Supporting Information

Substrate tunnel engineering aided by X-ray crystallography and functional dynamics swaps the function of MIO-enzymes

Zsófia Bata[1,2], Zsófia Molnár[1,2], Erzsébet Madaras[1,2], Bence Molnár[1,2], Evelin Sánta-Bell[1], Andrea Varga[3], Ibolya Leveles[2,4], Renzhe Qian[5], Friedrich Hammerschmidt[5], Csaba Paizs*[3], Beáta G. Vértessy*[2,4], László Poppe*[1,3]

[1] Department of Organic Chemistry and Technology, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary; [2] Institute of Enzymology, ELKH Research Center of Natural Sciences, Magyar tudósok krt. 2, H-1117 Budapest, Hungary; [3] Biocatalysis and Biotransformation Research Centre, Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University of Cluj-Napoca, Arany János Str. 11, RO-400028 Cluj-Napoca, Romania; [4] Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111 Budapest, Hungary; [5] Institute of Organic Chemistry, University of Vienna, Währinger Str. 38, A-1090 Vienna, Austria

*Correspondence and requests for materials should be addressed to C. P. (e-mail: csaba.paizs@ubbcluj.ro) B.G.V. (e-mail: vertessy.beata@vbk.bme.hu), and to L.P. (e-mail: poppe.laszlo@vbk.bme.hu)

Table of Contents

SUPPORTING METHODS

Cloning of MIO-enzymes. To allow removal of the His-tag, KkPAL, PcPAL and TcPAM were cloned to the pET15bTEV vector[1] using *NheI*/*SacI* sites in order to produce an *N*-terminal His_6-tagged protein, where the His_6-tag is cleavable by a tobacco etch virus protease (TEV). Primer and protein sequences are listed in Table S5 and S6.

Site directed mutagenesis of MIO-enzymes. All single and double mutant MIO-enzymes were created by site-directed mutagenesis following the protocol described by Naismith and Liu[2], using suitable mutagenic primers (Table S5) from the corresponding wild-type protein.

Preparation of the MIO-enzyme constructs. All proteins were produced and purified using the same procedure as described for PcPAL-His_{10}[3]. Briefly, the enzyme was produced in *E. coli* Rosetta BL21 (DE3) cells in Luria-Bertani (LB) media induced by IPTG at 28 °C for 6 h. After harvesting and cell lysis by sonication, the enzyme was separated from the rest of the supernatant by gravitational Ni-NTA chromatography. Composition of all fractions were verified by SDS-PAGE. Imidazole was removed by 16 h dialysis against 10-fold excess volume of TRIS buffer (pH 8.0, 50 mM). His-tagged TEV protease (1:10 molar ratio) was added to the His_6-tagged protein preparations during dialysis to cleave the His_6-tag. After dialysis, the non-cleaved His_6-tagged protein, His-tagged TEV and the cleaved His_6-tag was separated from the target enzyme by gravitational Ni-NTA chromatography. Enzyme preparations were stored at -80 °C until activity measurements.

For crystallizations, PcPAL-His_{10} and the tag-free PcPAL were concentrated to approximately 10 mg ml^{-1} with Amicon Ultra-15 Centrifugal Filter units. Stocks of 50-100 µl of protein solutions were frozen in liquid N_2 and stored at -80 °C until crystallization. After thawing, samples were ultra-centrifuged at 230 000 *g* for 20 min.

Phenylalanine aminomutase activity assay. The ammonia addition and ammonia elimination reactions catalyzed by native and mutant PALs and PAMs were analyzed by an Agilent 1100 HPLC using reverse phase HPLC column. Samples were taken at fixed intervals from the reaction mixtures. Samples from the enzymatic reaction mixture (50 µL) were stopped by heating at 95 °C for 10 min and centrifuged (13000 rpm, 2 min). The supernatant was transferred through a 0.22 µm filter, after filtration the samples were used directly for HPLC analysis.

Ammonia elimination reactions were carried out at 30 °C in 100 mM TRIS buffer at pH 8.5 with 10 µM enzyme concentration. The initial substrate concentrations in the reaction mixtures were 5 mM for α-L-Phe or 10 mM for racemic β-Phe.

Ammonia addition reactions were carried out at 30 °C in 3M $(NH_4)_2CO_3$ with 10 µM enzyme concentration and contained initially 5 mM cinnamic acid.

The molar fractions of the products were determined by an Agilent 1100 HPLC instrument, using *Gemini* NX-C-18 column (150 mm × 4.6 mm × 5 µm; *Phenomenex*, Torrance, CA, USA). Mobile phase: NH_4OH buffer (0.1 M, pH 9.0) and acetonitrile (92:8). Flow rate: 0.9 mL min^{-1}. Detection wavelengths and the retention times are given in Table S7. The molar fractions of the products were calculated from integrated peak areas using the appropriate response factor.

In a repeated series of experiments, a contaminant overlapping with the α-Phe peak on the *Gemini* NX-C-18 column was detected in the ammonia elimination reactions after three days reaction time (most likely due to enzyme degradation). The contamination overlapping with the α-Phe peak had strong absorption at 280 nm, where neither the α-Phe nor the β-Phe have significant absorption. Based on this behavior, the real amount of α-Phe could be easily corrected.

Separation of the enantiomers of α-Phe and β-Phe was performed by an Agilent 1100 HPLC instrument using Chiralpak ZWIX (+) column (250 × 4 mm × 3 µm, *Daicel Corporation*, Tokyo, Japan) and a mixture of MeOH, acetonitrile and H_2O as eluent supplemented with diethylamine (25 mM) and formic acid (50 mM), at a flow rate of 0.7 mL min^{-1}. Detection wavelengths and the retention times are given in Table S8.

Model preparation for molecular modeling. All MIO-enzymes exist and function as homotetramers, where three monomers participate in forming each one of the four active sites. Therefore, it is imperative to model them in their biological functional homotetrameric form. This results in very large systems that increase the simulation time, but it also allows to run four simultaneous experiments. Table S9 shows the setup of the different active sites of the four models.

The 6F6T crystal structure served as the basis for modelling of PcPAL. The tetrameric model was constructed using crystal symmetry data, as the asymmetric unit contains only two chains. Missing loop regions (A: 337-346, 549-557; B: 333-346, 551-554) were modelled based on the homology model by Phyre2[4]. Protonation state of residues were taken at pH 8.8. Hydrogen ions and protonation states were determined by Maestro[5] and PROPKA[6] , for pH 8.8 as it is the pH for enzyme kinetic experiments[7].

To obtain the most reliable initial positions for the substrates L-phenylalanine (Phe) and *trans*-cinnamate (CA) different methods were used to create the substrate bound protein models. Docking of the zwitterion form of Phe with Glide[8] determined the initial positions in active sites A and B, as there is no direct crystallographic data. Same docking settings were used as described previously[7]. From the docking results the initial Phe arrangement was selected by visual comparison to the crystal structure bound (*S*)-APPA in 6F6F and considering the docking score. Binding conformation of CA is known from a crystal structure of TcPAM, 4CQ5[9], therefore these coordinates were used for the initial CA position. Experimental results with HAL[10] suggest that in PAL the amino group remains also bound to the MIO, while CA exits from the active site, at least at the beginning of the egress. Therefore, MIO with bound amino group was modeled with the CA substrate. The amino group position was modeled based on the *N*-MIO intermediate observed in the 6F6T structure.

Substrate bound models of TcPAM were built similarly to those of PcPAL, based on the crystal structure 2YII[11], that was also used for previous molecular dynamics studies.[12] The coordinates for the missing regions (A: 114-119, 567-569, 606-616, 677-686; B: 55-56, 114-119, 569-571, 606-616, 678-686; C: 115-119, 567-571, 605-616, 677-686; D: 115-119, 568-576, 605-616, 677-686) were also modeled based on a Phyre2[4] model. All other steps were identical to those described above for PcPAL.

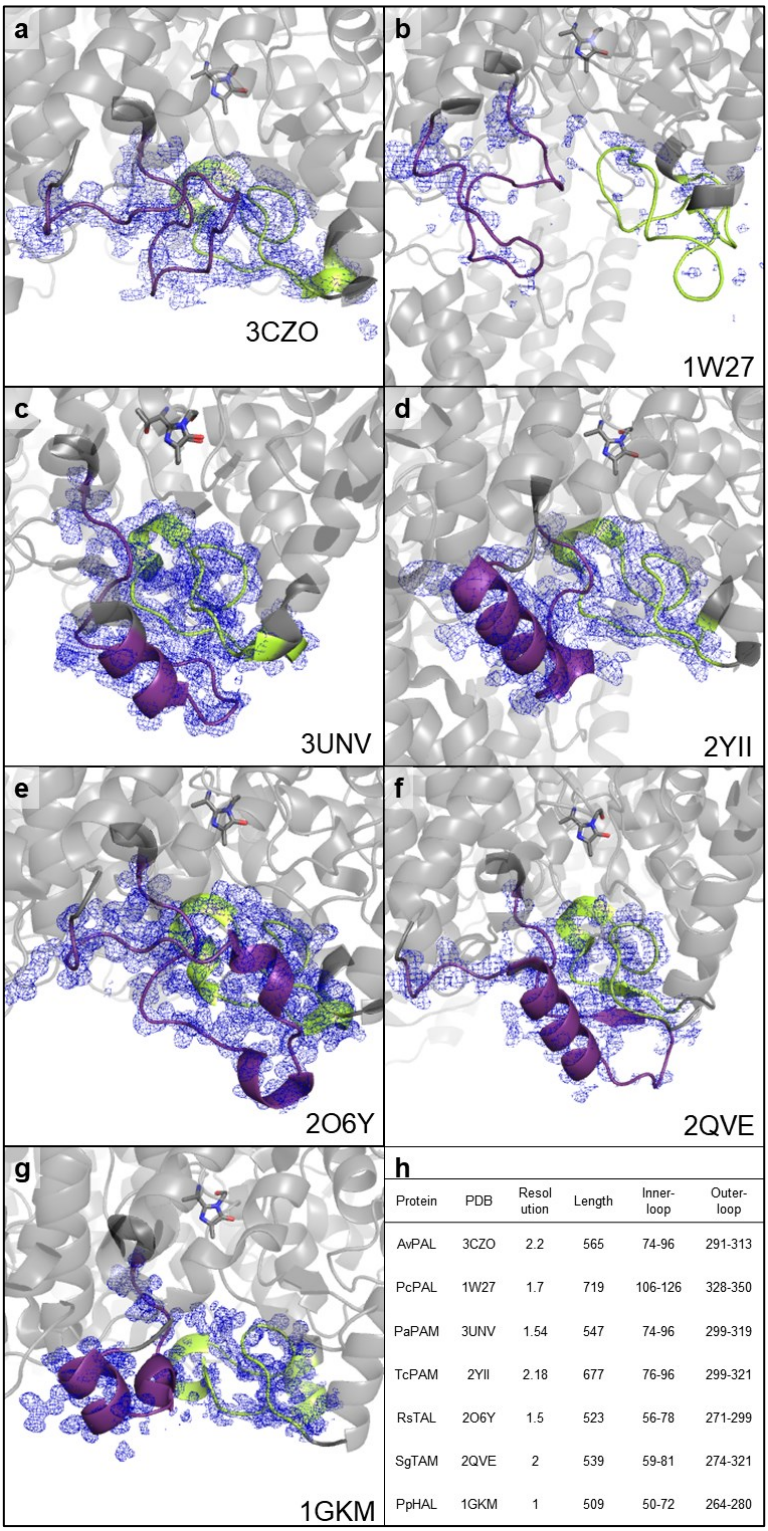
Parameters were determined using Antechamber[13] for residues not included in the standard ff14SB forcefield; MIO, amino-MIO, Phe, CA and deprotonated tyrosine. The initial values for the parameters were taken from crystal structures or from GAFF[14].

