Scientific report 2015-2017

PN-II-RU-TE-2014-4-1668

MIO-enzyme toolkit with defined and expanded substrate targetability

Ob1. Creation of MIO-enzyme toolbox, containing known and newly reported MIO-enzymes.

WP1. Molecular cloning of novel recombinant MIO- enzymes

The MIO-containing bacterial enzymes containing only the core HAL/PAL motif region (about 460 amino acids) are expected to be more stable than their eukaryotic analogues. Therefore, we hypothesized that bacterial PAL's are better candidates for being stable biocatalysts than their naturally destabilized eukaryotic analogues. PALs of marine origin – especially PAL from *Idomarina loihiensis* – are capable to catalyze the ammonia addition with high activity at elevated ammonia and substrate concentrations.[1] In this context our aim was to explore and develop novel thermophilic and stable PAL and PAM enzymes, focusing on microorganisms from marine and extremophile sources. Accordingly in the first stage of the project PALs from uncharacterized organisms: *Kangiella koreensis* (marine bacteria) and *Pseudozyma antaractica* (extremophile yeast) were selected. Furthermore we aimed to perform the molecular cloning of PAM from *Taxus canadensis*, having opposite enantioselectivity to our already existing PAM from *Pantoea agglomerans*, thus the synthesis of both (*R*)- and (*S*)-β-aminoacids can be accomplished.

During the molecular cloning process, first by database mining we found the sequences of the targeted genes, which were further codon optimized for expression in *E.coli.* At the end of the genes restriction sites of NcoI, NdeI, BamHI restriction enzymes were introduced (Table 1), for directional cloning into the pET19b expression vector, which allows the expression of the recombinant protein with an enterokinase cleavable *N*-terminal His-tag. Furthermore the native sequence was optimized by replacing the Cys residues from the surface of the protein (Table 1), in order to avoid the intermolecular disulphide bridge formation, which could initiate the aggregation process of the protein during the expression process, resulting in decreased enzyme activities. The designed genes were synthesized through the services of Genscript.

Name of enzyme and gene	Amino acid sequence of the recombinat protein	Nucleotide sequence of the gene
Taxus canadensis PAM (Uniprot code: Q6GZ04) Mutant C683S	MGFAVESRSHVKDILGLINTFNEV KKITVDGTTPITVAHVAALARRHD VKVALEAEQCRARVETCSSWVQR KAEDGADIYGVTTGFGACSSRRTN QLSELQESLIRCLLAGVFTKGCASS VDELPATATRSAMLLRLNSFTYGC SGIRWEVMEALEKLLNSNVSPKVP LRGSVSASGDLIPLAYIAGLLIGKP SVVARIGDDVEVPAPEALSRVGLR PFKLQAKEGLALVNGTSFATALAS TVMYDANVLLLLVETLCGMFCEV IFGREEFAHPLIHKVKPHPGQIESA ELLEWLLRSSPFQDLSREYYSIDKL KKPKQDRYALRSSPQWLAPLVQTI RDATTTVETEVNSANDNPIIDHAN DRALHGANFQGSAVGFYMDYVRI AVAGLGKLLFAQFTELMIEYYSNG LPGNLSLGPDLSVDYGLKGLDIAM AAYSSELQYLANPVTTHVHSAEQ HNQDINSLALISARKTEEALDILKL MIASHLTAMCQAVDLRQLEEALV KVVENVVSTLADECGLPNDTKAR LLYVAKAVPVYTYLESPCDPTLPL LLGLEQSCFGSILALHKKDGIETDT LVDRLAEFEKRLSDRLENEMTAV RVLYEKKGHKTADNNDALVRIQG SRFLPFYRFVREELDTGVMSARRE QTPQEDVQKVFDAIADGRITVPLL HCLQGFLGQPNGSANGVESFQSV WNKSA	5'AATACCATGGGCCATATGATGGGTTTCGCTGTTGAATCTCGTTCTCACGTTAAAGACATCCTG GGTCTGATCAACACCTTCAACGAAGTTAAAAAAATCACCGTTGACGGTACGACCCCGATCACCG TTGCTCACGTTGCTGCTCTGGCTCGTCGTCACGACGTTAAAGTTGCTCTGGAAGCTGAACAGTGC CGTGCTCGTGTTGAAACCTGCTCTTCTTGGGTTCAGCGTAAAGCTGAAGACGGTGCTGACATCT ACGGTGTTACCACCGGTTTCGGTGCTTGCTCTTCTCGTCGTACCAACCAGCTGTCTGAACTGCAG GAATCTCTGATCCGTTGCCTGCTGGCTGGTGTTTTCACCAAAGGTTGCGCTTCTTCTGTTGACGA ACTGCCGGCTACCGCTACCCGTTCTGCTATGCTGCTGCGTCTGAACTCTTTCACCTACGGTTGCT CTGGTATCCGTTGGGAAGTTATGGAAGCTCTGGAAAAACTGCTGAACTCTAACGTTTCTCCGAA AGTTCCGCTGCGTGGTTCTGTTTCTGCTTCTGGTGACCTGATCCCGCTGGCTTACATCGCTGGTCT GCTGATCGGTAAACCGTCTGTTGTTGCTCGTATCGGTGACGACGTTGAAGTTCCGGCTCCGGAA GCTCTGTCTCGTGTTGGTCTGCGTCCGTTCAAACTGCAGGCTAAAGAAGGTCTGGCTCTGGTTAA CGGTACGTCTTTCGCTACCGCTCTGGCTTCTACCGTTATGTACGACGCTAACGTTCTGCTGCTGC TGGTTGAAACCCTGTGCGGTATGTTCTGCGAAGTTATCTTCGGTCGTGAAGAATTTGCTCACCCG CTGATCCACAAAGTTAAACCGCACCCGGGTCAGATCGAATCTGCTGAACTGCTGGAATGGCTGC TGCGTTCTTCTCCGTTCCAGGACCTGTCTCGTGAATACTACTCTATCGACAAACTGAAAAAACCG AAACAGGACCGTTACGCTCTGCGTTCTTCTCCGCAGTGGCTGGCTCCGCTGGTTCAGACCATCCG TGACGCTACCACCACCGTTGAAACCGAAGTTAACTCTGCTAACGACAACCCGATCATCGACCAC GCTAACGACCGTGCTCTGCACGGTGCTAACTTCCAGGGTTCTGCTGTTGGTTTCTACATGGACTA CGTTCGTATCGCTGTTGCTGGTCTGGGTAAACTGCTGTTCGCTCAGTTCACCGAACTGATGATCG AATACTACTCTAACGGTCTGCCGGGTAACCTGTCTCTGGGTCCGGACCTGTCTGTTGACTACGGT CTGAAAGGTCTGGACATCGCTATGGCTGCTTACTCTTCTGAACTGCAGTACCTGGCTAACCCGGT TACCACCCACGTTCACTCTGCTGAACAGCACAACCAGGACATCAACTCTCTGGCTCTGATCTCTG CTCGTAAAACCGAAGAAGCTCTGGACATCCTGAAACTGATGATCGCTTCTCACCTGACCGCTAT GTGCCAGGCTGTTGACCTGCGTCAGCTGGAAGAAGCTCTGGTTAAAGTTGTTGAAAACGTTGTT TCTACCCTGGCTGACGAATGCGGTCTGCCGAACGACACCAAAGCTCGTCTGCTGTACGTTGCTA AAGCTGTTCCGGTTTACACCTACCTGGAATCTCCGTGCGACCCGACCCTGCCGCTGCTGCTGGGT CTGGAACAGTCTTGCTTCGGTTCTATCCTGGCTCTGCACAAAAAAGACGGTATCGAAACCGACA CCCTGGTTGACCGTCTGGCTGAATTTGAAAAACGTCTGTCTGACCGTCTGGAAAACGAAATGAC CGCTGTTCGTGTTCTGTACGAAAAAAAAGGTCACAAAACCGCTGACAACAACGACGCTCTGGTT CGTATCCAGGGTTCTCGTTTCCTGCCGTTCTACCGTTTCGTTCGTGAAGAACTGGACACCGGTGT TATGTCTGCTCGTCGTGAACAGACCCCGCAGGAAGACGTTCAGAAAGTTTTCGACGCTATCGCT GACGGTCGTATCACCGTTCCGCTGCTGCACTGCCTGCAGGGTTTCCTGGGTCAGCCGAACGGTT CTGCTAACGGTGTTGAATCTTTCCAGTCTGTTTGGAACAAATCTGCTTAAGGATCCATA
Pseudozyma	MAPTADVLPAAEASTRPGLLVQPS DTKLRKASSFRTEQVVIDGNNLKI	5'AATACCATGGGCCATATGATGGCTCCGACCGCTGACGTTCTGCCGGCTGCTGAAGCTAGCACC CGTCCGGGTCTGCTGGTTCAGCCGTCTGACACCAAACTGCGTAAAGCTAGCTCTTTCCGTACCG

antarctica PAM (Uniprot code: M9M0D4) mutant C655S

QGLVASARYGHIPILDNSPAVRKRI
DDSVNSLIAKLDRGESIYGINTGFG
GSADSRTANTRALQLALLQMQQC
GVLPVPSTFPTGEPSSAPFALPLTD
TETSLVMPEAWVRGAIVVRLSSL
MRGHSGVRWEVLDKMQKLFLQN
NVTPVVPVRSSISASGDLSPLSYVA
GALAGQRGIYCWVTDKKSGQRVK
VTADEACRMHGIEPVLYEPKEALG
LLNGTAFSASVAGLATYEAEKLA
QLTQLTTAMAVEALKGTDASFAP
FIHEVARPHPGQIKSARYIRALLSG
SKLAEHLENEKHVLFSEDNGTLRQ
DRYTLRTASQWVGPGLEDIENAK
RSVDIEINSTTDNPMIDPYDADGRI
HHGGNFQAMAMTNAVEKIRLALC
AMGKMTFQQMTELVNPAMNRGL
PANLASTPDLSLNFHAKGIDIALAS
VTSELMFLGNPVSTHVQSAEMAN
QAINSLALISGRQTLQAVECLSMIQ
AWSLYLLCQALDIRALQYKVAEQ
LPAMVLASINSHFGEWMDEAKQA
EIALLVFKSMSKRLDETSSKDLRD
RLVETYQDASSVLVKYFSELPSGG
GADPLRNIVKWRAAGVAETEKIY
RDVTVEFLDNPYASHASHLLGKT
KRAYEFVRKTLGVPMHGKENLYE
FKGEFSQWNTTGGYVSVIYASIRD
GELYDMLAELEKEIY

AACAGGTTGTTATCGACGGTAACAACCTGAAAATCCAGGGTCTGGTTGCTTCTGCTCGTTACGG
TCACATCCCGATCCTGGACAACTCTCCCGGCTGTTCGTAAACGTATCGACGACTCTGTTAACTCTC
TGATCGCTAAACTGGACCGTGGTGAATCTATCTACGGTATCAACACCGGTTTCGGTGGTTCTGCT
GACTCTCGTACCGCTAACACCCGTGCTCTGCAGCTGGCTCTGCTGCAGATGCAGCAGTGCGGTG
TTCTGCCGGTTCCGTCTACCTTCCCGACCGGTGAACCGTCTTCTGCTCCGTTCGCTCTGCCGCTG
ACCGACACCGAAACCTCTCTGGTTATGCCGGAAGCGTGGGTTCGTGGTGCTATCGTTGTTCGTCT
GTCTTCTCTGATGCGTGGTCACTCTGGTGTTCGTTGGGAAGTTCTGGACAAAATGCAGAAACTGT
TCCTGCAGAACAACGTTACCCCGGTTGTTCCGGTTCGTTCTTCTATCTCTGCTTCTGGTGACCTGT
CTCCGCTGTCTTACGTTGCTGGTGCTCTGGCTGGTCAGCGTGGTATCTACTGCTGGGTTACCGAC
AAAAAATCTGGTCAGCGTGTTAAAGTTACCGCTGACGAAGCGTGCCGTATGCACGGTATCGAACC
CGGTTCTGTACGAACCGAAAGAAGCTCTGGGTCTGCTGAACGGTACGGCTTTCTCTGCTTCTGTT
GCTGGTCTGGCTACCTACGAAGCTGAAAAACTGGCTCAGCTGACCCAGCTGACCACCGCTATGG
CTGTTGAAGCTCTGAAAGGTACGGACGCTTCTTTCGCTCCGTTCATCCACGAAGTTGCTCGTCCG
CACCCGGGTCAGATCAAATCTGCTCGTTACATCCGTGCTCTGCTGTCTGGTTCTAAACTGGCTGA
ACACCTGGAAAACGAAAAACACGTTCTGTTCTCTGAAGACAACGGTACGCTGCGTCAGGACCGT
TACACCCTGCGTACCGCTTCTCAGTGGGTTGGTCCGGGTCTGGAAGACATCGAAAACGCTAAAC
GTTCTGTTGACATCGAAATCAACTCTACCACCGACAACCCGATGATCGACCCGTACGACGCTGA
CGGTCGTATCCACCACGGTGGTAACTTCCAGGCTATGGCTATGACCAACGCTGTTGAAAAAATC
CGTCTGGCTCTGTGCGCTATGGGTAAAATGACCTTCCAGCAGATGACCGAACTGGTTAACCCGG
CTATGAACCGTGGTCTGCCGGCTAACCTGGCTTCTACCCCGGACCTGTCTCTGAACTTCCACGCT
AAAGGTATCGACATCGCTCTGGCTTCTGTTACCTCTGAACTGATGTTCCTGGGTAACCCGGTTTC
TACCCACGTTCAGTCTGCTGAAATGGCTAACCAGGCTATCAACTCTCTGGCTCTGATCTCTGGTC
GTCAGACCCTGCAGGCTGTTGAATGCCTGTCTATGATCCAGGCTTGGTCTCTGTACCTGCTGTGC
CAGGCTCTGGACATCCGTGCTCTGCAGTACAAAGTTGCTGAACAGCTGCCGGCTATGGTTCTGG
CTTCTATCAACTCTCACTTCGGTGAATGGATGGACGAAGCTAAACAGGCTGAAATCGCTCTGCT
GGTTTTCAAATCTATGTCTAAACGTCTGGACGAAACCTCTTCTAAAGACCTGCGTGACCGTCTGG
TTGAAACCTACCAGGACGCTTCTTCTGTTCTGGTTAAATACTTCTCTGAACTGCCGTCTGGTGGT
GGTGCTGACCCGCTGCGTAACATCGTTAAATGGCGTGCTGCTGGTGTTGCTGAAACCGAAAAAA
TCTACCGTGACGTTACCGTTGAATTTCTGGACAACCCGTACGCTTCTCACGCTTCTCACCTGCTG
GGTAAAACCAAACGTGCTTACGAATTTGTTCGTAAAACCCTGGGTGTTCCGATGCACGGTAAAG
AAAACCTGTACGAATTTAAAGGTGAATTTTCTCAGTGGAACACCACCGGTGGTTACGTTTCTGTT
ATCTACGCTTCTATCCGTGACGGTGAACTGTACGACATGCTGGCTGAACTGGAAAAAGAAATCT
ACTAAGGATCCATA

Kangiella koreensis PAL (Uniprot code: C7R9W9)

MTDTKTNITFGHSSLTIEQICQLAKG
NATAKLNSAPEFKHKIDQGADFIKEL
LREDGVIYGVTTGYGDSVTTPVPVQD
THELPLHLTRFHGCGLGSIFSAEHTR
AILATRLASLSQGYSGVSWSLLQQLE
LLLQKDILPRIPEEGSVGASGDLTPL
SYVAAALIGEREVLYKGQTQPTEQVF
KSLGIKPITLQPKEGLAIMNGTAVMT
ALACLAFQRADYLTQLCSRITSLCSI
ALQGNSAHFDELLFSVKPHPGQNQVA
AWIRDDLNHYKHPRNSDRLQDRYSIR
CAPHIIGALKDAMPWMRQTIETELNS
ANDNPIIDGAGQHVLHGGHFYGGHIA
MVMDSMKTGIANLADLMDRQMALLVD
SKFNNGLPNNLSAASEQRRPLNHGFK
AVQIGVSAWTAEALKLTMPASVFSRS
TECHNQDKVSMGTIAARDCLRILDLT
EQVAAASLMAATQAVTLRIKQSQLDK
SSLSDGVLSTLEQVFEHFELVSEDRP
LEHELRHFVALIQEQHWSTYAN

5′AATACCATGGGCCATATGATGACCGACACCAAAACCAACATCACCTTCGGTCACTCTTCTCTGACCATCGA
ACAGATCTGCCAGCTGGCTAAAGGTAACGCTACCGCTAAACTGAACTCTGCTCCGGAATTTAAACACAAAATC
GACCAGGGTGCTGACTTCATCAAAGAACTGCTGCGTGAAGACGGTGTTATCTACGGTGTTACCACCGGTTACG
GTGACTCTGTTACCACCCCGGTTCCGGTTCAGGACACCCACGAACTGCCGCTGCACCTGACCCGTTTCCACGG
TTGCGGTCTGGGTTCTATCTTCTCTGCTGAACACACCCGTGCTATCCTGGCTACCCGTCTGGCTTCTCTGTCT
CAGGGTTACTCTGGTGTTTCTTGGTCTCTGCTGCAGCAGCTGGAACTGCTGCTGCAGAAAGACATCCTGCCGC
GTATCCCGGAAGAAGGTTCTGTTGGTGCTTCTGGTGACCTGACCCCGCTGTCTTACGTTGCTGCTGCTCTGAT
CGGTGAACGTGAAGTTCTGTACAAAGGTCAGACCCAGCCGACCGAACAGGTTTTCAAATCTCTGGGTATCAAA
CCGATCACCCTGCAGCCGAAAGAAGGTCTGGCTATCATGAACGGTACGGCTGTTATGACCGCTCTGGCTTGCC
TGGCTTTCCAGCGTGCTGACTACCTGACCCAGCTGTGCTCTCGTATCACCTCTCTGTGCTCTATCGCTCTGCA
GGGTAACTCTGCTCACTTCGACGAACTGCTGTTCTCTGTTAAACCGCACCCGGGTCAGAACCAGGTTGCTGCT
TGGATTCGTGACGACCTGAACCACTACAAACACCCGCGTAACTCTGACCGTCTGCAGGACCGTTACTCTATCC
GTTGCGCTCCGCACATCATCGGTGCTCTGAAAGACGCTATGCCGTGGATGCGTCAGACCATCGAAACCGAACT
GAACTCTGCTAACGACAACCCGATCATCGACGGTGCTGGTCAGCACGTTCTGCACGGTGGTCACTTCTACGGT
GGTCACATCGCTATGGTTATGGACTCTATGAAAACCGGTATCGCTAACCTGGCTGACCTGATGGACCGTCAGA
TGGCTCTGCTGGTTGACTCTAAATTCAACAACGGTCTGCCGAACAACCTGTCTGCTGCTTCTGAACAGCGTCG
TCCGCTGAACCACGGTTTCAAAGCTGTTCAGATCGGTGTTTCTGCTTGGACCGCTGAAGCTCTGAAACTGACC
ATGCCGGCTTCTGTTTTCTCTCGTTCTACCGAATGCCACAACCAGGACAAAGTTTCTATGGGTACGATCGCTG
CTCGTGACTGCCTGCGTATCCTGGACCTGACCGAACAGGTTGCTGCTGCTTCTCTGATGGCTGCTACCCAGGC
TGTTACCCTGCGTATCAAACAGTCTCAGCTGGACAAATCTTCTCTGTCTGACGGTGTTCTGTCTACCCTGGAA
CAGGTTTTCGAACACTTCGAACTGGTTTCTGAAGACCGTCCGCTGGAACACGAACTGCGTCACTTCGTTGCTC
TGATCCAGGAACAGCACTGGTCTACCTACGCTAACTAAGGATCCATA

