



Horizon 2020 Work programme

Food Security, Sustainable Agriculture and Forestry, Marine, Maritime and Inland Water Research and the Bioeconomy

Call

H2020-FNR-2020: Food and Natural Resources

Topic name

FNR-16-2020: ENZYMES FOR MORE ENVIRONMENT-FRIENDLY CONSUMER PRODUCTS

FuturEnzyme:

Technologies of the Future for Low-Cost Enzymes for Environment-Friendly Products

Final ID: 101000327

26/05/2023



THE SET OF 18 MUTANTS GENERATED BY GENETIC ENGINEERING

D5.2

VICTOR GUALLAR

BSC

Jordi Girona 31, BARCELONA 08034, Spain

Document information sheet

Work package:	WP5, Enhancing enzymes through innovative engineering.
Authors:	Víctor Guallar (BSC)
Document version:	1
Date:	31/05/2023
WP starting date:	01/08/2021
WP duration:	40 months
WP lead:	FHNW
WP participant(s):	FHNW, CSIC, BSC, Bangor, UHAM, UDUS, Inofea AG, Eucodis
Deliv lead beneficiary:	BSC
Dissemination Level:	Confidential, only for members of the consortium (including the Commission Services)
Type:	Other
Due date (months):	24
Contact details:	Víctor Guallar (victor.guallar@bsc.es)

Summary

Set of 18 mutants generated by genetic engineering	3
1. Scope of Deliverable	3
2. Enzyme targets for engineering	3
2.1. FE_Lip9 lipase engineering: for detergent and textile applications	3
2.2. Engineering FE_Lip5 by lid domain remodelling: for detergent applications	10
2.3. PEH_Paes_PE-H engineering: for textile applications	11
2.4. PET46 from Candidatus Bathyarchaeota engineering: for textile applications	11
2.5. Gen0105 and Kest3 engineering for detergent application	13
2.6. Gen0095 engineering for textile applications	16
2.7. TB035 engineering for textile applications.....	16
3. Methodology	16
3.1. PELE simulations	16
3.2. Docking simulations.....	17
3.3. Hot spots analysis	17
3.4. Evolutionary trace	17
4. Conclusions	17
5. References	20

Set of 18 mutants generated by genetic engineering

1. Scope of Deliverable

This deliverable consists in 18 mutants generated by genetic engineering. This deliverable will be accompanied by a report detailing the methods used to generate the mutants and the datasets describing their performances and stabilities. This information will be linked to the corresponding QR code associated to each enzyme, which will be made available in the internal FuturEnzyme repository. As detailed in this report to date a total of 40 mutations were selected based on a number of approaches, from which the effect of 26 were experimentally validated and mutants with improved performance identified. Mutants include those containing single or multiple mutations, those with ancestral sequences reconstructed and those with domain redesign.

2. Enzyme targets for engineering

Enzyme engineering through novel advanced techniques is a valuable approach to address the engineering of enzymes with high application potential. In this deliverable several enzymes were selected for engineering, the reason and outcomes of which are detailed below.

2.1. FE_Lip9 lipase engineering: for detergent and textile applications

Partner: CSIC, BSC

Source: MarRef - Marine Metagenomics Database;

Nominated for detergent and textile application as it is capable to degrade all the stained fabrics tested (pigments with oil on polyester/cotton (PC-09), mayonnaise on cotton (C-S-05S), lipstick, pink on polyester/cotton (P-S-16), fluid make-up on cotton (C-S-17), high discriminative sebum BEY on polyester/cotton (PC-S-132), beef fat on cotton (C-S-61) and butterfat on cotton (C-S-10)), showing a preference for butterfat on cotton (C-S-10) (for details see **Figure 1**); it has the capacity to degrade and clean spinning oils all raw textiles (61488F1, 3X58, 67007, 61988F1, 5237-00 and E03130) (for details see **Figure 2**), and do show PETase and BHETase activities, being capable of degrading polyester-based fabrics (for details see **Figure 3**); T_{opt} : 25-40°C; T_d , $41.70 \pm 1.29^\circ\text{C}$; pH_{opt} , 7.0 to 10.0; $t_{1/2}$ in washing liquor, 3.5 h.

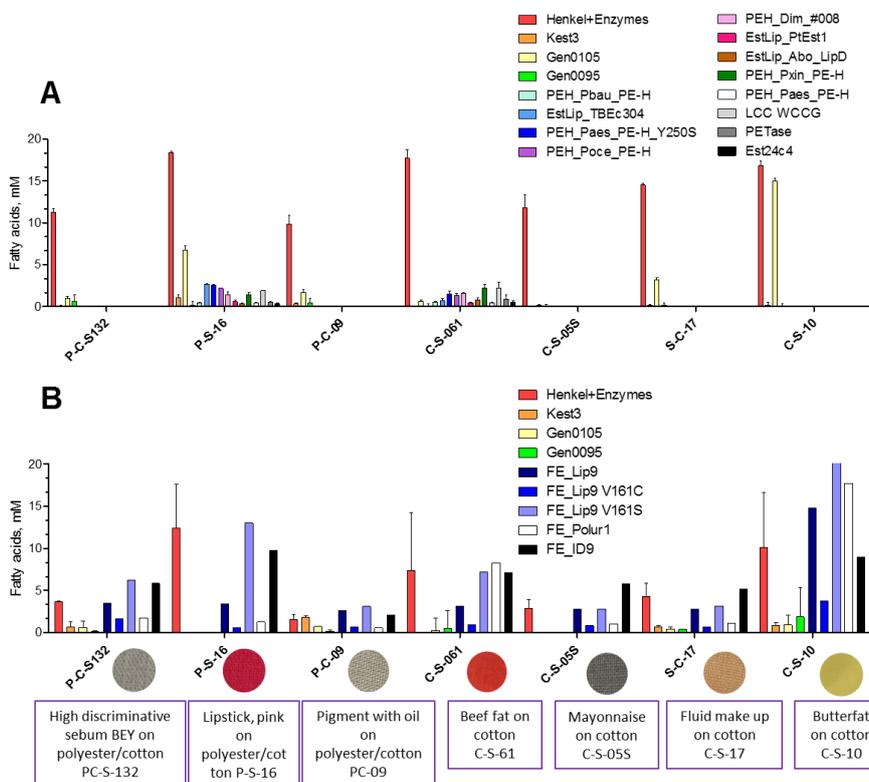


Figure 1. Enzymatic preparations capable of degrading stained swatches. Tests were performed as detailed in D3.2 using a buffer (A) or washing buffer (B).

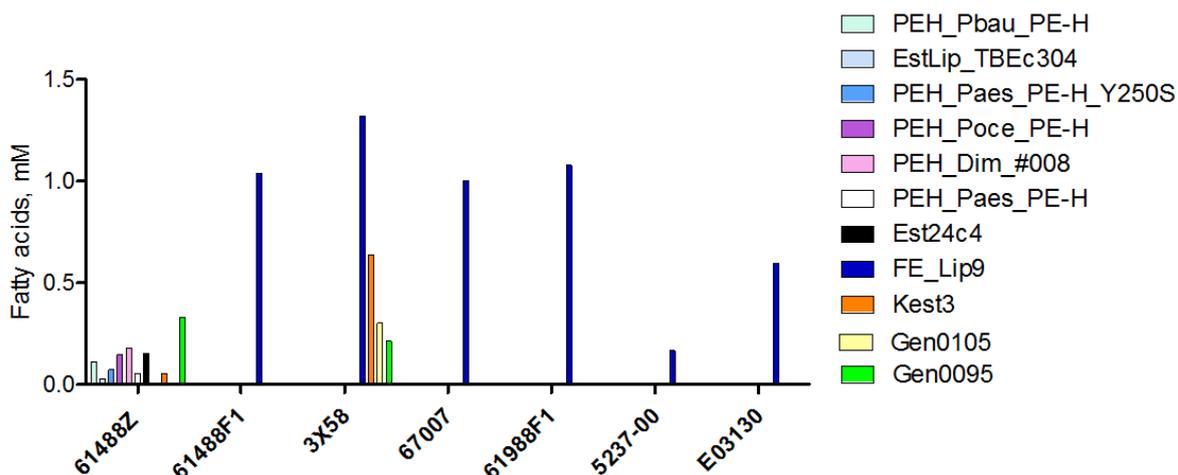


Figure 2. Enzymatic preparations capable of removing spinning oils from raw Schoeller fabrics. Tests performed as detailed in D3.2.