Amber ff14SB[15] force field was used to describe parameters for the proteins. Proteins were solvated in a truncated octahedral box with TIP3P[16] water molecules with 10 Å distances in all directions, to minimize the size of the systems. Na^+ ions neutralized the systems, and 0.1 M NaCl modeled the ionic strength of the experimental conditions.

Molecular dynamics simulations. All MD simulations were run with NAMD 2.11[17], as it is the only MD engine capable of running RAMD simulations. First 5000 step minimization removed clashes from the water box, with all protein and ligand atoms constrained using 10 kcal mol^{-1} $Å^{-1}$ force constant. Second the whole system was minimized for 10 000 steps. Third, 100 ps NVT simulations relaxed the water box (T=303 K, Langevin dynamics[18]), with all protein and ligand atoms constrained using 10 kcal mol^{-1} $Å^{-1}$ force constant. Finally, 20 ns-long NPT simulations (p=1 atm, T=300 K) probed the equilibration behaviors and relaxed the models before the RAMD simulations. All bonds involving

hydrogen atoms were constrained during the simulations to permit increasing the time step to 2 fs. Electrostatic interactions were cutoff at 12.0 Å, long-range interactions were treated with PME method[19]. Constant pressure and temperature were achieved with the Noosé-Hoover Langevin piston method[18,20]. Convergence of backbone root mean-square deviation (RMSD) showed that systems reached equilibrium after approximately 2 ns (Figure S14.). Comparison of the crystallographic B factors and the root mean-square flexibility (RMSF) observed during the simulations (Figure S15.) showed that the models and the crystal structures behaved similarly; the same regions of the proteins showed high flexibilities. Comparable patterns from the crystal structures and the molecular dynamics indicate that the simulations reproduced well the experimentally observed protein flexibility, and therefore the models are likely adequate.

SUPPORTING FIGURES



h

Protein	PDB	Resolution	Length	Inner-loop	Outer-loop
AvPAL	3CZO	2.2	565	74-96	291-313
PcPAL	1W27	1.7	719	106-126	328-350
PaPAM	3UNV	1.54	547	74-96	299-319
TcPAM	2YII	2.18	677	76-96	299-321
RsTAL	2O6Y	1.5	523	56-78	271-299
SgTAM	2QVE	2	539	59-81	274-321
PpHAL	1GKM	1	509	50-72	264-280

Figure S1. | Electron densities of the inner (green) and outer loop (purple) regions of representative structures of the different classes of MIO-enzymes**.** Electron densities (2fo-fc map) are shown by blue mesh at 1 sigma level. **a-g** Protein structures overlaid on the electron density data available from the PDB. **h** Summary of selected structures.

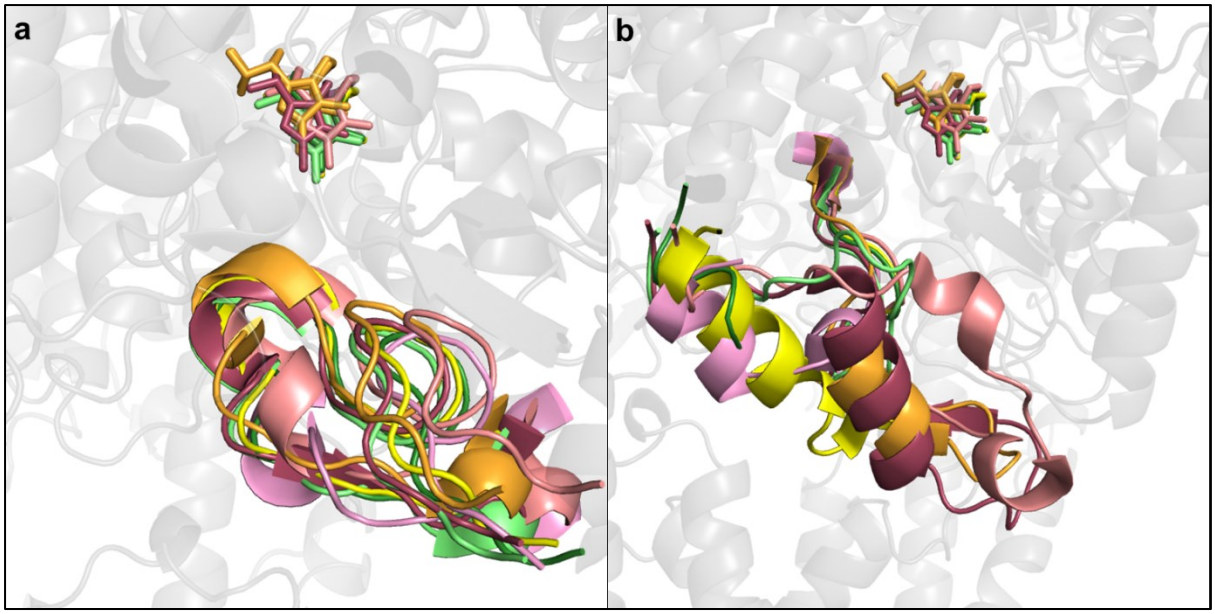


Figure S2. | Structural alignment of active site capping loop regions in selected MIO-containing class I lyase like enzymes. Structure PDB IDs and colors: AvPAL (3CZO, lime green); PcPAL (6F6T, green); PaPAM (3UNV, orange); TcPAM (2YII, yellow); RsTAL (2O6Y, salmon); SgTAM (2QVE, raspberry); PpHAL (1GKM, pink). **a** Inner loop conformations are highly similar (the catalytically component loop-in conformation), the most dissimilar is the loop of PpHAL (pink). **b** In contrast, outer loop conformations show high variability amongst the structures, both in sequence length and in conformation.

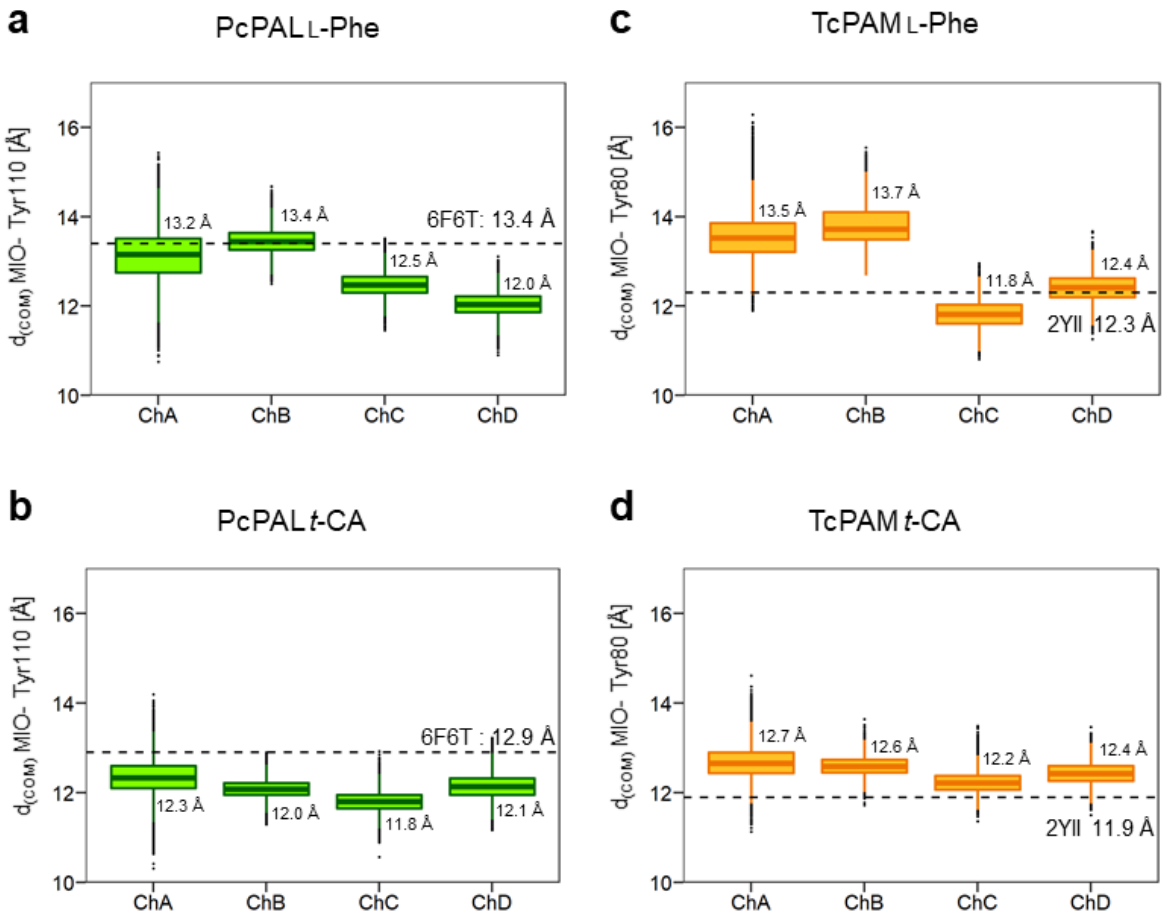


Figure S3. | Box plots showing Tyr110(80)-MIO center of mass (COM) distances in the different chains of the tetrameric models. **a** PcPAL model with L-phenylalanine (L-Phe) substrate. **b** PcPAL model with *trans*-cinnamate (*t*-CA) product, **c** TcPAM model with L-Phe substrate. **d** TcPAM model with *t*-CA product. Dashed lines show the reference values measured for chain A in the starting structures. Labels show the median distances in the different chains. Egress of *t*-CA was modeled from an NH_2-MIO containing active site, that reduces the COM distance between the catalytic residue and Tyr110(80) in these models. Table S9. provides detailed explanation of the configuration of the different sites.

Active site D was modeled equally for L-Phe and *t*-CA ligands, thus served as a control. In both PcPAL models (average values 12.0 Å and 12.1 Å) and in both TcPAM models (average values 12.4 Å in both models) active site D led to the same results, validating modeling and simulation protocols.

