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

30/11/2022

THE SHORTLIST OF AT LEAST 18 ENZYMES NOMINATED FOR ENGINEERING

D5.1

MANUEL FERRER
CSIC
Marie Curie 2, 28049, Cantoblanco, Madrid, Spain

Document information sheet

Work package:	WP5, Enhancing enzymes through innovative engineering.
Authors:	Manuel Ferrer (CSIC)
Document version:	1
Date:	30/11/2022
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:	CSIC
Dissemination Level:	Confidential, only for members of the consortium (including the Commission Services)
Type:	Report
Due date (months):	18
Contact details:	Manuel Ferrer (mferrer@icp.csic.es)

Summary

The shortlist of at least 18 enzymes nominated for engineering

1. Scope of Deliverable

This deliverable consists in a report detailing the best 18 enzymes nominated for engineering out of the initial set of enzymes subjected to characterization. This deliverable is accompanied by a report detailing the characteristics of each enzyme on the basis of which they were selected. This information is associated with a QR code, which has been made available in the internal FuturEnzyme repository.

2. Origin of the material

Along the already 18 months of the project, different deliverables have been accomplished from which the present one nourishes. To be mentioned:

Deliverables in the frame of WP2:

- D2.2_Set of 250 000 sequences pre-selected (November 2021; to be updated)
  In this deliverable, a set of about 3.2 sequences with interest for our project was retrieved by BLAST and HMM search, complemented with computational screens.
- D2.3_Set of 1000 enzymes selected using motif screens (May 2022; to be updated)
  In this deliverable sequences retrieved in the frame of WP2 (deliverable D2.2) and WP3 (deliverable D3.3) were subjected to a filtering pipeline comprising the following criteria: confirmation of full-length sequence, presence and conservation of all proper domains and catalytic residues (along with MSA), the 3D structure modeling using AlphaFold 2.0, substrates (specified by the manufacturers) docking using Glide software (Schrödinger company) in the active site of the enzymes, and the substrate positioning around the active site with PELE (Protein Energy Landscape Exploration) software from BSC. The pipeline includes also the analysis of sequence coverage, the homology with reported similar sequences, the pair-wise similarity, the network analysis using the MCL (Markov Cluster Algorithm) algorithm. The idea of this pipeline was to select the priority enzymes to work with in WP4.
- D2.4_Set of 180 enzymes for experimental focus (July 2022; to be updated)
  In this deliverable, at least 180 enzymes from the priority sequences retrieved in the frame of WP2 (deliverables D2.2, D2.3) and WP3 (deliverable D3.3), were preliminary selected to proceed with their cloning, synthesis, expression and characterization.

Deliverables in the frame of WP3:

- D3.3_Set of 100 clones, 10 isolates, 10 enzymes shortlisted for sequencing (March 2022; to be updated)
  In this deliverable, bio-resources available before the beginning of the project and newly generated during the project were screened by naïve/functional methods to identify those with interest for our project. Bio-resources include previous and new enzymes, environmental samples, isolates, enrichments, and clone libraries that were checked for the purpose of the present project, and best selected ones sequenced and sequences with interest for our project were retrieved.

Deliverables in the frame of WP4:

- D4.2_The FuturEnzyme portfolio of 1000 enzyme (recombinant/native/biomimetic) material, obtained (September 2022)
  In this deliverable, the expression, preparation and production of set of protein samples of about 1000 enzymatic materials were undertaken by members of the consortium in a variety of hosts (heterologous or native) and vectors, cell-free systems, biomimetic "metamorphosis" systems, to cite some, as well as genetically-engineered mutants and supramolecular-engineered (immobilized) enzymes generated in the frame of WP5.
- D4.3_Cell-free expression reported system developed (September 2022)
  In this deliverable, a cell-free expression system was developed that allow the production and detection of enzymatic activities in a high-throughput manner by skipping the step of recombinant expression.
- D4.4_Biomimetic protease production system, developed (September 2022, re-opened to be updated in month 30)

In this deliverable a green chemical-system was designed that allows the production of enzymes with inherent problems of expression, particularly, biomimetic proteases.

- D4.6_The metadata on expression yield, activity and stability available (November 2022)
  This deliverable consists on the datasets informing about the expression yield, activity and stability of all enzymes generated in the project until month 18.

