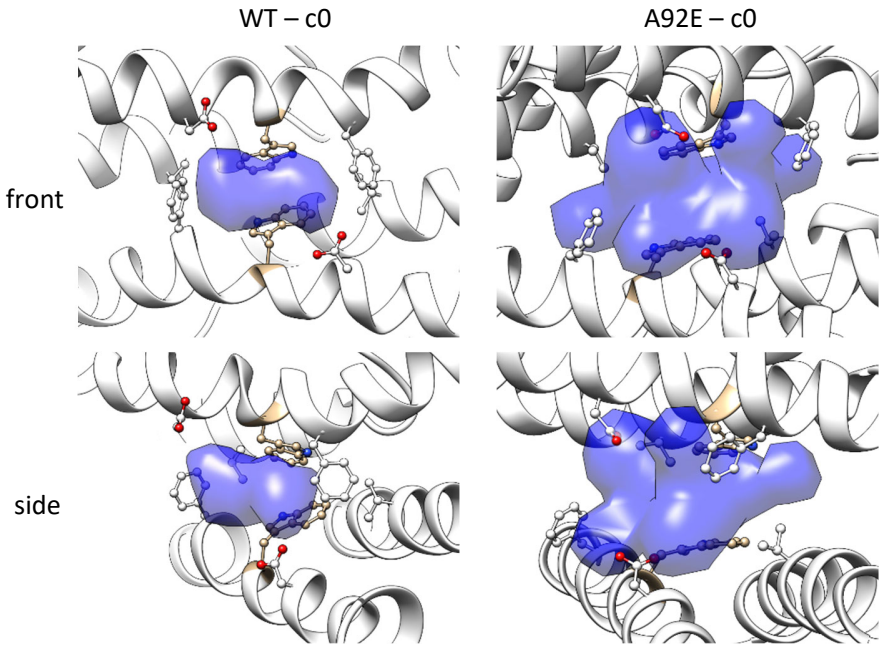


Figure S19. Measurement of the hydrophobic surface volume, calculated with Chimera, in the representative structures of the cluster c0 – c4 for WT and A92E variants of LmrR. Please notice that these clusters are not equally populated (cluster c0 for WT contains 20% of the analyzed frames while cluster c0 for A92E contains a 47%. For further details see **Figure S18**).

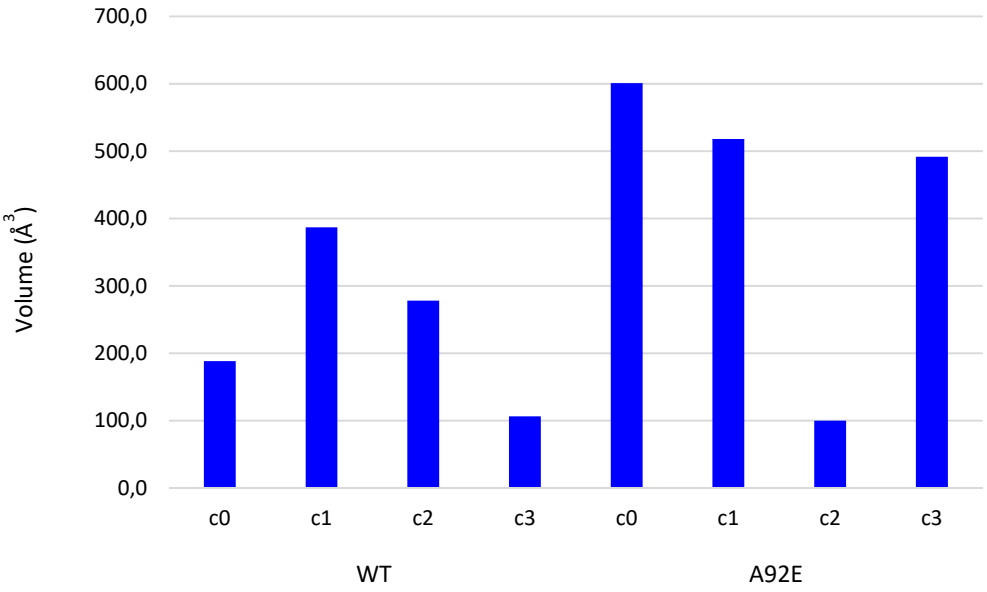


Figure S20. The columns named "Representative structure" show the protein-ligand interaction profile calculated with PLIP program[45] of the representative structure for clusters c0, c1, c2 and c3 calculated from 300 ns of MD simulation for both the WT and A92E variants of LmrR (see **Table S11**). Notice that for A92E variant, and NOT for WT, dual π-stacking by W96/W96' is found (c0, c1 and c3), which is possible due to the parallel arrangement of these tryptophans. The columns named "docking" shows the best five solutions of the Phen-Cu(II) cofactor bound to these representative structures after manually removing the Phen-Cu(II) cofactor. Docking results show good agreement with the binding mode of the Phen-Cu(II) complex found along the simulation which, for the A92E variant, results in slightly better Scoring due to an increase in the hydrophobicity resulting from the hydrophobic patch generated by W96/W96' (see **ChemScore** and $\mathbf{S_{lipo}}$ term in **Table S12**). Tryptophans W96/W96' are highlighted in tan.