Cleavage sites of the restriction enzymes: NcoI :CCATGG, NdeI :CATATG, BamHI: GGATCC
Cys-Ser mutations are marked with yellow colour in the amino acid sequence of the proteins

---

The synthesized genes from the pUC57 cloning vector were amplified by PCR using primers from Table 2, in order to introduce them into expression vector pET19b(+).The PCR products were further purified, using the *DNA Clean & Concentrator™-25Kit*. The purified PCR products and the recipient circular pET19b vector were digested with *NdeI* and *BamHI* restriction enzymes, and analyzed by agarose gel electrophoresis. The DNA bands were cut out from the agarose gel, and the recipient plasmid and insert at a ratio 1:3 were co-extracted using *Gen Elute Gel Extraction Kit* followed by ligation with T4 DNA ligase, which sticks the DNA ends together to form a single circular molecule that includes both the vector and the gene. The ligation products were transformed through heat-shock into *E.coli* XL-1 Blue competent cells and the presence of insert DNA vector was verified by colony PCR, using primers for T7 promoter and T7 terminator regions of the plasmid DNA (Table 2). The PCR reactions were analyzed by agarose gel electrophoresis (**Figure 1. a, b, c**).

During this procedure we succesfully cloned two novel (*Pza*PAM, *Kk*PAL) and known (*Tc*PAM) MIO eznymes, their vector maps being represented in in **Figure 2a-c.**

Table 2. Primers used for amplification of genes

Primers	Sequences	T_m (^{0}C)
TcPAM/FP	*5'* CTAGATAATACCATGGGCCATATG *3'*	62.3
TcPAM/RP	5'CCGATTATGGATCCTTAAGCAG 3'	63.7
PzaPAM/FP	5' CTAGATAATACCATGGGCCATATG3'	62.3
PzaPAM/RP	5' GGATCCGATTATGGATCCTTAGTAG3'	63.7
KkPAL/FP	5' CTAGATAATACCATGGGCCATATG3'	62.3
KkPAL/RP	5' CCGATTATGGATCCTTAGTTAGC3'	61.9
T7prom/FP	*5'*AATACGACTCACTATAGGGGAATTG *3'*	54
T7term/RP	5'TGCTAGTTATTGCTCAGCGG 3'	55

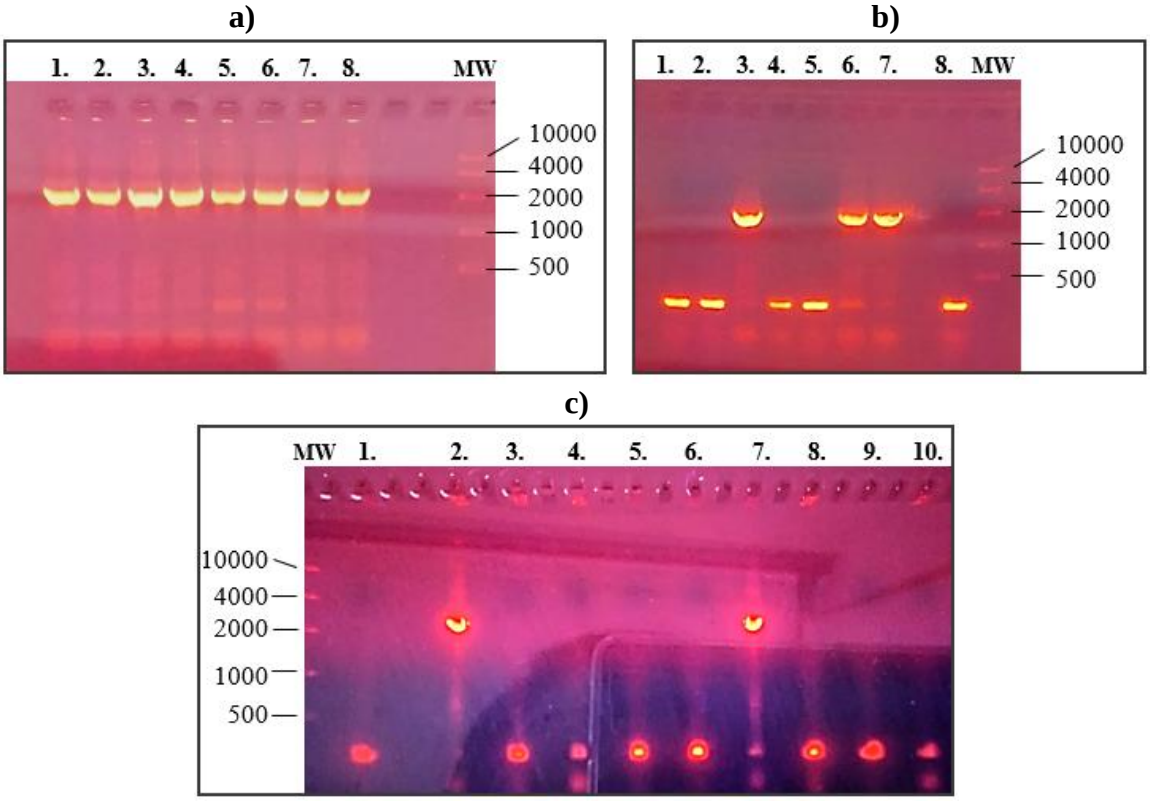


Figure 1. Agarose gel electrophoresis of the colony PCR products: **a.** *Tc*PAM ligated in pET19b and amplified - the insert gene with 2124bp is present in all 8 picked colonys. **b.** *Kk*PAL ligated in pET19b and amplified the insert gene with 1578 bp; 8 colonys were picked and in 3, 6, 7 *Kk*PAL is present in pET19b vector**,**1, 2, 4, 5, 8–*Kk*PAL is not present in pET19b vector.**c.** *Pza*PAM ligated in pET19b and amplified the insert gene with 2208 bp;10 colonys were picked and in 2, 7- *Pza*PAM is present in pET19b vector; 1, 3, 4, 5, 6, 8, 9 – *Pza*PAM is not present in pET19b vector.

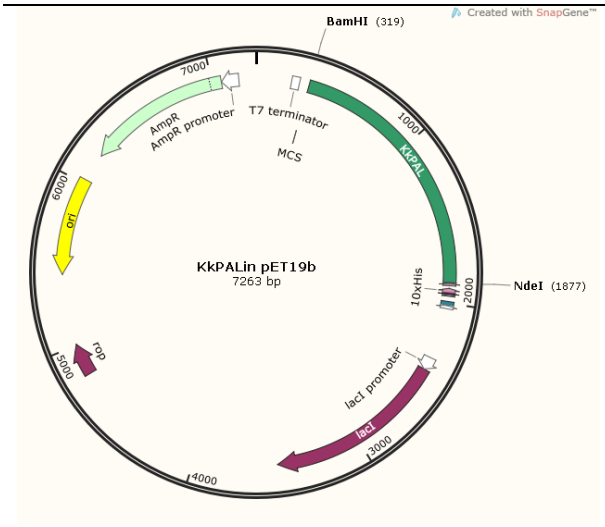
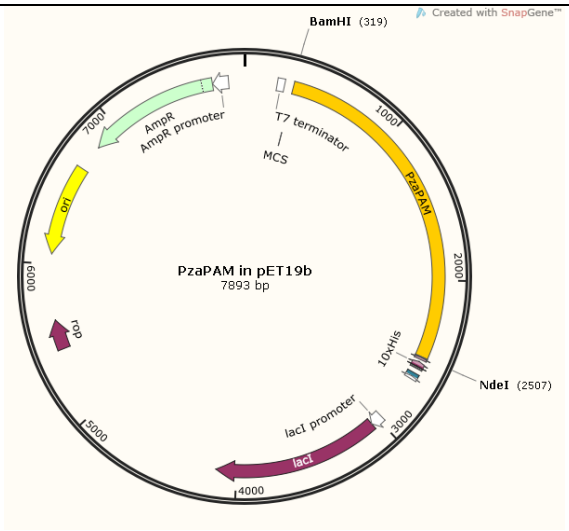
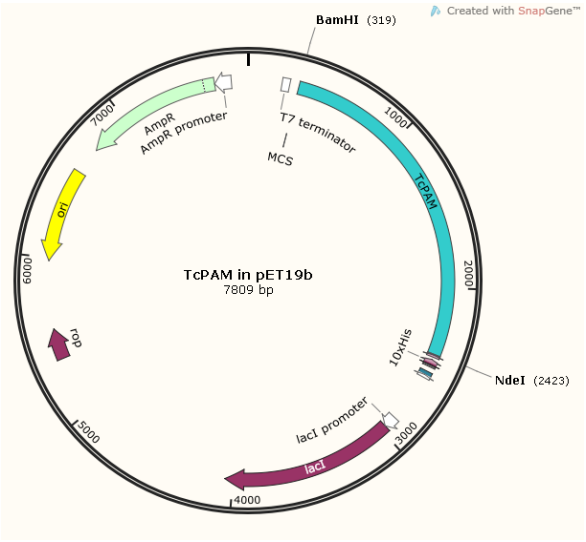


Figure 2: a) Plasmid map of the pET19b-*Kk*PAL.



b) Plasmid map of the pET19b-*Pza*PAM.



c) Plasmid map of the pET19b-*Tc*PAM

2. Molecular cloning of mutant MIO-enzymes through site-directed mutagenesis.

During this activity we envisaged rationally designed mutations in PcPAL and TcPAM (Table 3) influencing the aromatic substrate binding pocket of the enzyme (Figure 3) with the aim to extend the substrate scope of these MIO-enzymes towards bulky phenyalanine analogues. Accordingly PcPAL residues F137, L138, L134, L206, I460 and TcPAM residues L104, C107, L179 were mutated to smaller hydrophobic residues (Val, Ala) individually or in combinations (Table 3). The MIO-less PcPAL, (mutant S206A) was also prepared by molecular cloning, in order to investigate the presence of the competent, MIO-less reaction route, showing opposite enantiomerselectivity to the one observed in the generally occuring MIO-involved mechanism of action[2]. The mutations were introduced using the site-directed mutagenesis protocol developed by Naismith et al.[3] Therefore designed primers (Table 3) which contain non-overlapping sequence and primers which contain overlapping sequence (primer-primer complementary) with a $T_{m\ no}$ 5 to 10 ^{0}C higher than the $T_{m\ pp}$ so that the non-overlapping sequences can bridge the nick and bind to the newly synthesized DNA efficiently. Generally, the PCR amplification using these primers showed high efficiency (Figure 4.1 and 4.2), only in some cases low amplification eficiency could be observed, while no PCR

products were obtained in case of PaPAM mutants (Table 3, entries 32-35). In these cases optimization of the PCR procedures (addition of DMSO, modification of template and primer concentration) lead to succesfull mutation, except mutants of PaPAM, which could not be prepared even after optimization procedures. In all other cases the presence of desired mutations were confirmed by DNA sequencing, carried out using the sequencing services of Genomed and Antisel companies.

Further, the plasmids with the mutant genes were succesfully transformed into different *E.coli* host cells (*E.coli* Rossetta, BL21, C41 all with (DE3), pLysS modifications).

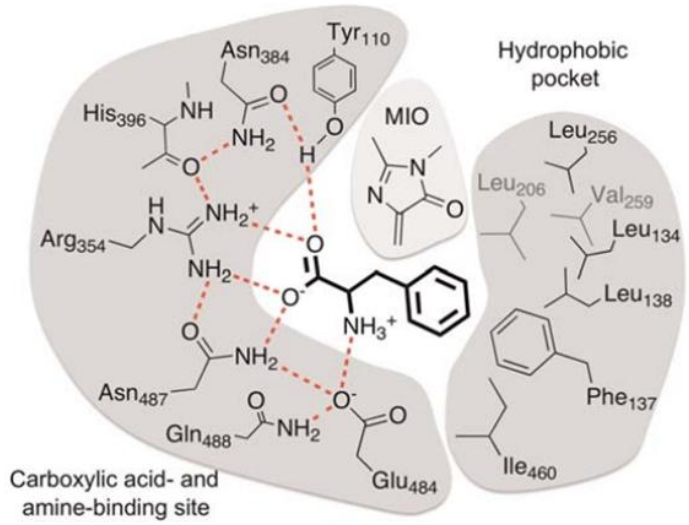


Figure 3. Catalytic site of *Pc*PAL (pdb-code: 1w27) with the docked substrate and binding pockets[4]

.

Table 3. List of prepared mutants and the corresponding primers and template genes used for mutagenesis

Entry	PcPAL mutant	Sequence (5'-3') of primers used for mutagenesis	Template DNA
1	L134A	L134A-for: CAGAAGGAAGCGATTAGATTCTTGAACGCCGGAATTTTCG	*wt-pcpal*
		L134A-rev: GAATCTAATCGCTTCCTTCTGCAAGGCTCCTCCCTGCTTAGTC	
2	L134V	L134V-for: GCAGAAGGAAGTGATTAGATTCTTGAACGCCGGAAT	*wt-pcpal*
		L134V-rev: CTAATCACTTCCTTCTGCAAGGCTCCTCCCTGCTTAG	
3	F137G	F137G-for: GATTAGAGGCTTGAACGCCGGAATTTTCGGAAACGGA	*wt-pcpal*
		F137G-rev: GCGTTCAAGCCTCTAATCAATTCCTTCTGCAAGGCTCCTC	
4	F137A	F137A-for: GATTAGAGCCTTGAACGCCGGAATTTTCGGAAACGGA	*wt-pcpal*
		F137A-rev: GCGTTCAAGGCTCTAATCAATTCCTTCTGCAAGGCTCCTC	
5	F137V	F137V-for: GATTAGAGTCTTGAACGCCGGAATTTTCGGAAACGGATC	*wt-pcpal*
		F137V-rev: GGCGTTCAAGACTCTAATCAATTCCTTCTGCAAGGCTCC	
6	L138A	L138A-for: TAGATTCGCGAACGCCGGAATTTTCGGAAACGG	*wt-pcpal*
		L138A-rev: GCGTTCGCGAATCTAATCAATTCCTTCTGCAAGGCTC	
7	L138V	L138V-for: TAGATTCGTGAACGCCGGAATTTTCGGAAACGGA	*wt-pcpal*
		L138V-rev: GGCGTTCACGAATCTAATCAATTCCTTCTGCAAGGCT	
8	S203A	S203A -for: ATTACTGCCGCCGGAGACTTGGTACCCTTGTCCTACATTGCCG	*wt-pcpal*
		S203A -rev: GTCTCCGGCGGCAGTAATAGTTCCTCTCAAGGGCAAGCAGG	
9	L206A	L206A-for: GGAGACGCGGTACCCTTGTCCTACATTGCC	*wt-pcpal*
		L206A-rev: GTACCGCGTCTCCGGAGGCAGTAATAGTTCCTCT	
10	L206V	L206V-for: CGGAGACGTGGTACCCTTGTCCTACATTGCC	*wt-pcpal*
		L206V-rev: GTACCACGTCTCCGGAGGCAGTAATAGTTCCTCTC	
11	L256A	L256A-for: GAAGGAGCGGCCTTGGTAAACGGAACTGCCGTAGG	*wt-pcpal*
		L256A-rev: AAGGCCGCTCCTTCCTTGGGCTGCAATTCGA	
12	L256V	L256V-for: GAAGGAGTGGCCTTGGTAAACGGAACTGCCGTAGG	*wt-pcpal*
		L256V-rev: CAAGGCCACTCCTTCCTTGGGCTGCAATTCGA	
13	Y351F	Y110F-rev: GTTACTCCGAAGGAGTCAGTTCCCTTGTTCATGGAGTCC	*wt-pcpal*
		Y351F-for: AGGACAGATTCGCCTTGAGAACTTCCCCCCAGTGGTTGG	
14	I460A	I460A-for: AGCCGAAGCTGCCATGGCCTCCTACTGCTCCGAATTGCA	*wt-pcpal*
		I460A-rev: CATGGCAGCTTCGGCTCCCTTGAATCCGTAGTCCAAGGAGG	
15	I460V	I460V-for: AGCCGAAGTTGCCATGGCCTCCTACTGCTCCGAATT	*wt-pcpal*

