Detailed Methodology

**1. Selection and Modeling**

* Review literature and select 3–5 ncAAs with suitable side chains for iminium formation.
* Build structural models of LmrR\_RMH variants with each ncAA at the catalytic site, both alone and with known distal mutations (F54L, N88Q, I62W).

**2. EVB Simulations**

* Parameterize the chosen ncAAs for use in EVB.
* Simulate the hydrazone formation reaction for each variant.
* Calculate activation energies and compare to pAF-containing enzyme.

**3. Prioritization**

* Select top candidates based on lowest activation energies and favorable reaction profiles.
* Prepare a shortlist for experimental validation.

**4. Experimental Work (in Groningen)**

* Clone, express, and purify the selected LmrR\_RMH variants.
* Perform kinetic assays for hydrazone formation (monitoring absorbance at 438 nm).
* Determine kcat, KM, and kcat/KM for each variant.

**5. Analysis and Reporting**

* Correlate computational predictions with experimental data.
* Write a detailed report on findings and propose further directions (e.g., additional mutations or reactions).

Key References

* Casilli, F., Canyelles-Niño, M., Roelfes, G., & Alonso-Cotchico, L. (2024). *Mutations in an artificial enzyme: the role of distal sites in catalysis and conformational dynamics.* Faraday Discussions, 252, 262–278.
* Roelfes, G. et al. (2023). *Biocatalytic cascades involving LmrR-based artificial enzyme.* [5](https://research.rug.nl/en/publications/biocatalytic-cascades-involving-lmrr-based-artificial-enzyme)
* Additional: Chemical Reviews 2024, "Noncanonical Amino Acids in Biocatalysis"; ACS Catalysis 2021, "Unlocking Iminium Catalysis in Artificial Enzymes"

**What is the model system in the paper?**

The paper focuses on **LmrR\_RMH**, which is:

* Based on the LmrR scaffold (Lactococcal multidrug resistance regulator)
* Contains the following mutations: **A92R, N19M, F93H**
* The catalytic residue is **V15pAF** (Valine 15 replaced by para-aminophenylalanine via amber suppression)
* Distal mutations of interest: **F54L, N88Q, I62W**

**Which PDB structure should you use?**

The authors and most LmrR artificial enzyme studies use **PDB ID: 3F8F** as the starting structure for modeling. This is the X-ray structure of wild-type LmrR.

* **PDB ID:** [3F8F](https://www.rcsb.org/structure/3F8F)
* **Title:** Crystal structure of LmrR from Lactococcus lactis
* **Resolution:** 1.9 Å

**How to prepare your model for EVB:**

1. **Download 3F8F** from the PDB.
2. **Introduce the mutations** (A92R, N19M, F93H, and your chosen ncAA at V15) using a molecular modeling tool (e.g., PyMOL, Chimera, or Modeller).
3. **Replace V15 with your selected ncAA** (not pAF), ensuring the side chain is correctly parameterized for EVB.
4. **Add any distal mutations** (F54L, N88Q, I62W) as desired.
5. **Prepare the substrate(s)** (e.g., 4-hydroxybenzaldehyde and NBD-H) and dock into the active site if needed.
6. **Parameterize the system** for EVB, including your ncAA.

**Summary Table**

| **System Name** | **Mutations** | **PDB Starting Model** |
| --- | --- | --- |
| LmrR\_RMH | A92R, N19M, F93H, V15pAF | 3F8F |
| LmrR\_RMH + Distal | + F54L, N88Q, I62W | 3F8F |
| Your Variant | A92R, N19M, F93H, V15X | 3F8F |

**Note:**

* There is no PDB structure with pAF or other ncAAs at V15; these must be modeled in silico.
* If you want to see how others have modeled pAF, check the **Supporting Information** of Roelfes' papers or contact the authors for their coordinates.

