Supporting Information

Computational Design of an Imine Reductase: Mechanism-Guided Stereoselectivity Reversion and Interface Stabilization

Kai Wu[a], Jinrong Yan[b,d], Qinde Liu[a,c], Xiaojing Wang[a], Piaoru Wu[a], Yiyang Cao[a], Xiuhong Lu[a], Yixin Xu[a], Junhai Huang[b,d]*, Lei Shao[a,d]*

a. School of Pharmacy, Shanghai University of Medicine & Health Sciences, 279 Zhouzhu Highway, Pudong New Area, Shanghai 201318, China

b. College of Chemistry and Chemical Engineering, Shanghai University of Engineering Science, 333 Longteng Road, Shanghai 201620, China

c. Shanghai University of Traditional Chinese Medicine, 1200 Cailun Road, Shanghai 201203, China

d. State Key Laboratory of New Drug and Pharmaceutical Process, Shanghai Institute of Pharmaceutical Industry, 285 Gebaini Rd., Shanghai 200040, China

1 Material and methods

1.1 General information

Substrate 2-(2,5-difluorophenyl)-pyrroline (2-DFPL) was purchased from Anhui Dexingjia biopharm Co., Ltd (Anhui, China). Substrate myosmine, 2-phenyl-1-pyrroline were purchased from Shanghai Bepharm Science & Technology Co., Ltd (Shanghai, China). ^{1}H NMR and ^{13}C NMR were measured on a Bruker Avance 600 MHz spectrometer.

1.2 Chemical synthesis of substrates and product racemic standards

Synthesis of 5-(2-fluorophenyl)-3,4-dihydro-2H-pyrrole: 1-Bromo-2-fluorobenzene (3 g) was dissolved in THF (30 mL), and the mixture was cooled to approximately 0 °C. Then, a solution of 2.0 M i-PrMgCl in THF (8.5 mL) was added over a 10-minute period while maintaining the reaction temperature at 0 °C. The solution was stirred at about 0 °C for 1 hour. Subsequently, a solution of tert-butyl 2-oxopyrrolidine-carboxylate (2.6 g) in 30 mL THF was added over approximately 15 minutes while keeping the reaction temperature at 0 °C. The reaction mixture was stirred below 25 °C for 4 hours and then quenched by adding saturated NH$_4$Cl solution while maintaining the reaction temperature at 0 °C. The resulting mixture was transferred to a separatory funnel and extracted three times with ethyl acetate. The organic layer was dried using anhydrous Na$_2$SO$_4$ and concentrated to an oil.
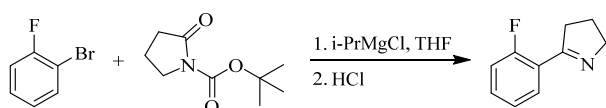
The oil was supplemented with 1,4-dioxane (3 mL) and 4.0 M HCl-1,4-dioxane solution (35 mL), prompting a mild exothermic reaction and gas evolution. The mixture was stirred at room temperature overnight. Subsequently, the mixture was concentrated to yield a solid, which was then washed with acetone. After filtration, the resulting solid was combined with MeOH (30 mL), and the pH was adjusted to 10 using a saturated NaOH solution. The mixture was stirred at room temperature for more than 5 hours. The solution was then subjected to vacuum concentration to remove the MeOH solvent, followed by the addition of H$_2$O (50 mL). The pH of the mixture was adjusted to 3

using an HCl solution and subsequently extracted three times with ethyl acetate. The pH of the aqueous phase was adjusted to 10 with a saturated NaOH solution, followed by another round of extraction with ethyl acetate. The organic layer was dried over anhydrous Na_2SO_4, and the solvent was then concentrated to yield an oil. The resultant oil was dissolved in petroleum ether and collected, ultimately yielding a concentrated yellow oil.

5-(2-fluorophenyl)-3,4-dihydro-2H-pyrrole, 43.6 % yield (1.0 g), yellow oil. ^{1}H-NMR (600 MHz, Chloroform-d): δ 7.94 (td, J = 7.7, 1.8 Hz, 1H), 7.41–7.34 (m, 1H), 7.19–7.14 (m, 1H), 7.12–7.05 (m, 1H), 4.04–3.98 (m, 2H), 3.04–2.97 (m, 2H), 2.05–1.98 (m, 2H). ^{13}C-NMR (151 MHz, Chloroform-d): δ 170.6, 161.4 (d, J = 250.8 Hz), 131.8 (d, J = 8.5 Hz), 130.1, 124.2 (d, J = 3.5 Hz), 122.8 (d, J = 12.0 Hz), 116.3 (d, J = 22.5 Hz), 60.7, 37.7, 22.9.

Synthesis of 5-(3-fluorophenyl)-3,4-dihydro-2H-pyrrole: Methyl 3-fluorobenzoate (12 g) was added to a solution of NaH (7.8 g, 60%) in anhydrous THF (200 mL) with mechanical agitation. The mixture was heated to 60 °C, and then N-vinylpyrrolidone (7.2 g) in 30 mL THF was added dropwise. The resulting mixture was further heated at 72 °C for 5 hours. After cooling the reaction mixture to room temperature, it was poured into ice water and extracted three times with ethyl acetate. The organic layer was concentrated, resulting in an oil.

The oil was mixed with THF (200 mL) and 6 M HCl solution (110 mL) under reflux conditions. After refluxing the reaction mixture for 12 hours, it was cooled to room temperature. The solution's pH was adjusted to 3 using saturated NaOH solution, and it was then extracted three times with ethyl acetate. The pH of the aqueous phase was adjusted to 10 using saturated NaOH solution and then extracted three times with ethyl acetate. The organic layer was dried using anhydrous Na_2SO_4 and concentrated to
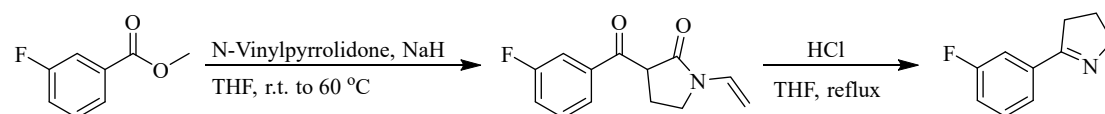
an oil. The oil was dissolved using petroleum ether and collected. After concentration, it resulted in a yellow oil.

5-(3-fluorophenyl)-3,4-dihydro-2H-pyrrole, 77.5 % yield (8.2 g), yellow oil. ^{1}H-NMR (600 MHz, Chloroform-d): δ 7.61–7.52 (m, 2H), 7.40–7.33 (m, 1H), 7.15–7.08 (m, 1H), 4.09–4.05 (m, 2H), 2.94–2.90 (m, 2H), 2.08–2.01 (m, 2H). ^{13}C-NMR (151 MHz, Chloroform-d): δ 172.3, 162.8 (d, J = 245.9 Hz), 136.9 (d, J = 7.4 Hz), 130.0 (d, J = 7.6 Hz), 123.4 (d, J = 2.4 Hz), 117.2 (d, J = 21.7 Hz), 114.4 (d, J = 22.6 Hz), 61.6, 35.0, 22.7.

Synthesis of 5-(4-fluorophenyl)-3,4-dihydro-2H-pyrrole: Methyl 4-fluorobenzoate (12 g) was added to a solution of NaH (7.8 g, 60%) in anhydrous THF (200 mL) under mechanical agitation. The mixture was heated to 60 °C, and then N-vinylpyrrolidone (7.2 g) in 30 mL THF was added dropwise. The resulting mixture was further heated to 72 °C for 5 hours. Afterward, the reaction mixture was cooled to room temperature and poured into ice water before being extracted with ethyl acetate. The organic layer was concentrated to yield an oil.
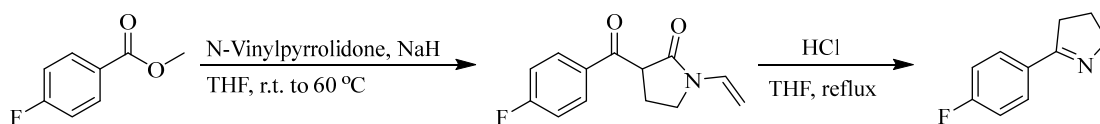

The oil was mixed with THF (200 mL) and 6 M HCl solution (110 mL) under reflux conditions. After refluxing the reaction mixture for 12 hours, it was cooled to room temperature. The solution's pH was adjusted to 3 using saturated NaOH solution, and it was then extracted three times with ethyl acetate. The pH of the aqueous phase was adjusted to 10 using saturated NaOH solution and then extracted three times with ethyl acetate. The organic layer was desiccated using anhydrous Na_2SO_4 and then concentrated to yield a solid. This solid was dissolved in petroleum ether and collected. After further concentration, a yellow solid was obtained.

5-(4-fluorophenyl)-3,4-dihydro-2H-pyrrole, 79.5 % yield (8.4 g), yellow solid. ^{1}H-NMR (600 MHz, Chloroform-d): δ 7.85–7.71 (m, 2H), 7.10–7.06 (m, 2H), 4.07–4.03

(m, 2H), 2.94–2.89 (m, 2H), 2.07–2.01 (m, 2H). ^{13}C-NMR (151 MHz, Chloroform-d): δ 172.1, 164.1 (d, J = 249.8 Hz), 130.9 (d, J = 3.5 Hz), 129.6 (d, J =8.5 Hz), 115.4 (d, J =21.7 Hz), 61.5, 35.0, 22.8.

Synthesis of 3-(3,4-dihydro-2H-pyrrol-5-yl)-5-fluoro-2-methoxypyridine: 3-Bromo-5-fluoro-2-methoxypyridine (10 g) was dissolved in THF (100 mL), and the mixture was cooled to approximately 0 °C. A solution of 2.0 M i-PrMgCl in THF (24 mL) was then added over 15 minutes while maintaining the reaction temperature at 0 °C. The solution was cooled to around 0 °C and stirred for 1 hour. Subsequently, a solution of tert-butyl 2-oxopyrrolidine-carboxylate (7.5 g) in 75 mL THF was added over approximately 30 minutes while maintaining the reaction temperature at 0 °C. The reaction was stirred below 25 °C for 6 hours and subsequently quenched with saturated NH$_4$Cl solution while keeping the reaction temperature at 0 °C. The mixture was then extracted with ethyl acetate. The organic layer was dried over anhydrous Na$_2$SO$_4$ and concentrated to yield an oil.
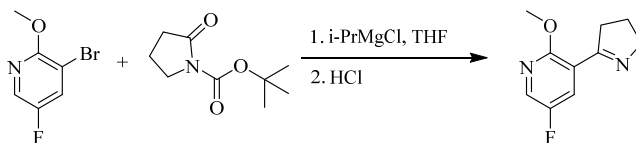
The oil was mixed with 1,4-dioxane (5 mL) and 4 M HCl-1,4-dioxane solution (100 mL) resulting in a mild exothermic reaction and gas evolution. The reaction mixture was stirred at room temperature overnight. Subsequently, the mixture was concentrated to an oil, which was then supplemented with MeOH (60 mL), and the pH was adjusted to 10 using saturated NaOH solution. The reaction was stirred at room temperature for over 5 hours. The mixture was concentrated under vacuum to remove the MeOH solvent and then supplemented with H$_2$O. The pH of the mixture was adjusted to 3 with HCl solution and extracted with ethyl acetate. The pH of the aqueous phase was adjusted to 10 with saturated NaOH solution and further extracted with ethyl acetate. The organic layer was dried over anhydrous Na$_2$SO$_4$ and concentrated using rotary evaporation. The resulting mixture was purified by silica gel column chromatography (petroleum ether/ethyl acetate = 2:1) and concentrated to yield a white solid.

3-(3,4-dihydro-2H-pyrrol-5-yl)-5-fluoro-2-methoxypyridine, 14.0% yield (1.1 g), white solid. ^{1}H-NMR (600 MHz, Chloroform-*d*): δ 8.06 (d, *J* = 2.5 Hz, 1H), 7.17 (d, *J* = 4.9 Hz, 1H), 4.06 (t, *J* = 7.5 Hz, 2H), 3.92 (s, 3H), 2.99–2.94 (m, 2H), 2.07–2.01 (m, 2H). ^{13}C-NMR (151 MHz, Chloroform-*d*): δ 169.1 (d, *J* = 3.4 Hz), 160.8, 153.4 (d, *J* = 250.8 Hz), 134.8 (d, *J* = 28.0 Hz), 133.1 (d, *J* = 12.4 Hz), 109.7, 61.4, 54.0, 37.4 (d, *J* = 5.2 Hz), 22.71.

Synthesis of racemic standards: The imine substrate was dissolved in methanol, and sodium borohydride was added under an ice-water bath. After stirring at room temperature overnight, the reaction mixture was acidified with HCl. Methanol was removed by rotary evaporation. The resulting solution was then made basic with NaOH and extracted three times with ethyl acetate. The combined organic extracts were dried with sodium sulfate, filtered, and concentrated to obtain the racemic product.

1.3 Enzyme Expression, Mutagenesis, and Purification

The *Pm*IR (GenBank: WP_016362380.1) from *Paenibacillus mucilaginosus* was synthesized by GENEWIZ Co. Ltd. and inserted into pET-28a (+). The resulting recombinant plasmid containing the *Pm*IR gene was then transformed into BL21(DE3) for expression. A single transformant was cultured at 37°C for 12 h and was then transferred to 100 mL fresh Luria–Bertani medium and cultured at 37°C. The culture was induced by adding 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when its optical density at 600 nm reached 0.6-0.8. After induction at 20°C for 20 h, resting cells were harvested by centrifuging at 8000 × g for 10 min and were then resuspended in 20 mM PBS (pH 7.4) with a cell concentration of 50 mg·mL^{-1}.

The recombinant plasmid was used as the PCR template for site directed mutagenesis. The obtained PCR products were digested using *Dpn*I to remove parent plasmids. The digested PCR products were transformed into *Escherichia coli* DH5α cells. Plasmids containing the mutant gene were then extracted and retransformed into *E. coli* BL21 (DE3) cells for mutant enzyme expression.

The *Pm*IR and mutants were purified for the determination of kinetic parameters, substrate scope, melting temperatures (T_m) evaluation. After disrupting the cells by

sonication and removing cell debris/inclusion bodies by centrifugation, the soluble cell-free extract was filtered (Millipore filtration, 0.22 μm) and loaded onto a nickel column preequilibrated with binding buffer (20 mM PBS, pH 7.4). After washing with binding buffer, the bound recombinant enzyme was eluted with binding buffer and increasing concentrations of imidazole (20–200 mM). The pooled elution fragments contained the desired proteins were concentrated by ultrafiltration (10 kDa, 4°C, Millipore). Samples was subsequently loaded on a Superdex 75 prepacked gel filtration column (120 mL, GE Healthcare), eluted with Superdex buffer (20 mM PBS, pH 7.4; NaCl, 150 mM; dithiothreitol, 1 mM) and concentrated to 20 mg/mL by ultrafiltration (MWCO: 10 kDa, Millipore) at 4°C for crystallization experiments.