3. Description of nominated enzymes and isolates

Deliverable 4.6 compiled all the relevant information regarding all the 678 enzymatic materials from which datasets were generated until month 18. Based on the data reported in this deliverable, a number of candidates (22 enzyme candidates and two isolates until month 18) were selected as having characteristics of interest for the three applications relevant for the FuturEnzyme project: detergent, textiles and cosmetics. Below we detailed the information about the candidate enzymes nominated for engineering (WP5). Further information is available in **Table 1** (access provided in the Annex section). The priority candidates for WP6 and WP7 are highlighted in Section 4. For the sake of clarity and continuity, all the figures are located at the end of the document (Section 5).

Candidate number: 1

Name: Kest3
Activity: Lipase
Partner: Bangor
Source: Abano Terme, Italy (45.362 N 11.789 E; isolated strain *Fervidobacterium pennivorans DSM 9078;* genome accession number CP003260.1)
Expression system: *Escherichia coli* De3 Lobstr(p15TVL)
Expression level: Soluble (expression +)
Sequence (no signal peptide):
MDEKRSVKFFNKLTHIAKIILFTKGILILSSFFLAFSNFLLFLTVVVILNVPLLRKSIFGRLPTDTKELRKSNVLSNQETYEYLPGLFL
DVFYPSFFTESKKSQSVKGIVLFAHGGGWISGYRRQPNNLSWYRYLVSKGFIVATIDYERGYKAGIEKLIEELLQAVVFLENHL
SSKLGINEKVSLMGLSAGGHLALLAASRIPERVKNVVAYYSPCDLLDIWHSASIFARFAAATTLKRLPTRARDVYERYSPINNIT
ENYPPTLLVHGLKDSVVPYFSSVKMFKTLREKGLAAKLLLHPKGDHGFEFVLRDRRTVDIIEKTAQFLEGKLW
Synthesized sequence:
MDEKRSVKFFNKLTHIAKIILFTKGILILSSFFLAFSNFLLFLTVVVILNVPLLRKSIFGRLPTDTKELRKSNVLSNQETYEYLPGLFL
DVFYPSFFTESKKSQSVKGIVLFAHGGGWISGYRRQPNNLSWYRYLVSKGFIVATIDYERGYKAGIEKLIEELLQAVVFLENHL
SSKLGINEKVSLMGLSAGGHLALLAASRIPERVKNVVAYYSPCDLLDIWHSASIFARFAAATTLKRLPTRARDVYERYSPINNIT
ENYPPTLLVHGLKDSVVPYFSSVKMFKTLREKGLAAKLLLHPKGDHGFEFVLRDRRTVDIIEKTAQFLEGKLW
Biochemical features:

- Kest3 was selected based on activity with coconut oil.
- Topt 70˚C (no significant activity reduction at temperatures up to 85˚C) (**Figure 1**). It showed comparable activity to commercial lipase from *Candida rugosa* (L1754, Sigma) (**Figure 1**).
- In the stained swatches tests Kest3 showed activity with Butterfat on cotton (C-S-10), Fluid make-up on cotton (C-S-17), Pigment with oil on polyester/cotton (PC-09) (**Figure 2**), in the not buffered washing liquor.
- Though Kest3 showed a rate decrease in the presence of detergent, the Km decreased too leading to the catalytic efficiency increase 2.2 times (4.1 to 9.3 s^{-1}*mM^{-1}) (**Figure 1**).

Nominated for: as priority for additive for Henkel detergent; nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations).

Nagoya protocol compliance: the enzyme was retrieved from the genome of an strain isolated from a sample from Abano Terme, Italy (45.362 N 11.789 E). Sampling done before 24.03.1994, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database):

- Sequence 38143 from patent US 7834146, 33.87% (accession ADT16738.1)