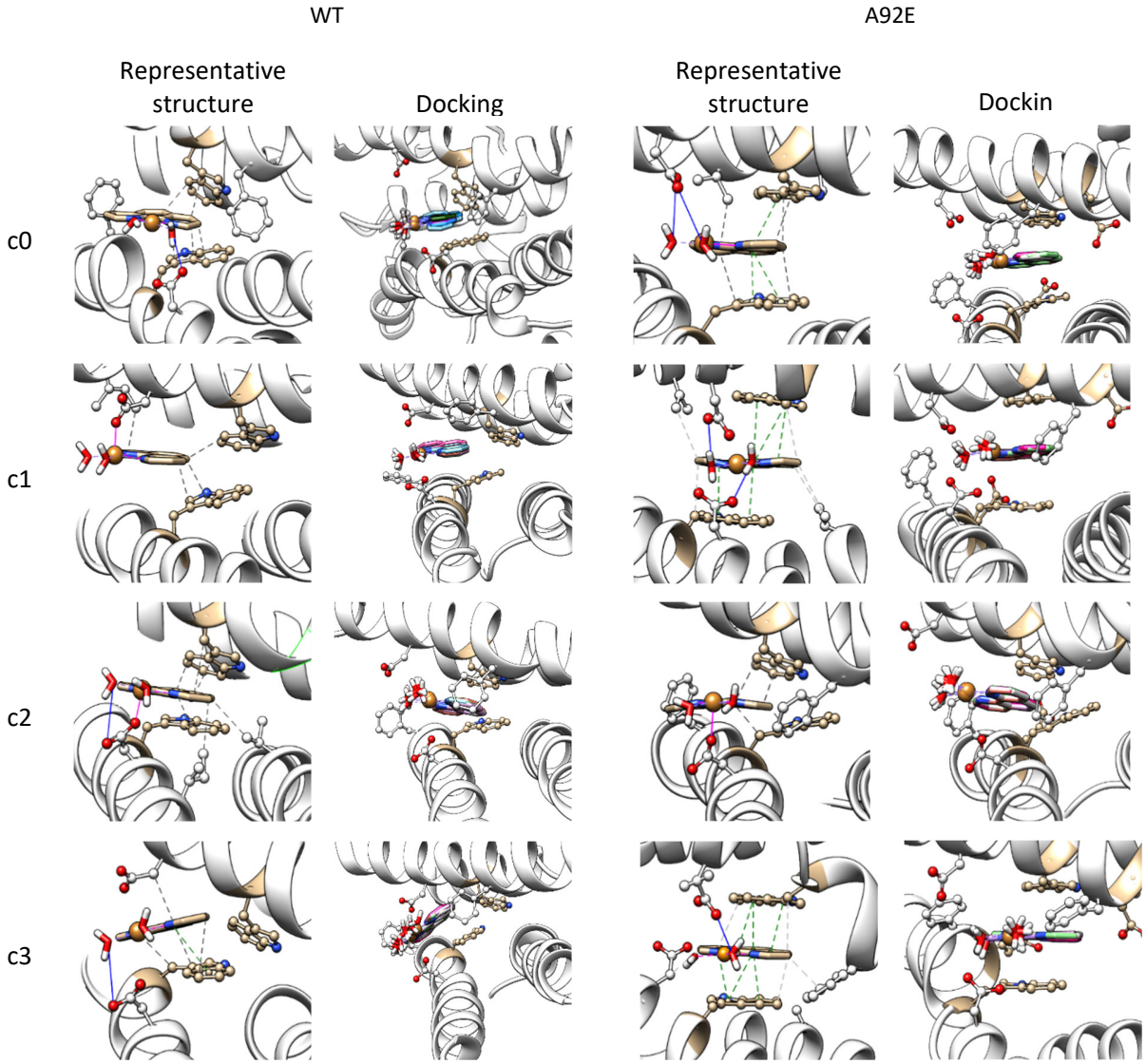


Table S11. Protein-ligand interaction analysis performed with PLIP program on the representative structures of most populated clusters along 300 ns MD simulation for both WT and A92E variants of LmrR.

