Computational and Experimental characterization of the catalytic activity of artificial enzymes based on the LmrR scaffold

Nayanika Das, Jordi Villà-Freixa
Universitat de Vic - Universitat Central de Catalunya

In collaboration with Roelfes Lab (University of Groningen)

In this document two possible projects are suggested for a collaboration with the group of Roelfes Lab at the University of Groningen, on the characterisation of the activity of newly created artificial enzymes based in the lactococcal multidrug resistance regulator (LmrR), a dimeric transcription factor with a large, hydrophobic binding pocket at its dimer interface. The pocket includes two tryptophan moieties, which are important for promiscuous binding of planar hydrophobic conjugated compounds by π-stacking, which can be transformed into an active enzyme by either incorporating transition metal complexes [2] or by incorporating non-natural amino acids such as p-aminophenylalanine (pAF) to catalyze reactions [3].Artificial enzymes are engineered biocatalysts that enable new-to-nature transformations with the precision and selectivity of natural enzymes, but often for reactions that are not found in biology [2] [4] [5]. LmrR is a dimeric transcription factor from Lactococcus lactis with a large hydrophobic pocket that can be engineered to host biological catalytic functionalities by incorporating non-natural amino acids such as p-aminophenylalanine (pAF) to catalyze reactions [2]. For example, LmrR_pAF efficiently catalyzes Friedel-Crafts alkylation and hydrazone formation, with pAF acting as the nucleophilic residue that forms iminium intermediates with carbonyl substrates [6]; however, the molecular mechanisms by which these mutations exert their effects remain incompletely understood [2]. The substrates include typical aromatic aldehydes and nucleophiles commonly used in bioconjugation chemistry, and the artificial enzyme is able to catalyze the formation of either a hydrazone or an oxime linkage, depending on the nucleophile present [8].

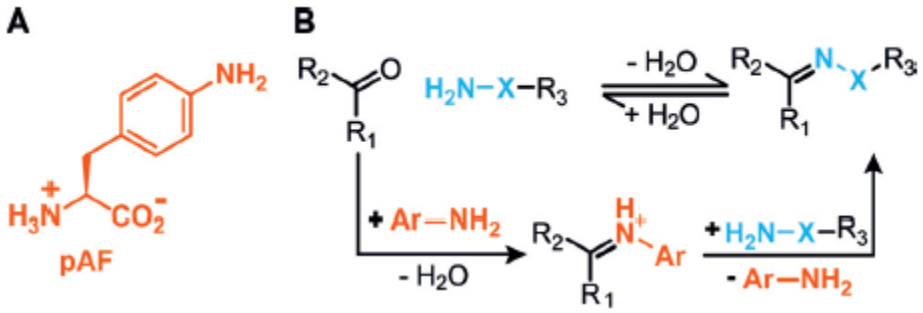
Synthesis of LmrR_pAF

Researchers used protein engineering to introduce a non-canonical amino acid (p-aminophenylalanine, pAF) at V15 in LmrR. This was accomplished by expanding the genetic code of the host organism, typically Escherichia coli, to allow site-specific incorporation of the unnatural amino acid during protein expression [2] [12]. This can be achieved using the stop codon suppression methodology introduced by Schultz et al., which relies on an orthogonal tRNA/aminoacyl-tRNA synthetase pair that recognizes a unique codon (often the amber stop codon, TAG) and specifically incorporates pAF at the desired site in the protein sequence [12]. The process begins with site-directed mutagenesis to introduce the TAG codon at the target position (such as V15) in the LmrR gene. The engineered gene is then co-expressed in cells along with the orthogonal tRNA/synthetase pair and supplied with pAF in the growth medium [12]. During translation, the cellular machinery incorporates pAF precisely at the engineered site, yielding a modified LmrR protein with a unique catalytic residue.The modified LmrR protein thus serves as a scaffold, with the newly introduced non-canonical amino acid providing the desired catalytic functionality[2] .