Active site B with bound ligand showed increased average Tyr-MIO COM distances compared to the unoccupied C active sites. This increase was larger for L-Phe (1.0 Å for PcPAL and 2.0 Å for TcPAM) then for *t*-CA (0.3 Å for PcPAL and 0.4 Å for TcPAM). The median values are smaller for both PcPAL models then the COM distance in the crystal structure (13.4 Å and 12.9 Å), in contrast median values are larger in almost all TcPAM active sites then the COM distance in the crystal structure (12.3 Å and 11.9 Å). These two observations indicate that the inner loop of PcPAL in 6F6T is slightly open due to the presence of the phosphonic acid inhibitor, while the inner loop is in closed conformation in 2YII as only a small β-mercaptoethanol molecule is bound here.

An increase in the width of the distribution can be observed in active site A in all four models, due to the application of the accelerating force. The difference in the median values compared to the ligand bound B active sites is most likely also due to the accelerating force.

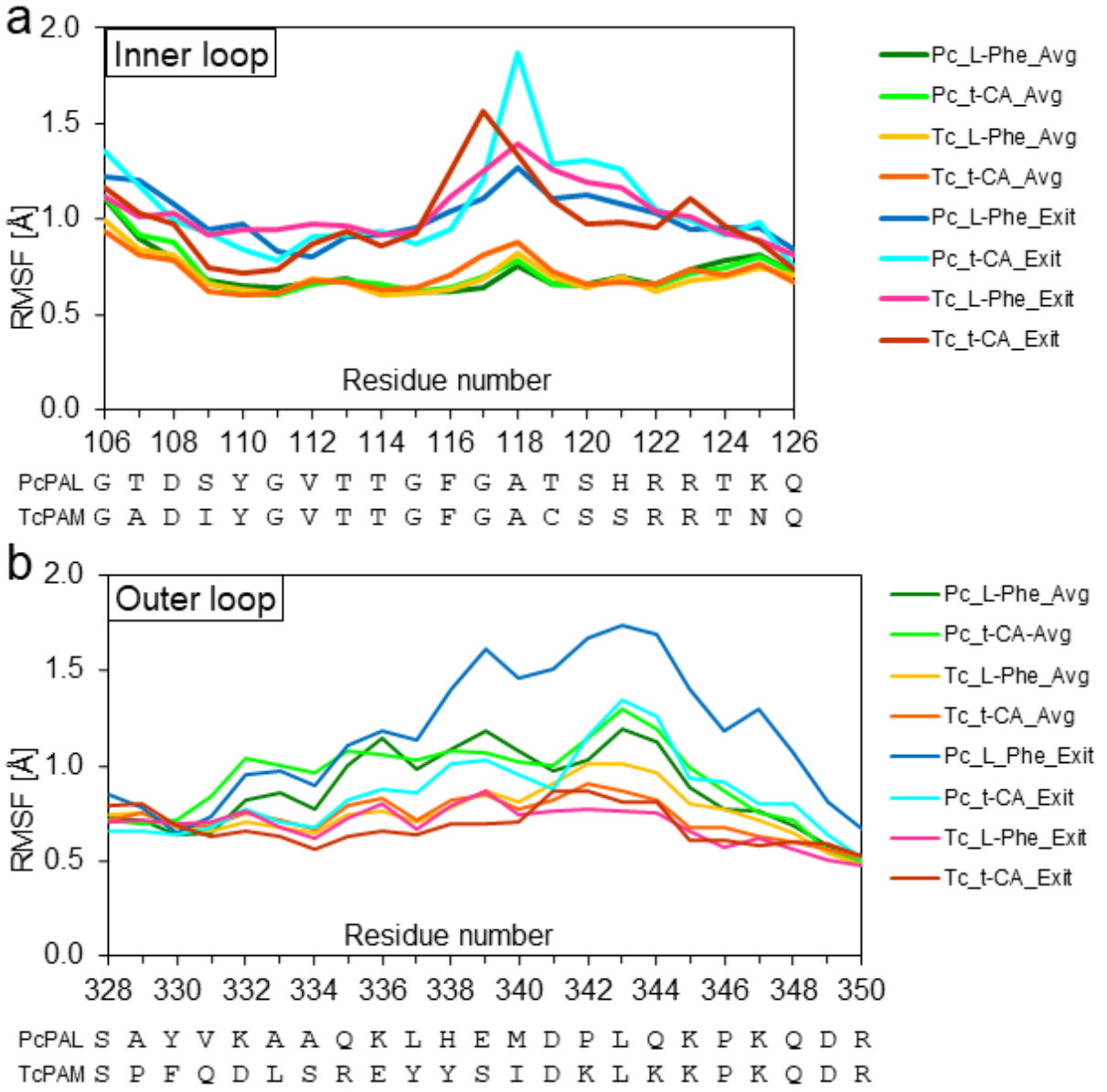


Figure S4. | Root-mean-square fluctuation (RMSF) of backbone atoms of the active site capping loops in PcPAL and TcPAM models. Data for inner loop (**a**) and of the outer loop (**b**) in the four different simulations (PcPAL with L-Phe and *t*-CA, TcPAM with L-Phe and *t*-CA). RMSF value shows the average root-mean-square deviation (RMSD) observed during the trajectory indicating the protein flexibility in a molecular dynamics simulation. Plots show average RMSF values for inner and outer loop residues averaged for all four chains during the RAMD simulations. Separate RMSF data sets were calculated for loops covering the RAMD active site for frames after ligand exit occurred (COM MIO-Lig>10Å).

Average inner loop movements (**a**) were identical in PcPAL and TcPAM (dark and bright green, yellow and orange lines in panel (**a**). Notable conformational differences of residues 116-122 being in direct contact with the outer loop occurred during ligand exit phase (blue, cyan, pink and red lines in panel (**a**).

Average outer loop movements (**b**) were more intense in PcPAL (dark and bright green lines in panel **b**) then in TcPAM (yellow and orange lines in panel (**b**). The largest conformational change occurred upon exit of L-Phe from PcPAL (blue line in panel **b**), while only minor differences occurred in all other cases (cyan, pink, and red lines in panel **b**).

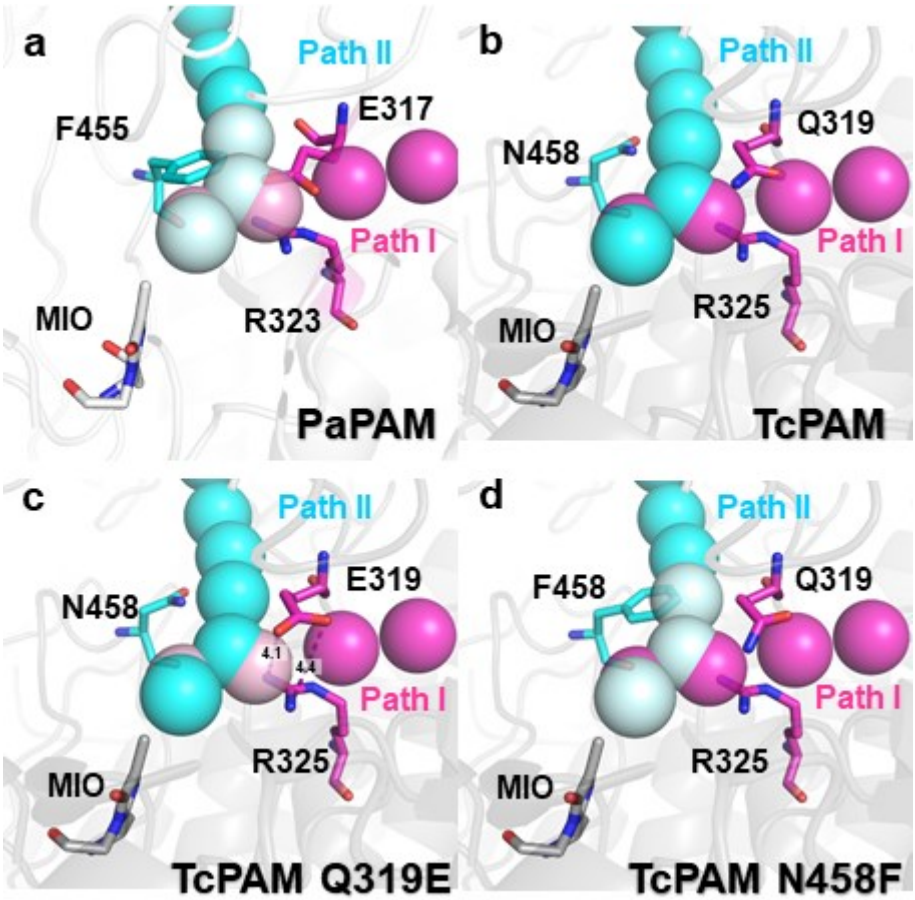


Figure S5. | Predicted effect of mutations on the accessibility of the active site in MIO-enzymes on the example of TcPAM. **a.** Both main ligand egress paths are blocked in the (*S*)-PAMs such as PaPAM by Glu317 and Phe455. **b.** The two most frequent egress paths in wt TcPAM. Paths are visualized by center of masses of the exiting substrates during the simulations, Path I is in magenta, Path II is in cyan. **c.** In mutant Q319E, Path I is blocked by the salt-bridge between the Glu319 and Arg325. **d.** In mutant N458F the bulky Phe residue blocks Path II.