Candidate number: 2

Name: FE_Lip9
Activity: Lipase
Partner: CSIC, BSC
Source: MarRef - Marine Metagenomics Database
Expression system: *Escherichia coli* (pET-45b(+))
Expression level: Soluble (expression +)
Sequence (signal peptide underlined):
MKVMFVKKRSLQILIALALVIGSMAFIQPKEVKAAEHNPVVMVHGIGGASYNFFSIKSYLATQGWDRNQLYAIDFIDKTGN
NRNNGPRLSRFVKDVLDKTGAKKVDIVAHSMGGANTLYYIKNLDGGDKIENVVTIGGANGLVSSRALPGTDPNQKILYTSV
YSSADLIVVNSLSRLIGARNVLIHGVGHIGLLTSSQVKGYIKEGLNGGGQNTN
Synthesized sequence once introduced in the vector pET-45b(+):
MAHHHHHHVGTGSNDDDDKSPDPMAEHNPVVMVHGIGGASYNFFSIKSYLATQGWDRNQLYAIDFIDKTGNNRNNGP
RLSRFVKDVLDKTGAKKVDIVAHSMGGANTLYYIKNLDGGDKIENVVTIGGANGLVSSRALPGTDPNQKILYTSVYSSADLIV
VNSLSRLIGARNVLIHGVGHIGLLTSSQVKGYIKEGLNGGGQNTN
Biochemical features:

- Most active at temperatures close to 25-40˚C, in agreement with its Td (41.70 ± 1.29°C)
- FE_Lip9 retains more than 80% of its maximal activity at pH from 7.0 to 10.0 (**Figure 3**).
- It also showed a half life time of 3.5 h in washing liquor.
- FE_Lip9 was selected based on activity with short to large triglycerides (**Figure 3**).
- FE_Lip9 is also capable to degrade all the stained fabrics tested (Pigment 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) (**Figure 4**).
- We also performed a free fatty acids assay with Schoeller's textiles where FE_Lip9 showed a broad capacity to clean the oils in all raw textiles (61488F1, 3X58, 67007, 61988F1, 5237-00 and E03130) (**Figure 5**).
- FE_Lip9 do show PETase and BHETase activities, being capable of degrading PET (**Figure 6**).
- The lipase has been successfully subjected to supramolecular engineering (FHNW) (**Figure 7**).
- The enzyme was successfully produced at 10 L scale: 3.96 grams of pure protein obtained.

Comments: this lipase is similar to one patented lipase (US 2009/0325240 A1 or EP 2260105 A2, US 5427936 A and maybe US 2007/0202566 A1). Currently we are using computational and HMM tools to identify sequences with lower level of homology to FE_Lip9. Also, BSC and CSIC are already selecting and producing a number of mutants to improve the activity towards long oils and the thermo-stability. Until now, two mutants have been designed (Val161Cys and Val161Ser), the synthesis of which is done and expression and activity level will be soon evaluated.
Nominated for: as priority for additive for Henkel detergent and for removal of spinning oils in Schoeller's fabrics; nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations).
Nagoya protocol compliance: the sequence was identified by homology screen in the MarRef - Marine Metagenomics Database, a manually curated marine microbial reference genome database that contains completely sequenced genomes (https://mmp2.sfb.uit.no/marref/). This protein is similar to one from a strain isolated from the marine sponge in the seawater in front of Seongsan-ri, Jeju Island, South Korea (33.38°N 126.53°E) (E-value of 8,56E-137 with WP_034624255.1_MMP06016472 MULTISPECIES: esterase [Bacillus] [mmp_id=MMP06016472] [mmp_db=marref]). Sample was collected 2011-11 (https://www.ebi.ac.uk/biosamples/samples/SAMN06016472), so that before the Nagoya Protocol entered into force on 12 October 2014.
Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >95% identity):

- Sequence 2 from patent US 5427936, 99.45% (accession AAA71046.1)

- Lipase [*Bacillus pumilus*], 99.45% (accession CAA02196.1)
- Sequence 1519 from patent US 10676751, 98.90% (accession QPT78109.1)
- Sequence 1537 from patent US 10676751 97.79% (accession QPT78118.1)
- Sequence 26 from patent US 8298799 96.69% (accession AFX20498.1)
- Sequence 73 from patent US 6858422 96.13% (accession AAY02152.1)
- Sequence 1517 from patent US 10676751 95.32% (accession QPT78108.1)
- Sequence 66 from patent US 6858422 95.03% (accession AAY02145.1)

Candidate number: 3

Name: FE_ID9

Activity: Lipase

Partner: CSIC

Source: Microbial assemblages from bone surface and the bone-eating worm *Osedax mucofloris* (BioProject ID PRJNA606180), Byfjorden, Bergen, Norway (60.397093N; 5.301293E) (collection date: from 01.2017 to 11.12.2017).