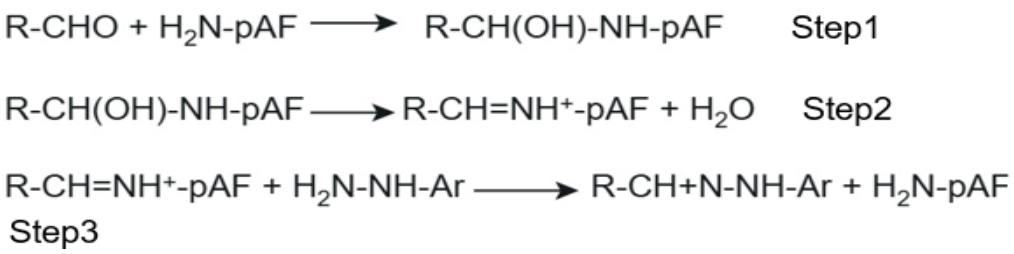
Catalytic properties of LmrR_pAF

The resulting artificial enzyme LmrR_pAf is a promiscuous catalyzer that has been shown to accelerate bothe the hydrazone formation and Friedel-Crafts alkylation reactions. The mechanisms of boh reactions differ and ar drafted below.

Hydrazone Reaction

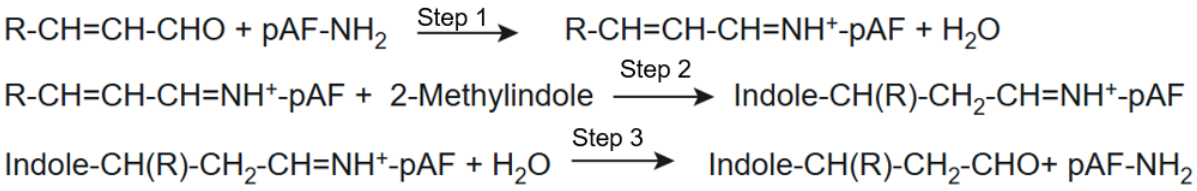


Anilines are well-known nucleophilic catalysts for hydrazone (X = NH) and oxime (X = O) formation. Primary aromatic amines accelerate these reactions by the formation of a protonated Schiff base intermediate, which subsequently undergoes transamination to yield the desired products (adapted from [11])



- Step 1: Hemiaminal Formation
  The aniline side chain of the unnatural amino acid p-aminophenylalanine (pAF) in the enzyme's active site acts as a nucleophile, attacking the carbonyl carbon of the aldehyde substrate (e.g., benzaldehyde). This forms a covalent hemiaminal intermediate [1]
  Barrier: ~15.2 kcal/mol (rate-determining step)[7].
- Step 2: Dehydration to Iminium Ion
  The hemiaminal intermediate undergoes dehydration (loss of water), resulting in the formation of an iminium ion covalently linked to the pAF residue [1]
  Barrier: ~8.7 kcal/mol [7].
- Step 3: Nucleophilic Attack by Hydrazine (Transimination)
  The hydrazine substrate (e.g., NBD-H) attacks the iminium intermediate, displacing the pAF side chain (regenerating the free pAF residue) and forming the hydrazone product and the hydrazone product dissociates from the enzyme's active site, completing the catalytic cycle [1] [7].

Friedel-Crafts Reaction



Friedel–Crafts alkylation between α,β-unsaturated aldehydes (enals) and indoles, containing the noncanonical amino acid para-aminophenylalanine (pAF) [6].

Reaction Sequence for FC alkylation Catalyzed by LmrR_pAF

- Step 1: Iminium Ion Formation
  pAF aniline side chain reacts with the enal substrate to form a Schiff base (iminium ion) intermediate [6].
- Step 2: Nucleophilic Attack by Indole
  Indole attacks the iminium intermediate, forming a new C–C bond (alkylated indole intermediate) [6].
- Step 3: Product Formation and Catalyst Regeneration
  The intermediate undergoes hydrolysis, releasing the product and regenerating the pAF residue [6].