1.4   Protein crystallization and structure determination

Initial screening of crystallization conditions was performed using the sitting drop vapor-diffusion method in 96-well plates. Well grown crystals were obtained by mixing 2 μL protein sample with 2 μL reservoir solution (2 M Ammonium sulfate, 0.1 M sodium acetate pH 5.5), after cultivation in 24-well sitting drop plates at 18°C, in dark for 2-3 weeks. Crystals were transferred into a pool which containing their corresponding reservoir solution and PEG 400 (20%, *v/v*), then flash-cooled in liquid nitrogen by using nylon CryoLoops[TM] (Hampton Research, USA).

Data were collected at 100 K, with the wavelength of 0.9792 Å by using a DECTRIS PILATUS3 6M detector at beamline BL19U1 of the Shanghai Synchrotron Radiation Facility (SSRF), data were processed with XDS[1]. The crystal structure of imine reductase *Pm*IR was solved by molecular replacement by employing CCP4 Suite[2], and the peptide chain A was extracted from the structure of imine reductase from BcSIRED from *Bacillus cereus* (PDB ID: 4D3F) as input template. Structure refinement was processed with COOT[3] and phenix[4]. Relevant statistics and X-ray Diffraction data quality analysis were summarized in Table S1.

1.5   Enzyme activity assay, biotransformation, and kinetic parameters determination

Specific activities and relative activities of *Pm*IR, *Pm*IR-Re and *Pm*IR-6P were

spectrophotometrically assayed by measuring the change in NADPH absorbance at 340 nm in 1 min. The reaction system contained the appropriate weights of the purified enzymes, 1.5 mM NADPH, 5 mM substrates, and 100 mM PBS (pH 6.0) buffer, at a final reaction volume of 200 μL

For comparation of the conversion, the reaction mixture (500 μL) included 275 μL crude enzyme, 50 μL of 2.5 mM NADP stock, 50 mM glucose, 10 mM substrate, 100 μL crude enzyme of recombinant glucose dehydrogenase (GDH) from *Bacillus subtilis* CGMCC 1.1398. The reaction mixture was incubated at 30°C with rotation at 200 rpm. The reaction was terminated by adding 1 mL isopropanol. The conversion and *ee* were analyzed by chiral HPLC (Table S7).

The kinetic parameters of *Pm*IR, *Pm*IR-Re and *Pm*IR-6P to 2-DFPL were determined by increasing the substrate concentration from 0.075 to 3.25 mM at their optimum temperature. Initial velocities at different substrate concentrations were used to generate Michaelis–Menten saturation curve.

1.6   Enzyme thermostability assay

The experimental validation was performed by comparison of the residual activity of WT and variants. The crude enzyme was firstly incubation in 60°C for 15 minutes and then compared the conversion at 30 °C. The variants with higher conversion were further purified to determine T_m.

The T_m was determined by the thermofluor method. Purified enzymes were incubated in100 mM PBS buffer pH 6.0 together with 0.1 μL of SYPRO® Orange dye at a total volume of 20 μL in a 384-Well PCR Reaction Plate. The plate was heated from 25 °C to 99 °C at 0.05 °C/s while monitoring fluorescence in Applied Biosystems QuantStudio 7 Flex RT-PCR machine.

The half-life of WT and *Pm*IR-6P were determined the residual activity over 120h incubation at 35 °C and 40 °C. The purified enzymes were used for residual activity determination using the standard activity assay method described above. The initial specific activity detected before incubation was normalized as 100%. The half-lives were calculated according to first-order exponential decay function.

1.7   Docking and molecular dynamic (MD) simulations

The PmIR-NADPH complex model was constructed by superimposing PmIR crystal structure (PDB ID: 8KFK) with the NADPH-bound BcSIRED crystal structure (sequence identity 75.8%) due to the conservative binding position of NADPH in IREDs. Discovery Studio 4.0 (Biovia, USA) were used for complex and substrate preparation, docking and ligand-receptor interaction analysis. To resolve steric clashes between PmIR and NADPH, energy minimization was further performed on NADPH and residues within 4 Å to achieve reasonable conformations. Substrate 2-DFPL in iminium cation form was docked into the binding pocket. Residues including D250 and Y187E were assigned as flexible residues. The pK_a of titratable residues was estimated by the pK_a predicting module in Discovery Studio 4.0. According to the optimal reaction pH of PmIR (pH 6.0) and pK_a of acidic residues, the acidic residue in binding pocket was set as deprotonated.

MD simulation was performed using the Particle Mesh Ewald Molecular Dynamics module implemented in the Amber 18 suite[6]. The ff14SB force field was used for the protein system as it was reported with increased accuracy of protein side chain[7], and the GAFF force field used for the ligands[8]. The ANTECHAMBER module and Gaussian 16 were used to calculate the substrate and NADPH RESP atom charges[8]. Hydrogen atoms and sodium ions (to neutralize the negative charges) were added to the protein using the tleap utility. Each simulation system was immersed in a cube of TIP3P explicit water, extending to 12 Å outside the protein on all sides. Water molecules were treated using the SHAKE algorithm, and the long-range electrostatic effects were considered using the particle mesh Ewald method. The non-bonded cutoff distance was set as 12 Å. A three-stage energy minimization was performed. The steepest descent algorithm for the first 9000 steps, then switches the conjugate gradient algorithm for another 1000 steps. The water molecules and ions were relaxed to minimize the energy during the 10,000 minimization steps with the protein and ligands restrained. Then the backbone of the protein was restrained with the other section relaxed to minimize the energy during the 10,000 minimization steps. In the last stage, the whole system was

minimized without the restraints during the 10,000 minimization steps. After energy minimization, the system was gradually heated in the NVT ensemble from 0 to 303 K over 200 ps. This procedure was followed by 200 ps of NPT simulation at 303 K and 1 atm pressure using the Langevin dynamics algorithm with the complex constrained. All the positional constraints used a force constant of 2.0 kcal mol^{-1} Å^{-2}.

In order to analyze the occurrence of reactive conformation, three independent productive simulations of the variants were conducted for 2 ns without any restraints. The trajectories obtained from these simulations were used to monitor two key distances: *d1*, which represents the distance between the imine nitrogen atom and the carboxyl carbon atom of D250 or E187, and *d2*, which represents the distance between the imine carbon atom and the NADPH-C4 atom. Conformations of 2-DFPL with *d1* ≤ 5 Å and *d2* ≤ 4 Å were considered reactive. The distribution plot of these distances was generated using OriginPro 2021. Variants that exhibited a higher proportion of reactive conformations included *Pm*IR-Re and variant_5352 were selected for further experimental validation.

To narrow down the variant number for experimental validation, we performed MD simulations and further analysis on variants predicted by FoldX and Rosetta_ddg below the cutoff value (refer to section 1.10). Each variant structure performed five independent 3 ns productive MD simulations at 303 K. It was reported multiple short trajectories provided a better sampling of conformations than a single long MD simulation[9-11]. Trajectories from 2 to 3 ns were gathered and analyzed for average structure and root mean square fluctuation (rmsf) calculation. The average structure was visually inspected and compared to that of the WT crystal structure using PyMOL with a python script described by Yinglu Cui et.al[15]. Variants were discarded if a loss of polar contacts (hydrogen bonds or salt bridges) or exposed hydrophobic residues was observed. Besides, the B-factor value, converted from the RMSF of each residue, was normalized into z-scores. The z-score formula is $z = (x-\mu)/\sigma$, where x is the B-factor of each residue, μ is the mean, and σ is the standard deviation. The value of the z-score indicates that how many standard deviations are away from the mean. A positive z-

score indicated residue flexibility higher than the mean and a negative z-score revealed it below the mean. The z-score, being standardized, allows for rough comparison of flexibility between a specific mutation and its template residue in different overall background movements, even when the two datasets have different means and standard deviations. Thus, a local mutation with a Δz-score > 0 (score$_{mutation}$ - score$_{wt}$) suggested a greater flexibility than average compared with original residue. The short MD simulations can serve as a preliminary filter to identify variants that significantly destabilize. Variants that did not reveal obvious structural undesirable mutations and with a Δz-score < 0 were selected for experimental characterization. (Table S6). Example of for stabilized variants screening is illustrated in figure S7.

To compare the compactness, flexibility of WT and *Pm*IR-6P. The temperature of the MD simulation was increased to 308 K. After a 200 ns productive MD simulation, the radius of gyration, RMSF was calculated with the cpptraj module. The distances between Cα of residue 87 and residue 257, as well as Cα of residue 217 and Cα of residue 176 of *Pm*IR-6P and WT were calculated with the cpptraj module.

1.8   QM calculation for transition state structure optimization

The resulted cluster model with reactive pro-*S* pose obtained from MD simulation was further used for transition state optimization with Gaussian 16[12]. Apart from the substrate and the NADPH cofactor, the model consists of the amino acids around substrate within 4 Å (F232, M254, S258, S112, V137, Q138, P140, E187, Q190, M191, F194, W195). At the pH 6.0 of the experiments, the most 2-DFPL was expected to be in the protonated form, and it was therefore modeled as iminium ion in the calculations while the side chain of E187 was also deprotonated. The NADPH cofactor and the amino acids were truncated as shown in Figure S4, and hydrogen atoms were added manually to saturate the carbon atoms. The carbon atom where the truncation was made, and one of its hydrogens, were kept fixed during geometry optimizations.

B3LYP was widely used in the search for transition state geometries for organic reactions, and it can generate results that are almost as reliable as much more expensive computational methods[13]. Polarization functions were added to light atoms. Geometry

structure optimization was performed using the B3LYP functional with the basis set 6-31G(d,p). Dispersion effects were evaluated using the B3LYP-D3(BJ) method. The obtained first-order saddle point was confirmed using vibrational frequency analysis and the intrinsic reaction coordinate calculation. The corresponding transition state structure was used as the reference for catalytic restraints in following Rosetta design (cartesian coordinates of TS structure in pro-*S* pose is listed in section 6).

1.9   Stereoselectivity reversion using Rosetta Design

The Y187E ternary complex with pro-*S* pose was further performed 2ns MD simulations for three times, and snapshots with reactive conformation was selected for transition state optimization and optimize the catalytic interactions with Rosetta Enzyme Design application. The geometric criteria of the transition state structure obtained from the *Pm*IR pro-*S* complex were used as a reference for geometric constraints determination (refer to section 1.8). The geometric criteria for catalytic constraint of distance, angle and dihedrals for stereoselectivity reversion were detailed in Table S3. The user-defined options were as follows: -detect_design_interface -cut1 0.0 -cut2 0.0 -cut3 8.0 -cut4 10.0 -cst_min -chi_min -bb_min -packing:use_input_sc -packing:soft_rep_design -extrachi_cutoff 1 -design_min_cycles 3 –ex1:level 4 –ex2: level 4 –ex1aro:level 4 –ex2aro:level 4 –no_optH false –no_his_his_pairE -flip_HNQ -ignore_unrecognized_res -extrachi_cutoff 1. The multiple conformations of substrate were produced with Open Babel. The residues of V249 (A), D250 (A), R251(A), Q138(B), P140(B), Q190(B) were selected for mutagenesis and repacking to optimize the interactions. The variant structures produced by Rosetta enzyme design were sorted, and designed structures with total score < -1645.0, all constraint energy < 10 kcal/mol and interface energy < -8.0 kcal/mol were selected for further screening with MD simulation to evaluate the frequency of reactive conformations.

1.10 Computational stabilization of *Pm*IR

A combinational stabilization strategy consisted of un/folding energy calculation and salt bridge design was used to predict beneficial mutations (Figure S6). For folding energy ($\Delta\Delta G^{fold}$) calculation, *Pm*IR crystal structure was used for virtual saturation

mutagenesis in each site of *Pm*IR. The relative folding free energy changes ($\Delta\Delta G^{Fold}$) predicted by the FoldX and Rosetta_ddg algorithms were calculated using as follows:

$$\Delta\Delta G^{Fold} = \Delta G^{Fold}_{mutation} - \Delta G^{Fold}_{WT}.$$

The G^{Fold} represents the free energy difference between the folded and unfolded structures. For FoldX[14], standard settings were used, and each calculation was repeated five times to obtain better averaging. We used the settings described by Yinglu Cui et.al previously for Rosetta_ddg (options -ddg::local_opt_only true -ddg::opt_radius 8.0 -ddg::weight_file soft_rep_design -ddg::iterations 50 -ddg::min_cst false -ddg::mean true -ddg::min false -ddg::sc_min_only false -ddg::ramp_repulsive false) [15].

The difference of $\Delta\Delta G^{fold}$ caused by mutation was calculated with FoldX (cut-off < -1.5 kJ/mol)[10] and rosetta cartesian_ddG (cut-off < -1.0 kJ/mol), respectively. To eliminate mutations that are unlikely to be stabilizing, the predicted variants were further screened by MD simulations and visual inspection (refer to section 1.7). The variants selected for further experimental validation were listed in table S6.

