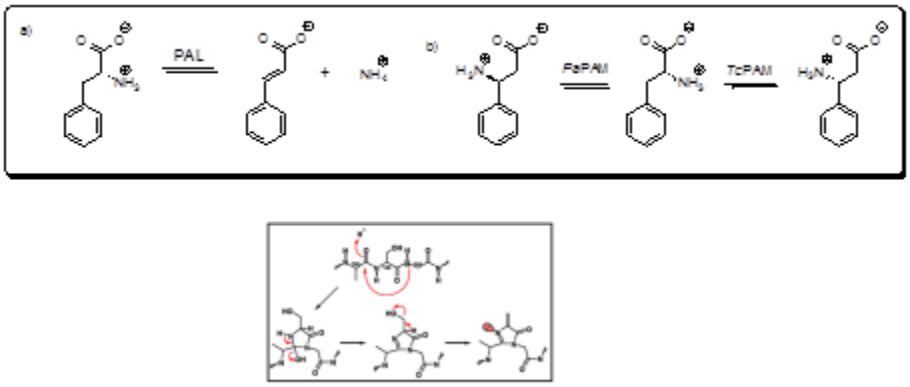
Scientific synthetic report of the project

IMMOBILIZED WILD-TYPE AND MUTANT AMMONIA-LYASES AND AMINOMUTASES FOR THE PRODUCTION OF ALPHA- AND BETA-PHENYLALANINE ANALOGUES

2012

Phenylalanine ammonia-lyases (PAL; EC 4.1.3.24) catalyse the non-oxidative deamination of L-phenylalanine in (*E*)-cinnamic acid, while phenylalanine 2,3-aminomutases (EC 5.4.3.x) catalyse the isomerization of L-phenylalanine to form L- or D-β-phenylalanine, depending on the origin of the enzyme (Scheme 1.). Even though PALs are frequently found in plants, where they have an essential role in forming phenylpropanoids, only a few bacterial PALs have been identified so far. PAL and PAM are used as biocatalysts for the synthesis of L-α-amino acids from acrylates (PAL), in kinetic resolution processes for obtaining D-α-amino acids starting from their racemates (PAL), or for the synthesis of L- or D-β-arylalanines (PAM).



Scheme 1. PAL and PAM catalysed reactions. Posttranslational forming of the strong electrophile MIO group.

PAL is a homotetramer formed of monomeric subunits, which can be divided into three domains. The first is the catalytic MIO domain, where the prosthetic group is found. The amino acids sequence which codifies this prosthetic group is strictly conserved. The second domain is the central domain and the third one is known as the "shielding-domain", which presumably has a defence and/or regulation role (**Figure 5.**). The domains are connected with mobile loops. The MIO group forms as a result of the autocatalytic, posttranslational cyclisation of an internal tripeptide Ala-Ser-Gly, by the elimination of two water molecules.

Initially, we expressed PAL from parsley *Pc*PAL(eucaryotic) and PAL from a radiotolerant thermophilic bacteria *Rubrobacter xylanophilus Rx*PAL (NCBI code: YP_644511.1, Uniprot code: Q1AV79, which codifies 540 AA) in *E. coli*, with an optimal growth at 65°C and we investigated their properties.

The latter was found using BLASTp program/algorithm in the non-redundant protein database NCBI and the sequence of PAL from *Photorhabdus luminescens* (*Pl*PAL) (Duchaud et al. 2003, Williams et al. 2005) (Uniprot code: Q7N4T3). Due to favorable results (code: YP_644511.1, which codifies 540 AA), it is assumed that this could be phenylalanine/hystidine ammonia-lyase from the thermophilic bacteria *Rubrobacter xylanophilus* DSM 9941.

The gene which codifies PAL from *Rubrobacter xylanophilus* was optimized at the codon level in order to be expressed in *Escherichia coli*. By using the PCR technique, the new synthesized gene - formed of 1632 base pairs -found in the cloning vector pMK and cloned in the expression vector pBAD-HisB, has a 6xHis-tag at the N-terminal, so it can be easily purified from the lysate by Ni-NTA chromatography. Other advantages offered by the new vector are the expression regulation and the concentration dependent induction, which allows the modulation of the expression levels. The first step in the purification of the plasmids consisted of obtaining the XL1Blue competent cells, which were transformed and selected by inoculation on solid LB medium in presence of tetracycline, resistence coresponding to the strain, and carbenicillin, respectively, according to the antibiotic resistance genes encoded in the plasmids. The obtained vector was confirmed by sequencing, using pBAD promotor, "forward" and "reverse" primers (5'–CCTGACGCTTTTTATCGCAACTC–3' and 5'–GAGGCATCCTGGTACCCCAG–3', respectively).