		Hydrophobic	**Hbond**	**π-stacking**	**Cu coordination**
WT	c0	W96/W96'	D100'	-	-
	c1	W96/W96', I103	-	-	D100'
	c2	W96/W96', V15', I103	D100	-	D100
	c3	W96, D100'	D100	W96	-
A92E	c0	W96/W96', V15'	D100	W96/W96'	-
	c1	W96/W96', V15/V15'	D100/D100'	W96/W96'	-
	c2	W96/W96', F93/F93'		-	D100'
	c3	W96/W96', F93', E92	D100'	W96/W96'	-

Table S12. Best scored docking solutions for the Phen-Cu(II)-2H_2O ligand bound to the representative structures of clusters c0, c1, c2 and c4 calculated from 300 ns of MD simulation for both the WT and A92E variants of LmrR. The ChemScore scoring function estimates the total free energy (**ΔG**) resulting from ligand binding by adding to it the magnitude of each particular physical contribution: lipophilicity (**S_{lipo}**), protein-ligand hydrogen bonding (**S_{hbond}**), metal binding (**S_{metal}**) and the loss of conformational entropy of the ligand after binding (**H_{rot}**). The final ChemScore value (**ChemScore** column) adds to the total free energy: clash penalty resulting from too close contacts between the ligand and the protein (**Clash**), internal ligand torsional strain penalty **(Internal**) and an internal ligand energy correction term (**Intcor**).

Variant	**Cluster**	**ChemScore**	**ΔG**	**S_{hbond}**	**S_{metal}**	**S_{lipo}**	**H_{rot}**	**Clash**	**Internal**	**Intcor**
WT	c0	28.74	-29.04	0.0	0.0	201.39	0.0	0.3	0.0	0.0
	c1	28.61	-29.14	0.0	0.0	202.26	0.0	0.54	0.0	0.0
	c2	29.61	-29.7	0.0	0.0	207.02	0.0	0.09	0.0	0.0
	c3	22.22	-24.59	0.0	0.0	163.31	0.0	2.36	0.0	0.0
A92E	c0	30.03	-30.39	0.0	0.0	212.89	0.0	0.36	0.0	0.0
	c1	29.96	-29.97	0.0	0.0	209.33	0.0	0.01	0.0	0.0
	c2	28.55	-29.16	0.0	0.0	202.38	0.0	0.61	0.0	0.0
	c3	31.34	-31.39	0.0	0.0	221.42	0.0	0.04	0.0	0.0

X. References

(1) Bos, J.; García-Herraiz, A.; Roelfes, G.; Thunnissen, A. W. H.; Driessen, A. J. M.; Watanabe, Y.; Wang, Z.; Guo, Z.; Lu, Y.; Ménage, S. An Enantioselective Artificial Metallo-Hydratase. *Chem. Sci.* **2013**, *4* (9), 3578–3582.

(2) van der Berg, J. P.; Madoori, P. K.; Komarudin, A. G.; Thunnissen, A.-M.; Driessen, A. J. M. Binding of the Lactococcal Drug Dependent Transcriptional Regulator LmrR to Its Ligands and Responsive Promoter Regions. *PLoS One* **2015**, *10* (8), e0135467.

(3) Kabsch, W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 133–144.

(4) Evans, P. Scaling and Assessment of Data Quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62* (1), 72–82.

(5) Collaborative Computational Project, Number 4, N. 4; IUCr. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50* (5), 760–763.

(6) McCoy, A. J. Solving Structures of Protein Complexes by Molecular Replacement with *Phaser*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2007**, *63* (1), 32–41.

(7) Emsley, P.; Cowtan, K. *Coot* : Model-Building Tools for Molecular Graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12), 2126–2132.

(8) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *REFMAC* 5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 355–367.

(9) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with *Phenix.Refine*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (4), 352–367.

(10) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C.; IUCr. *MolProbity* : All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.

(11) Madoori, P. K.; Agustiandari, H.; Driessen, A. J. M.; Thunnissen, A.-M. W. H. Structure of the Transcriptional Regulator LmrR and Its Mechanism of Multidrug Recognition. *EMBO J.* **2009**, *28* (2), 156–166.

(12) Roelfes, G.; Boersma, A. J.; Feringa, B. L. Highly Enantioselective DNA-Based Catalysis. *Chem. Commun.* **2006**, (6), 635–637.

(13) Navarro, M.; Cisneros-Fajardo, E. J.; Sierralta, A.; Fernández-Mestre, M.; Silva, P.; Arrieche, D.; Marchán, E. Design of Copper DNA Intercalators with Leishmanicidal

Activity. *JBIC J. Biol. Inorg. Chem.* **2003**, *8* (4), 401–408.

(14) Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G. Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold. *J. Am. Chem. Soc.* **2015**, *137* (31), 9796–9799.

