Dimer-assisted mechanism of (un)saturated fatty acid decarboxylation for alkene production

Leticia L. Rade[a,1], Wesley C. Generoso[a,1], Suman Das[b,1], Amanda S. Souza[a,1], Rodrigo L. Silveira[c], Mayara C. Avila[a], Plinio S. Vieira[a], Renan Y. Miyamoto[a], Ana B. B. Lima[c], Juliana A. Aricetti[a], Ricardo R. de Melo[a], Natalia Milan[a], Gabriela F. Persinoti[a], Antonio M. F. L. J. Bonomi[a], Mario T. Murakami[a], Thomas M. Makris[b,2], Leticia M. Zanphorlin[a,2]

[a]Brazilian Biorenewables National Laboratory, Brazilian Center for Research in Energy and Materials, Campinas, Brazil, 13083-100.

[b]Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, USA, 27695-7622.

[c]Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil, 21941-594.

(1) These authors have contributed equally to this work.
(2) **Correspondence should be addressed to LMZ (leticia.zanphorlin@lnbr.cnpem.br) or TMM (tmmakris@ncsu.edu)**

This PDF file includes:

Supplementary Materials and Methods

Tables S1 to S8

Figures S1 to S31

Supplementary References

SUPPLEMENTARY MATERIALS AND METHODS

Reagents and plasmids

The cytochrome P450 decarboxylase encoding ORF from *Rothia nassimurium* (GenBank accession number: WP_083091567) into the pET28a(+) vector was synthesized by Genscript (Piscataway, USA). Fatty acids (C10:0, C12:0, C14:0, C16:0, C17:0, C18:0, C20:0, C18:1 and C18:2), alkenes (1-nonene, 1-undecene, 1-tridecene, 1-pentadecene and 1-hexadecene) and methyl pentadecanoate were purchased from Sigma Aldrich Co. (St Louis, USA). α and β-hydroxyl acids derived from C12:0, C14:0, C16:0, C17:0, C18:0, and C20:0 were purchased from Larodan AB (Solna, Sweden). Hydrogen peroxide was obtained from Sigma Aldrich. Ni-NTA and Q-Sepharose resins were obtained from GE Healthcare Biosciences (Chicago, USA). Perdeuterated fatty acids were purchased from CDN isotopes (Quebec, CAN).

Bioinformatics and phylogenetic analyses

CYP152 sequences were retrieved from *Uniprot* database by HMMER/HMMSEARCH v.3.1b2 (E-value <1e-5) using the CYP152A HMM (KO K15629) from Kofam database (1) Retrieved sequences were subjected to multiple sequence alignment using mafft (2), and only those sequences that had conserved the residues, Phe79, His85, and Arg245, described for $OleT_{JE}$ as important for decarboxylase activity, were kept. These sequences were further clustered at 90% of sequence identity using cd-hit v 4.8.1 (3) and subjected to phylogenetic analysis using maximum likelihood methods as implemented in iqTree software (4). The branch encompassing the $OleT_{JE}$ was further inspected. One sequence of each genus was selected and used for an additional phylogenetic tree reconstruction using Genious software (https://www.geneious.com/). The final trees were annotated and visualized using iTOL software (https://itol.embl.de/).

Heterologous expression and purification of $OleTP_{RN}$ and mutants

The $OleTP_{RN}$-pET28a vector was co-transformed into *E. coli* BL21(DE3) with pG-TF2 plasmid (Takara Bio, Kusatsu, JPN), which encodes GroEL, GroES, and Tig chaperones. The expression and purification were performed as described by Amaya, J.A. et al. (5). Briefly, the transformed strain was grown in 500 mL of a selective (chloramphenicol and kanamycin) Terrific Broth medium (TB) supplemented with 125 mg L^{-1} of thiamine hydrochloride and 500 µL of trace metals solution (containing 50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnCl_2$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiSO_4$) at 37 °C. When the optical density (OD_{600nm}) of the culture reached 0.6, the growth temperature was reduced to 20 °C, and 200 µM isopropyl-β-D-thiogalactopyranoside (IPTG) (Invitrogen), 5 µM of $FeCl_3$, 100 µM of δ-aminolevulinic acid (ALA), and 10 µg/mL tetracycline were added. The culture was grown at 20 °C overnight. Cells were harvested by centrifugation at 7.000 x g (10 min at

4 °C) and subsequently, the cell pellets were resuspended in buffer A (100 mM potassium phosphate (pH 7.5), 300 mM NaCl, 10 mM imidazole, and 5 % (*v/v*) glycerol) and disrupted by sonication in a Vibracell VCX 500 device (Sonics and Materials, Newtown, USA), by performing 2 min of sonication (40% amplitude and no pause), followed by 5 minutes of stirring at 4 °C. This procedure was repeated 5 times. The cell lysate was centrifugated at 12.000 x g for 25 min and the clear supernatant was loaded onto a 5 mL His-Trap chelating HP column (GE Healthcare Biosciences), pre-equilibrated with buffer A. The target protein was eluted with a gradient of imidazole (5 to 300 mM) in buffer A. The eluted sample was dialyzed overnight against 100 mM phosphate buffer (pH 7.5) containing 50 mM NaCl and 5% (*v/v*) glycerol. Dialyzed $OleTP_{RN}$ was loaded onto a fast-flow Q-sepharose column (Sigma-Aldrich Co., St Louis, USA) and eluted by a NaCl gradient (50 mM to 1 M). The heterologous expression and enzyme purification steps were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, 12%).

Site-directed mutagenesis

$OleTP_{RN}$ mutants were generated by inverse PCR (**Table S7**). Primer pairs were constructed to hold a complementary sequence longer than 15 nt and/or 50°C, in which the codon changes were introduced. PCR amplicons were circularized by Gibson assembly (6).

Circular Dichroism

Spectra of far-UV circular dichroism (CD) were taken on a JASCO J-815 CD spectrometer controlled by a CDF-426S Peltier temperature control system (Jasco Analytical Instruments, Oklahoma, EUA). CD spectra were measured with 0.3 mg mL^{-1} of each enzyme separately (wild-type and mutants) in buffer GF (150 mM NaCl, 100 mM potassium phosphate, 5% (*w/v*) glycerol, pH 7.5) at 20 °C. A quartz cuvette with a 1-mm path length was used for all CD experiments. Each spectrum was an average of ten scans, measured from 190 to 260 nm, with a bandwidth of 2.0 nm and a response time of 4 s. Thermal unfolding experiments were determined by measuring the CD signal at 208, 215, and 222 nm with increasing temperature in a continuous ramp mode of 20 to 100 °C, at 1 °C min^{-1}. The observed ellipticity (θ $degree^{-1}$) was normalized against the protein concentration and reported as the mean residue molar ellipticity (θ $degree^{-1}$ cm^2 $dmol^{-1}$). The melting temperature (T_m) was determined according to Boltzmann sigmoidal fitting of the data obtained in the CD denaturation curve. Data were plotted on Origin 8.1 software (OriginLab, Northampton, USA).

UV–visible spectroscopic characterization

UV-visible spectroscopy was performed using a Thermo Scientific™ NanoDrop 2000 spectrophotometer for the P450 heme spectrum analyses, considering Soret band at 424 nm. The analysis of the ferrous-carbon monoxide complex (Fe^{+2}-CO) was performed under

anaerobic conditions. Carbon monoxide gas (~1 bubble/s for 30 seconds) was bubbled into a cuvette containing a solution of ferric $OleTP_{RN}$ (10 µM of $OleTP_{RN}$ diluted in storage buffer, composed of 100 mM potassium phosphate (pH 7.5), 100 mM NaCl, 5% (*v/v*) glycerol, and 3 µM of methylviolegene dichloride). $OleTP_{RN}$ was reduced with sodium dithionite after bubbling CO gas into the cuvette, and the UV-Vis spectra of the complex were recorded from 200 to 800 nm.

Substrate binding titrations and determination of dissociation constants

The substrate-binding titrations were performed with $OleTP_{RN}$ (and mutants) and several saturated fatty acids (C_n; n=10 to C20), and two unsaturated fatty acid (C18:1 and C18:2). All fatty acids were purchased from Sigma Aldrich Co. The fatty acids stocks (10 mM) were prepared in ethanol and triton X-100 solution (70/30 *v/v*). The titrations were performed by gradual additions of fatty acids into a cuvette containing 1 mL of 5 µM of the enzyme (reaction buffer: 100 mM potassium phosphate (pH 7.5), 100 mM NaCl, 5% (*v/v*) glycerol). A spectrum was collected after each addition of fatty acid aliquot (200-800 nm spectral range), and the changes in the absorption (ΔA), induced by substrate binding, were calculated by the difference between the absorbance maxima and minima at 422 nm (LS) and 399 nm (HS). ΔA at 399 and 422 nm were added together (ΔA_{max}) and were plotted as a function of the total concentration of the substrate and eventually fitted for Morrison's quadratic equation (7) to determine the dissociation constant (K_D).

Transient kinetics measurements and analysis

$OleTP_{RN}$ WT and mutants (S26F and E405F) were treated with 10 molar equivalents of hydrogen peroxide to turn over all the adventitiously bound fatty acids and passed through a PD-10 desalting column (GE Healthcare) to get rid of any unwanted small molecules. The enzyme-fatty acid (ES) complexes prepared for single turnover reactions consisted of 20 µM $OleTP_{RN}$ complexed with 100 µM (5x) fatty acid (final concentrations in the solution). The ES complexes were incubated on ice before the start of the experiments. The methods of kinetic experiments and data analysis have been described elsewhere (5, 8).

Ferrous oxidation xylenol-orange (FOX) assays

To investigate the optimum $OleTP_{RN}$ reaction conditions, ferrous oxidation xylenol-orange (FOX) (9) assays were performed, allowing the identification of peroxide oxidation in the reaction. The reactions with $OleTP_{RN}$ were carried out in 96-well plates, with a reaction volume of 100 µl composed of 0.5 µM of $OleTP_{RN}$ and 200 µM of myristic acid (prepared in pure ethanol) diluted in reaction buffer (100 mM potassium phosphate, pH 7.5, 100 mM NaCl, and 5% (*v/v*) glycerol). The reaction was incubated for 5 min at 30 °C, then 100 µM of hydrogen peroxide was added, and the system was incubated again for 10 min at 30 °C (Veriti Thermal Cycler - Applied Biosystems, Waltham, USA). After this period, the reactions

were quenched by temperature increasing to 70 °C for 5 min. Then, 25 µL of the reaction mixture was added to 75 µl of pure water and 100 µL of FOX solution (1 mL of FOX stock (10x), 300 µL of ethanol, and 8.7 mL of water). The FOX solution stock (10x) was prepared with 2.5 mM Xylenol orange, 2.5 mM Fe_2SO_4 and 0.5 M H_2SO_4. Posteriorly, the samples were incubated at 60 °C for 5 min, and 100 µL was used for the final absorbance measured at 560 nm using an Infinite 200 PRO microplate reader (TECAN Group Ltd., Mannedorf, CHE). The negative controls were performed without substrate. The optimum pH varied from 4 to 10, using the following buffers: citrate (pH 4.0, 4.5, 5.0, and 5.5), phosphate (pH 6.0, 6.5, 7.0 and 8.0), Tris–HCl (pH 9.0), and glycine (pH 10.0). The influence of temperature was evaluated by incubating the reaction in a range from 20 to 45 °C. The concentration tolerance of hydrogen peroxide and NaCl were evaluated from 0 to 40 mM and from 0 to 500 mM, respectively. The measurements were expressed as relative activity (%), considering the maximum catalytic activity observed for the best reactional condition. Experiments were carried out in quadruplicate.

Steady-state Turnover - Fatty acid substrate conversion

Activity assays were carried out with different fatty acid substrates (C10-C20, C18:1 and C18:1). 1.25 mL of the complex enzyme-substrate was prepared by mixing 500 µM (final concentration) of the fatty acid substrate and 2 µM (final concentration) of $OleTP_{RN}$ WT/mutants diluted in storage buffer and followed by 1-hour incubation at 4°C. The complex was divided into two aliquots of 500 µL for the positive and negative controls. Then, six additions of 1 µL of hydrogen peroxide (10 mM stock) were performed every 10 min on positive control over the course of 1 hr. For the negative control, the addition of buffer was performed instead of the H_2O_2. Reactions were carried out at 35 °C and 300 rpm for the activity evaluation of $OleTP_{RN}$ and mutations. The reaction was immediately stopped after 1 h with 4 µL of 6 M HCl (stock concentration). Then, internal standards such as 1-tetradecene and heptadecanoic acid (final concentration of 0.98 mg/mL) were added to the reaction mixtures. The products were extracted with the addition of an equal volume of chloroform, homogenization carried out in a tube revolver rotator (20 min at 4 °C), and centrifugation at 2000 rpm for 15 min. 400 µL of the organic phase was aliquoted and transferred to a 2 mL glass vial. For the derivatization step, 25 µL of HCl (2 M in methanol) and the second internal standard solution (final concentration of 44.78 µg/mL 1-hexadecene and 51.15 µg/mL methyl pentadecanoate) were added. The sample was then homogenized using a vortex shaker for 30 sec and maintained at 50 °C for 30 min. After this period, 40 µL of NaOH solution (1 M in methanol) was added, and the sample was rigorously homogenized in a vortex shaker for 3 min. Then, 500 µL of pure water was added, and the phases were separated by centrifugation at 2000 rpm for 5 min. The upper phase was discarded, and the sample was analyzed by gas chromatography.

Fatty acid competition assay

A fatty acid competition assay was performed in order to determine the substrate preference of $OleTP_{RN}$. A mixture of five fatty acids (100 µM each of C12, C14, C16, C18, and C20; total 500 µM final concentration) was complexed with 5 µM enzyme (final concentration), and this enzyme-substrates complex was reacted with 500 µM H_2O_2 (stock concentration) in small amounts over 2 hrs at 25 °C. The derivatization steps were the same as the above.

Analytical analyses for fatty acid and products quantification

The reaction products were determined by gas chromatography analysis (7890A, Agilent Technologies, Santa Clara, USA). The following conditions were used: RTx-5MS (30 m x 0,25 mm x 0,25 µm), detector temperature of 290 °C, injector temperature of 230 °C, split ratio 1:10, and oven temperature: 30 °C for 2 min, increase to 220 °C at a rate of 10 °C/min, increase to 240 °C at a rate of 2 °C/min and increase to 350 °C at 50 °C/min. To the identification of the peaks, the comparison of the retention time with authentic standards of fatty acids, alkenes, and hydroxylate products was performed. Quantification of the Compounds was carried out by the addition of internal standards at each sample (1-tetradecene for alkene quantification and heptadecanoic acid for fatty acids and hydroxylates quantification), and the exact quantification of the internal standards added was performed by using calibration curves for each internal standard (1-hexadecene and methyl pentadecanoate).