Detailed Computational Methodology

**Atomic Mechanism for LmrR\_pAF-Catalyzed Friedel–Crafts Alkylation**

**Stepwise Mechanism (Based on Literature)**

1. **Reactants:**
   * **α,β-unsaturated aldehyde** (e.g., trans-2-hexenal or methacrolein)
   * **Indole** (e.g., 2-methylindole)
   * **Catalytic residue:** pAF (para-aminophenylalanine) or your chosen ncAA at position 15
2. **Key Intermediates:**
   * **Iminium ion intermediate**: The aldehyde condenses with the pAF side chain to form an iminium ion (Schiff base).
   * **Enolate nucleophile:** Indole attacks the β-carbon of the iminium-activated enal.
3. **Product:**
   * **Friedel–Crafts alkylation product:** The indole is alkylated at the 3-position, and the iminium is hydrolyzed to regenerate the free amine on the ncAA.

**Atomic-Level Description for EVB (Qatoms)**

**Reactant State**

* **Protein:** LmrR\_RMH variant (with ncAA at position 15)
* **Substrate 1:** α,β-unsaturated aldehyde (e.g., methacrolein)
* **Substrate 2:** Indole (e.g., 2-methylindole)
* **Catalytic residue:** Free amine group (on pAF or your ncAA)

**Product State**

* **Protein:** LmrR\_RMH variant (with ncAA at position 15, amine regenerated)
* **Product:** 3-alkylated indole (Friedel–Crafts product)
* **Byproduct:** Water (from hydrolysis of iminium)

**Key Atoms (Qatoms) to Include**

* **Aldehyde carbonyl carbon and oxygen**
* **Aldehyde β-carbon (site of C–C bond formation)**
* **pAF (or ncAA) side chain nitrogen**
* **Indole C3 (site of nucleophilic attack)**
* **Hydrogen atoms involved in proton transfers**
* **Relevant backbone atoms if needed for geometry constraints**

**Mechanistic Scheme (from literature)**

1. **Iminium Formation:**

pAF-NH2+R-CH=CH-CHO→pAF-N=CH-CH=CHR+H2OpAF-NH2+R-CH=CH-CHO→pAF-N=CH-CH=CHR+H2O

1. **Friedel–Crafts Alkylation:**

pAF-N=CH-CH=CHR+Indole→pAF-NH2+Friedel–Crafts productpAF-N=CH-CH=CHR+Indole→pAF-NH2+Friedel–Crafts product

**References for Mechanism and Structures**

* **ACS Catalysis 2021, "Unlocking Iminium Catalysis in Artificial Enzymes to Create a Friedel–Crafts Alkylase"**[2](https://pubs.acs.org/doi/10.1021/acscatal.1c00996)[3](https://pmc.ncbi.nlm.nih.gov/articles/PMC8218303/)[6](https://chemrxiv.org/engage/api-gateway/chemrxiv/assets/orp/resource/item/60c7553df96a0041d12887fb/original/unlocking-iminium-catalysis-in-artificial-enzymes-to-create-a-friedel-crafts-alkylase.pdf)
* **Supporting Information of the above paper** contains atomistic models and sometimes coordinate files for the iminium intermediate and transition state analogs.
* **Roelfes group website** for more structures and mechanistic insights[5](https://sites.google.com/rug.nl/roelfesgroup/publications/artificial-enzymes).

**Summary Table: Reactant and Product States for EVB**

| **State** | **Key Species Involved** | **Description** |
| --- | --- | --- |
| Reactant | LmrR\_RMH (ncAA at 15), enal, indole | Free amine, aldehyde, indole |
| Product | LmrR\_RMH (ncAA at 15), Friedel–Crafts product | Amine regenerated, 3-alkylated indole, water |

**In summary:**  
You should model the **reactant state** as the free amine (on your ncAA) with the enal and indole, and the **product state** as the regenerated amine, with the Friedel–Crafts adduct and water. The **Qatoms** should include the atoms involved in iminium formation and C–C bond formation, as described above and in the cited papers

**1. Define the Chemical Mechanism**

**Hydrazone formation** is a nucleophilic addition–elimination reaction between an aldehyde (or ketone) and a hydrazine (or hydrazide):

* **Reactants:** Aldehyde (e.g., 4-hydroxybenzaldehyde) + Hydrazine derivative (e.g., NBD-H)
* **Product:** Hydrazone + Water

**Mechanistic Steps:**

1. Nucleophilic attack of the hydrazine nitrogen on the carbonyl carbon of the aldehyde, forming a tetrahedral intermediate.
2. Proton transfers and elimination of water, yielding the hydrazone.