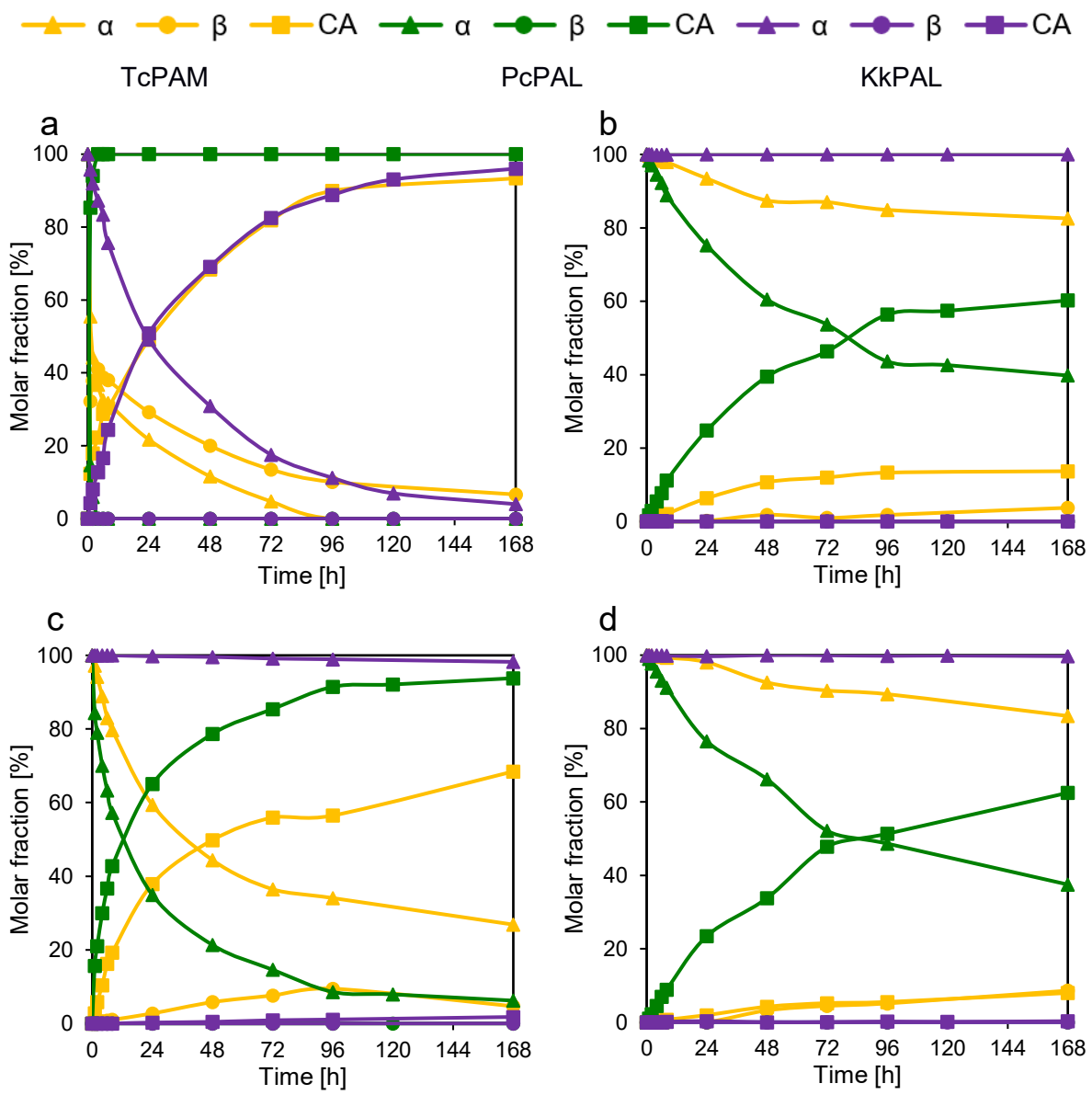


Figure S6. | Progress curves for biotransformations starting from L-α-Phe with wild type and mutant MIO-enzymes (TcPAM, PcPAL and KkPAL). During the reactions of the different enzymes (TcPAM - yellow, PcPAL - green, KkPAL - purple) the molar fractions of α-Phe (▲), β-Phe (●) and CA (■) were monitored. **a.** Wild-type enzymes. **b.** Mutation of outer loop residue Gln to Glu (TcPAM Q319E, PcPAL Q348E, KkPAL Q280E). **c.** Mutation of Asn to Phe (TcPAM N458F, PcPAL N487F, KkPAL N421F). **d.** Double mutants (TcPAM Q319E/N458F, PcPAL Q348E/N458F, KkPAL Q280E/N421F).

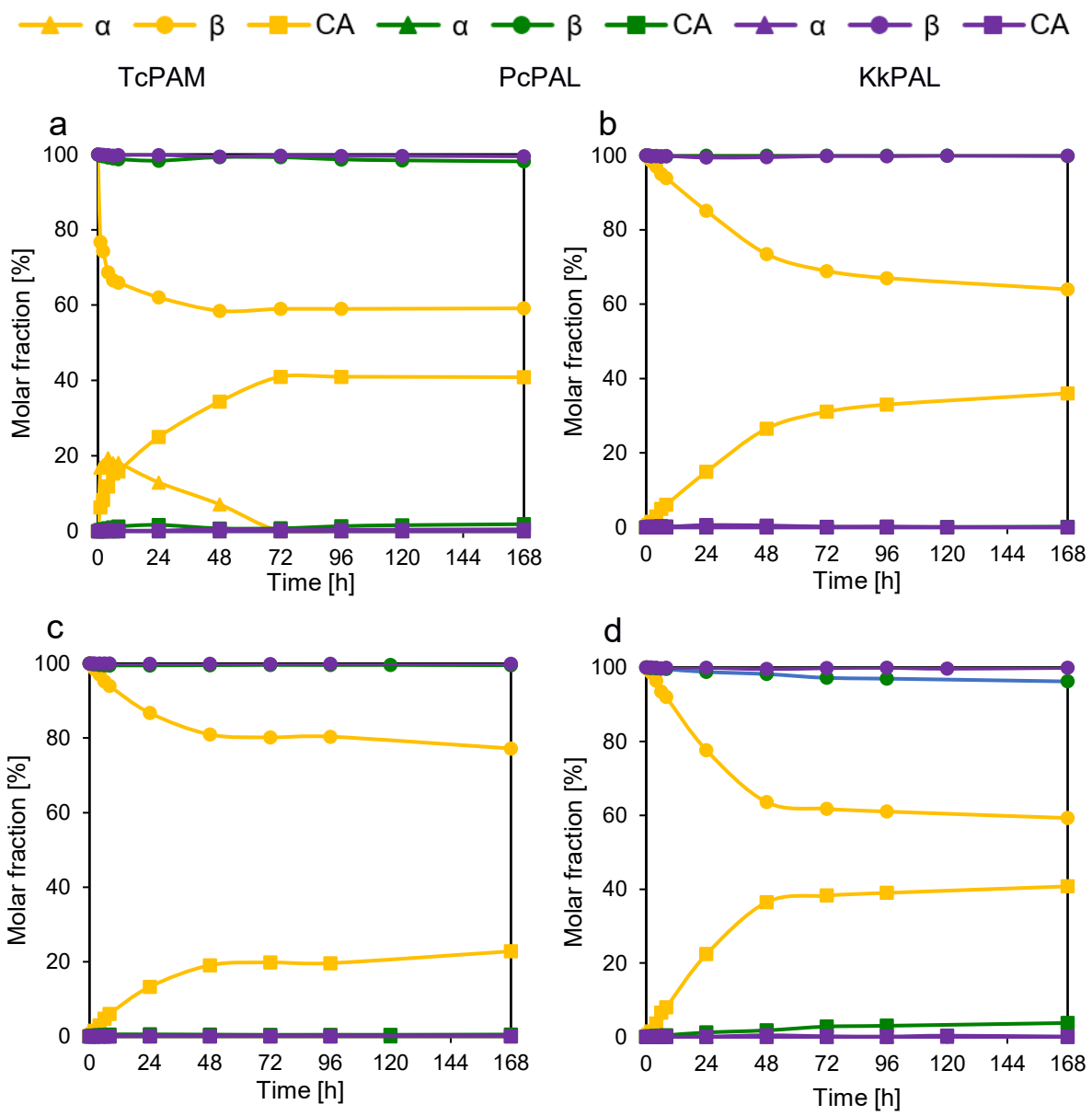


Figure S7. | Progress curves for biotransformations starting from racemic β-Phe with wild type and mutant MIO-enzymes (TcPAM, PcPAL and KkPAL). During the reactions of the different enzymes (TcPAM - yellow, PcPAL - green, KkPAL - purple) the molar fractions of α-Phe (▲), β-Phe (●) and CA (■) were monitored. **a.** Wild-type enzymes. **b.** Mutation of outer loop residue Gln to Glu (TcPAM Q319E, PcPAL Q348E, KkPAL Q280E). **c.** Mutation of Asn to Phe (TcPAM N458F, PcPAL N487F, KkPAL N421F). **d.** Double mutants (TcPAM Q319E/N458F, PcPAL Q348E/N458F, KkPAL Q280E/N421F).

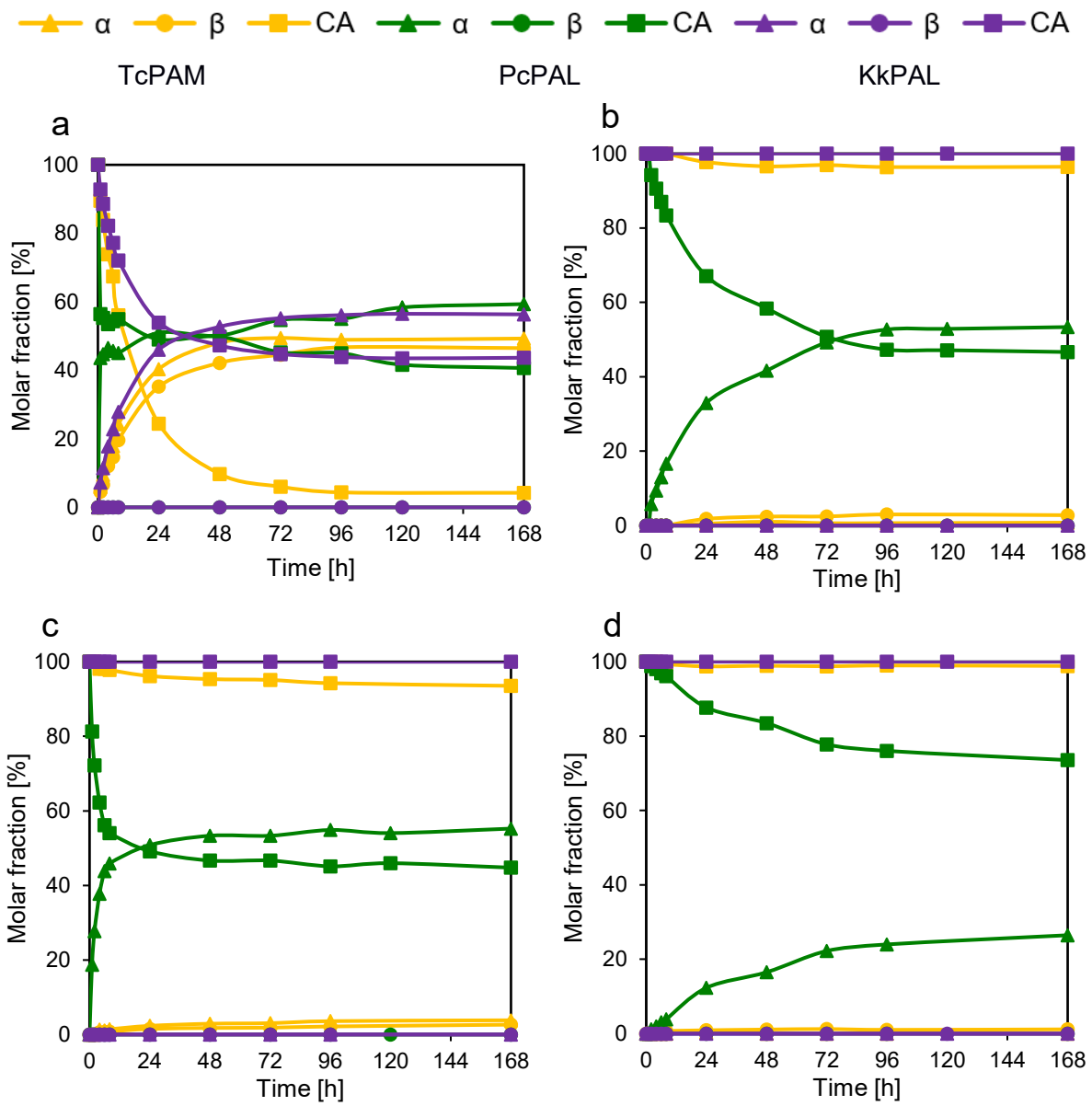


Figure S8. | Progress curves for biotransformations starting from cinnamic acid (CA) with wild type and mutant MIO-enzymes (TcPAM, PcPAL and KkPAL). During the reactions of the different enzymes (TcPAM - yellow, PcPAL - green, KkPAL - purple) the molar fractions of α-Phe (▲), β-Phe (●) and CA (■) were monitored. **a.** Wild-type enzymes. **b.** Mutation of outer loop residue Gln to Glu (TcPAM Q319E, PcPAL Q348E, KkPAL Q280E). **c.** Mutation of Asn to Phe (TcPAM N458F, PcPAL N487F, KkPAL N421F). **d.** Double mutants (TcPAM Q319E/N458F, PcPAL Q348E/N458F, KkPAL Q280E/N421F).