SEC-MALS and Analytical SEC

Size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS) experiments were performed at 25 °C using an eight-angle static light scattering detector DAWN8 and Optilab® refractive index monitor (Wyatt Technology, Santa Barbara, USA) coupled to an Agilent HPLC 1260 Infinity II system with a Superdex® 200 Increase 10/300 GL analytical size-exclusion column (GE Healthcare). 250 µL from purified $OleTP_{RN}$ WT at 50 µM was injected into the column and eluted in GF buffer (150 mM NaCl, 100 mM potassium phosphate, 5% (*w/v*) glycerol, pH 7.5). Once determined the oligomerization of $OleTP_{RN}$, the remaining mutants were evaluated through Analytical SEC using the same configuration without MALS, using WT- $OleTP_{RN}$ as control. To ensure that each further analysis had been performed with only the monomeric fraction, monomerized samples by mutagenesis have been prepared and collected from size exclusion chromatography prior functional experiments.

Small angle X-ray scattering

Small-angle X-ray scattering (SAXS) data-collection was performed with protein samples at 1 and 2 mg mL^{-1} in GF buffer (150 mM NaCl, 100 mM potassium phosphate, 5%

w/v glycerol, pH 7.5) at the D01A-SAXS2 beamline at the Brazilian Synchrotron Light Laboratory (LNLS-CNPEM, Campinas, Brazil), using a CCD-Mar165 detector. The sample-to-detector distance was set as 1000 mm. The samples were centrifuged at 20,000 x g for 15 min to remove any potential residual aggregates and then analyzed at 20 °C in 1-mm path length mica cells. The scattering profiles were recorded in 10 successive frames of 10 s to monitor radiation damage. The buffer contribution in each profile was subtracted considering the attenuation and integration of the sample using the fit2D software (10). Data were processed and analyzed with the ATSAS package v.4.8.6, using GNOM (11) v.5.0 to generate the distance distribution function, DAMMIF (12) v.1.1.2 to obtain *ab initio* models by simulated annealing, and DAMAVER (13) v.5.0 to reconstruct the final shape by averaging ten different *ab initio* models. SUPCOMB (14) v.2.3 was used to superpose the high-resolution structure into the SAXS envelope.

Protein crystallization, X-ray data collection and structure determination

$OleTP_{RN}$ samples were concentrated to 60 mg mL^{-1} and submitted to crystallization trials, using the vapor diffusion method in sitting drops. Single crystals were obtained from a solution containing 25% (*w/v*) PEG3500, 15% (*w/v*) 2-methyl-2,4-pentanediol, 0.1 M imidazole (pH 6.5), 0.2 M lithium sulfate. Diffraction data were collected at the Manacá beamline from the Brazilian Synchrotron Light Laboratory (LNLS, Campinas, Brazil), using a Dectris Pilatus 2M detector, at 100 K. Data were processed using XDS (15), and the structure was solved by the molecular replacement method using the program MOLREP (16). An electron density at the active site corresponding to palmitic acid (C16:0) was detected and modeled. The atomic coordinates were fitted based on the electron density using COOT (17) and refined using PHENIX refine (18) and REFMAC (19). The final atomic model was verified with MolProbity (20) and PDBRedo server (21). Detailed statistics of data collection and refinement are provided in **Table S8**.

Systems setup for molecular dynamics (MD) studies

Molecular dynamics simulations of $OleTP_{RN}$-C16 complex with heme in the pentacoordinate high spin state started from the crystal structure reported herein. From this structure, models of $OleTP_{RN}$ bound to C14, C12, and C10 fatty acids were obtained by progressively removing carbon atoms from C16. Models of $OleTP_{RN}$ bound to C18 and C20 fatty acid were built by first aligning the enzyme to the C20-bound $OleT_{JE}$ crystal structure (PDB entry: 5M0P) and then transposing 2 and 4 carbon atoms from this C20 fatty acid to build, respectively, C18 and C20 in $OleTP_{RN}$ from C16. Models of $OleTP_{RN}$ bound to terminal alkenes and β-hydroxylated fatty acids were built by removing the COO^- group from the parent substrate and by inserting a hydroxyl group into the C_β respectively. Ligand-free systems were built by removing the substrate from the crystal structure. Hydrogen atoms were inserted with tLeap (22), with titratable residues protonated according to pKa values estimated with H++ (23, 24) at pH=7.4. In such condition, all Asp/Glu residues were charged, His27, His127, and His220 were doubly protonated (+1 charge), and all other Histidine

residues were neutral with a proton attached to the Nε2 atom. Afterwards, the systems were solvated in truncated octahedral water boxes in such a way that the minimum distance between the enzyme and the edge of the box was 17 Å. Sodium ions were added to the solvent to make the system electrically neutral. Final systems contained about 62,000 and 156,000 atoms for the monomer and dimer, respectively. MD simulations of $OleT_{JE}$ in presence of C9, C13 and C19 alkenes were set up starting from the crystal structure taken from PDB 4l40, following similar procedures as described for the $OleTP_{RN}$ monomer.

Force fields and quantum chemical calculations

Force field parameters for fatty acids and alkenes were obtained from the Generalized Amber Force Field (25, 26). Quantum chemical calculations at the HF/6-31G(d) level of theory were performed to obtain Restrained Electrostatic Potential partial charges (27) of the ligands at their minimum energy geometries with Gaussian 09 (28). Parameters for the heme group in the pentacoordinated high-spin state, with Fe coordinated by a Cys residue, were obtained from Shahrhokh *et al*. (29). The ff14SB force field (30) was used for the protein and the TIP3P model (31) for water.

Simulation procedures

Periodic boundary conditions were employed in all simulations, with electrostatics handled with Particle Mesh Ewald (32) and short-range interactions truncated at 10.0 Å. Numerical integration of the equations of motions were carried out with a timestep of 2.0 fs, with bonds involving hydrogen atoms constrained at their equilibrium lengths. The Langevin thermostat was used with a collision frequency of 1.0 ps^{-1} to keep the temperature at 300 K, and the Berendsen barostat, with a relaxation time of 2.0 ps, was used to keep the pressure at 1 bar during equilibration steps. The following equilibration protocol was executed before the production runs: (i) 2000 steps of minimization of energy with protein, heme, and substrate harmonically restrained (k=10 kcal/mol.Å²); (ii) 200 ps of constant-volume molecular dynamics with protein, heme, and substrate harmonically restrained (k=10 kcal/mol.Å²), with the temperature raising from 0 to 300 K during the first 100 ps; (iii) 2000 steps of minimization of energy with the protein α-carbons enzyme harmonically restrained; (iv) 500 ps of constant-volume molecular dynamics with the protein α -carbons enzyme harmonically restrained and with the temperature raising from 0 to 300 K during the first 100 ps; (v) 2000 steps o minimization of energy with all atoms free; (vi) 400 ps of constant-pressure molecular dynamics with all atoms free and temperature ramp from 0 to 300 K during the first 100 ps; (vii) 100 ps of constant-volume molecular dynamics with all atoms free and temperature of 300 K. Force constants of 10 kcal/mol.Å² were used in all harmonic restraints. Production runs were carried out at constant volume and constant temperature for 300 ns. For each system, three independent 300-ns simulations starting from velocities drawn randomly from the Maxwell-Boltzmann distribution were performed. Analyses were performed with cpptraj (33) and in-house codes.

Binding free energies

Binding free energies were obtained with the Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) method. MM/PBSA calculations were performed considering ligand binding only to the protomer A of the dimer, with the protomer B being considered ligand-bound, as shown in the scheme of **Fig. S31**. The reported free energies correspond to the average based on 450 frames taken from 3 x 300 ns of simulation of each system, corresponding to a frame every 2 ns. The single trajectory protocol was used to estimate the free energies, as implemented in mmpbsa.py (34).

Umbrella Sampling

Potentials of mean forces were obtained with the Umbrella Sampling method (35) along the order parameter defined as the distance between the Arg246 CZ carbon and the fatty acid carboxyl carbon atom. The order parameter was sampled from 3.5 Å to 8.0 Å within 19 windows centered at the following values (in Å): 3.5, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 5.1, 5.5, 5.9, 6.3, 6.7, 7.1, 7.5, 7.9 and 8.3. All windows were biased using a harmonic potential of the form $k(r-r_0)^2$, with k = 50 kcal/mol.Å² for the windows centered between 4.0 Å and 4.7 Å, and k = 15 kcal/mol.Å² for the other windows. To obtain initial configurations for each window, the one centered at 3.9 Å, which is the closest one to the crystal structure, was minimized with 2000 steps and simulated for 200 ps. The final structure was then the starting point for the adjacent windows centered at 3.5 Å and 4.0 Å, where 2000 steps of energy minimization followed by 200 ps of molecular dynamics were performed. The final structure at 4.0 Å was then used as starting point for the window centered at 4.1 Å, where a cycle of 2000 steps of energy minimization followed by 200 ps of molecular dynamics was performed. The last structure at 4.1 Å was then used for the adjacent window at 4.2 Å and so on until a starting point for the window centered at 8.3 Å was obtained. After all starting points were obtained, an additional 4.8 ns of equilibration was performed for each window, totalizing 5.0 ns of equilibration per window. From that, 40 ns of production simulations were run for each window. Biased histograms of the order parameters obtained exhibited enough overlap **(Fig. S17)** to allow an estimate of the potential of mean force using the Weighed Histogram Analysis Method (WHAM). Convergence was assessed by comparing the time evolution of the calculated free energy profiles computed with 10 ns, 20 ns, 30 ns, and 40 ns of sampling per window. Reported error bars correspond to the standard deviation of mean free energy profiles calculated as the average of 4 free energy profiles calculated for blocks of 10 ns.

SUPPLEMENTARY TABLES

Table S1: Extent of the spin-state shift from low to high-spin (%) induced by substrates. Experiments were performed with OleTP$_{RN}$ WT and various mutants with C10H-C20H and C18:1H fatty acids. NP= Not performed.

Variant	C10H	C12H	C14H	C16H	C18H	C20H	C18:1H
WT	85.8±3.8	69.9±1.1	75.0±8.6	82.7±0.9	90.5±0.1	90.0±1.1	72.5±3.5
M305F	81.9±1.9	53.6±1.4	59.9±5.5	56.6±1.6	67.4±9.1	69.1±4.5	NP
F328A	92.0±3.0	79.9±5.9	77.7±7.7	76.0±4.4	91.3±5.1	99.2±3.6	NP
Loop	95.2±1.2	89.8±1.0	93.3±4.5	92.2±6.0	85.6±3.2	89.6±3.4	NP
F17A	<10	83.6±7.0	65.9±9.0	98.0±1.2	93.9±1.3	91.4±3.0	NP
M75A	<10	77.8±11.6	76.5±4.3	<10	81.2±4.4	<10	NP
M305A	<10	80.9±4.7	67.5±1.3	70.0±2.2	75.5±1.7	89.0±0.4	NP
S26F	NP	NP	99.0±1.1	NP	NP	NP	NP
E405F	NP	NP	95.2±1.8	NP	NP	NP	NP
H90Q	<10	<10	<10	<10	<10	<10	NP
V303F	<10	<10	<10	<10	<10	<10	NP
F84A	<10	<10	<10	<10	<10	<10	NP

Table S2: K_D (μM) of saturated C10-C20 and unsaturated C18:1 fatty acids with $OleTP_{RN}$ WT and various mutants. NP= Not performed and NA=Not available.

Variant	C10:0	C12:0	C14:0	C16:0	C18:0	C20:0	C18:1
WT	32.2±2.0	7.1±0.1	0.6±0.2	1.8±0.5	3.1±0.3	8.6±0.1	1.8±0.1
M305F	66.4±15.1	4.2±2.7	1.0±0.5	1.4±0.4	0.9±0.5	2.7±0.8	NP
F328A	31.2±10.9	15.0±2.4	15.9±2.2	12.1±3.0	15.4±6.6	16.9±5.9	NP
Loop	20.2±4.7	1.3±0.2	0.9±0.3	0.7±0.2	0.5±0.1	0.4±0.1	NP
F17A	NA	17.2±5.5	7.5±1.1	14.2±0.5	10.1±2.1	23.2±9.5	NP
M75A	NA	6.0±2.7	4.6±1.3	6.1±0.2	8.4±0.9	15.3±4.1	NP
M305A	NA	22.3±9.9	4.0±1.3	7.7±1.8	7.3±0.3	14.5±2.3	NP
S26F	NP	NP	0.9±0.5	NP	NP	NP	NP
E405F	NP	NP	2.1±0.3	NP	NP	NP	NP
H90Q	NA	NA	NA	NA	NA	NA	NP
V303F	NA	NA	NA	NA	NA	NA	NP
F84A	NA	NA	NA	NA	NA	NA	NP

Table S3: Melting temperature obtained from CD analysis. The buffers employed in the experiments were sodium citrate-phosphate (pH 5-6), sodium phosphate (pH 6.5-8.5) and glycine-NaOH buffer (pH 9-10), all at 50 mM final concentration.

pH	Tm (°C)	SD	Substrate
5.0	42.2	0.6	None
5.5	47.5	0.4	None
6.0	47.7	0.3	None
6.5	57.1	0.3	None
7.0	59.2	0.6	None
7.5	52.2	0.2	None
7.5	66.1	0.6	20 µM C14
7.5	67.6	1.4	200 µM C14
8.0	40.9	0.9	None
8.5	52.4	0.8	None
9.0	45.6	0.8	None
9.5	45.0	0.6	None
10.0	39.9	1.0	None

Table S4: Alkene production (%) among total fatty acid conversion.

Variant	C10:0	C12:0	C14:0	C16:0	C18:0	C20:0	C18:1
WT	70.8±2.5	73.9±3.3	68.6±1.8	66.6±0.3	69.2±1.8	65.9±2.0	71.4±0.7
F17A	70.8±0.2	72.5±4.5	71.3±2.0	67.9±0.9	69.4±2.9	64.4±2.9	n.d.
M305A	58.3±2.5	65.2±1.4	64.1±1.0	65.3±0.5	65.8±1.5	63.4±1.1	n.d.
M305F	63.5±3.6	72.0±4.3	68.1±3.0	64.6±3.5	64.6±0.3	64.2±3.8	n.d.
F328A	76.9±3.0	61.8±1.6	59.9±1.3	67.5±0.5	73.1±3.1	66.7±0.2	n.d.
Loop	45.0±0.3	50.7±1.8	51.1±1.9	57.2±0.2	63.8±3.9	53.4±0.7	n.d.
H90Q	n.d.	n.d.	29.5±0.4	n.d.	n.d.	n.d.	n.d.
F84A	n.d.	n.d.	35.5±2.4	n.d.	n.d.	n.d.	n.d.
N243A	n.d.	n.d.	37.0±1.1	n.d.	n.d.	n.d.	n.d.
S26F	n.d.	n.d.	68.5±1.3	n.d.	n.d.	n.d.	n.d.
E405F	n.d.	n.d.	68.5±1.4	n.d.	n.d.	n.d.	n.d.
V303F	n.d.	n.d.	30.9±1.0	n.d.	n.d.	n.d.	n.d.
Y22H	n.d.	n.d.	62.9±2.2	n.d.	n.d.	n.d.	n.d.
Y22W	n.d.	n.d.	1.4±0.1	n.d.	n.d.	n.d.	n.d.

Table S5: Ratio of alkenes to hydroxylated products obtained by $OleTP_{RN}$ in the reaction with saturated and unsaturated fatty acids and hydrogen peroxide.