Expression system: *Escherichia coli* (pET-45b(+))

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MTNLSKPIPNPREYPILPPDMNYIYFENAHLFPFEPEKRDYSPVNAWWLSECAFLVYCHPGFARMAMALVGFDHFHFFQG
KGTECMVSWNKDSIIVAFRGTEMKSLSAFHELRTDLNTAPVDFDKGSKVHKGFLKGLQEIWEGEEGLKLFLETLSAEAPSRS
MWICGHSLGGALAALCFARLEKASGLYIYGAPRIGDGEFVRICDNRPVWRVEHGRDPIPLVPPDVPALNFNFKDMGKLIYID
YRGEILFERPLVTVEEEKSKVLLNISQQRKRRESLSVEGFKGVLDKDRAKTLINGINEHIMQSRVEWKEYFDSLDKGIGLKIKD
HMPIYYCAKLWNILIEGL

Synthesized sequence once introduced in the expression vector:

MAHHHHHHVGTGSNDDDDKSPDPMTNLSKPIPNPREYPILPPDMNYIYFENAHLFPFEPEKRDYSPVNAWWLSECAFLV
YCHPGFARMAMALVGFDHFHFFQGKGTECMVSWNKDSIIVAFRGTEMKSLSAFHELRTDLNTAPVDFDKGSKVHKGFLK
GLQEIWEGEEGLKLFLETLSAEAPSRSMWICGHSLGGALAALCFARLEKASGLYIYGAPRIGDGEFVRICDNRPVWRVEHGR
DPIPLVPPDVPALNFNFKDMGKLIYIDYRGEILFERPLVTVEEEKSKVLLNISQQRKRRESLSVEGFKGVLDKDRAKTLINGINE
HIMQSRVEWKEYFDSLDKGIGLKIKDHMPIYYCAKLWNILIEGL

Biochemical features:

- Remarkable high activity at 40˚C and pH 9.5 (**Figure 3**).
- Capacity to degrade all the stained fabrics tested (Pigment 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) (**Figure 4**).
- Stable in the presence of washing liquor.
- The enzyme was successfully produced at 10 L scale: 10.2 grams of pure protein obtained.

Nominated for: as priority for additive for Henkel detergent; nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations).

Nagoya protocol compliance: ID9 (orignal name EstLip_NODE_494_length_56501_cov_3.272419_27) was isolated from the microbial assemblages from bone surface and the bone-eating worm *Osedax mucofloris* (BioProject ID PRJNA606180), collected at Byfjorden, Bergen, Norway (60.397093N; 5.301293E). The samples were collected from January to December in Norway, in the frame of the EraNet project ProBone. Norway was among the first when ratifying the Nagoya protocol so that Nagoya protocol applies; all documentation (including pictures) that shows where the samples were taken can be made available upon request.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >95.5% identity):

- Sequence 7 from patent US 9527906, 100.00% (accession ATJ64490.1)
- Sequence 62 from patent US 9884101, 100.00% (accession AVY22639.1)

Candidate number: 4

Name: FE_Polur1

Activity: Lipase

Partner: CSIC, BSC

Source: NCBI's database (WP_127927389.1 polyurethanase [Pseudomonas sp. RU47])

Expression system: *Escherichia coli* (pET-45b(+))