(15) Boersma, A. J.; Feringa, B. L.; Roelfes, G. Enantioselective Friedel–Crafts Reactions in Water Using a DNA-Based Catalyst. *Angew. Chemie Int. Ed.* **2009**, *48* (18), 3346–3348.

(16) Charles, M. D.; Schultz, P.; Buchwald, S. L. Efficient Pd-Catalyzed Amination of Heteroaryl Halides. *Org. Lett.* **2005**, *7* (18), 3965–3968.

(17) García-Fernández, A.; Megens, R. P.; Villarino, L.; Roelfes, G. DNA-Accelerated Copper Catalysis of Friedel-Crafts Conjugate Addition/Enantioselective Protonation Reactions in Water. *J. Am. Chem. Soc.* **2016**, *138* (50), 16308–16314.

(18) Zsila, F.; Bikádi, Z.; Simonyi, M. Induced Circular Dichroism Spectra Reveal Binding of the Antiinflammatory Curcumin to Human A1-Acid Glycoprotein. *Bioorg. Med. Chem.* **2004**, *12* (12), 3239–3245.

(19) Villarino, L.; Splan, K. E.; Reddem, E.; Alonso-Cotchico, L.; Gutiérrez de Souza, C.; Lledós, A.; Maréchal, J.-D.; Thunnissen, A.-M. W. H.; Roelfes, G. An Artificial Heme Enzyme for Cyclopropanation Reactions. *Angew. Chemie Int. Ed.* **2018**, *57* (26), 7785–7789.

(20) Dunbrack, R. L.; Karplus, M. Backbone-Dependent Rotamer Library for Proteins Application to Side-Chain Prediction. *J. Mol. Biol.* **1993**, *230* (2), 543–574.

(21) Dunbrack, R. L. Rotamer Libraries in the 21st Century. *Curr. Opin. Struct. Biol.* **2002**, *12* (4), 431–440.

(22) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

(23) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(24) 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* (45), 11623–11627.

(25) 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* (15), 154104.

(26) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. A Complete Basis Set Model Chemistry. I. The Total Energies of Closed-shell Atoms and Hydrides of the First-row Elements. *J. Chem. Phys.* **1988**, *89* (4), 2193–2218.

(27) Petersson, G. A.; Al-Laham, M. A. A Complete Basis Set Model Chemistry. II. Open-shell Systems and the Total Energies of the First-row Atoms. *J. Chem. Phys.* **1991**, *94* (9), 6081–6090.

(28) Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. 6-31G* Basis Set for Third-Row Atoms. *J. Comput. Chem.* **2001**, *22* (9), 976–984.

(29) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted *a b i n i t i o* Pseudopotentials for the First Row Transition Elements. *J. Chem. Phys.* **1987**, *86* (2), 866–872.

(30) Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D. Improved Protein-Ligand Docking Using GOLD. *Proteins Struct. Funct. Bioinforma.* **2003**, *52* (4), 609–623.

(31) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical Scoring Functions: I. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes. *J. Comput. Aided. Mol. Des.* **1997**, *11* (5), 425–445.

(32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179–5197.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(34) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.

(35) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25* (2), 247–260.

(36) Seminario, J. M. Calculation of Intramolecular Force Fields from Second-Derivative Tensors. *Int. J. Quantum Chem.* **1996**, *60* (7), 1271–1277.

(37) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

(38) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.

(39) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes **1977**, *23* (3), 327-341.

(40) Schneider, T.; Stoll, E. Molecular-Dynamics Study of a Three-Dimensional One-

Component Model for Distortive Phase Transitions. *Phys. Rev. B* **1978**, *17* (3), 1302–1322.

(41) Brünger, A.; Brooks, C. L.; Karplus, M. Stochastic Boundary Conditions for Molecular Dynamics Simulations of ST2 Water. *Chem. Phys. Lett.* **1984**, *105* (5), 495–500.

(42) Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. Hybrid Monte Carlo. *Phys. Lett. B* **1987**, *195* (2), 216–222.

(43) Eastman, P.; Pande, V. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Comput. Sci. Eng.* **2010**, *12* (4), 34–39.

(44) Rodríguez-Guerra, J.; Pedregal, L.; Alonso-Cotchico, L.; Velasco-Carneros, J.-D.; Maréchal, M. OMMProtocol: A Command Line Application to Launch Molecular Dynamics Simulations with OpenMM. https://doi.org/10.26434/chemrxiv.7059263.v1.

(45) Salentin, S.; Schreiber, S.; Haupt, V. J.; Adasme, M. F.; Schroeder, M. PLIP: Fully Automated Protein–Ligand Interaction Profiler. *Nucleic Acids Res.* **2015**, *43* (W1), W443–W447.