Based on the geometric requirement of salt-bridge (Figure S8)[16], the potential sites for salt bridge construction were confirmed by a python script listed below. Besides, the B-factor values of predicted sites from crystal structure were normalized to their respective Z-scores by a python script list below. The residues with Z-score >0 (more flexible than average) and located at surface of the two subunit were selected for salt-bridge construction and experimental validation.

```python
# python script for salt bridge search
import sys
import math
from Bio import PDB
parser=PDB.PDBParser()
s=parser.get_structure("name", "complex.pdb")
first_model=s[0]
chain_A=first_model["A"]
chain_B=first_model["B"]
file=open('salt_bridge.txt','a')
charpo=["SER","THR","CYS","ASN","GLN","LYS","ARG","ASP","GLU"]
for res1 in chain_A:
    if res1.get_resname()in charpo:
        for res2 in chain_B:
```

```python
            if res2.get_resname()!="GLY":
                d=res1["CA"]-res2["CA"]
                vector1=res1["CB"].get_vector()
                vector2=res1["CA"].get_vector()
                vector3=res2["CA"].get_vector()
                vector4=res2["CB"].get_vector()
                angle1=PDB.calc_angle(vector1, vector2, vector3)
                angle2=PDB.calc_angle(vector4, vector3, vector2)
                if 3.7<d<7 and 0<angle1<1.57 and 0<angle2<1.57:
                    print
(res1.resname,res1.get_id()[1],res2.resname,res2.get_id()[1],format(d,'
.2f'),format(math.degrees(angle1),'.2f'),format(math.degrees(angle2),'.
2f'))
                    file.write(res1.resname + str(res1.get_id()[1]) + "
" + res2.resname + str(res2.get_id()[1]) + " " + str(format(d,'.2f')) +
" "  + str(format(math.degrees(angle1),'.2f')) + " " +
str(format(math.degrees(angle2),'.2f'))+'\n')
                elif 7<d<14 and 0<angle1<0.9599 and 0<angle2<0.9599:
                    print
(res1.resname,res1.get_id()[1],res2.resname,res2.get_id()[1],format(d,'
.2f'),format(math.degrees(angle1),'.2f'),format(math.degrees(angle2),'.
2f'))
                    file.write(res1.resname + str(res1.get_id()[1]) + "
" + res2.resname + str(res2.get_id()[1]) + " " + str(format(d,'.2f')) +
" "  + str(format(math.degrees(angle1),'.2f')) + " " +
str(format(math.degrees(angle2),'.2f'))+'\n')
            else:
                n = res2["N"].get_vector()
                c = res2["C"].get_vector()
                ca = res2["CA"].get_vector()
                n = n - ca
                c = c - ca
                rot = PDB.rotaxis(-3.1415926 * 120.0/180.0, c)
                cb_at_origin = n.left_multiply(rot)
                cb_2 = cb_at_origin+ca
                d=res1["CA"]-res2["CA"]
                vector1=res1["CB"].get_vector()
                vector2=res1["CA"].get_vector()
                vector3=res2["CA"].get_vector()
                angle1=PDB.calc_angle(vector1, vector2, vector3)
                angle2=PDB.calc_angle(cb_2, vector3, vector2)
                if 3.7<d<7 and 0<angle1<1.57 and 0<angle2<1.57:
                    print
(res1.resname,res1.get_id()[1],res2.resname,res2.get_id()[1],format(d,'
```

```
.2f'),format(math.degrees(angle1),'.2f'),format(math.degrees(angle2),'.
2f'))
                    file.write(res1.resname + str(res1.get_id()[1]) + "
" + res2.resname + str(res2.get_id()[1]) + " " + str(format(d,'.2f')) +
" "  + str(format(math.degrees(angle1),'.2f')) + " " +
str(format(math.degrees(angle2),'.2f'))+'\n')
                elif 7<d<14 and 0<angle1<0.9599 and 0<angle2<0.9599:
                    print
(res1.resname,res1.get_id()[1],res2.resname,res2.get_id()[1],format(d,'
.2f'),format(math.degrees(angle1),'.2f'),format(math.degrees(angle2),'.
2f'))
                    file.write(res1.resname + str(res1.get_id()[1]) + "
" + res2.resname + str(res2.get_id()[1]) + " " + str(format(d,'.2f')) +
" "  + str(format(math.degrees(angle1),'.2f')) + " " +
str(format(math.degrees(angle2),'.2f'))+'\n')
file.close()


# python script for calculate B-factor and z-score from rmsf.agr
produced by MD
import os
import os.path
import re
import math
import statistics
path=os.getcwd()
wj=os.listdir(path)
for agr in wj:
    if str(agr).find('.agr')>-1:
        print (agr)
        resnum=[]
        rmsf=[]
        b=[]
        z=[]
        name=str(agr)[:-9]
        with open(agr,'r') as file:
            t=file.readlines()
        for line in t:
            if line[0:1]!='@':
                x=float(line[15:21])
                y=(math.pow(x,2)*8*math.pow(3.1415926,2))/3
                b.append(y)
                res=float((line[1:4]))
                resnum.append(res)
                rmsf.append(x) #
```

```
if len(b)==291:
    av=statistics.mean(b)
    dev=statistics.pstdev(b)
    for i in b:
        zscore=(i-av)/dev
        z.append(zscore)
    file.close()
    file1=open(str(name+'_BZ'),'a')
    file1.write(format('Residue','<10') +
format('RMSF','<10')+ format('B-factor','<10') + format('Z-
score','<10') + '\n')
    for i in range(len(resnum)):
        file1.write(str(format(resnum[i],'<10.3f')) +
str(format(rmsf[i],'<10.3f')) + str(format(b[i],'<10.3f')) +
str(format(z[i],'<10.3f')) + '\n')
file1.close()
```

1.11 Gram scale preparation

We scaled-up the production of (*R*)-2-DFPD to compare the practical application of the WT and engineered *Pm*IR-6P. The resting cells expressing WT and *Pm*IR-6P were used for the scale-up preparation. Recombinant glucose dehydrogenase (GDH) from *Bacillus subtilis* CGMCC 1.1398 and glucose was used for NADPH recycling. The pH was automatically adjusted to 6.0 by titration with 2.0 M NaOH solution.

The bioreduction of 150, 200 and 400 mM 2-DFPL was initiated at 30°C for WT and 35°C for *Pm*IR-6P by adding 2.70, 3.60 and 7.24 g 2-DFPL to 100 mM PBS (pH 6.0), which contained 0.25 mM NADP, 1.3 equivalent of glucose, 100 mg/mL resting cells expressing WT or *Pm*IR-6P, and 20 mg/mL resting cells expressing GDH, making the total reaction volume of 100 mL. The conversion ratio was monitored using HPLC analysis. After the reaction reached >99% conversion, the pH of reaction mixture was adjusted to >10 by addition of 2.0 M NaOH solution. The reaction mixture was extracted by ethyl acetate after reaction completion.

1.12 Analytical methods

The conversion and *ee* was analyzed using a Waters Alliance HPLC equipped with chiral analysis column maintained at 25 °C, using 0.1% diethylamine in hexane and isopropanol at a flow rate of 1 ml min^{-1}. The UV detection was performed at 256 nm

The detail of analysis conditions was listed in Table S7.

1.13   Protein sequence of *Pm*IR, *Pm*IR-Re and *Pm*IR-6P

>*Pm*IR

MKSSNRSENIRVGTENTVGKSKSVTVIGLGPMGKAMAAAFLEHGYKVTVWN
RTSNKADELITKGAVRASTVHEALAANELVILSLTDYDAMYTILEPASENLSGK
VLVNLSSDTPDKAREAAKWLANRGAGHITGGVQVPPSGIGKPESSTYYSGPK
EVFEANKETLEVLTGTDYRGEDPGLAALYYQIQMDMFWTAMLSYLHATAVA
QANGITAEQFLPYAAETMSSLPKFIEFYTPRINAGEYPGDVDRLAMGMASVEH
VVHTTQDAGIDITLPTAVLEVFRRGMENGHAGNSFTSLIEIFKKSDIRP

>*Pm*IR-Re

MKSSNRSENIRVGTENTVGKSKSVTVIGLGPMGKAMAAAFLEHGYKVTVWN
RTSNKADELITKGAVRASTVHEALAANELVILSLTDYDAMYTILEPASENLSGK
VLVNLSSDTPDKAREAAKWLANRGAGHITGGVMVMPSGIGKPESSTYYSGPK
EVFEANKETLEVLTGTDYRGEDPGLAALYEQIAMDMFWTAMLSYLHATAVAQ
ANGITAEQFLPYAAETMSSLPKFIEFYTPRINAGEYPGDVMNLAMGMASVEH
VVHTTQDAGIDITLPTAVLEVFRRGMENGHAGNSFTSLIEIFKKSDIRP

>*Pm*IR-6P

MKSSNRSENIRVGTENTVGKSKSVTVIGLGPMGKAMAAAFLEHGYKVTVWN
RTSNKADELITKGAVRASTVHEALAANELVILSLTDYDAMYTILEPASENLSGK
VLVNLSSDTPDKAREAAKWLANRGAGHITGGVMVMPSGIGKPESSTYYSGPK
EVFEANKETLEVLTGTDYRGEDPGLAALYEQIAMDMFWTAMLSYLHATAVAQ
ANGITAEEFLPYAAETMSSLPKFIEFYTPRINAGEYPGDVMNLAMGMRSVEHV
VHTTQDAGIDITLPMAVLEVFRRGMENGHAGNSFTSLIEIFKKSDIRP

2 Supplementary tables

Table S1. X-ray analysis statistics of *Pm*IR

	*Pm*IR
Data collection	
Space group	P 61
Cell dimens	
a, b, c (Å)	133.771 133.771 62.23
α, β, γ (deg)	90 90 120
Resolution range (Å)	50-2.50
Highest resolution shell (Å)	2.54-2.50
R_{merge}	0.150 (0.890)
I/σ[a] (I)	17.3 (4.9)
No. of unique reflections[a]	42733 (1977)
Completeness[a] (%)	99.3 (100)
Redundancy[a]	10.18 (10.99)
CC1/2[a]	0.997 (0.939)
Refinement	
Resolution (Å)	28.21-2.50
$R_{\text{work}}/R_{\text{free}}$ (%)	20.72/24.03
No. of water molecules	130
B factors	
protein (Å^2)	43.86
water (Å^2)	36.89
RMS deviations	
bond lengths (Å)	0.009
bond angles (deg)	1.07
Ramachandran plot	
favored (%)	95.04
allowed (%)	4.42
outliers (%)	0.53
PDB Code	8KFK

[a] The values in the parentheses refer to the highest resolution shell.

Table S2. Alanine scanning of interface binding pocket

Variant	Conversion (%)	*ee* (%), major enantiomer
WT	26.6	96, *R*
P140A	36.1	96, *R*
Q190A	28.4	96, *R*
R251A	32.3	95, *R*
Q138A	15	91, *R*
V139A	12.8	62.6, *R*
Y187A	3.4	4, *R*
M191A	0.5	n.d.
F194A	0.5	n.d.
W195A	n.d.[a]	n.d.
E224A	16.9	95, *R*
T225A	10.6	90, *R*
K231A	20.4	91, *R*
F232A	2.2	n.d.
F235A	1.2	n.d.
V249A	12.3	88.7, *R*
D250A	n.d.	n.d.
M254A	10.4	100, *R*
S258A	8.3	100, *R*

[a]n.d.: Not detected

Table S3. Constraints that specified for reactive conformation of pro-*S* pose

Geometric constraint	Constraint of Y187E and substrate	Tolerance	Geometric constraint	Constraint of substrate and NADPH	Tolerance
d1	3.3	0.5 Å	*d2*	3.2	0.5 Å
θ1	87.3	10 °	*θ3*	108.5	10 °
θ2	152.0	10 °	*θ4*	110.3	10 °
dihedral angle 1	100.3	20 °	dihedral angle 4	118.7	10 °
dihedral angle 2	-168.5	10 °	dihedral angle 5	-100.2	10 °
dihedral angle 3	-158.0	10 °	dihedral angle 6	-119.7	20 °
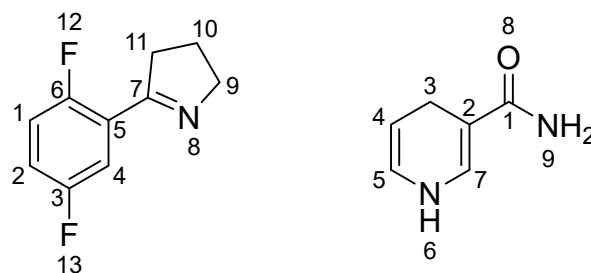
Table S4. Geometric feature of imine cation and NADPH in binding pocket.

Geometric feature of 2-DFPL and NADPH	Geometric feature in *pro-S* pose TS	Initial complex used for design	*Pm*IR-Re
distance: $C7_{2\text{-DFPL}}$-$C4_{NADPH}$	3.2 Å	3.6 Å	3.6 Å
angle 1: $C5_{2\text{-DFPL}}$ -$C7_{2\text{-DFPL}}$ -$C4_{NADPH}$	108.5 °	124.1 °	111.5 °
angle 2: $C7_{2\text{-DFPL}}$ -$C4_{NADPH}$ -$C3_{NADPH}$	110.3 °	131.4 °	115.3 °
dihedral angle 1: $C11_{2\text{-DFPL}}$ -$C5_{2\text{-DFPL}}$ -$C7_{2\text{-DFPL}}$ -$C4_{NADPH}$	118.7 °	105.9 °	118.5 °
dihedral angle 2: $C7_{2\text{-DFPL}}$ -$C4_{NADPH}$ -$C2_{NADPH}$ -$C1_{NADPH}$	-100.2 °	-79.6 °	-74.6 °
dihedral angle 3: $C5_{2\text{-DFPL}}$ -$C7_{2\text{-DFPL}}$ -$C4_{NADPH}$ -$C3_{NADPH}$	-119.7 °	-136.2 °	-126.5 °

Table S5. Comparison of conversion and stereoselectivity using WT and variants
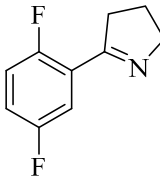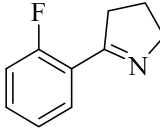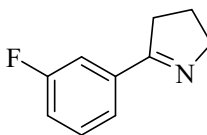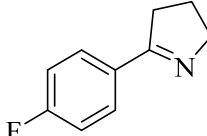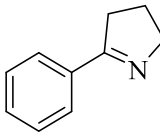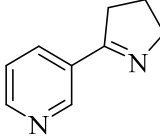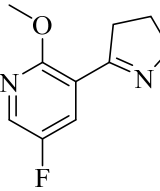
Variants	Conversion [a] (%)	*ee* [b] (%)
WT	22.3 ± 0.6	97.0
P140A	36.1 ± 1.8	97.7
Q190S	59.3 ± 4.0	98.0
R251N	46.0 ± 2.6	98.0
P140A/Q190S	43.0 ± 1.0	>99.0
P140A/R251N	35.7 ± 1.5	>99.0
Q190S/R251N	58.7 ± 2 .5	>99.0
P140A/Q190S/R251N	94.3 ± 0.6	>99.0

Table S6. Variants selected for experimental validation of improved stability

Variants	Average B-factor	ΔZscore	Foldx Energy (kcal·mol^{-1})	Rosetta Energy (kcal • mol^{-1})	Conversion (%)
Template (P140A/Q190S/R251N)	33.578	-	-	-	10.1
D192S	34.25	-1.622	-1.886	-4.283	n.d.
D268W	27.257	-2.304	-1.559	-3.978	8.0
D268Y	34.13	-1.461	-1.645	-4.249	n.d.
G143M	34.981	-0.711	-1.802	-3.595	34
G143W	33.445	-2.63	-1.674	-3.577	14
H132Y	28.856	-0.423	-1.892	-3.914	4
N211Y	32.397	-2.328	-2.6	-4.329	3
N109K	31.612	-1.575	-1.619	-4.539	n.d.
Q188L	30.389	-1.504	-1.527	-5.354	n.d.
S258T	34.783	-2.843	-2.761	-5.722	n.d.
T277L	29.956	-0.407	-2.486	-2.503	15
T277M	30.134	-0.255	-2.515	-1.655	40

[a]n.d.: Not detected

Table S7. The HPLC conditions for chiral analysis and conversion determination

Substrates	Chiral column	Chromatographic conditions[a]	Retention time of substrates (min)	Retention time of (R)-enantiomer (min)	Retention time of (S)-enantiomer (min)
	AD-H	95:5	4.5	5.1	5.7
	AD-H	98:2	8.8	6.2	6.5
	AD-H	95:5	5.4	6.3	7.0
	AD-H	98:2	8.3	6.6	7.1
	AD-H	95:5	6.3	6.7	6.5
	OD-H	95:5	15.8	21.4	23.9
	AD-H	95:5	6.3	6.7	7.9

