

Electronic Supplementary Information

Expanding the substrate scope of phenylalanine ammonia-lyase from *Petroselinum crispum* towards styrylalanines

László Csaba Bencze, Alina Filip, Gergely Bánóczy, Monica Ioana Toşa, Florin Dan Irimie,
Ákos Gellért, László Poppe, Csaba Paizs

Table of Contents

1. Materials	2
2. Instrumentation	2
3. Chemical synthesis of substrates <i>rac</i> - 1a-d and 2a-d	3
3.1. Synthesis of (2 <i>E</i> ,4 <i>E</i>)-styrylacrylates 2a-d (Scheme S1)	4
3.1.1. Synthesis of styrylacrylic esters 4a-d	4
3.1.2. Synthesis of the styryl allyl alcohols 5a-d	5
3.1.3. Synthesis of the styrylacroleins 6a-d	5
3.1.4. Synthesis of the styrylacrylates 2a-d	5
3.2. Synthesis of racemic amino acids <i>rac</i> - 1a-d (Scheme S2)	6
4. Michaelis-Menten curves with non-linear fitting for calculation of the kinetic parameters (K_M , v_{max} , k_{cat})	16
5. HPLC monitoring of the enzymatic reactions	25
5.1. Determination of the conversion by HPLC	25
5.2 HPLC chromatograms of the PcPAL-catalyzed ammonia elimination reaction from <i>rac</i> - 1a-d at the end-point of monitoring	26
5.3. HPLC methods to determine the enantiomeric excess (ee) of D- 1a-d	30
5.4 Chiral HPLC analysis of the final products from wt-PcPAL-catalyzed kinetic resolutions of <i>rac</i> - 1a-d	33
5.6. Investigation of the ammonia addition onto styrylacrylates 2a-d by the PcPAL variants	38
6. SDS-PAGE and size-exclusion chromatogram from the purification steps of F137V-PcPAL	39
7. Theoretical methods and derivations	40
7.1. Molecular modeling principles	40
7.2. Molecular modeling methods	41
7.3. Derivation and reasoning of Eq. 3 from the main text	45
References	47

1. Materials

The aldehydes **3a-d** used as starting materials, as well the other organic, inorganic reactants and solvents used in the chemical synthesis of the substrates were purchased from Sigma-Aldrich (St. Louis, MO, USA) and/or Alfa-Aesar (Haverhill, MA, USA). All solvents were purified and dried by standard methods as required. The primers used for the mutagenesis were purchased through the services of Genomed (Debrecen, Hungary). IPTG, Phusion Hot Start DNA Polymerase, dNTPs, DpnI, XhoI, Bpu1102I, agarose were all products of Thermo Fischer Scientific (Waltham, MA, USA). Plasmid extraction kit and ethidium bromide were purchased from Sigma-Aldrich (St. Louis, MO, USA). LB medium were from Liofilchem (Roseto, Italy), protease inhibitor cocktail from Hoffman La-Roche (Basel, Switzerland), while the Ni-NTA Superflow resin used for affinity chromatography was from Qiagen (Hilden, Germany).

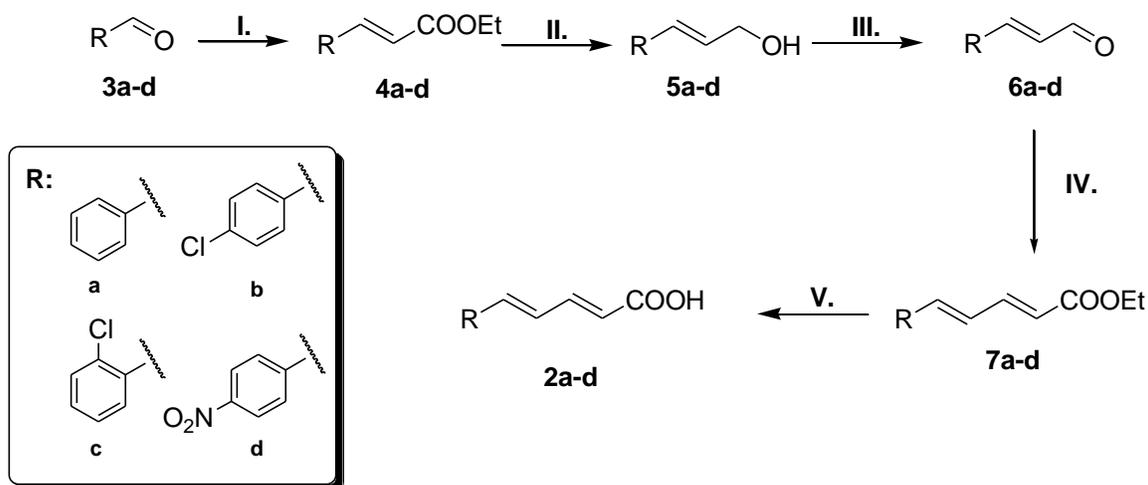
2. Instrumentation

The ^1H and ^{13}C NMR spectra were recorded on Bruker (Billerica, MA, USA) Avance spectrometers operating at 400 MHz and 101 MHz / 600 MHz and 151 MHz, respectively. Spectra were recorded at 25 °C in CDCl_3 , D_2O , MeOD, DMSO. ^1H and ^{13}C NMR spectra were referenced internally to the solvent signal. MS spectra and LC-MS analysis were recorded on Agilent 6410 Triple Quadrupole LC/MS mass spectrometry system. LC-MS measurements of *rac-1a-d* were performed using Phenomenex Kinetex 2.6 μm C18 100Å, 50×2.1mm column, acetonitrile 70%, water (0.1 HCOOH) 30% as mobile phase at 0.3 ml/min flow rate. The MS detector was operated in positive electrospray ionization mode, with source temp 35 °C, capillary voltage 4000 V, fragmentor 120 V and with MS2Scan mode, at least +/- 50 amu around molecular ion. High performance liquid chromatography (HPLC) analyses were conducted with an Agilent (Santa Clara, CA, USA) 1200 instrument. Kinetic measurements were performed on and Agilent 8453 UV-Vis spectrophotometer. For the PCR reactions, the Mastercycler from Eppendorf (Hamburg, Germany) was used. Gene sequencing services were performed through Genomed (Debrecen, Hungary). Protein purification with size exclusion chromatography were performed with an Äkta purifier FPLC (GE Healthcare, Sweden). Preparative scale enzymatic reactions were performed

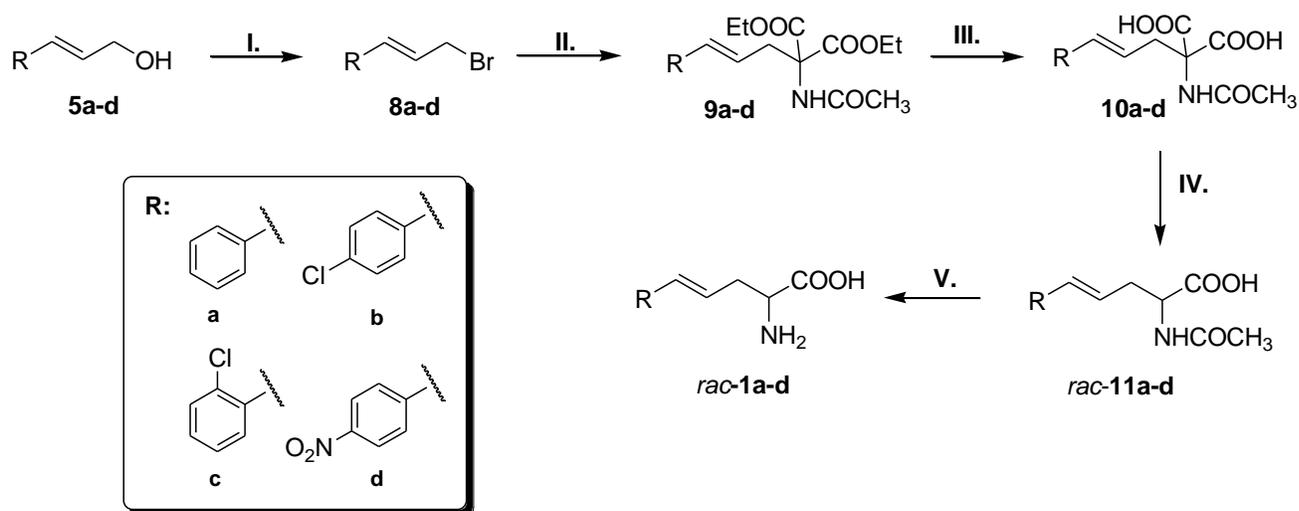
with Heidolph (Schwabach, Germany) Titramax 1100 equipped with incubator module. DNA concentrations were measured with Thermo Fischer Scientific (Waltham, MA, USA) Nanodrop 2000 instrument. SDS-PAGE was performed using the BioRad (Hercules, CA, USA) Mini-Protean Tetra Cell vertical electrophoresis unit, using 10% Tris-glycine Laemmli-gels.

3. Chemical synthesis of substrates *rac*-1a-d and 2a-d

Starting from the commercially available aldehydes **3a-d** acrylic esters **4a-d** were obtained through Wittig reaction, using the corresponding triphenyl-phosphoranylidine. The reduction of esters **4a-d** with DIBAL-H afforded the primary alcohol derivatives **5a-d**, which through an oxidation with manganese-dioxide, followed by another Wittig reaction with triphenyl-phosphoranylidene were converted into styrylacrylic esters **7a-d**. Mild alkaline hydrolysis from esters **7a-d** resulted in the styrylacrylates **2a-d** (Scheme S1).



Scheme S1. Synthesis of (*2E,4E*) styrylacrylates **2a-d**. Reagents and conditions: **I.** $\text{Ph}_3\text{P}=\text{CH-CO}_2\text{Et}$, toluene, reflux, 24 h; **II.** DIBAL-H/ CH_2Cl_2 at -60°C , 1 h; **III.** $\text{MnO}_2/\text{CH}_2\text{Cl}_2$, RT, 48 h; **IV.** $\text{Ph}_3\text{P}=\text{CH-CO}_2\text{Et}$, toluene, reflux, 24 h; **V.** 10% KOH, reflux, 20 h.



Scheme S2. Synthesis of racemic styrylalanines *rac-1a-d*. Reagents and conditions: **I.** $(\text{CH}_3)_3\text{SiCl}$, NaBr/MeCN , reflux, 5 h, $50\text{ }^\circ\text{C}$; **II.** NaH , $\text{CH}_3\text{CONHCH}(\text{COOEt})_2/\text{DMF}$, $60\text{ }^\circ\text{C}$, 3 h; **III.** 10% NaOH in water/ MeOH , $60\text{ }^\circ\text{C}$, 5 h; **IV.** toluene, reflux, 20 h; **V.** dioxane/ 18% HCl , reflux, 4 h.

The styrylic alcohols **5a-d**, obtained as described above (Scheme S1) were converted into the diethyl acetamidomalonates **9a-d** *via* malonic acid coupling of the brominated compounds **8a-d**. Further, through a mild alkaline hydrolysis of **9a-d**, followed by the subsequent decarboxylation of **10a-d**, the *N*-acylated amino acids *rac-11a-d* were obtained. Finally, deprotection of *rac-11a-d* afforded the racemic amino acids *rac-1a-d* (Scheme S2).

3.1. Synthesis of (2*E*,4*E*)-styrylacrylates **2a-d** (Scheme S1)

3.1.1. Synthesis of styrylacrylic esters **4a-d**

Into the stirred solution of aldehydes **3a-d** (18 mmol) in dry toluene (60 mL) triphenylphosphoranylidene (8.1 g, 23.4 mmol) was added in portions. The mixture was stirred at $110\text{ }^\circ\text{C}$ until completion of the reaction (approx. 12 h, checked by TLC), followed by the evaporation of the solvent. The residue was purified by silica gel column chromatography using CH_2Cl_2 as eluent leaving the product **4a-d** as colorless or yellow viscous oil in 85-90% yield. The ^1H , ^{13}C -NMR spectra of the products were in accordance with the reported data.¹

3.1.2. Synthesis of the styryl allyl alcohols **5a-d**

Into the cooled solution of acrylic esters **4a-d** (3.8 mmol) in dry CH₂Cl₂ (16 mL), at -60 °C under argon atmosphere, 10 mL of DIBAL-H (9.14 mmol, 1M solution in hexane) was added dropwise. The mixture was stirred for 1 h at 0 °C and after completion of reaction (checked by TLC), the excess of DIBAL-H was quenched with 1 mL of MeOH, followed by addition of 2 mL of saturated solution of NH₄Cl. The mixture was stirred again at 0 °C for 15 min, followed by extraction with CH₂Cl₂ (3 × 25 mL) and water (3 × 25 mL). The organic phase was dried over anhydrous MgSO₄, filtered and concentrated in vacuum. The residue was purified by silica gel column chromatography using CH₂Cl₂ as eluent leaving the product **5a-d** as colorless or yellow viscous oil in 90-95% yields. The ¹H-, ¹³C-NMR spectra of the products are in accordance with the reported data.²

3.1.3. Synthesis of the styryl acroleins **6a-d**

Into the stirred solution of alcohols **5a-d** (3.7 mmol) in dry CH₂Cl₂ (25 mL) was added manganese dioxide (1.6 g, 18.5 mmol) in portions. The mixture was stirred for 24 h until the reaction completed (checked by TLC). The mixture was filtered under vacuum through a short silicagel pad, using CH₂Cl₂ as eluent, followed by the removal of the solvent in vacuum to give the product **6a-d** as yellow or orange viscous oil in 84-89% yield. The ¹H-, ¹³C-NMR spectra of the products are in accordance with the reported data.³

3.1.4. Synthesis of the styryl acrylates **2a-d**

The reactions from aldehydes **6a-d** were performed similarly as described in Section 3.1.1. for the reactions from **3a-d** (Synthesis of acrylate esters **4a-d**). The products were purified by silica gel column chromatography using CH₂Cl₂ as eluent leaving the styryl acrylic esters **7a-d** as yellow-orange oils in 81-85% yields.

The esters **7a-d** were subsequently hydrolyzed by the addition of 10% NaOH (5 mL) and stirring under reflux until the reaction completed (4-20 h, checked by TLC). The resulted mixture was diluted with water (5 mL), and extracted with ethyl acetate (3×5 mL). The aqueous phase was

acidified with 10% HCl solution to pH 1-2, followed by extraction with ethyl acetate (3 × 10 mL). The combined organic phases were dried over anhydrous MgSO₄, filtered, followed by the evaporation of the solvent in vacuum to give **2a-d** as white or yellow-white solid in 70-76% yield.

(2*E*,4*E*)-5-Phenylpenta-2,4-dienoic acid (**2a**)

Yield: 76%, white solid; ¹H NMR (400 MHz, CDCl₃) δ: 7.55 (dd, *J* = 15.3, 9.7 Hz, 1H), 7.49 (d, *J* = 7.0 Hz, 2H), 7.35 (ddd, *J* = 10.7, 9.9, 5.3 Hz, 3H), 7.01–6.84 (m, 2H), 6.01 (d, *J* = 15.2 Hz, 1H). ¹³C NMR (100 MHz, CDCl₃, δ): 172.55, 147.13, 141.80, 135.95, 129.47, 129.00, 127.50, 126.09, 120.41. The NMR data are in accordance with data from literature.⁴

(2*E*,4*E*)-5-(4-Chlorophenyl)penta-2,4-dienoic acid (**2b**)

Yield: 70%, yellow-white solid; ¹H NMR (600 MHz, D₂O) δ: 7.70 (d, *J* = 7.5 Hz, 1H), 7.34-7.48 (m, 2H), 7.2-7.32 (m, 3H), 7.00 (m, 1H), 6.04 (d, *J* = 15.2 Hz, 1H); ¹³C NMR (151 MHz, D₂O): δ = 172.3, 145.8, 136.5, 132.4, 131.1, 130.9, 130.7, 130.1, 129.9, 128.3, 128.1. The NMR data are in accordance with data from literature.⁵

(2*E*,4*E*)-5-(2-Chlorophenyl)penta-2,4-dienoic acid (**2c**)

Yield: 73%, white solid; ¹H NMR (600 MHz, D₂O) δ: 7.42 (d, *J* = 8.1 Hz, 2H), 7.33 (d, *J* = 7.9 Hz, 2H), 7.13 – 7.00 (m, 1H), 6.96 – 6.85 (m, 1H), 6.80 (d, *J* = 15.6 Hz, 1H), 5.98 (d, *J* = 15.2 Hz, 1H). ¹³C NMR (151 MHz, D₂O): δ = 127.7, 127.8, 128.2, 128.9, 133.4, 135.1, 136.6, 140.9, 175.8. The NMR data are in accordance with data from literature.⁶

(2*E*,4*E*)-5-(4-Nitrophenyl)penta-2,4-dienoic acid (**2d**)

Yield: 72 %, brownish-yellow solid, ¹H NMR (600 MHz, D₂O) δ: 8.02 (d, *J* = 6.8 Hz, 2H), 7.47 (d, *J* = 6.7 Hz, 2H), 6.96 (m, 1H), 6.91 (m, 1H), 6.74 (d, *J* = 15.1 Hz, 1H), 5.99 (d, *J* = 14.9 Hz, 1H). ¹³C NMR (151 MHz, D₂O): 175.4, 146.6, 143.4, 140.1, 135.2, 131.5, 130.1, 127.4, 124.1. The NMR data are in accordance with data from literature.⁴

3.2. Synthesis of racemic amino acids *rac*-**1a-d** (Scheme S2)

Into the stirred solution of styryl allyl alcohol **5a-d** (10 mmol) in dry CH₂Cl₂ (50 mL) triphenylphosphine (5.5 g, 21 mmol) was added, followed by a portionwise addition of *N*-

bromosuccinimide (3.56 g, 20 mmol). The mixture was stirred for at room temperature, until the reaction was completed (0.5-2 h, checked by TLC). The solvent was evaporated in vacuum, followed by the purification of the residue by silica gel column chromatography using CH₂Cl₂ as eluent leaving the bromides **8a-d** which were directly used in the coupling reaction with malonic acid.

NaH (0.48 g, 11 mmol, 55% suspension in mineral oil) was added to dry *N,N*-dimethylformamide (20 mL) and stirred at room temperature under argon atmosphere. After 30 min, diethyl acetamidomalonate (2.23 g, 10 mmol) was added and the mixture was stirred for 30 min. To the ice-cooled mixture was added dropwise a solution of the bromide **8a-d** (10 mmol) dissolved in dry *N,N*-dimethylformamide (10 mL). The reaction mixture was stirred at room temperature for 3 h and at 60 °C for 4 h. The solution was cooled and poured on a water-ice mixture (200 mL). The formed precipitate was filtered off and dried under reduced pressure. The product was re-dissolved in CH₂Cl₂ (1-2 mL) and purified by silica gel column chromatography using CH₂Cl₂ as eluent leaving the diethyl 2-acetamido-2-cinnamyl malonic acid derivatives **9a-d**.

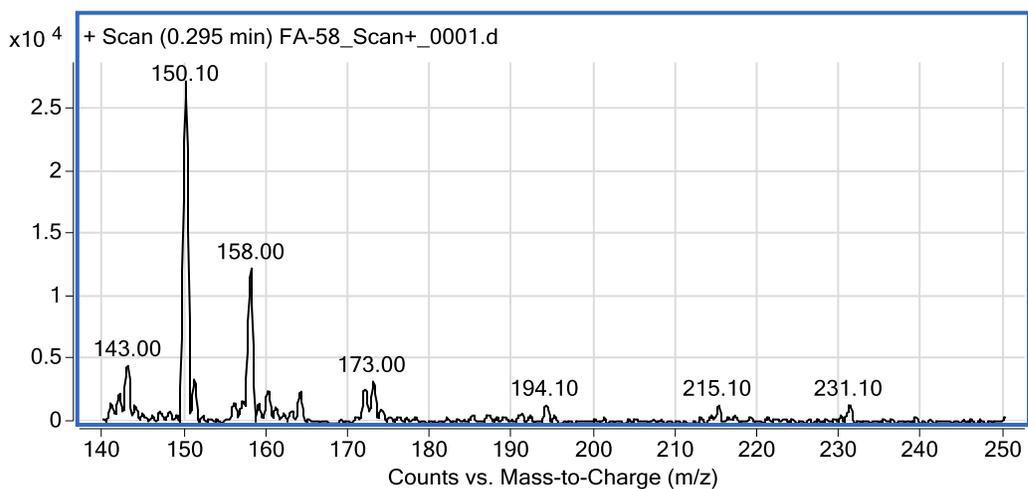
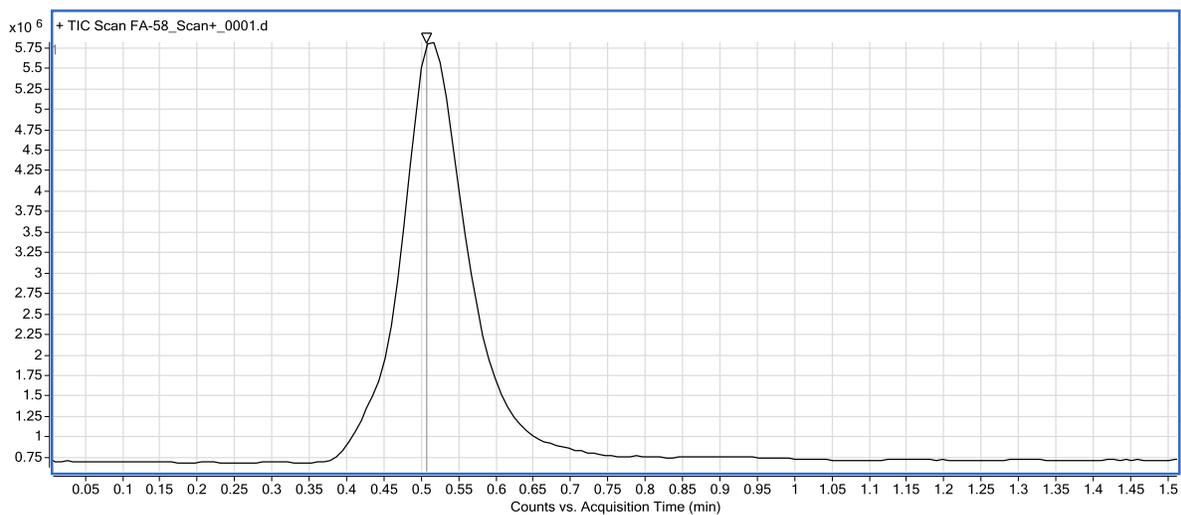
The obtained malonic acid **9a-d** (6 mmol) was dissolved in methanol (5 mL), followed by addition of 10% NaOH solution (5 mL). The resulted mixture was stirred at room temperature for 3 h, and at 60 °C until the reaction completed (2-3 h, checked by TLC). Subsequently, the methanol was removed in vacuum and the concentrated aqueous solution was diluted with water (10 mL) and extracted with ethyl acetate (3 × 10 mL). The pH of the aqueous phase was adjusted to 1 with 10% HCl solution, followed by extraction with ethyl acetate (3 × 15 mL). The combined organic phases were dried over Na₂SO₄, filtered, followed by the removal of the solvent in vacuum.