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MGVYDYKNFGTADSKALFSDAMAITLYSYHNLDNGFAAGYQHNGFGLGLPATLVTALLGGTDSQGVIPGIPWNPDSEKLA
LDAVKQAGWTPITASQLGYDGKTDARGTFFGEKAGYTTAQVEILGKYDAQGHLTELGIAFRGTSGPRENLILDSIGDVINDLL
AAFGPKDYAKNYVGEAFGNLLNDVVAFAKANGLSGKDVLVSGHSLGGLAVNSMADLSGGKWGGFFADSNYIAYASPTQS
STDKVLNVGYENDPVFRALDGSTFTGASVGVHDAPKESATDNIVSFNDHYASTAWNLLPYSILNIPTWISHLPTAYGDGMN
RVIDSKFYDLTSRDSTIIVANLSDPARANTWVQDLNRNAETHKGSTFIIGSDANDLIQGGSGNDYLEGRAGNDTFRDSGGY
NVILGGQGSNTLDLQSAVKNFDFANDGAGNLYVRDANGGISITRDIGSIVTKEPGFLWGLFKDDVTHSVTASGLKVGNNVT
AYESSVKGSAGADTLKAHSGGDWLFGLDGNDHLIGGAGNDVFVGGAGNDLMESGGGADTFLFNGAFGQDRVVGYTSN
DKLVFLGVQGVLPAEDFRAHAATVGQDTVLTFGNDSVTLVGVSLNSLSAEGVVIA

Synthesized sequence once introduced in the expression vector:

MAHHHHHHVGTGSNDDDDKSPDPMGVYDYKNFGTADSKALFSDAMAITLYSYHNLDNGFAAGYQHNGFGLGLPATLV
TALLGGTDSQGVIPGIPWNPDSEKLALDAVKQAGWTPITASQLGYDGKTDARGTFFGEKAGYTTAQVEILGKYDAQGHLT
ELGIAFRGTSGPRENLILDSIGDVINDLLAAFGPKDYAKNYVGEAFGNLLNDVVAFAKANGLSGKDVLVSGHSLGGLAVNSM
ADLSGGKWGGFFADSNYIAYASPTQSSTDKVLNVGYENDPVFRALDGSTFTGASVGVHDAPKESATDNIVSFNDHYASTA
WNLLPYSILNIPTWISHLPTAYGDGMNRVIDSKFYDLTSRDSTIIVANLSDPARANTWVQDLNRNAETHKGSTFIIGSDAND
LIQGGSGNDYLEGRAGNDTFRDSGGYNVILGGQGSNTLDLQSAVKNFDFANDGAGNLYVRDANGGISITRDIGSIVTKEPG
FLWGLFKDDVTHSVTASGLKVGNNVTAYESSVKGSAGADTLKAHSGGDWLFGLDGNDHLIGGAGNDVFVGGAGNDLM
ESGGGADTFLFNGAFGQDRVVGYTSNDKLVFLGVQGVLPAEDFRAHAATVGQDTVLTFGNDSVTLVGVSLNSLSAEGVVI
A

Biochemical features:

- Remarkable high activity at 50˚C and pH 9.5 (**Figure 3**).
- Stable in the presence of washing liquor.
- Capacity to degrade stained fabrics, particularly, Beef fat on cotton (C-S-61) and Butterfat on cotton (C-S-10)) (**Figure 4**).
- The lipase has been successfully subjected to supramolecular engineering (FHWM) (**Figure 7**).

Comments: this lipase is similar to one patented lipase (US 10676751), but not for detergent applications.

Nominated for: as priority for additive for Henkel detergent; nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations).

Nagoya protocol compliance: the sequence was identified by homology screen in the NCBI's database; it corresponds to the uncharacterized protein WP_127927389.1 from the genome of *Pseudomonas* sp. RU47. The genome of this strain, sequenced at the JGI, was released 2021-01-04. The strain was isolated from soil sample at Rothamsted, UK (Latitude, longitude: 51.8138, -0.3783) in 2007 (BioSample: SAMN07173783), before the Nagoya Protocol entered into force on 12 October 2014.

Link to BioSample SAMN07173783: https://www.ncbi.nlm.nih.gov/biosample/SAMN07173783

Link to JGI: https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=2913023320

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >94% identity):

- Sequence 1576 from patent US 10676751, 97.30% (accession QPT78139.1)
- Sequence 1572 from patent US 10676751, 94.98% (accession QPT78137.1)
- Sequence 1457 from patent US 10676751, 94.81% (accession QPT78076.1)
- Sequence 1630 from patent US 10676751, 94.65% (accession QPT78166.1)
- Sequence 1472 from patent US 10676751, 94.33% (accession QPT78084.1)
- Sequence 106 from patent US 8298799, 94.17% (accession AFX20538.1)

Candidate number: 5

Name: EstLip_Dim_#008

Activity: Lipase

Partner: UDUS

Source: from a metagenomic library of a dimenthyl succinate enrichment culture inoculated with material from a wastewater treatment plant in Germany. An identical enzyme was isolated by others from *Aneurinibacillus thermoaerophilu*s and described as AZ -lipase, which features moderate solvent tolerance and high thermal stability, but was not further explored.