Chemoselectivity (Ratio of alkene to total hydroxylated products)							
C10:0	C12:0	C14:0	C16:0	C18:0	C20:0	C18:1	C18:2
1.9	2.4	2.2	2.0	2.1	1.9	2.5	3.1

Experimental conditions: 2 µM enzyme, 500 µM substrate, 600 µM H_2O_2, 35 ºC and 1 hour of reaction.

Table S6: Rate constants from different single turnover reactions by $OleTP_{RN}$.

	370 nm		440 nm		690 nm
FA	**k_1, s^{-1} (a_1)**	**k_2, s^{-1} (a_2)**	**k_1, s^{-1} (a_1)**	**k_2, s^{-1} (a_2)**	**k_1, s^{-1} (a_1)**
C10D	144±7 (0.09)	1.4±0.4 (0.05)	151±10 (-0.2)	19±1 (0.1)	154±3 (0.03)
C12D	112±8 (0.07)	6.2±1.5 (0.03)	85±2 (0.2)	23.3±0.3 (0.1)	108±5 (0.03)
C14D	120±3 (0.1)	4.4±0.3 (0.03)	96±5 (-0.3)	19.1±3 (0.2)	120±5 (0.05)
C16D	150±4 (0.1)	5.1±0.1 (0.03)	119±6 (-0.3)	20.6±0.8 (0.2)	152±3 (0.05)
C18D	189±2 (0.1)	4.6±0.2 (0.04)	119±5 (-0.3)	25.4±0.1 (0.2)	161±6 (0.05)
C20D	216±6 (0.09)	5±1 (0.04)	125±3 (-0.2)	29.4±1.8 (0.1)	167±1 (0.04)

Table S7: List of primers employed for site-directed mutagenesis.

Primer name	Sequence (5' - 3')
RN_H90Q_F	CCGTGCAGACCCTCGACGGTCAG
RN_H90Q_R	CGAGGGTCTGCACGGCACCGTCG
RN_F17A_F	GCCCAATGCGATTAAGCAGGGCTACC
RN_F17A_R	CTTAATCGCATTGGGCAGCTGGTCATTC
RN_F328A_F	CGGGTGGCCCTTGATGTTTTAGGTAC
RN_F328A_R	ATCAAGGGCCACCCGCTGG
RN_M305A_F	GTCCCCGCGCTTCCCGCTATGG
RN_M305A_R	GGGAAGCGCGGGGACGAAGGGGTAC
RN_M305F_F	GTCCCCTTTCTTCCCGCTATGGTTAAGGAAG
RN_M305F_R	GGGAAGAAAGGGGACGAAGGGGTACTTGC
RN_loopOleT$_{JE}$_F	GCTCTGGGTGGTGCTTTCAAGCAGAACGTGCTCTCCTG
RN_loopOleT$_{JE}$_R	AGCACCACCCAGAGCACGGAACGTGTCGAGCAGGTGG
RN_S26F_F	TTTATGTTCCACCAGCGCCGTAAGGCC
RN_S26F_R	CTGGTGGAACATAAAAAGGTAGCCCTGCTTAATG
RN_E405F_F	TCTCAATTTACCGAAGACACCACTTTCAGTATG
RN_E405F_R	TTCGGTAAATTGAGAGATGGTGACGTTATCAG
RN_V303F_F	CCCTTCTTTCCCATGCTTCCCGCTATGG
RN_V303F_R	CATGGGAAAGAAGGGGTACTTGCGGCGG
RN_Y22H_F	GCAGGGCCACCTTTTTATGTCCCACC
RN_Y22H_R	AAAAGGTGGCCCTGCTTAATGAAATTGGGCAG
RN_Y22W_F	CAGGGCTGGCTTTTTATGTCCCACCAG
RN_Y22W_R	AAAAAGCCAGCCCTGCTTAATGAAATTGGGC

Table S8: Data collection and refinement statistics for crystallographic data. Values in parenthesis represent the higher resolution shell.

	OleTP$_{RN}$
PDB Code	8D8P
Data collection	
Wavelength (Å)	1.459
Detector	Dectris Pilatus 2M
Temperature (K)	100
Space group	F 2 2 2
Cell dimensions	
a, *b*, *c* (Å)	57.55, 203.38, 316.35
α=β=γ (°)	90.00
Resolution (Å)	48.41 - 2.75 (2.85 - 2.75)
R_{sym}	0.17 (1.18)
I / σ(I)	6.12 (0.92)
$CC_{1/2}$ (%)	0.996 (0.803)
Completeness (%)	99.81 (99.59)
Redundancy	5.86 (5.99)
Refinement	
Resolution (Å)	48.41 - 2.75
No. reflections	24494
R_{work} / R_{free}	0.228/0.269
No. atoms	
Protein	6514
Ligands (2 HEM, 2 PLM and 12 SO_4)	182
HEM - Protoporphyrin IX containing Fe (HEME)	
PLM – Palmitic acid (C16:0)	
Water	25
B-factors (Å^2)	
Protein	78.56
Ligand	78.64
Water	62.03
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	0.840
Ramachandran analysis	
Favored (%)	96.41
Allowed (%)	3.23
Outliers (%)	0.36

SUPPLEMENTARY FIGURES

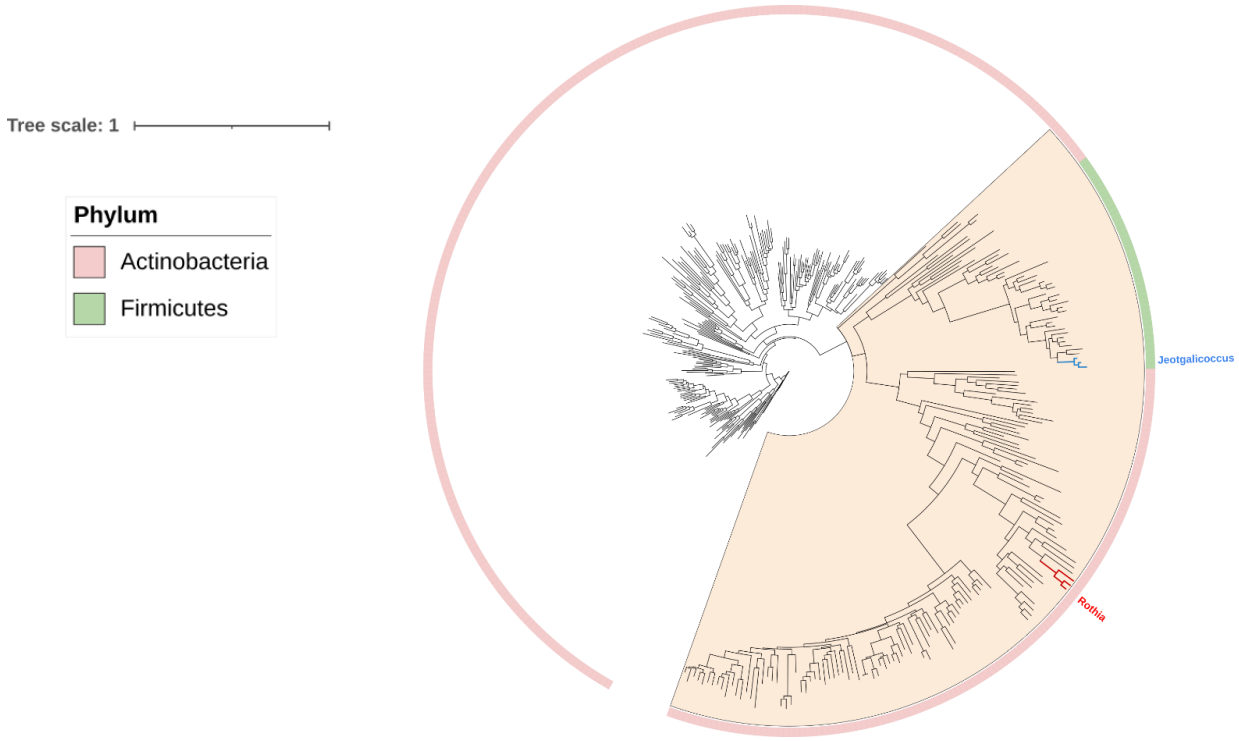


Figure S1: Full phylogenetic tree of putative fatty acid P450 decarboxylases obtained from HMM analysis. Sequences were trimmed by the conservation of amino acids important for decarboxylase activity (Phe79, His85, and Arg245 for $OleT_{JE}$). Nodes highlighted in salmon were employed for the phylogenetic relationship displayed in the main text Fig.1. *Jeotgaliccocus* and *Rothia* genera branches are indicated in blue and red, respectively.

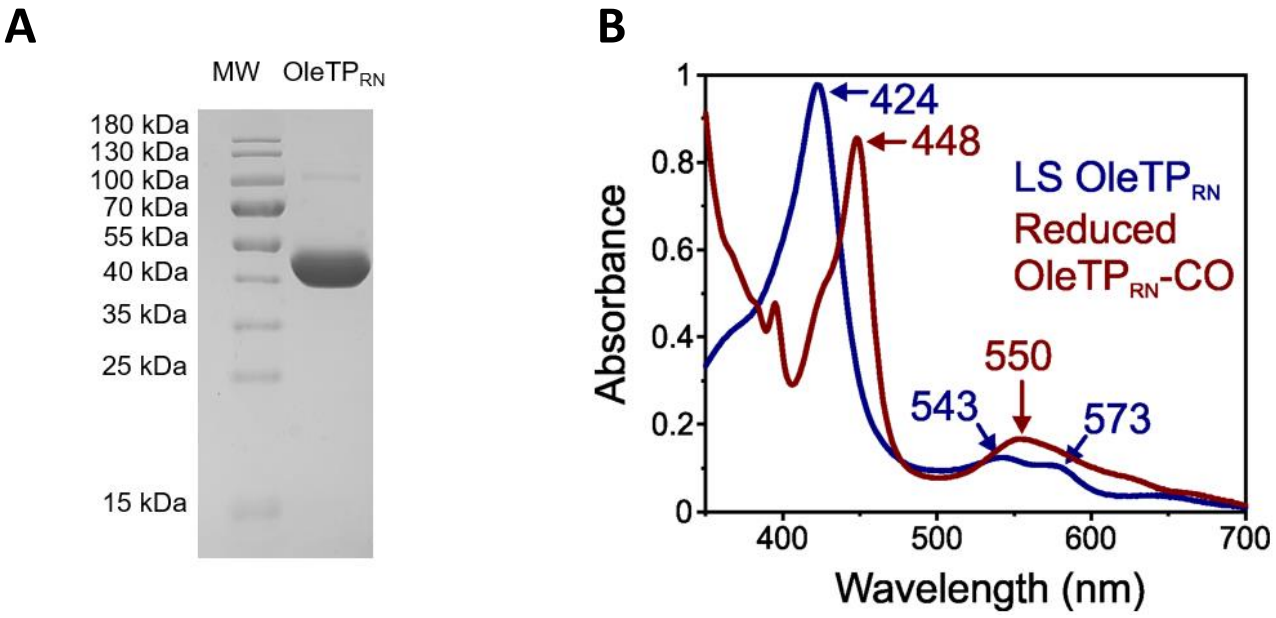


Figure S2: (A) SDS-PAGE gel image shows the purified forms of OleTP$_{RN}$. **(B)** Typical UV-visible absorption spectra for the ferric, substrate-free forms of OleTP$_{RN}$ (blue) showing the heme Soret (424 nm) and Q-bands (543 and 573 nm); the Fe^{II}-CO complex of OleTP$_{RN}$ (red) showing a characteristic Soret at 450 nm.

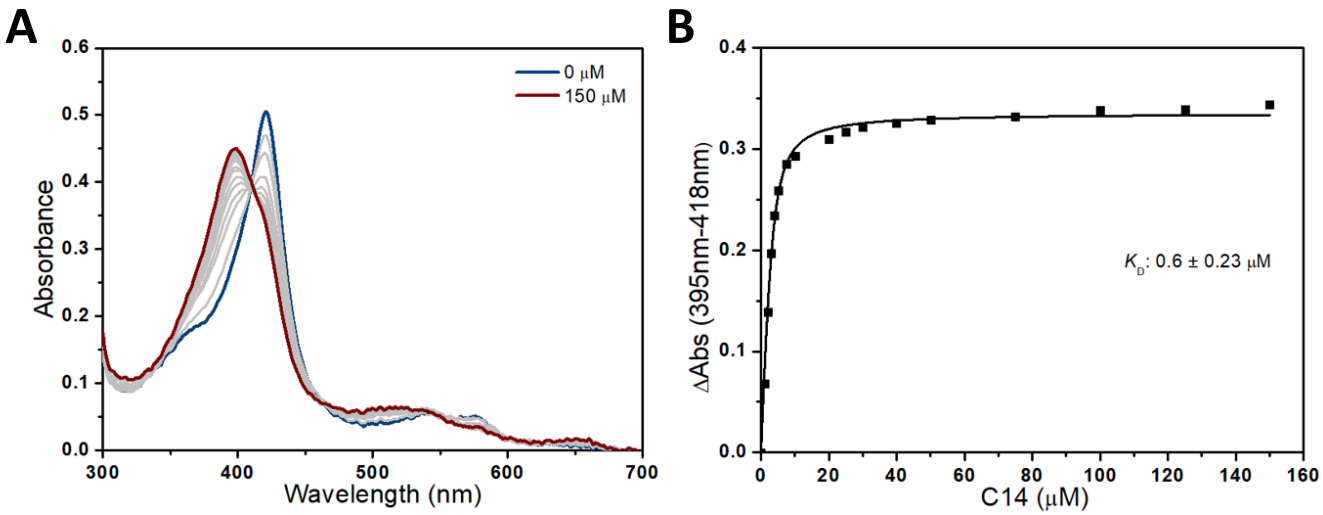


Figure S3: Representative fatty acid binding titration of $OleTP_{RN}$ with myristic acid (C14). **(A)** UV-Vis spectra obtained through substrate titration, demonstrating the conversion of LS (424 nm) to HS heme iron (395 nm). **(B)** Absorption changes (395-418 nm) values versus C14 concentration were fit using a hyperbolic function to determine a dissociation constant (K_D) ~0.6 μM.