		I460V-rev: ATGGCAACTTCGGCTCCCTTGAATCCGTAGTCCAAGG	
16	F137A/I460V	F137A-for and F137A-rev	*I460V-pcpal*
17	F137A/ I460A	F137A-for and F137A-rev	*I460A-pcpal*
18	F137V/I460A	F137V-for and F137V-rev	*I460A-pcpal*
19	F137V/I460V	F137V-for and F137V-rev	*I460V-pcpal*
20	F137A/L138V	F137A,L138V-for: TAGAGCCGTGAACGCCGGAATTTTCGGAAACGG	*F137A-pcpal*
		F137A,L138V-rev: GCGTTCACGGCTCTAATCAATTCCTTCTGCAAGGCTC	
22	F137A/L138A	F137A,L138A-for: TAGAGCCGCGAACGCCGGAATTTTCGGAAACGGATCC	*F137A-pcpal*
		F137A,L138A-rev:GCGTTCGCGGCTCTAATCAATTCCTTCTGCAAGGCTCCTC	
23	F137V/L138A	F137V,L138A-for: TAGAGCCGTGAACGCCGGAATTTTCGGAAACGG	*F137V-pcpal*
		F137V,L138A-rev: GCGTTCACGGCTCTAATCAATTCCTTCTGCAAGGCTC	
24	F137V/L138V	F137V,L138V-for: TAGAGTCGTGAACGCCGGAATTTTCGGAAACGGA	*F137V-pcpal*
		F137V,L138V-rev:GGCGTTCACGACTCTAATCAATTCCTTCTGCAAGGCT	
25	L134V/F137A	L134V,F137A-for: GCAGAAGGAAGTGATTAGAGCCTTGAACGCCGGAATTTT	*F137A-pcpal*
		L134V,F137A-rev: CTCTAATCACTTCCTTCTGCAAGGCTCCTCCCTGCTTAGTC	
26	F137V/L138V/I460V	I460V-for and I460V-rev	*F137V,L138V-pcpal*
27	F137A/L138V/I460V	I460V-for and I460V-rev	*F137A,L138V-pcpal*
28	F137A/L138A/I460V	I460V-for and I460V-rev	*F137A,L138A-pcpal*
29	F137A/L138A/I460A	I460A-for and I460A-rev	*F137A,L138A-pcpal*
30	F137A/L138V/I460A	I460A-for and I460A-rev	*F137A,L138V-pcpal*
31	F137A/S206A/I460A	S203A-for and S203A-rev	F137A,I460A-*pcpal*
32	Y78F	Y78F-for: GAGTAATTTTCGGAGTAAACACTTCCATGGGAGGATTCG	*wt-papam*
		Y78F-rev: GTGTTTACTCCGAAAATTACTCTTTCGTCGGATACCATGG	
33	S171A	S171A-for: TTGGGAACTGCCGGAGACTTGGGACCCTTGGCCGCCATT	*wt-papam*
		S171A-rev: AAGTCTCCGGCAGTTCCCAAGGATCCCTTTTCGGGAATGCAGG	
34	T167A	T167A-for: TTGGGAGCTTCCGGAGACTTGGGACCCTTGG	*wt-papam*
		T167A-rev: CGGAAGCTCCCAAGGATCCCTTTTCGGGAA	
35	Y320F	Y320F -for: GAAGACGCCTTCTCCATTAGATGCACTCCCCAGATTTTGG	*wt-papam*
		Y320F -rev: CTAATGGAGAAGGCGTCTTCAATCTGATGGTTGGAGGCCTTTAC	
36	L104A	L104A -for: GAATCTGCGATCCGTTGCCTGCTGGCTGGT	*wt-tcpam*
		L104A -rev: ACGGATCGCAGATTCCTGCAGTTCAGACAGCTGG	
37	C107V	C107V -for: ATCCGTGTCCTGCTGGCTGGTGTTTTCACCAAAGG	*wt-tcpam*
		C107V -rev: CAGCAGGACACGGATCGCAGATTCCTGCAGTTC	
38	C107A	C107A -for: ATCCGTGCCCTGCTGGCTGGTGTTTTCACCAAAG	*wt-tcpam*
		C107A -rev: CAGCAGGCAACGGATCGCAGATTCCTGCAGTTC	
39	L108A/L104A	L108A/L104A -for: GTTGTGCGCTGGCTGGTGTTTTCACCAAAGGTTGC	*wt-tcpam*
		L108A/L104A -rev: AGCCAGCGCACAACGGATCGCAGATTCCTGCA	
40	L179A	L179A -for: GTGACGCGATCCCGCTGGCTTACATCGCTG	*wt-tcpam*
		L179A -rev: GGGATCGCGTCACCAGAAGCAGAAACAGAACCA	
41	Q319M	Q319M -for: CCGAAAATGGACCGTTACGCTCTGAATTCTTCTCCG	*wt-tcpam*
		Q319M -rev: CGGTCCATTTTCGGTTTTTTCAGTTTGTCGATAGAGTAGTATTC	
42	Q319N	Q319N -for: CCGAAAAACGACCGTTACGCTCTGAATTCTTCTCCG	*wt-tcpam*
		Q319N -rev: CGGTCGTTTTTCGGTTTTTTCAGTTTGTCGATAGAGTAGTATTC	
43	R325K	R325K -for: GCTCTGAAGTCTTCTCCGCAGTGGCTGGCTCCG	*wt-tcpam*
		R325K -rev: CGGAGAAGACTTCAGAGCGTAACGGTCCTGTTTCGGTTT	
44	R325N	R325N -for: GCTCTGAATTCTTCTCCGCAGTGGCTGGCTCCG	*wt-tcpam*
		R325N -rev: CGGAGAAGAATTCAGAGCGTAACGGTCCTGTTTCGGTTT	

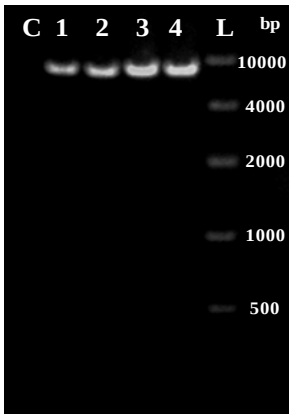


Figure 4.1. Agarose gel electrophoresis of the PCR products: **C.** DNA control before PCR, **1.** *Pc*PAL F137V mutant **2.** *Pc*PAL I460V mutant. **3**. *Pc*PAL L138V mutant product of PCR **4**. *Tc*PAM C107V mutant **L.** DNA Ladder

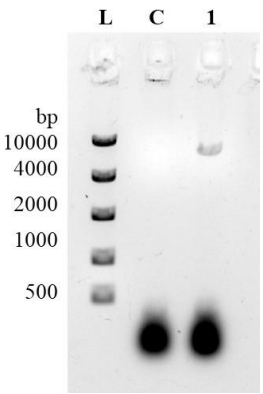


Figure 4.2. Agarose gel electrophoresis of the PCR products: **L.** DNA Ladder**, C.** negative control, PCR reactions for *Pc*PAL **L256V.** The detection of low intensity band (lane 1) was provided by ChemiDoc™ Imaging System from Bio-RAD.

WP 2,3. Expression, purification of the recombinant MIO-enzymes of the existing collection

The MIO-enyzme collection created by recombinant technologies through **WP1** consists of 47 enzymes: *i*) two novel, unexplored enzymes (*Kk*PAL, *Pza*PAM), *ii*) 31 mutant *Pc*PAL, from which 26 are novel mutant variants, and 9 mutant TcPAM variant from which 5 are novel mutants *iii*) 5 clones of the known phenylalanine ammonia lyases (PAL) from *Petroselinum crispum*, *Rhodosporidium toruloides*, *Anabeana variabilis* and phenylalanine aminomutase (PAM) from *Pantoea agglomerans* provided by the Biocatalysis and Biotransformation Research Center of the Host Institute and the newly cloned PAM from *Taxus canadesis*).

Further we optimized the expression, isolation and purification process of these enzymes. The recombinant plasmids were transformed through heat-shock into different *E.coli* competent cells (*E.coli* Rossetta, BL21, C41 all with (DE3), pLysS modifications) in order to study the influence of the nature of host cells. Furthermore we studied the inducer (IPTG) concentration and of the temperature upon the expression levels of the enzymes. The protein expression, isolation and purification process was optimized for three model enzymes, wt-*Pc*PAL, wt-*Rt*PAL,wt-*Pa*PAM, while the optimal conditions were employed for all other enzymes. In case of low isolation yields or purity, the optimization process was retaken for that particular case.

Generally for all three model proteins the best isolation yields and protein purities were obtained using Rosetta(DE3)pLysS host cells.

- ***The effect of different concentrations of inducer*** (0.1, 0.5 and 1 mM IPTG) was tested on the growing culture of cells containing the recombinant plasmids and the expression levels of the proteins were determined by SDS-PAGE (see Figure 6 in case of *Pa*PAM). The SDS-PAGE bands were similar for 0.5 and 1 mM IPTG concentrations (**Figure 5,** Lane C and D), indicating the same level of expression of the 72 kDa *Pa*PAM, despite the increasing inducer concentration. Based on these results, the final concentration of IPTG was set up to 0.1 mM.

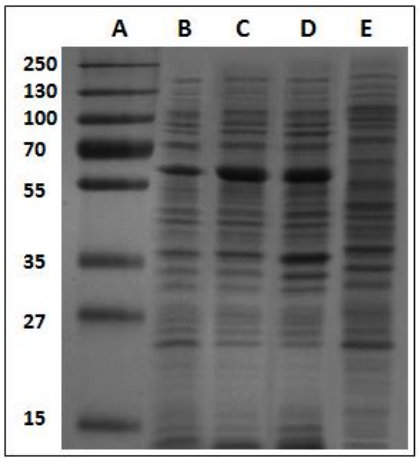


Figure 5. Induction of the expression of *Pa*PAM by different concentrations of IPTG in *E.coli* BL21(DE3)pLysS cells, after 4h. Lane **A**: protein ladder, Lane **B**: induction with 0.1 mM IPTG, Lane **C**: induction with 0.5 mM IPTG, Lane **D**: induction with 1 mM IPTG, Lane **E**: control (0 mM IPTG). The samples were prepared as described in experimental section.

- To evaluate ***the effect of growth temperature*** on the expression of recombinant enzymes, the cultures after induction with 0.1 M IPTG were incubated at different temperatures (18, 25, and 37°C). Initially the cell cultures were incubated at 37°C. After the density of cells reached OD_{600} ~0.6 (approx. 4 h) the temperature was reduced (to 18 or 25°C) and the cultures were induced with 0.1 mM IPTG. The density of the cells was monitored in time. Due to the reduced incubation temperature the protein synthesis rate is slower at 18°C than at 25 or 37°C, longer induction times are necessary for cells growing. At higher incubation temperature (25 or 37°C), the protein synthesis is faster and the stationary phase is

reached sooner than at lower temperatures. After 20 h the cells were harvested by centrifugation, followed by sonication and the protein was purified by metal affinity chromatography on Ni-NTA resin. The maximum yield of enzyme was obtained in case of 25°C incubation temperature for all the enzymes. The optimal post-induction time on the expression of *Pa*PAM was 15-16 hours and 11-12 hours for *Rt*PAL and *Pc*PAL. The determined optimal conditions were also been used for the large-scale fermentations (6x0,5 L in 2 L Erlenmeyer flasks).

- **Chromatographic purification of the recombinant enzymes**

Ni–NTA chromatography system is a rapid and easy purification technique. Proteins fused with His-tag at either ends (*N*- or *C*-terminus) bind tightly with high affinity on immobilized nickel ions. The strong binding between His-tag and matrix allows easy washing and efficient elution of bounded His-tagged protein by competition with imidazole.

In the pET-19b vector a His_{10}-tag at the *N*-terminus is included which is longer than the usual His_6-tag. Lengthening the His-tag increases the affinity of the enzyme to the Ni-NTA resin. Consequently, higher imidazole concentrations are required to elute all the bounded enzyme from the resin (from 250 mM up to 500 mM). Accordingly **Figure 6** depicts the monitorization of the purification process of the *Pa*PAM with or without the use of protease inhibitor cocktail tablets during the isolation, purification steps. The SDS-PAGE gels suggests that the use of protease inhibitors is important to obtain a protein fraction with higher purity (**Figure 7**, lane I). Furthermore we found also that 350 mM imidazole concentration is optimal for the elution of the protein from the Ni-NTA column.

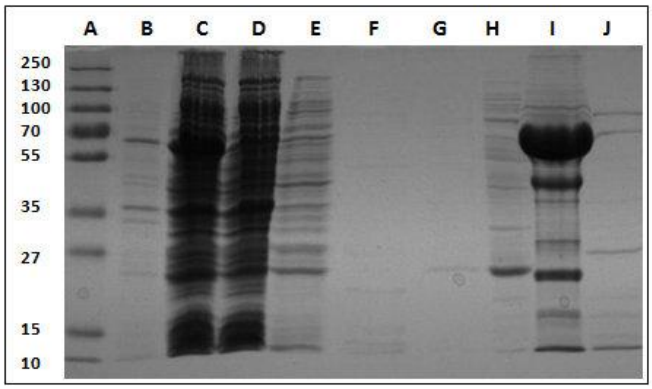


Figure 6. Purification of *Pa*PAM with Ni-NTA, in absence of protease inhibitor cocktail. Lane **A**: protein ladder, Lane **B**: supernatant, Lane **C**:pellet, Lane **D**: flow through Lane **E**: LS1, Lane **F**: HS, Lane **G**:LS2, Lane **H**: 20 mM Imidazole, Lane **I**: 500 mM Imidazole, Lane **J**: 1 mM Imidazole. The samples were prepared as described in experimental section.

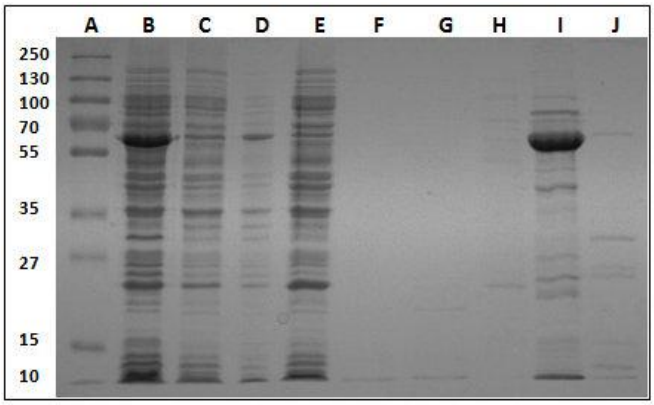
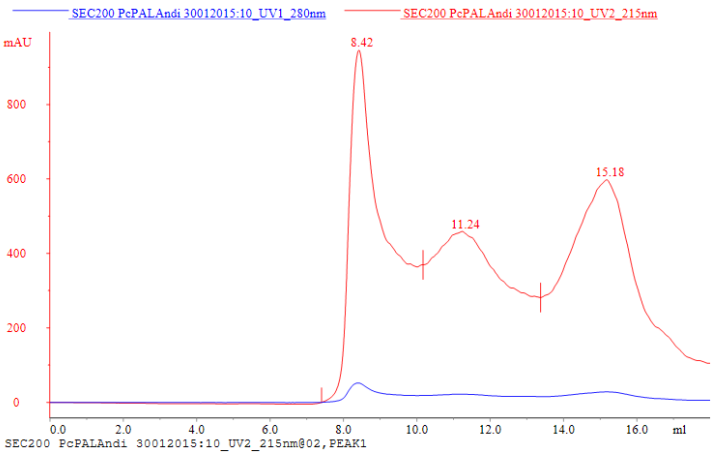


Figure 7. Purification of *Pa*PAM with Ni-NTA, in presence of protease inhibitor cocktail. Lane **A**: protein ladder, Lane **B**: supernatant, Lane **C**: flow through, Lane **D**: pellet, Lane **E**: LS1, Lane **F**: HS, Lane **G**:LS2, Lane **H**: 20 mM Imidazole, Lane **I**: 350 mM Imidazole, Lane **J**: 1 mM Imidazole. The samples were prepared as described in experimental section.

In some cases especially for purification of recombinant PcPALs in order to increase the purity of the isolated *Pc*PAL we introduced additional purification step, using size-exclusion chromatography, with a Superdex 200 10/300 GL. This chromatography step also allows the determination of the oligomerization state of the protein, and thus the homogeneity

of the isolated protein. This is crucial since different oligomerization states of the same protein (including aggregated forms of the protein) on SDS-PAGE cannot be distinguished, they migrate at the same molecular weight corresponding to the monomer however their enzyme activity can significantly differ. Thus, besides the separation of the protein impurities, the SEC method will provide additional results related to the correct, native-like oligomerization state of the isolated protein.

Performing the SEC purification of the fraction eluted with 250 mM imidazole from the Ni-NTA column, several protein peaks were eluted (**Figure 8**). The identity of the protein peaks eluted was determined by SDS-PAGE. These results showed that the protein fractions eluted at 8.4 and 11.2 ml retention volumes are pure *Pc*PAL proteins, presumably having different oligomerization states.



SEC200 PcPALAndi 30012015:10_UV2_215nm@02,PEAK1

No	Ret ml	Area mAU*ml	Height mAU
1	-2.02	0.0928	10.989
2	-1.86	37.0035	38.647
3	8.42	1152.3995	946.250
4	11.24	1191.1324	459.374
5	15.18	1512.1388	598.230
6	18.45	87.8232	148.133

Figure 8. Purification of *Pc*PAL on Superdex 200 10/300 GL size-exclusion column

In order to determine the molecular weight of the eluted fractions the calibration of the Superdex column was done, using the elution volume parameter, such as as K_{av}[5] of known calibration standards. Accordingly the molecular weight of the *Pc*PAL protein fractions eluted were estimated to be 355 kDa for the peak eluted at 11,2 mL and > 600 kDa for the fraction eluted at 8,4 ml, corresponding to the presumed native tetrameric and higher oligomerization states (octameric or even higher aggregated forms) of the protein, respectively.

Furthermore we compared the enzyme activities of the presumed largely aggregated and tetrameric forms of *Pc*PAL, separated by size-exclusion chromatography, towards the natural substrate L-Phe, observing three times higher activity for the presumed native, tetrameric form of the protein. These results directed us to optimize the isolation and purification procedure in order to achieve high yields for the production of the *Pc*PAL tetramer, introducing several modification in the isolation and purification protocol: *i*) the imidazole concentration used for the elution at the Ni-affinity chromatography was decreased, *ii*) the concentration steps of the solutions with high imidazole content was avoided, performing the protein concentration step after the removal of the imidazole with SEC chromatography or dialysis, *iii*) the use of protease inhibitor cocktail (mini-tablets from Roche) during all isolation and purification steps. These modifications provided an optimal isolation and purification procedure for the recombinant *Pc*PAL, resulting in protein with high purity in its native, tetrameric fold. This optimized procedure was also succesfully applied for the purification of the other MIO-enzymes from the collection, obtaining the tetramers (Fig. 9) in high purity (Fig.10).