After the optimization of the expression, *Rx*PAL was overexpressed in *E. coli* TOP 10 in a soluble form, with a 6xHis-tag at the N-terminal (*Rx*PAL). The recombined protein *Rx*PAL was purified by affinity chromatography using Ni-NTA. In order to analyse the expression (samples: before induction, before lysis, cellular debris and supernatant after sonication) and the purification, polyacrylamide gel electrophoresis (SDS-PAGE) was performed. As can be observed on the images of the gels (**Figure 1.**), *Rx*PAL enzyme, with a molecular mass of 58.4 kDa, was successfully produced and purified.

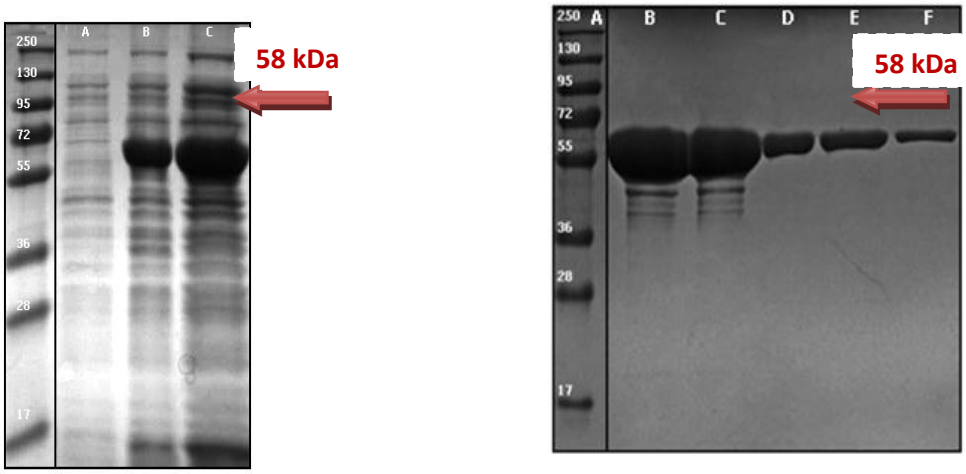


Figure 1. Expression (a) and purification (b) verification of the phenylalanine ammonia-lyase protein from *Rubrobacter xylanophilus* (*Rx*PAL) in *E. coli* TOP10 strain a. ProSieveTM QuadColor™ protein marker, 17 kDa – 250 kDa; **A**- cell culture before induction **B**- bacterial lysate; **C**-supernatant containing the recombinant protein b. **A**- ProSieveTM QuadColor™ protein marker, 17 kDa–250 kDa; **B - F** –fractions collected after the elution of the protein off the column

The enzyme activity was monitored by spectrophotometry, at a 290nm wavelength, by detection the product formation ((*E*)- cinnamic acid) starting from its natural substrate, L-phenylalanine. First, we focused on finding the optimal pH of the enzyme, measurements were performed in buffer solutions, maintaining constant the ionic strength and the temperature, in the pH range of 6.5-12, for 7 minutes. The pH profile in **Figure 2.** was obtained. On the representation of the slopes as a function of pH (the graphic on the right side), two values of optimal pH can be observed, 8.5 and 11.2, respectively. The data found in the scientific literature are similar with the first optimal pH at 8.5, but in this case the highest activity was registered at pH 11.2.