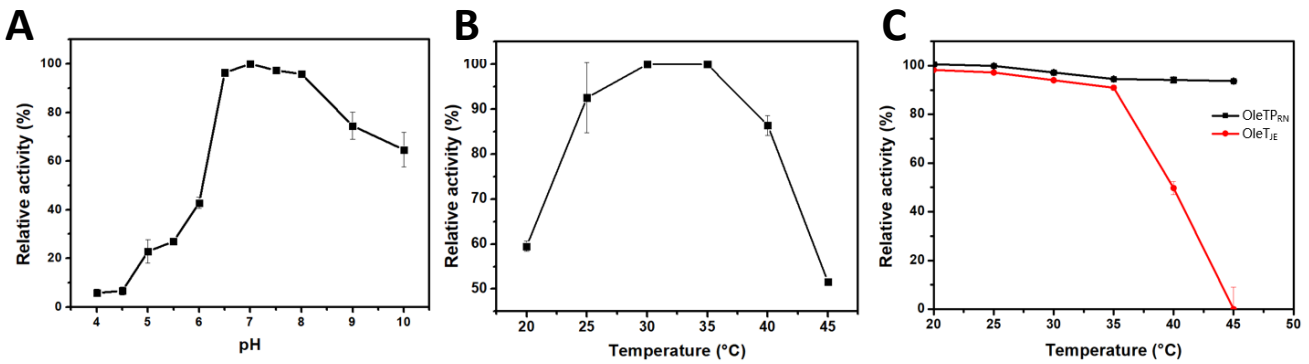


Figure S4: Biochemical characterization of the OleTP$_{RN}$. Effects of pH **(A)** and temperature **(B)** on the relative activity of the enzyme. Thermo stability of OleTP$_{RN}$ and OleT$_{JE}$ (1 µM) after one hour pre-incubation at different temperatures in the best-defined buffer for enzymatic activity **(C)**. Experiments were conducted with 200 µM stearic acid, 0.5 µM enzyme and 100 µM H_2O_2, for 5 min. The remaining H_2O_2 was measured by the ferrous oxidation−xylenol orange (FOX) assay. Results are expressed as mean ± SD from three independent experiments (n=3). For the pH profile, reactions were performed at 30 °C, and 50 mM of each following buffers: citrate (pH 4.0, 4.5, 5.0, and 5.5), phosphate (pH 6.0, 6.5, 7.0 and 8.0), Tris–HCl (pH 9.0), and glycine (pH 10.0). OleT$_{JE}$ enzymatic conditions followed the described by (36) [100 mM KPhosphate, 750 mM NaCl, pH 8.0].

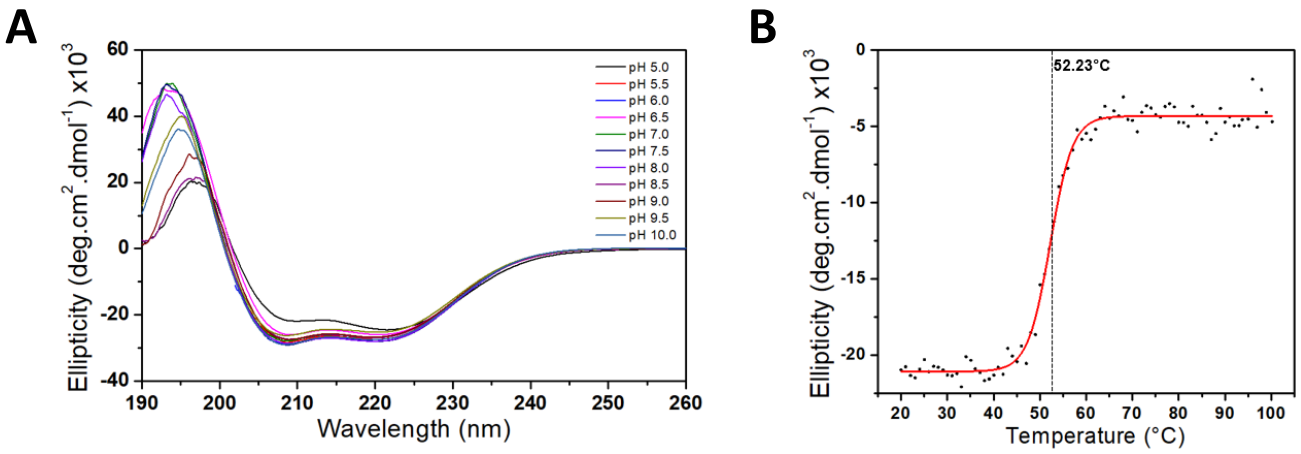


Figure S5: Circular dichroism (CD) analysis of OleTP$_{RN}$ regarding the employed pH buffer. **(A)** Ellipticity over a wavelength range of 190–260 nm at 20°C, with a pH range from 5 to 10. **(B)** Thermal unfolding profile of OleTP$_{RN}$ following ellipticity at 222 nm. The buffers employed in the experiments were sodium citrate-phosphate (pH 5-6), sodium phosphate (pH 6.5-8.5) and glycine-NaOH buffer (pH 9-10), all at 50 mM final concentration.

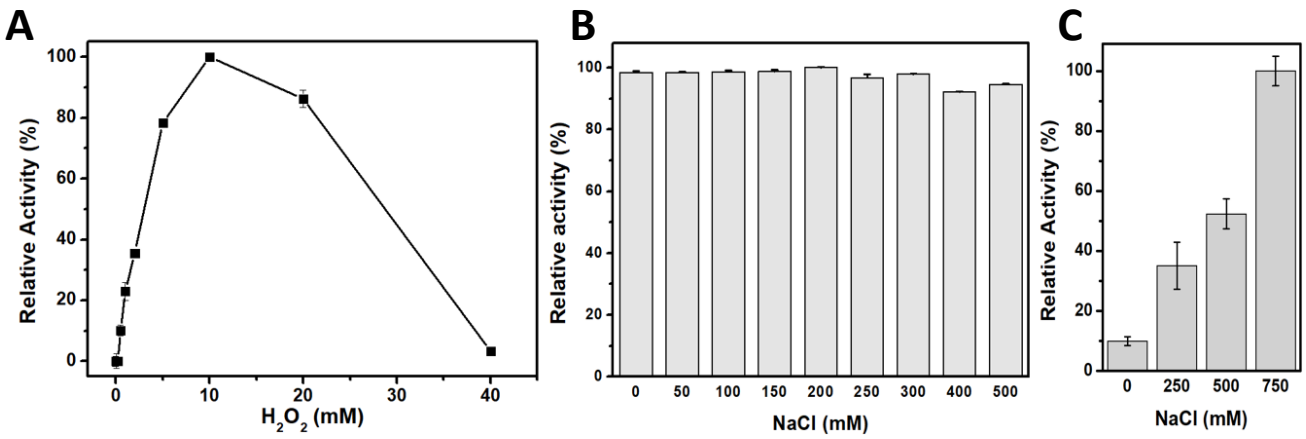


Figure S6: Oxygen peroxide tolerance of $OleTP_{RN}$ **(A)** and salt dependence of $OleTP_{RN}$ **(B)** and $OleT_{JE}$ **(C)**. Effects of H_2O_2 and NaCl concentration on the relative activity of the enzymes. Experiments were conducted with 100 µM stearic acid, 1 µM enzyme and 100 µM H_2O_2, if not specified. The remaining H_2O_2 was measured by ferrous oxidation−xylenol orange (FOX) assay. Results are expressed as mean ± SD from three independent experiments (n=3). The measurements were expressed as relative activity (%), considering the highest catalytic activity measured (10 mM H_2O_2, 200 mM NaCl for $OleTP_{RN}$ and 750 mM for $OleT_{JE}$).

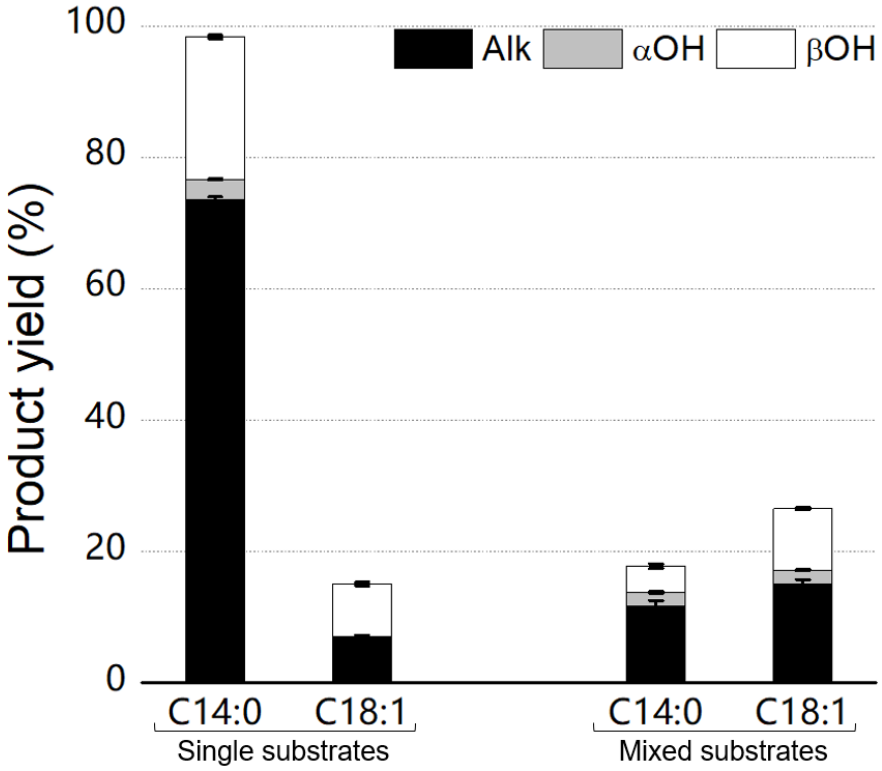


Figure S7: The inhibitory effect of oleic acid on $OleT_{JE}$ activity. The product yield of turnover reactions of fatty acid substrates with $OleT_{JE}$. The total oleic acid consumption was around 15% (TON = 32), of which only 47 % for alkene production for the single substrate. For the reactions performed with mixed myristic and oleic acids, the total conversion of myristic drastically decreased. Enzyme reactions were performed following conditions defined in (36) with 500 µM myristic and oleic acid as single substrates and 250 µM each substrate in one reaction. Reactions were performed with 2 µM enzyme, 500 µM total substrate, 1 mM final H_2O_2 (rate 100 µM per 5 min) at 35°C.

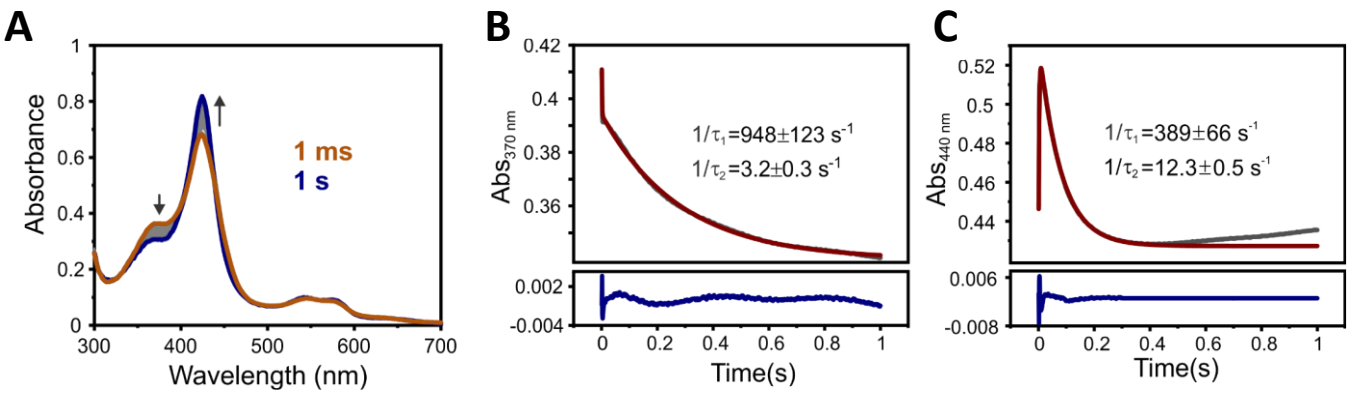


Figure S8: (A) Photodiode array spectra for 10 µM C14H bound $OleTP_{RN}$ WT and 10 mM H_2O_2 monitored for 1 s at 4°C. Arrows indicate the direction of decreasing and increasing absorbances during the reactions. Exponential fits of 370 nm **(B)** and 440 nm **(C)** PMT traces of the single turnover reaction. Actual PMT data is in dark grey, two exponential fittings are in red, and the residuals are colored in navy blue. Numbers in the figures represent the reciprocal relaxation time (RRT) constants associated with Compound I decay **(B)** and Compound II formation and decay **(C)**. 370 nm PMT trace was fitted up to 1 s **(B)** and 440 nm PMT trace was fitted up to only 300 ms **(C)**.

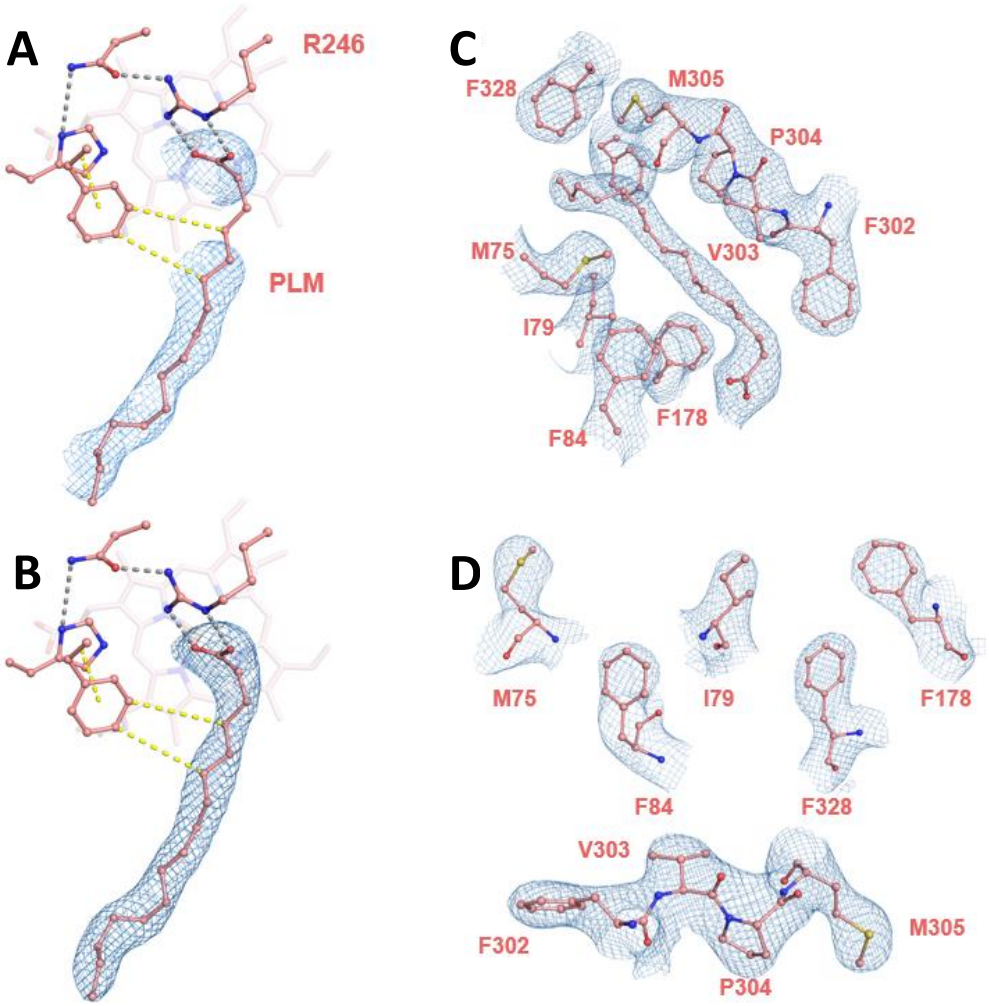


Figure S9: Electron density map of OleTP$_{RN}$ active site accommodating a palmitate substrate (PLM). The composite omit **(A)** and 2mFo – dFc **(B)** maps highlight the ligand density fit. 2mFo – dFc map highlighting the catalytic residues **(C)** and the amino acids that closely interact with palmitate **(D)**. All maps are displayed in blue and contoured at 1.0 σ.