[a]Eluted using n-hexane and isopropanol with 0.1% diethylamine

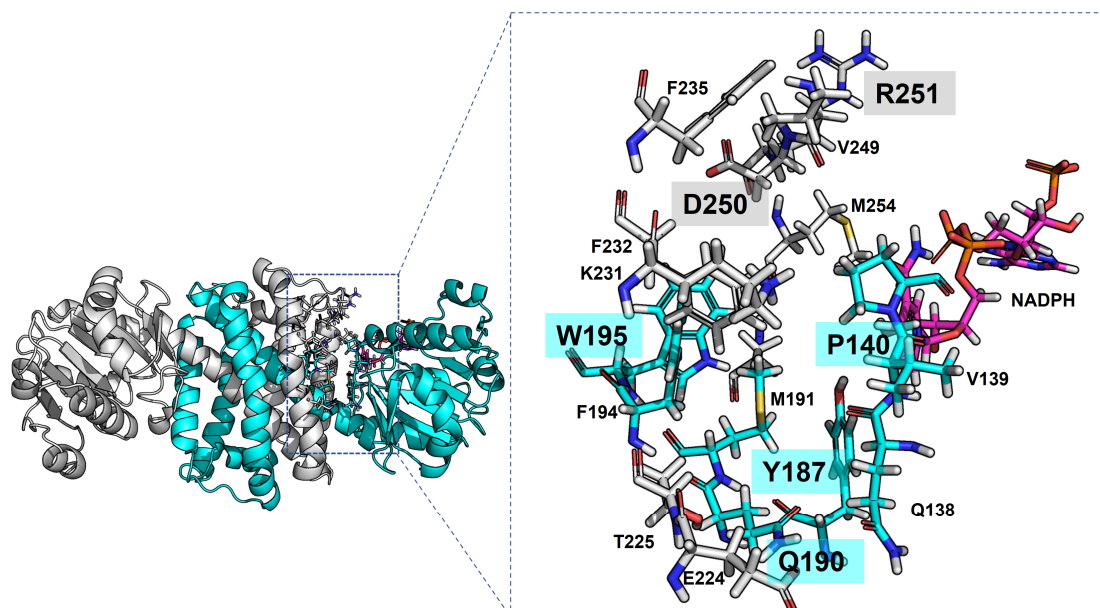# 3   Supplementary figures

Figure S1. Residues selected for alanine scanning. Mutation sites W195A, D250A with complete activity loss, Y187A with low activity, and Q190A, P140A, R251A with increased activity were marked in grey (chain A) and cyan (chain B).
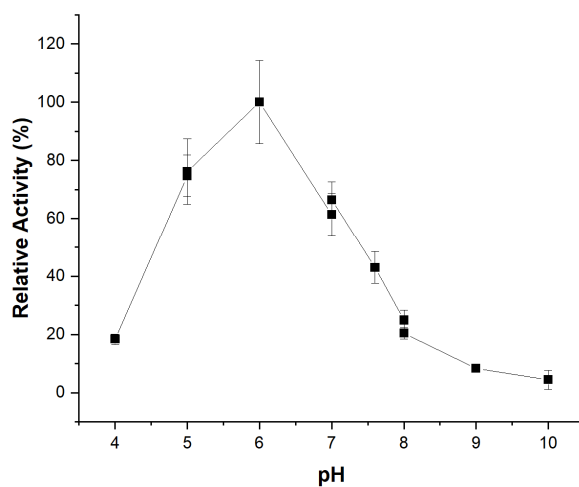

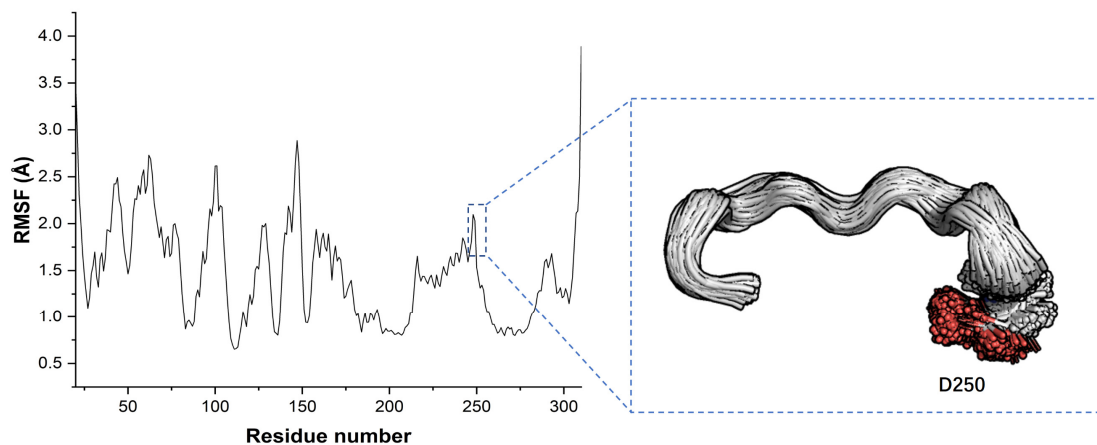**Figure S2.** The pH optimal of *Pm*IR wild type.

Figure S3. RMSF analysis of *Pm*IR. The 242-251 binding loop with a relatively high fexibility and reorientations of D250 side chain was observed in MD simulations.
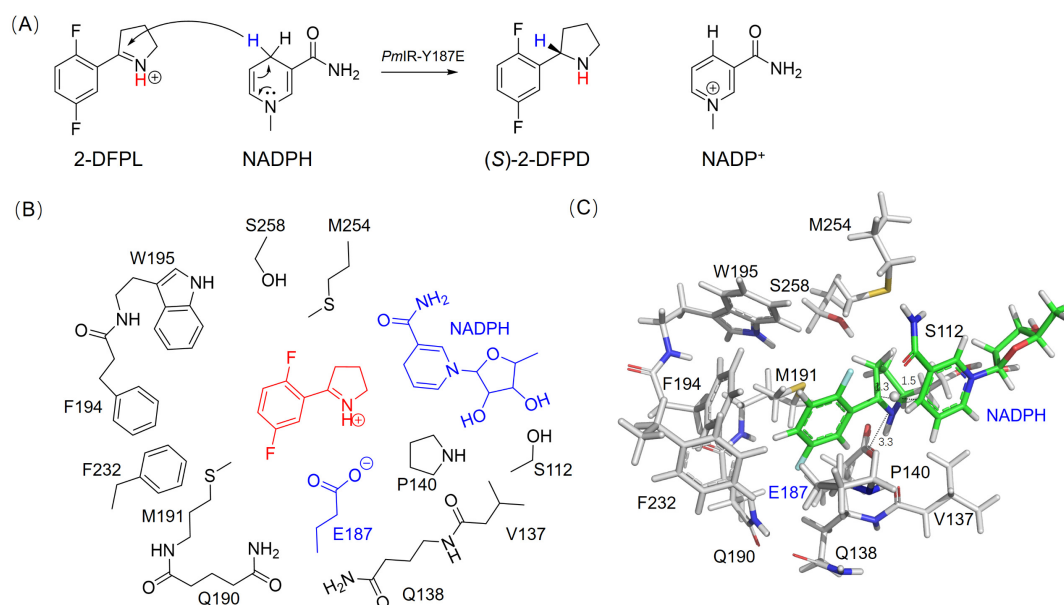

Figure S4. (A) Reaction investigated for transition state optimization. (B) Schematic illustration of the active site model of Y187E used in the current study. (C) The transition state for hydride transfer from NADPH to imine cation. The substrate and NADPH was colored in green.
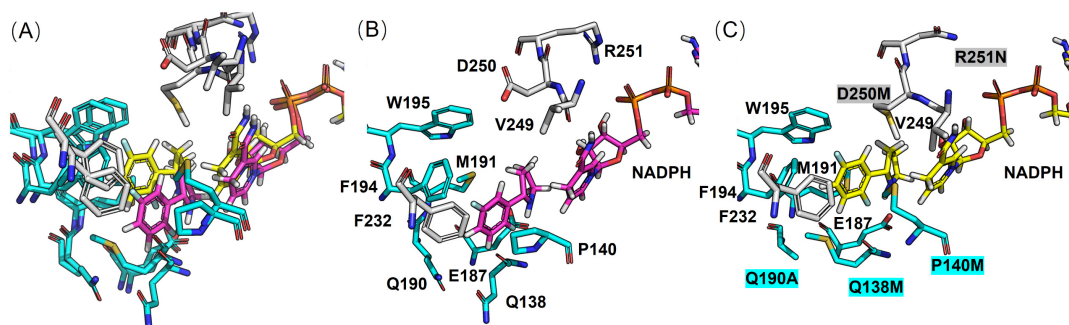
Figure S5. (A) Superimposition of initial Y187E complex for design and *Pm*IR-Re complex. (B) Initial Y187E complex for design. The substrate 2-DFPL and NADPH are colored in purple. (C) The *Pm*IR-Re complex designed with Rosetta. The substrate 2-DFPL and NADPH are colored in yellow. The designed mutations were marked. The residues in chain A and chain B are colored in grey and cyan respectively.
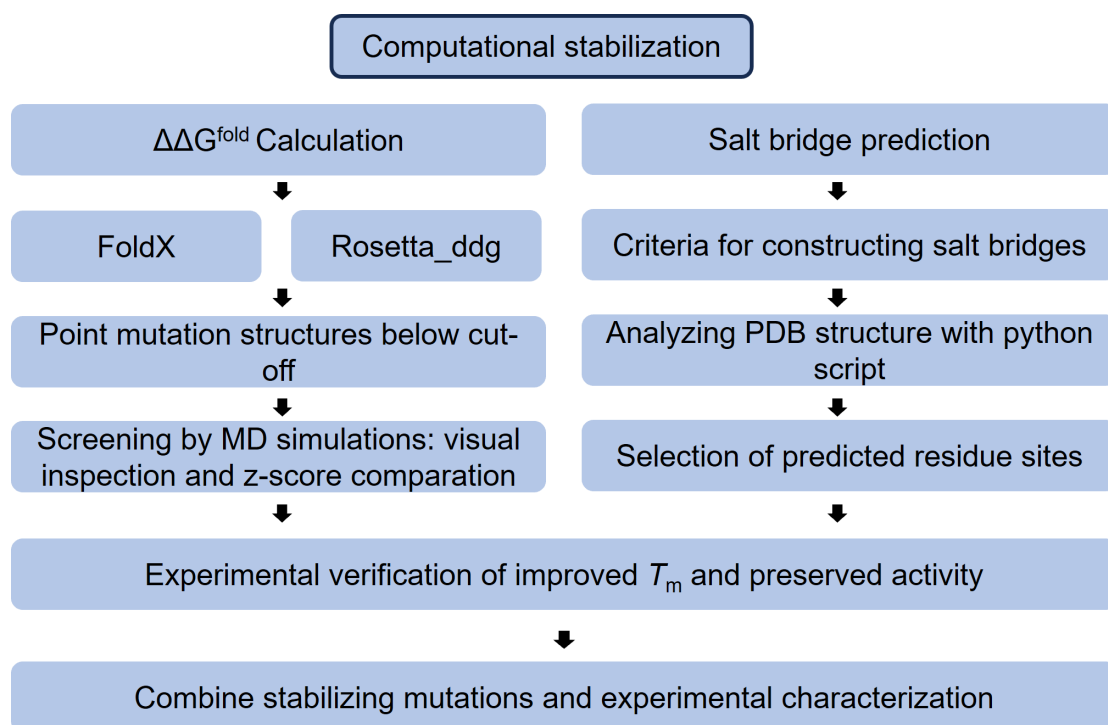