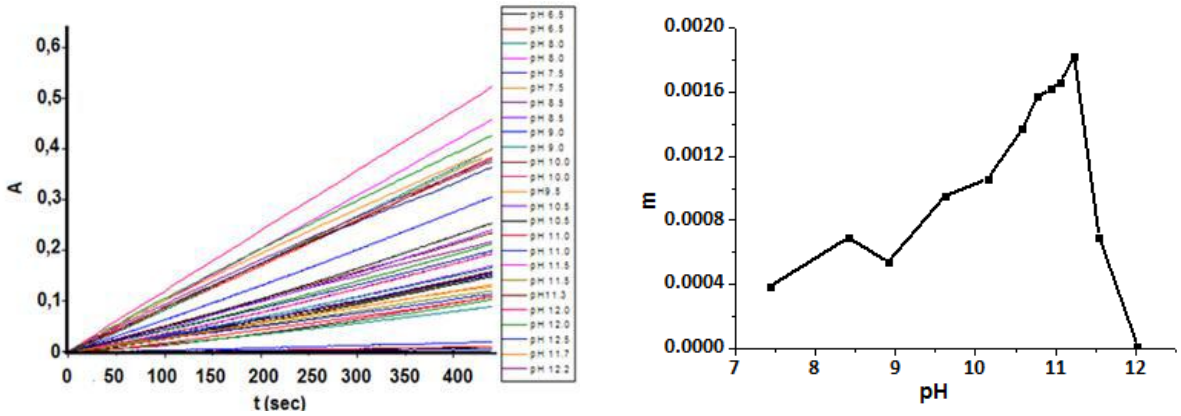
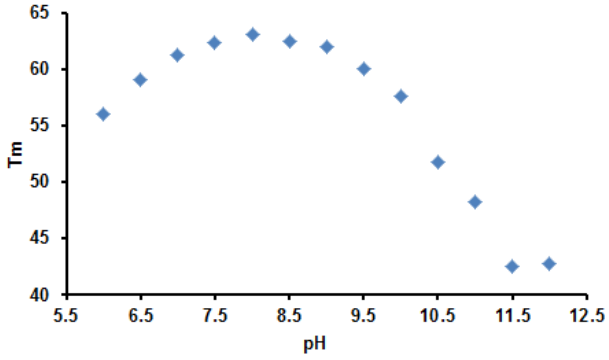
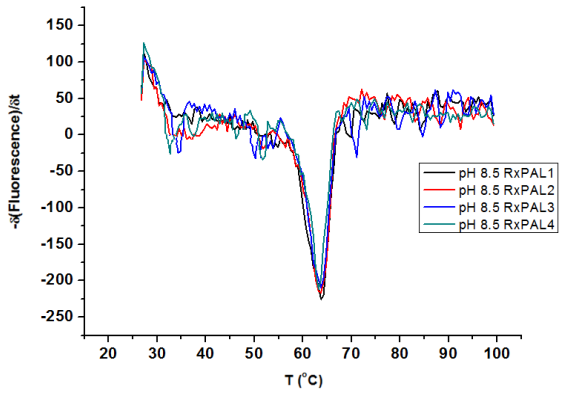


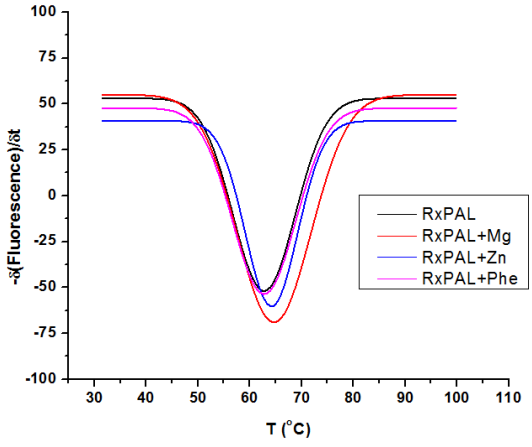
Figure 2. Activity profile of *Rx*PAL in different pH conditions



a



b



c

Figure 3. Melting temperature profile of *Rx*PAL in different pH conditions

The thermal stability of the enzyme was determined by thermofluorometric measurements at a concentration of 2 mg/mL *Rx*PAL. First, in order to determine the optimal pH regarding the melting temperature (Tm), measurements were performed in buffer solutions 100mM, in the pH range 6-12. From the pH profile (**Figure 3**.**a**), it can be observed that the maximum melting temperature is reached in the 8-8.5 pH range and is approximately 63°C.

The following measurements were performed in Tris-HCl buffer solution, pH 8.5 and different substrate concentrations (**Figure 3.b)**, and further on in the presence of metal ions (Mg^{2+}, Zn^{2+}) or phenylalanine (**Figure 3.c)**. The melting temperature was read from the negative curve of the first derivative of the experimental curve. By comparing the measurements results obtained in presence and in absence of the substrate, it can be observed that the melting temperature (63°C) is not modified in presence of phenylalanine, but in exchange, the Mg^{2+} ion in a 0.3mM concentration increases with 1.9°C the melting temperature of the protein, while the Zn^{2+} ion in 0.2 µM concentration increases it with 1,4°C.

For the purpose of analyzing the **enzyme structure stability**, circular dichroism (CD) spectroscopy measurements were performed, in the presence of the natural product of the enzymatic reaction, cinnamic acid, due to the fact that the phenylalanine spectrum overlays on the spectrum of the protein. Measurements were performed in buffer solutions keeping a constant temperature (20°C), in the pH range 6.5-12, at a *Rx*PAL concentration of 2 mg/mL.