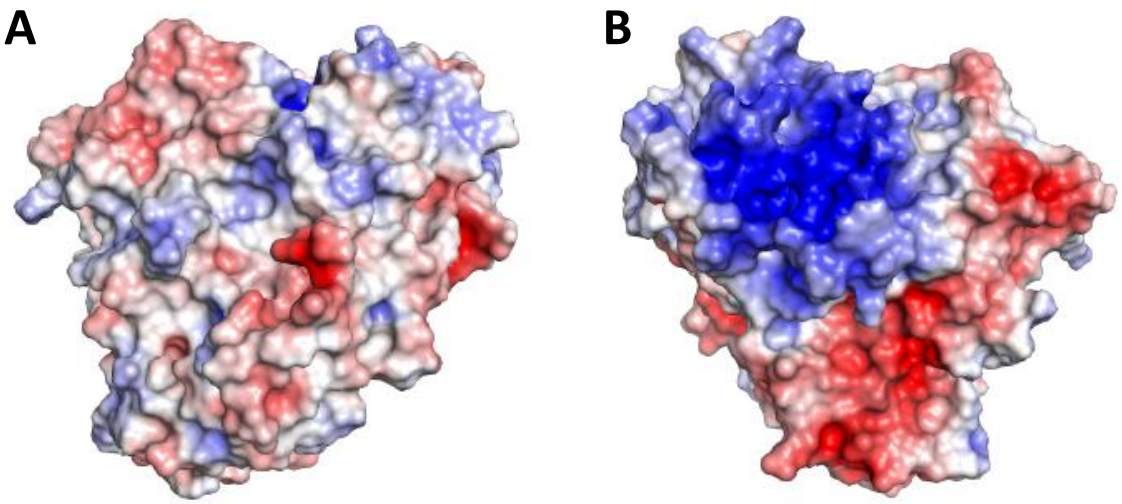


Figure S10: Electrostatic potential surface of $OleTP_{RN}$. Proximal **(A)** and distal **(B)** sides are shown. The negatively charged regions are represented in red, while the positively charged regions are in blue.

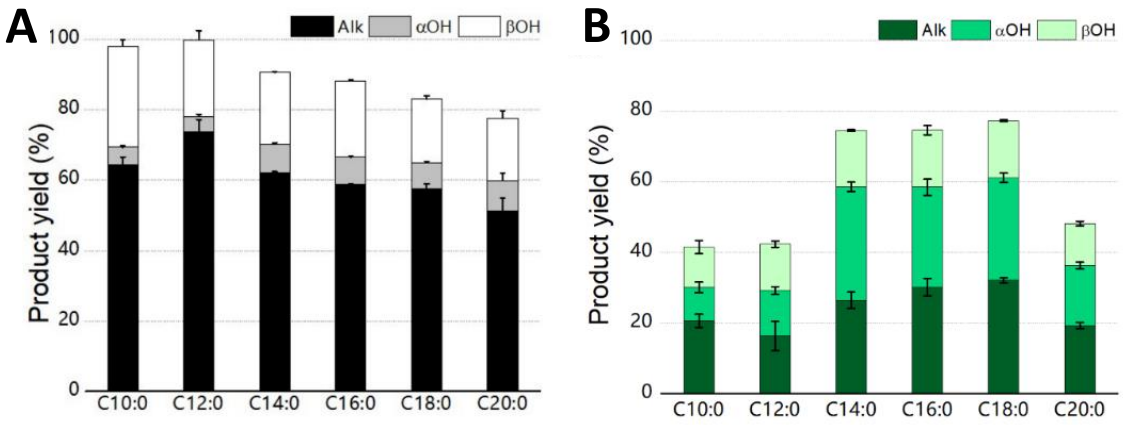


Figure S11: Product yield of turnover reactions with the $OleTP_{RN}$ wild-type **(A)** and F84A mutant **(B)**. Results for the wild-type enzyme were repeated out of the main text to facilitate visualization of the conclusions.

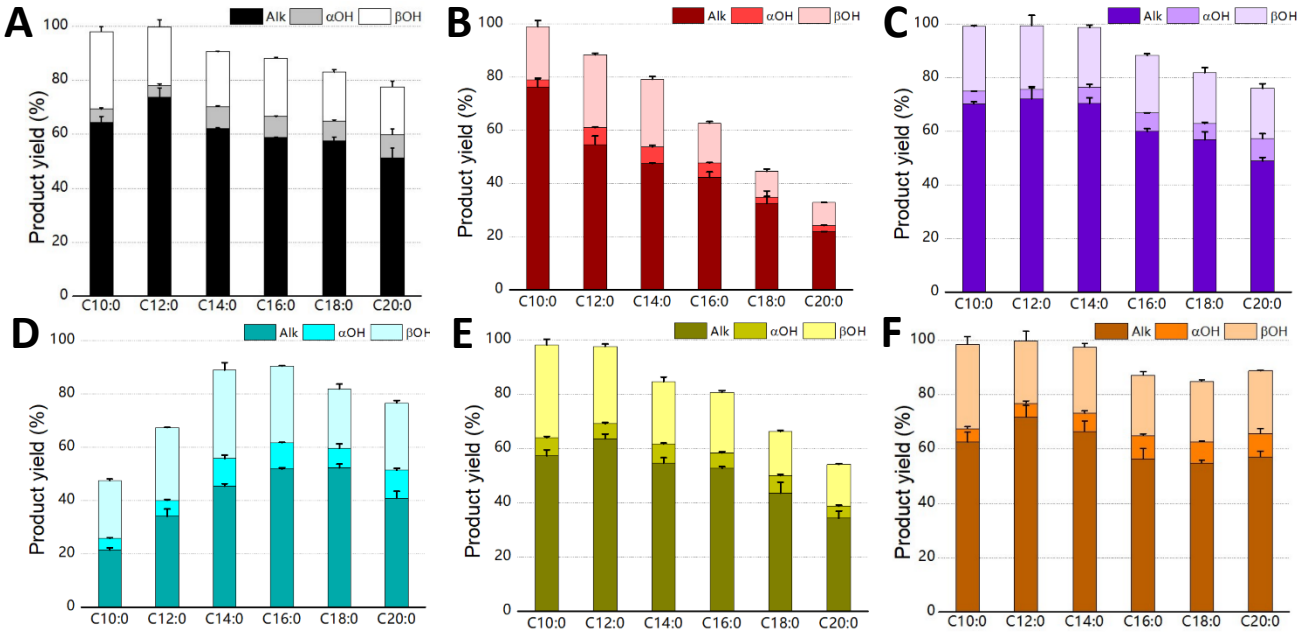


Figure S12: Product yield of turnover reactions with the OleTP$_{RN}$ wild-type and mutants involved in substrate binding. WT **(A)**, F328A **(B)**, F17A **(C)**, Loop **(D)**, M305A **(E)**, M305F **(F)**. Results for the wild-type enzyme were repeated out of the main text to facilitate visualization of the conclusions.

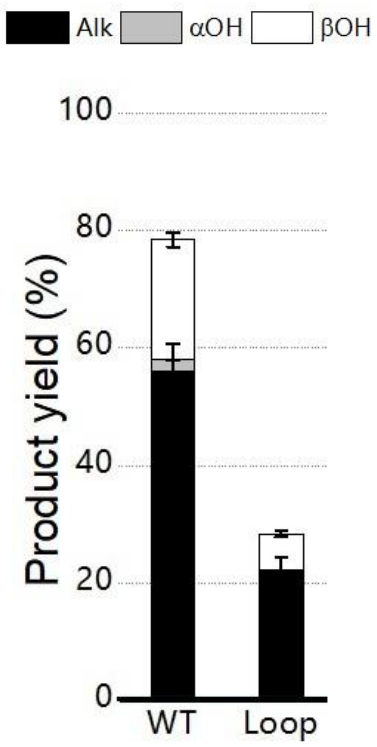


Figure S13: Comparison of the product yield of turnover reactions of oleic acid with the $OleTP_{RN}$ wild-type and mutant with the native F-G loop replaced by the FRALGGAFK sequence.

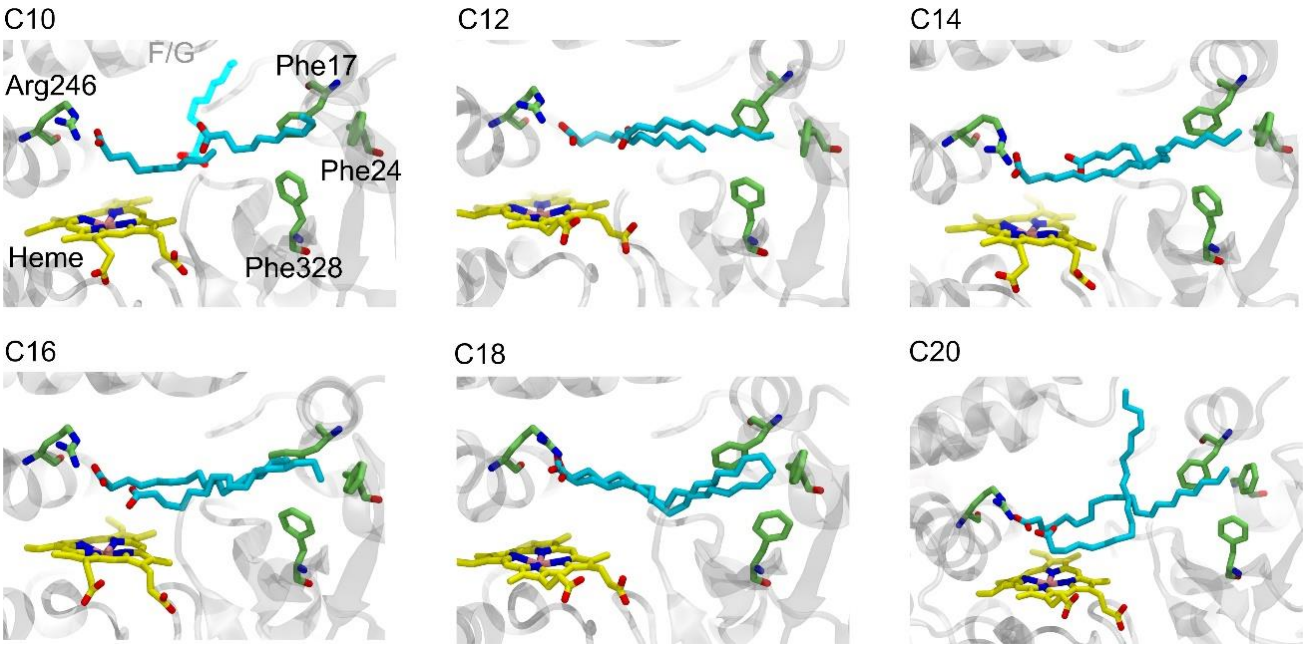


Figure S14: Representative snapshots taken from MD simulations of OleTP$_{RN}$ complexed to C10-C20 substrates, showing substrate configurations similar to the one observed in the crystal structure (where the FA interacts with Arg246) and other configurations that diverge from the crystal structure, which include pseudo-bound states (where the interaction Arg246-FA is disrupted), especially for C10-C14 substrates, and bound states with the hydrophobic chain projected towards the exit channel.

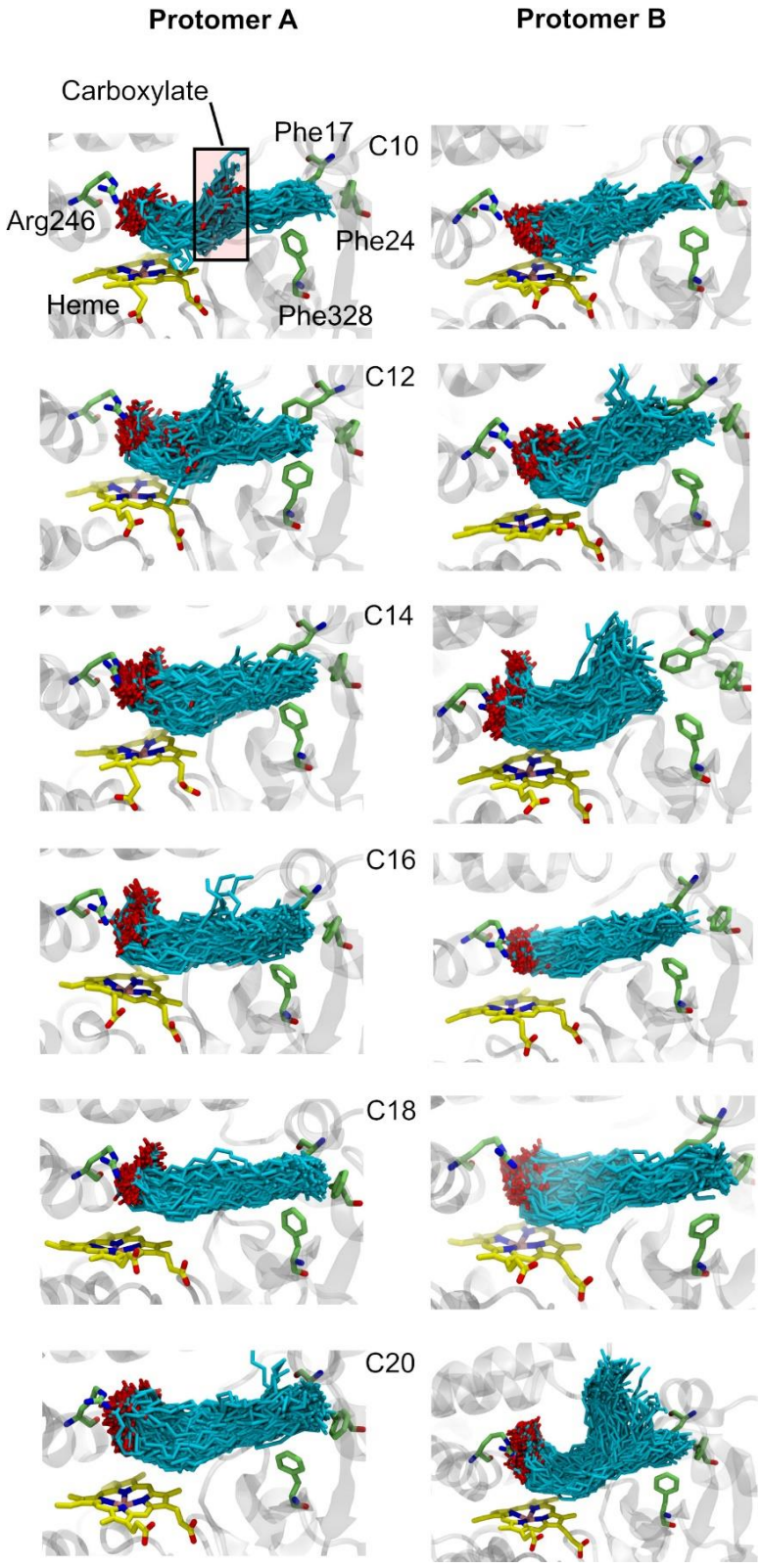


Figure S15: Dynamic view of different fatty acids within the binding site of OleTP$_{RN}$. Conformations of the fatty acids visited during the simulations are shown superposed in cyan.