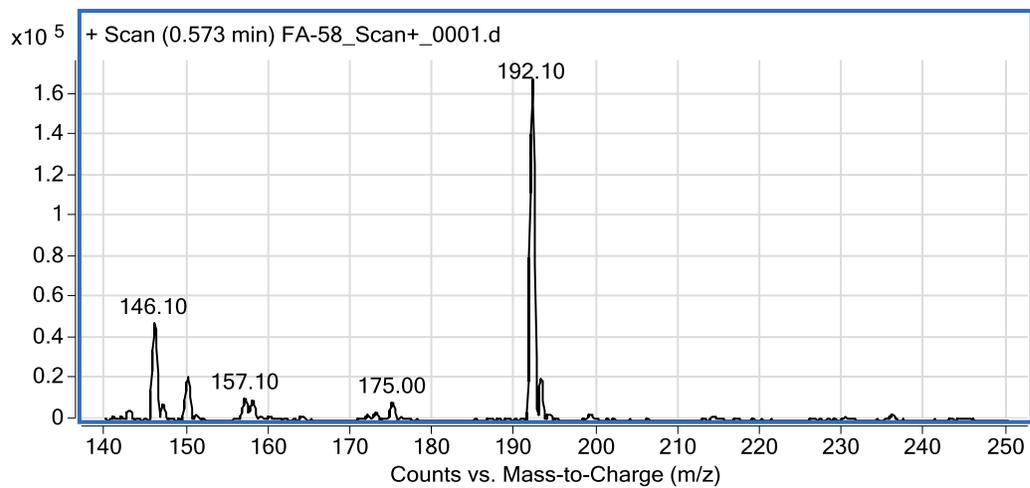
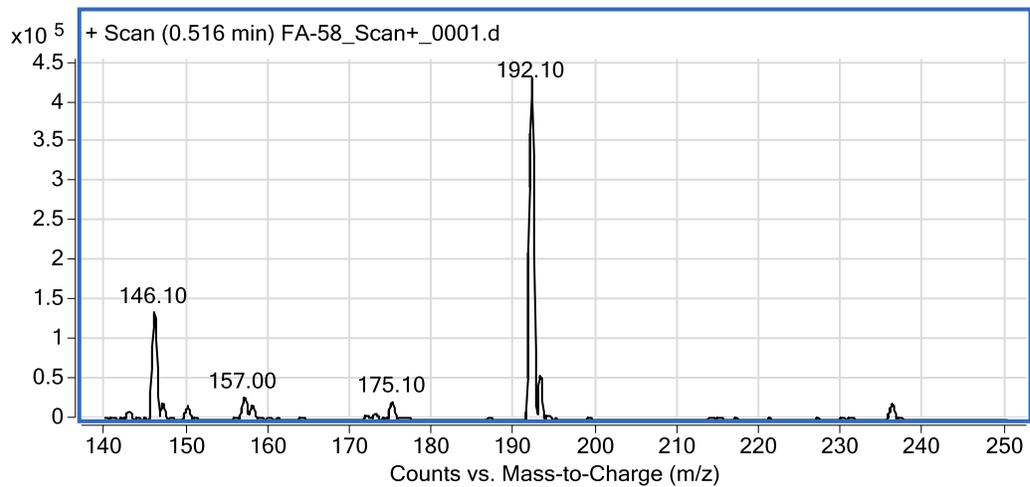
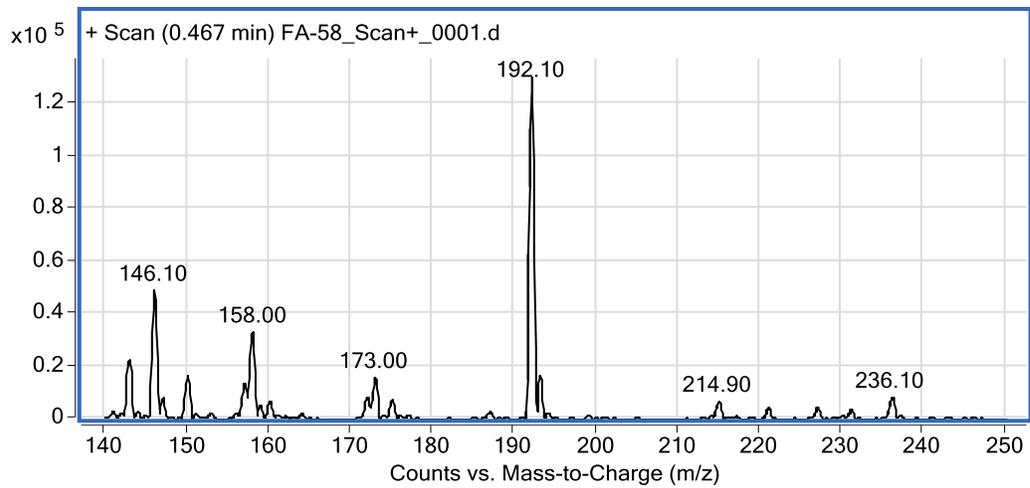
The formed dicarboxylic acid derivative **10a-d** was suspended in toluene (10 mL) and heated under reflux for 20 h. The solvent was removed in vacuum to leave the products **11a-d**.

The *N*-acetyl styrylalanine **11a-d** was suspended in dry 1,4-dioxane (10 mL) followed by the addition of concentrated HCl (0.5 mL). The resulted mixture was heated under reflux until the reaction completed (4-12 h, checked by TLC). The organic solvent was evaporated in vacuum and the obtained precipitate was washed with anhydrous diethyl ether (3 × 5 mL) to give the pure racemic styrylalanines *rac*-**1a-d**.

rac-(*E*)-2-Amino-5-phenylpent-4-enoic acid (*rac*-**1a**)

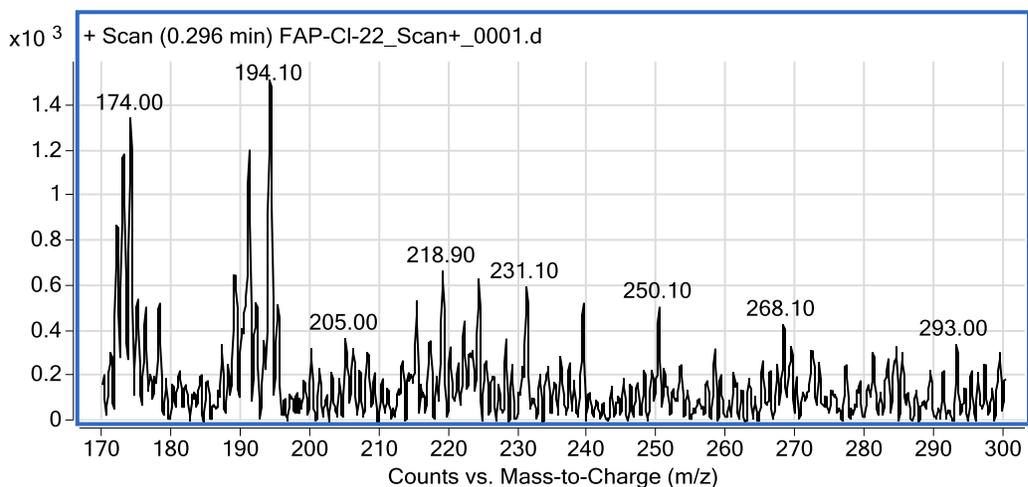
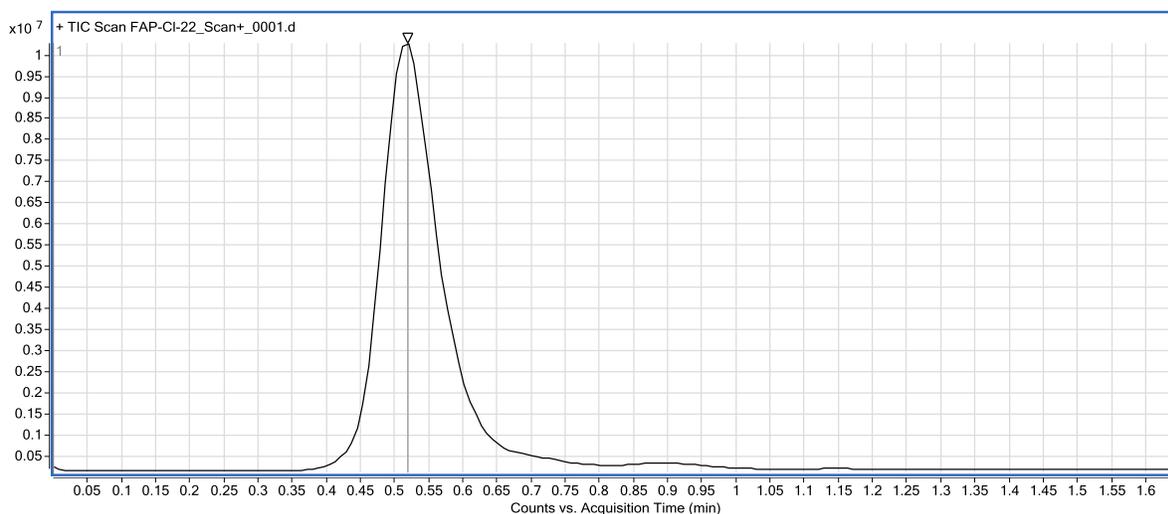
^1H NMR (400 MHz, D_2O) δ : 7.41-7.51 (m, 2H), 7.32-7.40 (m, 2H), 7.24-7.27 (m, 1H), 6.63 (d, J = 15.4 Hz 1H), 6.14-6.22 (m, 1H), 4.21-4.24 (m, 1H), 2.86 (m, 2H); ^{13}C -NMR: 171.4, 136.3, 135.5, 128.9, 128.1, 126.3, 121.6, 52.5, 33.5; LC-MS: positive ionization mode m/z : 192.1 ($[\text{M}+\text{H}]^+$, (calculated for $\text{C}_{11}\text{H}_{13}\text{NO}_2\text{S}$: 192.0946 ($[\text{M}+\text{H}]^+$); 146.1 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]$). The four illustrative scans represent the spectral background of the measurement and the MS spectra at half peak height of peak front, at retention time and at half peak height of peak tail, correspondingly.

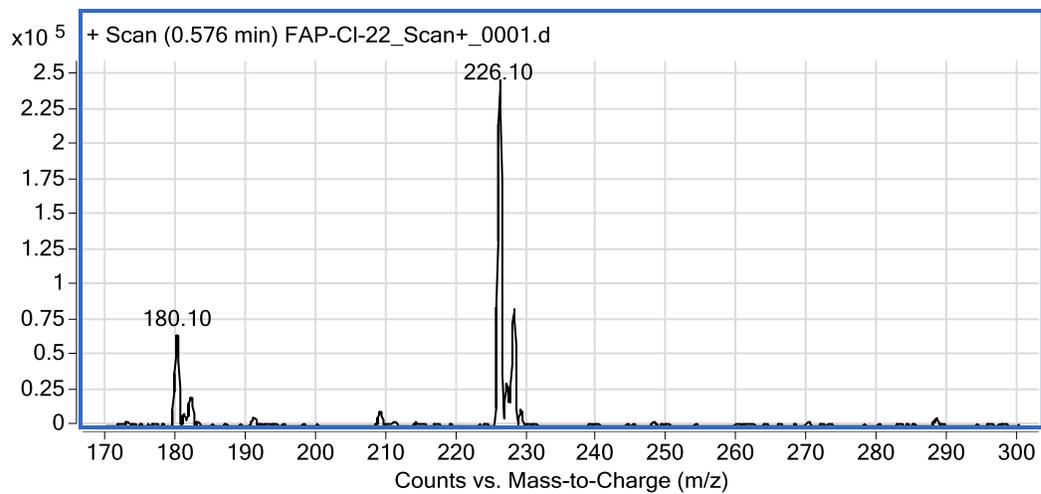
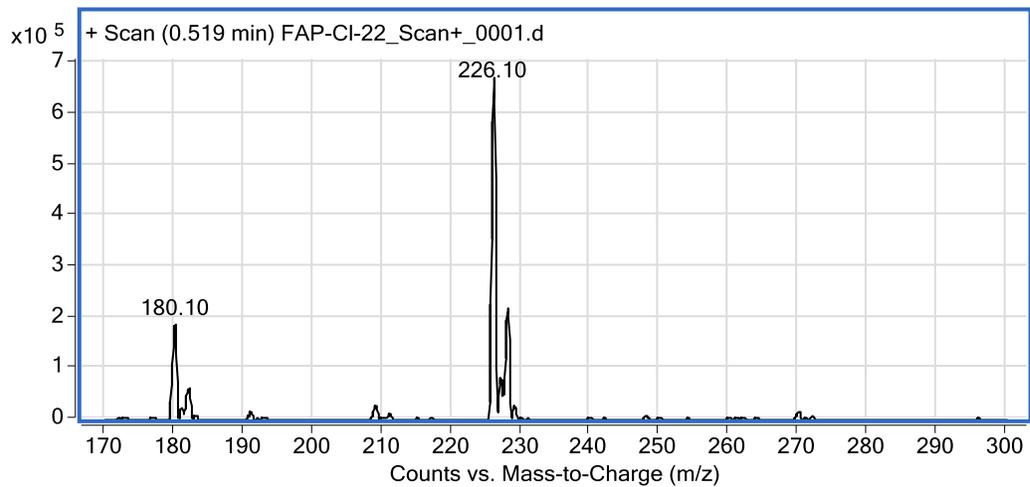
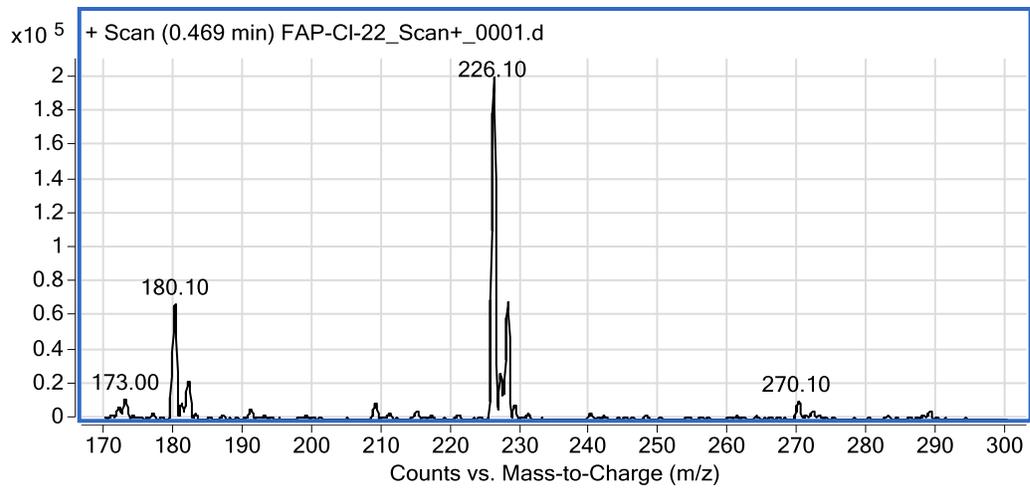




rac-(*E*)-2-Amino-5-(4-chlorophenyl)pent-4-enoic acid (*rac*-**1b**)

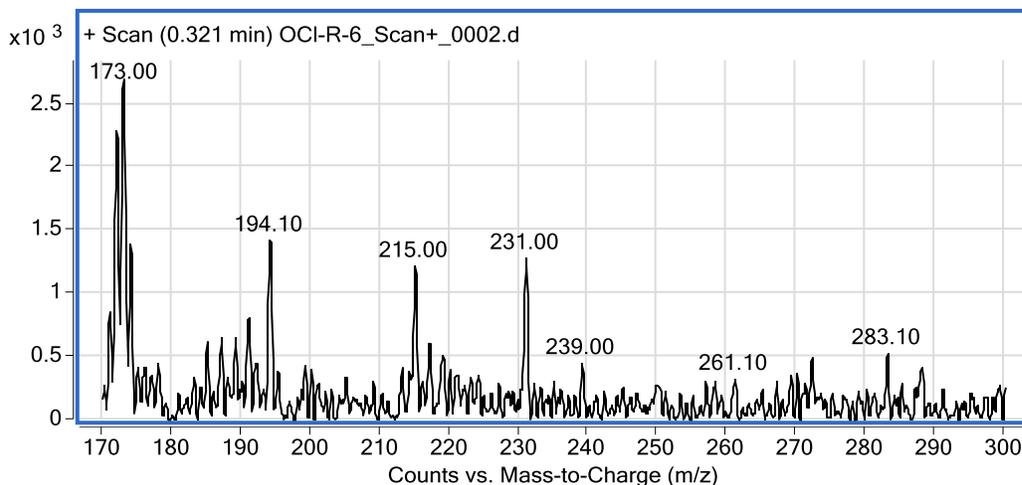
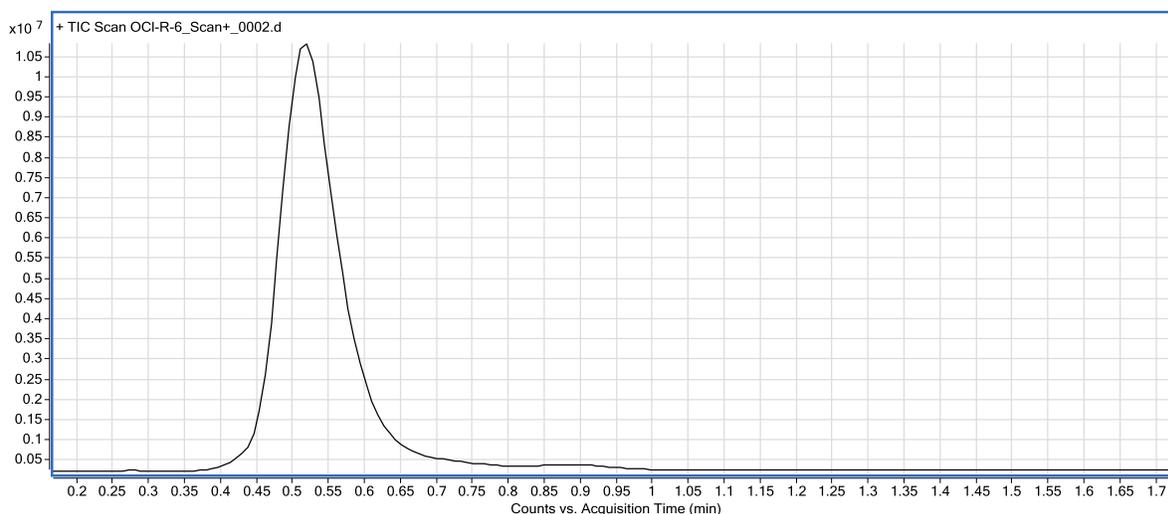
^1H NMR (600 MHz, D_2O) δ : 7.34-7.41 (m, 2H), 7.27-7.33 (m, 2H), 6.43 (d, $J = 14.7$ Hz, 1H), 6.18 (m, 1H), 3.32 (m, 1H), 2.39-2.49 (m, 2H). ^{13}C -NMR: 182.6, 136, 132.2, 131.4, 128.6, 127.5, 127.3, 55.6, 38.3; LC-MS: positive ionization mode, m/z : 226.1 ($[\text{M}+\text{H}]^+$, ^{35}Cl), 228.1 ($[\text{M}+\text{H}]^+$, ^{37}Cl) (calculated for $\text{C}_{11}\text{H}_{12}\text{ClNO}_2$ 226.0557 ($[\text{M}+\text{H}]^+$, ^{35}Cl), 228.0527 ($[\text{M}+\text{H}]^+$, ^{37}Cl), 180.1 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]$, ^{35}Cl), 182.1 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]$, ^{37}Cl). The four illustrative scans represent the spectral background of the measurement and the MS spectra at half peak height of peak front, at retention time and at half peak height of peak tail, correspondingly.