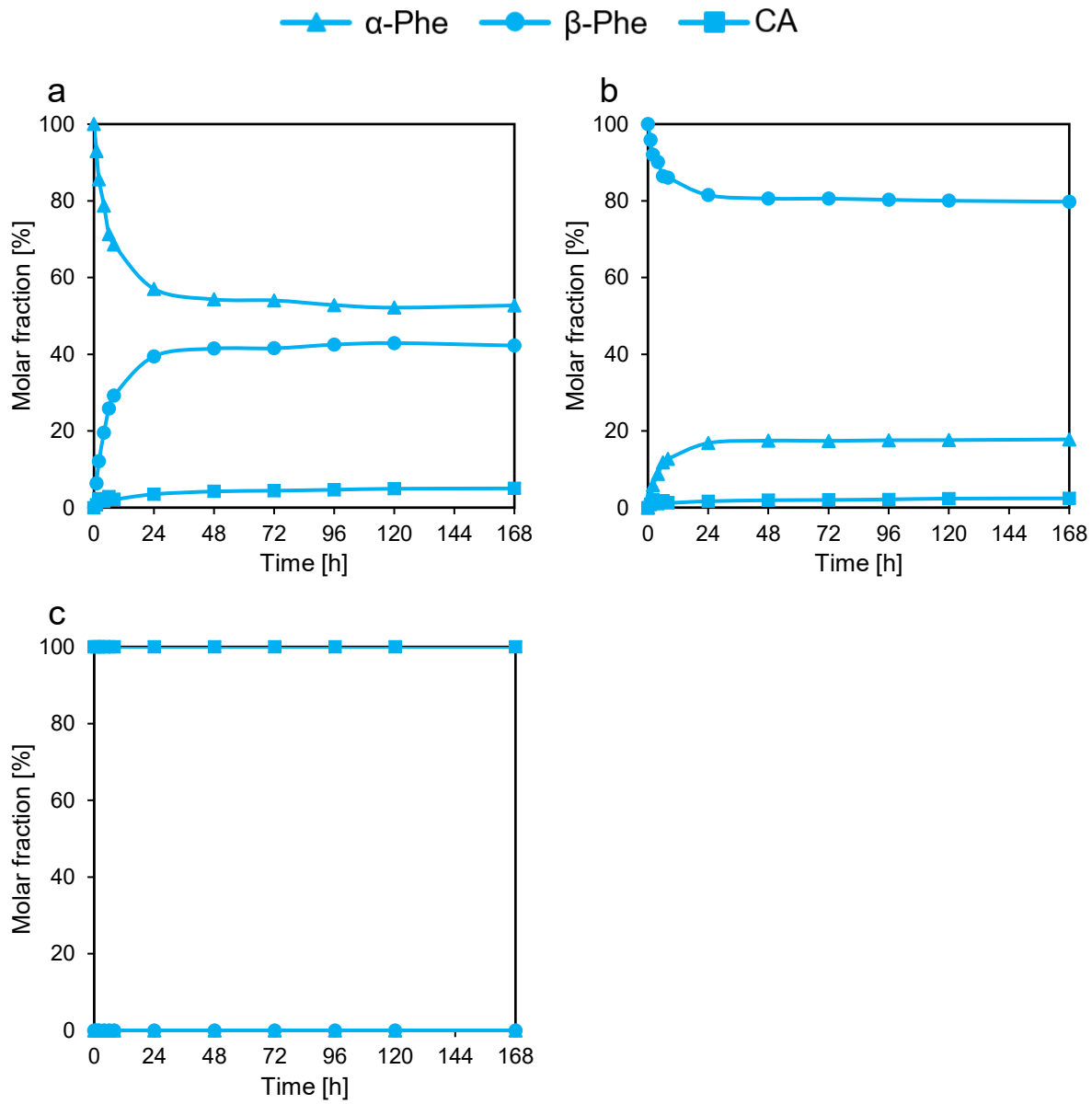


Figure S9. | Progress curves for biotransformations by PfPAM starting from (**a**) L-α-Phe, (**b**) β-Phe and (**c**) CA.

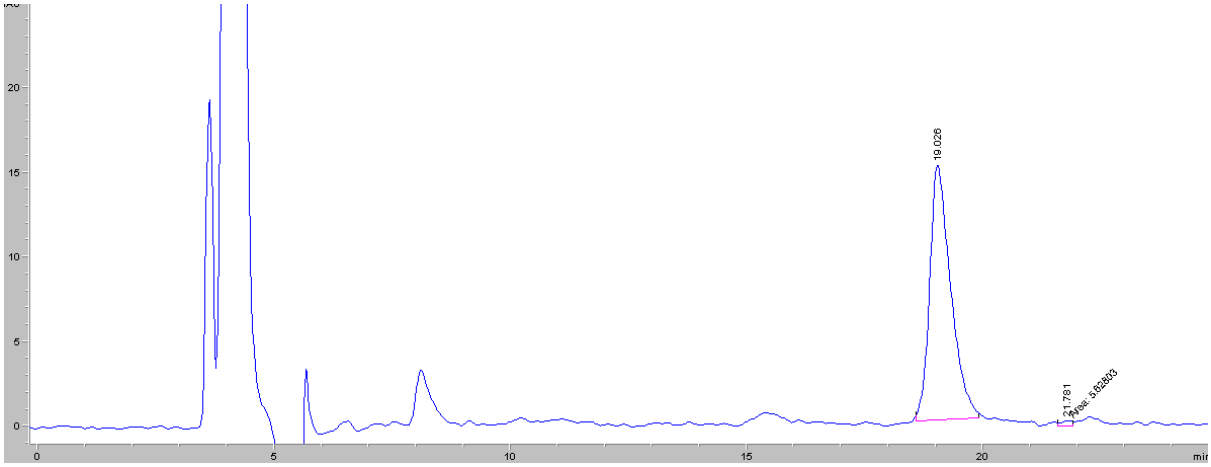


Figure S10. | HPLC analysis of the biotransformation from β-Phe after 168 h with wt TcPAM.

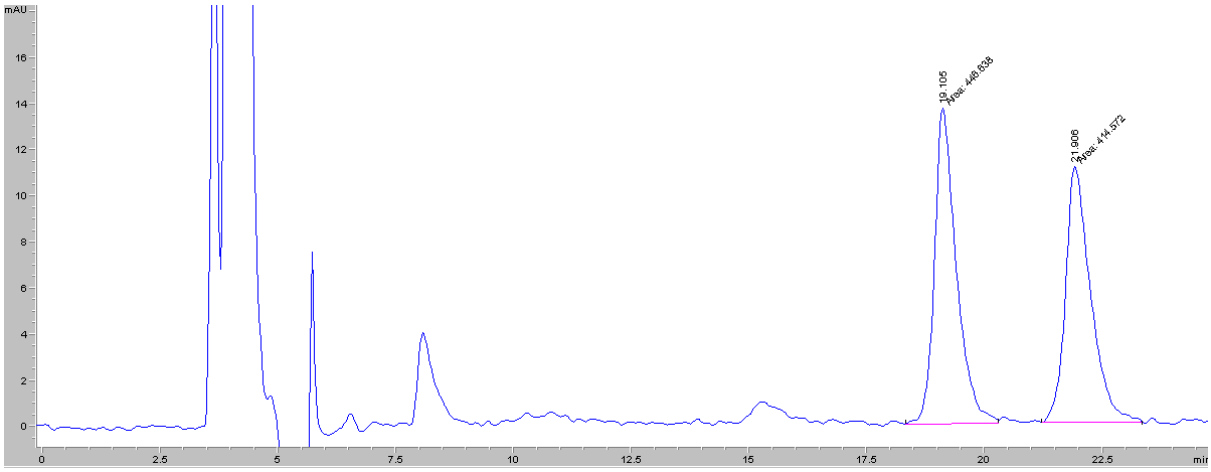


Figure S11. | HPLC analysis of the biotransformation from β-Phe after 168 h with TcPAM Q319E.

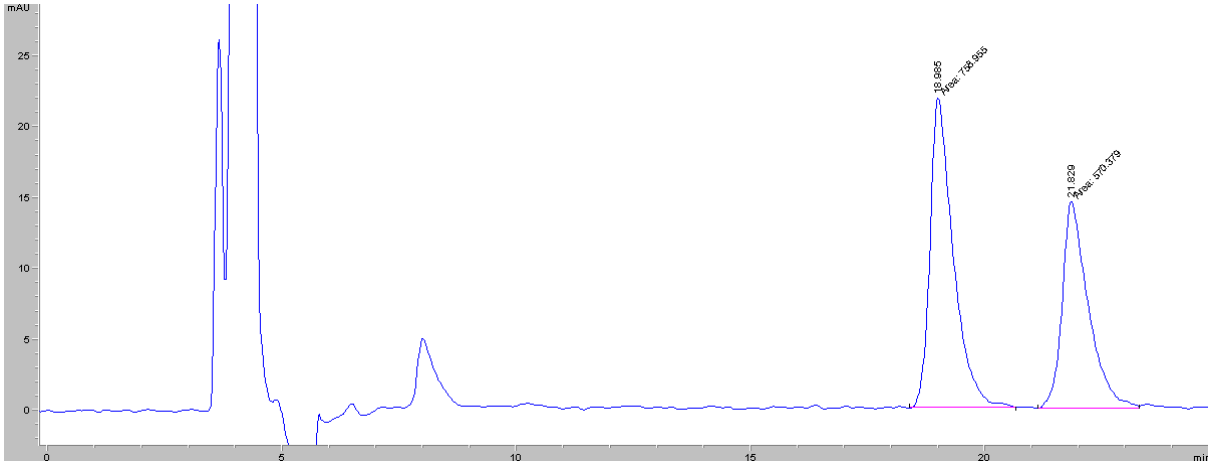


Figure S12. | HPLC analysis of the biotransformation from β-Phe after 168 h with TcPAM N458F.