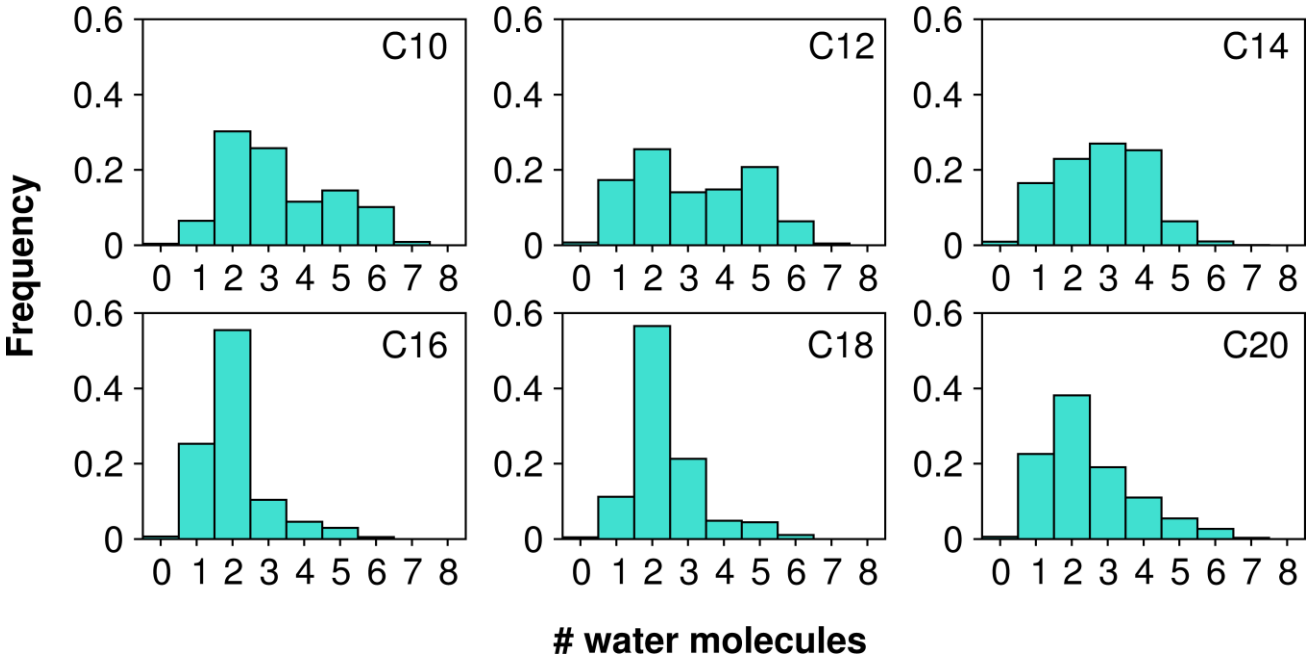
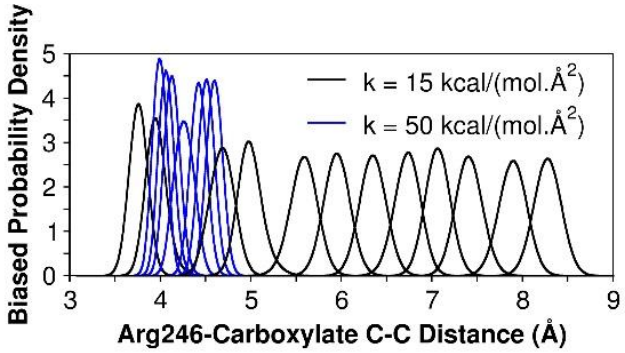
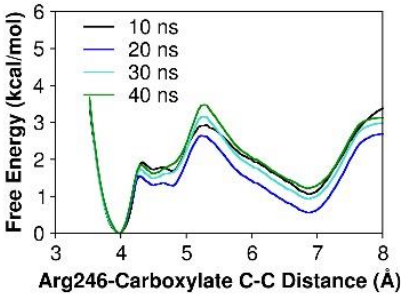


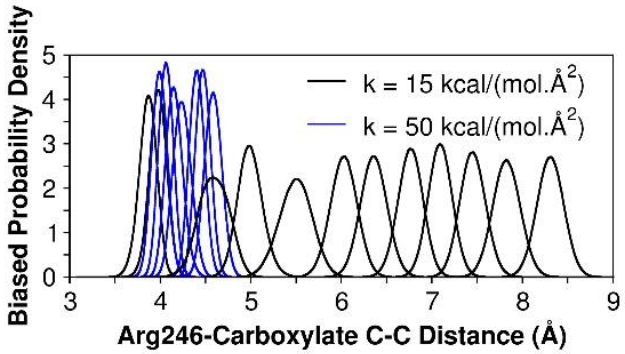
Figure S16: Number of water molecules found within 3.5 Å of the carboxylate group of fatty acids.

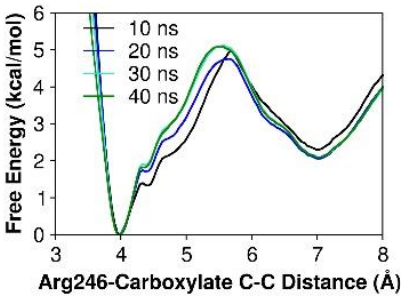
C10



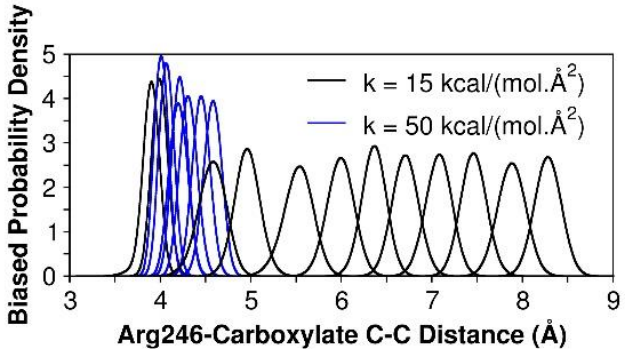


C14





C18



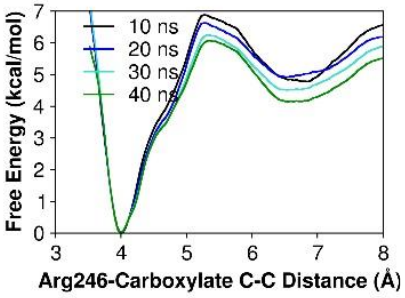


Figure S17: Biased histograms obtained with Umbrella Sampling simulations, showing enough overlap between the distribution of adjacent windows, and free energies profiles computed for 10, 20, 30 and 40 ns of sampling, showing the convergence of the results.

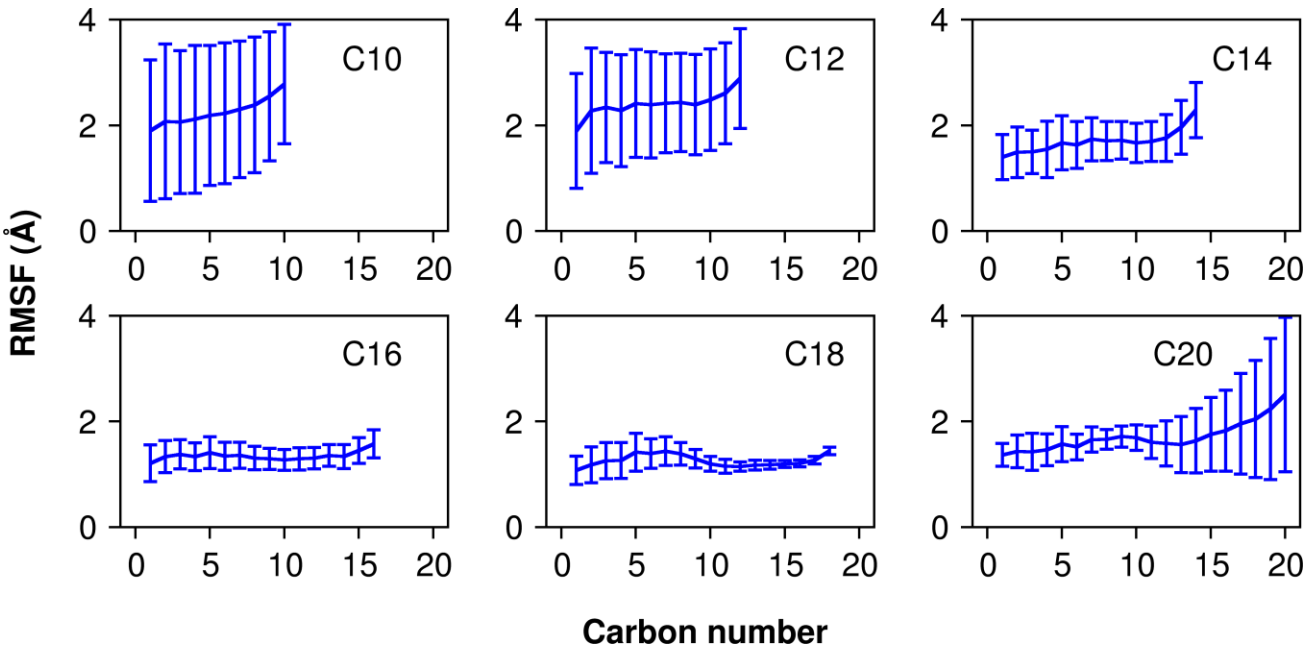


Figure S18: RMSF (root mean squared fluctuations) calculated for each carbon atom of each fatty acid. Carbon number 1 is the one closest to the carboxylate group. Error bars correspond to the standard deviation for three independent simulations. It is observed that C10 and C12 exhibit RMSF mostly above 2 Å and large error bars, while C14 – C20 exhibit RMSF between 1-2 Å for most of the carbon chain, indicating that these fatty acids are better accommodated in the enzyme binding site. C20, while very well accommodated in the region closer to the polar part of the substrate, exhibits large fluctuations at the end of the hydrophobic chain, suggesting limited room for accommodating a 20-carbon chain.

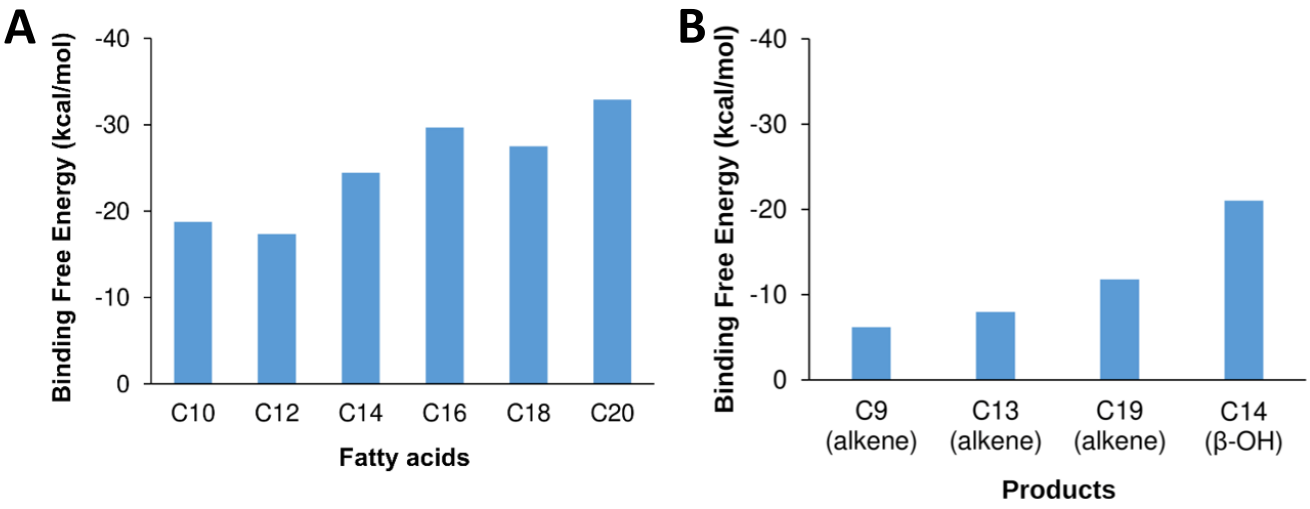


Figure S19: MM/PBSA binding free energies estimated for fatty acids **(A)** and products **(B)**.

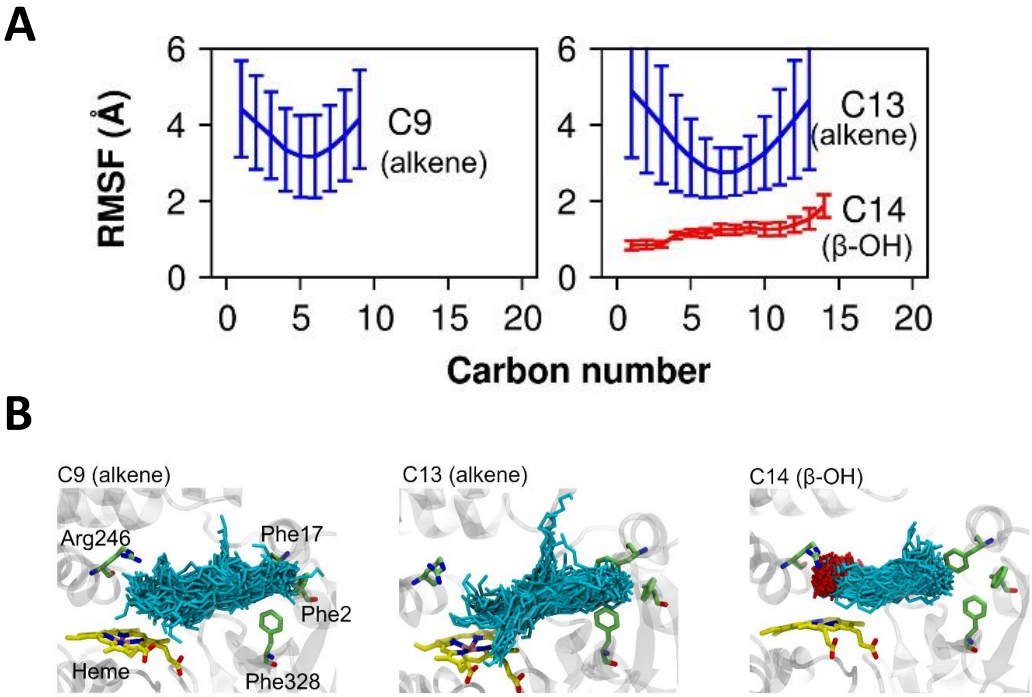


Figure S20: (A) Root mean squared fluctuations (RMSF) of each carbon atom of the alkene and β-hydroxylated fatty acids products. **(B)** Dynamic pictures showing a superposition of ligand conformations as visited during the MD simulations. While alkenes are mobile and can visit the exit channel, the C14 β-hydroxylated fatty acid keeps tightly bound to the enzyme.