Figure S6. Flowchart illustrating the computational stabilization process of *Pm*IR.
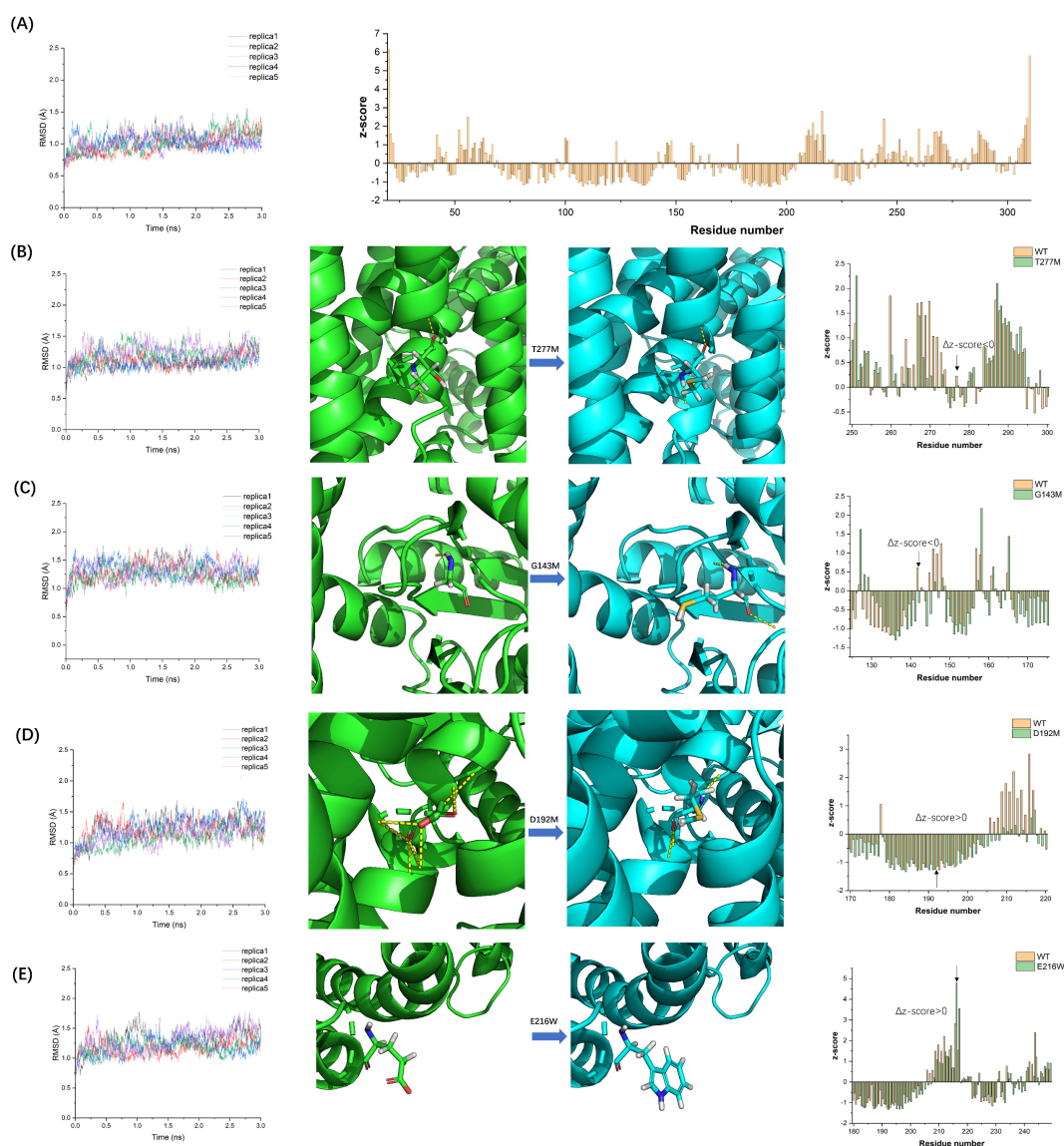
Figure S7. Example for screening variants predicted by $\Delta\Delta G^{Fold}$ calculation. The RMSDs and RMSFs were calculated with protein backbone atoms. The RMSF were converted to z-score. Polar contacts are marked with yellow dashed lines. Variants that did not reveal obvious structural undesirable mutations and with a Δz-score < 0 were selected for experimental characterization. (A) The RMSD of WT during 3 ns simulations (left). Z-score of each residue of WT (right). (B) The RMSD of T277M during 3 ns simulations (left). The T277M without obvious hydrogen bond loss (middle). The mutation T277M with a Δz-score < 0 (right). (C) The RMSD of G143M during 3 ns simulations (left). G143M with additional hydrogen bond (middle). The mutation G143M with a Δz-score < 0 (right). (D) The RMSD of D192M during 3 ns simulations (left). D192M with an obvious loss of polar contacts (middle). The mutation D192M with a Δz-score > 0 (right). (E) The RMSD of E216W during 3 ns simulations (left). Hydrophobic residue E216W that is solvent exposed (middle). The mutation E216W with a Δz-score > 0 (right).
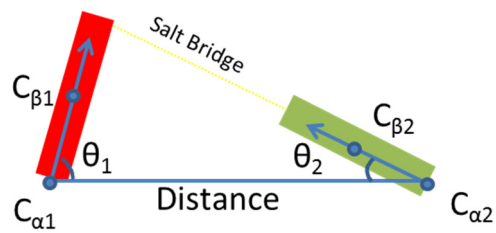
Figure S8. Criteria for constructing salt bridges were established as follows: 1. The distance between Cα1 and Cα2 of the residues should be in the range ·of 5-14 Å, and the angles $\angle C_{\beta 1}C_{\alpha 1}C_{\alpha 2}$ (θ1) and $\angle C_{\beta 2}C_{\alpha 2}C_{\alpha 1}$ (θ2) should range from 0 to 180°. 2. If the distance between $C_{\alpha 1}$ and $C_{\alpha 2}$ is greater than 7 Å, both $\angle C_{\beta 1}C_{\alpha 1}C_{\alpha 2}$ (θ1) and $\angle C_{\beta 2}C_{\alpha 2}C_{\alpha 1}$ (θ2) must be smaller than 110°. Residue pairs that satisfy these criteria are considered as potential candidates for constructing salt bridges through mutagenesis.
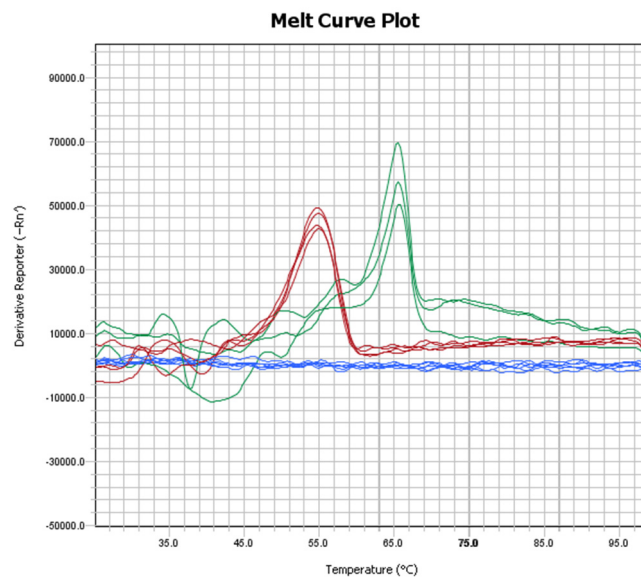
Figure S9. The melt curve plot for the determination of T_m of WT (red), *Pm*IR-6P (green) and control (blue).
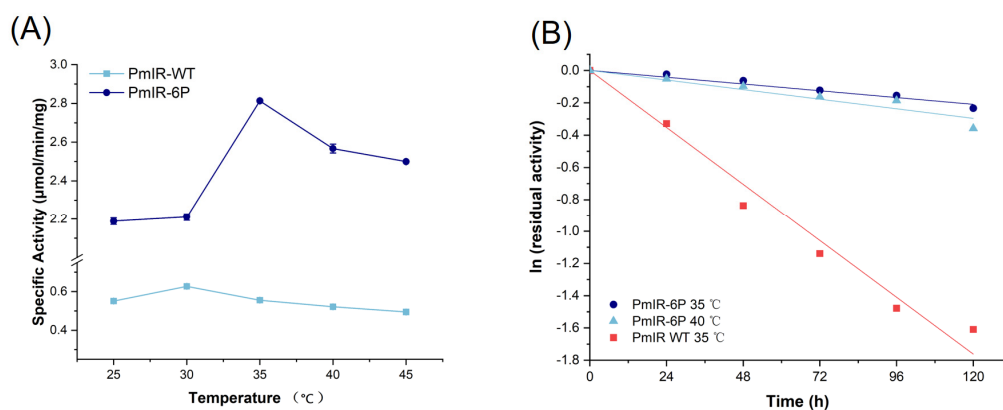
Figure S10. Comparison of WT and stabilized *Pm*IR-6P. (A) Optimal reaction temperature. (B) Kinetic deactivation curve.
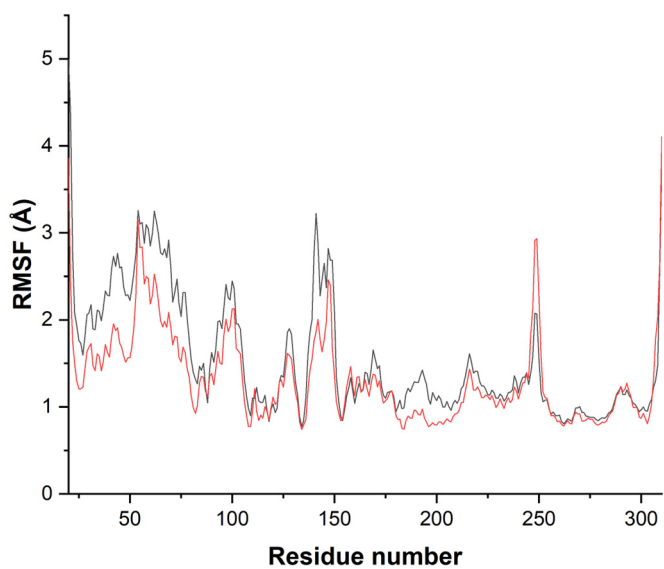
Figure S11. Comparison of RMSF calculated from 200 ns simulation at 308K. The *Pm*IR wild type is in black line and the *Pm*IR-6P is in red line.