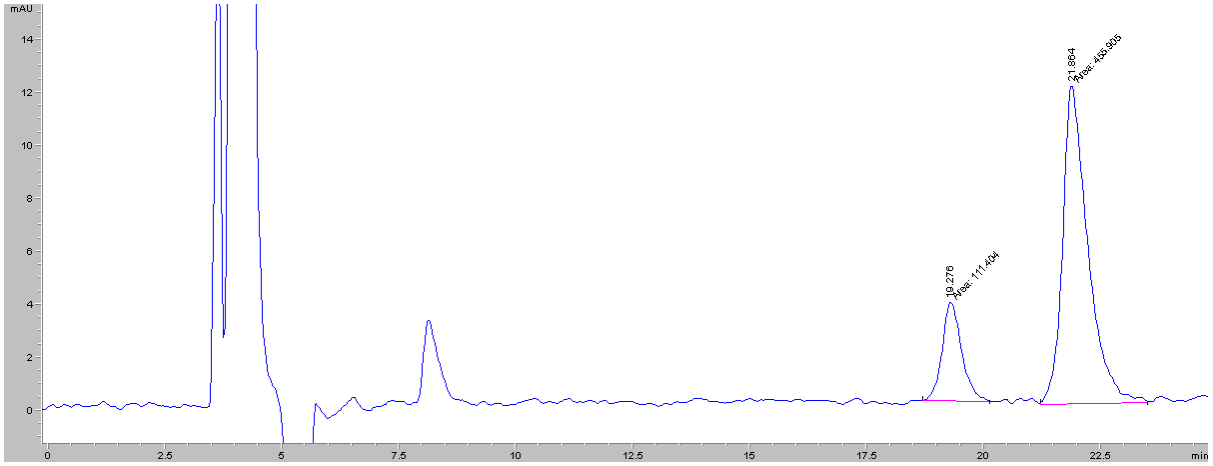


Figure S13. | HPLC analysis of the biotransformation from β-Phe after 168 h with TcPAM Q319E/N458F.

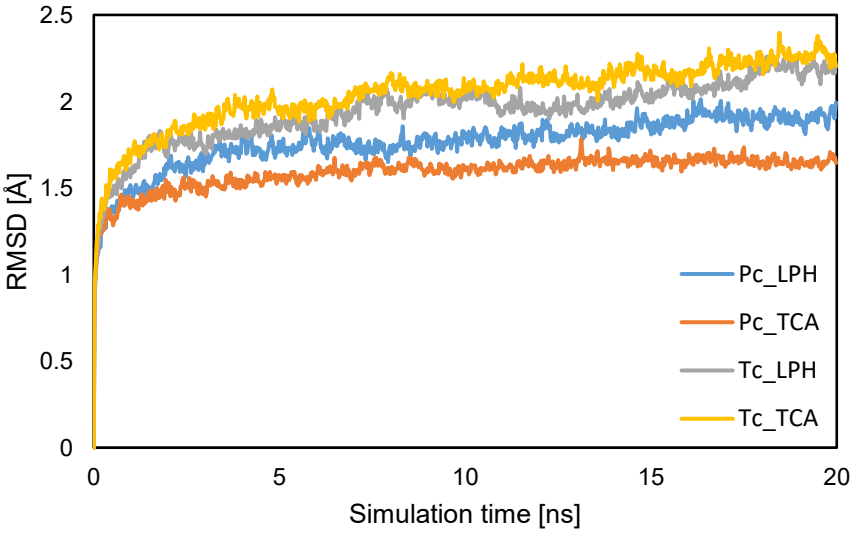


Figure S14. | Root mean square deviation (RMSD) calculated for all protein backbone atoms in PcPAL and TcPAM during MD equilibration. It is visible that all four equilibration simulations converged after ~2ns.

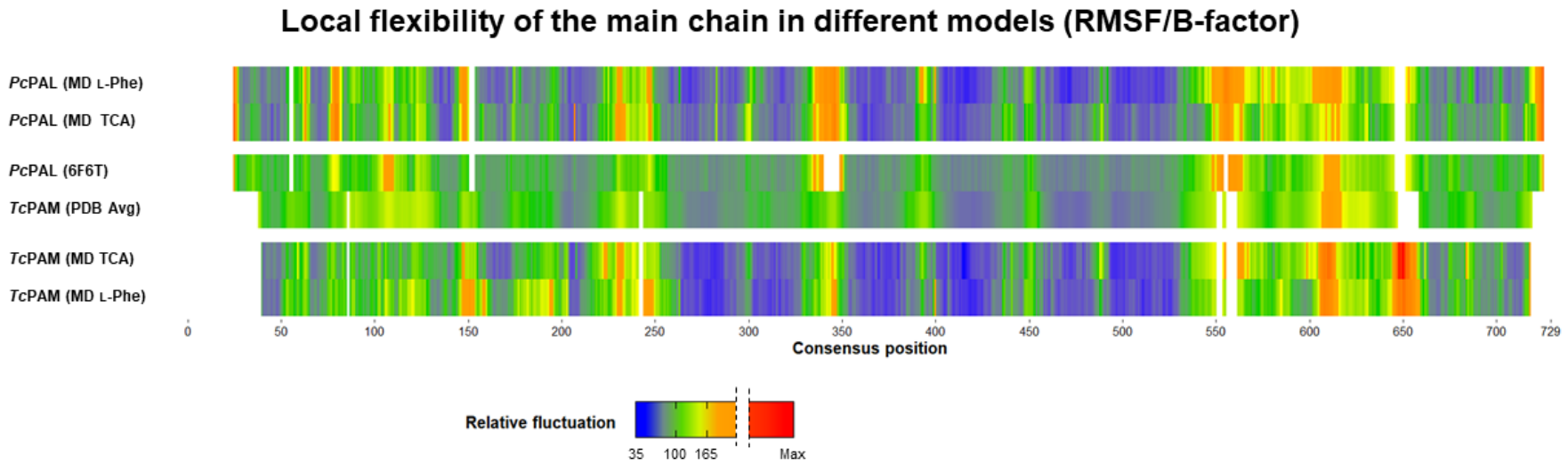


Figure S15. | Comparison of local flexibility of PcPAL and TcPAM by molecular dynamics simulation and crystal structure B-factors. As values correspond to the local flexibility of the proteins, the relative RMSF values in molecular dynamics simulations and relative B-factors in the crystal structures (for PcPAL as values in the structure 6F6T, for TcPAM as an average of the structures 2YII, 3NZ4, 4BAB, 4C5R) are plotted by the consensus positions of the sequence alignment. Relative RMSF values compared to relative B-factors suggest that local dynamics of PcPAL and TcPAM are quite similar. Comparable patterns from the crystal structures and the molecular dynamics indicate that the simulations reproduced well the experimentally observed protein flexibility. Residues with average flexibility are plotted in green, more rigid residues are shown in blue. Yellow indicates residues whose movement differs from the average by the same magnitude as the less flexible residues. Extremely flexible residues are represented by orange and red colors. Insertions in the sequence alignment and missing residues in the crystal structures are represented by white blanks.

SUPPORTING TABLES

Table S1. | Protein sequences of MIO-enzymes mentioned in this study.

Abbreviation	Organism	Enzyme function[a]	UniProt code
PcPAL	*Petroselinum crispum*	PAL	P24481
RtPAL	*Rhodosporidium toruloides*	PAL	P11544
SbPAL	*Sorghum bicolor*	PAL	C5XXT8
NpPAL	*Nostoc punctiforme*	PAL	B2J528
AvPAL	*Anabaena variabilis*	PAL	Q3M5Z3
KkPAL	*Kangiella koreensis*	PAL	C7R9W9
RsTAL	*Rhodobacter sphaeroides*	TAL	Q3IWB0
TcPAM	*Taxus canadensis*	PAM	Q6GZ04
TwPAM	*Taxus wallichiana* var. *chinensis*	PAM	Q68G84
PaPAM	*Pantoea agglomerans*	PAM	Q84FL5
SmPAM	*Streptomyces maritimus*	PAM	Q9KHJ9
PpHAL	*Pseudomonas putida*	HAL	P21310

[a] PAL: phenylalanine ammonia-lyase; PAM: phenylalanine 2,3-aminomutase; TAL: tyrosine ammonia-lyase; TAM: tyrosine 2,3-aminomutase; HAL: histidine ammonia-lyase.

Table S2. | Inner and outer loop sequence alignment of MIO containing class I lyase like enzymes which have crystal structure in the PDB.

Enzyme	Inner loop[a]	Outer loop[a]
	106 **110** 126	327 350
PcPAL	GTDS**Y**GVTTGF**G**ATSHRR--TKQ	GSAYVKA-AQKLHEM----------DPLQKPK**Q**DR
RtPAL	SMSV**Y**GVTTGF**G**GSADTR--TED	GSRFAVHHEEEVKVK----------DDEGILR**Q**DR
SbPAL	GGDI**Y**GVTTGF**G**GTSHRR--TKD	GSAFMKH-AKKVNEL----------DPLLKPK**Q**DR
NpPAL	AQPI**Y**GVTSGF**G**GMADVVISREQ	DSSLVRE-ELDGKHE----------YRGKDLI**Q**DR
AvPAL	GEPI**Y**GVTSGF**G**GMANVAISREQ	NSQLVRD-ELDGKHD----------YRDHELI**Q**DR
RsTAL	ARHV**Y**GLTTGF**G**PLANRLISGEN	GSARVVRHVIAERRLDA-----GDIGTEPEAG**Q**DA
TcPAM	GADI**Y**GVTTGF**G**ACSSRR--TNQ	SSPFQDL-SREYYSI----------DKLKKPK**Q**DR
TwPAM	GADI**Y**GVTTGF**G**ACSSRR--TNR	SSPFQEL-SREYYSI----------DKLKKPK**Q**DR
PaPAM	ERVI**Y**GVNTSM**G**GFVNYIVPIAK	DSSLAVNEHEVEKLIAE--EMDGLVKASNHQI**E**DA
SmPAM	ERVI**Y**GVNTSM**G**GFVDHLVPVSQ	DSHLAVNELDTEQTLAG--EMGTVAKAGSLAI**E**DA
PpHAL	DRTA**Y**GINTGF**G**LLASTRIASHD	DSSEVSLSH----------KNCDKV**Q**DP

[a] Absolute numbering of PcPAL sequence is shown.

Table S3. | Data collection and refinement statistics for PcPAL structures.