During the project all succesfully cloned MIO-enzymes (Table 3) were isolated with high yields and purity, creating the collection of purified enzymes. Furthermore glycerol stocks of bacterial cells, containing the different plasmids are kept at -80 ^{o}C, representing the ready to use whole cell collection of MIO-enzymes. Since *Pa*PAM mutants (Table 3) could not be obtained through site-directed mutagenesis (WP1), therefore only the wt-PaPAM is present in both libraries.

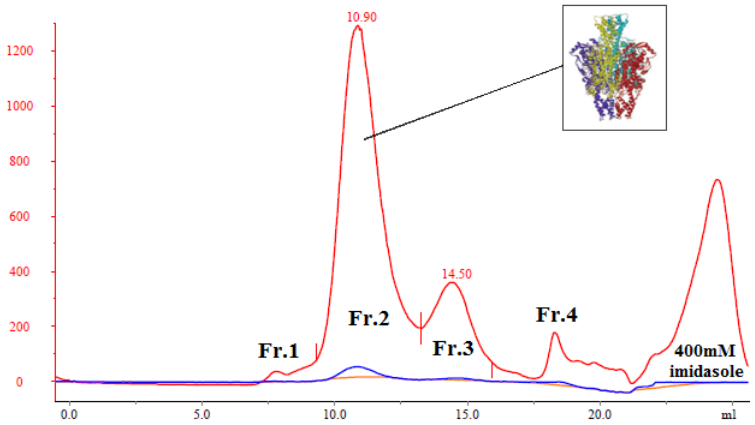


Figure 9. Chromatogram from *Tc*PAM purification by Superdex 200 10/300 GL size-exclusion column

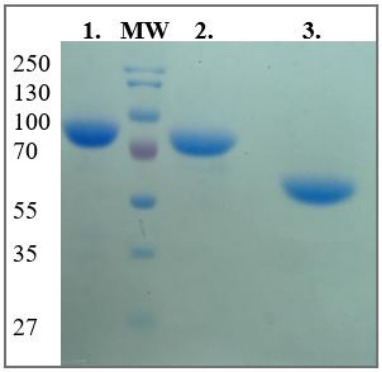


Figure 10. Representative SDS-PAGE gel containing samples from the purification steps of *Pza*PAM, *Tc*PAM and *Kk*PAL: lane **1**: *Pza*PAM fraction, lane **2**:*Tc*PAM fraction, lane **3**: *Kk*PAL fraction eluted from the SEC column at the molecular weight corresponding to the molecular weight of the tetramer

Ob2. Development of high throughput activity screening assay

WP1. Molecular cloning, production and purification of phenylacrylic acid decarboxilase system.

Phenolic acid decarboxilases[6] and phenylacrylic acid decarboxilases[7] (PAD) were found to transform a wide range of phenylacrylate analogues[8]. The mechanism of action of PADs is not elucidated, recent papers suggesting that the decarboxylation process in *A. niger* also requires 4-hydroxybenzoic acid decarboxylase (OhbA1) and the regulator SdrA, while in *S. cerevisiae* the phenylacrylic acid decarboxylase (PAD1) and the ferulic acid decarboxylase (FDC1) genes are essential for the decarboxylation process[9]. Recently the recombinant FDC1-PAD1 protein complex was patented[10] as efficient decarboxylation system for various phenylacrylic acid derivatives for producing valuable styrenes.

The plasmid pTfdc1Sc, containing the FDC1 gene of *Saccharomyces cerevisiae* in Ptpad1Sc vector was a generous gift from Prof. David R. Nielsen from the Arizona State University (USA).

The molecular cloning steps aiming to introduce the fdc1 gene is pET19b vector, were similar with those employed for novel enzymes *Kk*PAL, *Pza*PAM, described at Objective 1. Accordingly, the insert gene was PCR amplified using the primers:

FDC1_NdeI_FOR1: gaaacagcatatggaattcggg;

FDC1_REV1 : gatgcctggcagtttatggc.

The amplified gene, the PCR product was digested with restriction enzymes NdeI and BamHI, co- extracted with NdeI and BamHI digested pET19b(+) vector from the agarose gel and ligated. The ligation mixture was transformed and the grown colonies were picked up and the presence of the desired insert (fdc1) gene was proved by colony PCR. Therefore we sucesfully obtained the clone for the production of FDC1 with an *N*-terminal His-tag (**Figure 11**)

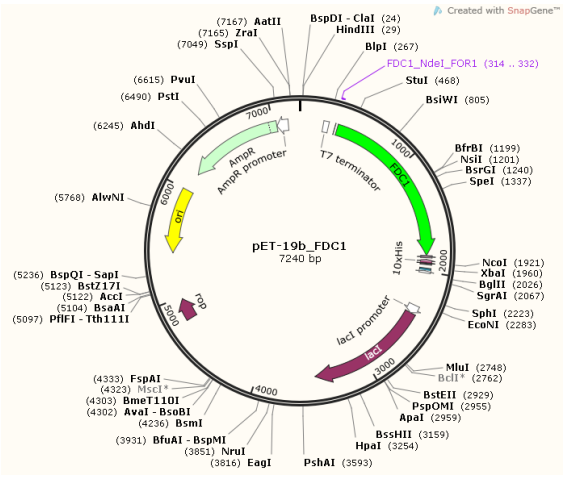


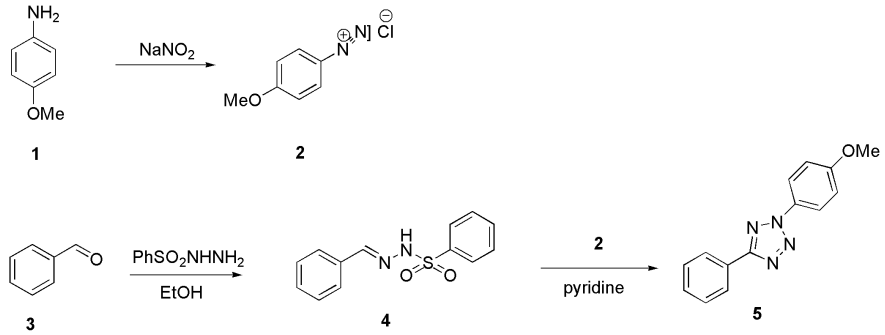
Figure 11.
Plasmid map of the pET-19b-fdc1.

Further we were interested in the expression the FDC1 enzyme in order determine its activity and substrate scope at the whole cell level.

Accordingly the recombinant plasmids (in pTfdc1Sc and pET19b_fdc1 vector) were transformed through heat-shock into *E.coli* BL21 DE3 pLysS competent cells. In the first round for the cell growth the optimal conditions described in the literature[11] were employed.

WP2) Synthesis of diaryltetrazole based probes

The synthesis of diaryltetrazole probe was performed according to existing procedures[12], using as starting material p-anisidine **1** and benzaldehyde **3** (**Scheme 1**).

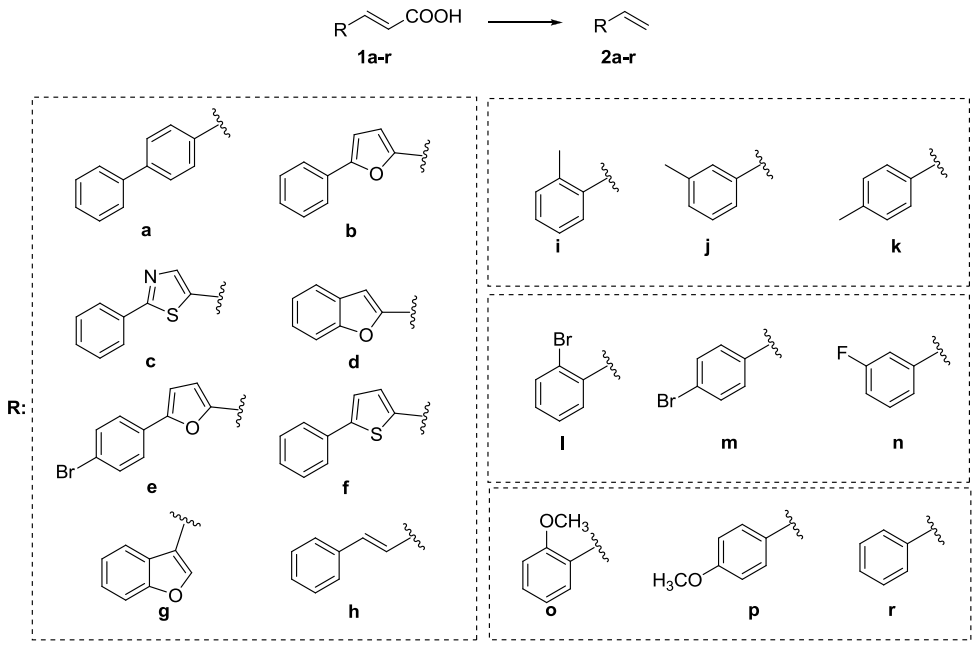


Scheme 1. The synthesis of diaryltetrazole probe **5**

The synthesized diaryltetrazole based product **5** was characterized by ^{1}H- and ^{13}C-NMR [^{1}H NMR (400 MHz, $CDCl_3$) δ 8.23 (dd, J= 7.9, 1.7 Hz, 2H), 8.09 (d, J= 9.1 Hz, 2H), 7.53 –7.47 (m, 3H), 7.04 (d, J= 9.1 Hz, 2H), 3.87 (s, 3H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.68, 159.20, 129.10, 127.60, 125.68, 120.09, 113.37, 54.35]

WP3) Activity assay set-up for the whole cell and purified protein MIO-enzyme kit.

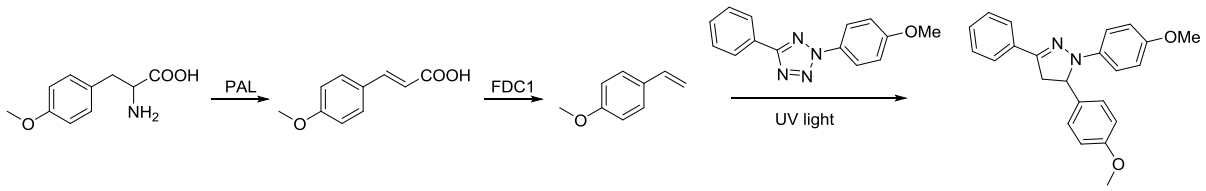
The decarboxylation of cinnamic acid **1r** resulting styrene **2r** is catalyzed by FDC1 **(Scheme 2)**, however the substrate scope of of FDC1 or/and their functionally similar PAD enzymes was not yet described[11]. In order to develop the high-throughput MIO-enzyme activity assay, based on the fluorescens detection of the styrene derivatives **2a**-**r**, we were interested whether the FDC1 enzyme is accepting unnatural, bulkier substrates in the decarboxylation process. Therefore we performed the decarboxylation of various acrylic acid derivatives **1a**-**r (Scheme 2)**, detecting the formation of the products **2a**-**r** by GC-MS.



Scheme 2. Activity tests for FDC1 using various unnatural acrylic acids in order to determine the substrate scope of the enzyme

In the majority of cases styrene derivatives could be detected succesfully, except in case of **1a** and **1e** where probably the sterical demand of the substrate is too high. Based on these results we aim to describe for the first time the substrate scope of ferrulic acid decarboxylase (FDC1).

Further we focused on the optimal conditions for the reaction of the produced styrene with diaryltetrazole probe, developing the structural and fluorescent analysis of the product formed. First we used as model substrate (pMeO-phenylalanine), and as model enzyme (*Pc*PAL), the optimal conditions for FDC1 activity and the fluorescent probe in order to test its functionality in 96-well deep plate format.



Scheme 3. The model reaction tested in microplate format

First the enzymatic reactions were performed with different isolated *Pc*PAL variants (I460A, I460V, wilde-type) and with whole cell ferrulic acid decarboxylase (FDC1) in 96-well deep plate format, using 2.5mM substrate concentration, Tris buffer (20 mM, pH 8) 50µg *Pc*PAL isolated enzymes followed by overnight incubated at 30˚C for PAL reation. Then FDC1 whole cells were grown and induced with IPTG and added to the reaction (20mg/mL) followed by for 4-6 h incubation, at 30°C. The styrene product was extracted in n-hexane and the n-hexane solution of the fluorescent diaryltetrazoleprobe diluted was added to the samples. The UV reaction was carried out at 302nm for 1min, and the appearance of the resulting fluorescent pyrazoline product was excitated 368 nm and at detected at 460nm emission wavelength by TECAN Spark 10M microplate reader. Excitation and emission spectra of the styrene (negative control) and of its reaction with fluoroprobe with and without UV was scanned by Shimadzu RF-6000 spectrofluorophotometer (see Figure 12-14). Excitation and emission maximum of the produced pyrazoline found to be at 368nm and 460nm).

The fluorescence signal and accordingly the enzyme activity is highest with PcPAL mutant I460V, and the signal intensity is significantly higher than those of the negative control (Figure 15). Further the PAL-activity of the same isolated enzymes were determined by our reported HPLC assay[13], developed within this project, which confirmed that indeed mutant I460V has the highest activity from the tested enzymes (Figure 16), therefore validating our assay performed with isolated PALs.

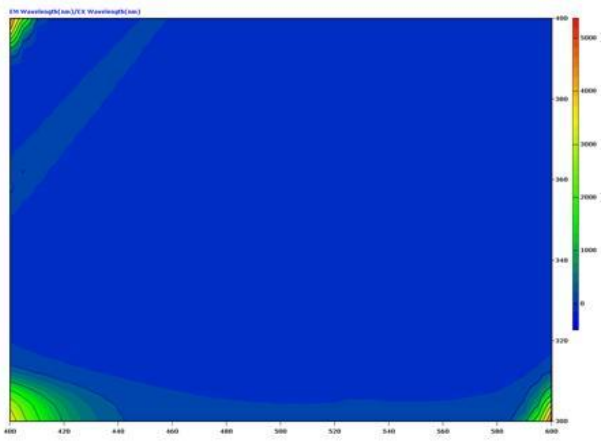


Figure 12. 3D-Excitation (OY axis) and emisson (OX axis) spectra of the *p*-MeO-phenylstyrene –no emission between 368 and 460 nm (negative control)

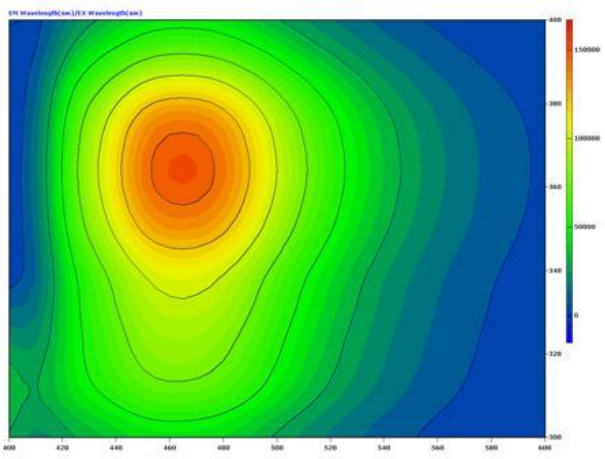


Figure 13. 3D-Excitation (OY axis) and emisson spectra (OX axis) pyrazoline fluorescent product, from reaction of the *p*-MeO-phenylstyrene and the diaryltetrazoleprobe with UV light. Excitation maximum at 368 nm and emission maxima at 460-470 nm.

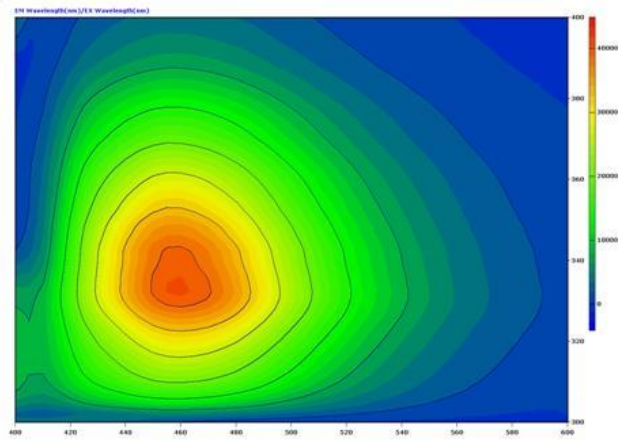


Figure 14. 3D-Excitation (OY axis) and emisson (OX axis) spectra of the of the *p*-MeO-phenylstyrene and the diaryltetrazoleprobe without UV light– excitation maximum at 325 nm, emission maximum at 450-460 nm (negative control 2)

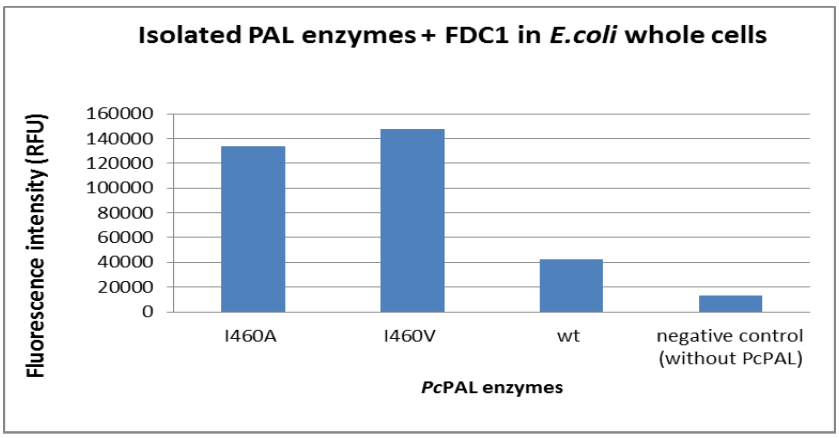


Figure 15. Fluorescence and enzyme activity with isolated *Pc*PAL enzymes

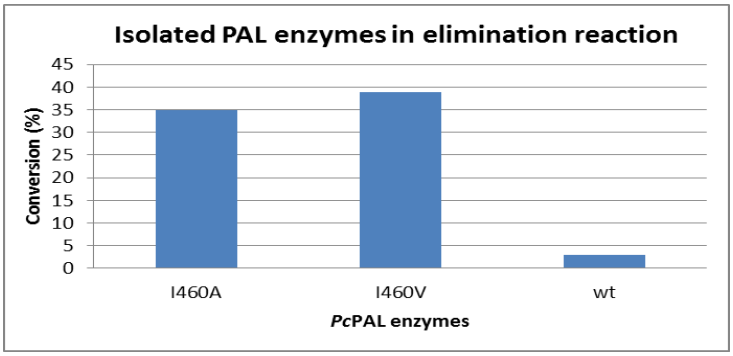


Figure 16. PAL activity determined by HPLC assay

In the next step the reactions were repeated using *Pc*PAL (I460A, I460V, wild-type) whole cells and the same reaction conditions as with the isolated PAL enzymes, only in this case the *Pc*PAL whole cells and the FDC1 whole cells were added in the same time and the reactions were incubated at 30˚C over night. Styrene product was extracted in n-hexane and the fluorescent diaryltetrazoleprobe diluted in hexane was added to the samples. UV reaction was carried out at 302nm, 1min, and the appearance of the resulting fluorescent pyrazoline product was excited 368nm and at detected at 460nm emission wavelength by TECAN Spark 10M microplate reader . The activity of the same PAL reactions were monitored by the HPLC assay, and the activity order among the different PAL enzymes tested was the same, and also correlated with the results obtained with the purified enzymes (Figure 15 and Figure 17). Accordingly PcPAL I460V cells proved to have the highest activity, followed by mutant F137V and wt-PcPAL. The background provided by the negative control (reactions performed without the presence of PAL cells) is somewhat higher than those observed in case isolated enzyme, probably due to the presence of whole cells, which upon lysis might provide metabolites that by reaction with fluoroprobe, or other UV-initiated reaction could generate fluorescent signals. However the background intensity is still considered low enough to unambigously detect PAL activity of the whole cells.