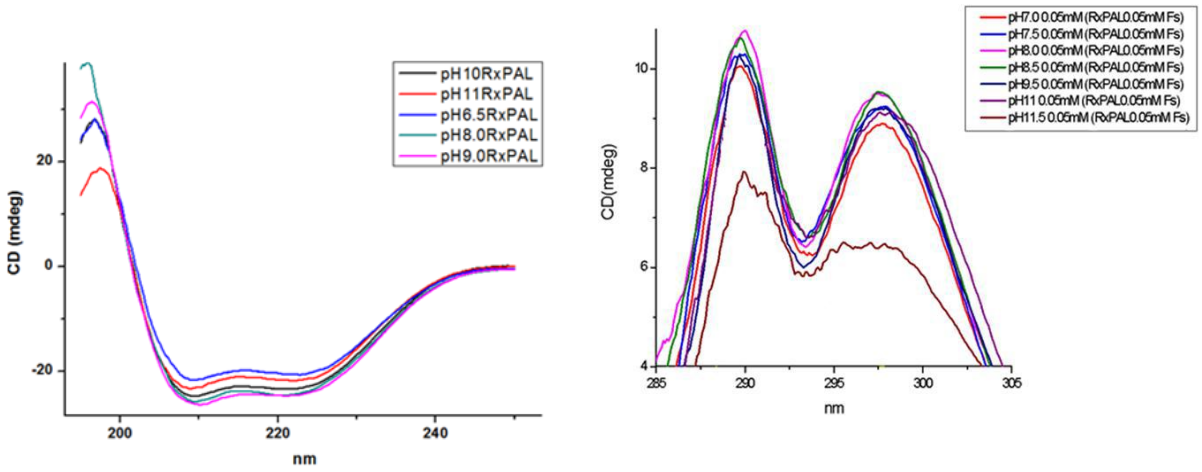


Figure 4. The CD spectra of *Rx*PAL in far UV and near UV at different pH values

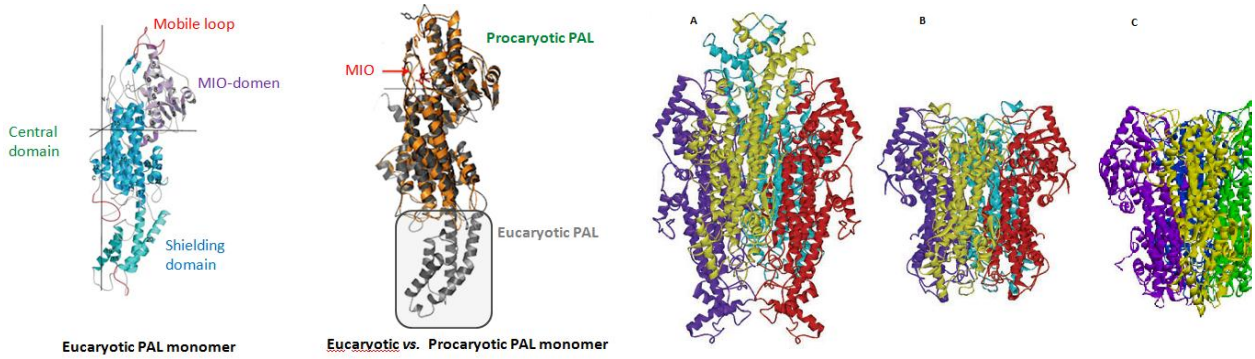
The CD spectra (**Figure 4**) indicate the fact that the enzyme doesn't present significant modifications regarding the secondary, tertiary and cuaternary structure in the 6.5-11.2 pH range, which means that the structure is stable in this pH range. Also, it can be observed that, at pH values higher than 11.5, a part of the enzyme structure deteriorates.

The activity of the enzyme was determined by monitoring the formation of the (*E*)-cinamic acid from phenylalanine in UV, at 290nm, using different concentrations of the substrate (10-40 mM). The measurements were performed in CAPS solution with a concentration of 100mM and pH 11.4, the enzyme concentration being 2 µM. The found Michaelis-Menten constants were K_m= 1 mM and k_{cat} $10^{-4}s^{-1}$. A possible explanation for these values could be that the nucleophile 2-sulfanylethanol used in the purification interacts with the strongly electrophile prosthetic group MIO, forming an adduct which irreversibly inhibts the enzyme. In absence of the reduction

agent, the thiolic groups interact intermoleculalrly, forming enzymatic aggregates(cross-linked enzyme aggregates-CLEA), which leads to precipitation[1].

By going through all the steps mentioned above, **eucaryotic PAL from *Petroselinum crispum*** (parsley, *Pc*PAL) was characterized. In contrast to *Rx*PAL, the activity of *Pc*PAL is maximum at 37°C and pH 8.8. Furthermore, the K_m 30mM and k_{cat} $34s^{-1}$ values, plus the fact that *Pc*PAL proved to be a useful biocatalyst for the biotransformation of unnatural analogues of phenylalanine, qualify this enzyme to have a superior potential compared to *Rx*PAL. Nevertheless, a relatively quick inactivation of *Pc*PAL was revealed by measurements of enzymatic activity, CD and thermofluorometry.

By comparing the eucaryotic to the procaryotic counterparts, it can be stated that the latter contain approximately 20% less amino acids in their structure. The approximately 200 amino acids C-terminal multihelix loop, found in eucaryotic PALs, is positioned around the catalytic site, restricting the attachement of the substrate and the detachment of the reaction product from the enzyme (**Figure 5**). Furthermore, it is assumed that these loops structurally destabilize the enzymes.