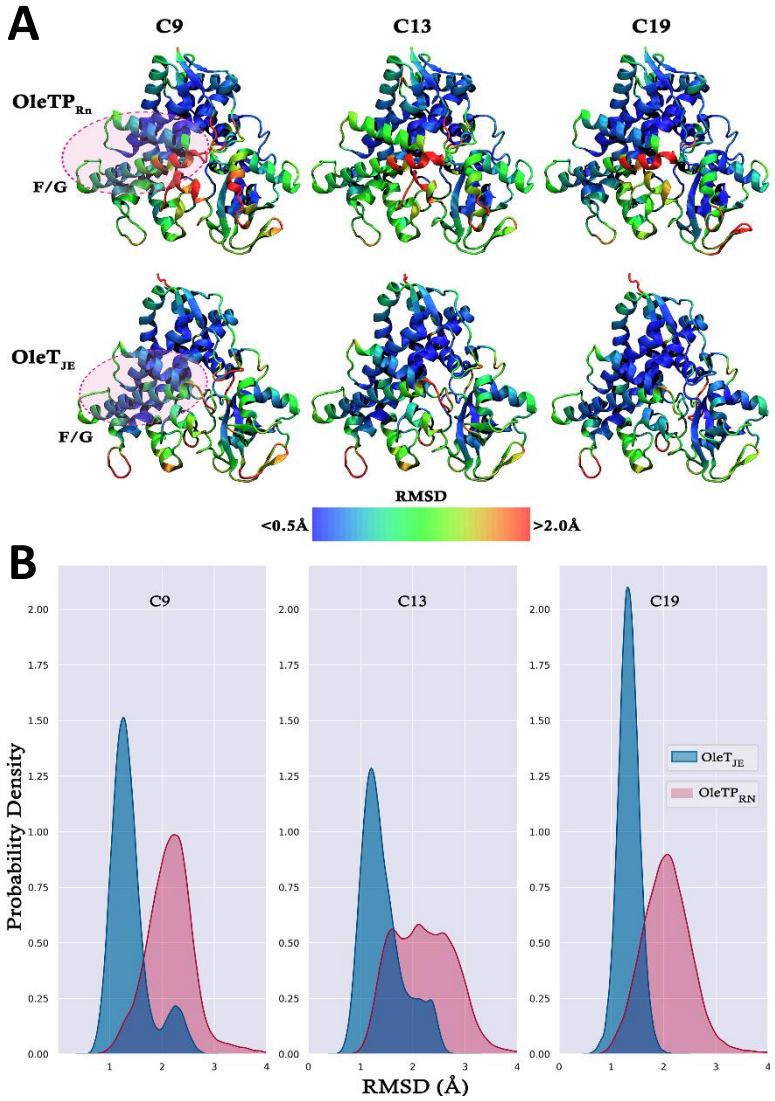
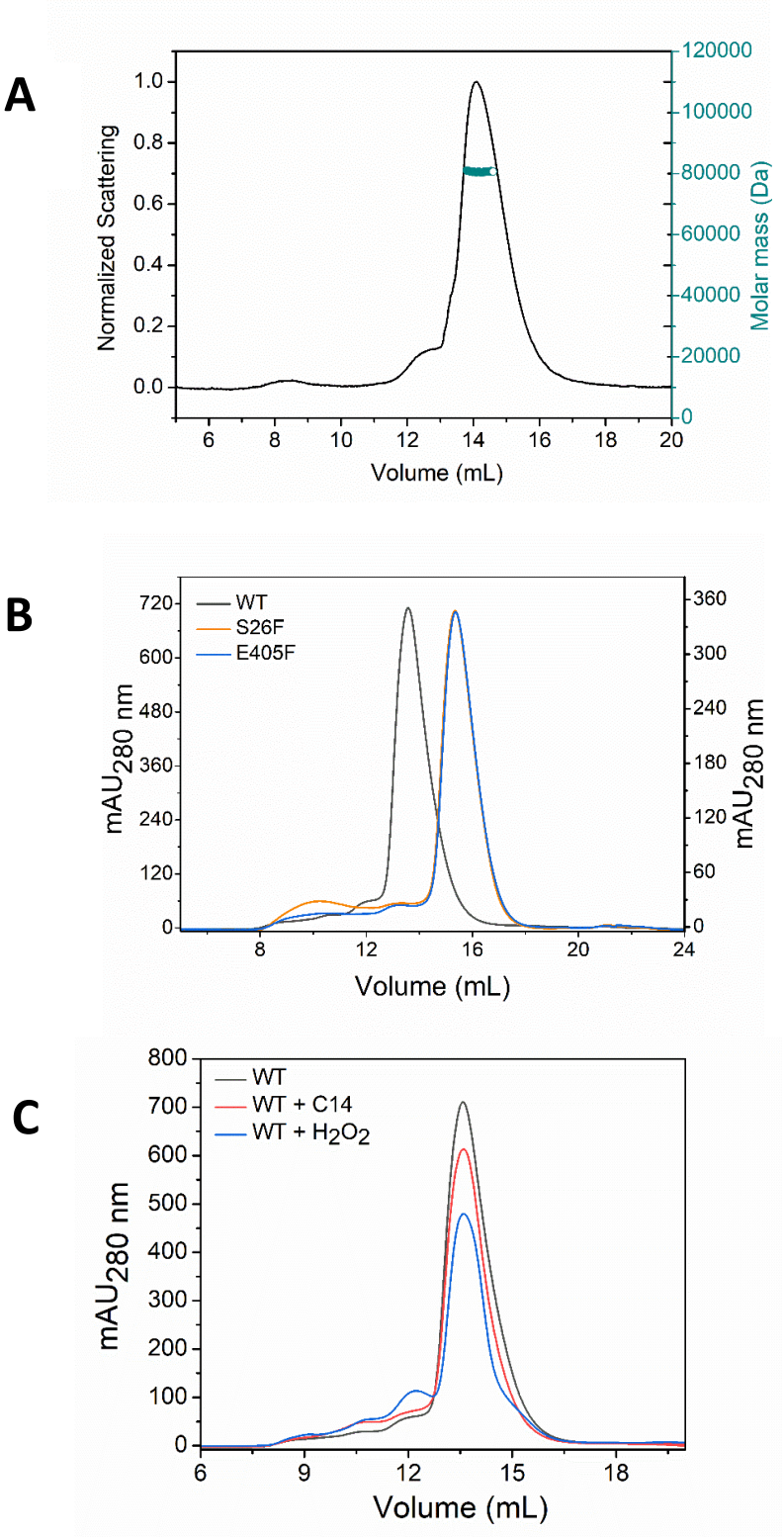


Figure S21: (A) Average structures of OleTP$_{RN}$ and OleT$_{JE}$ in the presence of C9, C13, and C19 alkenes colored according to the alpha carbon fluctuations (RMSD). From blue to red, the colors represent RMSD values from less than 0.5 Å to greater than 2.0 Å. **(B)** RMSD distribution (probability density) computed specifically for the F/G helices and loops in relation to their crystal structures, showing a higher mobility of this region in OleTP$_{RN}$ when compared to OleT$_{JE}$.



Supplementary Fig. 22: (A) Multi-angle light scattering coupled with size exclusion chromatography (SEC-MALS) analysis of OleTP$_{RN}$, showing MALS estimated molecular weight matching to protein dimer (cyan line). **(B)** Size exclusion chromatography of the OleTP$_{RN}$ WT (black), S26F (orange) and E405F (blue) mutants. Protein profile was measured at 280 nm. **(C)** OleTP$_{RN}$ elution from SEC in the presence of 500 µM C14 substrate (red) and 10 mM hydrogen peroxide (blue).

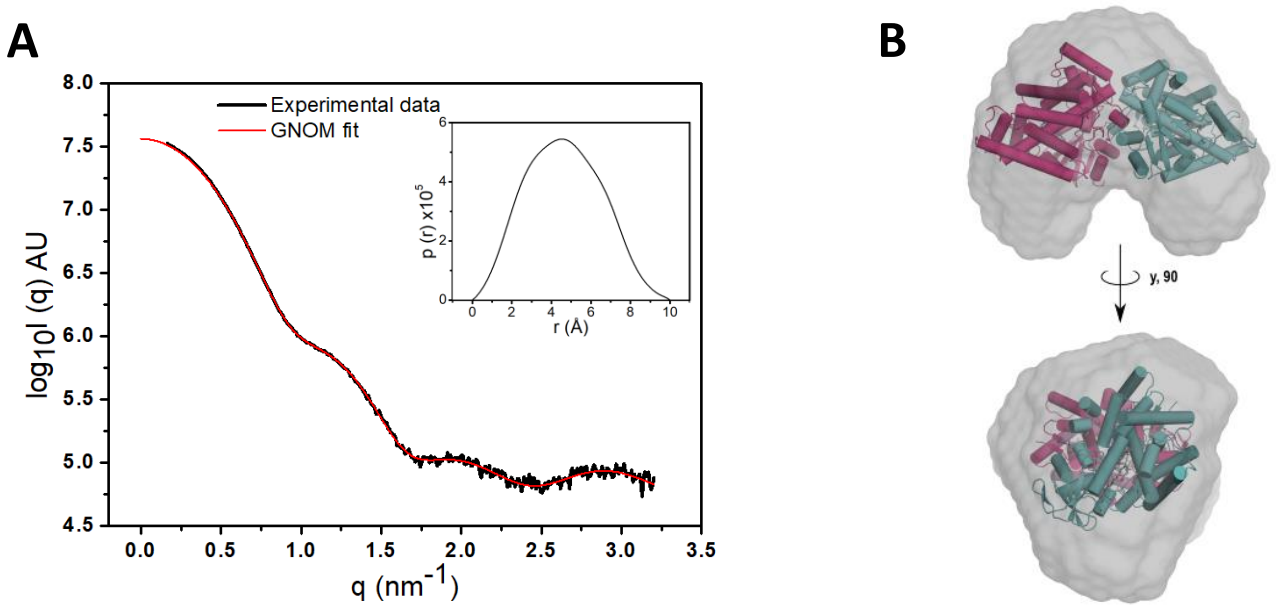


Figure S23: SAXS data of OleTP$_{RN}$. **(A)** Experimental SAXS curve (red) and theoretical scattering profile (black line) computed from the p(r) function, obtained with GNOM program. **(B)** Crystal structure of the OleTP$_{RN}$ fitted into the molecular envelope calculated from SAXS data.

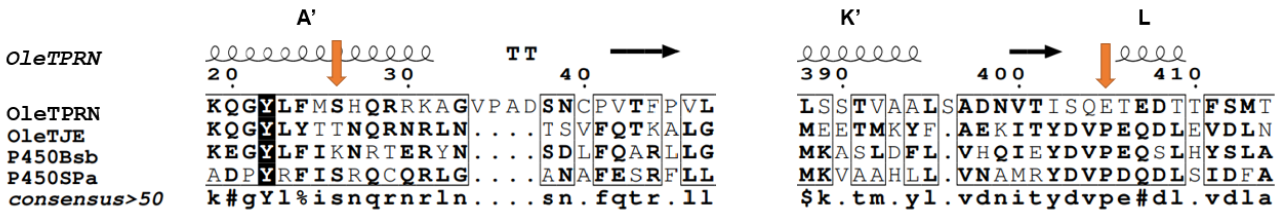


Figure S24: Sequence alignment of OleTP$_{RN}$ and members of CYP152 with oligomerization already described. Identical residues are highlighted by a black background. Orange arrows indicate residue mutations that hindered protein dimerization. For clarity purposes, only relevant sequence fragments were displayed.

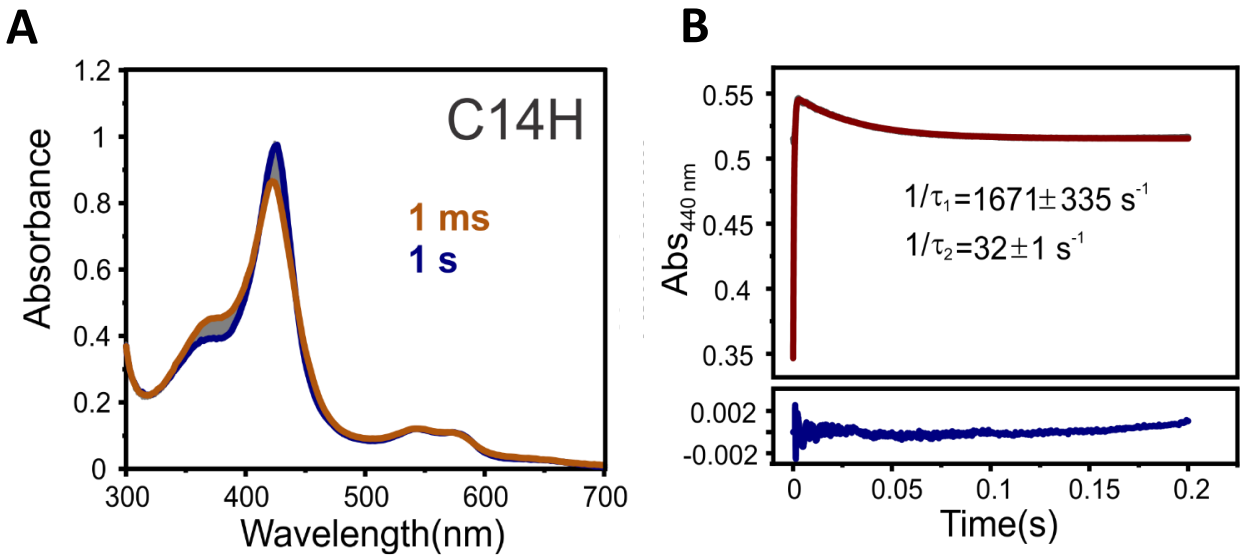


Figure S25: (A) Photodiode array (PDA) spectra of the single turnover reaction between 10 µM C14H bound $OleTP_{RN}$ S26F and 10 mM H_2O_2 monitored for 1 s at 4°C. Arrows indicate the direction of decreasing and increasing absorbances during the reactions. **(B)** Two-exponential fit (up to 200 ms) of the 440 nm PMT trace collected during the single turnover reaction of C14H single turnover by the same mutant.