4 HPLC spectrums

Figure S12. HPLC chromatograms of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 2-DFPL reduction. (A) 2-DFPL substrate standard.(B) Racemic 2-DFPD standard. (C) Bioreduction of 2-DFPL with *Pm*IR-6P. (D) Bioreduction of 2-DFPL with *Pm*IR-Re.
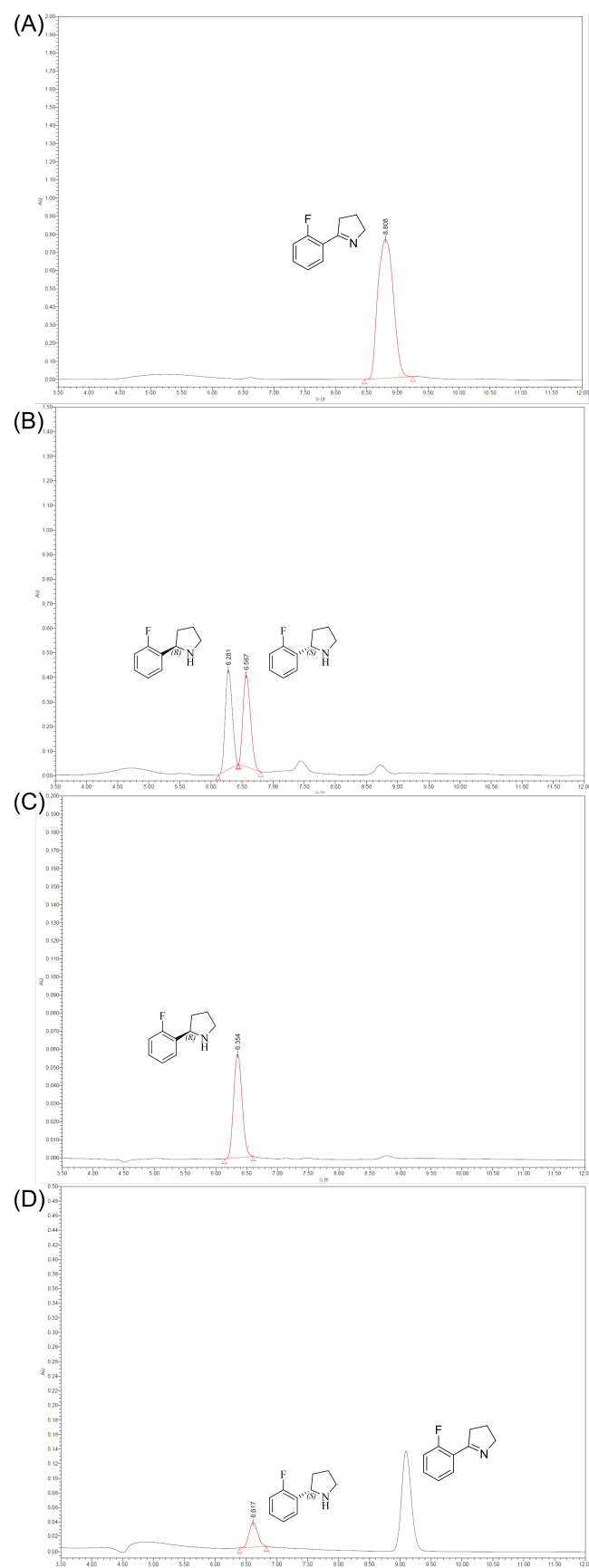
Figure S13. HPLC chromatograms of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 5-(2-fluorophenyl)-3,4-dihydro-2H-pyrrole reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.
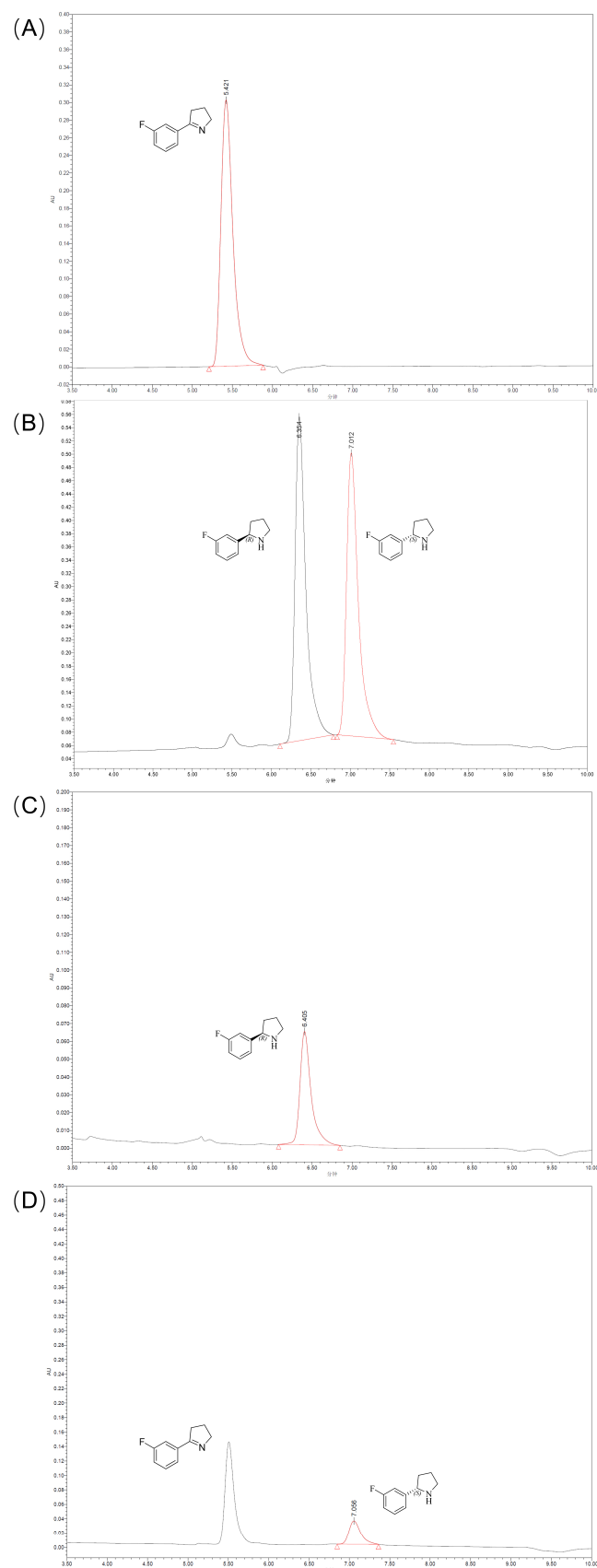
Figure S14. HPLC chromatogram of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 5-(3-fluorophenyl)-3,4-dihydro-2H-pyrrole reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.
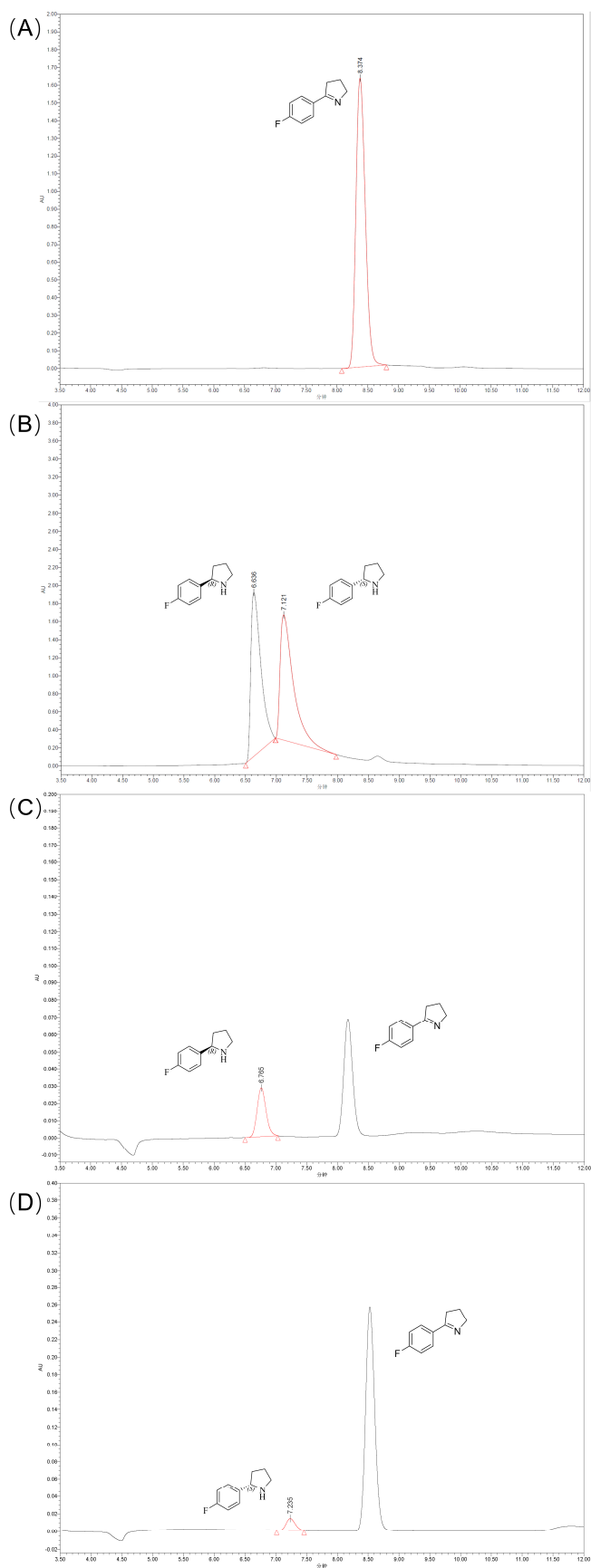
Figure S15. HPLC chromatogram of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 5-(4-fluorophenyl)-3,4-dihydro-2H-pyrrole reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.
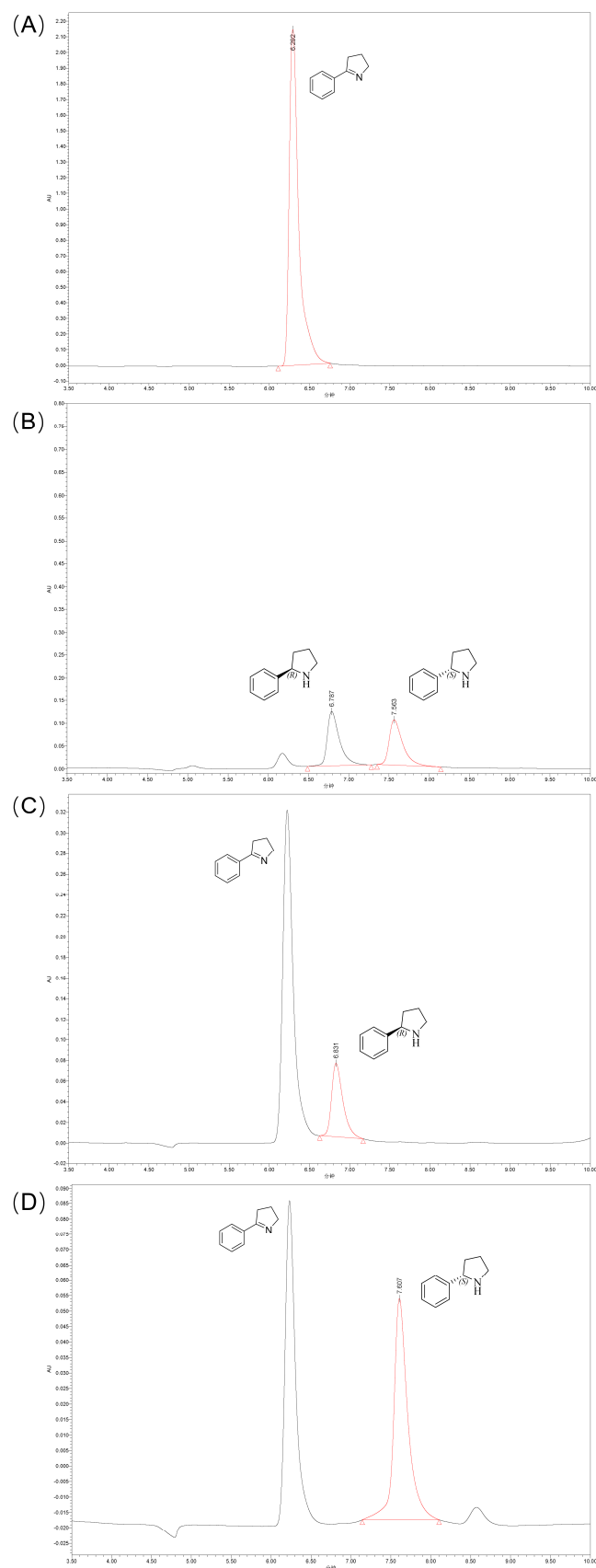
Figure S16. HPLC chromatogram of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 2-phenyl-1-pyrroline reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.
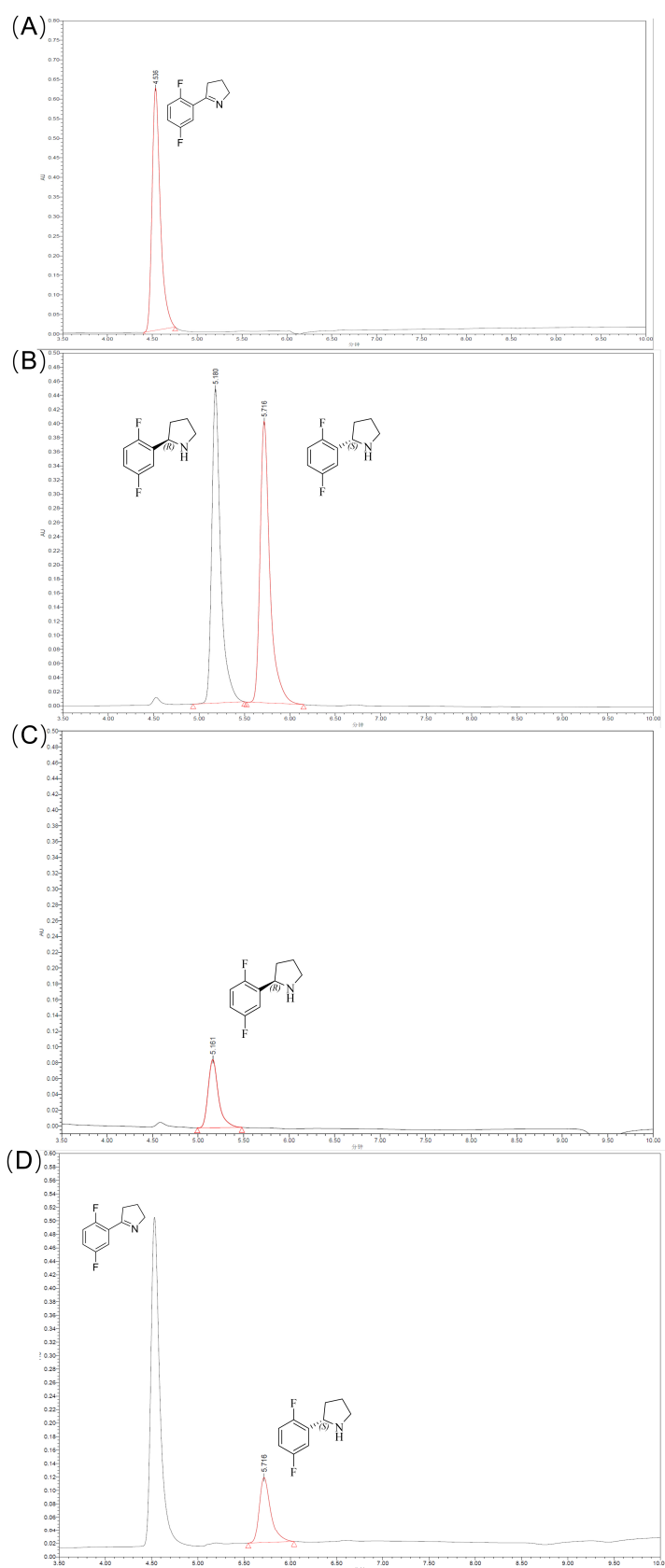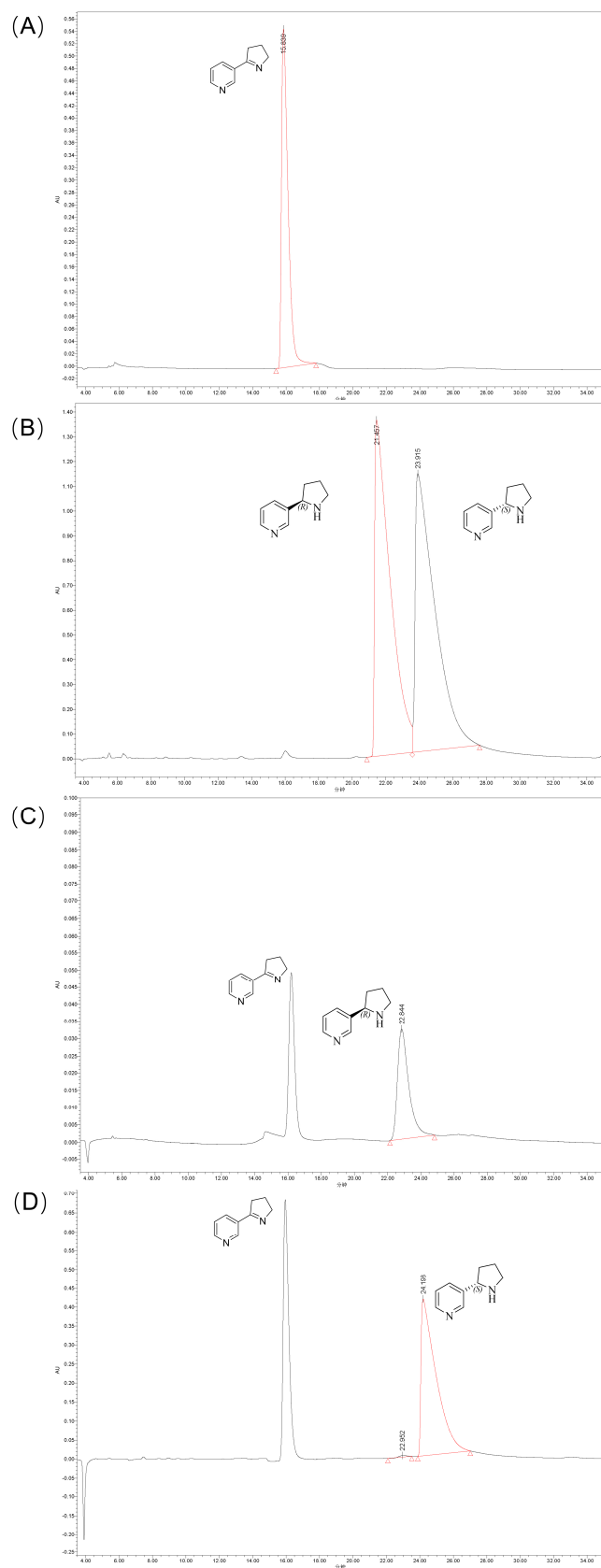
Figure S17. HPLC chromatogram of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed myosmine reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.
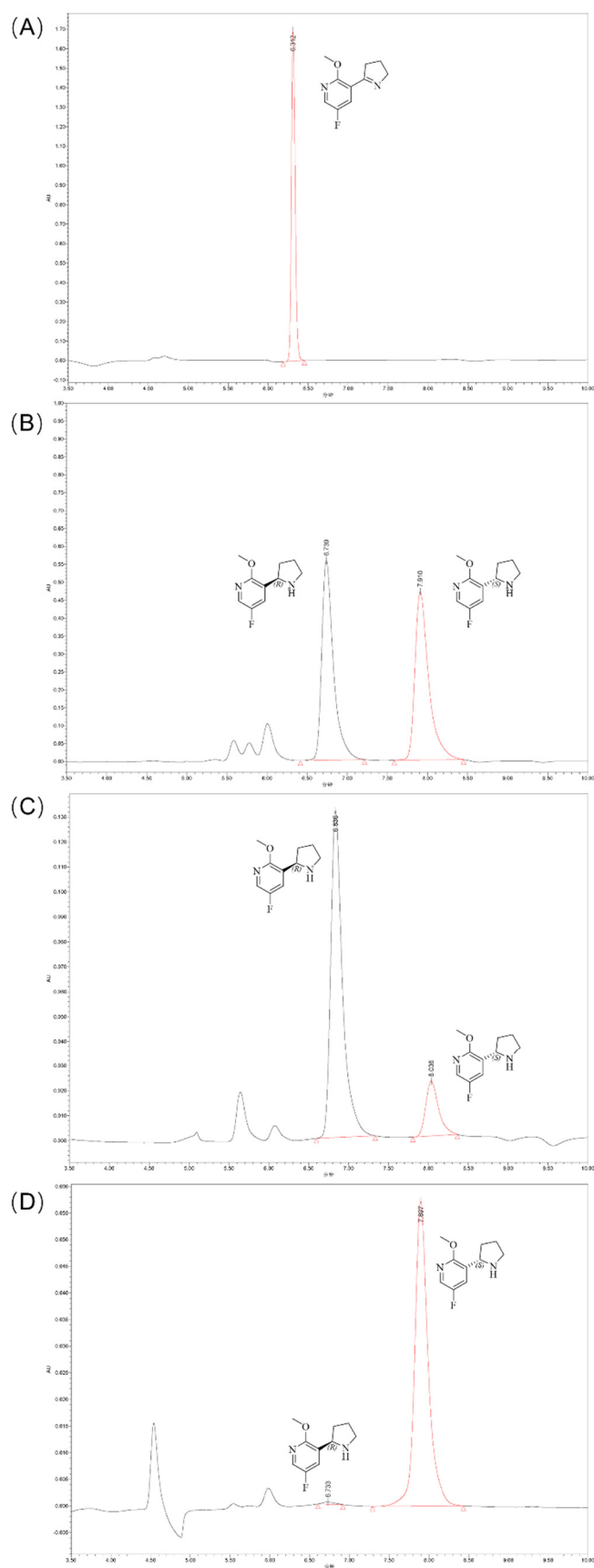
Figure S18. HPLC chromatogram of chiral analysis of *Pm*IR-6P and *Pm*IR-Re catalyzed 3-(3,4-dihydro-2H-pyrrol-5-yl)-5-fluoro-2-methoxypyridine reduction. (A) Substrate standard.(B) Racemic product standard. (C) Bioreduction with *Pm*IR-6P. (D) Bioreduction with *Pm*IR-Re.