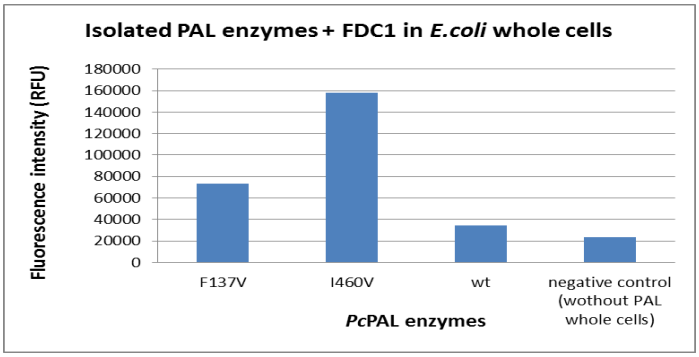


Figure 17. Fluorescence and enzyme activity with whole cell *Pc*PAL enzymes

All these results demonstrate the functionality of the developed assay for PAL activity screen in a HTS format, in 96 microtiter plates, of both isolated enzyme or whole cell collection. Further validation of the assay on the whole MIO-enzyme collection is undergoing and will lead to a high impact publication in the bioorganic chemistry field.

Ob 3. Exploring and broadening the substrate scope of the MIO-enzyme toolbox

WP1) Creation of substrate library for the MIO-enzyme toolbox

Several differently substituted phenylalanine analogues and its acrylic counterparts were available from our earlier works and/or from our collaborators, therefore we focused on the synthesis of novel substrates, such as stryrylalanines, biphenylalanines and phenyl-thiophene based compounds, or on poor PAL substrates, such as napthyl, *o*-, *m*-, *p*-methoxy and bromo substituted derivatives (Figure 18.).

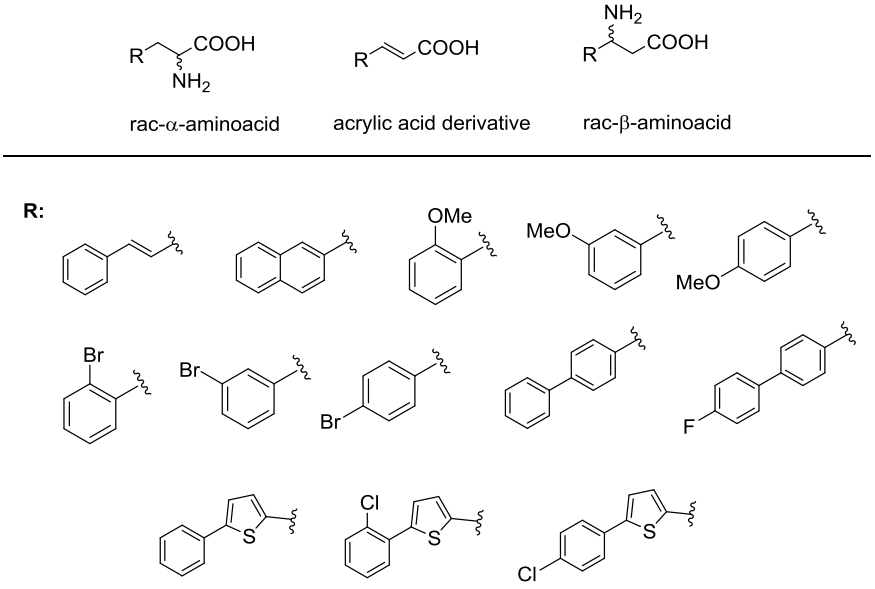
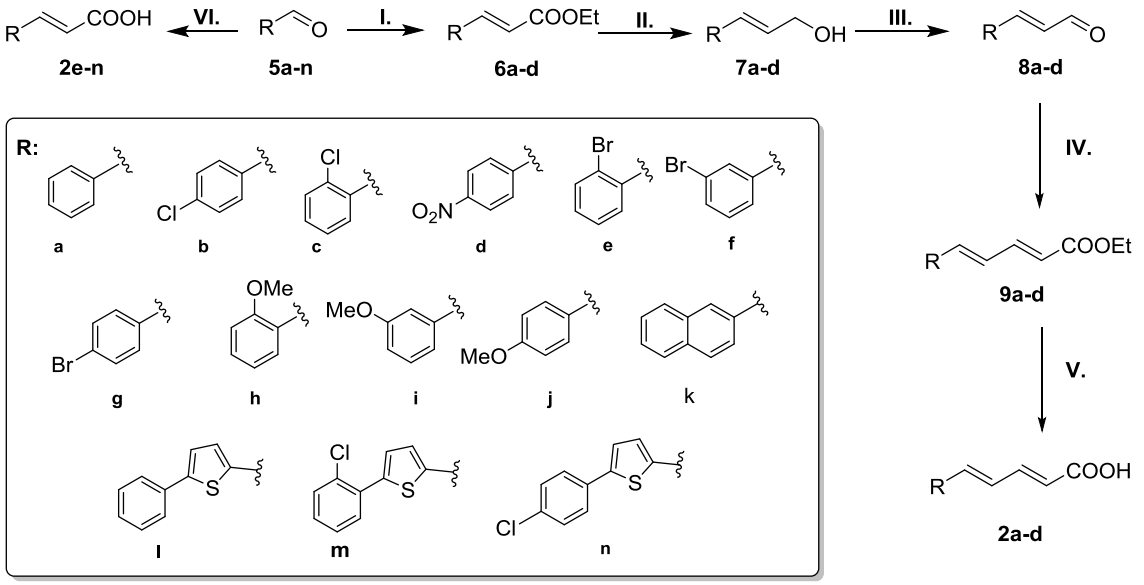


Figure 18. Synthesized never tested or poor substrates of MIO-enyzmes

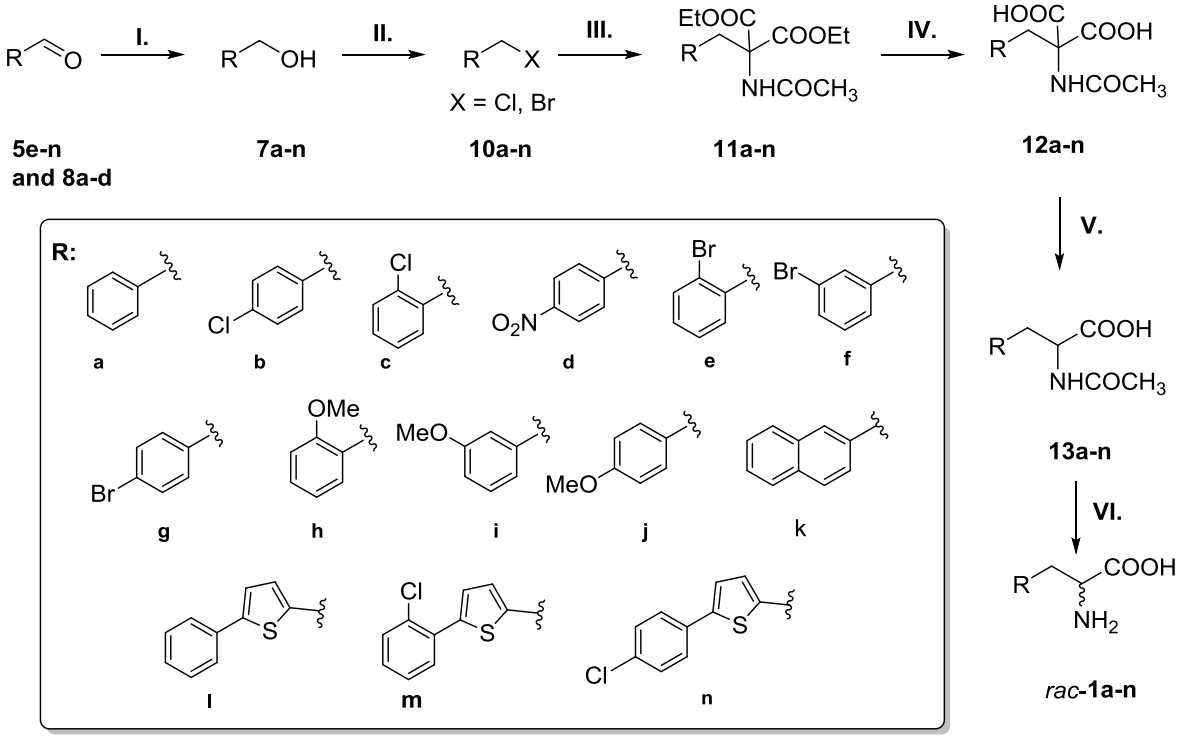
Theferore starting from the commercially available aldehydes **3a-d** acrylic esters **4a-d** were obtained through Wittig reaction, using the corresponding triphenyl-phosphoranylidine. The reduction of esters **4a-d** with DIBAL-H afforded the primary alcohol derivatives **5a-d**, which through an oxidation with manganese-dioxide, followed by another Wittig reaction with triphenyl-phosphoranylidene were converted into styrylacrylic ester derivatives **7a-d**, from which through mild alkaline hydrolysis the styrylic acrylates **2a-d** were obtained (**Scheme 4**). Acrylic acid derivatives **2e-n** were obtained through Knoevenagel-Dubner reaction (Scheme 3, reaction VI) starting from aldehydes **5e-n** using malonic acid as reactant and pyridine as solvent.

The synthesis of the corresponding racemic α-amino acids rac-**1a-n** started from the corresponding aldehydes **8a-d** and **5e-n**, followed by chemical reduction and the convertion of the alcohol derivatives **7a-n,** into the diethyl-acetamido malonate derivatives **11a-n** *via* malonic acid coupling of the brominated compounds **10a-n**. Further, through a mild alkaline hydrolysis of **11a-n**, followed by the subsequent decarboxylation of **12a-n**, the *N*-acylated amino acids *rac*-**13a-n** were obtained. Finally, deprotection of *rac*-**13a-n**afforded the racemic amino acids *rac*-**1a-n** (**Scheme 5**).



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

I. Ph_3P=CH-CO_2Et, toluene, reflux, 24h; **II.** DIBAL-H/CH_2Cl_2 at -60^{o}C, 1h; **III.** MnO_2/CH_2Cl_2, r.t, 48h; **IV.** Ph_3P=CH-CO_2Et, toluene, reflux, 24h; **V.** 10% KOH, reflux, 20h. **VI.** $CH_3(COOH)_2$, pyridine, 80˚C



Scheme 5. Synthesis of racemic amino acids *rac*-**1a-n.**

I. $NaBH_4$, MeOH, r.t; **II. A)** $(CH_3)_3SiCl$, NaBr /MeCN, reflux, 5h, 50^{o}C; **B)** *N*-bromosuccinimide, PPh_3, CH_2Cl_2, r.t; **III.** NaH, $CH_3CONHCH(COOEt)_2$/DMF, 60^{o}C, 3h; **IV.** 10% NaOH in water /MeOH, 60^{o}C, 5h; **V.** toluene, reflux, 20h; **VI.** dioxane/18% HCl, reflux, 4h.

- **Spectral analysis results of acrylic acid derivatives**

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

Yield: 76%, white solid; ^{1}H NMR (400 MHz, $CDCl_3$) 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). ^{13}C NMR (100 MHz, $CDCl_3$, δ): 172.55, 147.13, 141.80, 135.95, 129.47, 129.00, 127.50, 126.09, 120.41.

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

Yield: 70%, yellow-white powder; ^{1}H NMR (600 MHz, D_2O) 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); ^{13}C 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.

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

Yield: 73% %, white powder; ^{1}H 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). ^{13}C NMR (151 MHz, D2O): δ= 127.7, 127.8 128.2, 128.9, 133.4, 135.1, 136.6, 140.9, 175.8.

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

Yield: 72 % yellow-brown solid, ^{1}H NMR (600 MHz, D_2O) 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). ^{13}C NMR (151 MHz, D2O): 175.4, 146.6, 143.4, 140.1, 135.2, 131.5, 130.1, 127.4, 124.1.

(E)-3-(2-bromophenyl)acrylic acid (**2e**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.97 (d, *J* = 15.9 Hz, 1H), 7.69 (d, *J* = 7.8 Hz, 1H), 7.59 (d, *J* = 8.1 Hz, 1H), 7.33 (t, *J* = 7.1 Hz, 1H), 7.23 (t, *J* = 6.9 Hz, 1H), 6.41 (d, *J* = 15.9 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 169.64, 144.15, 135.55, 134.46, 132.61, 129.15, 129.06, 126.04, 122.38.

(E)-3-(3-bromophenyl)acrylic acid (**2f**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.65 (s, 1H), 7.60 – 7.37 (m, 3H), 7.21 (t, *J* = 7.8 Hz, 1H), 6.40 (d, *J* = 16.0 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 169.81, 144.41, 138.15, 134.06, 131.85, 131.72, 127.84, 123.91, 121.10.

(E)-3-(4-bromophenyl)acrylic acid (**2g**)

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.68 – 7.51 (m, 5H), 6.56 (d, *J* = 16.0 Hz, 1H). ^{13}C NMR (101 MHz, DMSO) δ 167.49, 142.67, 133.57, 131.90, 130.21, 123.59, 120.17.

(E)-3-(2-methoxyphenyl)acrylic acid (**2h**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.94 (d, *J* = 16.1 Hz, 1H), 7.53 (d, *J* = 7.8 Hz, 1H), 7.34 (t, *J* = 7.1 Hz, 1H), 6.99 (d, *J* = 8.3 Hz, 1H), 6.93 (t, *J* = 7.5 Hz, 1H), 6.47 (d, *J* = 16.2 Hz, 1H), 3.86 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 170.92, 159.74, 141.62, 132.88, 129.66, 124.30, 121.80, 119.34, 112.39, 56.05, 49.64, 49.43, 49.21, 49.00, 48.79, 48.57, 48.36.

(E)-3-(3-methoxyphenyl)acrylic acid (**2i**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.53 (d, *J* = 16.0 Hz, 1H), 7.20 (t, *J* = 7.9 Hz, 1H), 7.12 – 6.95 (m, 2H), 6.86 (d, *J* = 8.2 Hz, 1H), 6.36 (d, *J* = 16.0 Hz, 1H), 3.71 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 170.06, 161.29, 146.04, 136.91, 130.74, 121.48, 119.34, 117.01, 113.78, 55.54, 49.43, 49.21, 49.00, 48.79, 48.57, 48.36, 48.15.

(E)-3-(4-methoxyphenyl)acrylic acid (**2j**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.56 (d, J = 16.0 Hz, 1H), 7.46 (d, J = 8.9 Hz, 2H), 6.88 (d, J = 8.9 Hz, 2H), 6.26 (d, J = 15.9 Hz, 1H), 3.75 (s, 2H). ^{13}C NMR (101 MHz, MeOD) δ 170.78, 163.06, 146.21, 130.90, 128.35, 116.52, 115.38, 55.83, 49.64, 49.43, 49.21, 49.00, 48.79, 48.57, 48.36.

(E)-3-(naphthalen-2-yl)acrylic acid(**2k**)

^{1}H NMR (600 MHz, DMSO-d_6): 8.02 – 7.76 (m, 4H), 7.71 (d, J = 8.3 Hz, 1H), 7.48 (t, J = 6.9 Hz, 2H), 7.27 (d, J = 15.8 Hz, 1H), 6.53 (d, J = 15.9 Hz, 1H); ^{13}C NMR (151 MHz, DDMSO-d_6): 170.49, 135.35, 134.43, 133.26, 132.77, 130.91, 128.13, 127.98, 127.57, 126.92, 126.37, 126.07, 124.04.

(E)-3-(5-phenylthiophen-2-yl)acrylic acid(**2l**)

^{1}H NMR (600 MHz, DMSO-d_6): 7.65 (d, J = 7.6 Hz, 2H), 7.56 – 7.34 (m, 3H), 7.30 (t, J = 7.3 Hz, 1H), 7.23 – 6.92 (m, 2H), 6.12 (d, J = 15.6 Hz, 1H); ^{13}C NMR (151 MHz, DMSO-d_6): 169.25, 142.17, 141.51, 133.58, 130.71, 129.16, 129.04, 128.73, 127.95, 127.79, 127.76, 125.43, 125.27, 125.14, 124.33, 39.94, 39.80, 39.66, 39.52, 39.38, 39.24, 39.10.