Kinetic Parameters

Table 1: Comparison of Kinetic Parameters and Activation Barriers for Hydrazone Formation and Friedel–Crafts Reactions

Reaction	System/Catalyst	k_{cat} (s^{-1} $\times 10^{-2}$)	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})	$\Delta G^{\ddagger}$, QM (kcal/mol)	$\Delta G^{\ddagger}$, exp (kcal/mol)	Reference
Hydrazone formation	Buffer only	n.a.	n.a.	~0.001	23.0 (uncatalyzed)	—	Drienovská et al., *Nat Chem* 2018
Hydrazone formation	LmrR (no pAF)	n.a.	n.a.	~0.002	—	—	Drienovská et al., *Nat Chem* 2018
Hydrazone formation	LmrR_pAF	1.2	0.017	0.72	15.2 (hemiaminal) 8.7 (dehydration)	—	Drienovská et al., *Nat Chem* 2018
Friedel–Crafts alkylation	Buffer only	n.a.	n.a.	0.01	—	24.7	Leveson-Gower et al., *ACS Catal* 2021
Friedel–Crafts alkylation	LmrR (no pAF)	n.a.	n.a.	0.02	—	23.4	Leveson-Gower et al., *ACS Catal* 2021
Friedel–Crafts alkylation	LmrR_pAF	1.45	18.2	0.80	—	20.8	Leveson-Gower et al., *ACS Catal* 2021

The table summarizes the kinetic parameters, comparing the uncatalyzed processes, wild-type LmrR, and the engineered LmrR_pAF variant [1][6][7]. Catalyst lowers the activation barrier (ΔG‡), leading to increased reaction rates. This is evident in the dramatic improvements in kcat and catalytic efficiency for LmrR_pAF. Agreement between QM and experimental barriers in buffer/water validates the mechanistic understanding of the uncatalyzed process.

Questions

The use of unnatural amino acids to generate artificial enzymes based on the LmrR dimer scaffold has been shown to produce functional catalyzers for both hydrazone and Friedel-Crafts reactions. Two questions arise that we would like to pursue in this collaboration, as alternative projects from which one should be chosen:

1. PROJECT 1: Can a non-natural amino acid other than pAF, when incorporated at the catalytic site of LmrR enhance iminium-based catalysis of hydrazone formation?
2. PROJECT 2: Do computational and experimental alanine scanning analyses fit for either hydrazone formation or Friedel-Crafts alkylation?

State of the Art on non-canonical amino acids

Roelfes' group has shown that swapping the active site ncAA (e.g., from p-aminophenylalanine to 3-amino tyrosine) in FC alkylation [10] can dramatically change the enzyme's enantioselectivity and activity.

How a new ncAA can Contribute

If we use the same mechanism as with pAF (p-acetylphenylalanine) for a new ncAA like p-propionyl phenylalanine in Hydrazone Reaction, the general reactivity such as hydrazone formation essentially should be preserved, since both side chains feature a ketone group at the para position.
However, the outcome will depend on the specific context:

- The propionyl group ($–COCH_2CH_3$) is slightly more bulky, the longer side chain could affect how the ncAA sits in the active site or interacts with other residues.

Property	p-Acetylphenylalanine (pAF)	p-Propionyl Phenylalanine
Side chain functional group	para-acetyl ($–COCH_3$)	para-propionyl ($–COCH_2CH_3$)
Carbonyl group type	Ketone (The ketone group reacts with hydrazines to form hydrazones)	By analogy, its ketone group would also react with hydrazines to form hydrazones (though explicit references are not present in the search results).
Reactivity in hydrazone ligation	High (widely used)	High (slightly less than pAF)

Steric bulk	Small	Larger (longer alkyl chain)

State of the Art on Alanine Scanning

In the paper *"Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel−Crafts Alkylase"* [6], the Roelfes lab performed experimental alanine scanning mutagenesis to systematically replace active site residues with alanine. This approach identifies amino acids critical for catalysis by measuring changes in enzymatic activity after each mutation. It is a powerful method for mapping functional hotspots in enzymes, particularly in artificial systems where non canonical residues are introduced to enable new catalytic functions.