5 ^{1}H NMR and ^{13}C NMR spectrums

7.95 7.95 7.94 7.93 7.92 7.92 7.39 7.38 7.37 7.36 7.28 7.18 7.17 7.15 7.10 7.08 7.08 7.07   4.03 4.01 4.00   3.02 3.00 2.98   2.04 2.03 2.02 2.00 1.99   0.00

0.93 0.93 0.96 0.93   2.00   2.02   2.12

^{1}H NMR spectrum of 5-(2-fluorophenyl)-3,4-dihydro-2H-pyrrole (600 MHz, CDCl$_3$)

170.59 162.26 160.59   131.78 131.73 130.06 124.17 124.14 122.88 122.80 116.33 116.18   77.30 77.08 76.87   60.67   37.74   22.92

^{13}C NMR spectrum of 5-(2-fluorophenyl)-3,4-dihydro-2H-pyrrole (150 MHz, CDCl$_3$)

38

^{1}H NMR spectrum of 5-(3-fluorophenyl)-3,4-dihydro-2H-pyrrole (600 MHz, CDCl$_3$)

^{13}C NMR spectrum of 5-(3-fluorophenyl)-3,4-dihydro-2H-pyrrole (150 MHz, CDCl$_3$)

^{1}H NMR spectrum of 5-(4-fluorophenyl)-3,4-dihydro-2H-pyrrole (600 MHz, CDCl$_3$)

^{13}C NMR spectrum of 5-(4-fluorophenyl)-3,4-dihydro-2H-pyrrole (150 MHz, CDCl$_3$)

^{1}H NMR spectrum of 3-(3,4-dihydro-2H-pyrrol-5-yl)-5-fluoro-2-methoxypyridine (600 MHz, CDCl₃)

^{13}C NMR spectrum of 3-(3,4-dihydro-2H-pyrrol-5-yl)-5-fluoro-2-methoxypyridine (150 MHz, CDCl₃)

^{1}H NMR spectrum of (*R*)-2-DFPD (600 MHz, CDCl$_3$)

6 Cartesian coordinates of TS structure in pro-*S* pose

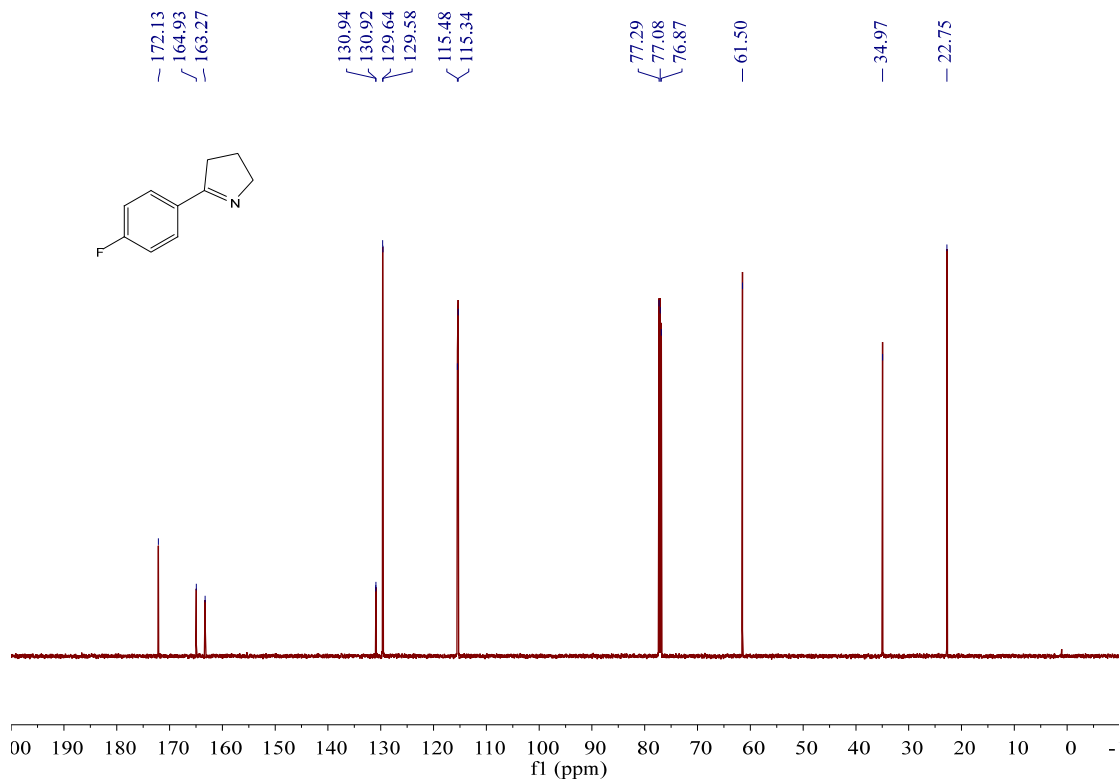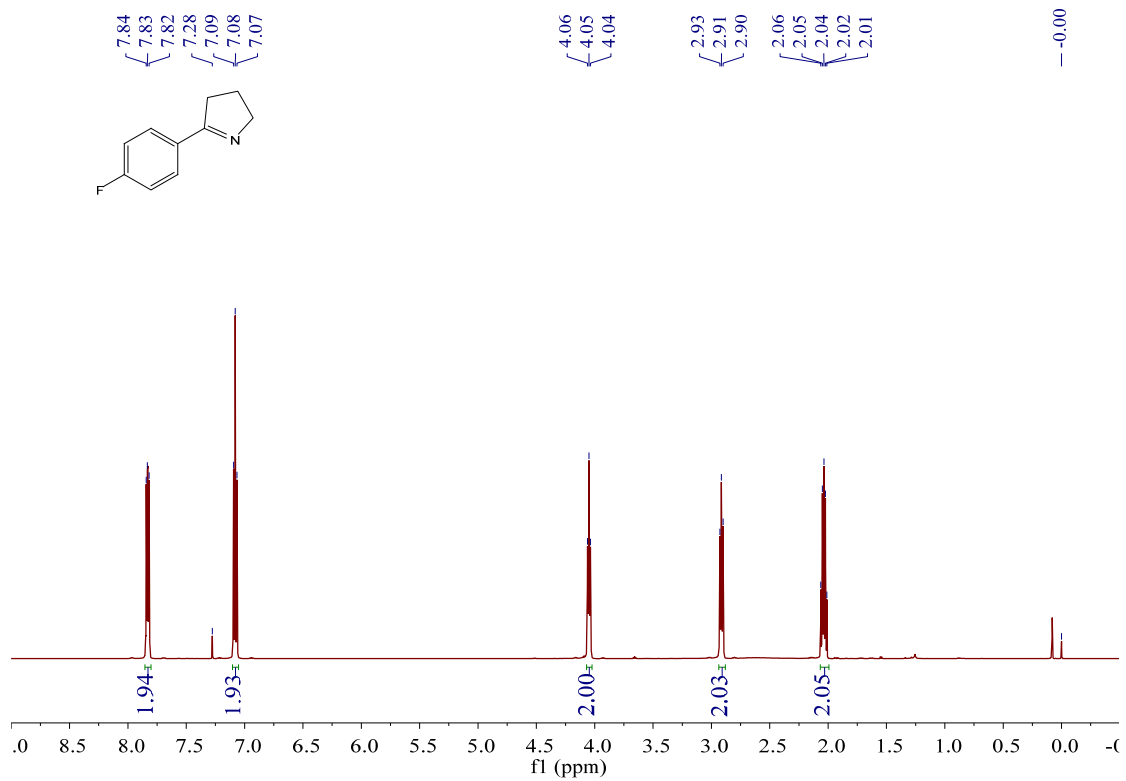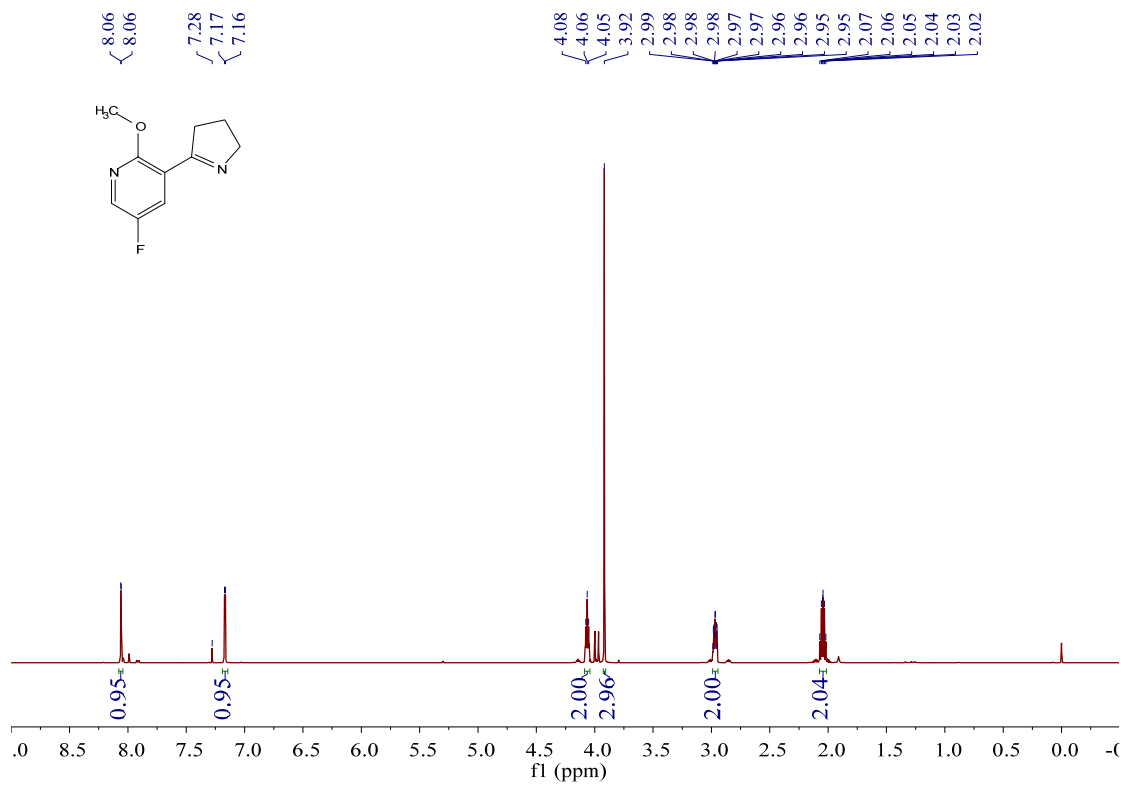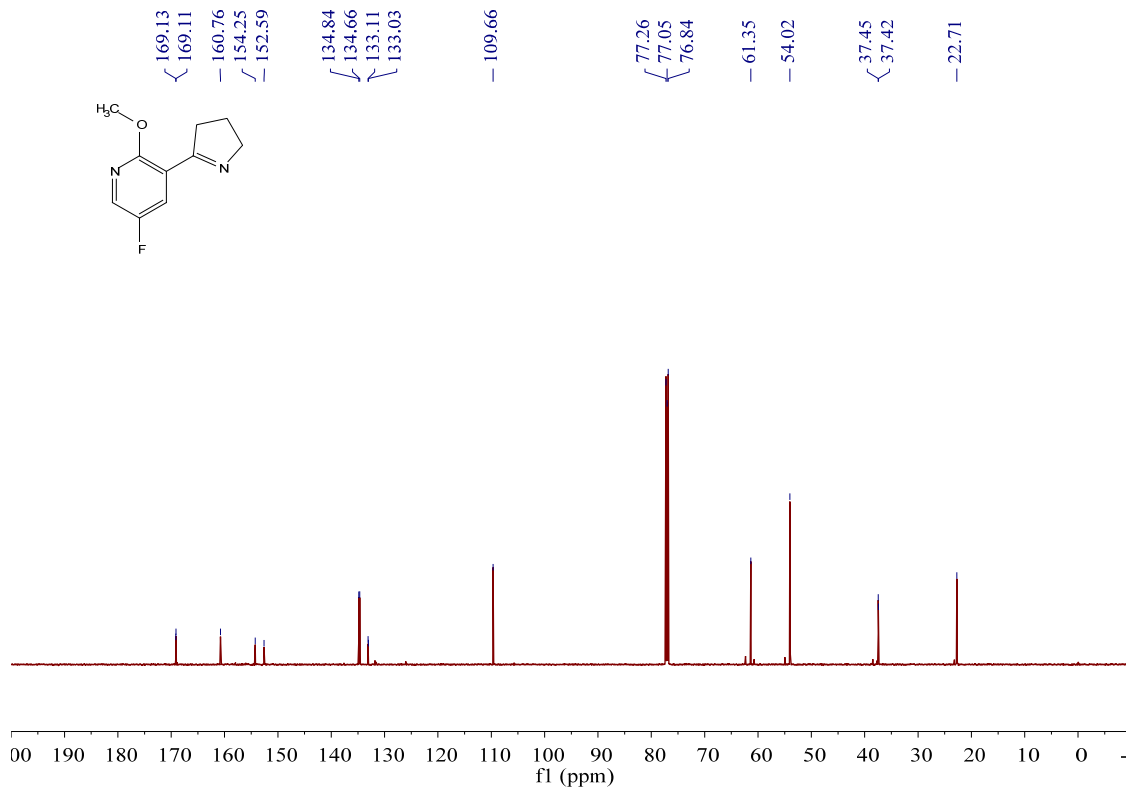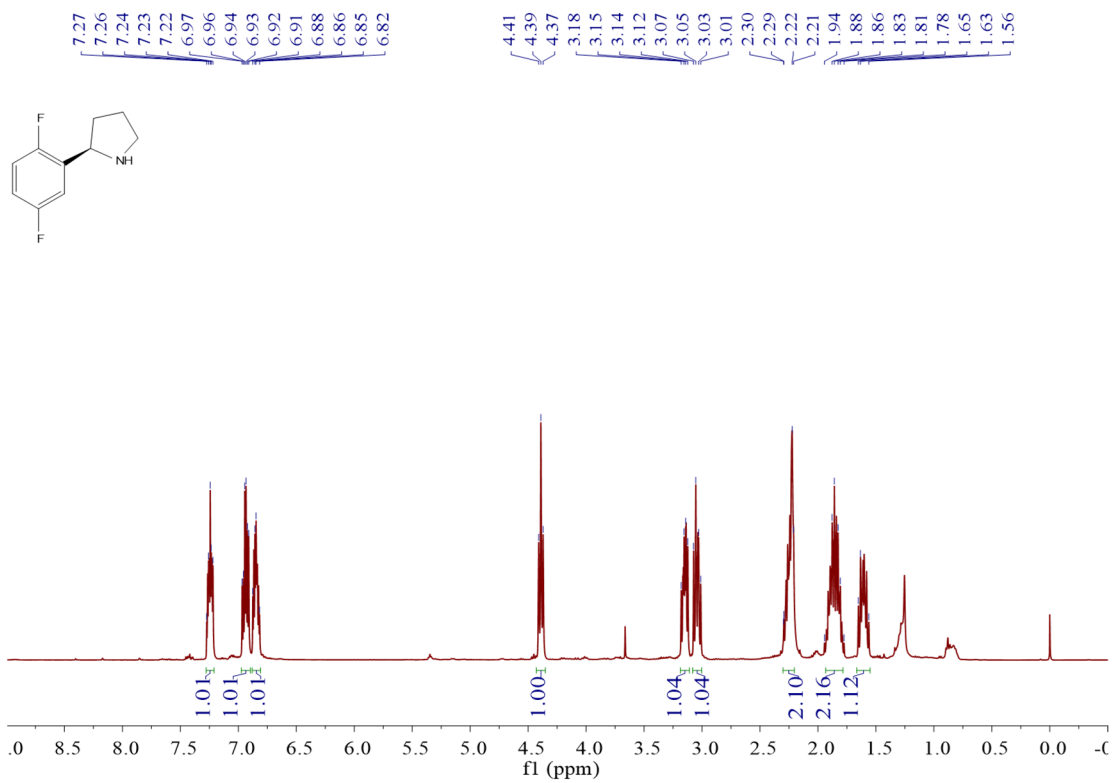|							
C	58.17097400	60.07702000	35.01804400	C	39.89503400	55.69098400	36.84205100
C	57.02756300	59.18440200	35.53458300	C	39.86327500	57.16403500	37.20158700
C	55.82187800	59.20043300	34.62320800	O	38.58164900	57.51981700	37.70821600
C	55.73335500	58.31574300	33.54111500	H	40.87982100	55.39211500	36.46933600
C	54.79030300	60.12936200	34.80756100	H	40.64544500	57.37957900	37.94937500
C	54.65088400	58.36093500	32.66283200	H	40.10155100	57.76337300	36.31162700
C	53.70505000	60.18020900	33.93194600	H	38.57730300	58.48844600	37.70135100
C	53.63236900	59.29623000	32.85412400	H	39.63897000	55.07799200	37.70598500
H	57.85282700	61.12243800	34.95629000	H	39.15630100	55.48298400	36.06125700
H	57.39133700	58.15591500	35.63879800	C	42.65523400	58.18032000	30.78498700
H	56.73384200	59.51428300	36.53758000	C	44.10651500	58.02888700	31.29218100
H	56.52127600	57.58246300	33.38955800	O	44.64901500	58.82993600	32.04685300
H	54.84169900	60.82025900	35.64539700	C	41.83769600	59.25953900	31.48738700
H	54.59868400	57.66351600	31.83210000	C	40.65713600	59.70745700	30.61918500
H	52.91723500	60.91105100	34.08811300	C	41.36017900	58.74718900	32.85259700
H	52.78662300	59.33086600	32.17631100	H	42.83275300	58.36871000	29.72690700
H	59.03602400	60.01698200	35.67896700	H	42.50381600	60.11629300	31.64841900
H	58.48322200	59.76561300	34.01627500	H	40.99506900	60.11132800	29.65889600
C	44.55622600	60.11379200	44.52098400	H	39.98296400	58.86821700	30.41058100
C	45.68812200	60.62302600	43.61832800	H	40.06997400	60.48329200	31.12280700
C	45.26995000	60.98298800	42.19732600	H	42.19704600	58.41880400	33.47065400
S	44.49975400	59.62682700	41.23201500	H	40.79700300	59.52050500	33.38831700
C	45.96718600	58.57594900	40.95104700	H	40.68778700	57.89044600	32.72721900
H	43.61670900	60.63704500	44.31255400	H	42.14740100	57.20975300	30.85091300
H	46.11993300	61.53332200	44.05497700	N	44.78762100	57.00401700	30.67107100
H	46.50470600	59.89262900	43.59394600	C	46.09402300	56.54923400	31.16228900
H	44.51914500	61.78050900	42.21783200	C	46.58280400	55.28694000	30.46426700
H	46.13383800	61.35661300	41.63707700	C	47.10472300	55.50651800	29.02981200
H	46.62710900	59.02702000	40.20547700	C	48.08315200	56.66786100	29.01421500
H	46.50963200	58.42241000	41.88538300	N	47.60604100	57.83791800	28.50077800
H	45.60190100	57.60674900	40.60776000	O	49.22559300	56.58795400	29.46762000
H	44.37845200	59.04319000	44.40954900	H	44.19380200	56.26308600	30.32024900
H	44.84181900	60.34517500	45.54701100	H	46.01742300	56.37251500	32.23499200
C	43.88798400	53.85497800	42.04499700	H	47.39836000	54.88719800	31.06684800
C	44.96194800	54.39830600	41.10292900	H	45.80305000	54.51591800	30.45142200
O	44.60235500	55.63832700	40.49437100	H	47.63853900	54.60529400	28.71770600
H	42.94882600	53.68009900	41.50820200	H	46.27946800	55.69284000	28.33439800
H	45.88586200	54.59861800	41.65373400	H	48.15883700	58.66714400	28.67071700
H	45.20591800	53.65148700	40.33269800	H	46.61011100	57.96171000	28.39655400
H	43.94548500	55.46804000	39.80729900	H	46.82803800	57.33793900	30.99787600
H	44.21199600	52.90500600	42.47001600	N	47.97008700	61.28004300	31.91289500
H	43.69143200	54.56054600	42.85673600	C	48.58505300	62.52194300	32.40486900