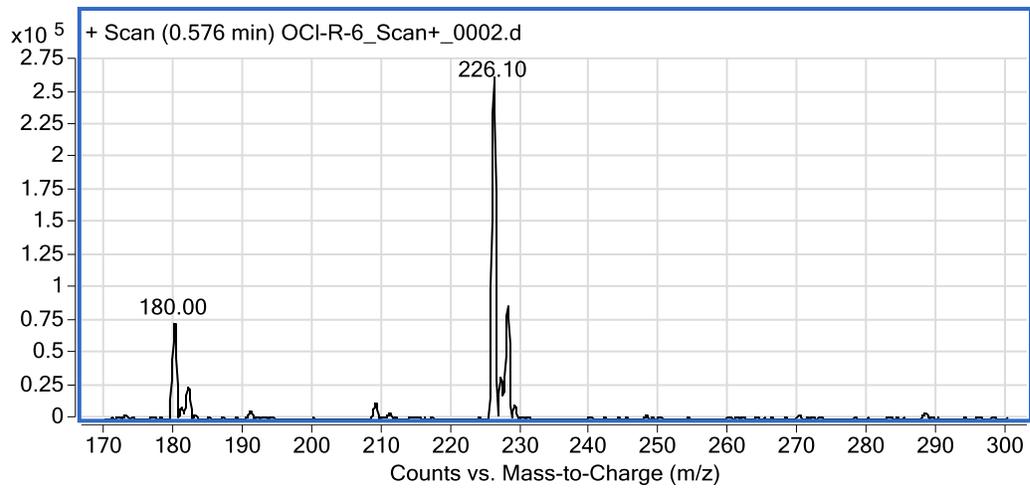
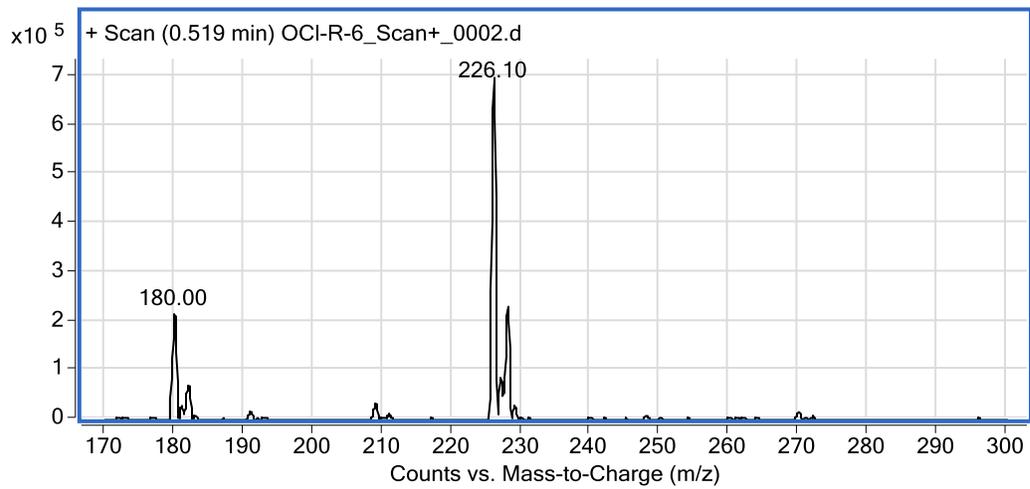
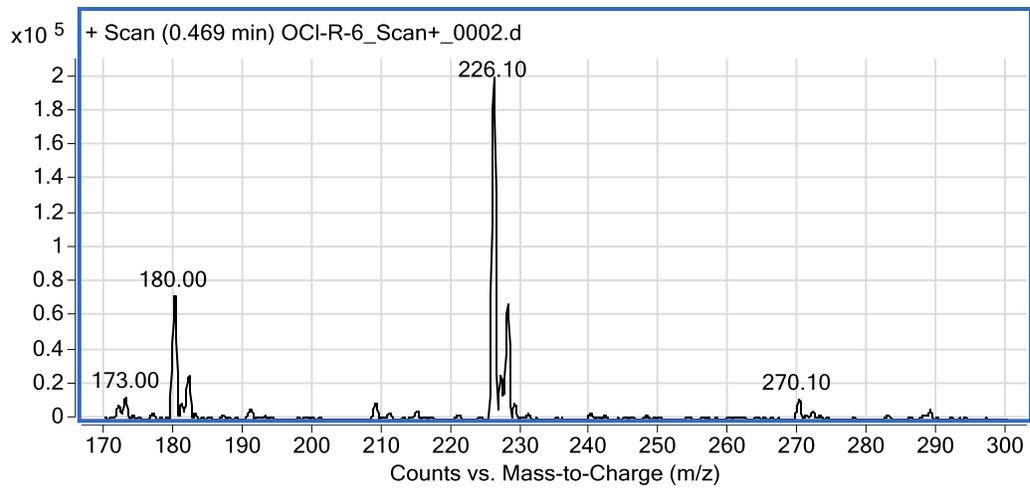




rac-(*E*)-2-Amino-5-(2-chlorophenyl)pent-4-enoic acid (*rac*-**1c**)

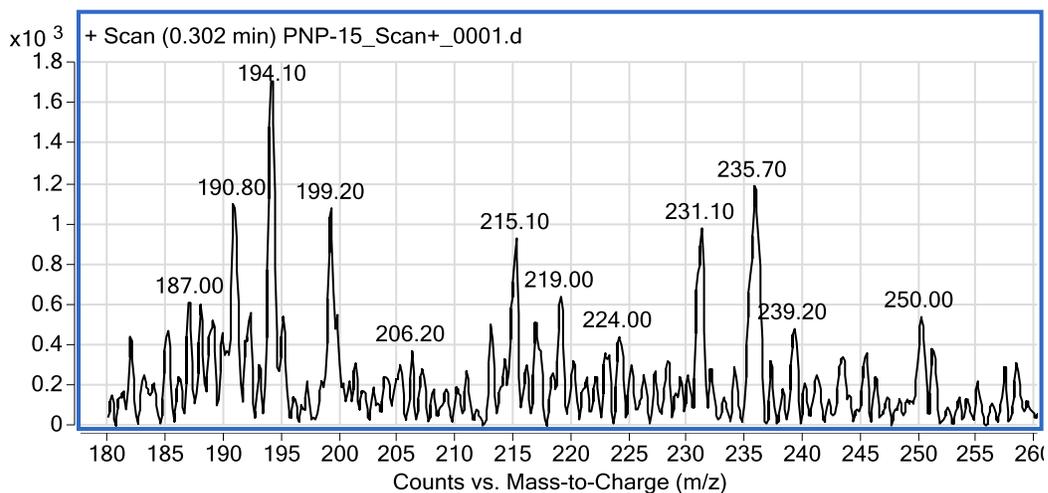
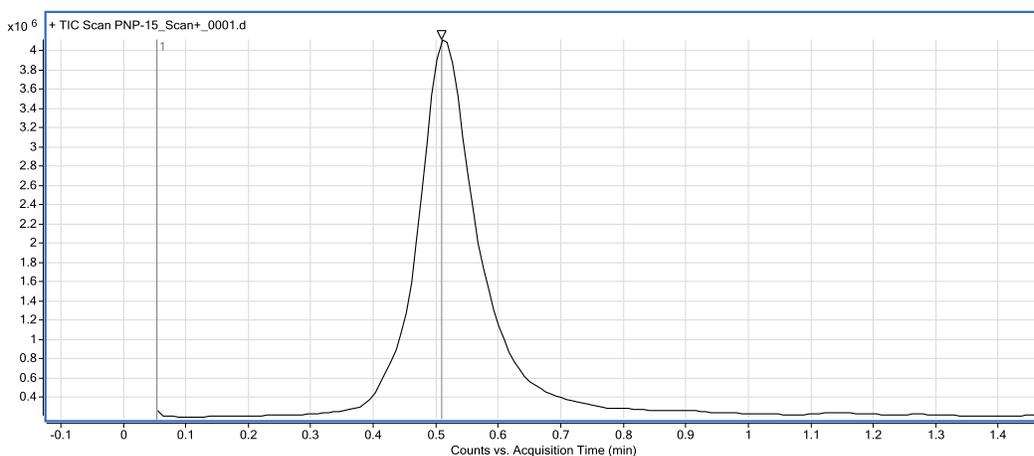
^1H NMR (600 MHz, D_2O) δ 7.55 (m, 1H), 7.34 (m, 1H), 7.15-7.23 (m, 2H), 6.79 (d, $J = 15.3$ Hz, 1H), 6.16 (m, 1H), 3.33 (m, 1H), 2.49 (m, 2H). ^{13}C NMR (151 MHz, D_2O): 182.5, 135.2, 131.9, 129.9, 129.5, 128.7, 128.6, 127.3, 126.9, 55.6, 38.5; LC-MS: positive ionization mode, m/z : 226.1 ($[\text{M}+\text{H}]^+$, ^{35}Cl), 228.1 ($[\text{M}+\text{H}]^+$, ^{37}Cl) (calculated for $\text{C}_{11}\text{H}_{12}\text{ClNO}_2$ 226.0557 ($[\text{M}+\text{H}]^+$, ^{35}Cl), 228.0527 ($[\text{M}+\text{H}]^+$, ^{37}Cl); 180.0 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]^+$, ^{35}Cl), 182.0 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]^+$, ^{37}Cl). The four illustrative scans represent the spectral background of the measurement and the MS spectra at half peak height of peak front, at retention time and at half peak height of peak tail, correspondingly.

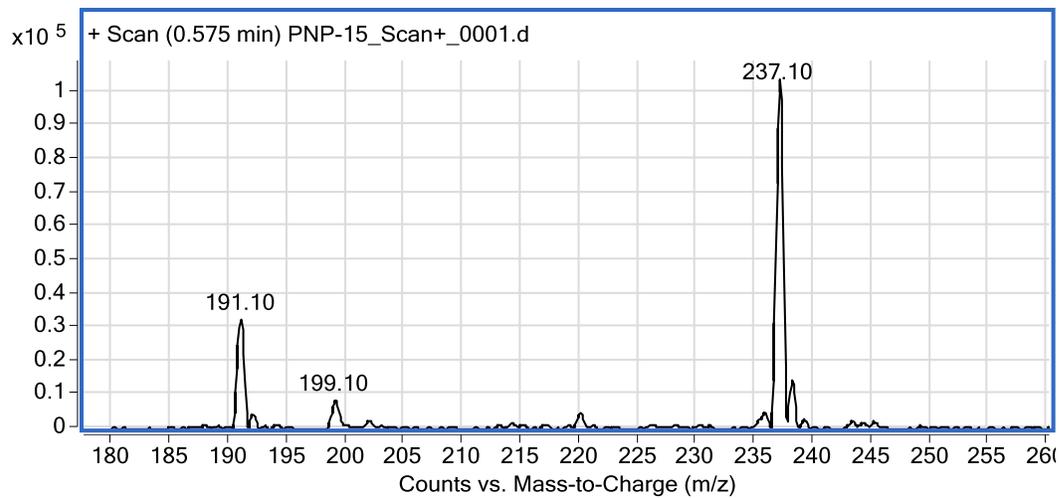
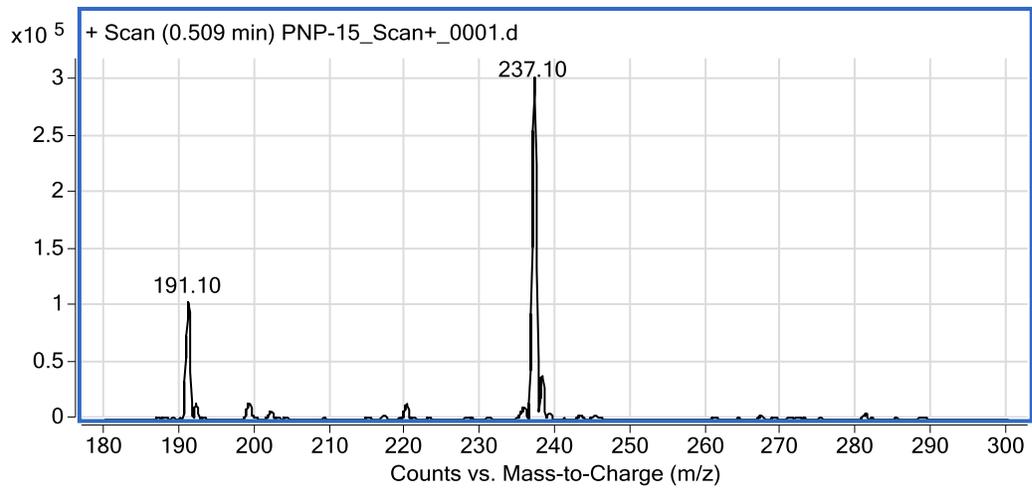
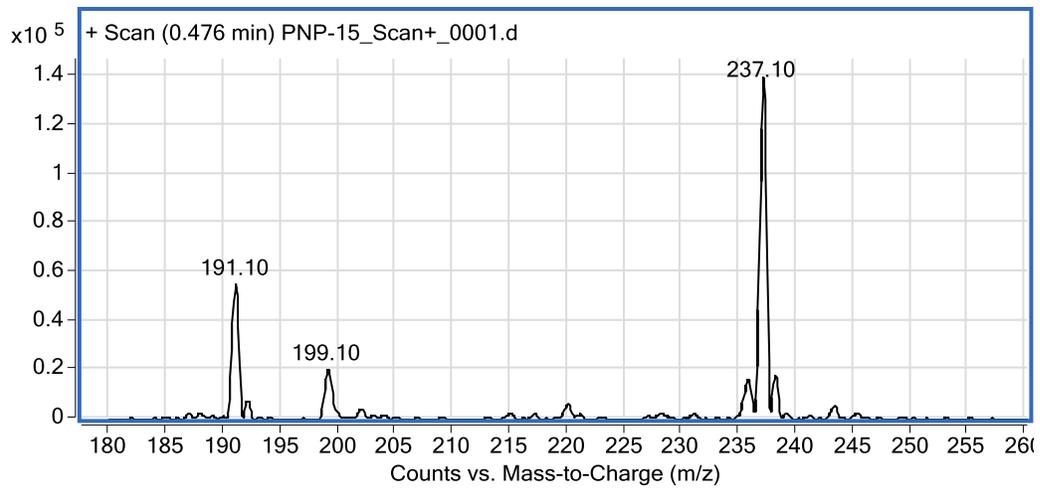




rac-(*E*)-2-Amino-5-(4-nitrophenyl)pent-4-enoic acid (*rac*-**1d**)

^1H NMR (400 MHz, D_2O) δ 8.08 (d, $J = 8.7$ Hz, 2H), 7.5 (d, $J = 8.8$ Hz, 2H), 6.54 (d, $J = 15.9$ Hz, 1H), 6.35-6.45 (m, 1H), 3.36 (t, $J = 6$ Hz, 1H), 2.43-2.57 (m, 2H). ^{13}C NMR (151 MHz, D_2O): 182.3, 145, 144.3, 131.8, 130.9, 126.7, 123.9, 55.5, 38.6; LC-MS: positive ionization mode, m/z : 236.1 ($[\text{M}+\text{H}]^+$) (calculated for $\text{C}_{11}\text{H}_{12}\text{N}_2\text{O}_4$ 236.0797 ($[\text{M}+\text{H}]^+$); 190.1 ($[\text{M}-\text{COOH}, -\text{H}, +\text{H}]$). The four illustrative scans represent the spectral background of the measurement and the MS spectra at half peak height of peak front, at retention time and at half peak height of peak tail, correspondingly.





4. Michaelis-Menten curves with non-linear fitting for calculation of the kinetic parameters (K_M , v_{max} , k_{cat})

The kinetic measurements were based on UV-spectroscopy by monitoring the production of the acrylic derivative **2a-d** at wavelengths where the corresponding amino acids *rac-1a-d* showed no absorption

Table S1. The wavelengths used for monitoring the production of acrylic derivatives **2a-d** and molar extinction coefficients (ϵ) at the corresponding wavelengths

Substrate	Product	λ (nm)	ϵ ($M^{-1}\cdot cm^{-1}$)
L-Phe	cinnamic acid	290	8810
L- 1a , <i>rac-1a</i>	2a	302	26900
<i>rac-1b</i>	2b	306	37500
<i>rac-1c</i>	2c	300	25800
<i>rac-1d</i>	2d	360	13000

The values of v_{max} and K_M were obtained from non-linear regression fitting of the Michaelis-Menten curves by MATLAB using the equation $v = v_{max}[S]/(K_M + [S])$. The k_{cat} values were calculated according to equation $k_{cat} = v_{max}/[E]$, where $[E]$ - represents the enzyme concentration used in the kinetic assay (μM) – Table S2.

Table S2. Enzyme concentrations used in the kinetic assays

Substrate	F137V (μM)	wt-PcPAL (μM)	F137A (μM)	F137G (μM)
L-Phe	0.3125	0.3125	0.3125	0.3125
L- 1a	0.3125	0.3125	0.3125	0.3125
1a	0.3125	0.625	-	-
1b	0.3125	9.375	-	-
1c	0.3125	2.5	-	-
1d	2.5	2.5	-	-

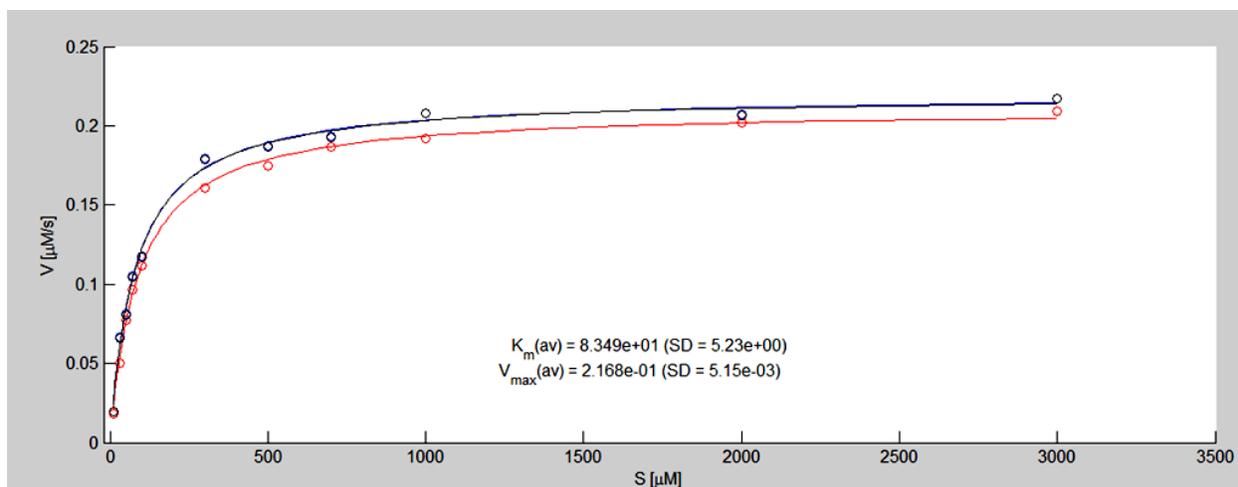


Figure S1. Non-linear fitting curves for the ammonia elimination from L-Phe catalyzed by wt-PcPAL; measured in triplicate.

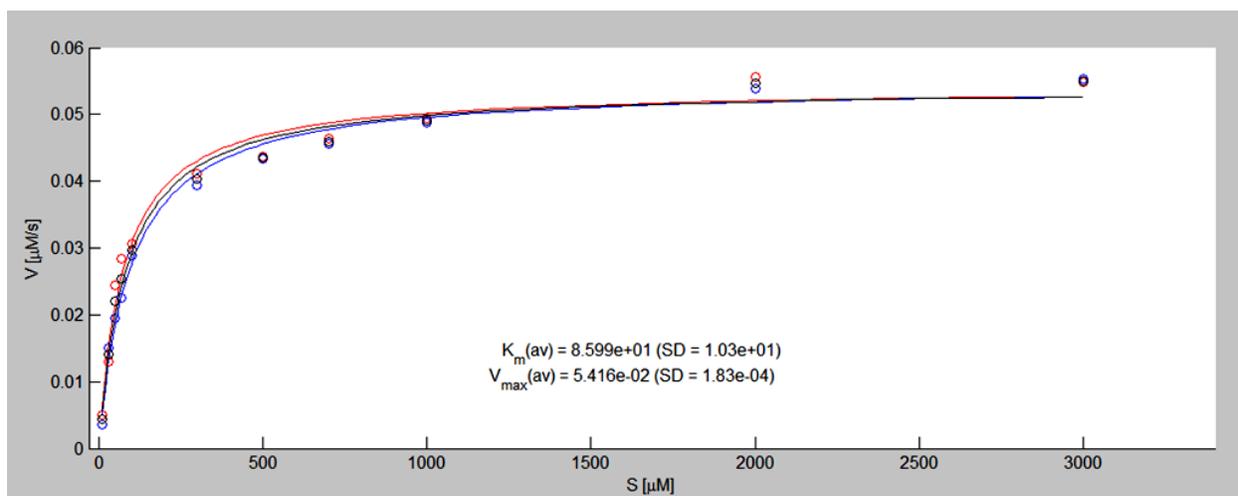


Figure S2. Non-linear fitting curves for the ammonia elimination from L-Phe catalyzed by F137V-PcPAL; measured in triplicate.

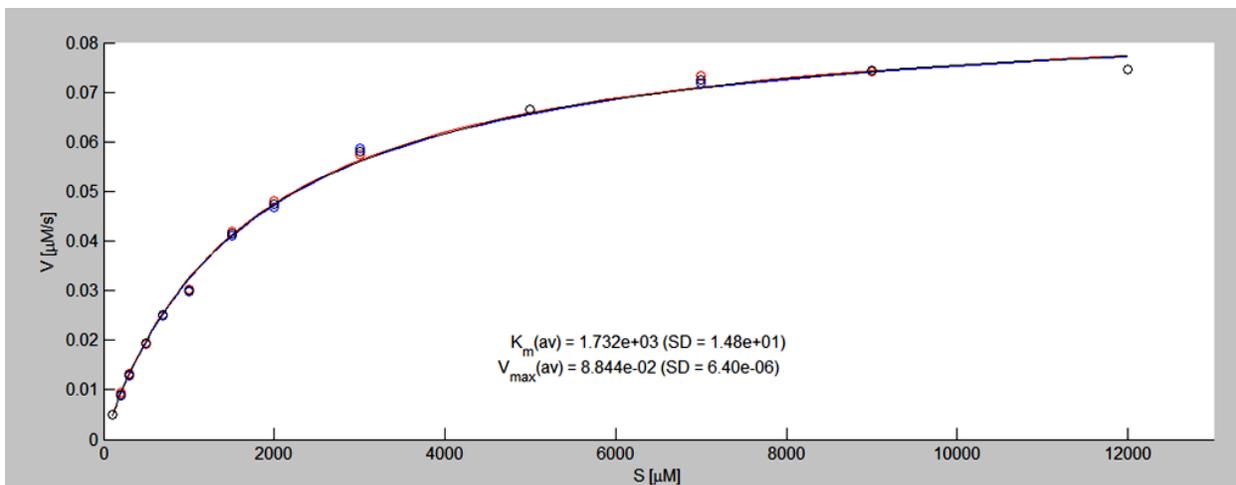


Figure S3. Non-linear fitting curves for the ammonia elimination from L-Phe catalyzed by F137A-PcPAL; measured in triplicate.