Difference between the monomer in eucaryotic PAL and procaryotic PAL, respectively.

A. *Pc*PAL **B.** *Pl*PAL (model) **C.** *Rx*PAL

Figure 5. Eucaryotic and procaryotic PAL structures

Regarding the thermostability of *Rx*PAL, two factors were considered: the possibility of forming disulfide and ionic bridges. Three possibilities of forming disulfide bridges in a monomer were detected (**Figure 6a**): Cys_{35}-Cys_{116}, Cys_{321}-Cys_{478}, Cys_{231}-Cys_{228}, the last one being the intrahelical one, the least important. Cys_{35} and Cys_{116} are placed at the surface, consequently, these cysteins were transformed in serines, in order to prevent the formation of interproteic aggregates. The other cysteins are disadvantaged in forming aggregates not only because of the large number of ionizable superficial amino acids (Glu 8.5%, Arg 8.32% reported to the entire sequence), but also because of their even distribution on the surface (**Figure 6**).

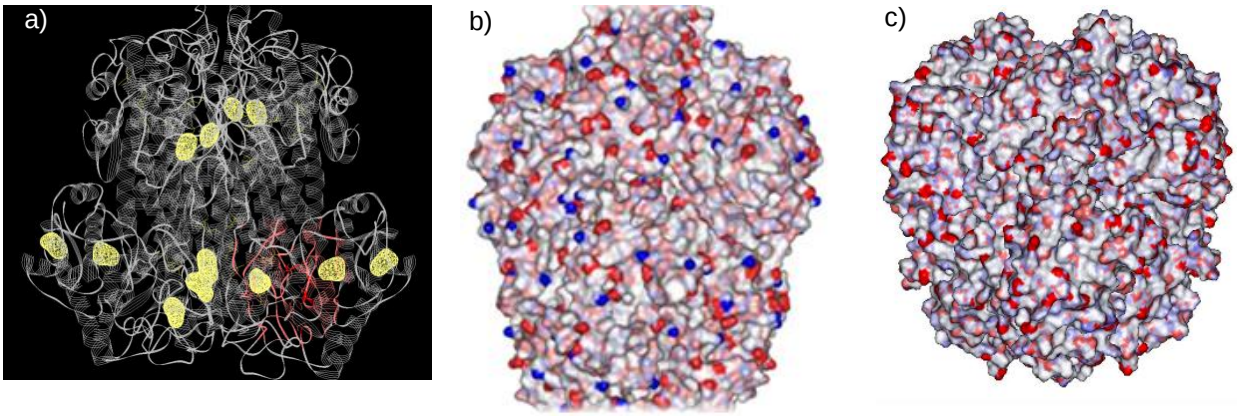


Figure 6 (a) Proximal cysteins in *Rx*PAL structure (yellow domains) one of the four catalytic sites marked with red.The model of the surfaces of (b) *Pc*PAL and (c) *Rx*PAL. Acid residues (red), basic residues (blue). Isolated charges for *Pc*PAL and the even, yet acid, distribution on the surface of *Rx*PAL.

Although the aggregation possibility of *Rx*PAL was removed, the catalytic parameters of the enzyme weren't improved. This fact led us to the conclusion that, at least in the case of *Rx*PAL, the formation of disulfide bridges in the expression process is essential. Since the expression in the pBAD-HisB vector and *E. coli* TOP 10 don't allow this, we focused our attention on designing native genes which can be expressed in pET19b vector and are compatible with *E. coli* Rosetta (DE3)plysS competent cells, which assist posttranslational disulfide bridges forming. Besides the *Pc*PAL and *Rx*PAL genes, a series of native phenylalanine ammonia-lyases in the pET19b expression vectors were ordered, like the ones from *Anabaena variabilis* (wt-*Av*PAL), *Streptomyces maritimus* (wt-*Sm*PAL) and *Rhodosporidium toruloides* (wt-*Rt*PAL), but also native phenylalanine aminomutases, like the ones from *Taxus canadensis* (wt-*Tc*PAM) and *Pantoea agglomerans* (wt-*Pa*PAM). It is worth mentioning that, according to the case, the singular cysteine codons from the shielding domain (which are not involved in disulfide bridges) were exchanged with the serine codons.