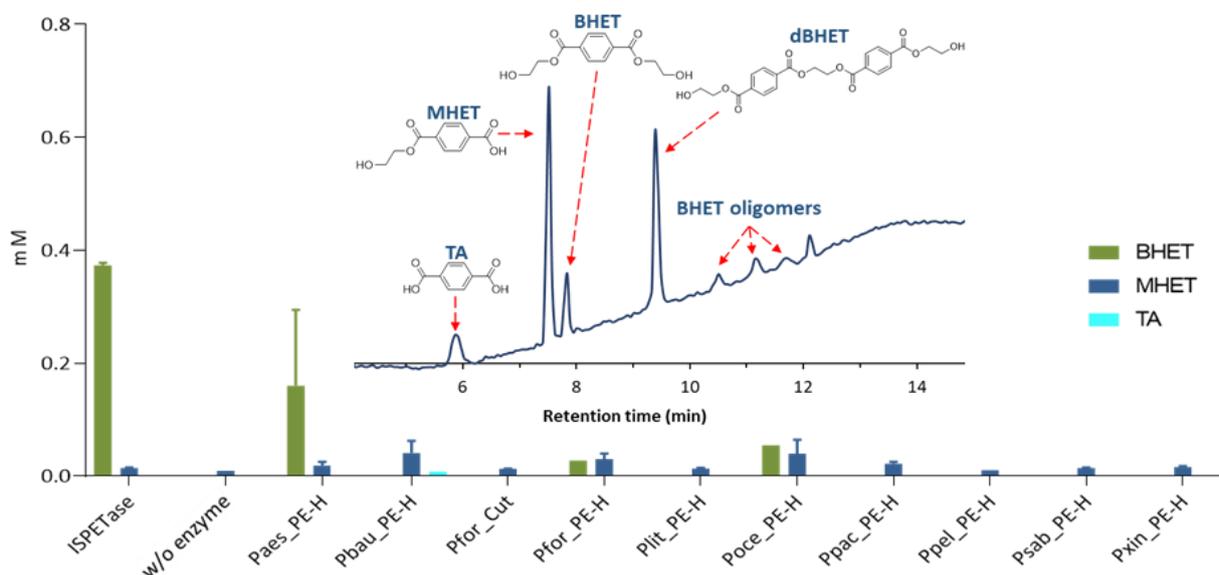


Figure 3. Enzymatic preparations capable of degrading sample textile 4-b 3X58 (VORB, 100% PES 100g/m²) pretreated by alkaline boiling, provided by partner Schoeller. Inset, represent the capacity of FE_Lip9 to degrade this fabric with the indication of the HPLC chromatograms with degradation products observed. Tests, by which the amount of degradation products (BHET, MHET and TA) could be quantified, were performed as detailed in D3.2.

As detailed in Deliverable **D5.1**, FE_Lip9 was found as having characteristics of interest for the detergent and textile sectors. Indeed, FE_Lip9 was among the lipases most efficient at degrading dyed fabrics tested (see Deliverable **D5.1**, and above). After obtaining these results, a classical engineering approach was also implemented with the aim of improving their catalytic activity against these substances. The engineering of FE_Lip9 was first approached by the analysis of functional hot spots found through hotspot wizard server (<https://loschmidt.chemi.muni.cz/hotspotwizard>), using PET₄ for docking with SwissDock server (<http://www.swissdock.ch/>). Two mutants, V161S and V161C, were selected because their priority positions in the hot spot ranking and proximity to catalytic pocket (see **Figure 4**), synthesized, expressed and characterized, as for the wild type variant. The results revealed that the V161S mutant was more active than the native enzyme (Lip9) and the V161C mutant, for most of the stains (see **Figure 1**), being capable of degrading pigment with oil on Polyester/Cotton - 90 cm width (PC-09) 2.4-fold more efficient than the benchmark product at a concentration as low as 0.07% w/v.

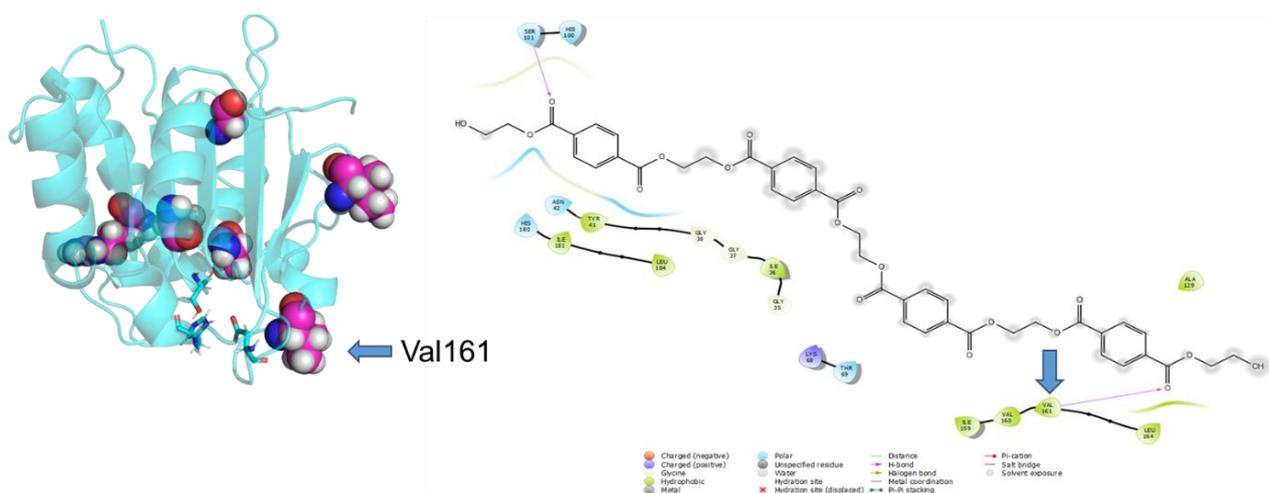


Figure 4. 3D structure of FE_Lip9 representing the main hotspots. The position of Val161 is highlighted. The target substrate used for docking, PET₄, and its ligand diagram (that represent 4 Å distant residues) are shown on the right side.

We further used molecular mechanics simulations of substrate binding with the Monte Carlo code PELE to explore the active site; considering the conservation of the residues and the amino acid frequencies in related sequences, we suggested some potential mutants that could improve the activity of the lipase. For this, we used molecular mechanics simulations of substrate binding with the Monte Carlo code PELE. We used triolein or glyceryl trioleate (TOL) and a tetramer of polyethylene terephthalate (PET₄ or PET for simplification) as substrates to evaluate the affinity of the active sites for both types of substrates. Triolein was selected as a target lipid for engineering the lipase character of FE_Lip9 for detergent application, and tetramer of polyethylene terephthalate from engineering the PETase character for textile application. The protein structure model of Fe_Lip9 obtained from AlphaFold 2.0 (Model 1.1.1) was prepared with PrepWizard from Maestro (Schrodinger), to check all hydrogens atoms and set the correct protonation states. We docked the substrates to the active site and selected the best pose for each protein-ligand combination, in order to have a starting pose for PELE’s simulations. This selection was based on distance and Glide score.

As shown in **Figure 5**, there are several goods candidates that improve the energetics when compared with wild type when PELE simulations were performed. Substrates were constrained into a spherical box with a 25 Å radius to compare the affinity to the modified active sites. In brief, the designed mutants using triolein were: G11S, I12Y, I12N, N48V, H76N, H76S, A81V, A105S, L108Y, I135N, I135Y, V136F, V137C, V137S, L160A, L160F, and the double mutants I135Y/I12N and I135Y/I12Y. As shown in **Figure 6**, there are several goods candidates that improve the energetics when compared with wild type when PELE simulations were performed. Substrates were constrained into a spherical box with a 25 Å radius to compare the affinity to the modified active sites. In brief, the designed mutants using triolein were: G11S, I12Y, I12N, N48V, H76N, H76S, A81V, A105S, L108Y, I135N, I135Y, V136F, V137C, V137S, L160A, L160F, and the double mutants I135Y/I12N and I135Y/I12Y.

In order to minimise the number of mutations to be experimentally validated, the residues and mutations selected by PELE, were subjected to HotSpot wizard calculations to screen the beneficial mutation points in the context of the evolutionary trace of Lip9. Combining all calculations, a number of residues, from those selected by PELE calculations, were selected as best candidate to focus on (**Figure 7; Table 1**).

PELE results for Felip9 Rational Design Mutants and Triolein (TOL)