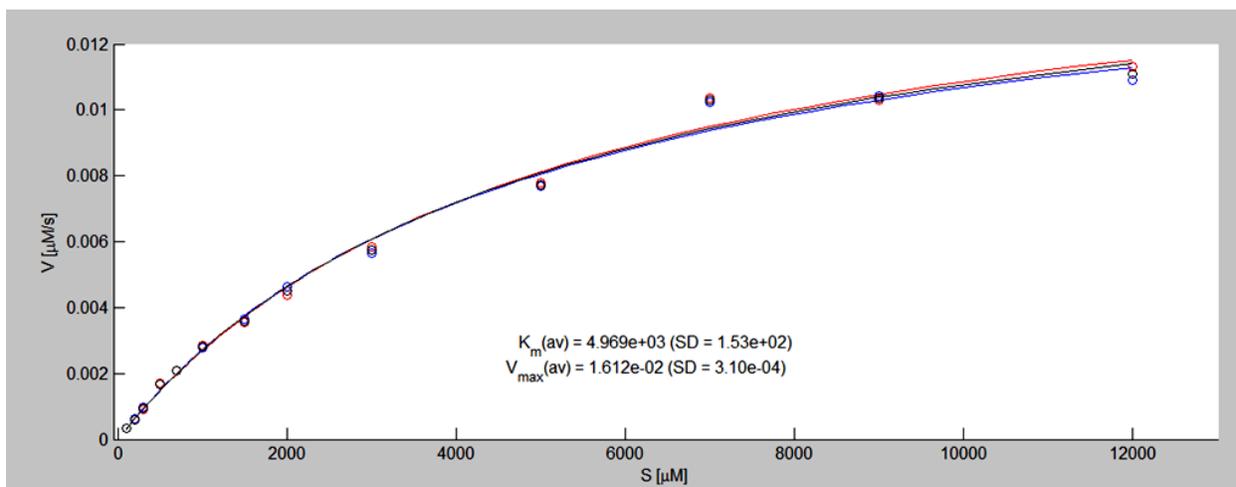


Figure S4. Non-linear fitting curves for the ammonia elimination from L-Phe catalyzed by F137G-PcPAL; measured in triplicate.

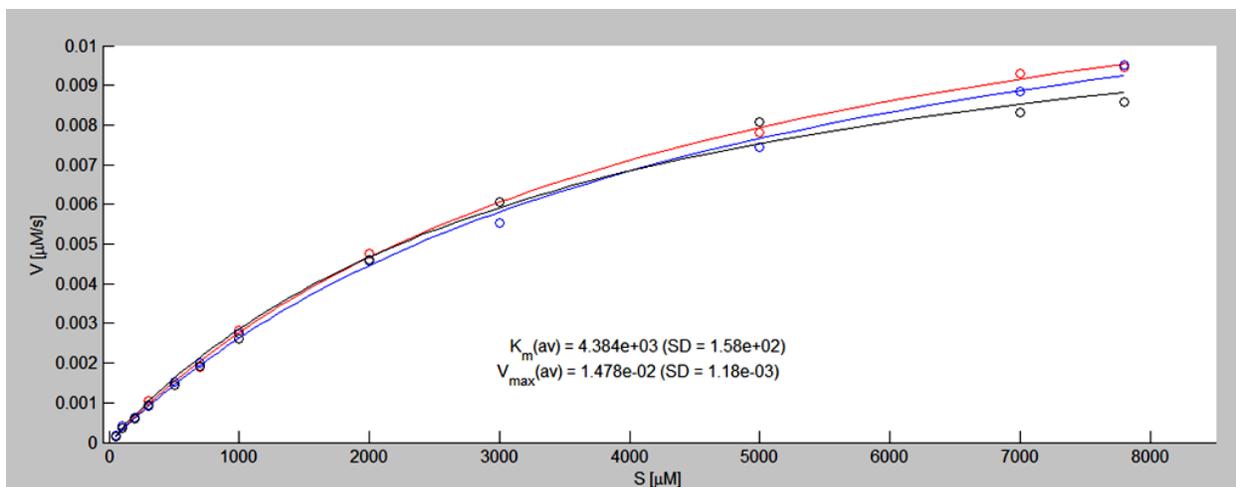


Figure S5. Non-linear fitting curves for the ammonia elimination from L-1a catalyzed by wt-PcPAL; measured in triplicate.

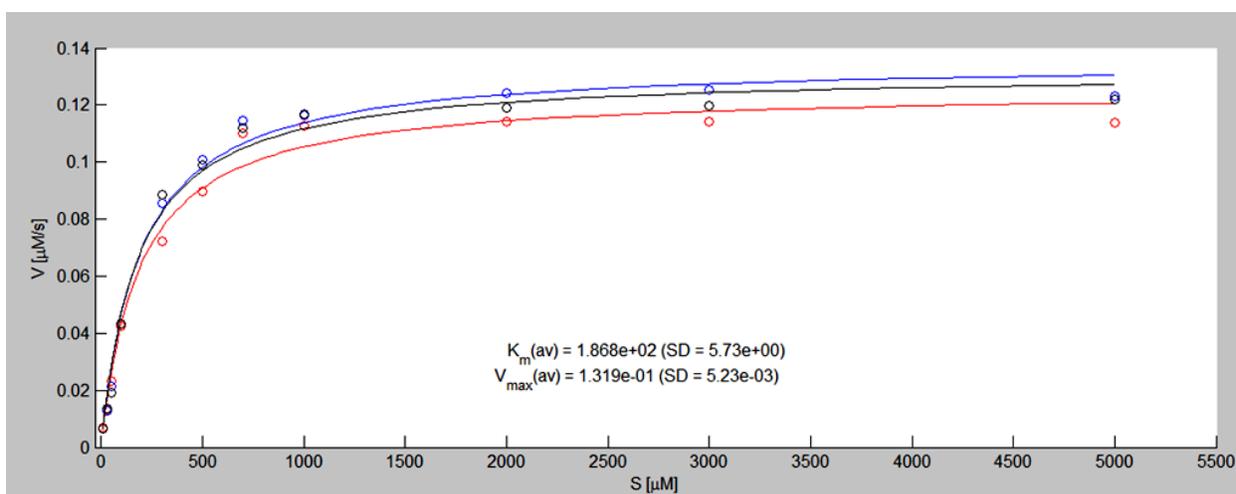


Figure S6. Non-linear fitting curves for the ammonia elimination from L-1a catalyzed by F137V-PcPAL, measured in triplicate.

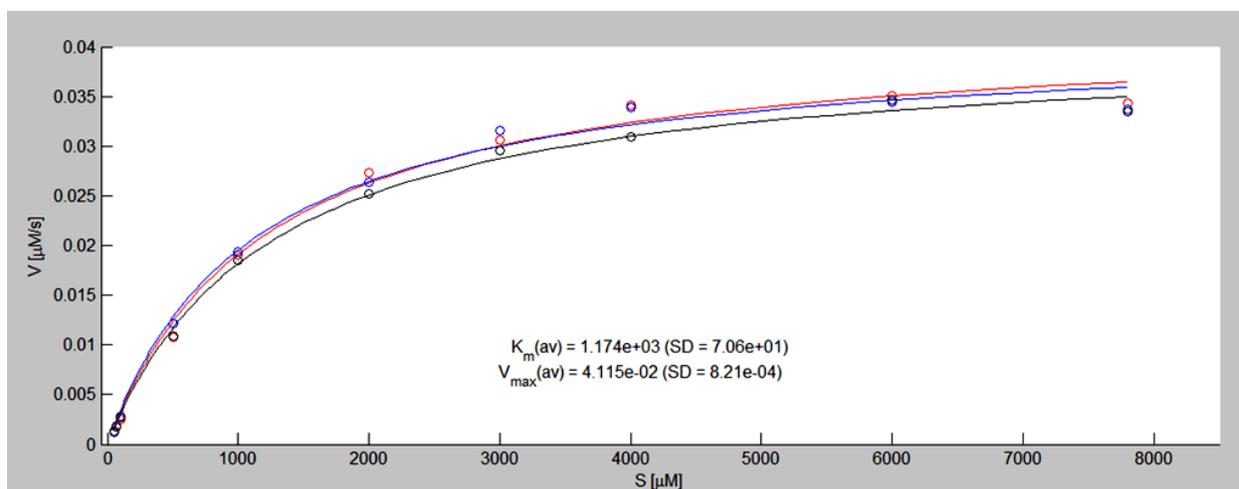


Figure S7. Non-linear fitting curves for the ammonia elimination from L-1a catalyzed by F137A-PcPAL; measured in triplicate.

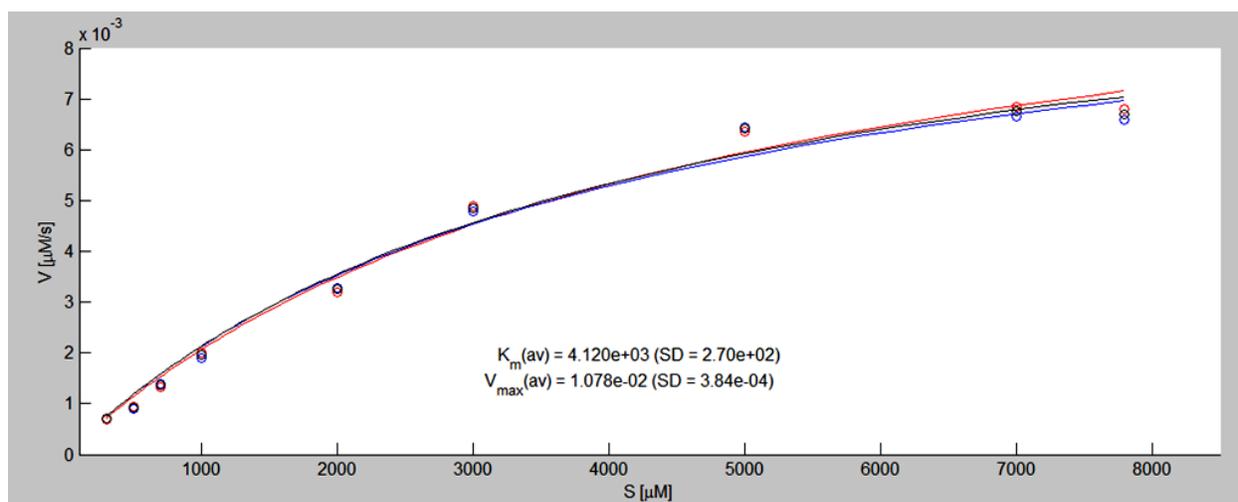


Figure S8. Non-linear fitting curves for the ammonia elimination from L-1a catalyzed by F137G-PcPAL; measured in triplicate.

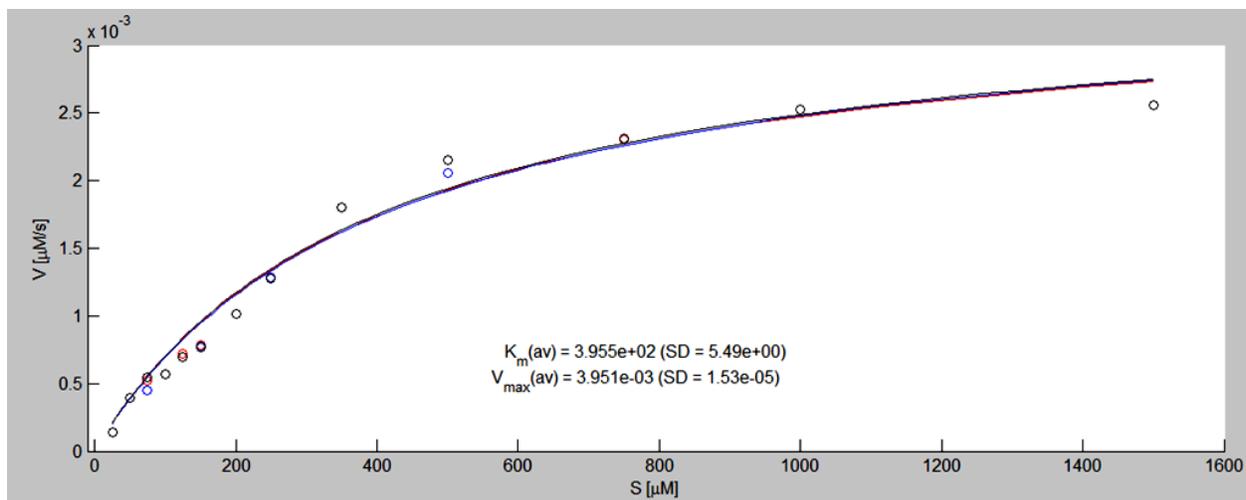


Figure S9. Non-linear fitting curves for the ammonia elimination from *rac-1a* catalyzed by wt-PcPAL; measured in triplicate.

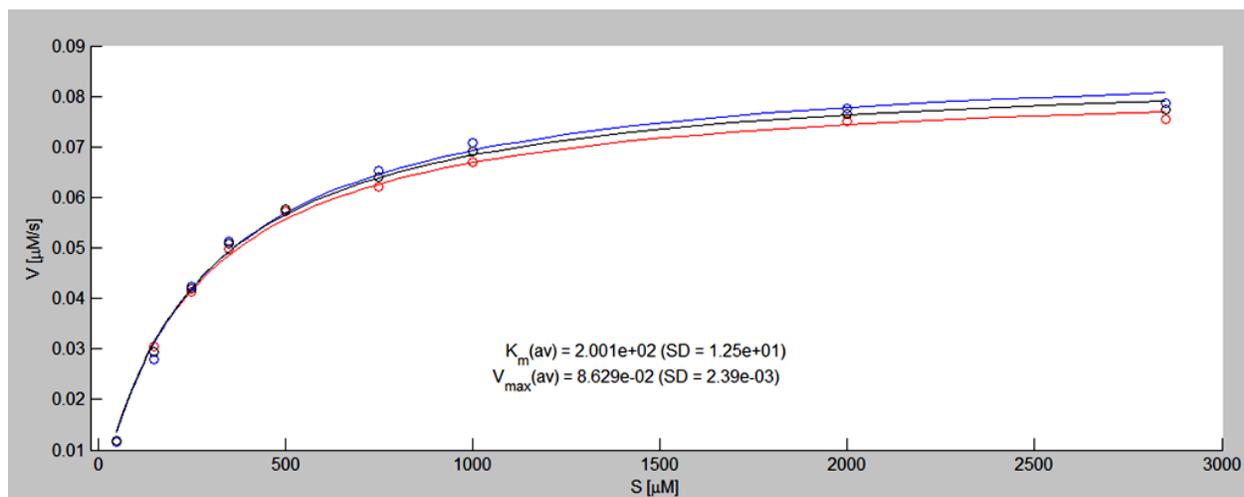


Figure S10. Non-linear fitting curves for the ammonia elimination from *rac-1a* catalyzed by F137V-PcPAL; measured in triplicate.

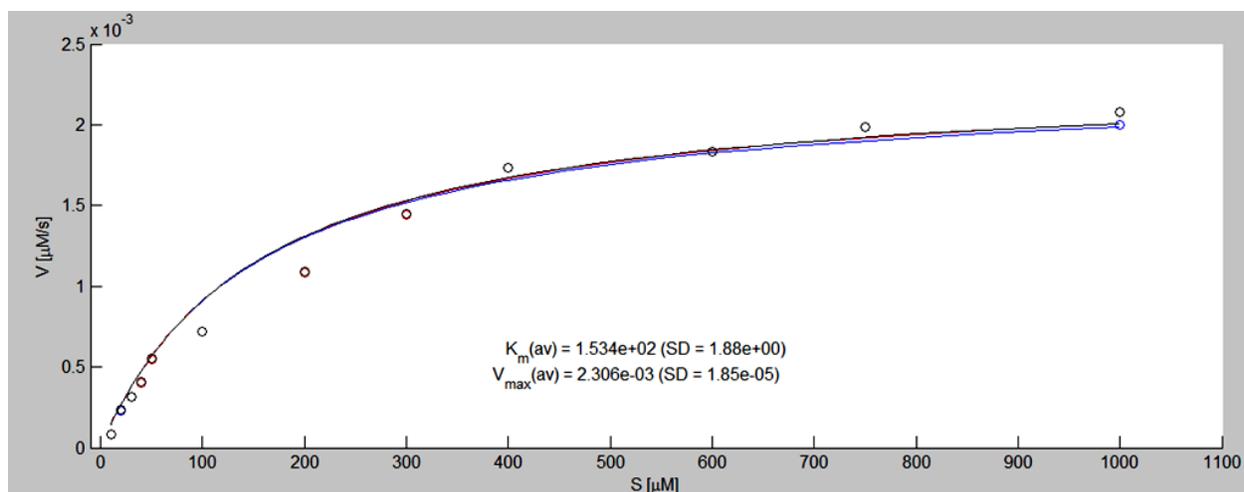


Figure S11. Non-linear fitting curves for the ammonia elimination from *rac-1b* catalyzed by wt-PcPAL; measured in triplicate.

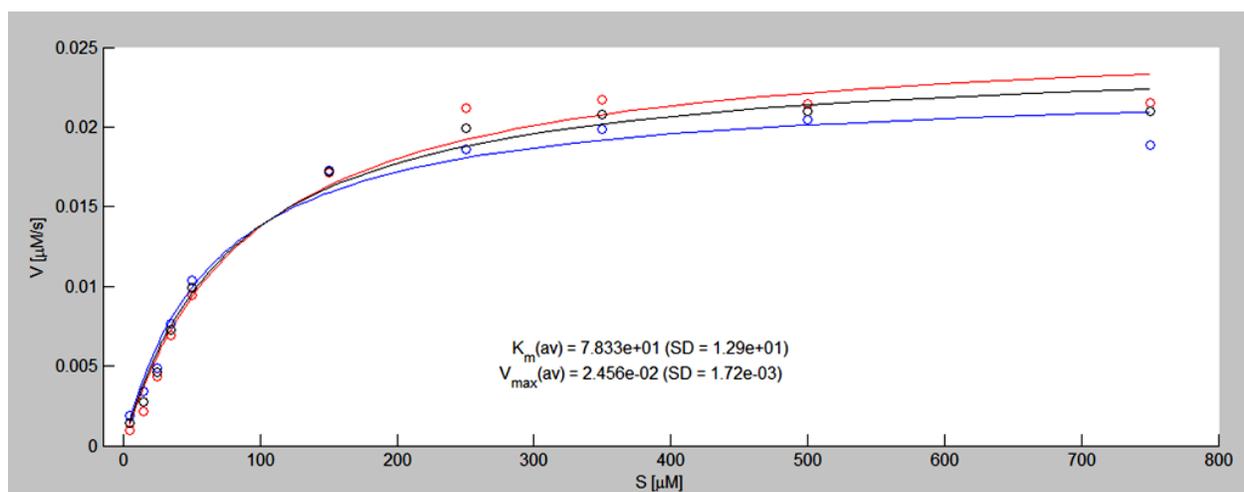


Figure S12. Non-linear fitting curves for the ammonia elimination from *rac-1b* catalyzed by F137V-PcPAL; measured in triplicate.

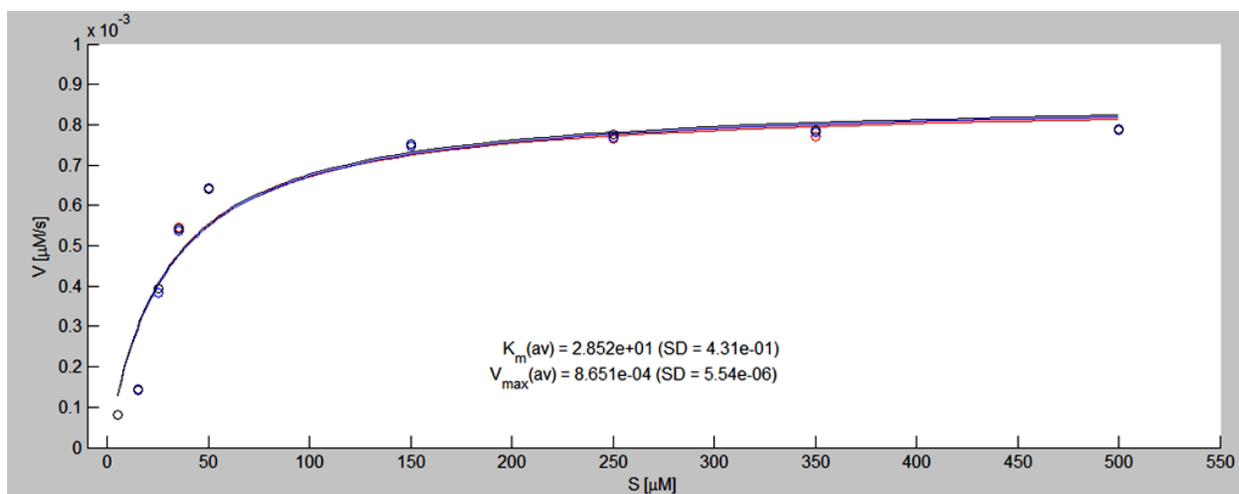


Figure S13. Non-linear fitting curves for the ammonia elimination from *rac-1c* catalyzed by wt-PcPAL; measured in triplicate.

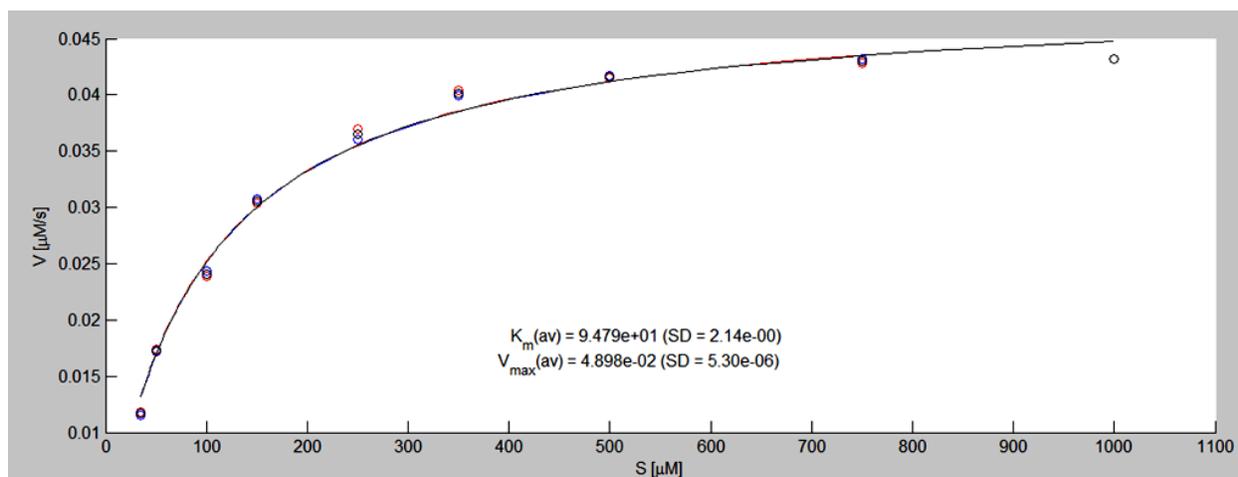


Figure S14. Non-linear fitting curves for the ammonia elimination from *rac-1c* catalyzed by F137V-PcPAL; measured in triplicate.