How This Proposal Advances Previous Enzyme Engineering Efforts

Predicting the Impact of New Mutations
Previous studies have relied on experimental screening to find beneficial mutations in LmrR active site, but our approach uses EVB (Empirical Valence Bond) simulations to predict how alanine substitutions at various positions in the active site will affect enzyme activity. By building on existing experimental data from alanine scanning of LmrR in the Friedel Crafts alkylation reaction, we can test and refine our EVB model to accurately forecast changes in activation free energy for this enzyme with a non-natural catalyst. Later if our predictions go well, we can try to apply the same approach on Hydrazone reaction studied on the same enzyme by the lab before. This integrated approach allows us to move from qualitative to quantitative understanding of enzyme catalysis, supporting the design of improved biocatalysts based on mechanistic insight

Goal

For the Alanine Scanning Project (PROJECT 2): :We will use EVB in order to model the FC reaction mechanism. We will perform computational alanine scanning by EVB (Empirical Valence Bond Simulations) using the Q5/Q6 program [13]. EVB is used to calculate activation free energies and reaction profiles for wild-type and mutant enzymes. It will be first used on the FC alkylation reaction to reproduce the existing results [6] and later on hydrazone reaction, prior to stay at Roelfes lab. At Roelfes Lab we would test all these mutants experimentally using the procedure as described in ref [6].

For incorporation of new ncAA (PROJECT 1): We will use EVB to model Hydrazone reaction. Obtain activation free energies first for the LmrR_pAF and then at the same position V15 we will replace the pAF to p-Propionyl Phenylalanine and carry out EVB, to predict if it catalyses the HYdrazone Reaction prior to the stay. In Roefles lab we would want to follow the experimental procedure as stated in ref [1] to carry out gene mutagenesis, plasmid preparation Protein Expression, Purification and Activity Assays and measure how efficiently they catalyze hydrazone formation

Objectives (Project 1)

1. EVB Simulations (UVIC-UCC)
    - Predict and Compare Catalytic Potential:
      Use (Empirical Valence Bond, EVB simulations) to predict how well LmrR with different non-canonical amino acid (ncAAs) at position V15 can catalyse hydrazone formation.
    - Compare the activation energies for the original enzyme (wild type), the known variant with pAF, and new variant (such as p-Propionyl Phenylalanine).

2. Experimental Validation (Groningen)
    - Engineer Enzyme Variants:
      Incorporate the selected ncAAs into the LmrR enzyme at position V15 using genetic engineering techniques that the lab has already developed
    - Test and Compare Catalytic Activity:
      Purify the engineered enzymes and measure how efficiently they catalyze hydrazone formation

Objectives (Project 2)

1. EVB Simulations (UVIC-UCC)
    - We will use EVB (Empirical Valence Bond) free energy calculations in combination with systematic alanine scanning to identify residues essential for catalysis. Residues whose mutation to alanine leads to significant increases in the predicted activation energy will be considered critical for the reaction
    - Our first goal is to validate the EVB model by reproducing results from previous alanine scanning experiments. Once validated, we will apply the model to predict the energetic effects of mutations on the hydrazone formation reaction.
2. Experimental Validation (Groningen)
    - These computational predictions will be tested experimentally in Groningen.

Workflow (Project 1)

1. Computational Selection

- Use LmrR structure (6I8N) [11] to model incorporation of new (selected) ncAAs at V15.
- Parameterize EVB for hydrazone mechanism [1][7]
- Perform Empirical Valence Bond (EVB) simulations on pAF_LmrR and with the new ncAAs to estimate and compare activation energies

Experimental Workflow for ncAA Incorporation into LmrR [1]

1. Gene Design and Mutagenesis

- Template: Use the LmrR gene from (PDB: 6I8N) as the starting sequence.
- Site Selection: Target residue V15 for ncAA incorporation [1]
- Mutagenesis: Introduce an amber stop codon (TAG) at position V15 using site-directed mutagenesis (e.g., QuikChange protocol).

2. Plasmid Construction and Preparation

- Expression Vector: Clone the mutated LmrR gene (V15TAG) into pET17b or pET21b, as used in Roelfes' studies.
- Orthogonal tRNA/Synthetase System: Use the pEVOL-pAF plasmid for para-aminophenylalanine (pAF) incorporation, or an analogous pEVOL vector for other ncAAs.
- Co-Transformation: Prepare both plasmids for co-transformation into *E. coli* BL21(DE3).