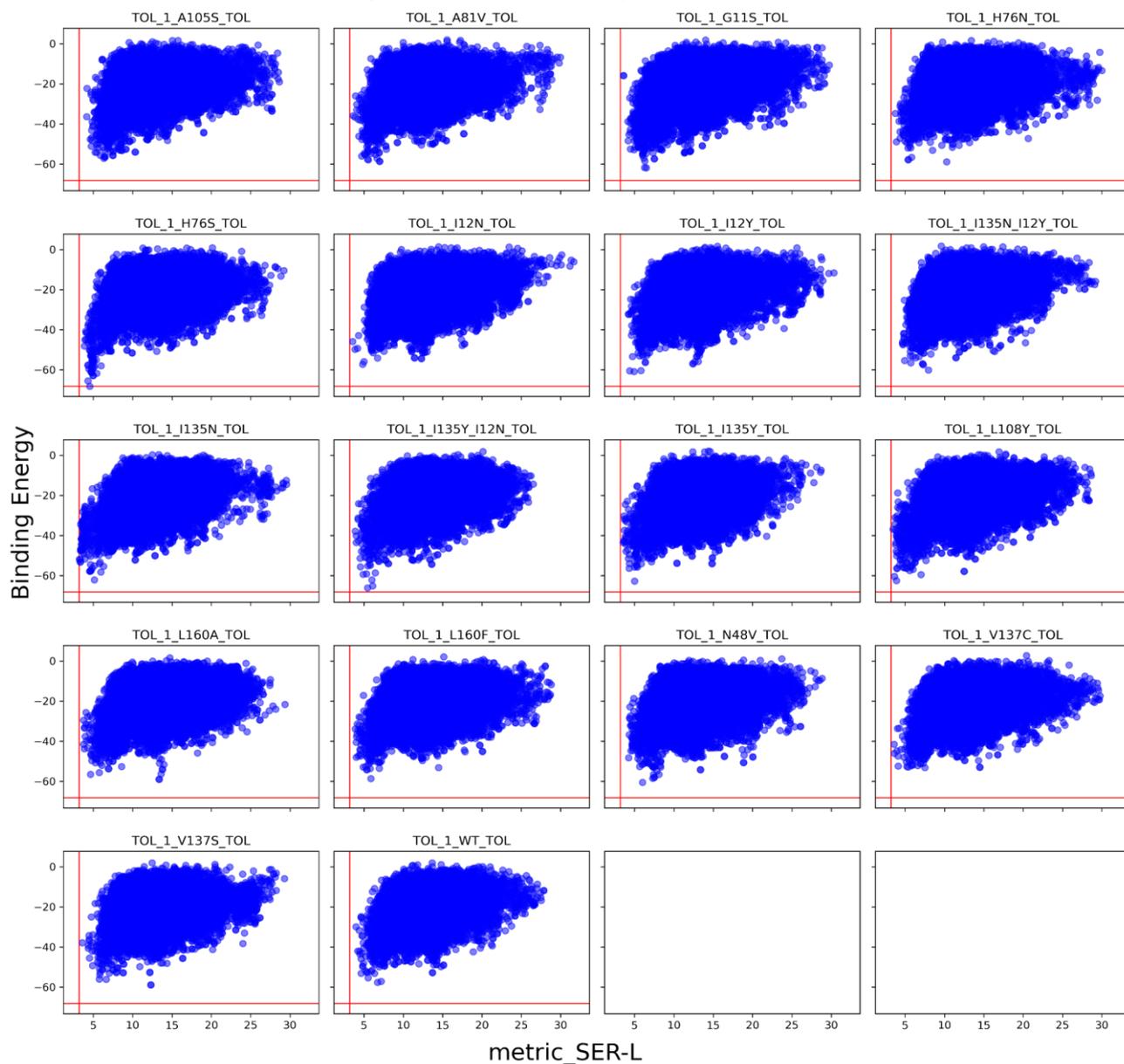


Figure 5. Multiplot of the energy profiles of the active site local exploration with PELE. It is highlighted with a red line the minimum catalytic distance serine-ligand (A) from all the simulations and the minimum binding energy or interaction energy (kcal/mol).

PELE results for Felip9 Rational Design Mutants and PET tetramer

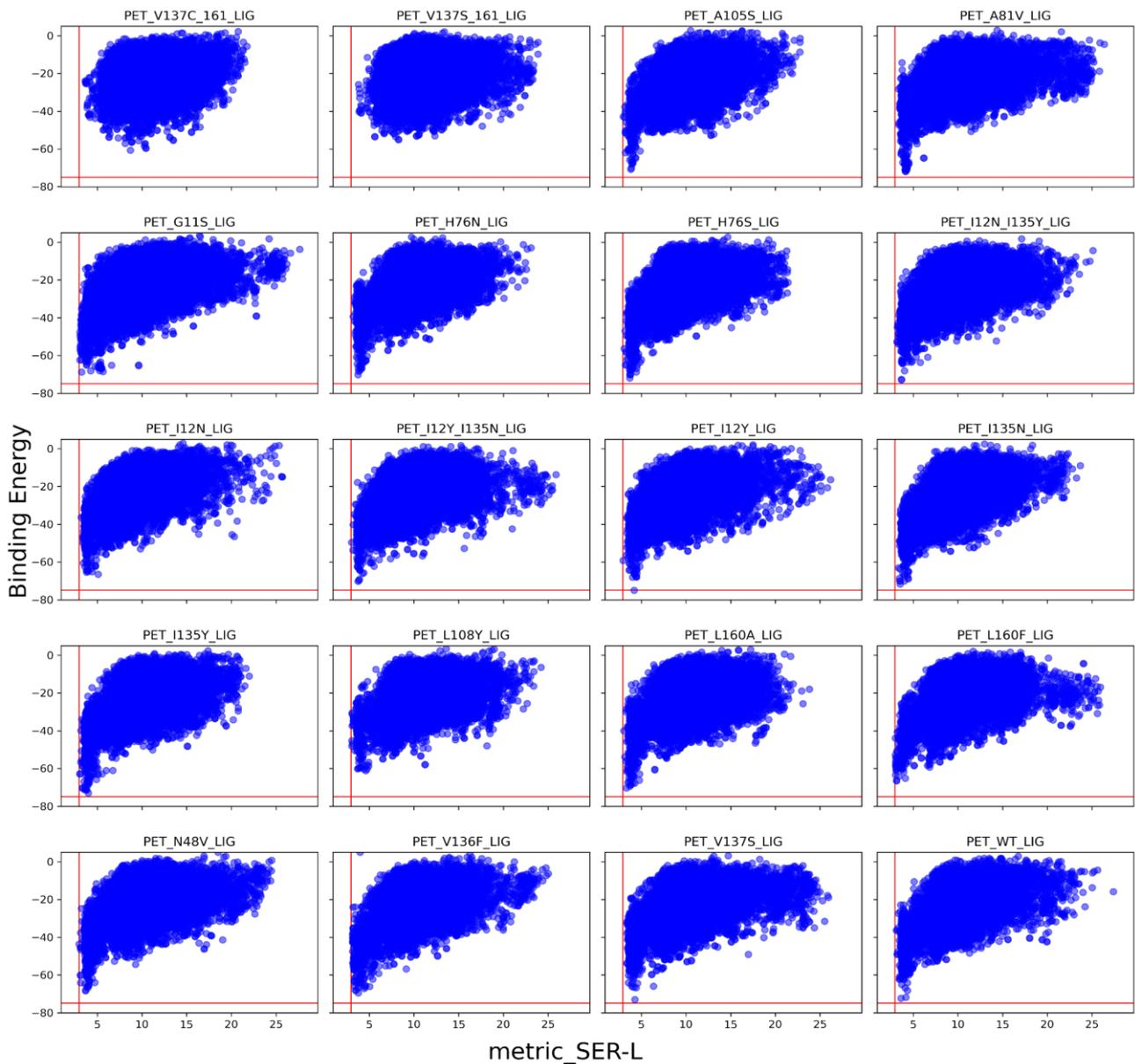


Figure 6. Multiplot of the energy profiles of the active site local exploration with PELE. It is highlighted with a red line the minimum catalytic distance serine-ligand (A) from all the simulations and the minimum binding energy or interaction energy (kcal/mol).

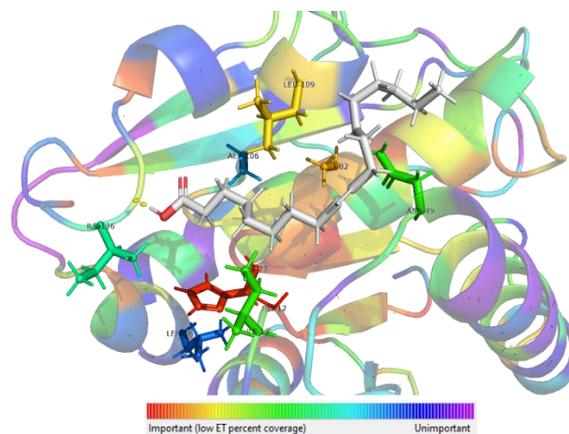


Figure 7. Lip9 3D model colored by evolutionary trace coverage value (red-important, violet-less important), oleic acid (white) is positioned by swissdock, the positions suggested are shown as sticks.

Table 1. Selected positions in FE_Lip9 by HotSpot wizard calculations, with the coverage value obtained by evolutionary trace and the aminoacids present in the MSA of trace.

Residue	AA type	Coverage	Variability
106	A	0.69444	.TMNSQLGDCIYAVP
136	I	0.54444	TEVILK.MAFQ
13	I	0.38333	VLINQTFYMW
109	L	0.17778	MLNF.REQA
82	A	0.15000	LPTNSVIAM
12	G	0.01667	GA
77	H	0.04444	HYNFLW
161	L	0.75000	VGIERPKMAT.LCNDS
49	N	0.40000	AINSTVHLQRKM

Finally, in order to obtain thermostable mutants of FE_Lip9 we performed a stability hot spot assay (hotspot wizard webserver) by sequence consensus. **Figure 8** shows the residues the modification of which may have an effect on thermal stability. In brief, these residues include I120, I126, S134, V152 and Y192 (**Table 2**).