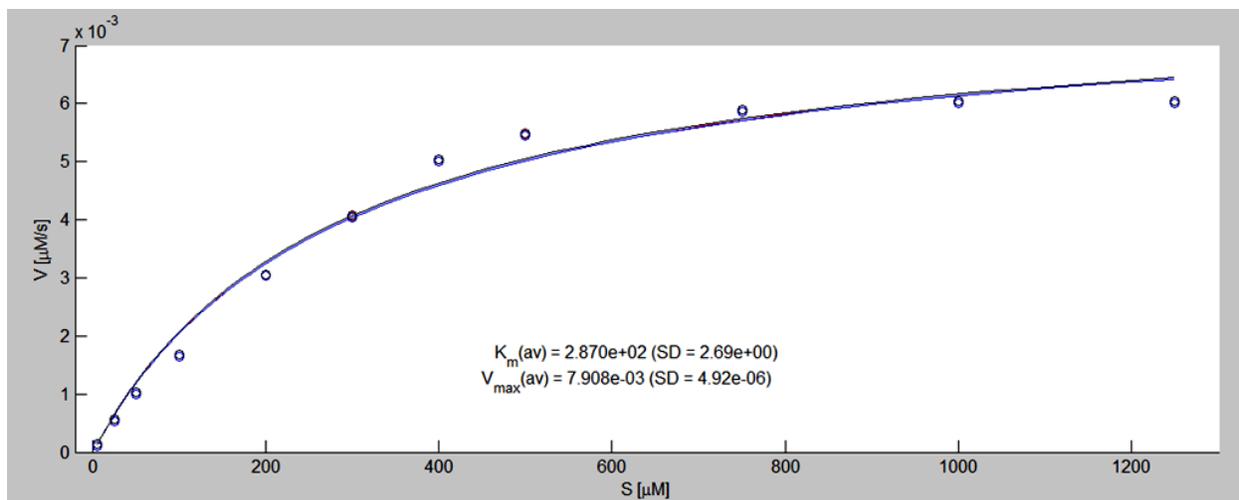


Figure S15. Non-linear fitting curves for the ammonia elimination from *rac-1d* catalyzed by wt-PcPAL; measured in triplicate

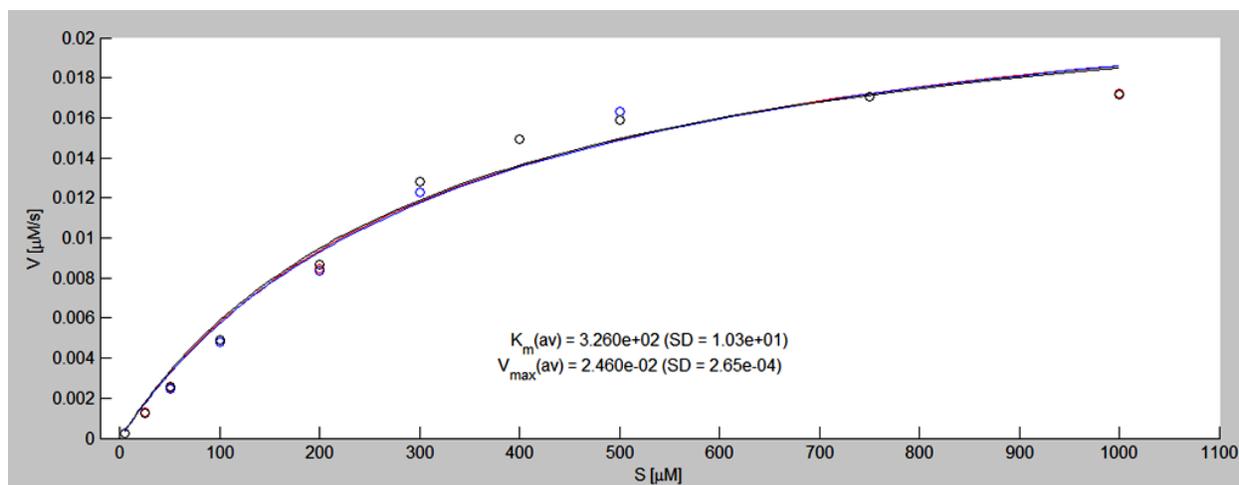


Figure S16. Non-linear fitting curves for the ammonia elimination from *rac-1d* catalyzed by F137V-PcPAL measured in triplicate.

5. HPLC monitoring of the enzymatic reactions

5.1. Determination of the conversion by HPLC

In order to determine the conversion in the PcPAL-catalyzed enzymatic transformations, the relative response factor of the acrylic derivatives **2a-d** compared to the amino acids *rac-1a-d* was determined by injecting the mixture of known composition of the corresponding amino acid *rac-1a-d* and the corresponding acrylic derivative **2a-d** onto an Agilent Zorbax Eclipse XDB-C8 column (150 × 4.6 mm; 5 μm).

Table S3. HPLC methods and response factors, used for the conversion value determinations

Compound	Eluent*[% B]	Retention time (min)		Response factor
		<i>rac-1</i>	2	2 vs. <i>rac-1</i>
<i>rac-1a</i> , 2a	15 to 50 in 15 min	9.5	13.5	2.63
<i>rac-1b</i> , 2b	25 to 60 in 12 min	12.0	15.1	1.41
<i>rac-1c</i> , 2c	15 to 50 in 12 min	10.6	13.3	1.54
<i>rac-1d</i> , 2d	20 to 60 in 12 min	9.6	12.2	0.20

* Mobile phase: A: NH₄OH buffer (0.1 M, pH 9.0) / B: MeOH; flow rate: 1.0 mL min⁻¹, measurements performed at 20 °C, wavelength used for UV detection: 260 nm.

5.2 HPLC chromatograms of the PcPAL-catalyzed ammonia elimination reaction from *rac-1a-d* at the end-point of monitoring

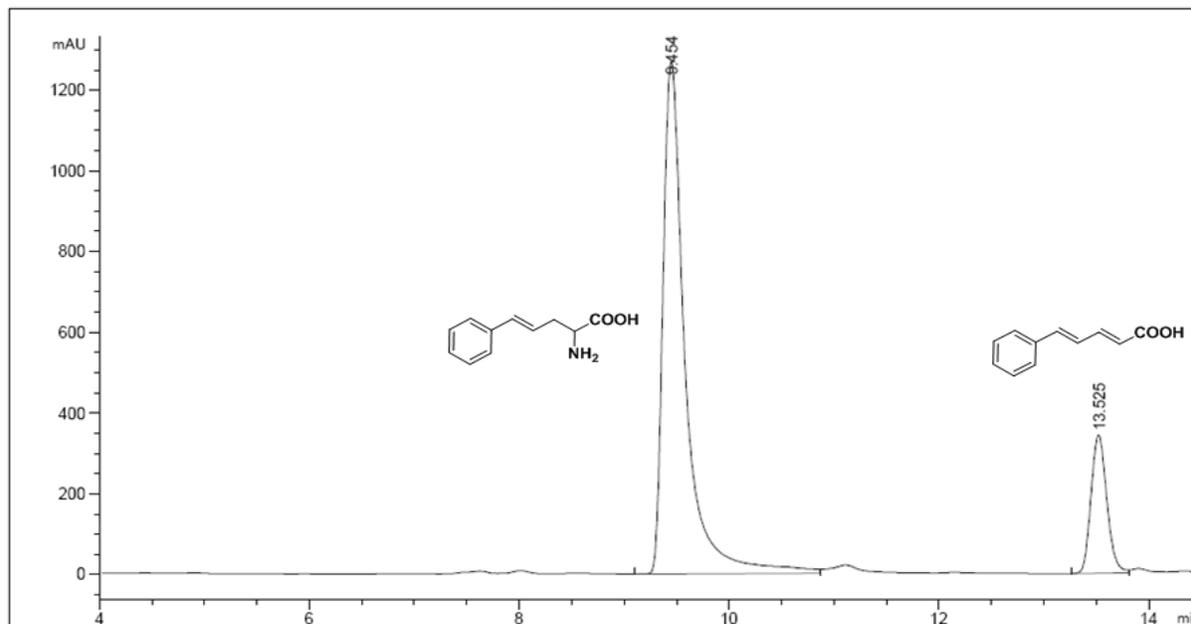


Figure S17. HPLC chromatogram of the ammonia elimination from *rac-1a* catalyzed by wt-PcPAL (after 274 h reaction time; conversion value: 50%).

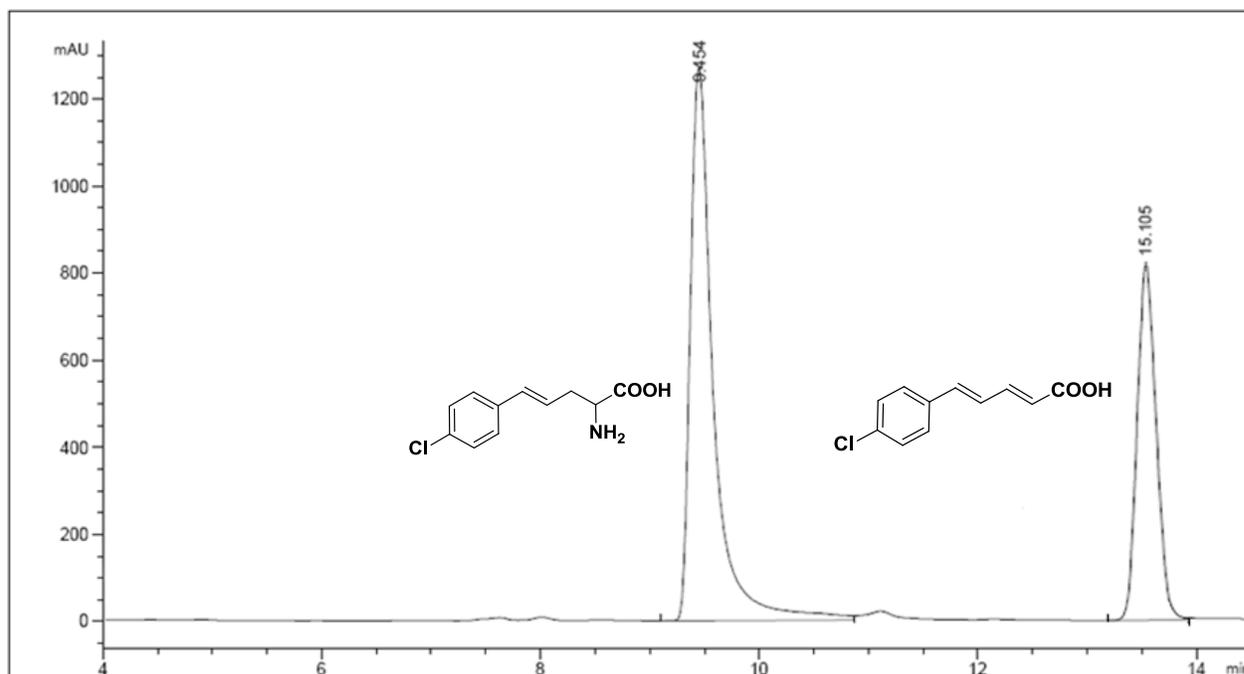


Figure S18. HPLC chromatogram of the ammonia elimination from *rac-1b* catalyzed by wt-PcPAL (after 504 h reaction time; conversion value: 29%).

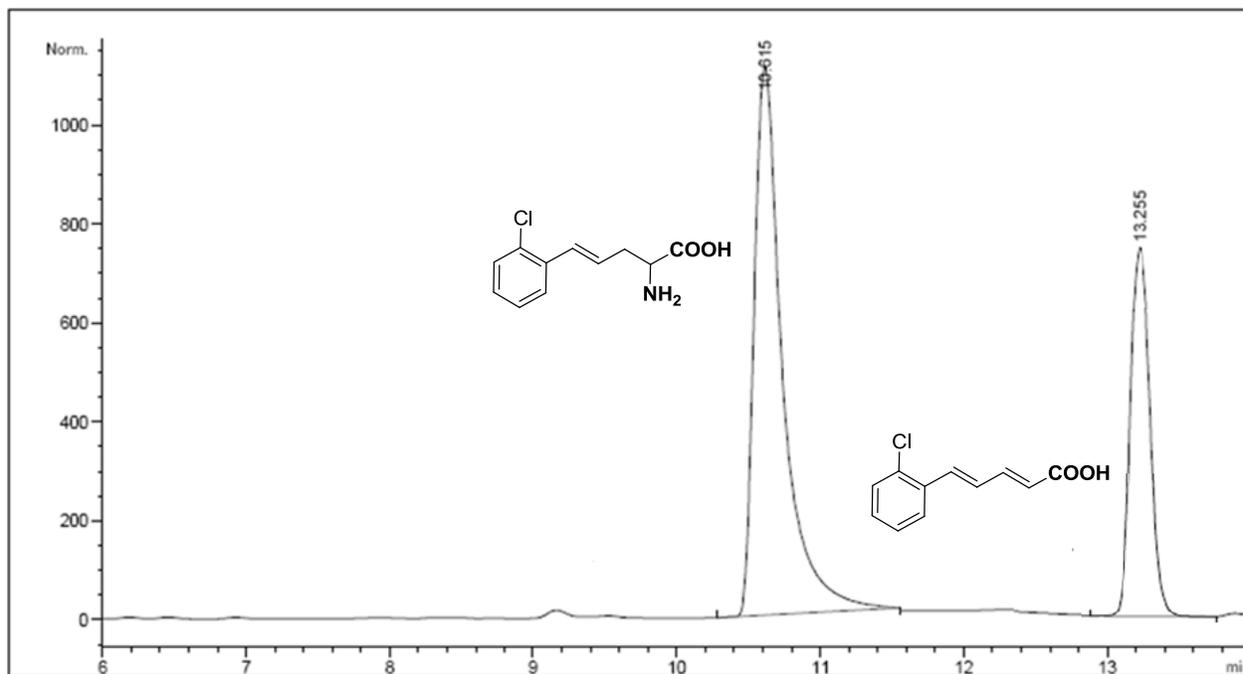


Figure S19. HPLC chromatogram of the ammonia elimination from *rac-1c* catalyzed by wt-PcPAL (after 600 h reaction time; conversion value: 36%).

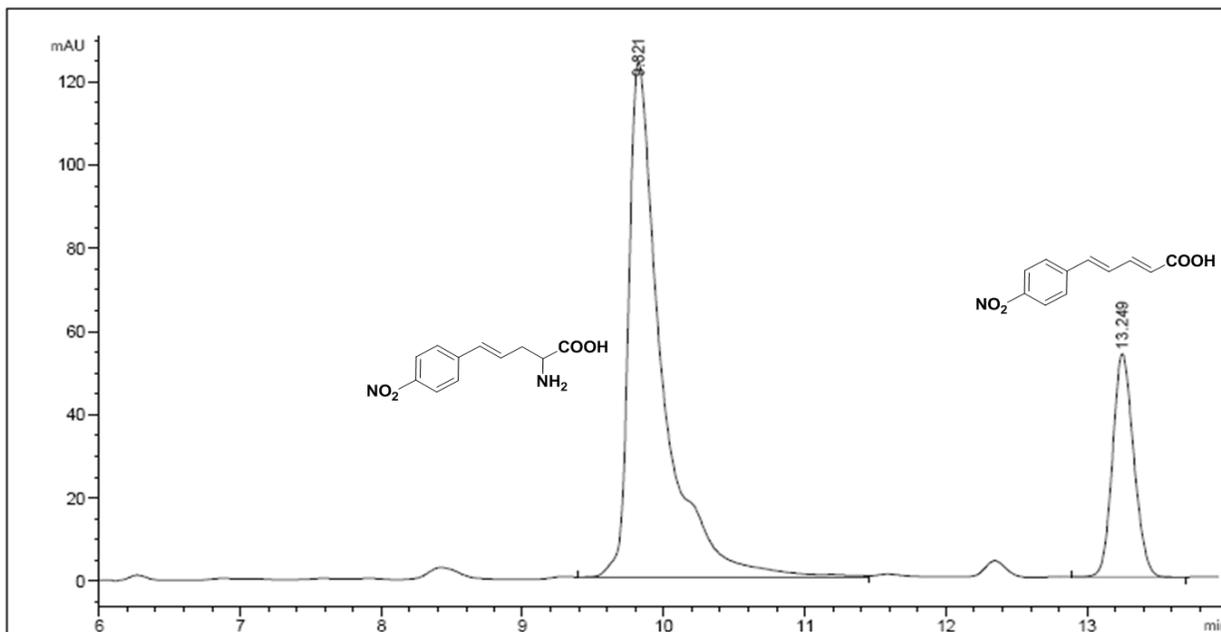


Figure S20. HPLC chromatogram of the ammonia elimination from *rac-1d* catalyzed by wt-PcPAL (after 600 h reaction time; conversion value: 31%).

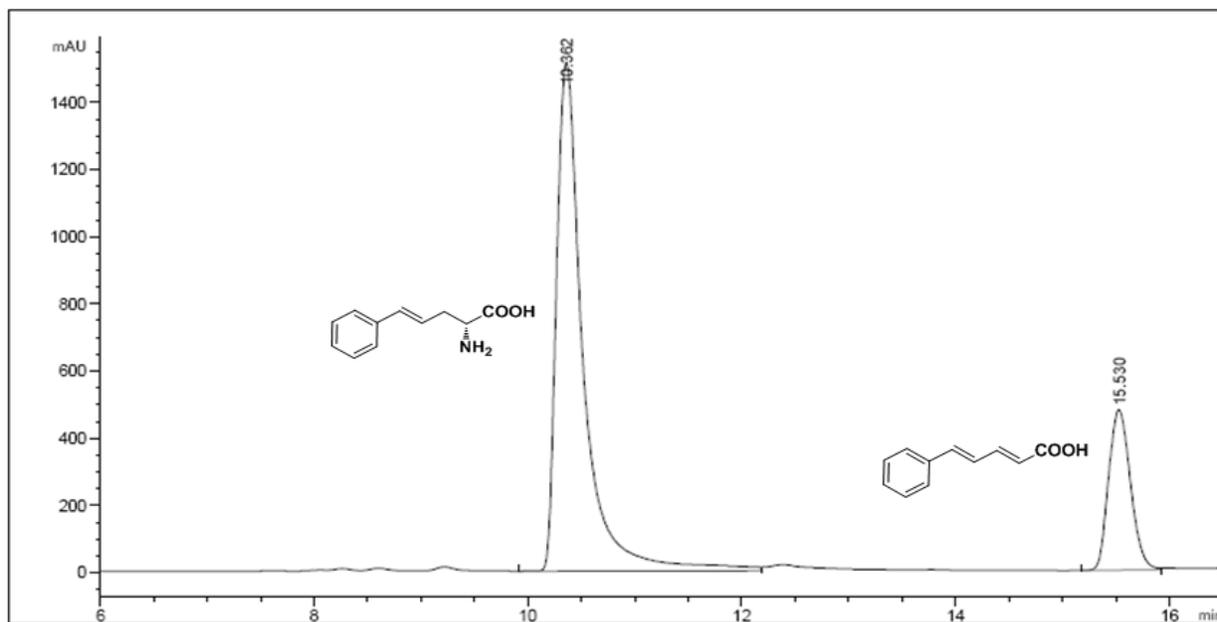


Figure S21. HPLC chromatogram of the ammonia elimination from *rac-1a* catalyzed by F137V-PcPAL (after 24 h reaction time; conversion value: 50%).

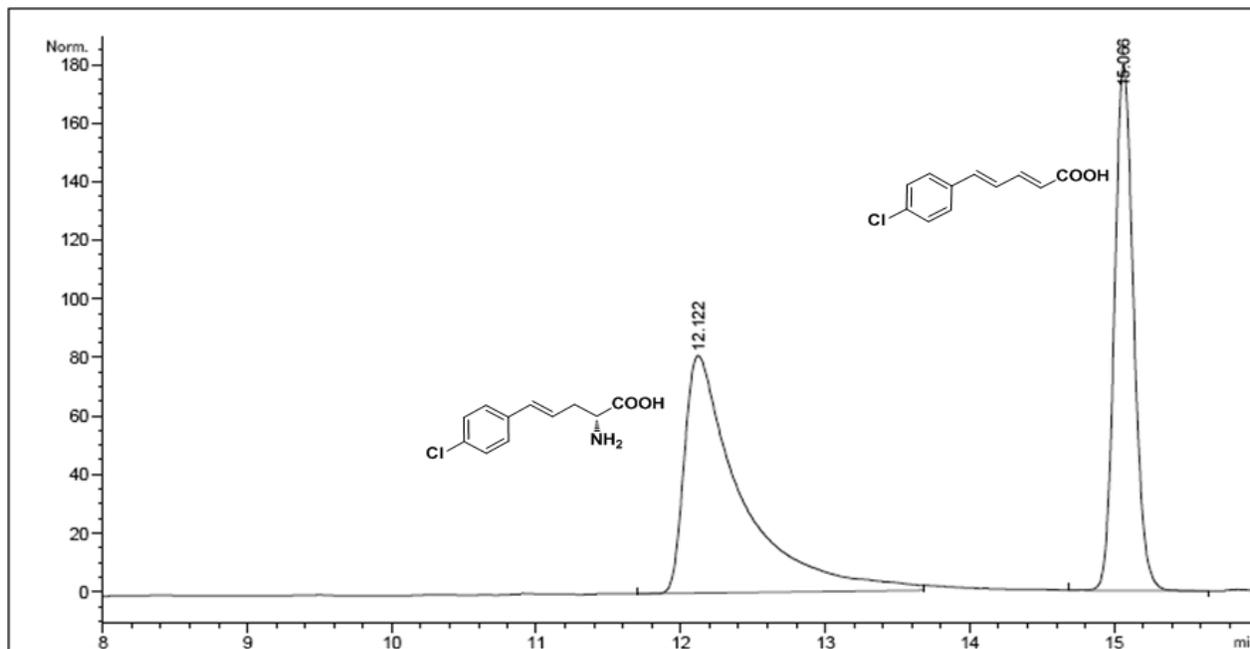


Figure S22. HPLC chromatogram of the ammonia elimination from *rac-1b* catalyzed by F137V-PcPAL (after 274 h reaction time; conversion value: 50%).