We present below the amino acids sequence from the wt-PAL monomer (cysteine exchanged with serine is marked with yellow):

MENGNGATTNGHVNGNGMDFCMKTEDPLYWGIAAEAMTGSHLDEVKKMVAEYRKPVVKLGGETLTISQVAAISARDGSGVTVELSE
AARAGVKASSDWVMDSMNKGTDSYGVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGNGSDNTLPHSATRAAMLVRINTLQGYSGI
RFEILEAITKFLNQNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPTGVILSPEEAFKLAGVEGGFFELQPKEGLALVNGTAVGS
GMASMVLFEANILAVLAEVMSAIFAEVMQGKPEFTDHLTHKLKHHPGQIEAAAIMEHILDGSAYVKAAQKLHEMDPLQKPKQDRYAL
RTSPQWLGPQIEVIRSSTKMIEREINSVNDNPLIDVSRNKAIHGGNFQGTPIGVSMDNTRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLS
GGRNPSLDYGFKGAEIAMASYCSELQFLANPVTNHVQSAEQHNQDVNSLGLISSRKTSEAVEILKLMSTTFVGLCQAIDLRHLEENLKST
VKNTVSSVAKRVLTMGVNGELHPSRFCEKDLLRVVDREYIFAYIDDPCSATYPLMQKLRQTLVEHALKNGDNERNLSTSIFQKIATFED
ELKALLPKEVESARAALESGNPAIPNRIEECRSYPLYKFVRKELGTEYLTGEKVTSPGEEFEKVFIASKGEIIDPLLESLESWNGAPLPIS

The pET19b plasmids contain a His10-tag N-terminal sequence, a clivage site with an enterokinase, a gene resistant to ampicillin, a promotor T7*lac*, PAL and PAM genes, respectively. The expressions induced using IPTG in *E. coli* Rosetta (DE3)pLysS gave good yields, the purification of the proteins using Ni-NTA chromatography was performed using a 500mM imidazole solution as eluent. In **Figure 7** the genetic map of the plasmids and the electrophoretogram of different fractions obtained after the lysis of the cell wall and from the eluents of the chromatographic purification, respectively.

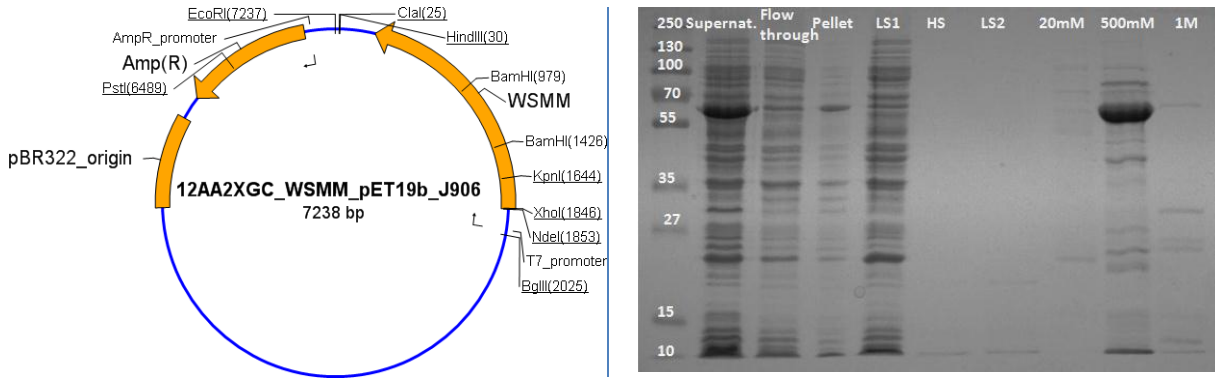


Figure 7. The genetic map of the plasmid used for the production of wt-*Pa*PAM and the electrophoretogram which certifies the expression and purity of the enzyme in different stages of the purification process.

Lyases and mutases obtained this way, presented similar catalytic parameters to the ones obtained previously by us (for *Pc*PAL) or presented in literature for the other proteins. As expected, *Rx*PAL activity improved, K_m 5mM and k_{cat} $0.5s^{-1}$, but unfortunately, these values are not promising enough to qualify this enzyme among the preparative scale useful biocatalysts. The stability of the eucaryotic enzymes grew with approximately 10-15%, the destabilising effect of the helices from the shielding domain continue to persist.

In conclusion, it was proved that procaryotic enzymes are stable, but catalytically inefficient, while eucaryotic enzymes are structurally unstable, but possess an important biocatalytic capacity.