	PcPAL / Apo	**PcPAL / (*S*)-APPA**	**PcPAL / (*R*)-APEP**
PDB ID Code	6H2O	6F6T	6HQF
Data Collection			
Space group	C 2 2 21	C 2 2 21	C 2 2 21
Cell Dimensions			
a,b,c(Å)	119.89, 161.12, 141.43	118.74, 161.10, 141.65	118.88, 160.95, 141.58
α,β,γ(°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)[a]	48.09-1.90 (2.00-1.90)	95.58-1.90 (2.20-1.90)	38.70-1.72 (1.77-1.72)
Unique reflections	104777 (13880)	106555 (37601)	141190 (9135)
Redundancy	4.5 (4.6)	6.8 (6.9)	6.9
Completeness (%)	97.5 (91.9)	99.9 (99.9)	98.7 (87.0)
I/σI	8.78 (1.39)	8.88 (2.50)	8.52 (0.60)
R_{means}	0.144 (1.485)	0.150 (0.787)	0.161 (3.010)
$CC_{1/2}$	0.99 (0.52)	1.00 (0.85)	0.99 (0.17)
Refinement			Resolution: 38.70-1.76
Wilson B-Factor (Å^2)	28.0	19.2	30.8
R_{work} (%)	0.170	0.192	0.168
R_{free} (%)	0.206	0.239	0.205
RMS deviations			
Bond lengths (Å)	0.0071	0.0083	0.0078
Bond angle (°)	0.78	0.94	0.82
Ramachandran[b]			
Favored (%)	98.42	98.58	98.34
Outliers (%)	0.08	0.00	0.23
Clashscore	2.37	2.90	1.80
Average B-factor	33.0	28.0	36.0

[a] Statistics for the highest-resolution shell are shown in parentheses. [b] Categories were defined by PHENIX[21]. All non-glycine residues are included for this analysis.

Table S4. | Conversion of L-α-Phe, *rac*-β-Phe and CA by TcPAM variants after 168 h

TcPAM variant	**Starting substrate**	**α-Phe[a]**	**β-Phe**			**CA**
		Molar fraction[b] [%]	**Molar fraction[b] [%]**	**Config.[c]**	***ee*[c] [%]**	**Molar fraction[b] [%]**
wt	L-α-Phe	0.0	6.6	*R*	>95%	93.4
Q319E		82.6	3.7	*R*	n.d.	13.7
N458F		26.9	4.7	*R*	33%	68.5
Q319E/N458F		83.4	8.6	*S*	>95%	8.0
wt	*rac*-β-Phe	0.0	59.1	*S*	94%	40.9
Q319E		0.0	64.0	*S*	5%	36.0
N458F		0.0	77.2	*S*	15%	22.8
Q319E/N458F		0.0	59.2	*R*	61%	40.8
wt	CA	49.3	46.5	*R*	98%	4.2
Q319E		0.8	2.8	*S*	1%	96.4
N458F		3.8	2.6	*R*	71%	93.5
Q319E/N458F		0.0	1.2	*S*	>95%	98.8

[a] Enantiomeric excess of α-Phe (product, or residual) was >99% in all cases. [b] Molar fractions were determined with achiral HPLC. [c] Config.: Configuration of the β-Phe fraction, *ee*: enantiomeric excess of the β-Phe fraction, determined with chiral HPLC.

Table S5. | Primers used for cloning and mutagenesis of MIO-enzymes.

Primer name	Sequence[a]
Kk_H6_for	TTATGCTAGCATGACCGACACCAAAAC
Kk_H6_rev	TATTGAGCTCTTAGTTAGCGTAGGTAGAC
Pc_H6_for	TTATGCTAGCATGGAAAACGGAAACGGAG
Pc_H6_rev	TATTTGAGCTCTCAGGAAATGGGCAAGG
Tc_H6_for	TTATGCTAGCATGGGTTTCGCTGTTG
Tc_H6_rev	TATTGAGCTCTTAAGCAGATTTGTTCCAAAC
Kk_Q280E_for	CGTCTGGAAGACCGTTACTCTATCCGTTGCG
Kk_Q280E_rev	CGGTCTTCCAGACGGTCAGAGTTACGCGG
Kk_N421F_for	ATGCCACTTCCAGGACAAAGTTTCTATGGGTACGATCG
Kk_N421F_rev	CCTGGAAGTGGCATTCGGTAGAACGAGAGAAAACAG
Pc_Q348E_for	CCAAGGAAGACAGATACGCCTTGAGAACTTCCC
Pc_Q348E_rev	GTATCTGTCTTCCTTGGGCTTCTGCAAGGGG
Pc_N487F_for	GAACAGCATTTCCAGGACGTAAACTCCTTGGG
Pc_N487F_rev	CTGGAAATGCTGTTCGGCGGACTGTACATG
Tc_Q319E_for	CCGAAAGAAGACCGTTACGCTCTGCGTTCT
Tc_Q319_rev	CGGTCTTCTTTCGGTTTTTTCAGTTTGTCGATAGAGTAG
Tc_N458F_for	CAGCACTTCCAGGACATCAACTCTCTGGCT
Tc_N458F_rev	CCTGGAAGTGCTGTTCAGCAGAGTGAACGT

[a] Codons for the mutated residues are highlighted in red. The non-overlapping parts of the primers are indicated in green.

Table S6. | Protein sequences of MIO-enzyme constructs for mutation studies.

Enzyme	Sequence[a]
KkPAL-H10	MG**HHHHHHHHHH**SSGHIDDDDKHMMTDTKTNITFGHSSLTIEQICQLAKGNATAKLNSAPEFKHKIDQGAD FIKELLREDGVIYGVTTGYGDSVTTPVPVQDTHELPLHLTRFHGCGLGSIFSAEHTRAILATRLASLSQGY SGVSWSLLQQLELLLQKDILPRIPEEGSVGASGDLTPLSYVAAALIGEREVLYKGQTQPTEQVFKSLGIKP ITLQPKEGLAIMNGTAVMTALACLAFQRADYLTQLCSRITSLCSIALQGNSAHFDELLFSVKPHPGQNQVA AWIRDDLNHYKHPRNSDRLQDRYSIRCAPHIIGALKDAMPWMRQTIETELNSANDNPIIDGAGQHVLHGGH FYGGHIAMVMDSMKTGIANLADLMDRQMALLVDSKFNNGLPNNLSAASEQRRPLNHGFKAVQIGVSAWTAE ALKLTMPASVFSRSTECHNQDKVSMGTIAARDCLRILDLTEQVAAASLMAATQAVTLRIKQSQLDKSSLSD GVLSTLEQVFEHFELVSEDRPLEHELRHFVALIQEQHWSTYAN
KkPAL-H6	MGSS**HHHHHH**SSGR**ENLYFQG**HMASMTDTKTNITFGHSSLTIEQICQLAKGNATAKLNSAPEFKHKIDQGA DFIKELLREDGVIYGVTTGYGDSVTTPVPVQDTHELPLHLTRFHGCGLGSIFSAEHTRAILATRLASLSQG YSGVSWSLLQQLELLLQKDILPRIPEEGSVGASGDLTPLSYVAAALIGEREVLYKGQTQPTEQVFKSLGIK PITLQPKEGLAIMNGTAVMTALACLAFQRADYLTQLCSRITSLCSIALQGNSAHFDELLFSVKPHPGQNQV AAWIRDDLNHYKHPRNSDRLQDRYSIRCAPHIIGALKDAMPWMRQTIETELNSANDNPIIDGAGQHVLHGG HFYGGHIAMVMDSMKTGIANLADLMDRQMALLVDSKFNNGLPNNLSAASEQRRPLNHGFKAVQIGVSAWTA EALKLTMPASVFSRSTECHNQDKVSMGTIAARDCLRILDLTEQVAAASLMAATQAVTLRIKQSQLDKSSLS DGVLSTLEQVFEHFELVSEDRPLEHELRHFVALIQEQHWSTYAN
PcPAL-H10	MG**HHHHHHHHHH**SSGHIDDDDKHMLEMENGNGATTNGHVNGNGMDFCMKTEDPLYWGIAAEAMTGSHLDEV KKMVAEYRKPVVKLGGETLTISQVAAISARDGSGVTVELSEAARAGVKASSDWVMDSMNKGTDSYGVTTGF GATSHRRTKQGGALQKELIRFLNAGIFGNGSDNTLPHSATRAAMLVRINTLLQGYSGIRFEILEAITKFLN QNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPTGVILSPEEAFKLAGVEGGFFELQPKEGLAL VNGTAVGSGMASMVLFEANILAVLAEVMSAIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSA YVKAAQKLHEMDPLQKPKQDRYALRTSPQWLGPQIEVIRSSTKMIEREINSVNDNPLIDVSRNKAIHGGNF QGTPIGVSMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAMASYCSELQF LANPVTNHVQSAEQHNQDVNSLGLISSRKTSEAVEILKLMSTTFLVGLCQAIDLRHLEENLKSTVKNTVSS VAKRVLTMGVNGELHPSRFCEKDLLRVVDREYIFAYIDDPCSATYPLMQKLRQTLVEHALKNGDNERNLST SIFQKIATFEDELKALLPKEVESARAALESGNPAIPNRIEECRSYPLYKFVRKELGTEYLTGEKVTSPGEE FEKVFIAMSKGEIIDPLLESLESWNGAPLPIS
PcPAL-H6	MGSS**HHHHHH**SSGR**ENLYFQG**HMASMENGNGATTNGHVNGNGMDFCMKTEDPLYWGIAAEAMTGSHLDEVK KMVAEYRKPVVKLGGETLTISQVAAISARDGSGVTVELSEAARAGVKASSDWVMDSMNKGTDSYGVTTGFG ATSHRRTKQGGALQKELIRFLNAGIFGNGSDNTLPHSATRAAMLVRINTLLQGYSGIRFEILEAITKFLNQ NITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPTGVILSPEEAFKLAGVEGGFFELQPKEGLALV NGTAVGSGMASMVLFEANILAVLAEVMSAIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSAY VKAAQKLHEMDPLQKPKQDRYALRTSPQWLGPQIEVIRSSTKMIEREINSVNDNPLIDVSRNKAIHGGNFQ GTPIGVSMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAMASYCSELQFL ANPVTNHVQSAEQHNQDVNSLGLISSRKTSEAVEILKLMSTTFLVGLCQAIDLRHLEENLKSTVKNTVSSV AKRVLTMGVNGELHPSRFCEKDLLRVVDREYIFAYIDDPCSATYPLMQKLRQTLVEHALKNGDNERNLSTS IFQKIATFEDELKALLPKEVESARAALESGNPAIPNRIEECRSYPLYKFVRKELGTEYLTGEKVTSPGEEF EKVFIAMSKGEIIDPLLESLESWNGAPLPIS
TcPAM-H10	MG**HHHHHHHHHH**SSGHIDDDDKHMMGFAVESRSHVKDILGLINTFNEVKKITVDGTTPITVAHVAALARRH DVKVALEAEQCRARVETCSSWVQRKAEDGADIYGVTTGFGACSSRRTNQLSELQESLIRCLLAGVFTKGCA SSVDELPATATRSAMLLRLNSFTYGCSGIRWEVMEALEKLLNSNVSPKVPLRGSVSASGDLIPLAYIAGLL IGKPSVVARIGDDVEVPAPEALSRVGLRPFKLQAKEGLALVNGTSFATALASTVMYDANVLLLLVETLCGM FCEVIFGREEFAHPLIHKVKPHPGQIESAELLEWLLRSSPFQDLSREYYSIDKLKKPKQDRYALRSSPQWL APLVQTIRDATTTVETEVNSANDNPIIDHANDRALHGANFQGSAVGFYMDYVRIAVAGLGKLLFAQFTELM IEYYSNGLPGNLSLGPDLSVDYGLKGLDIAMAAYSSELQYLANPVTTHVHSAEQHNQDINSLALISARKTE EALDILKLMIASHLTAMCQAVDLRQLEEALVKVVENVVSTLADECGLPNDTKARLLYVAKAVPVYTYLESP CDPTLPLLLGLEQSCFGSILALHKKDGIETDTLVDRLAEFEKRLSDRLENEMTAVRVLYEKKGHKTADNND ALVRIQGSRFLPFYRFVREELDTGVMSARREQTPQEDVQKVFDAIADGRITVPLLHCLQGFLGQPNGSANG VESFQSVWNKSA
TcPAM-H6	MGSS**HHHHHH**SSGR**ENLYFQG**HMASMGFAVESRSHVKDILGLINTFNEVKKITVDGTTPITVAHVAALARR HDVKVALEAEQCRARVETCSSWVQRKAEDGADIYGVTTGFGACSSRRTNQLSELQESLIRCLLAGVFTKGC ASSVDELPATATRSAMLLRLNSFTYGCSGIRWEVMEALEKLLNSNVSPKVPLRGSVSASGDLIPLAYIAGL LIGKPSVVARIGDDVEVPAPEALSRVGLRPFKLQAKEGLALVNGTSFATALASTVMYDANVLLLLVETLCG MFCEVIFGREEFAHPLIHKVKPHPGQIESAELLEWLLRSSPFQDLSREYYSIDKLKKPKQDRYALRSSPQW LAPLVQTIRDATTTVETEVNSANDNPIIDHANDRALHGANFQGSAVGFYMDYVRIAVAGLGKLLFAQFTEL MIEYYSNGLPGNLSLGPDLSVDYGLKGLDIAMAAYSSELQYLANPVTTHVHSAEQHNQDINSLALISARKT EEALDILKLMIASHLTAMCQAVDLRQLEEALVKVVENVVSTLADECGLPNDTKARLLYVAKAVPVYTYLES PCDPTLPLLLGLEQSCFGSILALHKKDGIETDTLVDRLAEFEKRLSDRLENEMTAVRVLYEKKGHKTADNN DALVRIQGSRFLPFYRFVREELDTGVMSARREQTPQEDVQKVFDAIADGRITVPLLHCLQGFLGQPNGSAN GVESFQSVWNKSA