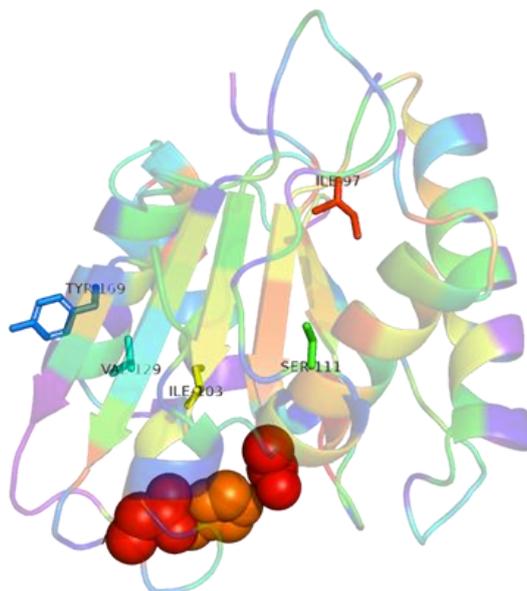


Figure 8. Lip9 3D model colored by evolutionary trace coverage value (red-important, violet-less important), oleic acid (white) is positioned by swissdock, the positions suggested are shown as sticks.

Table 2. Selected positions in the MSA of trace by HotSpot wizard calculations, that may have effects on thermal stability of FE_Lip9. Suggested mutations are highlighted.

Position	Mutant suggested	Variability
152 Val	Trp/Asp	VLEIMFYW
134 Ser	His/Ala	VIACPG.LRTSYQ
126 Ile	Phe	VLIGTFM
120 Ile	Val	VIA
192 Tyr	Gln	GNYK.HAVTSILERQDM

From all the mutations selected to increase the lipase character, the PETase character and the thermostability, 13 were selected for experimental validations, besides the 2 mutations (Val161Ser and Val161Cys) discussed above. They are listed in **Table 3**.

Table 3. Selected mutations to increase the PETase and lipase character and thermostability of FE_Lip9.

Enzyme	Target	Solubility	Expression (mg/L)	Relative activity (%)							Degradation products (μ M)	
				Tri-C ₈	Tri-C ₁₀	Tri-C ₁₂	Tri-C ₁₄	Olive oil	Palm oil	Coconut oil	BHET	nPETc
Lip9 WT		Yes	1.0	100	100	100	100	100	100	100	2971±24	2238±40
Lip9 Val161Cys	Engineering PETase character	Yes	2.0	118.9	115.1	84.1	111.4	80.0	192.8	150.7	2880±13	1581±41
Lip9 Val161Ser	Engineering PETase character	Yes	3.0	123.9	98.1	91.7	102.3	168.9	204.8	179.2	3003±29	3342±63
Lip9 I36N-I159Y	Engineering lipase character	Yes	0.4	2.2	5.3	8.8	0.0	0.0	8.5	2.0	77±22	0
Lip9 H100S	Engineering lipase character	Yes	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0
Lip9 A105F	Engineering lipase character	Yes	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0
Lip9 I36Y	Engineering lipase character	Yes	1.5	7.9	30.4	21.2	14.4	0.0	30.7	21.8	0	0
Lip9 N72Y	Engineering lipase character	Yes	2,1	117.6	95.8	183.6	57.6	35.6	236.4	140.0	251±10	110±10
Lip9 A105V	Engineering lipase character	Yes	4.8	3.9	0.4	0.3	0.0	0.0	0.0	0.6	0	0
Lip9V152W	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 V152E	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 S134H	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 S134A	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 I126F	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 I120V	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Lip9 Y192Q	Thermostability	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

The mutant genes were synthesized, and the mutants expressed and characterized, as for the wild type variant. In addition to the two mutants described above (Val161Ser and Val161Cys), the expression and characterisation of other 6 mutants has been completed. As shown in **Table 3**, all eight mutants were found to be produced in soluble form, albeit at different yields, with only 6 being active. Only V161S, V161C, and N72Y did show activities higher than the wild type. Interestingly, whereas mutations all three mutations increased the lipase character by 1.4-2.3-fold by meaning of higher activity toward long triglycerides such as palm and coconut oil, V161S was the only one increasing the PETase character, by meaning of the degradation of nano-sized PET particles (nPETc).

We will continue testing the other mutants, so that to deepen into the effect of the selected mutations in the lipase character, the PETase character and the thermostability of FE_Lip9.

2.2. Engineering FE_Lip5 by lid domain remodelling: for detergent applications

With the idea of opening new doors in lipase enzyme engineering, we tested the possibility to enhance the hydrolytic activity towards larger triglycerides through lid domain design. Indeed, lipases display a preference for hydrophobic substrates, and thus, they work at the oil-water interface. To efficiently cleave the ester bonds of liposoluble molecules, lipases share a common feature, known as the lid domain. This structural motif is placed over the active site and protects the hydrophobic cavity from the polar solvent when the substrate is absent. At organic-aqueous interphases, the lid domain is opened, allowing the access of the substrate to the active site, and thus, having the lipase in an active conformation. This activity can be achieved due to the amphipathic nature of the lid domain, where the hydrophilic residues face the solvent, while the hydrophobic ones are turned towards the catalytic site in the closed conformation. Once the enzyme opens the lid domain, the hydrophobic part of the lid helps in binding lipo-soluble substrates around the active site. The specific residues in the lid domain are extremely correlated with the activity and specificity of lipases. This has been widely proven by different articles on lid engineering where they show how dramatically the activity and specificity of lipases can change. Thus, the lid domain is an essential “hot spot” for tailoring lipases toward the user’s needs. In FuturEnzyme we used the engineering of the lid domain of the esterase FE_Lip5 (retrieved from MarRef - Marine Metagenomics Database; T_{opt} , 25-45°C, pH_{opt} 6) to switch from the hydrolysis of small-chain triglycerides to long-chain ones through lid swapping with the lid domain of *Rhizopus oryzae* lipase. Since the engineering of the lid domain can lead to drastic changes in the activity and specificity of the lipase, we visually inspected this structural motif (**Figure 9**). The lid domain (FRGTEITQIKDWLTDA) seemed to have a tryptophan residue at the inner side, hindering its full opening to bind medium and long-chain triglycerides, which was confirmed by molecular simulation with the PELE software and further experimentally confirmed. We found that substituting this lid domain by the one of *Rhizopus oryzae* lipase different (FRGTNSFRSAITDIVF) that does not contain any tryptophan residue enable the opening of this structural motif, allowing the hydrolysis of bulky triglycerides, as it was experimentally confirmed, as the wild type enzyme did not convert glyceryl tridecanoate and tridodecanoate, coconut oil and palm oil, but the FE_Lip5-lid did (specific activity (U/mg) from 534 to 14.4 units/mg). This provided experimental evidences that true lipases could be designed from an esterase by lid swapping.

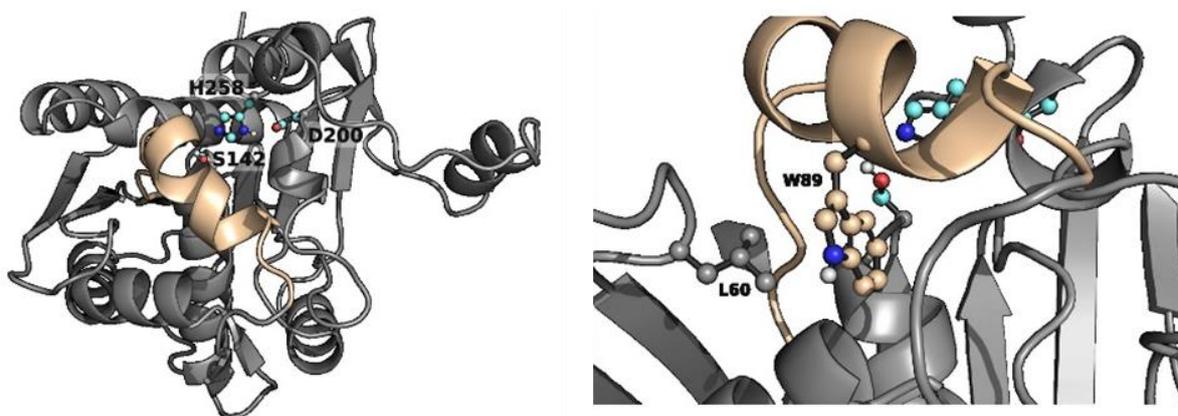


Figure 9. Left, FELip5 3D model highlighting the catalytic triad (with C atoms colored in cyan) and the lid domain (colored in wheat). Zoom to the positions of W89 and L60 in the lid domain of FELip5 is shown on the right side.

In addition to the lid swapping approach, the results of the simulations showed that the FE_Lip5 lipase had a closed metastable state at 8.99 Å and an opened one at 14.43 Å, meaning the lid domain opens up to 5.44 Å. We created a variant aiming at opening more the lid domain of the lipase (**Figure 9**). The designed mutant replaced W89 for a less bulky residue, but still a reasonable change according to the BLOSUM62 matrix, namely methionine. To compensate for this change and avoid the access of water molecules in the active site to the closed conformation, L60, a residue not placed in the lid domain (**Figure 9**), was mutated to phenylalanine. Then, the same type of simulation was performed on the double mutant. The results gave a similar closed metastable conformation (at 8.81 Å), but a more opened one (at 15.78 Å). Thus, the difference in the opening distance between the double mutant and the WT was around 1.5 Å, meaning the variant could accommodate bulkier triglycerides. The analysis of this mutant (FE_Lip5L60F-W89M) is under progress.