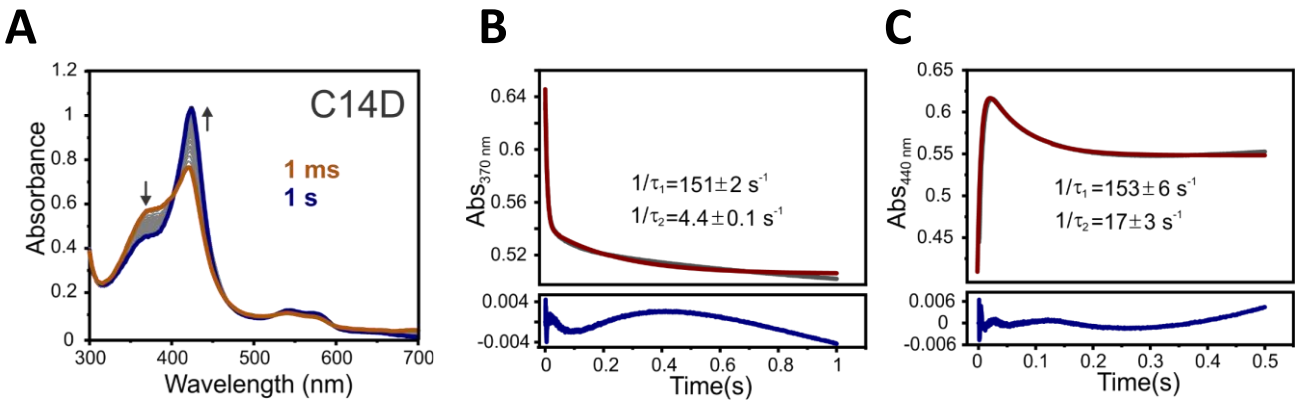


Figure S26: (A) PDA spectra of the single turnover reaction between 10 µM C14D- $OleTP_{RN}$ S26F and 10 mM H_2O_2 monitored for 1 s at 4°C. Arrows indicate the direction of decreasing and increasing absorbances during the reactions. Two exponential fits of 370 nm **(B)** (fitted up to 1 s) and 440 nm **(C)** (fitted up to 500 ms) PMT traces were collected during the single turnover reaction of C14D bound $OleTP_{RN}$ S26F mutant with hydrogen peroxide.

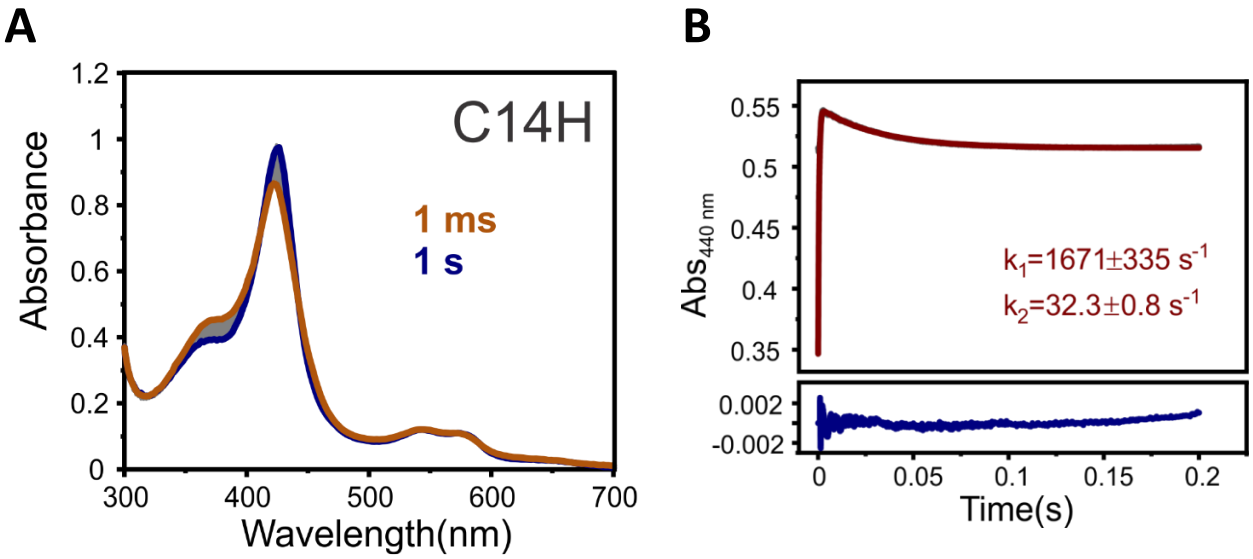


Figure S27: (A) Photodiode array (PDA) spectra of the single turnover reaction between 10 µM C14H bound OleTP$_{RN}$ E405F and 10 mM H_2O_2 monitored for 1 s at 4°C. Arrows indicate the direction of decreasing and increasing absorbances during the reactions. **(B)** Two-exponential fit (up to 200 ms) of 440 nm PMT trace collected during the single turnover reaction of C14H single turnover by the OleTP$_{RN}$ E405F with 10 mM H_2O_2.

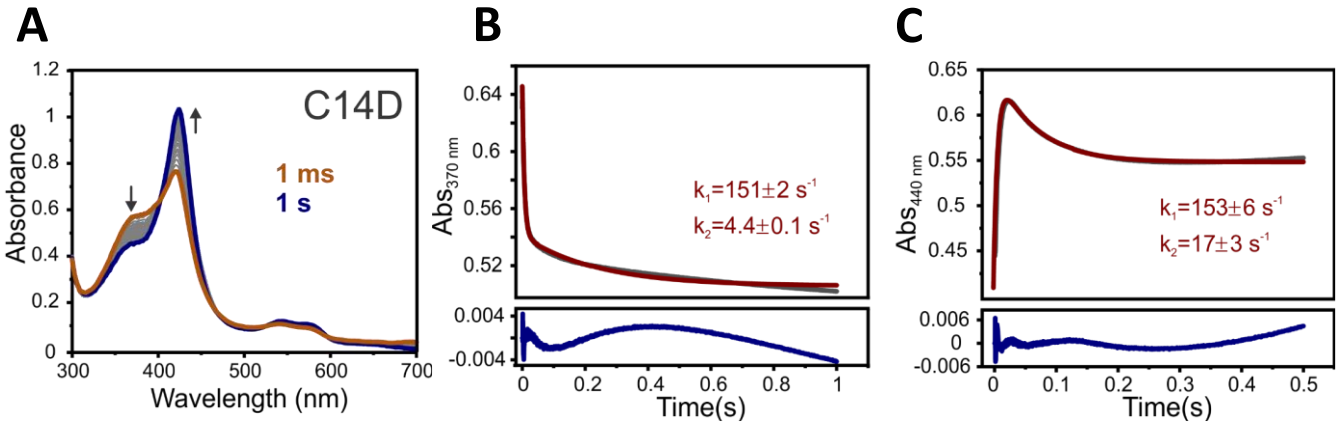


Figure S28: (A) PDA spectra of the single turnover reaction between 10 µM C14D-OleTP$_{RN}$ E405F and 10 mM H_2O_2 monitored for 1 s at 4°C. Arrows indicate the direction of decreasing and increasing absorbances during the reactions. Two exponential fits of 370 nm **(B)** (fitted up to 1 s) and 440 nm **(C)** (fitted up to 500 ms) PMT traces collected during the single turnover reaction of C14D bound OleTP$_{RN}$ E405F mutant with hydrogen peroxide.

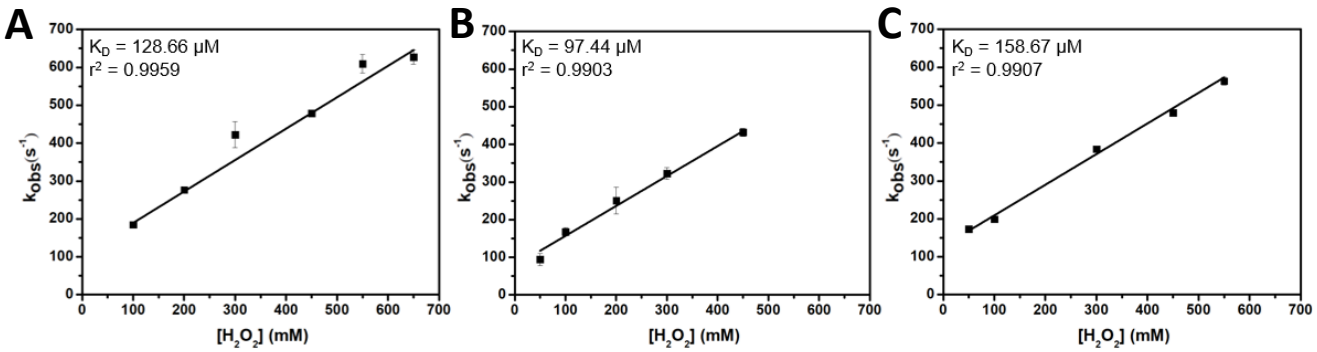


Figure S29: Dependence of RRT on H_2O_2 concentration for $OleTP_{RN}$ wild type **(A)**, monomeric mutation E405F **(B)**, and S26F **(C)**. Black lines represent the linear fitting. The forward and reverse rate constants for the peroxide binding step are provided by the slope and intercept of the plot, respectively.

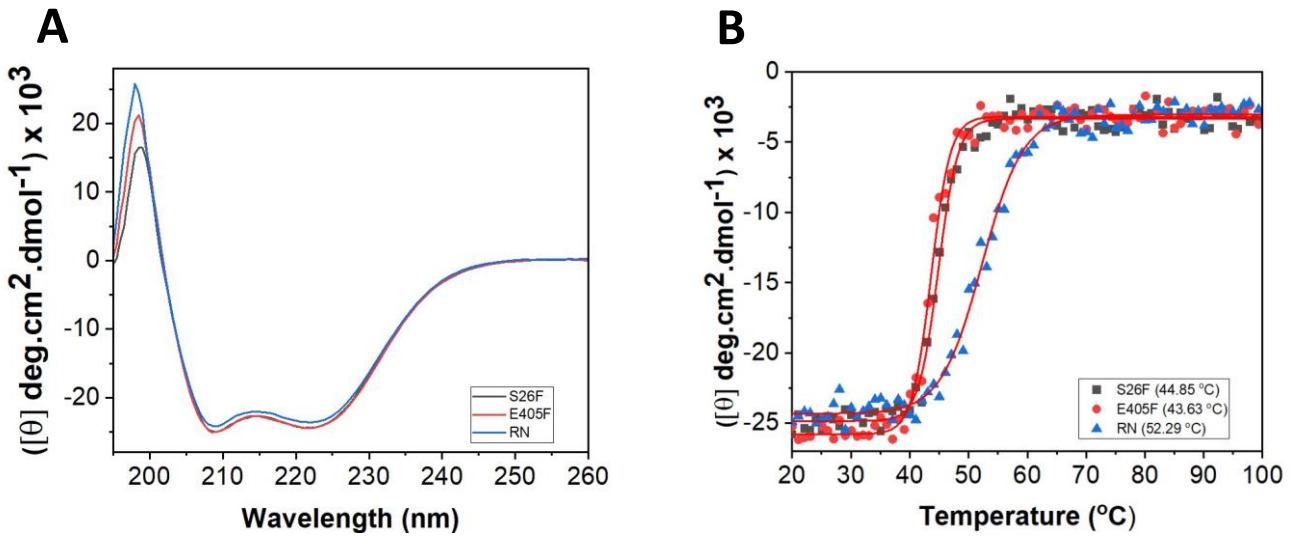


Figure S30: Circular dichroism (CD) analysis of OleTP$_{RN}$ and monomeric mutants. **(A)** Ellipticity over a wavelength range of 190–260 nm at 20°C. **(B)** Thermal unfolding profile following ellipticity at 222 nm.

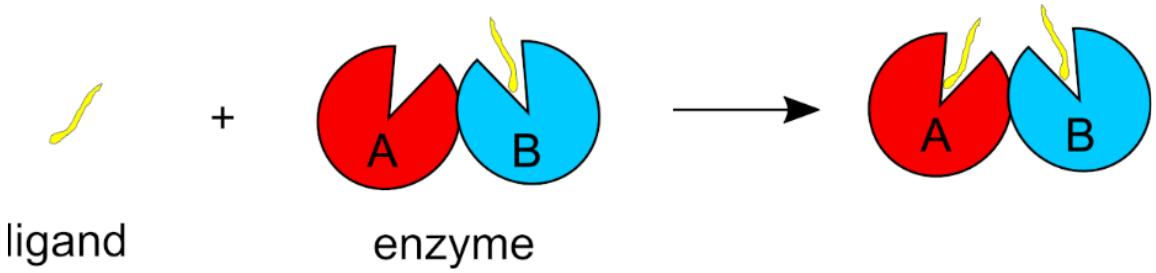


Figure S31: Binding process used to compute binding free energies with MM/PBSA.

SUPPLEMENTARY REFERENCES

1. T. Aramaki, *et al.*, KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).

2. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

3. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).

4. L. T. Nguyen, H. A. Schmidt, A. von Haeseler, B. Q. Minh, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

5. J. A. Amaya, C. D. Rutland, N. Leschinsky, T. M. Makris, A Distal Loop Controls Product Release and Chemo- and Regioselectivity in Cytochrome P450 Decarboxylases. *Biochemistry* **57**, 344–353 (2018).

6. D. G. Gibson, Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol.* **498**, 349–361 (2011).

7. J. F. Morrison, Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. *Biochim. Biophys. Acta* **185**, 269–286 (1969).

8. J. L. Grant, M. E. Mitchell, T. M. Makris, Catalytic strategy for carbon-carbon bond scission by the cytochrome p450 oleT. *Proc. Natl. Acad. Sci. U S A* **113**, 10049–10054 (2016).

9. J. M. DeLong, *et al.*, Using a modified ferrous oxidation-xylenol orange (FOX) assay for detection of lipid hydroperoxides in plant tissue. *J. Agric. Food Chem.* **50**, 248–254 (2002).

10. A. P. Hammersley, FIT2D: a multi-purpose data reduction, analysis and visualization program. *J. Appl. Cryst.* **49**, 646–652 (2016).

11. D. I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* **25**, 495–503 (1992).

12. D. Franke, D. I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **42**, 342–346 (2009).

13. V. v. Volkov, D. I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Cryst.* **36**, 860–864 (2003).

14. M. B. Kozin, D. I. Svergun, Automated matching of high- and low-resolution structural models. *J. Appl. Cryst.* **34**, 33–41 (2001).

15. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

16. A. Vagin, A. Teplyakov, MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).

17. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

18. P. v. Afonine, *et al.*, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).

19. G. N. Murshudov, *et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).

20. V. B. Chen, *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

21. R. P. Joosten, F. Long, G. N. Murshudov, A. Perrakis, The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* **1**, 213–220 (2014).

22. R. Salomon-Ferrer, D. A. Case, R.C. Walker. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **3**, 198-210 (2013).

23. J. Myers, G. Grothaus, S. Narayanan, A. Onufriev, A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* **63**, 928–938 (2006).

24. J. C. Gordon, *et al.*, H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **33,** W368–371 (2005).

25. J. Wang, W. Wang, P. A. Kollman, D. A. Case, Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).

26. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

27. C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *J. Phys. Chem.* **97**, 10269–10280 (1993).

28. J. E. Carpenter and F. Weinhold, Analysis of the geometry of the hydroxymethyl radical by the different hybrids for different spins natural bond orbital procedure, *J. Mol. Struct. (Theochem),* **139**, 41-62 (1988).

29. K. Shahrokh, A. Orendt, G. S. Yost, T. E. Cheatham, Quantum mechanically derived AMBER-compatible heme parameters for various states of the cytochrome P450 catalytic cycle. *J. Comput. Chem.* **33**, 119–133 (2012).

30. J. A. Maier, *et al.*, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

31. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1998).

32. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089 (1998).

33. D. R. Roe, T. E. Cheatham, PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

34. B. R. Miller, *et al.*, MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).

35. J. Kästner, Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 932–942 (2011).

36. Matthews *et al.*, Catalytic determinants of alkene production by the cytochrome P450 peroxygenase $OleT_{JE}$. *J. Biol. Chem.* **292**, 5128–5143 (2017)