[a] The six/ten histidine residues are emphasized in bold. The tobacco etch virus protease recognition site (ENLYFQG) is highlighted in orange and the actual cleavage site (between Q and G) is shown in green.

Table S7. | HPLC conditions and retention times for molar fraction determinations on achiral C-18 phase

Eluent[a] [A:B%]	***T*** [°C]	**λ** [nm]	**Retention time**[b] [min]			**Molar response factor**[b]		
			1	**2**	**3**	**2*vs*1**	**3*vs*1**	**3*vs*2**
92:8	25	220	3.4	2.9	9.8	1.91	0.13	0.06

[a] Eluent components: A, NH_4OH buffer (0.1 M, pH 9.0); B, acetonitrile. Flow: 0.9 mL/min
[b] α-Phenylalanine: **1**, *rac*-β-phenylalanine: **2**, cinnamic acid: **3**.

Table S8. | HPLC conditions and retention times for enantiomeric excess determination on chiral ZWIX (+) phase

Eluent[a] [A:B:C%]	***T*** [°C]	**λ** [nm]	**Retention time**[b] [min]			
			(*S*)-**1**	(*R*)-**1**	(*S*)-**2**	(*R*)-**2**
49:49:2	25	220	9.2	10.6	19.0	21.8

[a] Eluent components: A: MeOH; B: acetonitrile; C: water; 25 mM diethylamine, 50 mM formic acid. Flow: 0.7 mL/min
[b] α-Phenylalanine: **1**, β-phenylalanine: **2**.

Table S9. | Active site configurations in the four PcPAL / TcPAM models.[a]

Enzyme model	**Active site A**	**Active site B**	**Active site C**	**Active site D**
PcPAL / Phe (RAMD)	Phe (RAMD) MIO TYD-110	Phe MIO TYD-110	no ligand MIO TYD-351	no ligand MIO TYD-110
PcPAL / CA (RAMD)	CA (RAMD) NH_2-MIO	CA NH_2-MIO	no ligand NH_2-MIO	no ligand MIO TYD-110
TcPAM / Phe (RAMD)	Phe (RAMD) MIO TYD-110	Phe MIO TYD-110	no ligand MIO TYD-351	no ligand MIO TYD-110
TcPAM / CA (RAMD)	CA (RAMD) NH_2-MIO	CA NH_2-MIO	no ligand NH_2-MIO	no ligand MIO TYD-110

[a] Phe: L-phenylalanine (zwitterionic), CA: *trans*-cinnamate, MIO: non-modified MIO, NH_2-MIO: amino-MIO state, TYD-110: deprotonated state of Tyr110, TYD-351: deprotonated state of Tyr351, RAMD: The random acceleration was applied on the ligand indicated.

Table S10. | Dependence of ligand exit time on acceleration and displacement settings of the RAMD simulations in PcPAL and TcPAM.[a]

Acceleration [kcal mol^{-1} Å^{-1}]	**Displacement** [Å]	t_{exit} **in TcPAM**[b] [ps]		t_{exit} **in PcPAL**[b] [ps]	
		Phe	CA	Phe	CA
0.15	0.05	44/56/91	78/25/46	164/149/40	62/47/243
0.14	0.05	38/21/57	81/44/92	40/63/50	51/224/171
0.13	0.05	52/102/116	196/82/211	286/309/41	135/109/287
	0.01	671/314/253	1881/×/382	216/464/225	×/468/755
0.12	0.05	117/97/240	77/53/182	88/820/279	196/198/364
	0.1	369/479/577	388/136/790	716/185/367	1867/1692/65
	0.01	×/80/145	72/1361/×	×/104/432	×/535/571
0.11	**0.05**	**164/182/668**	**391/162/232**	**225/411/791**	**407/917/728**
	0.1	1096/868/1226	85/121/1210	1337/1921/×	×/×/1642
	0.01	×/118/53	200/281/×	356/×/761	×/589/×
0.10	0.05	815/361/319	194/88/506	1038/×/222	1750/1039/459
	0.1	1281/1723/×	481/428/×	×/×/×	×/×/×
0.09	0.05	839/851/377	434/223/452	×/×/×	1892/×/×
0.08	0.05	692/407/1519	×/×/×	×/×/×	×/×/×
0.07	0.05	×/×/×	×/×/×	×/×/×	×/×/×

[a] Values selected as optimal setting are emphasized by bold typesetting in dark red color.
[b] The three values correspond to exit times from three different equilibrated starting model structures; × means no exit occurred within 2000 ps simulation.

SUPPORTING REFERENCES

(1) Guca, E.; Nagy, G. N.; Hajdú, F.; Marton, L.; Izrael, R.; Hoh, F.; Yang, Y.; Vial, H.; Vértessy, B. G.; Guichou, J.-F.; Cerdan, R. Structural Determinants of the Catalytic Mechanism of Plasmodium CCT, a Key Enzyme of Malaria Lipid Biosynthesis. *Sci. Rep.* **2018**, *8* (1), 11215.

(2) Liu, H.; Naismith, J. H. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008**, *8*, 91.

(3) Dima, N. A.; Filip, A.; Bencze, L. C.; Oláh, M.; Sátorhelyi, P.; Vértessy, B. G.; Poppe, L.; Paizs, C. Expression and Purification of Recombinant Phenylalanine Ammonia-Lyase from *Petroselinum crispum*. *Stud. Univ. Babes-Bolyai Chem.* **2016**, *61* (2), 21–34.

(4) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858.

(5) Schrödinger Release 2017-4: Maestro, Schrödinger, LLC, New York, NY, 2017.

(6) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537.

(7) Bata, Z.; Qian, R.; Roller, A.; Horak, J.; Bencze, L. C.; Paizs, C.; Hammerschmidt, F.; Vértessy, B. G.; Poppe, L. A Methylidene Group in the Phosphonic Acid Analogue of Phenylalanine Reverses the Enantiopreference of Binding to Phenylalanine Ammonia-Lyases. *Adv. Synth. Catal.* **2017**, *359* (12), 2109–2120.

(8) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes. *J. Med. Chem.* **2006**, *49* (21), 6177–6196.

(9) Wybenga, G. G.; Szymanski, W.; Wu, B.; Feringa, B. L.; Janssen, D. B.; Dijkstra, B. W. Structural Investigations into the Stereochemistry and Activity of a Phenylalanine-2,3-Aminomutase from *Taxus chinensis*. *Biochemistry* **2014**, *53*, 3187–3198.

(10) Peterkofsky, A. The Mechanism of Action of Histidase: Amino-Enzyme Formation and Partial Reactions. *J. Biol. Chem.* **1962**, *237* (3), 787–795.

(11) Wu, B.; Szymański, W.; Wybenga, G. G.; Heberling, M. M.; Bartsch, S.; de Wildeman, S.; Poelarends, G. J.; Feringa, B. L.; Dijkstra, B. W.; Janssen, D. B. Mechanism-Inspired Engineering of Phenylalanine Aminomutase for Enhanced β-Regioselective Asymmetric Amination of Cinnamates. *Angew. Chem. Int. Ed.* **2012**, *51* (2), 482–486.

(12) Heberling, M. M.; Masman, M. F.; Bartsch, S.; Wybenga, G. G.; Dijkstra, B. W.; Marrink, S. J.; Janssen, D. B. Ironing out Their Differences: Dissecting the Structural Determinants of a Phenylalanine Aminomutase and Ammonia Lyase. *ACS Chem. Biol.* **2015**, *10* (4), 989–997.

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

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

(15) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.

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

(17) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(18) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621.

(19) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N*log(N) Method for Ewald Sums in Large

Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.

(20) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant Pressure Molecular Dynamics Algorithms. *J. Chem. Phys.* **1994**, *101* (5), 4177–4189.

(21) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.