(E)-3-(5-(2-chlorophenyl)thiophen-2-yl)acrylic acid (**2m**)

^{1}H NMR (400 MHz, D_2O): 7.07 (d, J = 15.7 Hz, 1H), 6.75 – 6.30 (m, 6H), 5.87 (d, J = 15.7 Hz, 1H); ^{13}C NMR (101 MHz, D_2O): 174.71, 140.65, 140.41, 133.03, 131.79, 130.96, 130.59, 129.93, 129.59, 128.16, 126.61, 123.58;HRMS: $[M-H]^-$ found 262.9950 (calculated: 262.9939 for $C_{13}H_8ClO_2S^-$; Δm_i = 4.18 ppm)

(E)-3-(5-(4-chlorophenyl)thiophen-2-yl)acrylic acid (**2n**)

^{1}H NMR (600 MHz, Acetic Acid-d_4): 7.90 (d, J = 15.6 Hz, 1H), 7.67 (d, J = 8.5 Hz, 2H), 7.46 – 7.36 (m, 4H), 6.30 (d, J = 15.7 Hz, 1H); ^{13}C NMR (151 MHz, Acetic Acid-d_4): 171.53, 146.56, 139.10, 138.69, 134.04, 133.51, 132.09, 129.09, 127.08, 124.64, 115.43

- **Spectral analysis results of racemic α-amino acids**

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

^{1}H 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 ($[M+H]^+$, (calculated for $C_{11}H_{13}NO_2S$: 192.0946 ($[M+H]^+$).

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

^{1}H 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 ($[M+H]^+$, ^{35}Cl), 228.1 ($[M+H]^+$, ^{37}Cl) (calculated for $C_{11}H_{12}ClNO_2$ 226.0557 ($[M+H]^+$, ^{35}Cl), 228.0527 ($[M+H]^+$, ^{37}Cl).

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

^{1}H 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, D2O): 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 ($[M+H]^+$, ^{35}Cl), 228.1 ($[M+H]^+$, ^{37}Cl) (calculated for $C_{11}H_{12}ClNO_2$ 226.0557 ($[M+H]^+$, ^{35}Cl), 228.0527 ($[M+H]^+$, ^{37}Cl).

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

^{1}H 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,

D2O): 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 ($[M+H]^+$) (calculated for $C_{11}H_{12}N_2O_4$ 236.0797 ($[M+H]^+$)

rac- 2-amino-3-(2-bromophenyl)propanoic acid (*rac*-**1e**)

^{1}H NMR (600 MHz, Deuterium Oxide) δ 7.65 (d, *J* = 8.5 Hz, 1H), 7.29 (d, *J* = 60.4 Hz, 2H), 4.35 (s, 1H), 3.37 (d, *J* = 140.8 Hz, 1H). ^{13}C NMR (151 MHz, D_2O) δ 171.35, 133.64, 133.26, 131.79, 129.87, 128.24, 124.22, 52.88, 36.26.

rac- 2-amino-3-(3-bromophenyl)propanoic acid (*rac*-**1f**)

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.62 – 7.37 (m, 2H), 7.36 – 7.17 (m, 2H), 3.94 (dd, *J* = 7.9, 5.3 Hz, 1H), 3.21 (dd, *J* = 14.5, 5.3 Hz, 1H), 3.06 (dd, *J* = 14.5, 7.9 Hz, 1H). ^{13}C NMR (101 MHz, D_2O) δ 173.48, 137.37, 132.04, 130.64, 130.56, 128.09, 122.18, 55.72, 35.87.

rac-2-amino-3-(4-bromophenyl)propanoic acid (*rac*-**1g**)

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.61 (s, 2H), 7.27 (s, 2H), 4.34 (s, 1H), 3.28 (d, *J* = 34.4 Hz, 2H). ^{13}C NMR (101 MHz, D_2O) δ 171.40, 133.17, 132.04, 131.21, 121.31, 53.98, 35.06.

rac- 2-amino-3-(2-methoxyphenyl)propanoic acid (*rac*-**1h**)

^{1}H NMR (400 MHz, Deuterium Oxide) δ 7.37 (ddd, *J* = 8.4, 7.5, 1.7 Hz, 1H), 7.23 (dd, *J* = 7.5, 1.7 Hz, 1H), 7.11 – 6.91 (m, 2H), 4.31 (dd, *J* = 7.2, 5.5 Hz, 1H), 3.83 (s, 3H), 3.36 (dd, *J* = 14.4, 5.6 Hz, 1H), 3.15 (dd, *J* = 14.4, 7.2 Hz, 1H). ^{13}C NMR (101 MHz, D_2O) δ 171.74, 157.49, 131.46, 129.75, 121.99, 121.04, 111.23, 55.14, 53.34, 30.83.

rac- 2-amino-3-(3-methoxyphenyl)propanoic acid (*rac*-**1i**)

^{1}H NMR (600 MHz, DMSO-d_6) δ 7.23 (t, *J* = 7.9 Hz, 1H), 6.99 – 6.74 (m, 3H), 4.13 (t, *J* = 6.2 Hz, 1H), 3.74 (s, 3H), 3.13 (d, *J* = 6.3 Hz, 2H). ^{13}C NMR (151 MHz, DMSO) δ 170.23, 159.28, 136.43, 129.56, 121.69, 115.22, 112.74, 54.98, 53.11, 39.94, 39.80, 39.66, 39.52, 39.38, 39.24, 39.10, 35.59.

rac- 2-amino-3-(4-methoxyphenyl)propanoic acid (*rac*-**1j**)

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.21 – 7.08 (m, 2H), 6.90 – 6.76 (m, 2H), 4.12 (dd, *J* = 7.6, 5.4 Hz, 1H), 3.70 (s, 3H), 3.17 (dd, *J* = 14.6, 5.4 Hz, 1H), 3.03 (dd, *J* = 14.6, 7.6 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 169.88, 159.43, 130.20, 125.82, 114.10, 54.33, 53.85, 48.25, 48.04, 47.82, 47.61, 47.40, 47.18, 46.97, 35.08.

rac-2-amino-3-(naphthalen-2-yl)propionic acid *(rac*-**1k**)

^{1}H NMR (600 MHz, D_2O): 7.95 (ddd, *J* = 16.4, 8.8, 5.9 Hz, 3H), 7.84 (s, 1H), 7.65 – 7.52 (m, 2H), 7.47 (dd, *J* = 8.4, 1.7 Hz, 1H), 4.33 (dd, *J* = 7.8, 5.6 Hz, 1H), 3.51 (dd, *J* = 14.6, 5.5 Hz, 1H), 3.36 (dd, *J* = 14.6, 8.0 Hz, 1H); ^{13}C NMR (151 MHz, D_2O): 172.23, 133.17, 132.44, 132.08, 128.88, 128.37, 127.73, 127.67, 127.10, 126.75, 126.48, 54.63, 36.05

rac-2-amino-3-(5-phenylthiophen-2-yl)propanoic acid *(rac*-**1l**)

^{1}H NMR (600 MHz, 5% NaOD): 7.73 – 7.56 (d, *J* = 7.2 Hz, 2H), 7.47 – 7.36 (m, 2H), 7.36 – 7.21 (m, 2H), 6.87 (dd, *J* = 3.5, 1.8 Hz, 1H), 3.54 – 3.39 (t, *J* =5.1 Hz, 1H), 3.21 – 3.02 (m, 2H); ^{13}C NMR (151 MHz, 5% NaOD):181.7, 142.4, 140.4, 134.1, 129.3, 127.7, 127.6, 125.4, 123.4, 57.2, 35.2; HRMS: $[M+H]^+$ found 248.0747 (calculated: 248.0740 for $C_{13}H_{14}NO_2S^+$; Δm_i = 2.82 ppm)

rac-2-amino-3-(5-(2-chlorophenyl)thiophen-2-yl)propanoic acid *(rac*-**1m**)

^{1}H NMR (600 MHz, 5% NaOD): 7.27 (m, 1H), 7.18 (m, 1H), 7.00 (m, 2H), 6.94 (m, 1H), 6.69 (m, 1H), 3.33 (t, *J* =7.8 Hz, 1H), 2.81-3.06 (m, 2H); ^{13}C NMR (151 MHz, 5% NaOD):181.44, 141.82, 138.32, 132.45, 131.16, 130.97, 130.28, 128.73, 127.88, 127.21,

126.4, 57.26, 35.37; HRMS: $[M-H]^-$ found 280.0214 (calculated: 280.0205 for $C_{13}H_{11}ClNO_2S^-$; Δm_i = 3.21 ppm)

rac-2-amino-3-(5-(4-chlorophenyl)thiophen-2-yl)propanoic acid *(rac*-**1n**)

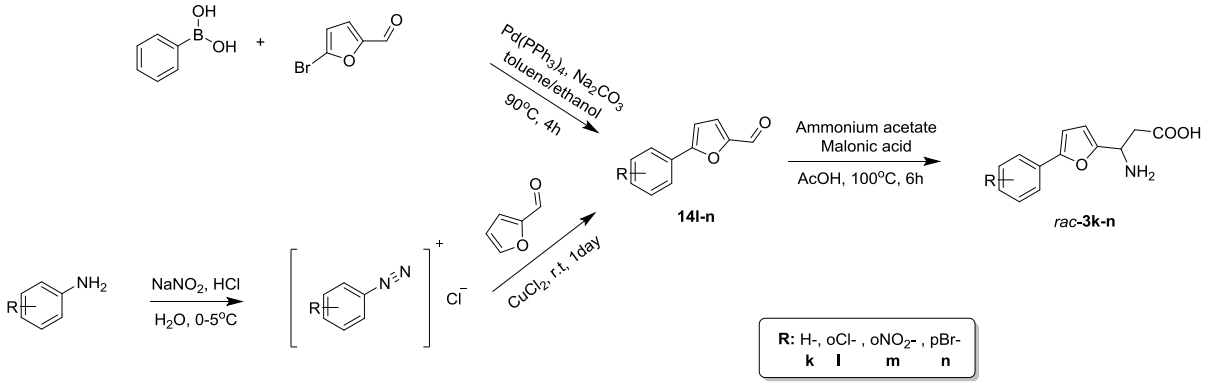
^{1}H NMR (600 MHz, 5% NaOD) δ 7.00 (d, *J* = 8.2 Hz, 2H), 6.82 (d, *J* = 8.2 Hz, 2H), 6.71 (d, *J* = 3.6 Hz, 1H), 6.42 (d, *J* = 3.7 Hz, 1H), 3.53 (t, *J* = 6.0 Hz, 1H), 2.93 (dd, *J* = 15.9, 5.0 Hz, 1H), 2.87 (dd, *J* = 15.7, 7.1 Hz, 1H);^{13}C NMR (151 MHz, DMSO-d_6): 170.26, 141.82, 136.50, 132.85, 132.31, 129.47, 129.35, 127.05, 124.74, 53.21, 30.46; HRMS: $[M-H]^-$ found 280.0212 (calculated: 280.0205 for $C_{13}H_{11}ClNO_2S^-$; Δm_i = 2.49 ppm)

- **Synthesis of racemic β-amino acids and their esters**

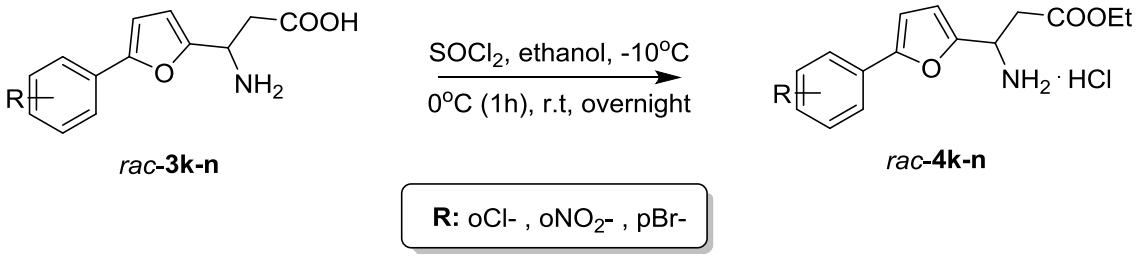
Based on the analogy with PcPAL catalyzed biotransformations of phenylfurane derivatives[14] we were interested whether TcPAM accespt phenylfurane based β-amino acids as substrates.

The synthesis of the corresponding β-amino acids *rac*-**3k-n** were obtained via the synthesis of the corresponding aldehydes **14l-n** and the reaction of aldehydes with malonic acid and ammonium carbamate gave the products *rac*-**3k-n**. (Scheme 6)

The synthesis of the racemic ethyl 3-amino-3-(5-phenylfuran-2-yl)propanoate hydrochlorides rac-**4k-n**·HCl were obtained by the reaction of β-amino acids *rac*-**3k-n** with $SOCl_2$ in anhydrous ethanol (Scheme 7).



Scheme 6. Synthesis of β-amino acids



Scheme 7. *Synthesis of racemic ethyl 3-amino-3-(5-phenylfuran-2-yl)propanoate hydrochlorides rac-***4k-n**·*HCl*

- **Spectral analysis results of racemic β-amino acids and their ester derivatives**

rac- 3-amino-3-(5-phenylfuran-2-yl)propanoic acid (**3k**)

1H NMR (400 MHz, D2O) δ 7.81 – 7.76 (m, 2H), 7.50 (t, J = 7.7 Hz, 2H), 7.38 (t, J = 7.5 Hz, 1H), 6.80 (d, J = 3.4 Hz, 1H), 6.39 (d, J = 3.3 Hz, 1H), 4.38 (t, J = 7.2 Hz, 1H), 2.76 (dd, J = 15.0, 7.1 Hz, 1H), 2.66 (dd, J = 15.0, 7.4 Hz, 1H); 13C NMR (101 MHz, D2O) δ 179.8,

157.2, 152.5, 130.5, 129.1, 127.6, 123.5, 107.0, 106.1, 46.8, 43.9. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C13H11O3: 215.07027; found, 215.07001.

rac- 3-amino-3-(5-(2-chlorophenyl)furan-2-yl)propanoic acid (**3l**)

1H NMR (400 MHz, D2O) δ 7.84 (dd, J = 7.9, 1.7 Hz, 1H), 7.47 (dd, J = 8.0, 1.3 Hz, 1H), 7.36 (td, J = 7.7, 1.4 Hz, 1H), 7.24 (td, J = 7.7, 1.7 Hz, 1H), 7.06 (d, J = 3.4 Hz, 1H), 6.36 (d, J = 3.4 Hz, 1H), 4.31 (t, J = 7.2 Hz, 1H), 2.67 (dd, J = 15.0, 7.1 Hz, 1H), 2.57 (dd, J = 15.0, 7.4 Hz, 1H); 13C NMR (101 MHz, D2O) δ 179.6, 157.4, 148.9, 130.6, 129.3, 128.5, 128.4, 127.8, 127.2, 111.5, 106.8, 46.7, 43.8. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C13H10ClO3: 249.03130; found, 249.03106.

rac- 3-amino-3-(5-(2-nitrophenyl)furan-2-yl)propanoic acid (**3m**)

1H NMR (400 MHz, D2O) δ 7.62 (d, J = 8.1 Hz, 1H), 7.59 – 7.47 (m, 2H), 7.34 (t, J = 7.9 Hz, 1H), 6.55 (d, J = 3.5 Hz, 1H), 6.26 (d, J = 3.4 Hz, 1H), 4.19 (t, J = 7.2 Hz, 1H), 2.57 (dd, J = 15.3, 6.2 Hz, 1H), 2.45 (dd, J = 15.2, 8.0 Hz, 1H); 13C NMR (151 MHz, D2O) δ 179.3, 158.9, 147.2, 146.3, 132.7, 128.7, 128.5, 123.9, 123.4, 110.4, 107.0, 46.7, 43.6. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C13H10NO5: 260.05535; found, 260.05512.

rac- 3-amino-3-(5-(4-bromophenyl)furan-2-yl)propanoic acid (**3n**)

1H NMR (600 MHz, D2O) δ 7.53 (t, J = 6.7 Hz, 1H), 6.68 (d, J = 3.3 Hz, 0H), 6.29 (d, J = 3.3 Hz, 0H), 4.28 (t, J = 7.4 Hz, 0H), 2.65 (dd, J = 15.0, 7.0 Hz, 0H), 2.55 (dd, J = 15.0, 7.5 Hz, 0H); 13C NMR (151 MHz, D2O) δ 179.4, 157.4, 151.4, 131.6, 129.3, 124.8, 120.5, 106.8, 106.4, 46.7, 43.6. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C13H10BrO3: 292.98078; found, 292.98065.

rac- 3-amino-3-(5-phenylfuran-2-yl)propanoate hydrochloride (**4k**)

1H NMR (600 MHz, D2O) δ 7.77 (dd, J = 8.4, 1.2 Hz, 2H), 7.49 (t, J = 7.8 Hz, 2H), 7.42 – 7.36 (m, 1H), 6.85 (d, J = 3.5 Hz, 1H), 6.68 (d, J = 3.3 Hz, 1H), 5.01 (t, J = 7.1 Hz, 1H), 4.19 (qd, J = 7.2, 1.0 Hz, 2H), 3.27 (dd, J = 16.8, 7.3 Hz, 1H), 3.23 (dd, J = 16.8, 6.9 Hz, 1H), 1.20 (t, J = 7.1 Hz, 3H); 13C NMR (151 MHz, D2O) δ 171.4, 154.8, 147.2, 129.8, 129.1, 128.4, 123.9, 112.1, 106.3, 62.6, 45.2, 35.8, 13.2. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C15H15O3: 243.10157; found, 243.10128.

rac-3-amino-3-(5-(2-chlorophenyl)furan-2-yl)propanoate hydrochloride (**4l**)

1H NMR (600 MHz, D2O) δ 7.82 (dt, J = 8.0, 1.4 Hz, 1H), 7.53 (d, J = 8.5 Hz, 1H), 7.41 (t, J = 7.6 Hz, 1H), 7.36 – 7.30 (m, 1H), 7.12 (d, J = 3.3 Hz, 1H), 6.70 (d, J = 3.5 Hz, 1H), 5.00 (t, J = 7.1 Hz, 1H), 4.17 (q, J = 7.0 Hz, 2H), 3.36 – 3.13 (m, 2H), 1.18 (t, J = 7.1 Hz, 3H); 13C NMR (151 MHz, D2O) δ 171.2, 151.3, 147.3, 130.7, 129.9, 129.3, 128.1, 127.9, 127.3, 111.8, 111.6, 62.5, 45.0, 35.7, 13.1. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C15H14ClO3: 277.06260; found, 243.06240.

rac-3-amino-3-(5-(2-nitrophenyl)furan-2-yl)propanoate hydrochloride (**4m**)

1H NMR (600 MHz, D2O) δ 7.88 (d, J = 8.1 Hz, 1H), 7.76 – 7.69 (m, 2H), 7.58 (ddd, J = 8.5, 7.0, 1.8 Hz, 1H), 6.80 (d, J = 3.6 Hz, 1H), 6.71 (d, J = 3.5 Hz, 1H), 4.96 (t, J = 7.1 Hz, 1H), 4.17 (q, J = 7.2 Hz, 2H), 3.15 (d, J = 7.1 Hz, 2H), 1.19 (t, J = 7.1 Hz, 3H); 13C NMR (151 MHz, D2O) δ 170.9, 149.9, 148.8, 146.7, 133.1, 129.7, 129.5, 124.3, 123.0, 112.1, 110.5, 62.5, 44.8, 35.5, 13.1. HRMS-ESI (m/z) [M+H]+ Calcd for C15H17N2O5: 305.11320; found, 305.11305.

rac- -3-amino-3-(5-(4-bromophenyl)furan-2yl)propanoate hydrochloride (**4n**)

1H NMR (600 MHz, D2O) δ 7.60 (s, 4H), 6.81 (d, J = 3.5 Hz, 1H), 6.65 (d, J = 3.5 Hz, 1H), 4.98 (t, J = 7.1 Hz, 0H), 4.17 (q, J = 7.1 Hz, 2H), 3.33 – 3.16 (m, 2H), 1.19 (t, J = 7.1 Hz, 3H);

13C NMR (151 MHz, D2O) δ 171.3, 153.9, 147.4, 131.9, 128.7, 125.4, 121.4, 112.1, 106.8, 62.5, 45.1, 35.7, 13.2. HRMS-ESI (m/z) [M+H-NH2]+ Calcd for C15H14BrO3: 321.01208; found, 321.01160.