2.3. PEH_Paes_PE-H engineering: for textile applications

As detailed in Deliverable **D5.1**, one of the enzymes was found priority because having characteristics of interest for the detergent and textile sectors, PEH_Paes_PE-H: Activity, lipase, PETase; Partner, UDUS; Source, *Halopseudomonas aestusnigri*; Nominated as additive for Henkel detergent and for end-of-life fabric recycling, because high specific activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (see **Figure 1**; **Figure 10A**); although, it is not the enzyme with the highest levels of activity against model lipase substrates (**Figure 10B**), the capacity to degrade fatty stains and its full or high percentage of residual activity in the presence of surfactants, i.e. washing liquor, the availability of experimental structural data (structural features of cutinases/PETases), and activity on PES-fabrics (see **Figure 3**) makes an interesting candidate of it. Literature-informed mutagenesis of enzyme Paes_PE-H yielded a variant that showed increased monomer-release from PET polyester fabrics, outcompeting the benchmark enzyme ISPETase under assay conditions (**Figure 10B**). In addition, the mutation Y250S improved the degradation of a PET-based sample textile 4-b 3X58 (VORB, 100% PES 100g/m²) pretreated by alkaline boiling. PET monomer release from Schoeller sample textile increased by Paes_PE-H mutant (**Figure 10C**). In addition, we detected for Y250S a 2.5x higher release of fatty acids during the first 2h of incubation than for the wild type. With Lipstick stain, the activity of the mutant for even 4.4x higher.

2.4. PET46 from Candidatus Bathyarchaeota engineering: for textile applications

As detailed in Deliverable **D5.1**, an additional enzyme was found priority because having characteristics of interest for textile sectors, PET46: Activity, PETase; Partner, UHAM; Source, *Candidatus Bathyarchaeota*; Nominated for end-of-life fabric recycling because availability of experimental structural data and activity on PET polyester fabric (**Figure 3**). We designed and characterized PET46 (NCBI accession RLI42440.1), the first enzyme from archaeal origin reported to hydrolyze PET polymer. The enzyme is encoded in the metagenome-assembled genome (MAG) of the *Candidatus Bathyarchaeota* archaeon B1_G2, a member of the TACK group that was found at the Guaymas Basin. The experimentally established crystal structure of the protein is similar to bacterial PET-degrading enzymes, but reveals several unique features (**Figure 11**). PET46 is a promiscuous feruloyl esterase that hydrolyzes MHET, BHET, PET-trimer (3PET) and PET polymers.

Based on the docking results, we identified two amino acids, A46 and A140, nearby both predominant docking poses that might be relevant for the substrate accessibility and binding. Introducing the larger substitutions A46V and A140I should thus impact the catalytic activity. We further identified K147, which possibly interacts with docked poses from the second-largest cluster. Variant K147A abolishes this interaction and widens the binding groove. We then proceeded to incubate PET46 WT and all the constructed variants (including the PET46 Δ lid) on 3PET at 30, 60 and 70°C. At the two highest temperatures, we observed a very similar activity pattern, where PET46 WT, K147A, and A46V degraded all the 3PET to MHET and TPA within the first 3 h (**Figure 11**).

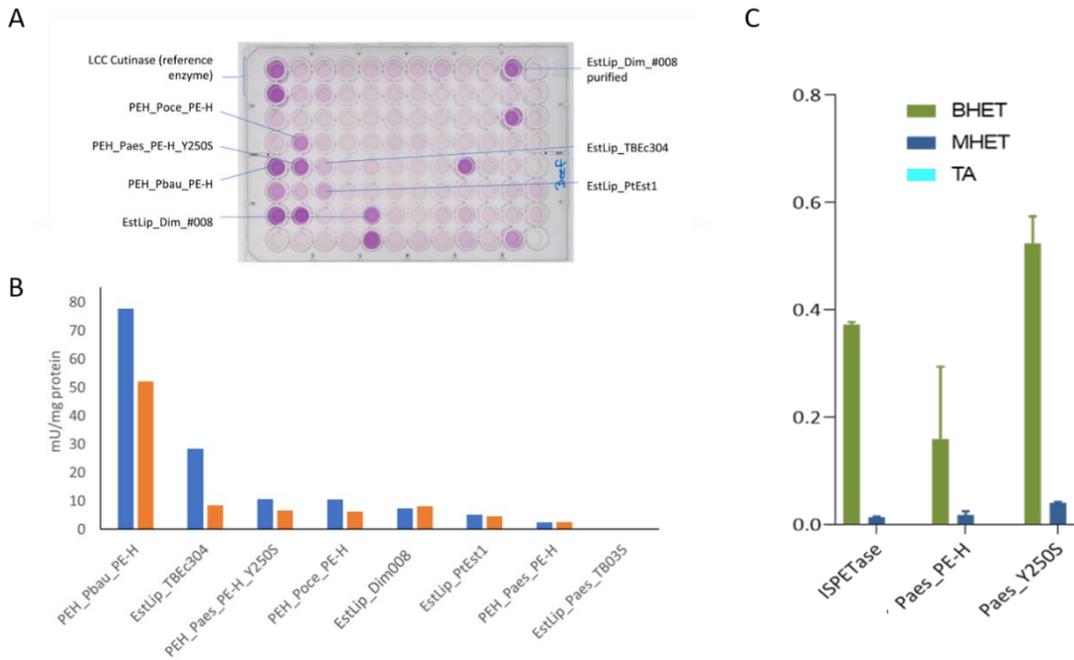


Figure 10. Hydrolases activities on stained swatches. Activity of the selected PEH_Paes_PE-H hydrolase compared to other priority hydrolases upon incubation with standard stained fabric material, determined by measuring the released fatty acids. (A) Screening plate assaying the fatty acid concentration after overnight incubation with beef fat-stained fabric. The darker the violet, the higher the fatty acid concentration. (B) Approximated specific activity of purified enzymes [1 U= 1 μ mol fatty acid released per minute, determined with a serial dilution with oleic acid) after 2 h. (C) Release of building blocks after incubation of pre-treated PET fabrics of wildtype Paes_PE-H and the engineered variant Y250S-H in comparison to Is_PETase (equimolar amounts of enzymes).

PET46 A140I performed slightly worse, while PET46 Δ lid could only convert half of the 3PET after 72 h incubation. Interestingly, the A46V mutant was 3.17-fold better than the WT enzyme at a temperature of 30°C after 24h using 3PET as substrate. In all experiments, we were not able to detect any BHET. Together with the previously obtained MHET-TPA profiles over time, we assume degradation happens at the polymer chain end (exo-activity), where 3PET is hydrolyzed to MHET units, which are subsequently converted to TPA and ethylene glycol (EG). Interestingly, after incubation with PET46 WT and A46V, no MHET was measured (**Figure 11**), suggesting these two enzymes being more effective.

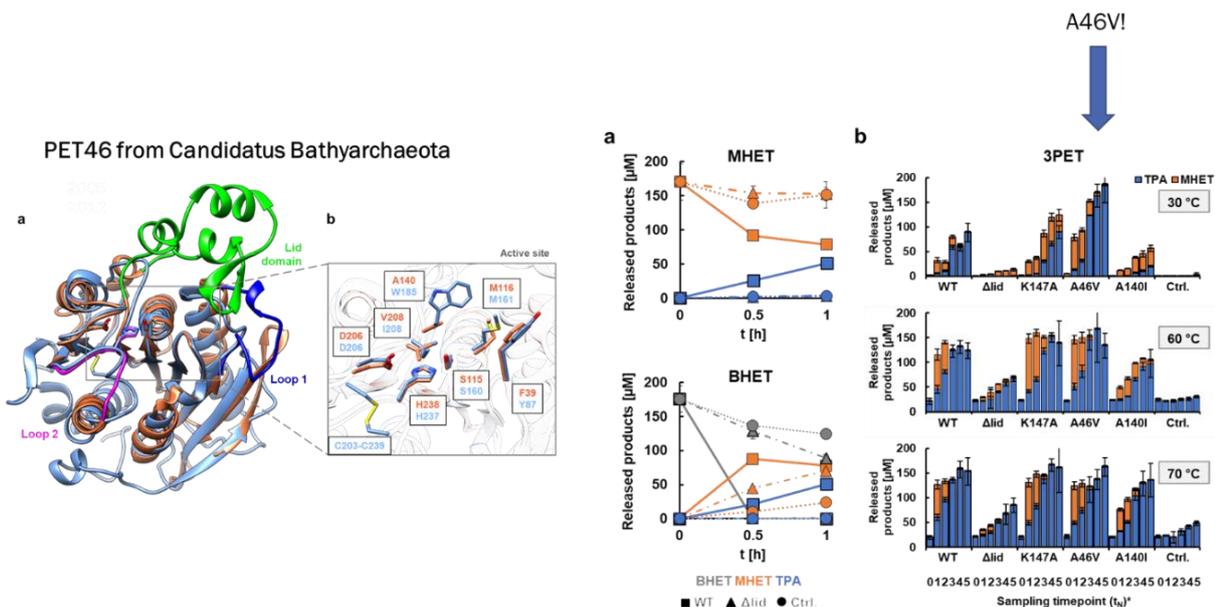


Figure 11. PET46 uses the lid domain to effectively degrade MHET, BHET and 3PET. PET46 WT can degrade both BHET and MHET to TPA and EG at 70°C, but the lid-less variant PET46 Δ lid can only convert BHET to MHET (A). PET46 and the produced variants –1 degrade 3PET at 30, 60 and 70°C (B). * $t_0=0$ h; $t_1=3$ h; $t_2=6$ h; $t_3=24$ h; $t_4=48$ h; $t_5=72$ h.