C	50.07949900	62.24533600	32.18457300	S	46.59583000	54.74555000	37.50210500
C	50.20298000	60.70610400	32.39364900	C	48.33139300	54.99755100	36.99309400
C	48.74677200	60.20909600	32.53932400	H	47.16090700	50.53884500	35.38977500
H	48.39812300	62.69624500	33.48216500	H	48.84007700	52.35622000	36.93341500
H	50.36667900	62.51297300	31.16433300	H	46.11967400	51.02549000	37.39248300
H	50.70959600	62.81829600	32.87033200	H	46.83483100	52.06999300	38.59388500
H	50.68909200	60.24031000	31.53241400	H	46.49135900	53.10668700	35.73706000
H	50.78854300	60.44996500	33.28006800	H	45.10021500	53.08551300	36.80073300
H	48.56942700	59.24910500	32.04385700	H	48.57163600	56.03938600	37.20166700
H	48.50530000	60.07909800	33.60719900	H	49.02028500	54.37261900	37.55935700
H	48.16894800	63.36704600	31.85611100	H	48.44654800	54.81934000	35.92263000
C	44.08988000	51.69105800	32.45965800	H	48.69801200	50.77802700	37.65102000
C	43.28491900	52.88835900	32.97186500	C	53.11099000	52.59303700	37.10803200
C	44.27794100	53.97824900	33.40969800	C	53.52480200	52.58540200	38.61571000
C	43.70867000	55.16093500	34.21212600	O	54.70692800	52.50495700	38.93116400
O	43.93689500	56.31723300	33.70392000	C	51.86940500	53.41266700	36.70995400
O	43.12996500	54.91371500	35.29132600	C	52.04415000	54.91316400	36.76508600
H	44.90352500	52.05532100	31.81696200	C	52.56852800	55.59789700	35.65949100
H	42.62426600	53.28358200	32.19007600	C	51.69534700	55.65507200	37.90062700
H	42.64948200	52.61449900	33.82020000	C	52.75042700	56.98069000	35.68508800
H	45.03567500	53.50838800	34.05210200	C	51.86305600	57.04095800	37.93005400
H	44.80343200	54.36006600	32.53092100	C	52.39376200	57.70686700	36.82436700
H	44.58519100	51.19560100	33.30606400	H	54.01300800	52.91895600	36.58998700
H	43.61506500	50.90078700	31.87923600	H	51.58451100	53.10825700	35.69927800
C	48.82091300	50.08400100	33.47777400	H	51.02158000	53.11097600	37.33644500
C	49.13255900	50.68953500	34.84923000	H	52.84339700	55.03558600	34.77009600
O	50.27049300	51.01180200	35.18457800	H	51.29286200	55.15253900	38.77370800
C	48.56684300	51.18137100	32.42535100	H	53.16916500	57.49090100	34.82517600
C	49.68809400	52.21481900	32.30079500	H	51.57992400	57.58760000	38.82365200
C	49.42910500	53.17769300	31.14221300	H	52.54144500	58.78276600	36.84318200
N	49.89955800	54.43889400	31.31306700	H	52.93217600	51.54717800	36.83012600
O	48.85828600	52.81109300	30.11466500	N	52.53136100	52.59024900	39.56264200
H	49.66404400	49.44298800	33.22117500	C	52.74494900	52.54810500	41.02487100
H	48.41114600	50.72350400	31.44517600	C	51.38684700	52.31712500	41.71630700
H	47.63033300	51.69694600	32.67315100	C	50.29875600	53.22576600	41.21280500
H	49.82556600	52.76347700	33.23780300	C	49.16824100	52.83170100	40.53056400
H	50.64222600	51.70786200	32.10776000	C	50.24960200	54.67004600	41.27036000
H	49.72005500	55.13996700	30.58986500	C	49.05494900	55.07246200	40.60725300
H	50.17477500	54.76264800	32.22672100	C	51.09923200	55.66479700	41.78543200
H	47.93448200	49.44041100	33.53722900	N	48.40495100	53.93235900	40.18933600
N	48.05570900	50.92139100	35.65820400	C	48.70135400	56.41419100	40.43617000
C	48.18303000	51.48795800	37.00403500	C	50.75117500	57.00008100	41.62247200
C	46.79508200	51.88420700	37.51625900	C	49.56594900	57.37130600	40.95240600
C	46.19357800	53.10262000	36.79007500	H	51.57965800	52.75769200	39.27041400

44

H	53.19502700	53.47294400	41.38710900	C	43.62114300	57.45829000	37.77016300
H	51.07660800	51.27805900	41.56004800	C	42.89112400	57.82351400	36.47545200
H	51.53084200	52.44511900	42.79614900	N	43.97223800	58.20929900	35.56773500
H	48.83964500	51.83293000	40.27893200	C	45.06651600	58.66682200	36.26480900
H	47.59828000	53.95182500	39.57094700	C	46.45080800	58.36948900	35.77347400
H	52.01785700	55.39599500	42.29698600	C	46.72023400	57.44790700	34.75453000
H	47.78793500	56.68674700	39.91959300	C	48.04270200	57.16433400	34.44101500
H	51.40436600	57.77428000	42.01262200	C	49.12189200	57.74993500	35.08892300
H	49.32670500	58.42384000	40.83417700	C	48.86075800	58.67564000	36.10045900
H	53.44430500	51.74021700	41.25619000	C	47.54231100	58.96709100	36.41565700
C	40.56098000	62.26820100	37.04988000	F	47.29163100	59.85613900	37.41621200
N	41.93294500	61.89452800	36.68144700	F	48.29689200	56.22717300	33.48743000
C	39.80534600	61.25386600	37.95133900	H	44.54399200	59.37211800	38.27747000
C	43.00116500	62.19703500	37.48934300	H	45.69032900	58.03439900	38.26305900
N	45.30954800	62.37791600	39.24333700	H	43.98770200	56.43190600	37.67004600
O	39.01600600	60.36721200	37.18844500	H	42.99937900	57.53966000	38.66677800
C	38.92676900	62.19285500	38.84477700	H	42.36740300	56.97813000	36.03127400
C	44.28396200	61.93037800	37.09153200	H	42.19891400	58.65950700	36.63242700
O	37.54371800	61.88969600	38.71752200	H	45.90793000	56.93475200	34.24945800
C	39.17830600	63.59118200	38.25596500	H	50.13835700	57.46952900	34.83894200
C	44.55604900	61.26398200	35.84302600	H	49.66754600	59.14455700	36.65009200
O	40.53333300	63.49904800	37.75465000	H	44.03460500	57.58996400	34.71528400
C	38.94103000	64.76501800	39.20509700	H	46.98271100	61.20634100	32.14537400
C	43.40663700	60.96450400	35.02358900				
O	46.45131500	62.85962500	37.31782400				
C	42.15528700	61.27392100	35.46912500				
C	45.46078800	62.41868900	37.88187900				
H	37.89397400	64.70199900	39.52691400				
H	39.10377800	65.71948900	38.70092000				
H	38.50019700	63.69437500	37.39349600				
H	39.24834800	62.15242000	39.89046300				
H	37.30263100	61.22924600	39.38084700				
H	40.51822200	60.66762700	38.53811100				
H	38.10810000	60.72097100	37.23290500				
H	40.00308700	62.34754000	36.10693200				
H	42.76731700	62.73528000	38.39746300				
H	44.77866600	61.60826600	39.64457500				
H	46.14505800	62.63506200	39.75155500				
H	45.02632700	59.93009100	36.19210900				
H	45.51062700	61.49310900	35.37310300				
H	43.53137700	60.50434000	34.05078800				
H	41.26833300	61.06347400	34.88636000				
H	39.59292700	64.70671600	40.08141000				
C	44.80434700	58.43463400	37.77454500				

7   References

1.   Krug, M.; Weiss, M. S.; Heinemann, U.; Mueller, U., XDSAPP: a graphical user interface for the convenient processing of diffraction data using XDS. *J. Appl. Crystallogr.* **2012,** *45*, 568-572.

2.   Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Structural Biology* **2011,** *67*, 235-242.

3.   Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. *Acta Crystallographica Section D-Biological Crystallography* **2010,** *66*, 486-501.

4.   Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkoczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L.-W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D., Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D-Structural Biology* **2019,** *75*, 861-877.

5.   Aleku, G. A.; Man, H.; France, S. P.; Leipold, F.; Hussain, S.; Toca-Gonzalez, L.; Marchington, R.; Hart, S.; Turkenburg, J. P.; Grogan, G.; Turner, N. J., Stereoselectivity and Structural Characterization of an Imine Reductase (IRED) from Amycolatopsis orientalis. *ACS Catalysis* **2016,** *6* (6), 3880-3889.

6.   Lee, T.-S.; Cerutti, D. S.; Mermelstein, D.; Lin, C.; LeGrand, S.; Giese, T. J.; Roitberg, A.; Case, D. A.; Walker, R. C.; York, D. M., GPU-Accelerated Molecular Dynamics and Free Energy Methods in Amber18: Performance Enhancements and New Features. *Journal of Chemical Information and Modeling* **2018,** *58* (10), 2043-2050.

7.   Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015**, 11(8):3696-713.

8.   Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of computational chemistry* **2004,** *25* (9), 1157-1174.

9.   Caves, L. S. D.; Evanseck, J. D.; Karplus, M., Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin. Protein Science 1998, 7.

10.   Wijma, H. J.; Floor, R. J.; Jekel, P. A.; Baker, D.; Marrink, S. J.; Janssen, D. B., Computationally designed libraries for rapid enzyme stabilization. Protein engineering, design & selection : PEDS 2014, 27 (2), 49-58.

11.   Wullich, S. C.; Wijma, H. J.; Janssen, D. B.; Fetzner, S., Stabilizing AqdC, a Pseudomonas Quinolone Signal-Cleaving Dioxygenase from Mycobacteria, by FRESCO-Based Protein Engineering. Chembiochem: a European journal of chemical biology 2020, 22 (4), 733-742.

12.   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene,

M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. C.01*, Wallingford, CT, 2016.

13. Simón, L.; Goodman, J. M., How reliable are DFT transition structures? Comparison of GGA, hybrid-meta-GGA and meta-GGA functionals. *Organic & biomolecular chemistry* **2011,** *9* (3), 689-700.

14. Buss, O.; Rudat, J.; Ochsenreither, K., FoldX as Protein Engineering Tool: Better Than Random Based Approaches? *Computational and structural biotechnology journal* **2018,** *16*, 25-33.

15. Cui, Y.; Chen, Y.; Liu, X.; Dong, S.; Tian, Y. e.; Qiao, Y.; Mitra, R.; Han, J.; Li, C.; Han, X.; Liu, W.; Chen, Q.; Wei, W.; Wang, X.; Du, W.; Tang, S.; Xiang, H.; Liu, H.; Liang, Y.; Houk, K. N.; Wu, B., Computational Redesign of a PETase for Plastic Biodegradation under Ambient Condition by the GRAPE Strategy. *ACS Catalysis* **2021,** *11* (3), 1340-1350.

16. Lee, C. W.; Wang, H. J.; Hwang, J. K.; Tseng, C. P., Protein thermal stability enhancement by designing salt bridges: a combined computational and experimental study. *PloS one* **2014,** *9* (11), e112751.