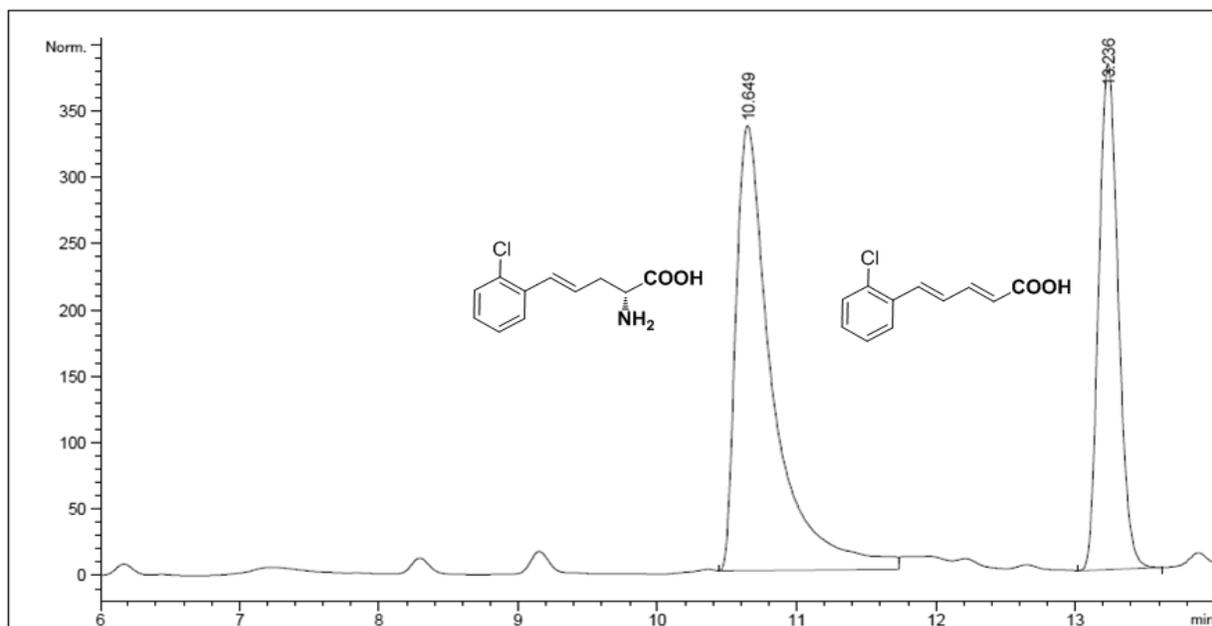


Figure S23. HPLC chromatogram of the ammonia elimination from *rac*-1c catalyzed by F137V-PcPAL (after 134 h reaction time; conversion value: 50%).

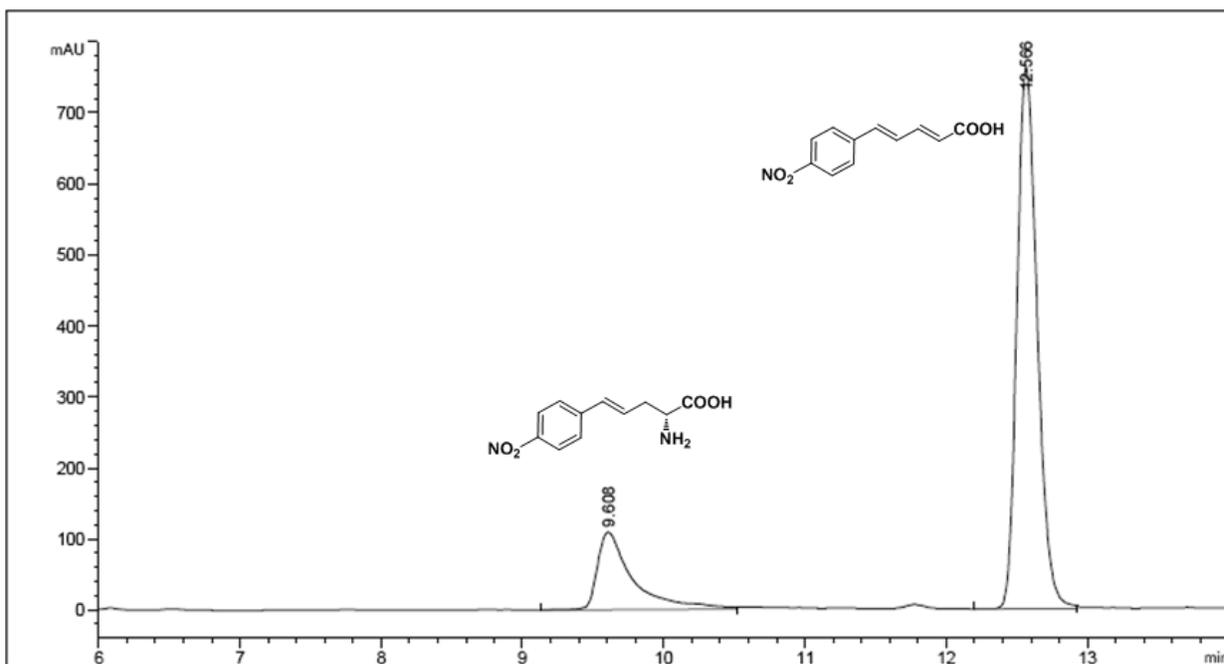


Figure S24. HPLC chromatogram of the ammonia elimination from *rac*-1d catalyzed by F137V-PcPAL (after 300 h reaction time; conversion value: 50%).

5.3. HPLC methods to determine the enantiomeric excess (ee) of D-**1a-d**

The chiral HPLC separation of the *rac*-**1a-d** was developed using Chiralpak ZWIX (+) chiral column and MeOH (6% CH₃COOH + 50 mM DEA): MeCN: H₂O 49:49:2 (v/v/v) as mobile phase, at a flow rate of 1 mL min⁻¹.

Table S4. Retention times of the enantiomers of *rac*-**1a-d** obtained on Chiralpak ZWIX (+) chiral column

Compound	Retention time (min)	
	<i>t</i> _{rL}	<i>t</i> _{rD}
<i>rac</i> - 1a	6.34	7.51
<i>rac</i> - 1b	7.36	8.90
<i>rac</i> - 1c	6.71	8.27
<i>rac</i> - 1d	7.97	10.13

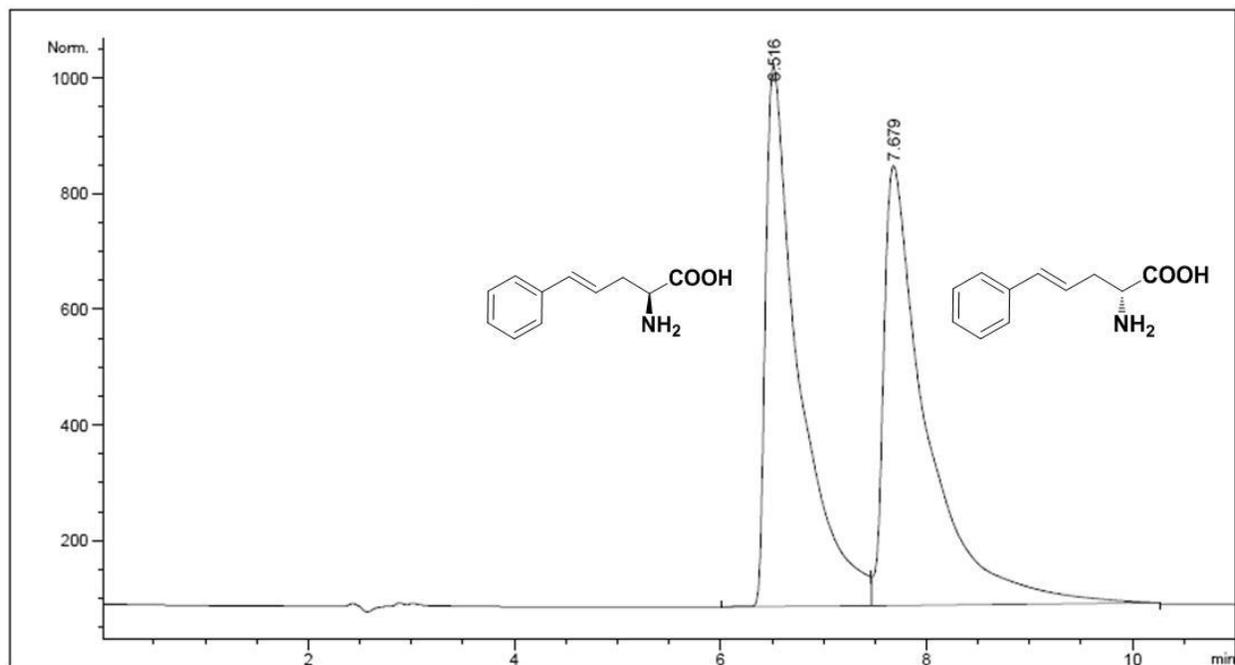


Figure S25. HPLC chromatogram of *rac*-**1a** obtained on Chiralpak ZWIX (+) chiral column

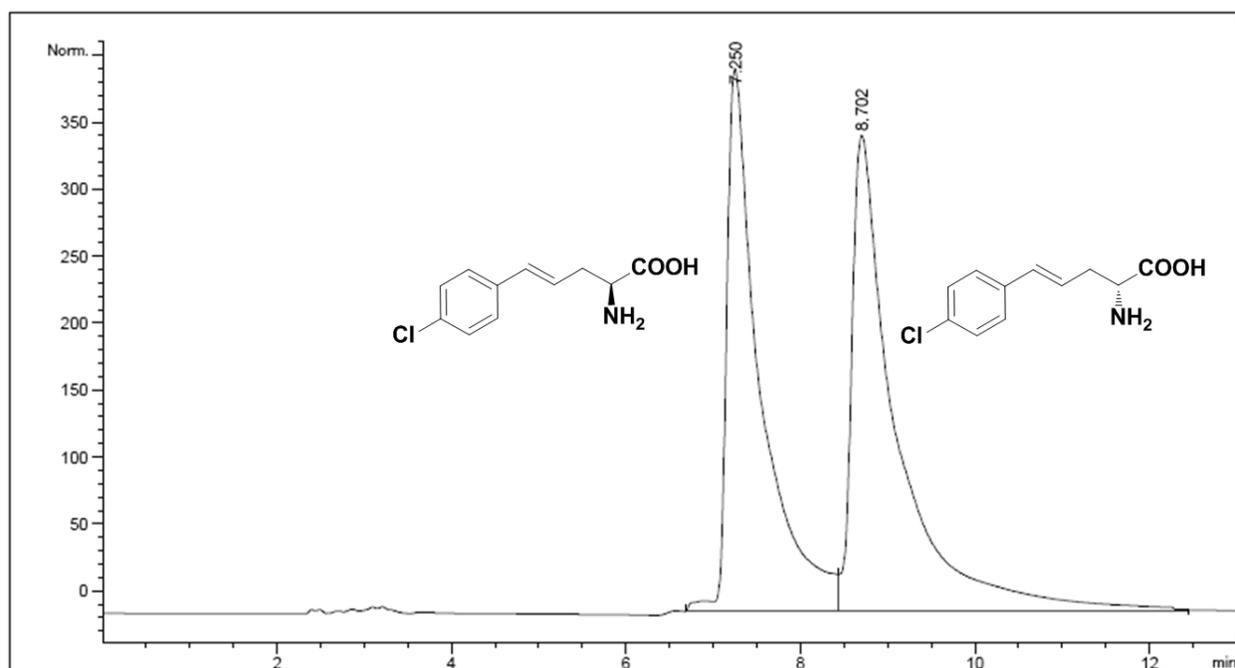


Figure S26. HPLC chromatogram of *rac-1b* obtained on Chiralpak ZwiX (+) chiral column

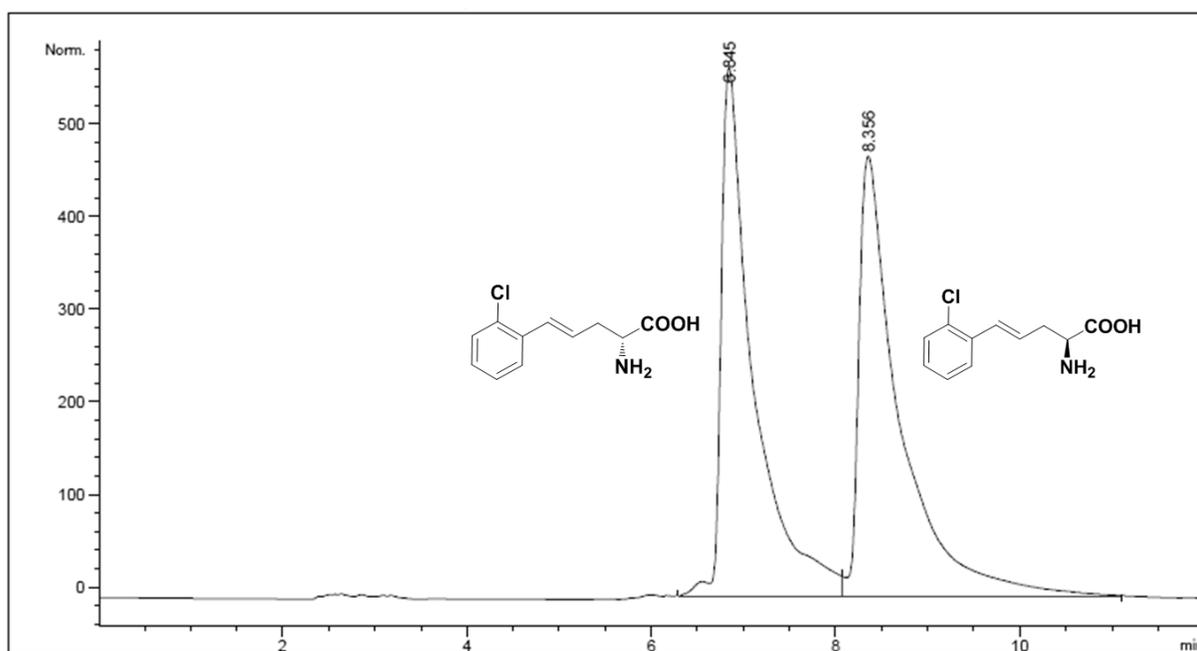


Figure S27. HPLC chromatogram of *rac-1c* obtained on Chiralpak ZwiX (+) chiral column

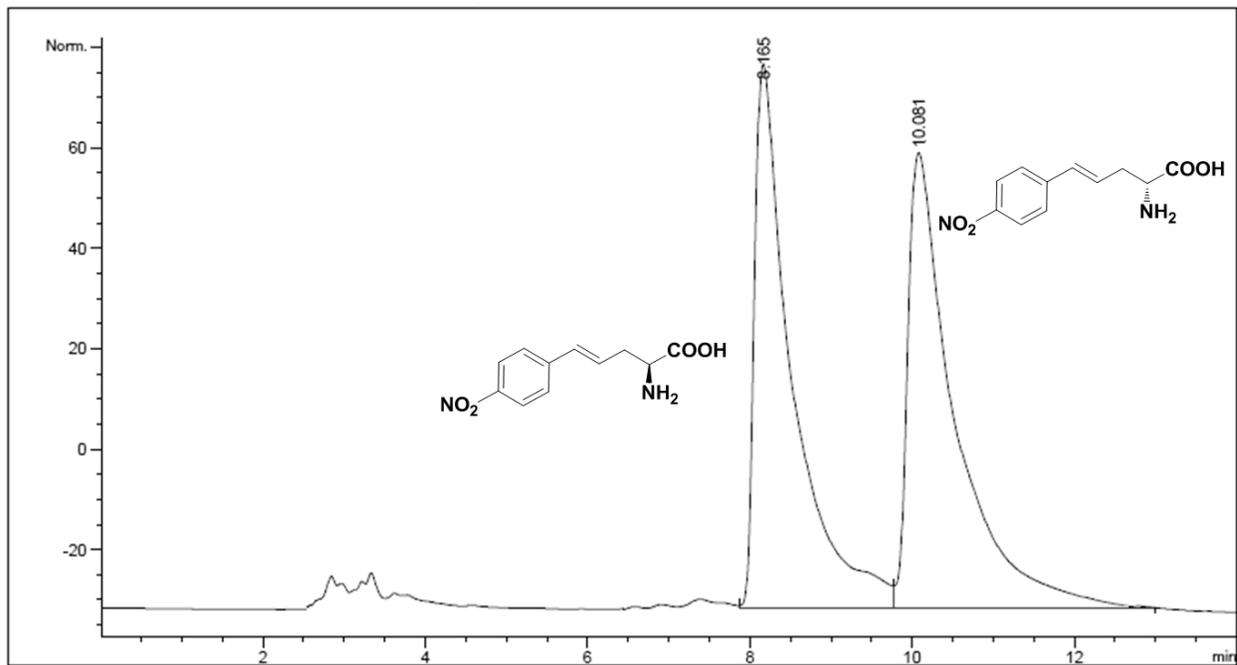


Figure S28. HPLC chromatogram of *rac*-1d obtained on Chiralpak ZWIX (+) chiral column

5.4 Chiral HPLC analysis of the final products from wt-PcPAL-catalyzed kinetic resolutions of *rac-1a-d*

Samples were taken from the enzymatic reactions after different time intervals as described in the main manuscript (Experimental part; ammonia elimination reactions from *rac-1a-d*) followed by the HPLC analysis using the chiral separation method indicated above.

The peaks appearing at $R_t = 2.9-3.2$ min, represent the elution front, which contain the produced acrylic derivatives **2a-d**.

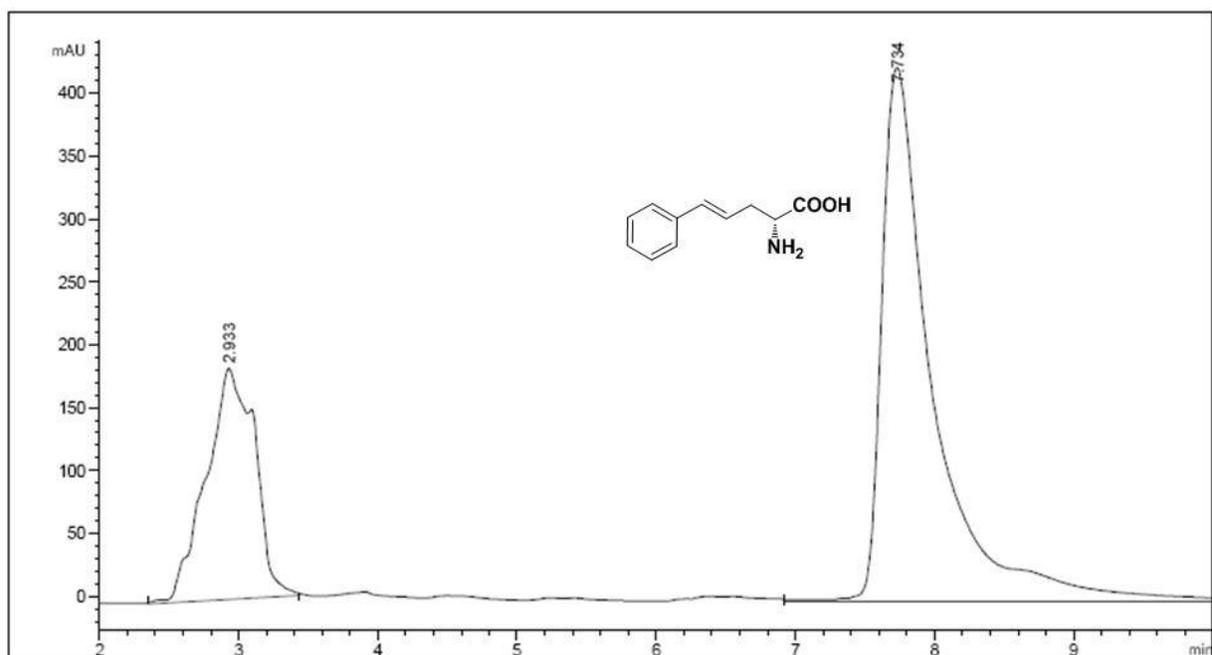


Figure S29 .HPLC analysis of the ammonia elimination from *rac-1a* catalyzed by wt-PcPAL on Chiralpak ZWIX (+) chiral column (after 274 h reaction time; ee >98%).

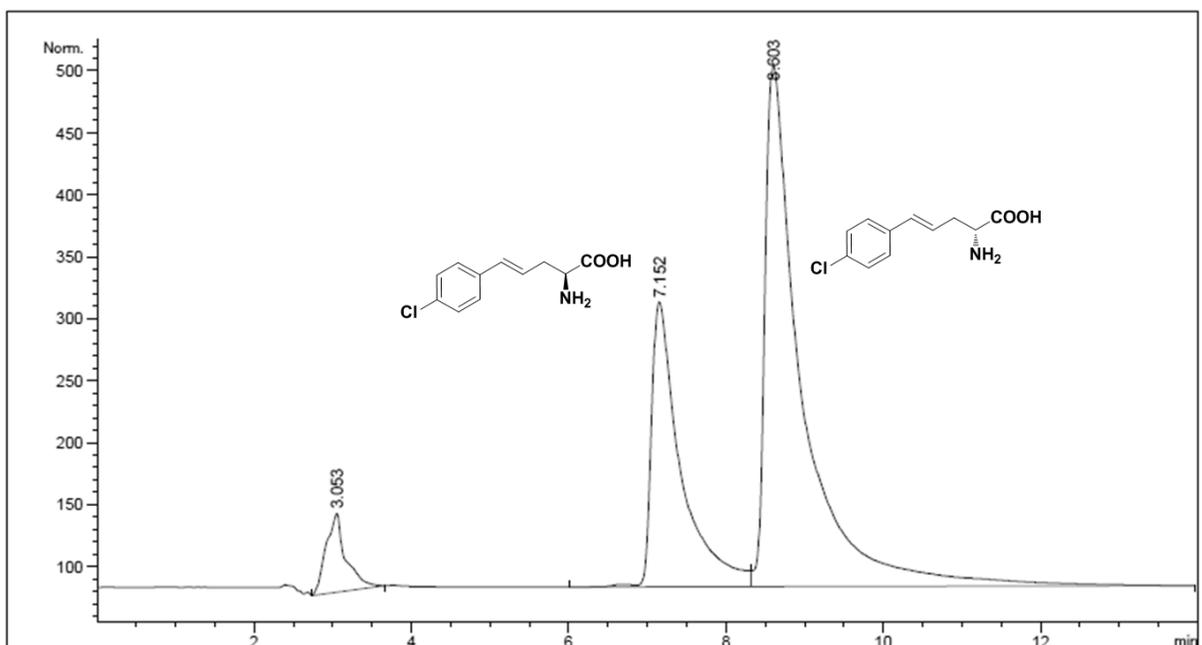


Figure S30. HPLC analysis of the ammonia elimination from *rac-1b* catalyzed by wt-PcPAL on Chiralpak ZwiX (+) chiral column (after 504 h reaction time; ee= 41%).