2.5. Gen0105 and Kest3 engineering for detergent application

As detailed in Deliverable **D5.1**, two additional enzymes were found priority because having characteristics for the detergent sector, Gen0105 and Kest3: Activity, lipase; Partner, Bangor; Source, Abano Terme, Italy (isolated strain *Fervidobacterium pennivorans* DSM 9078; genome accession number CP003260.1) (Kest3) and metagenome of a mesophilic anaerobic digester, Evry, France (Gen0105); Nominated based on the initial screens with natural coconut oil and/or stained swatches that include lipstick, butterfat, fluid makeup, pigment with oil and sebum, comparable with Henkel enzymes mix (**Figure 1**); however, activities significantly decreased in washing liquid.

Two approaches were applied for enzyme stabilization: ancestral phylogenetic reconstruction and energy-based approach, evaluating change in free energy upon mutation using FireProt tools of the Loschmidt laboratories (<https://loschmidt.chemi.muni.cz>). Based on energy optimization approach 6 mutations were introduced to the Kest3 lipase: G106P, V118I, V135I, T187L, S225F, A275M. Kest3 mutant protein with all 6 mutations combined resulted in -10kcal/mol less deltaG than wild type. 8 mutations were introduced to the Gen0105: A12P, Q54W, A57W, A65P, N81Y, T84I, G115A, and Q288M, resulting mutant decreased deltaG for -28.3kcal/mol in comparison with wild type enzyme. For details see **Figure 12A-B**. The analysis of the mutants in under progress.

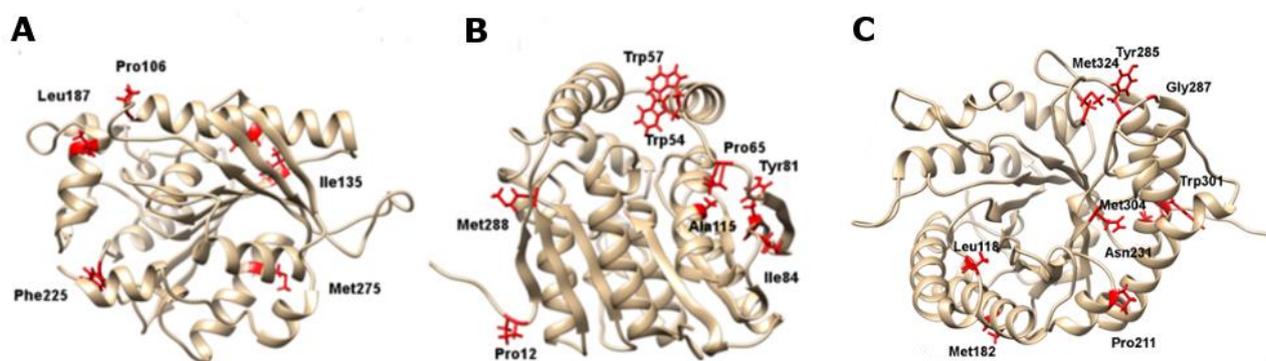


Figure 12. Selection of mutants in Kest3, Gen0105 and Gen0095. 3D structure of Kest3 (A) and Gen0105 (B) with mutated residues (FireProt) highlighted in red. (C) 3D structure of Gen0095 (see below). The mutated residues (Energy optimization, FireProt) are highlighted in red.

Further, mutated sequences were selected by using the ancestral sequence reconstruction and FireProt energy minimisation approaches. One mutated sequence was synthesized through energy minimisation, while two ancestral reconstructed sequences, were selected and subsequently synthesized for each Kest3 and Gen0105. The ancestral reconstructed sequences for Kest3 (Node58 and Node88; **Figure 13**) exhibited 63.4% and 40.1% of identity to the Kest3 wild type. As for the ancestral reconstructed sequences for Gen0105 (Node 152 and Node 186; **Figure 14**) they demonstrated 42% and 51.8% identical to the Gen0105 wild type.

Protein expression in 1 litre medium (Luria-Bertani) identified 2 ancestral Gen0105 sequences (Node152 and Node186) and Kest3 (Kest3_EMF) energy minimisation mutant being soluble (**Figure 15A**). However, the activity test revealed that Kest3_EMF mutant was inactive when screened with *p*-nitrophenyl-substrate (**Figure 15B**), and no thermostable mutant was identified (**Figure 16**). The Gen0105 ancestral sequences of Node 152 and Node 186 exhibited activity with *p*-nitrophenyl-substrates ranging from C₆-C₁₂. Node186 displayed a similar activity pattern to Gen0105 wild type, whereas Node152 significantly lost activity with *p*-nitrophenyl-C₆-C₁₂ substrates (**Figure 17**). The thermostability screening of the Gen0105 ancestral sequences demonstrated enhanced properties with model *p*-nitrophenyl-C₈ substrate. After 1h of incubation at 80°C, Node152 retained up to 38% activity, and Node 186 retained 40% of activity during 1h incubation at 60°C (**Figure 17**). However, tests using the real substrates 3PET, amorphous PET (data not shown) and contaminated textile screens did not confirm higher efficiency of the ancestral reconstructed Gen0105 enzymes (**Figure 18**). Nevertheless, further investigations are currently being undertaken.

Kest3 Ancestral reconstruction

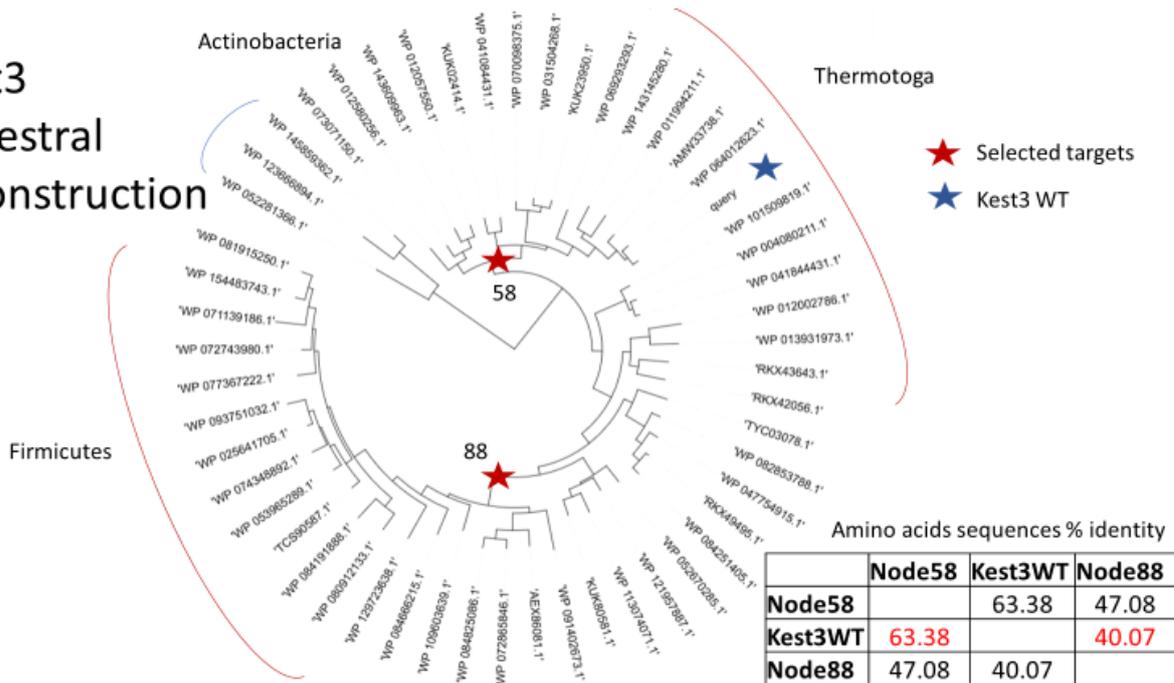


Figure 13. Selection of mutated sequences of Kest3 by the ancestral sequence reconstruction.