3. Expression of LmrR Variants

- Host Strain: Transform *E. coli* BL21(DE3) with both the LmrR expression vector and the pEVOL-tRNA/synthetase plasmid.
- Culture Conditions:
    - Grow cells in LB or M9 minimal medium containing appropriate antibiotics (ampicillin for pET, chloramphenicol for pEVOL).
    - Supplement medium with the selected ncAA (e.g., 1 mM final concentration).
    - Induce protein expression with IPTG (0.5 mM) and arabinose (0.02% w/v for pEVOL), typically at 18–25°C overnight.

4. Protein Purification

- Harvesting: Collect cells by centrifugation, resuspended in a lysis buffer.
- Lysis: Lyse cells by sonication or chemical lysis.
- Purification: Purify His-tagged LmrR variants by Ni-NTA affinity chromatography.
- Analysis:
    - Assess purity by SDS-PAGE.
    - Confirm ncAA incorporation by ESI-MS or MALDI-TOF mass spectrometry.

5. Catalytic Activity Assays

- Reaction Setup:
    - Use standard hydrazone formation substrates (e.g., benzaldehyde and phenylhydrazine).
    - Typical buffer: 50 mM sodium phosphate, pH 7.0.
- Assay Conditions:
    - Incubate enzymes with substrates at room temperature.
    - Monitor product formation by UV/Vis spectroscopy (e.g., at 350 nm) or by HPLC.
- Comparison:
    - Determine turnover frequency (TOF) and compare activity of new ncAA-LmrR variants to pAF-LmrR and wild-type LmrR.

Timeline

Weeks	Step	Key Tasks
8 (prior to stay)	Computational Selection	Model ncAA at V15 in LmrR (PDB: 6I8N); EVB simulations for pAF and new ncAAs; select top candidates
3	Gene Mutagenesis	Introduce TAG at V15 (site-directed mutagenesis); order primers; verify mutation

2	Plasmid Preparation	Clone LmrR(V15TAG) into pET17b/pET21b; prep pEVOL-tRNA/synthetase plasmid for ncAA
3	Protein Expression	Co-transform BL21(DE3); grow with ncAA; induce with IPTG/arabinose at 18–25°C
3	Purification	Lyse cells; purify His-tagged LmrR by Ni-NTA; confirm by SDS-PAGE and MS
2	Activity Assays	Hydrazone assays with standard substrates; monitor by UV/Vis or HPLC; compare activities

Problems

- Specialized Resources: Incorporation of non-canonical amino acids (ncAAs) requires access to specific tRNA/synthetase systems.
- The tight schedule leaves little room for troubleshooting, optimization, or repeating experiments.

Workflow (Project 2)

Computational Alanine Scanning

- Target active site residues for alanine scanning on LmrR structure (6I8N) [10] [6]
- Parameterize and run EVB for Friedel Craft mechanism and later the Hydrazone mechanism [1][6][7]
- Rank mutants by predicted catalytic impact

Then preparing for a stay at Gerard Roelfes lab in Groningen, with the aim to validate our predicted EVB simulations.

Experimental Workflow for Alanine Scanning [6]

1. Mutagenesis & Cloning
    - Perform site-directed mutagenesis to generate active site variants of LmrR.
    - Clone the mutant LmrR genes into the pET-17b expression vector.
    - Co-transform *E. coli* with pET-17b-LmrR and a pEVOL plasmid encoding the orthogonal tRNA/synthetase for ncAA incorporation.
2. Expression & Purification
    - Express LmrR variants in *E. coli* using small-scale cultures.
    - Purify proteins using Ni-NTA affinity chromatography.
3. Activity Measurement & Data Analysis
    - Perform Hydrazone and Friedel–Crafts alkylation assays for all LmrR variants and controls in parallel.
    - Monitor product formation by UV/Vis spectroscopy or HPLC.
    - Analyze assay data to determine kinetic parameters (e.g., reaction rates, k_cat, K_M) using standard software, enabling direct comparison of mutant activities.