WP 2,3,4) Activity of MIO-enzyme toolbox, docking studies

- **Activity of PALs**

In frame of our activity to create a library of recombinant MIO-enzymes consisting of PALs with enhanced substrate scope, our first activity study focused on expanding the substrate acceptance of *Pc*PAL towards styrylalanine derivatives which are synthetically challenging phenylalanine analogues with enlarged distance between the aromatic moiety and the stereogenic α-carbon.[15] First, the commercially available L-styrylalanine L-**1a** was tested as substrate in the ammonia elimination reaction catalyzed by wild-type *Pc*PAL (wt-*Pc*PAL) using our HPLC, or newly fluorescent assays (**Scheme 8**). By the ^{1}H-NMR spectra, formation of (2*E*,4*E*)-styrylacrylate **2a** as product was observed in the ammonia elimination reaction. Kinetic studies with L-styrylalanine L-**1a** resulted in 14.7 fold decreased k_{cat} value in comparison with the ammonia elimination reaction from the natural substrate L-Phe (**Table 4**).

The low enzyme activity observed with L-styrylalanine L-**1a** was rationalized by molecular modeling studies using induced-fit covalent docking, performed using the corresponding *N*-MIO intermediate states of L-Phe and L-styrylalanine within the active center of *Pc*PAL but containing the Tyr110-loop region in catalytically active loop conformation corrected by partial homology modelling.[16] The modeling studies suggested that the reaction rate might be improved by reducing the steric clash between the aromatic moieties of the substrates and the F137 residue in the hydrophobic binding pocket (**Figure 18**).

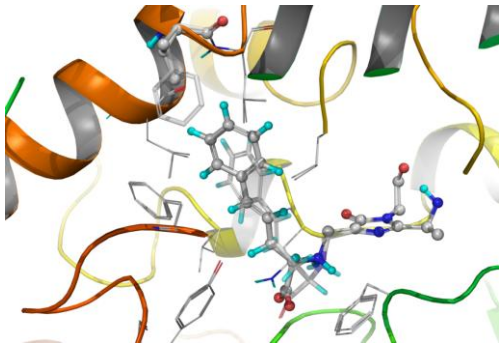


Figure 18. The docking studies of styryl alanine *rac*-**1a** using the homology model of *Pc*PAL suggested that residue F137 directs towards the phenyl ring of the substrate and sterical hindrance might occur, which force the substrate in conformations, where the MIO-group distances from the NH_2 group of the substrate (in front – L-styrylalanine-*N*-MIO intermediate and mutant enzyme F137V; in back, dashed lines – L-phenylalanine-*N*-MIO intermediate and wild-type *Pc*PAL)

Therefore, to provide more space for the aromatic moiety of L-styrylalanine, the mutant PcPAL library was screened and F137V, F137A provided the most intensive fluorescence signals, suggestinh highest enzyme activity.

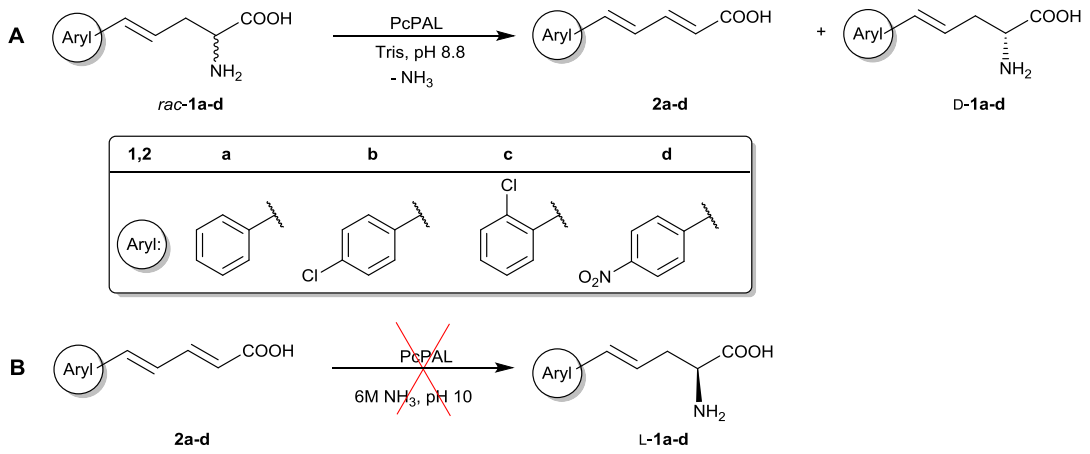
Therefore in the next step the kinetic parameters of wild-type and mutant *Pc*PALs were determined for the ammonia elimination reaction of natural substrate L-Phe and for L-styrylalanine L-**1a** (Table 4). Not surprising that all mutants showed decreased catalytic efficiency in the ammonia elimination reaction of the natural substrate L-Phe in comparison with the wt-*Pc*PAL. However, it is clearly shown that F137V had superior catalytic efficiency for L-**1a** as substrate with nearly 240-fold increased k_{cat}/K_M value compared to the wt-*Pc*PAL.

The F137A mutation also provided increased k_{cat}/K_M value with L-**1a** compared to wt-*Pc*PAL (10.5-fold enhancement), while the F137G mutation showed no beneficial effect on the transformation of L-**1a**.

Table 4. Kinetic parameters for wild-type *Pc*PAL and for F137X *Pc*PAL mutants in ammonia elimination from L-phenylalanine and from L-styrylalanine

*Pc*PAL	L-Phenylalanine			L-Styrylalanine		
	K_M (μM)	$k_{cat} \times 10^{-3}$ (s^{-1})	$k_{cat}/K_M \times 10^{-3}$ (nM^{-1} s^{-1})	K_M (μM)	$k_{cat} \times 10^{-3}$ (s^{-1})	$k_{cat}/K_M \times 10^{-3}$ (nM^{-1} s^{-1})
wt	83 ± 5	694 ± 20	8310 ± 291	4384 ± 158	47.2 ± 0.3	10.7 ± 0.4
F137V	86 ± 10	173 ± 1	2015 ± 131	186 ± 6	422 ± 28	2270 ± 168
F137A	1732 ± 15	283 ± 1	163 ± 2	1173 ± 70	132 ± 2.6	112.2 ± 7.2
F137G	4969 ± 153	52 ± 3	10.4 ± 0.9	4120 ± 270	34.5 ± 3.8	8.4 ± 0.2

Next, we wanted to explore the usefulness of F137V *Pc*PAL in the synthetically valuable ammonia elimination and addition reactions of various styrylalanines *rac*-**1a-d** and their acrylic counterparts **2a-d**, respectively (Scheme 8). The enzymatic ammonia elimination reactions (Scheme 8A) were performed in Tris buffer at pH 8.8, while 6M ammonia at pH 10 as reaction medium was applied to test the enzymatic ammonia addition (Scheme 8B). In all cases, the progress of the reactions was monitored by HPLC. The styrylalanines *rac*-**1a-d** were all accepted as substrates by F137V *Pc*PAL as well as by wt-*Pc*PAL in the ammonia eliminations. Regrettably, no product formation could be observed in the attempted ammonia addition reactions of styrylacrylates **2a-d** with either wt-*Pc*PAL or F137V *Pc*PAL.



Scheme 8. The (A) ammonia elimination reactions from racemic styrylalanines **1a-d** and the (B) attempted ammonia addition reactions onto styrylacrylates **2a-d** catalyzed by wt- and mutant *Pc*PALs

The kinetic parameters for the ammonia elimination reaction of *rac*-**1a-d**, with PAL and wt- and F137V *Pc*PAL were determined by UV-based enzymatic assays (Table 5). The fact that kinetic parameters of L-styrylalanine (Table 4) and racemic styrylalanine *rac*-**1a** (Table 5) were comparable with F137V *Pc*PAL supported that D-styrylalanine was not a

significant inhibitor. In case of all substrates *rac*-**1a-d**, the F137V mutant showed highly improved catalytic activity (k_{cat}/K_M) in comparison with the wild-type enzyme (Table 6).

Table 5. Kinetic parameters for wt- and F137V *Pc*PALs in the ammonia elimination of racemic styrylalanines *rac*-**1a-d**

Substrate	wt-*Pc*PAL			F137V *Pc*PAL		
	K_M (μM)	$k_{cat} \times 10^{-3}$ (s^{-1})	$k_{cat}/K_M \times 10^{-3}$ ($nM^{-1}\ s^{-1}$)	K_M (μM)	$k_{cat} \times 10^{-3}$ (s^{-1})	$k_{cat}/K_M \times 10^{-3}$ ($nM^{-1}\ s^{-1}$)
rac-**1a**	395 ± 6	6.2 ± 0.7	16.1 ± 0.9	200 ± 12	276 ± 13	1373 ± 97
rac-**1b**	154 ± 7	0.24 ± 0.01	1.6 ± 0.03	78.3 ± 12	78.6 ± 4.4	1004 ± 86
rac-**1c**	28 ± 1	0.34 ± 0.02	11.9 ± 0.6	94.7 ± 2	156.7 ± 0.1	1660 ± 10
rac-**1d**	287 ± 3	3.2 ± 0.03	10.9 ± 0.5	326 ± 10	9.9 ± 0.01	30.2 ± 1.9

To confirm the synthetic usefulness of mutant F137V towards styrylalanines, the ammonia elimination reactions were performed with F137V *Pc*PAL and *wt-Pc*PAL on semi-preparative scale (0.1 mmol substrate at 5 mM substrate concentration with 0.5-1 mg purified enzyme; Table 6). The conversion of the *Pc*PAL-catalyzed reactions from racemic styrylalanines *rac*-**1a-d** (Figure 19) and enantiomeric excess of the residual substrates D-**1a-d** (Table 6) were monitored by HPLC. In case of longer reaction times (>48 h) the increase of conversion stopped. The observation that conversion could be further increased by adding another batch of enzyme excluded that stopping the conversion was due to product inhibition. Therefore, in reactions exceeding 48 h reaction time, a fresh enzyme batch was added to the reaction mixture at the end of each 48 h period.

Time course analysis of the *Pc*PAL-catalyzed conversions of racemic styrylalanines *rac*-**1a-d** (Figure 19) also supported the beneficial catalytic properties of F137V *Pc*PAL compared to the wild type enzyme. With the F137V mutant the theoretically possible conversion in a selective kinetic resolution (50%) could be achieved in reactions of *rac*-**1a,c** after moderate reaction time (*c* >49% after 25 h and 134 h, respectively) and even from *rac*-**1b,d** after longer incubation times (*c* >49% after 274 h and 300 h, respectively). On the other hand, in reactions of racemic styrylalanines *rac*-**1a-d** with wt-*Pc*PAL, complete conversion of the L-enantiomer could be obtained only for *rac*-**1a** (*c* >49% after 274 h), while the conversions of *rac*-**1b-d** could not approach the theoretically possible value within 274 h.

Analysis of the enantiomeric composition of the products D-**1a-d** indicated that by the F137V mutation in *Pc*PAL not only the reaction velocities were increased, but the enantiomer selectivity of the reaction from *rac*-**1d** was also enhanced (Table 6).With wt-*Pc*PAL the enantiomeric excess for the residual amino acid D -**1a** was high only in ammonia elimination from styrylalanine *rac*-**1a** (ee_{obs}>99%). The enantiomeric excess observed for the residual amino acids d-**1b,c** in the reactions from styrylalanines *rac*-**1b,c** was in agreement with the theoretically achievable one [Table 6, from *rac*-**1b**: ee_{obs}= 23% (ee_{theor}= 22% at *c*= 18%), from *rac*-**1c**: ee_{obs}= 30%, (ee_{theor}= 30% at *c*= 23%)]indicating high enantiomer selectivity of the ammonia elimination even with wt-*Pc*PAL. However, the significantly lower than theoretically achievable enantiomeric excess observed for the residual amino acid D-**1d** [ee_{obs}= 47% (ee_{theor}= 70% at *c*= 41%)] in the ammonia elimination from *rac*-**1d** indicated that

wt-*Pc*PAL could catalyze this reaction not only slowly but also with low enantiomer selectivity.

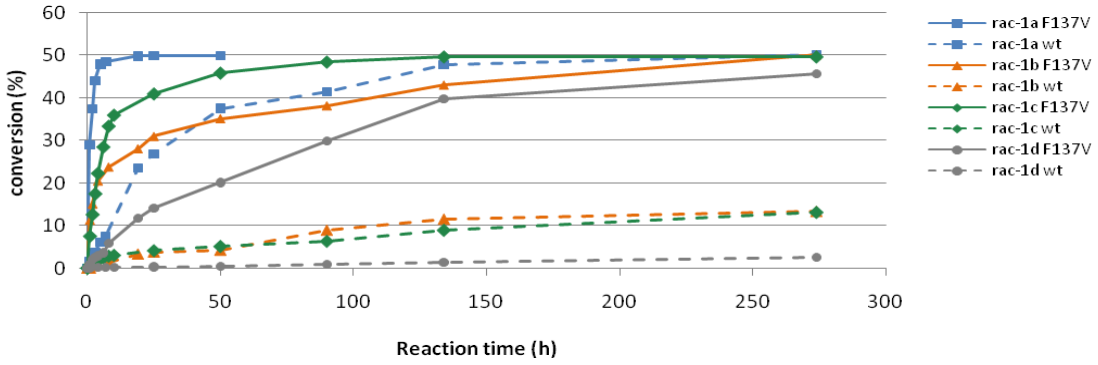


Figure 19. Comparative time profiles of conversions of *rac*-**1a-d** in the ammonia elimination reactions catalyzed by F137V-*Pc*PAL (continuous lines) and wt-*Pc*PAL (dashed lines).

Table 6. Conversion of styrylalanines *rac*-**1a-d** and enantiomeric excess of the residual D-**1a-d** in the semi-preparative scale ammonia elimination reactions catalyzed by *Pc*PAL variants

	*Pc*PAL	Substrate	Time (h)	*c* (%)	Ee_{theor}(%)[a]	Ee_{obs}(%)
1	F137V	*rac*-**1a**	24	50	100	>99
2		*rac*-**1b**	274	50	100	>99
3		*rac*-**1c**	94	50	100	>99
4		*rac*-**1d**	300	50	100	>99
5	*wild-type*	*rac*-**1a**	60	38	60	56
6		*rac*-**1a**	100	50	100	>99
7		*rac*-**1b**	274	19	23	29
8		*rac*-**1b**	504	28	39	41
9		*rac*-**1c**	274	13	15	10
10		*rac*-**1c**	600[b]	37	59	55
11		*rac*-**1d**	274	4	4	3
12		*rac*-**1d**	600[b]	31	45	41

[a]Ee_{theor}= *c*/(1-*c*) for a fully selective kinetic resolution**;** [b]Elongated batch with higher amounts of wt-*Pc*PAL

The occurance of the non-stereoselective MIO-independent reaction pathway2 was ruled out performing the ammonia elimination reactions of *rac*-**1a-d** with the S206A MIO-less *Pc*PAL mutant. The MIO-less enzyme showed no activity, while in the MIO-independent, positive control reaction using *p*-NO_2-phenylalanineas substrate, product formation was observed.

Molecular modeling revealed manifold reasons of the poor enantioselectivity. One reason could be the unusually high affinities of the D-enantiomers towards *wt-Pc*PAL (**Table 7**). Another reason was the unusually favorable arrangements of the catalytically active *N*-MIO intermediates of D-enantiomers. This was understandable because the alanine substructures of the D-enantiomers were essentially in perfect mirror image arrangement of the corresponding L-counterpart (**Figure 20, panel C**). The latter phenomenon implied that the rate limiting activation energies for both enantiomers should be in the same order of magnitude. The low reaction velocities can be explained by the worse affinities of all investigated intermediates of both L-**1a-d** and D-**1a-d** relative to the corresponding intermediates of D-Phe, and the distorted structures of the catalytically active *N*-MIO intermediates **(Figure 20A**).