Expression system: *Escherichia coli* (pET22b)

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MQKERKNQYPIVLVHGFAGWGRDEMLGVKYWGGMHDIQEDLKQYGYETHTAVVGPFSSNWDRACELYAQLVGGTVD
YGAAHAEKYGHDRFGRTYPGLLKNWDGEHKIHLIGHSMGGQTVRVLTQLLKEGSQEEREYAKKHGVQLSPLFEGGKSWV
HSVTTIATPNDGTTLADVVTQLIPAAQQIMGLAAAVSGNTNVPVYDFKLDQWGLKRKAGESFVHYADRVWNSGIWTNTK
DISAWDLKPEGAKELNNWVKAQPDVYYFSYSGEATFRSLITGHHLPDLTMNKLITPFGIFLGCYGSDEKWWQNDGIVNTIS
MNGPKLGSTDEIVPYDGTPKIGKWNDMGIQENWDHADYIGLSLSYVLGIEKIEDFYRGVADMLGSLSVR

Synthesized sequence once introduced in the expression vector:

MQKERKNQYPIVLVHGFAGWGRDEMLGVKYWGGMHDIQEDLKQYGYETHTAVVGPFSSNWDRACELYAQLVGGTVD
YGAAHAEKYGHDRFGRTYPGLLKNWDGEHKIHLIGHSMGGQTVRVLTQLLKEGSQEEREYAKKHGVQLSPLFEGGKSWV
HSVTTIATPNDGTTLADVVTQLIPAAQQIMGLAAAVSGNTNVPVYDFKLDQWGLKRKAGESFVHYADRVWNSGIWTNTK
DISAWDLKPEGAKELNNWVKAQPDVYYFSYSGEATFRSLITGHHLPDLTMNKLITPFGIFLGCYGSDEKWWQNDGIVNTIS
MNGPKLGSTDEIVPYDGTPKIGKWNDMGIQENWDHADYIGLSLSYVLGIEKIEDFYRGVADMLGSLSVR

Comments:

- It has lipase activity and thermostability and excellent expression level.
- Melting points beyond 60°C (**Figure 8**).
- High specific activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (**Figure 9**).
- It features moderate activity in washing liquor, which may constitute a leverage point for protein engineering.

Nominated for: as additive for Henkel detergent; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), although it may be not priority as it is not the most active lipase towards fatty standard stains.

Nagoya protocol compliance: The sequence was attributed to *Aneurinibacilus thermoaerophilus* that was isolated before 2009 (https://doi.org/10.1016/j.procbio.2012.11.002) because it corresponds to protein ADC84241.1. An identical gene was recovered in 2015 from a sample taken at site of Bayer AG in Germany and provided by Bayer in the frame of the H2020 project INMARE.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >61% identity):

- Sequence 18 from patent US 9303264, 100.00% (accession APN37501.1)

Candidate number: 6

Name: EstLip_Paes_TB035

Activity: Lipase

Partner: UDUS

Source: *Halopseudomonas aestusnigri*

Expression system: *Escherichia coli* (pET22b)

Expression level: Soluble (expression +)

Sequence (signal peptide underlined):

MSTARLTPTTASTLKLLSVAVAALTLSACLSGGGGGRADGPNPLAIETSEGKVVGISNDGIRVFRGIPYAAPPVGDLRLAPPQ
PPASRSETLRLSEEFGNSCPQSDLTTGQQVGNEDCLYLNVYAPAEAEDLPVMVWIHGGAFVFGNGGGEYDPTRLVEQDVI
VVTLNYRLGNLGFLAHPALESDAGNFALMDQQLALAWVKENIAAFGGDPANVTIFGESAGGHSVMSHIVSPRAEEADLFQ