Timeline

Week	Step	Actions
6 (prior to the stay)	Computational Alanine Scanning	Model, simulate
1	Mutagenesis & Cloning	Generate and verify alanine mutants
1	Expression & Purification	Express and purify mutants
2	Activity Assays and Data Analysis	Run catalytic assays,collect data and write reports

References:

1. Drienovská I, Mayer C, Dulson C, Roelfes G. A designer enzyme for hydrazone and oxime formation featuring an unnatural catalytic aniline residue. Nat Chem. 2018 Sep;10(9):946-952. doi: 10.1038/s41557-018-0108-2. PMID: 30104623.

2. Roelfes G. LmrR: A Privileged Scaffold for Artificial Metalloenzymes. Acc Chem Res. 2019 Mar 19;52(3):545-556. doi: 10.1021/acs.accounts.8b00647. PMID: 30799687.
3. Drienovská I, Roelfes G. Expanding the enzyme universe with genetically encoded unnatural amino acids. Nat Catal. 2020 Feb;3:193-202. doi: 10.1038/s41929-019-0412-7.
4. Ubbink M, Roelfes G. Artificial Enzymes Based on the Lactococcal Multidrug Resistance Regulator LmrR. ChemBioChem. 2023 Jan 3;24(1):e202200486. doi: 10.1002/cbic.202200486. PMID: 36198638.
5. Casilli F, Canyelles-Niño M, Roelfes G, Alonso-Cotchico L. Computation-guided engineering of distal mutations in an artificial enzyme. Faraday Discuss. 2024 Sep 11;252(0):262-278. doi: 10.1039/d4fd00069b. PMID: 38836699; PMCID: PMC11389854.
6. Leveson-Gower RB, Zhou Z, Drienovská I, Roelfes G. Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel-Crafts Alkylase. ACS Catal. 2021 Jun 18;11(12):6763-6770. doi: 10.1021/acscatal.1c00996. Epub 2021 May 26. PMID: 34168902; PMCID: PMC8218303.
7. Leveson-Gower RB, Tiessler-Sala L, Rozeboom HJ, Thunnissen AWH, Maréchal JD, Roelfes G. Evolutionary Specialization of a Promiscuous Designer Enzyme. ACS Catal. 2025 Jan 13;15(3):1544-1552. doi: 10.1021/acscatal.4c06409. PMID: 39944761; PMCID: PMC11811923.
8. Kölmel DK, Kool ET. Oximes and Hydrazones in Bioconjugation: Mechanism and Catalysis. Chem Rev. 2017;117(15):10358-10376. doi:10.1021/acs.chemrev.7b00090
9. Madoori PK, Agustiandari H, Driessen AJ, Thunnissen AM. Structure of the transcriptional regulator LmrR and its mechanism of multidrug recognition. EMBO J. 2009 Jan 21;28(2):156-66. doi: 10.1038/emboj.2008.263. Epub 2008 Dec 18. PMID: 19096365; PMCID: PMC2634732.
10. Brouwer B, Della-Felice F, Thunnissen AWH, Roelfes G. Genetically encoded 3-aminotyrosine as catalytic residue in a designer Friedel-Crafts alkylase. Chem Sci. 2025;16(20):8721-8728. Published 2025 Mar 31. doi:10.1039/d5sc01055a
11. Mayer C, Dulson C, Reddem E, Thunnissen AWH, Roelfes G. Directed Evolution of a Designer Enzyme Featuring an Unnatural Catalytic Amino Acid. Angew Chem Int Ed Engl. 2019;58(7):2083-2087. doi:10.1002/anie.201813499
12. Wang L, Brock A, Herberich B, Schultz PG. Expanding the genetic code of Escherichia coli. Science. 2001 Apr 20;292(5516):498-500. doi: 10.1126/science.1060077. PMID: 11313494.
13. https://doi.org/10.1016/j.softx.2017.12.001 Q6: A comprehensive toolkit for empirical valence bond and related free energy calculations