GEN0105: Ancestral reconstruction

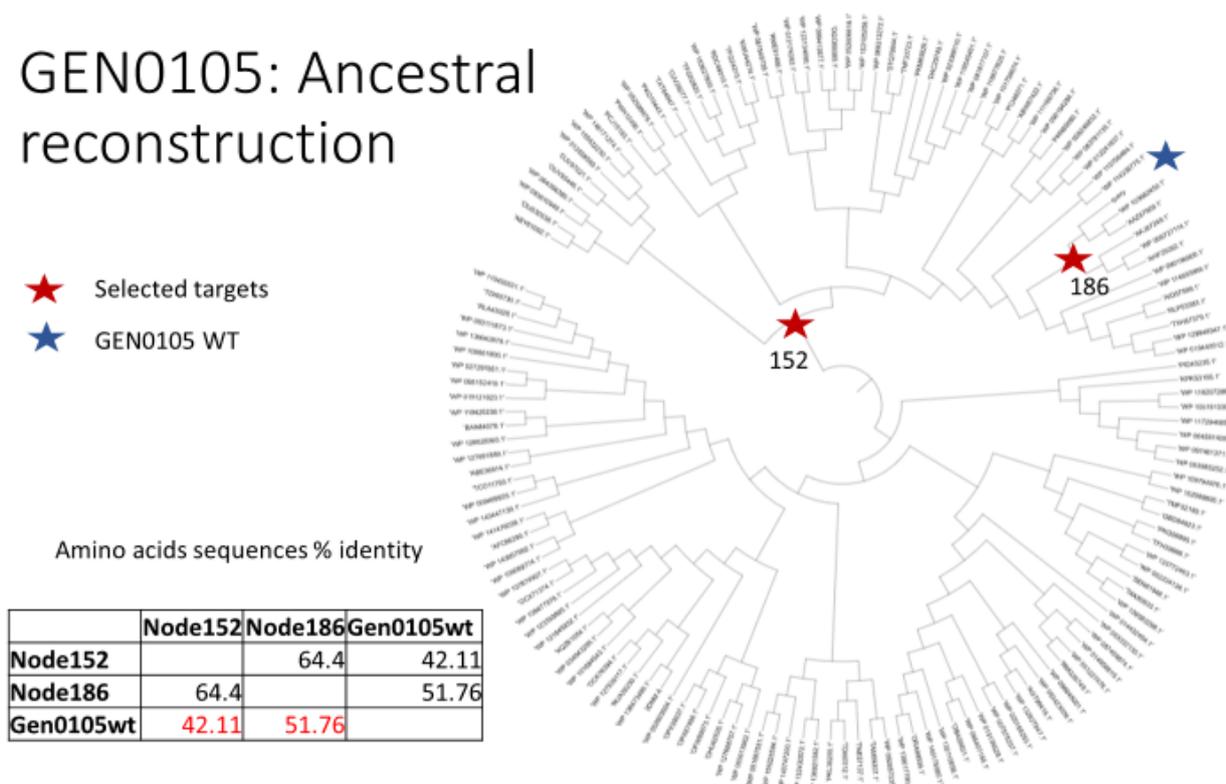


Figure 14. Selection of mutated sequences of Kest3 by the ancestral sequence reconstruction.

A

Protein name	Function	Stabilization method	Solubility
Gen0105_EMF	PETase	Energy minimization	-
Node152	PETase	Ancestral reconstruction	+
Node186	PETase	Ancestral reconstruction	+
Kest3_EMF	Lipase	Energy minimization	+
Node58	Lipase	Ancestral reconstruction	-
Node88	Lipase	Ancestral reconstruction	-

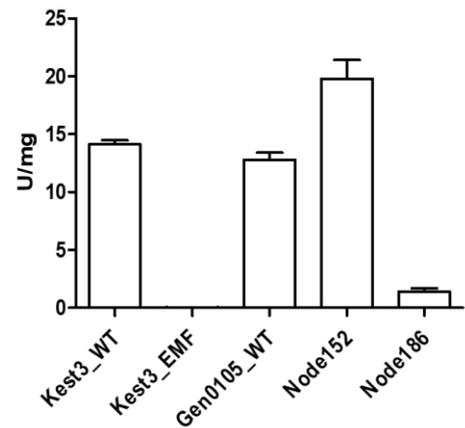
B

Figure 15. Protein expression level of the 2 ancestral Gen0105 (Node152 and Node186) and Kest3 (Node58 and Node88) proteins (A). Activity level of the 2 ancestral Gen0105 (Node152 and Node186) and Kest3 (Node58 and Node88) proteins, once expressed in *E. coli* (B).

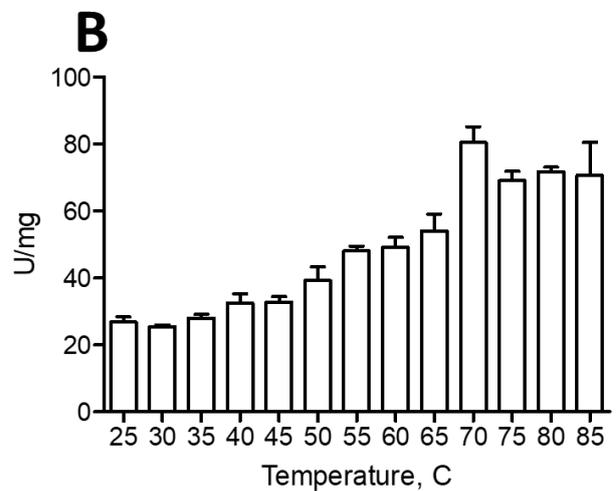
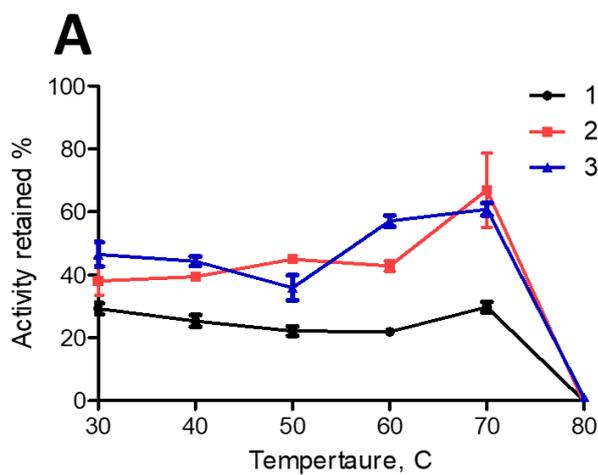


Figure 16. Thermostability (A) and optimal temperature (B) for activity of Kest3 protein. None of the mutants (Node58 and Node88) turned thermostable, and data are not shown.

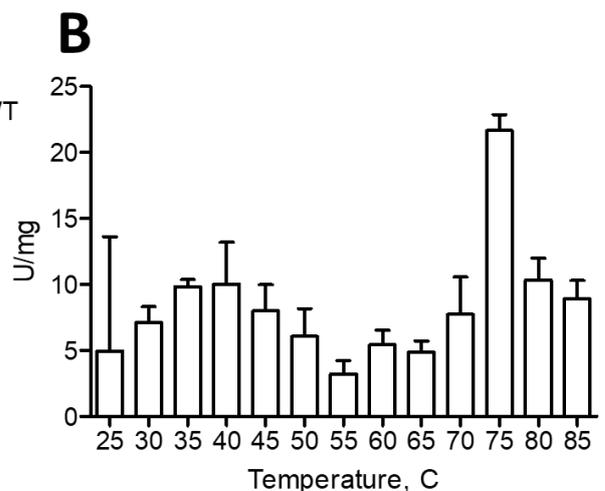
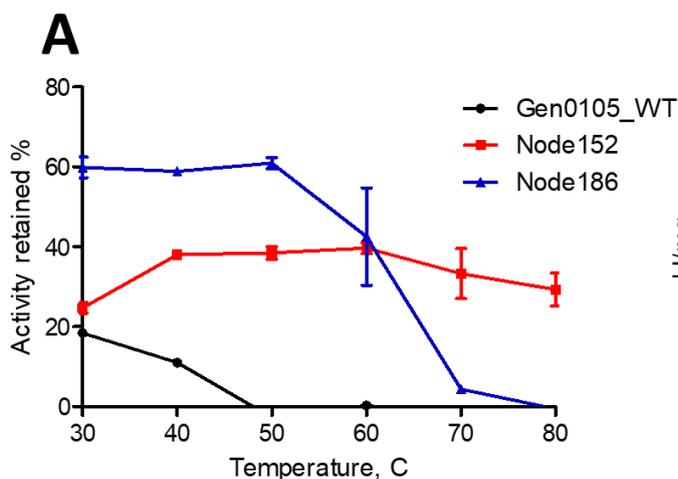


Figure 17. Thermostability of the Gen0105 WT and 2 ancestral Gen0105 (Node152 and Node186) (A). The optimal temperature for activity of the Gen0105 wild type protein (B).

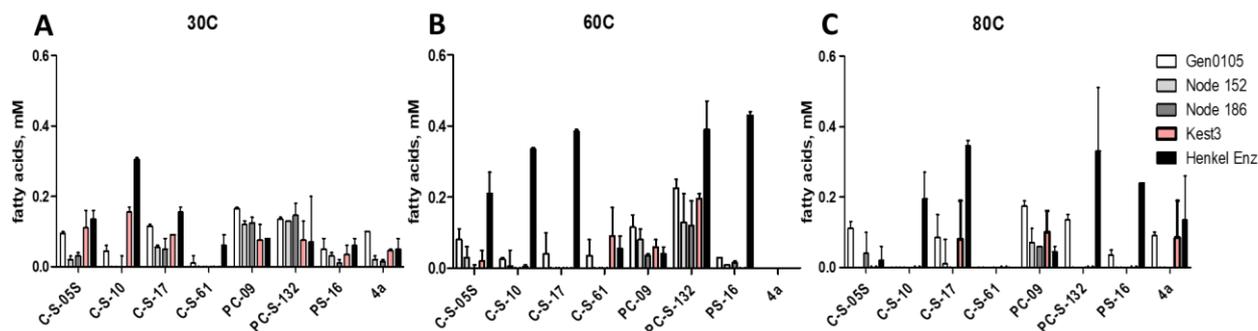


Figure 18. Gen0105 WT, ancestral reconstructed proteins Node152 and Node186, Kest3 WT and Henkel Enzymes textile decontamination screening.