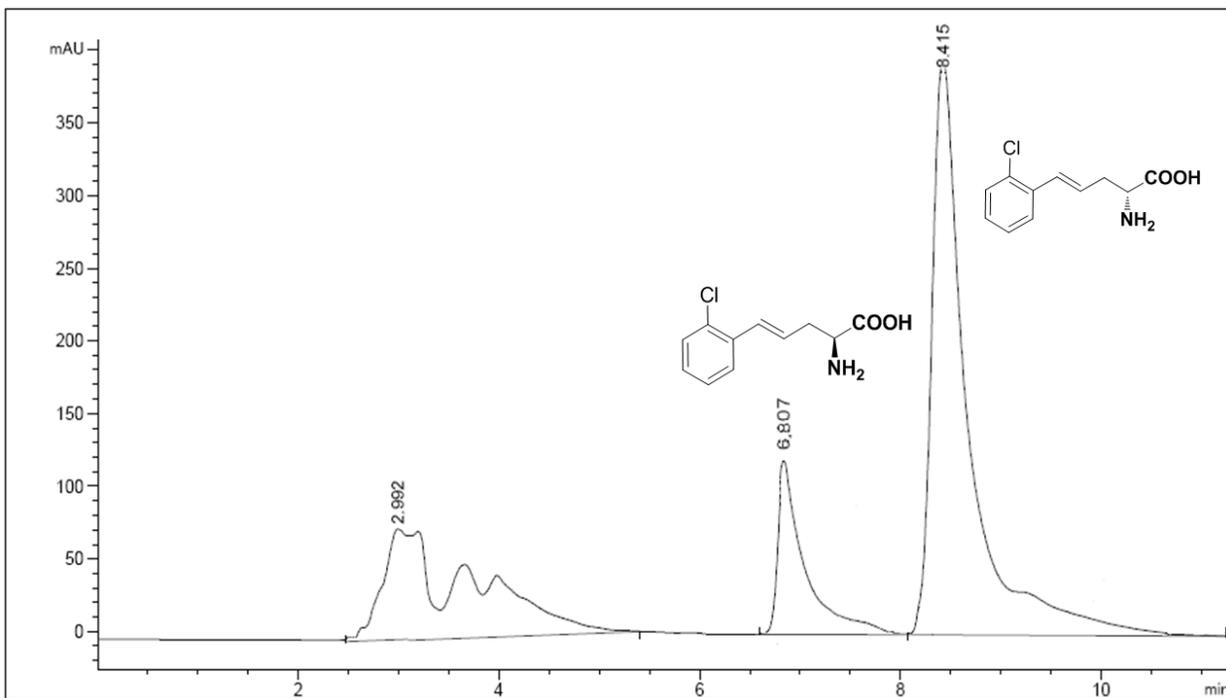


Figure S31. HPLC analysis of the ammonia elimination from *rac-1c* catalyzed by wt-PcPAL on Chiralpak ZwiX (+) chiral column (after 600 h reaction time; ee= 55%).

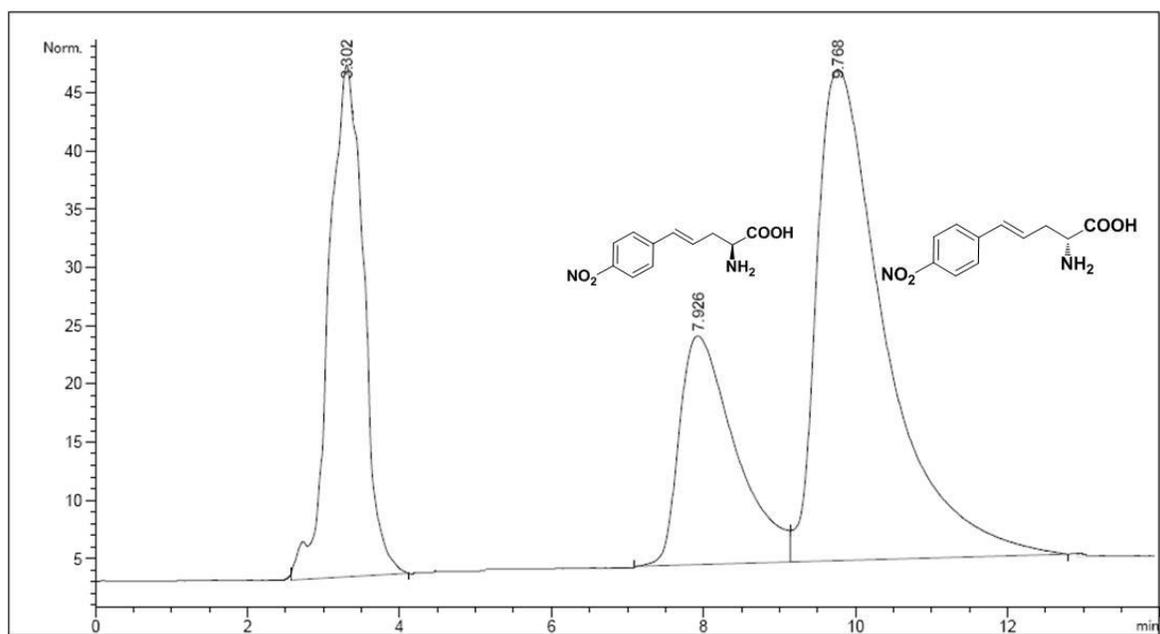


Figure S32. HPLC analysis of the ammonia elimination from *rac*-**1d** catalyzed by wt-PcPAL on Chiralpak ZWIX (+) chiral column (after 600 h reaction time; ee= 41%).

5.5 Chiral HPLC analysis of the final products from F137V-PcPAL-catalyzed kinetic resolutions of *rac*-**1a-d**

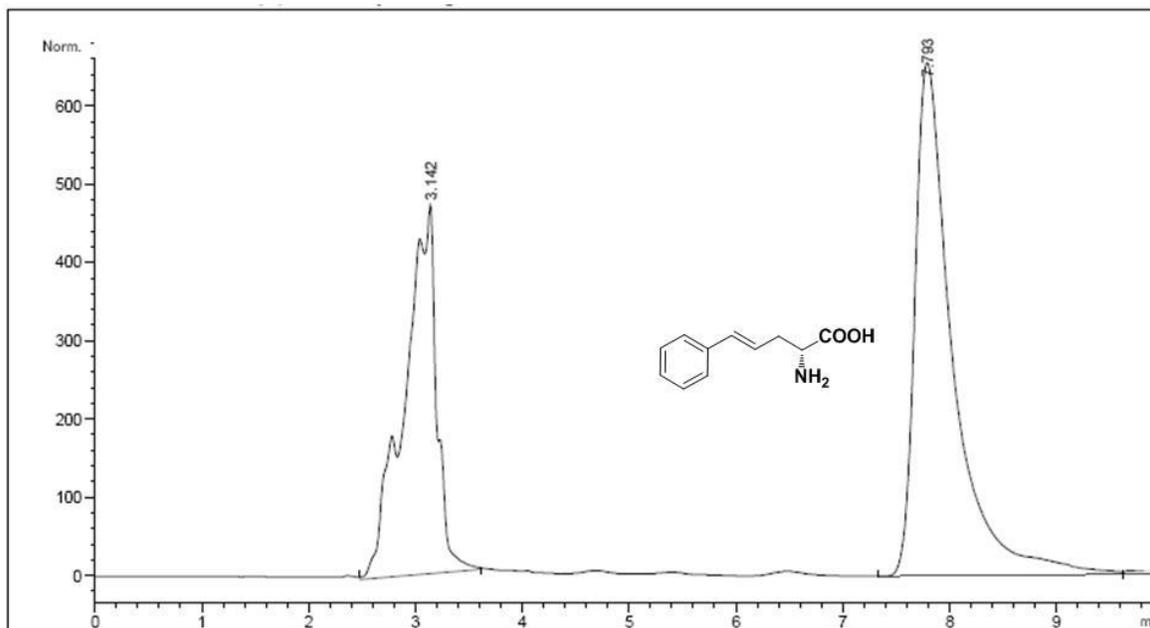


Figure S33. HPLC analysis of the ammonia elimination from *rac*-**1a** catalyzed by F137V-PcPAL on Chiralpak ZwiX (+) chiral column (after 24h reaction time; ee >99%).

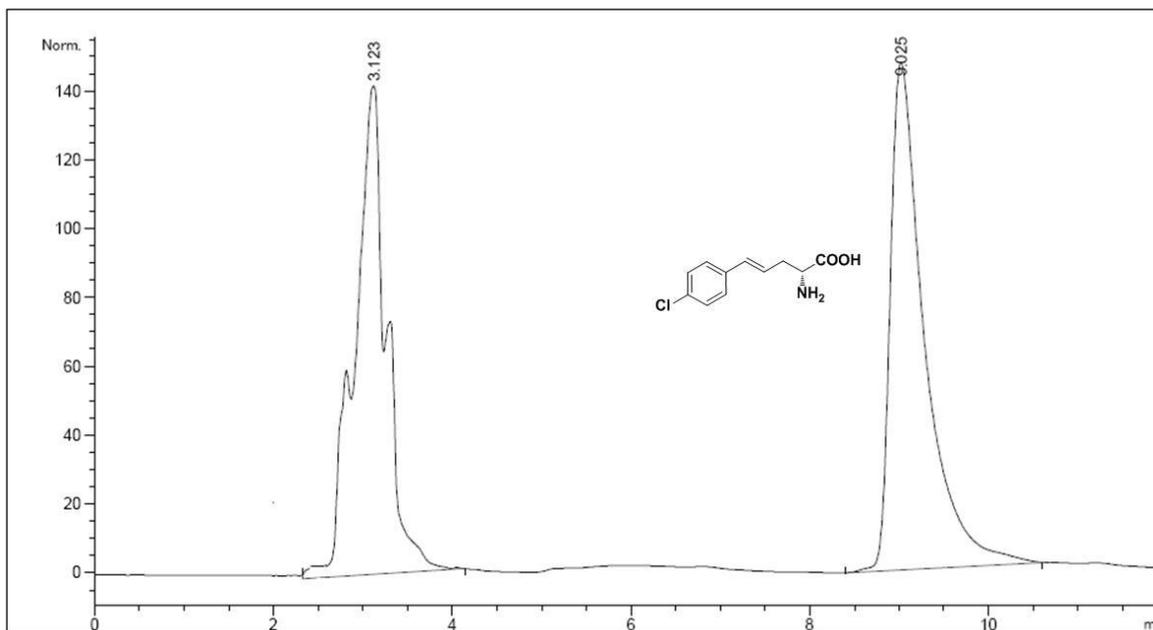


Figure S34. HPLC analysis of the ammonia elimination from *rac*-**1b** catalyzed by F137V-PcPAL on Chiralpak ZwiX (+) chiral column (after 274 h reaction time; ee >99%).

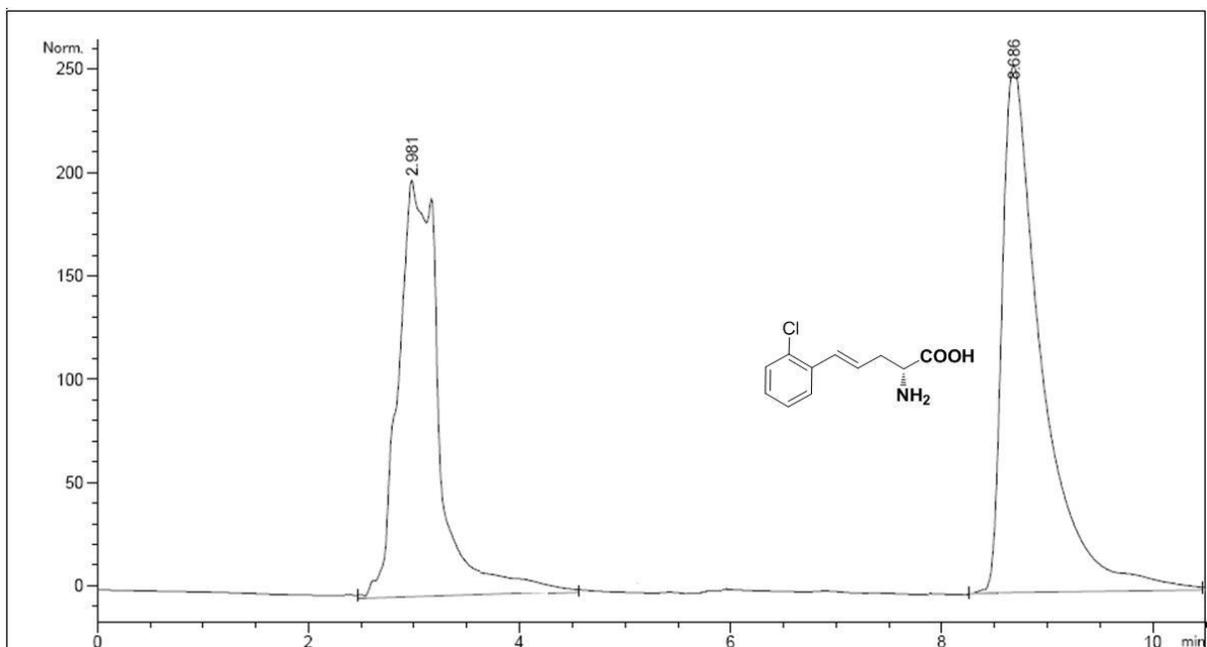


Figure S35. HPLC analysis of the ammonia elimination from *rac*-1c catalyzed by F137V-PcPAL on Chiralpak ZwiX (+) chiral column (after 134 h reaction time; ee >99%).

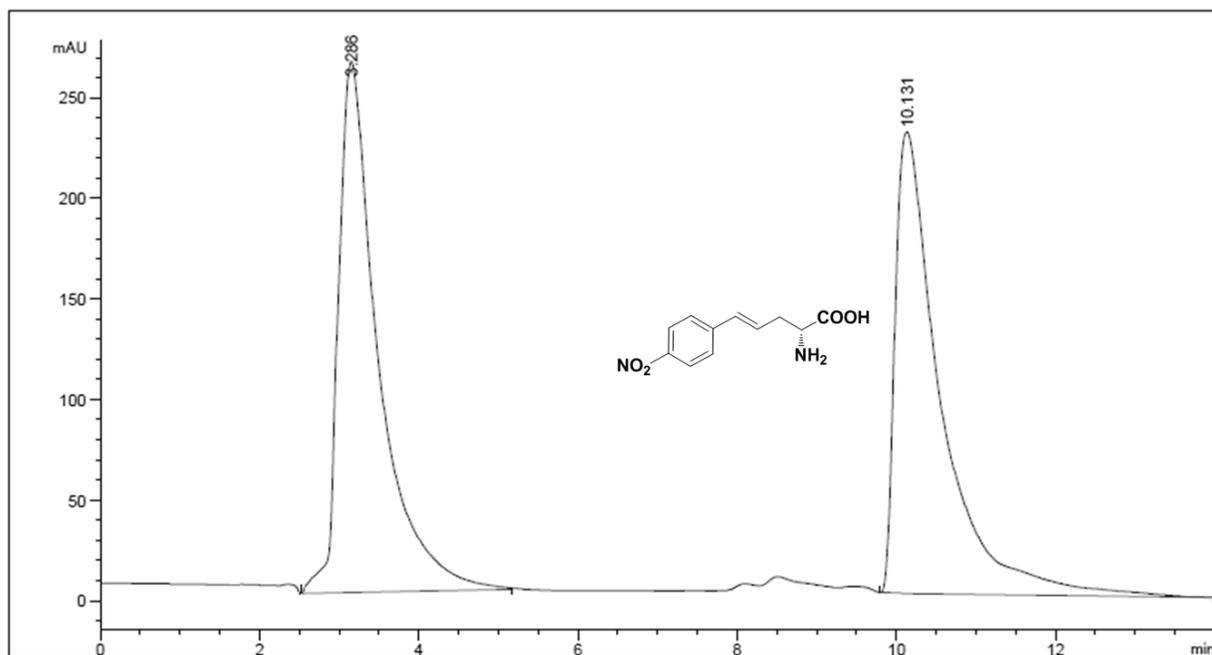


Figure S36. HPLC analysis of the ammonia elimination from *rac*-1d catalyzed by F137V-PcPAL on Chiralpak ZwiX (+) chiral column (after 300 h reaction time; ee >99%).

In order to rule out that the asymmetrical peak form is caused by the presence of impurities, besides the LC-MS analysis of *rac*-**1a-d** (see Supporting information, chapter 3.2), which supports the high purity level of the compounds, we also analyzed commercially available L-styrylalanine **L-1a** using the developed chiral method. The obtained asymmetrical peak, similar with those obtained from the enzymatic reactions or from the separation of *rac*-**1a-d** supports that the peak shape of the chiral HPLC chromatograms may be due to variable binding poses to the chiral stationary phase.

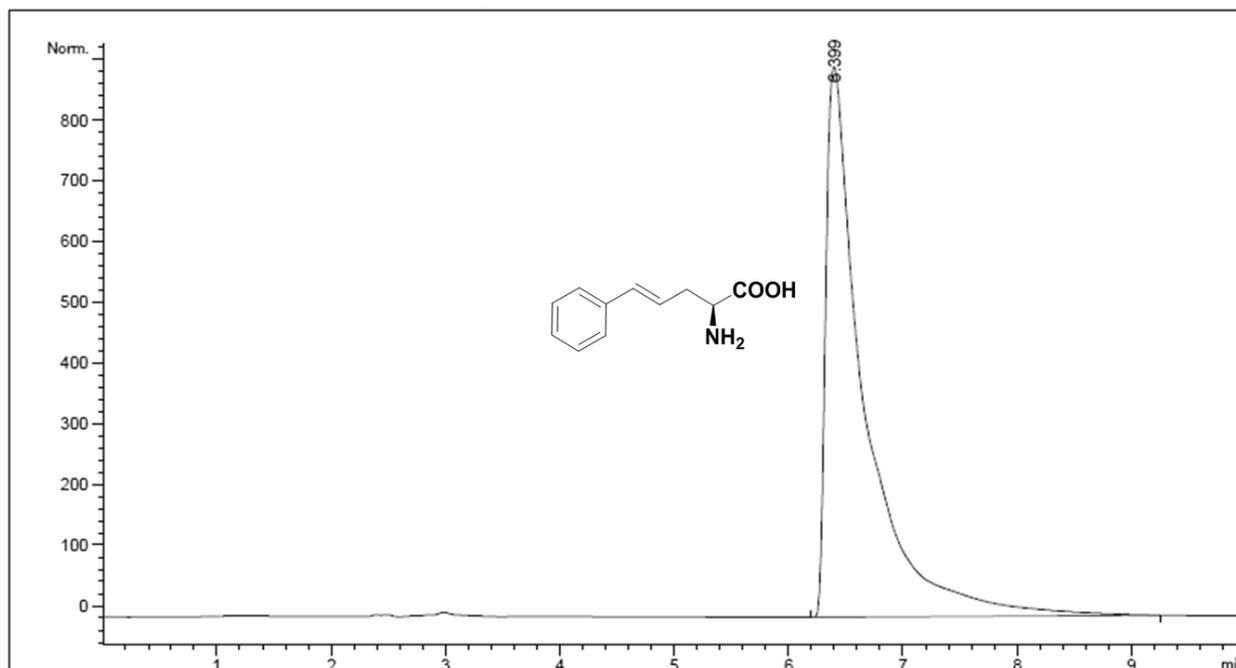


Figure S37. HPLC chromatogram of **L-1a** obtained on Chiralpak ZwiX (+) chiral column

5.6. Investigation of the ammonia addition onto styrylacrylates **2a-d** by the PcPAL variants

The enzymatic ammonia addition reactions were performed in triplicate as follows:

The 6M NH₃-solutions (1 mL, pH 10, adjusted with CO₂), containing styrylacrylates **2a-d** at a final concentration of 5 mM, and purified PcPAL (50 μg of F137V-, F137A-, F137G-, or wt-PcPAL), were incubated at 30 °C with shaking at 250 rpm. Conversions were monitored using reverse phase high performance liquid chromatography (HPLC) as described earlier in section 5.1. Samples (30 μL) were taken from the reaction mixtures after 24 h time intervals, quenched by adding an equal volume of MeOH, vortexed and centrifuged (13400 rpm, 12100 g, 2 min). The supernatant

was used directly for HPLC analysis after transferring through a 0.22 μm nylon membrane filter. Despite monitoring the reactions over 20 days, no product (L-**1a-d**) formation could be detected by HPLC. In order to rule out enzyme deactivation during the 20 days reaction monitoring, fresh batch of enzyme was added to the mixture after each 48 h period.