The rate-limiting step is typically the breakdown of the tetrahedral intermediate to release water and form the C=N bond[7](https://pubs.acs.org/doi/10.1021/ol500262y).

**2. Define the EVB States**

You need to define at least two EVB states:

* **State 1 (Reactant State):**
  + Aldehyde and hydrazine are not covalently bonded (or have only a weak association).
* **State 2 (Product State):**
  + Hydrazone is formed (C=N bond), water is released.

**Qatoms (Quantum Atoms):**  
Include all atoms directly involved in bond making and breaking:

* Carbonyl carbon and oxygen of the aldehyde
* Nitrogens of the hydrazine
* Hydrogens involved in proton transfer
* Any side-chain atoms of the catalytic residue (e.g., pAF or other ncAA) if they participate directly

**3. Build the Structural Models**

* **Start from the enzyme structure (e.g., LmrR variant with ncAA at position 15, using PDB 3F8F or 6I8N as template)**
* **Dock the aldehyde and hydrazine substrates into the active site** so that the nucleophilic nitrogen is close to the carbonyl carbon.
* **Generate coordinates for both EVB states** (reactant and product) using quantum chemical calculations or by manually editing the structure.

**4. Parameterize the EVB States**

* **Assign force field parameters** for both states (bond lengths, angles, charges, etc.).
* **Fit the energy gap between the two states** to reproduce quantum chemical or experimental activation and reaction energies[3](http://ndl.ethernet.edu.et/bitstream/123456789/5191/1/Fernanda%20Duarte_2017.pdf)[4](https://lirias.kuleuven.be/retrieve/478543).
* **Include the relevant solvent model** (e.g., TIP3P water sphere as in [1](https://pubs.acs.org/doi/10.1021/acsomega.8b00346)).

**5. Simulation Setup**

* **Solvate the system** (e.g., in a water sphere or box).
* **Minimize and equilibrate** the system.
* **Run EVB simulations** along the reaction coordinate (energy gap between states), typically using umbrella sampling or similar techniques to map the free energy profile[1](https://pubs.acs.org/doi/10.1021/acsomega.8b00346)[2](https://www.diva-portal.org/smash/get/diva2:1217357/FULLTEXT01.pdf)[3](http://ndl.ethernet.edu.et/bitstream/123456789/5191/1/Fernanda%20Duarte_2017.pdf)[4](https://lirias.kuleuven.be/retrieve/478543).
* **Analyze the free energy surface** to extract activation barriers and reaction free energies.

**6. Key References and Methodology**

* **General EVB methodology:**
  + Duarte, F. (2017). *Theory and Applications of the Empirical Valence Bond Approach*[3](http://ndl.ethernet.edu.et/bitstream/123456789/5191/1/Fernanda%20Duarte_2017.pdf)
  + Harvey et al. (2014). *Empirical Valence Bond Methods for Exploring Reaction Dynamics*[4](https://lirias.kuleuven.be/retrieve/478543)
* **Practical enzyme example:**
  + Kamerlin et al. (2018). *Empirical Valence Bond Simulations Suggest a Direct Hydride Transfer Mechanism for Human Diamine Oxidase* (ACS Omega)[1](https://pubs.acs.org/doi/10.1021/acsomega.8b00346)[2](https://www.diva-portal.org/smash/get/diva2:1217357/FULLTEXT01.pdf)
* **Hydrazone mechanism:**
  + Fast Alpha Nucleophiles: Structures that Undergo Rapid Hydrazone Formation (ACS Org Lett, 2014)[7](https://pubs.acs.org/doi/10.1021/ol500262y)

**Summary Table: EVB Setup for Hydrazone Formation**

| **Step** | **Details** |
| --- | --- |
| Mechanism | Aldehyde + hydrazine → hydrazone + water |
| EVB States | State 1: Reactants (separate), State 2: Product (hydrazone + water) |
| Qatoms | Aldehyde C, O; hydrazine N, H; any directly involved side-chain atoms |
| Structure | Enzyme + substrates (reactant); enzyme + hydrazone + water (product) |
| Parameterization | Fit to reproduce quantum chemical/experimental energies |
| Simulation | Solvate, minimize, sample along reaction coordinate, analyze free energy profile |