2.6. Gen0095 engineering for textile applications

As detailed in Deliverable **D5.1**, an additional enzyme was found priority because having characteristics for the detergent sector, Gen0095: Activity, cellulase with activity comparable with benchmark from *Trichoderma* (Sigma 912-54-8); Partner, Bangor; Source, metagenome of a mesophilic anaerobic digester, Evry, France; Nominated based on the screens with stained swatches, and capacity to degrade butterfat on cotton (**Figure 1**). Based on energy optimization approach (<https://loschmidt.chemi.muni.cz/fireprotweb/>) 9 mutations were selected: N118L, S182M, Q211P, S231N, N285Y, N287G, S301W, T304M, N324M (**Figure 12C**), with all mutations resulting in -20.2kcal/mol less deltaG than the wild type protein. We will continue characterizing these mutants.

2.7. TB035 engineering for textile applications

EstLip_TB035 was suggested for further exploration in the first place because of its remarkable resistance towards denaturing agents, including surfactants. Further, it shows a rather high structural melting point, although unfolding already starts at a temperature below 40°C. Despite being highly substrate-promiscuous, this enzyme, unfortunately, is not able to hydrolyze long-chain triglycerides. The yet unpublished structure reveals an uncommon topology with two cap domains that may be targeted for engineering. The enzyme originates from *Halopseudomonas aestusnigri*. It is membrane-bound in its original form, linked to membrane lipids by the Sec Typell Lipoprotein secretion machinery which leads to low yields in recombinant expression and hampers purification. To improve expression and purification, a variant was constructed by PCR and restriction/ligation in which the Typell-Sec signal peptide predicted by SignalP5.0 was exchanged for a SecI signal peptide sequence from the *Escherichia coli* protein PelB. The fusion protein could be obtained in a mg /L culture scale.

3. Methodology

3.1. PELE simulations

The starting position of both substrates in the active site of felip9 were obtained using Glide Docking from the Schrodinger Suite with extra precision (XP) settings. We filtered and selected the best poses based on the catalytic distance between the ester carbons of the substrates and the catalytic oxygen of the reactive serine in the active site. The binding interaction between Felip9 and the substrates (TOL and PET) was explored using the PELE (Protein Energy Landscape Exploration) software. PELE is a heuristic Monte Carlo-based algorithm that incorporates protein structure prediction methods. The software begins by sampling various microstates of the ligand through small rotations and translations. The flexibility of the protein is also considered by applying the anisotropic network model (ANM) approach (1). To avoid steric clashes, the side chains of residues near the ligand are also sampled. Subsequently, a truncated Newton minimization is performed using the OPLS2005 force field (2), and the acceptance or rejection of the new microstate is determined using the Metropolis criterion. The Variable Dielectric Generalized Born Non-Polar (VDGBNP) implicit solvent model (3) is used to simulate the influence of water molecules surrounding the protein. We evaluated the impact of different mutations on substrate binding in the active site using a constraint box

radius of 25Å along 10 iterations of 100 PELE steps, utilizing a total of 96 CPUs. To measure the differences in substrate binding, we monitored the number of near-attack conformations adopted by the substrate during the independent PELE-induced fit simulations, as well as the average energy of the ligand in the active site. We considered a near-attack conformation when the distance between one of the ester carbons of the substrates was within 4.5Å of the catalytic serine residue, while maintaining reasonable distances for the remaining hydrogen bonds of the catalytic triad. This can also be observed in scatterplots as an increased minimum compared to the wildtype (**Figures 5 and 6**).

3.2. Docking simulations

For docking studies, we used SwissDock (4) webserver limiting the poses of the ligand to a 10Å radius box, giving the oxygen of catalytic serine atomic coordinates as the center of the box. SwissDock is a web service to predict the molecular interactions that may occur between a target protein and a small molecule. SwissDock is based on the docking software EADock DSS, whose algorithm consists of the following steps: many binding modes are generated either in a box (local docking) or in the vicinity of all target cavities (blind docking). simultaneously, their CHARMM energies are estimated on a grid; the binding modes with the most favorable energies are evaluated with FACTS, and clustered; and the most favorable clusters can be visualized online and downloaded on your computer.

3.3. Hot spots analysis

For designing mutant designing we used HotSpot Wizard v3.1 webserver (5) which is a server developed by Loschmidt Laboratories for automated identification of ‘hot spots’ and design of smart libraries for engineering protein stability, catalytic activity, substrate specificity and enantioselectivity. It implements four different established protein engineering strategies, enabling users to selectively target sites affecting protein stability and catalytic properties. The aminoacid sequences and 3D structures of the target proteins were used as different input data.

3.4. Evolutionary trace

For the selection of the variants once we have a hotspot position to mutate, we make a second analysis to evaluate the impact that this mutation can induce in the stability and folding of the enzyme in the production process. For this study we used the Evolutionary trace server (6) from Lichtarge Lab. With this study we minimize the experimental efforts by focusing in evolutionary relevant positions avoiding critical positions such as catalytic triads or other positions that can have a deleterious effect and picking variants that are present in the MSA performed by the server.

4. Conclusions

This deliverable reported the 8 proteins subjected to protein engineering: Lip9, Lip5, PEH_Paes_PE-H, PET46, Gen0105, Kest3, Gen0095, and TB035. A total of 32 mutants were selected based on a number of approaches (**Table 4**), from which the effect of 26 were experimentally validated and mutants with improved performance identified. We will continue characterizing the designed mutants and transfer to WP6 and WP7 the best selected candidates once all engineering work is completed. That said, we will continue expanding the engineering to additional enzymes which revealed interesting properties. Note that the artificial enzymes designed by the PluriZyme approach are not listed here but in the Deliverable **D5.3**.

Table 4. List of mutants designed per each of the 8 enzyme subjected to engineering.

Nr	ID, name
	Lip9 WT
1	Lip9 Val161Cys

Nr	ID, name
2	Lip9 Val161Ser
3	Lip9 I36N-I159Y
4	Lip9 H100S
5	Lip9 A105F
6	Lip9 I36Y
7	Lip9 N72Y
8	Lip9 A105V
9	Lip9V152W
10	Lip9 V152E
11	Lip9 S134H
12	Lip9 S134A
13	Lip9 I126F
14	Lip9 I120V
15	Lip9 Y192Q
	FE_Lip5 WT
16	FE_Lip5-lid (FRGTEITQIKDWLTDA → FRGTNSFRSAITDIVF)
17	FE_Lip5-L60F-W89M
	PEH_Paes_PE-H WT

Nr	ID, name
18	PEH_Paes_PE-H-Y250S
	PET46 WT
19	PET46Δlid
20	PET46-A46V
21	PET46-A140I
22	PET46-K147A
	Kest3 WT
23	Kest3- G106P, V118I, V135I, T187L, S225F, A275M (Kest3-EMF)
24	Kest3-Node58 (ancestral reconstructed sequence)
25	Kest3-Node88 (ancestral reconstructed sequence)
26	Kest3-EMF (ancestral reconstructed sequence)
	Gen0105 WT
27	Gen0105-A12P, Q54W, A57W, A65P, N81Y, T84I, G115A, and Q288M (Gen0105-EMF)
28	Gen0105-Node152 (ancestral reconstructed sequence)
29	Gen0105-Node186 (ancestral reconstructed sequence)
30	Gen0105-EMF
	Gen0095 WT

Nr		ID, name
31		Gen0095-N118L, S182M, Q211P, S231N, N285Y, N287G, S301W, T304M, N324M (Gen0095-EMF)
TB035 WT		
32		TB035- Secl (Secl signal peptide sequence added)

5. References

1. Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of Fluctuation Dynamics of Proteins with an Elastic Network Model. *Biophys. J.* 2001, 80, 505– 515, DOI: 10.1016/S0006-3495(01)76033-X
2. Banks, J. L.; Beard, H. S.; Cao, Y.; Cho, A. E.; Damm, W.; Farid, R.; Felts, A. K.; Halgren, T. A.; Mainz, D. T.; Maple, J. R.; Murphy, R. et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem.* 2005, 26, 1752– 1780, DOI: 10.1002/jcc.20292
3. Bashford, D.; Case, D. A. Generalized Born Models of Macromolecular Solvation Effects. *Annu. Rev. Phys. Chem.* 2000, 51, 129– 152, DOI: 10.1146/annurev.physchem.51.1.129
4. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res.* 2011 Jul;39(Web Server issue):W270-7. doi: 10.1093/nar/gkr366. Epub 2011 May 29. PMID: 21624888; PMCID: PMC3125772.
5. Sumbalova, L., Stourac, J., Martinek, T., Bednar, D., Damborsky, J., 2018: HotSpot Wizard 3.0: Web Server for Automated Design of Mutations and Smart Libraries based on Sequence Input Information. *Nucleic Acids Research* 46 (W1): W356-W362.
6. Lichtarge, O., H.R. Bourne and F.E. Cohen (1996). "An evolutionary trace method defines binding surfaces common to protein families." *J. Mol. Biol.* 257(2): 342-58.