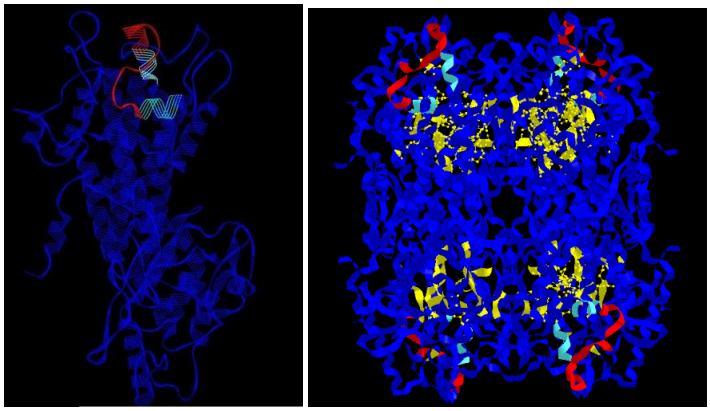
Next, we designed hybrid mutants, by combining the catalytic domain of eucaryotic enzymes with the shielding domain of the procaryotic enzymes, aiming to produce stable and catalytically efficient enzymes. The cleavage section was determined by aligning the sequences and visually inspecting them, taking into account two possibilities: performing one intersection including the excision of several amino acids from the host sequence (eucaryotic catalytic domain) or performing two intersections by inserting the procaryotic loop between the catalytic domain and the short C-terminal sequence, both belonging to the host domain. The second option was selected, due to the fact that an important hydrophobic interaction between the central domain and the short C-terminal sequence of *Pc*PAL was noticed. Furthermore, the two cleavage sections are well conserved and close to the junction place (**Figure 8**).

```
PcPAL VEILKLMSTTFLVGLCQAIDLRHLEENLKSTVK 537
AvPAL VDIFQNYVAIALMFGVQAVDLRTYKKTGHYDAR 501
*:*::  : *:  **:*** ::. : .:
PcPAL RASLSPATERLYSAVRHVVGQKPTSDRPYIWND 683
AvPAL RIEECRSYP-LYKFVRKELG------TEYLTGE 527
* . . :  **. **: :*      *: .:
```

Figure 8. Parts from the sequence alignment of *Pc*PAL and *Av*PAL. The intersection conserved domains are marked with green.

Further on, we evaluated the possible hybrids and designed the amino acids sequences for the *Pc*PAL(catalytic domain)-*Av*PAL(shielding domain), *Tc*PAM(cd)-*Av*PAL(sd), *Rt*PAL(cd)-*Av*PAL(sd) and *Av*PAL(cd)-*Pc*PAL(sd). For the latter, a drop in stability is expected due to the assembling of the procaryotic catalytic domain of *Av*PAL and the highly destabilising shielding domain of *Pc*PAL. The amino acids sequence for one of themonomers of the *Pc*PAL(cd)-*Av*PAL(sd) hybrid tetramer is given in **Figure 9a**. In **Figure 9b, c** we can see the model monomer and the tetramer of the hybrid mutant, respectively, where in dark blue is represented the catalytic domain of *Pc*PAL, in light blue is the intersection conserved domain from *Pc*PAL, in red is the shielding domain of *Av*PAL and in yellow are the serine which replaced the singular cysteins, which are not involved in forming disulfide bridges. The design of the gene and the plasmid, respectively, took place in the same manner as described above. The expression and purification of all hybrids was performed successfully, by using the method described before. The first activity and stability determinations look promising; furthermore, the *Pc*PAL(cd)-*Av*PAL(sd) mutant hybrid accepts as substrates, unnatural phenylalanine analogues. We estimate that by the end of the year we will succeed in completely characterising the activity and stability of the hybrid mutants.

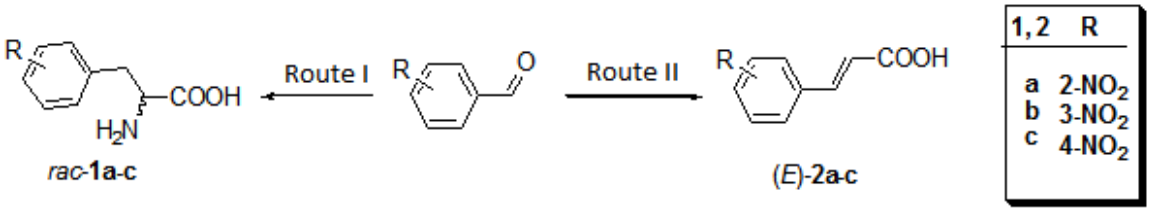
MENGNGATTNGHVNGNGMDFCMKTEDPLYWGIAAEA
MTGSHLDEVKKMVAEYRKPVVKLGGETLTISQVAAISA
RDGSGVTVELSEAARAGVKASSDWVMDSMNKGTDSY
GVTTGFGATSHRRTKQGGALQKELIRFLNAGIFGNGSDN
TLPHSATRAAMLVRINTLQGYSGIRFEILEAITKFLNQNIT
PCLPLRGTITASGDLVPLSYIAGLLTGRPNSKAVGPTGVI
LSPEEAFKLAGVEGGFFELQPKEGLALVNGTAVGSGMA
SMVLFEANILAVLAEVMSAIFAEVMQGKPEFTDHLTHK
LKHHPGQIEAAAIMEHILDGSAYVKAAQKLHEMDPLQK
PKQDRYALRTSPQWLGPQIEVIRSSTKMIEREINSVNDNP
LIDVSRNKAIHGGNFQGTPIGVSMDNTRLAIAAIGKLMF
AQFSELVNDFYNNGLPSNLSGGRNPSLDYGFKGAEIAM
ASYCSELQFLANPVTNHVQSAEQHNQDVNSLGLISSRKT
SEAVEILKLMSTTFLVGLCQAIDLRTYKKTGHYDARASL
SPATERLYKFVRKELGTEYLTGEKVTSPGEEFEKVFIAM
SKGEIIDPLLESLESWNGAPLPIS