6. SDS-PAGE and size-exclusion chromatogram from the purification steps of F137V-PcPAL

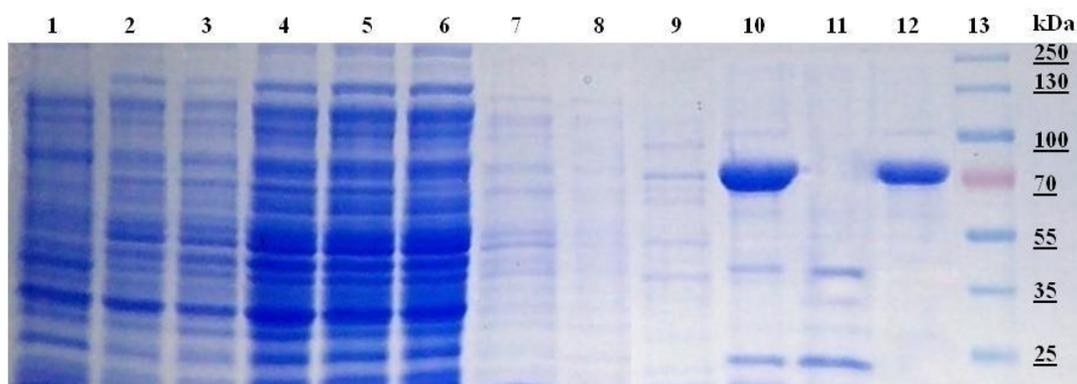


Fig S37. SDS-PAGE gel containing samples from the purification steps of F137V-PcPAL: **1:** before induction; **2:** after induction; **3:** supernatant of cells; **4:** cell lysate **5:** supernatant after centrifugation of the lysate; **6:** flow-through from the Ni-NTA column; **7:** low salt washing step (from Ni-NTA); **8:** high salt washing step (from Ni-NTA); **9:** fraction eluted with 25mM of imidazole (from Ni-NTA); **10:** fraction eluted with 400 mM of imidazole (from Ni-NTA); **11:** fraction eluted at 18-20 mL from the Superdex 200 10/300 GL column, **12:** fraction eluted at 10-12 mL from the Superdex 200 10/300 GL column, **13:** marker.

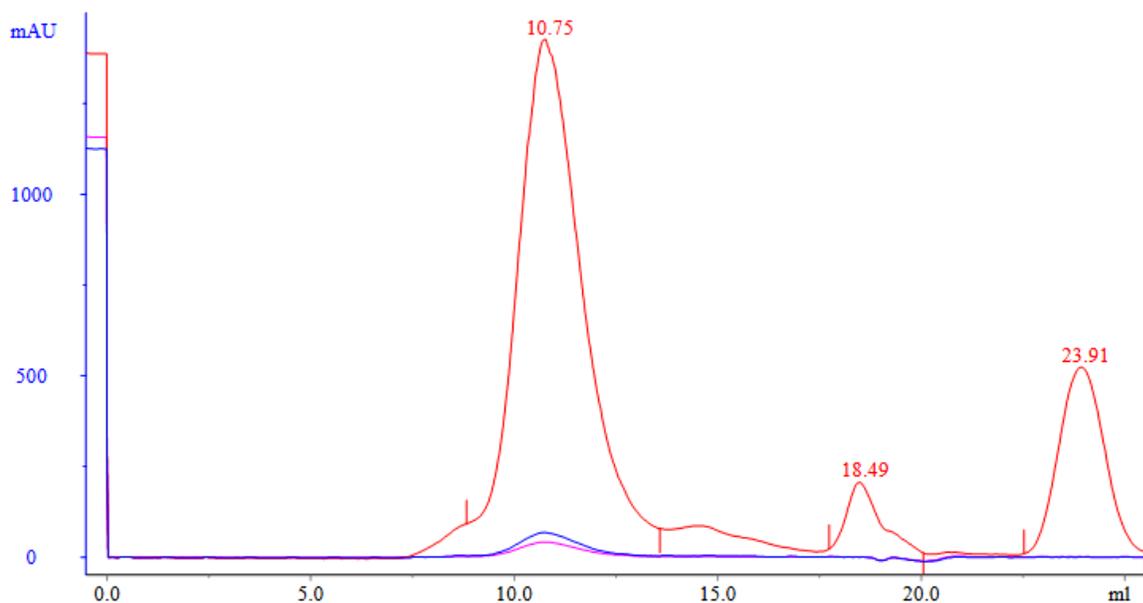


Fig S38. Chromatogram obtained by size-exclusion chromatography of F137V-PcPAL using a Superdex 200 10/300 GL column

7. Theoretical methods and derivations

7.1. Molecular modeling principles

Besides the practical work, we tried to rationalize the experimental data with additional theoretical work. Our research philosophy comprised of a straightforward, physics-based approach. We aimed to acquire the ligand pose/conformation sets of the important biocatalytic sequence members and corresponding energy values *in silico*. These computational data were used to approximate changes on the potential energy surface and thus characterize certain phenomena of the catalysis. Furthermore, the enumeration of the possible ligand conformations could reveal energetically preferred, but unproductive states which can slow down or block the catalytic process.

Following this principle, three major properties were calculated. One was the ligand binding energy (${}_b\Delta E$), calculated with a modified MM-GBSA methodology, appreciated as one of the most useful and meaningful theoretical quantities, from the standpoint of experimental work. To approximate the Gibbs energy of activation, two descriptors were used. One was associated with

the acidity of the β -proton ($_{cB}\Delta E$), and the other one was associated with the steric strain ($_{Ster}\Delta E$). Moreover, these energy differences were transformed relative to the same property of L-Phe or (*E*)-cinnamate with wt-PcPAL that makes these values not just more intelligible, but enables us to compare intermediate states with different connectivity with molecular mechanics.⁹ However, one should be warned that the absolute values of these calculated properties should not be compared directly to experimental values, and should be even scaled in some cases.

For such investigations an atomistic model of the enzyme – representing its structure at the assay conditions, with special attention paid to the active site – and knowledge about the reaction mechanism is needed. Unfortunately, the only crystal structure determined for PcPAL so far did not contain a substrate or substrate analogue.⁷ Moreover, the essential Y110-containing loop in this structure was not in the active conformation.⁸ Thus, a structure with a catalytically active loop conformation required partial homology modeling^{8,10} and refinement of the Y110-loop region.

Next, the enzyme-substrate complexes within the active site were constructed. Committing to these assumptions, the corresponding covalent *N*-MIO and preceding substrate-binding, non-covalent intermediate states of all studied substrates were created introducing our induced-fit covalent docking protocol which was previously applied in the case of another MIO enzyme, *Pantoea agglomerans* PAM.⁹ This protocol enumerates the possible ligand poses in the active site while accounting for the conformational change of the enzyme too. Moreover, explicit solvation of the complexes was performed with the aid of Grand canonical Monte Carlo simulations. Lastly, the final, refined enzyme-substrate complexes were energetically evaluated with Schrödinger's Prime function. Docking did not indicate significant, energetically preferred unproductive states with any substrate and PcPAL mutant combinations in the case of the substrate-binding, non-covalent intermediate directly leading to the subsequent covalent *N*-MIO state. Thus this intermediate state wasn't considered problematic and is not discussed throughout this text.

7.2. Molecular modeling methods

For molecular modeling studies, the tetrameric partial homology model of PcPAL¹⁰ was used as a starting structure. The crude model was completed and adjusted using the Protein Preparation Wizard¹¹ in two steps: i) hydrogen atoms were added and bond orders were assigned, and ii) the hydrogen bond network, tautomeric states, side chain conformations of selected amino acids, and ionization states were determined and optimized corresponding to the experimental assay conditions. In all four active centers, Y110 was set deprotonated and Y351 protonated. Protein pKa were predicted using PROPKA.¹² Next, residues 103-126 of the catalytically essential Y110 containing loop and further residues within 3 Å distance were subjected to loop modeling. This step consisted of two repetitions of side chain prediction of the previously mentioned residues with C_α-C_β vectoring incorporated, thus simulating also the small scale movement of the backbone (implicit water solvation model: VSGB, number of steps:2) using Prime.¹³ The final model was minimized with Prime¹³ [RMSG: 0.1 kcal mol⁻¹ Å⁻¹, algorithm: TNCG, implicit water solvation model: VSGB).

The refined and completed structure served as a starting point to create an overall protein model corresponding to the experimental assay conditions. The buffer solution solvated model was created by the Desmond program suite.¹⁴ The buffer solvated model was then equilibrated with a slightly modified default equilibration protocol, applying harmonic constraints to the Cartesian coordinates of buried hydrogen atoms and all protein heavy atoms. A spherical model of the active site with a radius of 27 Å, centered on the exocyclic methylene carbon of the MIO prosthetic group of chain C, was cut off and capped with acetyl and *N*-methylamino groups.

N-MIO type covalent complexes of substrates *rac*-**1a-d** and **2a-d** were constructed by our induced-fit covalent docking protocol. This involved the creation of initial conformations of compounds by docking with Glide¹⁵ into a modified and artificially enlarged active site in which the residues L134, F137, L138, L256, V259, and I460 were exchanged to Ala residues, further the MIO prosthetic group was reduced to Ala + Gly and three water molecules in the active site were removed.

After having docked into the enlarged active site, all side chains and the MIO group were restored. Residue 137 was in all cases changed from Ala to the residue corresponding to the actual enzyme mutant. For compounds D- and L-**1a-d**, a covalent bond between the nitrogen

atom of the amino group and the exocyclic carbon of MIO was created, for compounds **2a-d**, the MIO group was restored as an aminated NH₂-MIO group, and Y110 was set protonated. The ligands and the residues in close proximity were minimized, and finally, redundant conformations were eliminated with MacroModel.¹⁶ After this step, side chain conformations of residues L134, F137, L138, L256, V259, I460, and N260 were predicted with Prime (with C_α-C_β vectoring, implicit water solvation model: VSGB, number of steps:2). Next, water molecules in the active center and its immediate vicinity were predicted and placed with Grand Canonical Monte Carlo simulation using Desmond.¹⁴ Finally, a minimization of all ligands and the same set of atoms in the 6 Å proximity resulted in the final models and energies using Prime. The OPLS3 force field was applied in all molecular mechanics calculations and simulations.

Binding energy values (${}_b\Delta E$) were calculated with a modified MM-GBSA methodology. The standard MM-GBSA method involves the minimization, incorporating also implicit solvation, of the ligand-receptor complex, followed by the subsequent minimization of the receptor and the ligand individually after separation. Finally, the MM-GBSA score (${}_b\Delta E$) is calculated according to

$${}^{X137}{}_b\Delta E = E_{complex} - E_{receptor} - E_{ligand} \quad (S1)$$

Eq. S1,

where the three terms are the final energies of the previously mentioned minimizations, respectively, and X137 refers to the F137V mutant (X=V) or wild-type (X=F) PcPAL enzyme. Our modifications involved i) the substitution of the ligand energy term in Eq.S1 with a value obtained after mixed Monte Carlo/low-mode conformational search with MacroModel¹⁶ [solvation model: VSGB, number of steps: 2000, energy window: 31 kJ mol⁻¹] and a re-optimization with Prime [used with the same settings as before], ii) the substitution of the receptor energy term in Eq. S1 uniformly for all ligands with the corresponding apoenzyme structure's energy, and iii) the solvation of the *N*-MIO intermediate complexes were done with explicit solvation rather than implicit solvation with the aid of Grand Canonical Monte Carlo simulation using Desmond.¹⁴ This mixed use of implicit and explicit solvation is undesired in

general, however the application of implicit solvation to the minimization of the complexes introduces unacceptably large errors, and in addition, the favorable cancellation of errors made the mixed scheme convenient to provide relevant molecular structures and energies.

Values of the descriptor ${}_{ster}\Delta E$ were calculated according to Eq. S1, with the small change that the term E_{ligand} was obtained with minimization without the application of any solvation model. Values of descriptor ${}_{cB}^R\Delta E$ were calculated according to Eq. S2,

$${}_{cB}^R\Delta E = (E_{cB} - E_{N-MIO}) - (E_{cB} - E_{N-MIO})_{L-Phe;wt-PcPAL} \quad (S2)$$

where L-Phe;wt-PcPAL in the subscript refers to the same quantity calculated with L-Phe in the wt-PcPAL active center, E_{cB} is the single point energy of the model of the conjugated base, and E_{N-MIO} is the single point energy of the N-MIO model, both calculated at the B3LYP/6-311+G(dp) level with the GD3 dispersion correction model using the program suite MRCC (www.mrcc.hu).¹⁷ For the N-MIO model, the previously described enzyme-substrate models were used such as the atoms of the enzyme were deleted, and the dangling substrate-MIO bond was capped with a methyl group. For the conjugated base model, the *pro-S* β -proton was removed from the N-MIO model, and the remaining one was adjusted to help the β -carbon of the alanyl substructure to form sp^2 hybrid state arrangement.

Quantities ${}_{ster}\Delta E$ and ${}_b\Delta E$ (denoted here as Y in general) were also evaluated relative to the corresponding values of L-Phe or (E)-cinnamate with wt-PcPAL throughout the text, according to

$${}_{Y}^R\Delta E = {}_{Y}^{X137}\Delta E - {}_{Y}^{F137}\Delta E_{L-Phe;(E)-cinn} + {}_{Y}^{F137X}\Delta E \quad (S3)$$

Eq.S3.

${}_{Y}^{F137X}\Delta E$ has the purpose to incorporate the change when the mutation F137V is introduced, and it is calculated according to Eq. S4.

$${}_{Y}^{F137X}\Delta E = (E_{complex} - E_{receptor})_{44}^{X137} - (E_{complex} - E_{receptor})_{F137} \quad (S4)$$

For the necessary relativization of the previously discussed quantities, see ref.9.

7.3. Derivation and reasoning of Eq. 3 from the main text

At substrate saturation, when the concentration of the free enzyme is negligible, we assumed that four species of enzyme–substrate intermediate complexes (E_{L-psc} , E_{L-pst} , E_{L-u} and E_D), can be present in thermal equilibrium (not considering allosteric inhibition). As discussed previously, from the L-enantiomer two types of catalytically active enzyme–substrate complexes can be formed (the *s-cis* or *s-trans* **2a-d** yielding L-**1a-d**-N-MIO states, abbreviated as E_{L-psc} , E_{L-pst}) and an unproductive one (the unproductive L-**1a-d**-N-MIO states, abbreviated as E_{L-u}), while another unproductive complex could be formed from the D-enantiomers (the unproductive D-**1a-d**-N-MIO states, abbreviated as E_D). Thus, according to Eq. S5, the catalytic activity is the sum of the activities of E_{L-psc} and E_{L-pst} ,

$$v_{\max} = k_{\text{cat}}[E^*] = k_{\text{cat,L-psc}}[E_{\text{psc}}] + k_{\text{cat,L-pst}}[E_{\text{pst}}] \quad (\text{S5})$$

where $[E^*]$ is the total concentration of all the enzyme species. After dividing by $[E^*]$, the fractions of the concentrations of E_{L-psc} and E_{L-pst} can be rewritten as molar fractions. Further, substituting the Eyring equation (Eq. S6) yields Eq. S7,

$$k = \frac{k_{\text{B}}T}{h} e^{\frac{\Delta G^\ddagger}{RT}} \quad (\text{S6})$$

$$k_{\text{cat}} = \frac{k_{\text{B}}T}{h} e^{\Delta G_{L-psc}^\ddagger/RT} x_{\text{psc}} + \frac{k_{\text{B}}T}{h} e^{\Delta G_{L-pst}^\ddagger/RT} x_{\text{pst}} \quad (\text{S7})$$

where T is the temperature, k_{B} is the Boltzmann constant, h is the Planck constant, ΔG^\ddagger is the corresponding Gibbs energy of activation, and x is the corresponding molar fraction. To approximate the Gibbs energy of activation, it was decomposed into two parts, into an electronic

term that is treated with quantum chemical calculations, and a term specific for the steric stress that is treated with molecular mechanics calculations (Eq. S8).

$$\Delta G^\# \approx \Delta E^\# = \Delta E_{Electronic}^\# + \Delta E_{Steric}^\# \quad (S8)$$

Numerous evidences suggest that the elimination step happens with an E1cB mechanism. Thus, Hammond's postulate was applied, and a descriptor – associated with the acidity of the *pro-S* β -proton (${}_{cb}^R\Delta E$) – in a linear relationship was used to calculate $\Delta E_{Electronic}^\#$ (Eq. S9). In the case of steric strain, a simple rule was applied stating that *N*-MIO type enzyme–substrate complexes with better interaction energies are able to “provide the energy needed for activation proportionally easier”, thus $\Delta E_{Steric}^\#$ is linearly dependent on ${}_{ster}^R\Delta E$ (Eq. S10). This assumption is supported by the fact that the active center is very compact in the *N*-MIO intermediate form also almost entirely desolvating the substrate.

Because **L-1a-d** are congeneric – and thus it is expected to exhibit fairly similar properties in many cases –, it was assumed that irrespectively of the true dependence of the activation energy terms on the descriptors, linear dependence should give a satisfactory description. However, using both constants *b* and *d* from Eq. S9 and Eq. S10, respectively, in a non-linear estimation would lead eventually to overfitting, thus, as a compromise, *b* was considered equal to *d*.

$$\Delta E_{Electronic}^\# = b * {}_{cb}^R\Delta E + c \quad (S9)$$

$$\Delta E_{Steric}^\# = d * {}_{ster}^R\Delta E + e \quad (S10)$$

Finally, substituting Eq. S8, Eq. S9, and Eq. S10 into Eq. S7, bringing the molar fractions into the exponents, and applying that *b=d*, Eq. S11 (Eq. 3 in the main text) is obtained, where *a* and *b* are constants.

$$k_{cat} = a(e^{b*\Delta E_{L-psc}^\#/RT + \ln x_{L-psc}} + e^{b*\Delta E_{L-pst}^\#/RT + \ln x_{L-pst}}) \quad (S11)$$

Non-linear estimation of the constants in Eq. S11 gave statistically significant regression ($R^2=0.82$, $p=0.002$, $F=21.8$, $df=6$) and parameters ($a=90.632$, $p=0.023$; $b=0.0791$, $p=0.014$). The

temperature was taken to be 303 K. In a few cases, multiple *p*sc and/or *p*st conformations were obtained (**3a** with wt- and F137V-PcPAL, **4a** with F137V-PcPAL). In these cases, only the two lowest energy active conformations were taken into account. Strictly speaking, they were two *p*sc conformations in the case of **3a** with F137V-PcPAL, but they were used in the non-linear estimation nevertheless. Data in Table 4. of the main text show the exact lowest energy *p*sc and *p*st conformations of **3a** with F137V-PcPAL.

Descriptor values and molar fractions were obtained for 4-fluoro-, 4-trifluoromethyl-, 4-bromo, and 4-cyanostyrylalanine with the same procedure as discussed earlier. Predicted k_{cat} values ($k_{\text{cat,pred}}$) were determined according to Eq. S12. Statistical analysis was conducted with the use of Statistica.¹⁸

$$k_{\text{cat,pred}} = 90.632(e^{0.014*\Delta E_{L-p\text{sc}}^{\#}/RT+\ln x_{L-p\text{sc}}} + e^{0.014*\Delta E_{L-p\text{st}}^{\#}/RT+\ln x_{L-p\text{st}}}) \quad (\text{S12})$$

References

- 1 W. Ying, D. Guang-Fen, G. Cheng-Zhi, X. Fen, D. Bin, H. Lin, *Tetrahedron*, 2016, **72**, 472-478.
- 2 (a) M. Abe, K. Nishikawa, H. Fukuda, K. Nakanishi, Y. Tazawa, T. Taniguchi, M. Shindo, *Phytochemistry*, 2012, **84**, 56-67; (b) L. P.B. Beaulieu, L.E. Zimmer, A. Gagnon, A.B. Charette, *Chem. Eur. J.*, 2012, **18**, 14784-14791.
- 3 a) H. W. Jeong, M. R.Kim, K.H. Son, M. Y. Han, J.H. Ha, M. Garnier, L. Meijer, B.M. Kwon, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1819 – 1822; b) M. Pedro, A. Pastor, M.J. Vilaplana, *Tetrahedron Lett.*, 1993, **34**, 3773-3776.
- 4 E. Pontiki, D. Hadjipavlou-Litina, K. Litinas, K., G. Geromichalos, *Molecules*, 2014, **19**, 9655-9674.
- 5 L. Crombie, W.M.L. Crombie, *J. Chem. Soc. Perk. T. 1*, 1994, **10**, 1267-1274.
- 6 M. Perez., M. Lamothe, B. Le Grand, R. Letienne, US Patent No.: US8022064B2, 2011.
- 7 H. Ritter, G.E. Schulz, *Plant Cell* 2004, **16**, 3426–3436.
- 8 S.Pilbák, A.Tomin, J. Rétey, L. Poppe, *FEBS J.* 2006, **273(5)**, 1004–1019.
- 9 A. Varga, G. Bánóczy, B. Nagy, L.C. Bencze, M.I. Toşa, Á. Gellért, F. D. Irimie, J. Rétey, L. Poppe, *RSC Adv.*, 2016, **6**, 56412-56420.
- 10 G. Bánóczy, C. Szabó, Z. Bata, G. Hornyánszky, L. Poppe, *Stud. Univ. Babeş-Bol. Ser. Chem*, 2015, **60**, 213-228.
- 11 a) Protein Preparation Wizard 2015-2: Epik, Version 2.4; Impact, Version 5.9; b) G. M. Satry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J. Comput. Aid. Mol. Des.*, 2013, **27**, 221-234.

- 12 M. H. M. Olsson, C. R. Søndergard, M. Rostkowski and J. H. Jensen, *J. Chem. Theor. Comput.*, 2011, **7**, 525-537.
- 13 Prime, Version 4.0, Schrödinger, LLC, New York, NY, USA
- 14 (a) Desmond Molecular Dynamics System, Version 4.2, D. E. Shaw Research, New York, NY, USA; (b) Maestro-Desmond Interoperability Tools, Version 4.2, Schrödinger, New York, NY, USA; (c) K. J. Bowers, E. Chow, H. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. Shan and D. E. Shaw, *Proceedings of the ACM/IEEE Conference on Super-computing (SC06)*, Tampa, FL, 2006, November 11-17.
- 15 (a) Glide 2015, Version 6.7, Schrödinger, LLC, New York, NY, USA; (b) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, D. E. Shaw, M. Shelley, J. K. Perry, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739-1749; (c) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin and D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177-6196; (d) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750–1759.
- 16 MacroModel, version 10.8, Schrödinger, LLC, New York, NY, 2015.
- 17 MRCC, a quantum chemical program suite written by M. Kállay, Z. Rolik, J. Csontos, I. Ladjánszki, L. Szegedy, B. Ladóczki, G. Samu, K. Petrov, M. Farkas, P. Nagy, D. Mester, B. Hégyel. Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki, and M. Kállay, *J. Chem. Phys.*, 2013, **139**, 094105, www.mrcc.hu.
- 18 Statistica 12 (<http://www.statsoft.com/Products/STATISTICA-Features/Version-12>, retrieved 20. 07. 2016.)