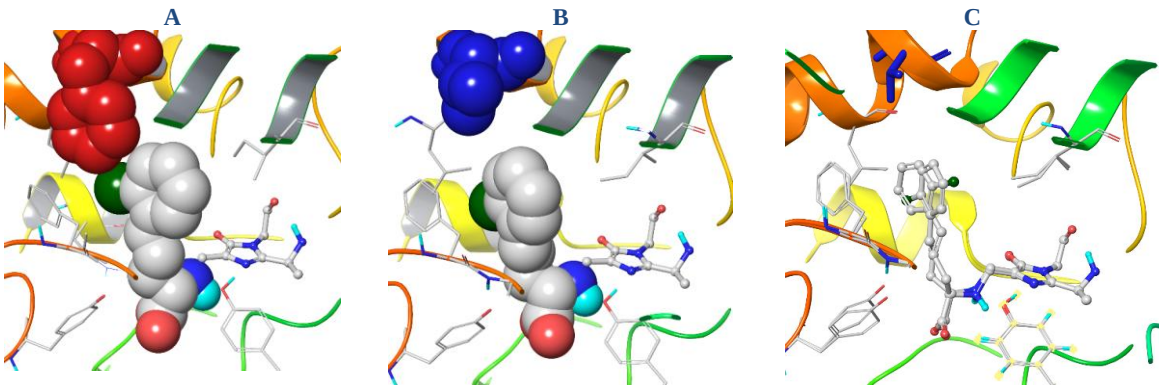


Figure 20. *N*-MIO covalent intermediates of **1c** in the active centre of *Pc*PAL. Panel A: *N*-MIO covalent intermediate of L-**1c** in *wt*-*Pc*PAL. F137 (in red) is in close contact with the substrate. Panel B: *N*-MIO covalent intermediate of L-**1c** in F137V-*Pc*PAL. V137 (in blue) is no more in close contact with substrate, maintaining a more relaxed complex. Panel C: *N*-MIO covalent complexes of L-**1c** and D-**1c** in wt-*Pc*PAL.

Table 7. Calculated molecular properties of compounds **1a-d.** Legend: A_{NC}– affinity of the non-covalent intermediate, A_C - affinity of the covalent *N*-MIO intermediate. L and D in the subscripts correspond to the actual enantiomer. The unit is kcal mol^{-1} for all quantities.

***Pc*PAL**	**Substrate**	$A_{NC,L}$	$A_{NC,D}$	$A_{NC,L}$-$A_{NC,D}$	$A_{C,L}$	$A_{C,D}$	$A_{C,L}$-$A_{C,D}$
wild-type	*rac*-**1a**	8.1	3.0	5.1	1.0	10.0	-9.0
	rac-**1b**	10.1	8.7	1.4	9.5	17.5	-8.0
	rac-**1c**	17.1	7.6	9.5	16.8	12.6	4.2
	rac-**1d**	15.5	12.8	2.7	30.9	40.4	-9.5
F137V	*rac*-**1a**	-4.7	-2.0	-2.7	-6.6	5.5	-12.1
	rac-**1b**	-12.0	-7.4	-4.6	-7.0	-3.0	-4.0
	rac-**1c**	-12.8	-4.3	-8.5	-8.0	2.9	-10.9
	rac-**1d**	-15.6	-5.5	-10.1	5.8	14.0	-8.2

In case of using the F137V-*Pc*PAL at conversion values appropriate to 50%, the enantiomeric excess values of the D-**1a-d** were high in all cases (Table 7, Entries 1-4). Further addition of enzyme batches and longer reaction times did not result in increased conversion values, supporting that the ammonia elimination reactions of *rac*-**1a-d** catalyzed by the F137V-*Pc*PAL are highly enantiomer selective. The calculated affinity values support this observation, as the F137V-*Pc*PAL favors heavily the L-enantiomers, practically excluding the D-enantiomers (Table 7).

The failure of ammonia addition onto styrylacrylates **2a-d,** presumably might derive from their low affinity towards the catalytic site, similarly as observed in case of *Tc*PAM, structurally closely related with *Pc*PAL, where the release of stryrylacrylate product (**2a**) obtained from the ammonia elimination from (*S*)-α-styrylalanine (*S*)-**1a** preceeded the NH_2-MIO deamination, thus the formation of β-styrylalanine was impaired.[17]

In our next study we focused on determining how the spatial position of the residues of the hydrophobic pocket of PcPAL influence its activity towards differently aromatic ring substituted, bulky phenylalanine analogues, such as *o*-, *m*-, *p*- methoxy- or bromo-phenylalanines. Testing the activity of our PcPAL enzyme collection we could identify exchange of I460 residues to Val, or Ala as recommended exchange to increase or even to

gain enzyme activity for *p*-substituted position, while exchange of residue L256 is also favourable for *o*-substituted derivatives (Table 8,9).

Table 8. Relative PAL activities measured by the novel fluorescent assay using *o*-, *m*-, *p*-phenylalanines as substrates.

***o*-Br-Phe**		***m*-Br-Phe**		***p*-Br-Phe**	
PcPAL	**Relative activity**	**Enzyme code**	**Relative activity**	**Enzyme code**	**Relative activity**
I460V	100.0	PcPAL *wt*	100.0	I460V	100.0
PcPAL *wt*	97.4	I460V	86.6	PcPAL *wt*	52.9
L256A	85.2	L134A	79.9	F137V/I460V	34.8
L256V	50.6	F137V	55.0	F137V	28.1
F137V	37.6	F137V/I460V	29.9	F137A/L138V	21.4
L256C	30.8	I460A	28.7	F137A	17.2
I460A	29.8	L256V	21.4	I460A	15.1
F137V/I460V	27.3	L138A	11.8	F137A/I460V	13.2
L134A	18.1	F137A/L138V	9.7	L134A	5.1
F137A/L138V	9.2	L138V	6.6	L138V	4.8
F137A	5.8	F137A	6.2	L256V	3.6
F137A/I460V	2.4	L134V	5.5	L134V	2.2
L134V	1.7	F137V/L138V	4.1	F137V/L138V	0.7
F137V/L138V	1.0	F137A/I460V	3.9	L138A	0.7
L138V	1.4	L256A	0.98	L256C	0.7
L138A	0.6	L256C	0.34	L256A	0.0

Table 9. Relative PAL activities measured by the novel fluorescent assay using *o*-, *m*-, *p*-methoxy-phenylalanines as substrates

***o*-MeO –Phe**		***m*-MeO –Phe**		***p*-MeO –Phe**	
Enzyme code	**Relative Activity**	**Enzyme code**	**Relative Activity**	**Enzyme code**	**Relative Activity**
L134A	100.0	PcPAL wild	100.0	I460V	100.0
I460V	89.9	I460V	86.6	PcPAL wt	52.9
F137A/I460V	52.4	L134A	79.9	F137V/I460V	34.8
I460A	50.3	F137V	55.0	F137V	28.1
L256A	49.0	F137V/I460V	29.9	F137A/L138V	21.4
L138A	48.5	I460A	28.7	F137A	17.2
L134V	47.4	L256V	21.4	I460A	15.1
L138V	46.8	L138A	11.8	F137A/I460V	13.2
PcPAL wild	42.8	F137A/L138V	9.7	L134A	5.1
F137V/I460V	35.4	L138V	6.6	L138V	4.8
F137A	31.1	F137A	6.2	L256V	3.6
F137A/L138V	28.8	L134V	5.5	L134V	2.2
L256V	28.4	F137V/L138V	4.1	F137V/L138V	0.7
F137V	23.3	F137A/I460V	3.9	L138A	0.7
F137V/L138V	22.8	L256A	0.98	L256C	0.7
L256C	13.1	L256C	0.34	L256A	0.0

Molecular modeling supporting our activity determinations is undergoing and will complete the documentation for the manuscript in preparation.

Further work in collaboration with other members of the Biocatalysis and Biotransformation Research Center focused succesfull immobilization of the *wt* -and mutant

PcPALs of the the developed MIO-enzyme collection on amine-functionalized single-walled carbon nanotubes, resulting in robust biocatalysts preparations which were tested in microreactor system for the synthesis of enantiopure aminoacids (see dissemation of the results).

- **PAM activity**

The succesfull expression and isolation of TcPAMs was hindered by the measured low activity of this enzyme in both aminomutase and ammonia addition reactions. The reason for the inactivity was not found, structurally (oligomerization state, melting temperature) the isolated enzymes proved to be in order and according to data from literature, purity assesed by SDS-PAGE and SEC chromatography was also fine. One possible cause might be an unsuccesfull codon optimization leading to non-correctly translated protein sequence which might result in enzyme activity loss. In order to remediate new TcPAM gene is in cloning process and will be tested in later stages. However to valorify the correlated synthetic work, the prepared racemic β-aminoacids (**3k-n**) and their corresponding ester derivatives (**4k-n**) were resolved succesfully by lipase mediated kinetic resolution in collaboration with other members the Biocatalytic and Biotransformation Research Center, yielding the optically pure D-**3k-n** and L-**3k-n** (see results dissemination).

References

1. van Assema, FBJ, Sereinig N (DSM): Method for producing optically active phenylalanine compounds from cinnamic acid derivatives employing a phenylalanine ammonia lyase derived from *Idomarina loihiensis*. *WO 2008/031578,* **2008** and *PCT/EP 2007/007945*, **2007**.

2. Parmeggiani F,Lovelock SL, Weise NJ, Ahmed ST, Turner NJ: Synthesis of d- and l-Phenylalanine Derivatives by Phenylalanine Ammonia Lyases: A Multienzymatic Cascade Process, *Angew. Chem. Int. Ed.* **2015**, *54*, 1-5.

3. Liu, H., Naismith, J. H.: An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology*, **2008**, 8:91

4. Bartsch S, Bornscheuer UT. Mutational analysis of phenylalanine ammonia lyase to improve reactions rates for various substrates. *PEDS* **2010**, *23*, 929-933

5. GE Healthcare - *Gel Filtration Calibration Kits*, Product booklet

6. Frank, A., Eborall, W., Hyde, R., Hart, S., Turkenburg, J.P., Grogan, G.: Mutational analysis of phenolic acid decarboxylase from *Bacillus subtilis* (*Bs*PAD), which converts bio-derived phenolic acids to styrene derivatives, *Catal. Sci. Technol.*, **2012**, *2*, 1568-1574

7. Plumridge, A., Stratford, M., Lowe, K.C., Archer, D.B. The weak-acid preservative, sorbic acid, is decarboxylated and detoxified by a phenylacrylic acid decarboxylase, PadA1, in the spoilage mold *Aspergillus niger*, *Appl Environ Microbiol*, **2008**, 74, 550–552

8. Stradford, M., Plumridge, A., Pleasants, M.W., Novodvorska, M., Baker-Glenn, C.A.G., Pattenden, G., Archer, D.B.: Mapping the structural requirements of inducers and substrates for decarboxylation of weak acid preservatives by the food spoilage mould *Aspergillus niger*, *International Journal of Food Microbiology*, **2012**, *157*, 375-383

9. Mukai, N., Masaki, K., Fujii, T., Kawamukai, M., Iefuji, H.: PAD1 and FDC1 are essential for the decarboxylation of phenylacrylic acids in *Saccharomyces cerevisiae*, *Journal of Bioscience and Bioengineering*, **2010**, *109*, 564–569
10. Leys, D., Payne, K.A.P., Scrutton, N.S., Parker, D.A, Murphy, A.J.: Methods for preparing hydrocarbon, **2013**, *US Patent* US 2013/0330795 A1
11. McKenna, R., Nielsen, D.R.: Styrene biosynthesis from glucose by engineered *E.coli*, *Metabolic Engineering*, **2011**, *13*, 544-554
12. Lee, Z.H., Raghavan, S.S., Ghadessy, F.J., Teo, N.Y.: Rapid and sensitive detection of acrylic acid using a novel fluorescence assay, *RCS Adv*, **2014**, *4*, 60216
13. Bencze, L.C., Filip, A, Banoczi, G., Tosa, M.I., Irimie, F.D., Poppe, L., Paizs, C.: Expanding the substrate range of phenylalanine ammonia lyase from *Petroselinum crispum* towards styryl-alanines, *Org. Biomol. Chem.*, 2017, **15**, 3717-3727.
14. C. Paizs, M. I. Toşa, L. C. Bencze, J. Brem, F. D. Irimie, *Heterocycles*, 2011, **82**, 1217–1228.
15. W. Qiu, V.A. Soloshonok, C. Cai, X. Tang, V.J. Hruby: Convenient, large-scale asymmetric synthesis of enantiomerically pure trans-cinnamylglycine and -α-alanine**,** *Tetrahedron*, **2000**, *56*, 2577–2582.
16. G. Bánóczi**,** Cs. Szabó, Zs. Bata, G. Hornyánszky, L. Poppe, Structural modeling of phenylalanine ammonia-lyases and related mio-containing enzymes – An insight into thermostability and ionic interactions, *Stud. Univ. Babes-Bol. Ser. Chem*, **2015**, *60*, 213–228.
17. Wanninayake, U., DePorre, Y., Ondari, M., Walker, K.D.: (*S*)-Styryl-α-alanine used to probe the intermolecular mechanism of an intramolecular MIO-aminomutase, *Biochemistry*, **2011**, *50*, 10082–10090

Disemination of the result from stage 2015-2017

Scientific publications:

1. Nagy B., Galla Z., Bencze L.C., Toșa M.I., Paizs C., Forró E., Fülöp F.:Covalently immobilized lipases are efficient stereoselective catalysts for the kinetic resolution of rac-(5-Phenylfuran-2-yl)-β-alanine ethyl ester hydrochlorides, *European Journal of Organic Chemistry*, **2017**, *20*, 2878-2882.

2. Varga A., Bata Z., Csuka P., Bordea D.M., Vértessy B.G., Marcovici A., Irimie F.D., Poppe L., Bencze L.C.: A novel phenylalanine ammonia-lyase from Kangiella koreensis, *Studia Universitatis Babes-Bolyai Chemia*, **2017**, accepted manuscript.

3. J. H. Bartha-Vári, L. C. Bencze, E. Bell, L. Poppe, G. Katona, F. D. Irimie, C. Paizs and M. I. Toşa,: Aminated single-walled carbon nanotubes as carrier for covalent immobilization of phenylalanine ammonia-lyase, *Per. Polytechn. Chem. Eng*. **2017**, *61*, 59.

4. Bencze, L.C., Filip, A, Banoczi, G., Tosa, M.I., Irimie, F.D., Poppe, L., Paizs, C.: Expanding the substrate range of phenylalanine ammonia lyase from *Petroselinum crispum* towards styryl-alanines, *Org. Biomol. Chem.*, **2017**, **15**, 3717-3727.

5. Dima, N.A., Filip, A., Bencze, L.C., Olah, M., Satorhelyi, P., Vertessy, B., Poppe, L., Paizs, C.: Expression and purification of recombinant phenyalanine-ammonia lyase from *Petroselinum crispum*. *Stud. Univ. Babes-Bol. Chemia*, **2016**, *61 (2)*, 21-34.

6. Filip, A., Bencze, L.C., Paizs, C., Poppe, L., Irimie, F.D.: MIO-enzyme toolbox: cloning, expression and purification of recombinant *Rt*PAL, *Stud. Univ. Babes-Bol.*, Sp. Iss., **2015**, 39-43.

International conferences

1. Bencze, L.C., Filip, A., Nagy, E.Zs.A., Nagy, B., Banoczi, G., Poppe, L., Paizs, C.: Rational design of phenylalanine ammonia lyase from Petroselinum crispum (PcPAL)towards sterically demanding phenylalanine analogues, *VII EFMC International Symposium on Advances in Synthetic and Medicinal Chemistry, Viena (Austria), **2017*** - Poster Section

2. Bordea D. M., Varga A., Bencze, L.C.**:** Molecular Cloning, isolation and characterization of novel, highly active phenylalanine ammonia lyase from *Pseudozyma antartica*, XIVth Edition of the International Conference "Students for Students" Cluj-Napoca, **2017 -** Award Certificate – for the best undergraduate poster at the Internal Conference Student for Student, 2017 - Poster Section

3. Tork S.D., Filip A, Nagy Z.: *Novel PcPAL mutants exhibiting broadened substrate scope*, XIVth Edition of the International Conference "Students for Students" Cluj-Napoca, **2017 -** THE GO GREEN! APPROACH award granted by SPI (Special Pro Installations) **-** Oral presentation

4. Tork S.D., Filip A., Bartha-Vári J.H.,Bencze L.C., Irimie F.D., *Immobilization and substrate domain expansion of RtPAL*, XIIIth Edition of the International Conference "Students for Students" Cluj-Napoca, **2016 -** Oral presentation

5. Filip, A, Banoczi, G., Poppe, L., Paizs, C., Irimie, F.D., Bencze, L.C.: Expanding the substrate range of phenylalanine ammonia lyase from *Petroselinum crispum* towards styryl-alanines, *Biotransformations for Pharmaceutical and cosmetic industry*, Varsovia, **2016** – gold medal award - Poster Section

6. Filip, A, Tork, D.S., Bencze, L.C., Banoczi, G., Irimie, F.D., Poppe, L., Paizs, C.: MIO-enzyme toolbox: F137V mutant of phenylalanine ammonia-lyase from *Petroselinum crispum* with improved substrate tolerance and stereoselectivity, *Action COST CM1303 SysBiocat Training School*, Siena, **2016**

7. Nagy, E.Zs.A., Filip, A., Bencze, L.C., Paizs, C., Irimie, F.D. Expanding the substrate range of *Pc*PAL towards biphenyl-alanines, 16th CEEPUS Symposium and Summer School on Bioanalysis, Varsovia, **2016** - Poster Section

8. Filip, A, Tork, D.S., Bencze, L.C., Banoczi, G., Irimie, F.D., Poppe, L., Paizs, C.: MIO-enzyme toolbox: F137V mutant of phenylalanine ammonia-lyase from *Petroselinum crispum* with improved substrate tolerance and stereoselectivity, *Action COST CM1303 SysBiocat Training School*, Siena, **2016** - Poster Section

9. Bencze, L.C., Filip, A, Banoczi, G., Tosa, M.I., Irimie, F.D., Poppe, L., Paizs, C.: Styryl-alanines as novel substrates for phenylalanine ammonia lyases, *International Symposium on Synthesis and Catalysis (ISYSYCAT)*, Evora (Portugal), **2015**, poster awarded with runner-up prize in the Catalysis section. - Poster Section

30.09.2017 Lect. Dr. Laszlo-Csaba Bencze