A **B** **C**

Figure 9. A. Primary structure of the hybrid mutant *Pc*PAL(dc)-*Av*PAL(ds). **B.** Model structure of the monomer. **C**. Quaternary structure of the homotetramer

In paralell, a series of experiments were performed, using unnatural nitro-phenylalaninic substrates with wt-*Pc*PAL with increased stability, where the two singular cysteins (which are not involved in disulfide bridges) from the shielding domain were exchanged with serines. Plus, by moving Ser143→Ala143, a new mutant was created, in which the posttranslational formation of the MIO prosthetic group becomes impossible (*Pc*PAL MIO deficitary). The synthesis of the substrates is given in **Scheme 2**.



I. **a**. $NaBH_4$ / CH_3OH; **b**. $SOCl_2$/CH_2Cl_2, 0-5°C- r.t.; **c**. NaH, $CH_3CONH(COOEt)_2$ / DMF, 60°C; **d**. 18% HCl-dioxan 1:1 (V/V), reflux, 12h;
II. **a**.Ph_3P=CHCOOEt / toluen, reflux, 2 h; **b**. 10% KOH, reflux.

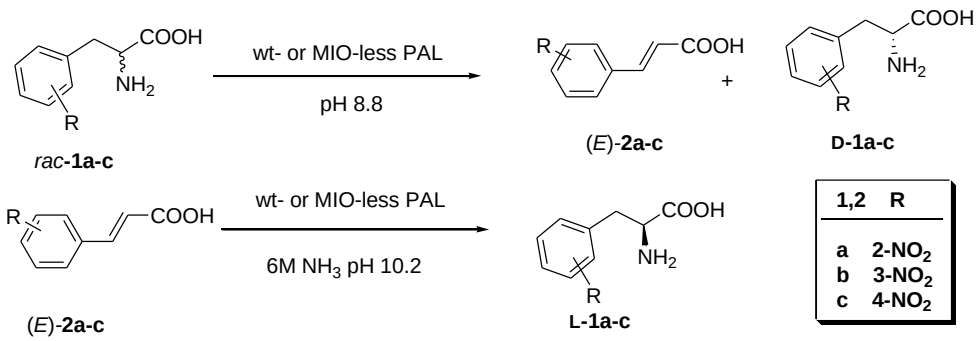
Scheme 2. The synthesis of nitro-phenylalanines and their corresponding acrylates

Kinetic measurements with wt-*Pc*PAL and *Pc*PAL-MIO deficitary highlighted, on one side, the fact that wt-*Pc*PAL can successfully mediate the deamination of all investigated nitro-phenylalanines, on the other side, the fact that the lack of the strongly electrophile prosthetic group in *Pc*PAL MIO-deficitary is counterbalanced by the strongly electron withdrawing character of the nitro group, thus the deamination of these substrates can take place with acceptable kinetic parameters if it is mediated by PAL with a quasi-inexistent catalytic capacity as opposed to its natural substrate (**Scheme 3**, **Table 1**).

Table 1. Kinetic parameters for PAL mediated biotransformations

Substrate	K_m (mM)	$V_{max}/V_{max\ Phe}$	$\lambda_{\mu\alpha\xi}$ (nm)	V_{max} wt-PAL / V_{max} PAL MIO-deficitary
l-Phe	33.3	1	290	413
rac-**1a**	268.1	0.57	243	244
rac-**1b**	64.6	0.21	260	52
rac-**1c**	295.5	0.86	340	411

0.53 mg PAL/ml solution; 26.5 mg PAL/sample



Scheme 3. Preparative scale synthesis of both enantiomers of nitro-phenylalanines

The *Pc*PAL stabilised by replacing singular cysteins with serines and by forming disulfide bridges (expression in *E. coli* Rosetta (DE3) pLysS), respectively, allowed the investigation of nitro- containing substrates interaction with the MIO-deficitary mutant, which besides the practical utility of enzymatic reactions was also used to elucidate the PAL-substrate interaction mechanism.

Assoc. Prof. Dr. Eng. Paizs Csaba