RAIVQSGSYAPFQMPKATAQFLGTSVANGLGCTDPETAASCLRSLPVSAFLAAQGSQSIPVVDPDDDLLPKSIQQALADGD
FNSSLDIMIGSNQNEGTLFVALDEVGGDPIDDEAEYRERVAEFFQPYQASIPFDDDQIATDYLDFVDGAAKPFEAALSGIWT
DFMFACNAYSQASTFAGASMNTFQYWFRDEDAPWTLVPPFAVSFPLGATHAGEIPYVLYPQAIMEQRYTGDPDDLNSLA
GEMVDYWTQFAKTGDPNTTDGVAAAWQQAATGNLLTLDVPNASNANTLGFLGYHHCSYWADPPLVLP
Synthesized sequence once introduced in the expression vector:
CLSGGGGGRADGPNPLAIETSEGKVVGISNDGIRVFRGIPYAAPPVGDLRLAPPQPPASRSETLRLSEEFGNSCPQSDLTTGQ
QVGNEDCLYLNVYAPAEAEDLPVMVWIHGGAFVFGNGGGEYDPTRLVEQDVIVVTLNYRLGNLGFLAHPALESDAGNFA
LMDQQLALAWVKENIAAFGGDPANVTIFGESAGGHSVMSHIVSPRAEEADLFQRAIVQSGSYAPFQMPKATAQFLGTSV
ANGLGCTDPETAASCLRSLPVSAFLAAQGSQSIPVVDPDDDLLPKSIQQALADGDFNSSLDIMIGSNQNEGTLFVALDEVGG
DPIDDEAEYRERVAEFFQPYQASIPFDDDQIATDYLDFVDGAAKPFEAALSGIWTDFMFACNAYSQASTFAGASMNTFQY
WFRDEDAPWTLVPPFAVSFPLGATHAGEIPYVLYPQAIMEQRYTGDPDDLNSLAGEMVDYWTQFAKTGDPNTTDGVAA
AWQQAATGNLLTLDVPNASNANTLGFLGYHHCSYWADPPLVLP
Comments:

- This enzyme originates from *Halopseudomonas aestusnigri*. It is membrane-bound in its original form, but soluble expression with good yields is possible with a variant lacking the lipoprotein signal peptide. It was suggested for further exploration in the first place because of its remarkable resistance towards denaturing agents, including surfactants (**Figure 3**, and 10.1128/AEM.00106-20).
- For an increased expression, the original Sec-TypII Signalpeptide for a lipoprotein was replaced with the *E. coli* PelB -Signal Peptide.
- It shows a rather high structural melting point (melting point beyond 60°C; **Figure 8**), 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.
- Availability of experimental structural data. The yet unpublished structure reveals an uncommon topology with two cap domains that may be utilized for PluriZyme construction.
- Full or high percentage residual activity in the presence of surfactants like in washing liquor (**Figure 10**).

Nominated for: as additive for Henkel detergent.

Nagoya protocol compliance: This is an artificial genetic fusion. The original gene was retrieved from *Pseudomonas aestusnigri* (now *Halopseudomonas aestusnigri*) that was isolated before 2005 in Carnota, Galicia, Spain, so that before the Nagoya Protocol entered into force on 12 October 2014. Strain and sampling information were obtained from DSMZ.

Patent match information (BLASTP against the Patented Protein Sequences Database):

- Sequence 648 from patent US 8298799, 41.98% (accession AFX20809.1)

Candidate number: 7

Name: EstLip_PtEst1

Activity: Lipase

Partner: UDUS

Source: from the metagenomic library of a cyclohexanol succinate enrichment culture inoculated with material from a wastewater treatment plant in Germany. The enzyme is identical to a database protein sequence of *Pseudonocardia thermophila.*

Expression system: *Escherichia coli* (pET22b)

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MTTTSNSTETGRRPGRLGDPDRCLRTDPRTDPRTVEALAPFGLDVNAAPAPIGPDAPREQQLEYAMGAEAAFEGVFAAL
MDGLDPVPGIERRTETISGPAGNEIKLYVHRPAGAVGPLPGIFHIHGGGMVILQAAGPVYVRFRDELAATGTVVVGVEYRN
GAGVLGPHPFPAGLHDCAVALDWVHARRAELGISTLTVAGESGGGNLTLATAIRAKREGRLDAIDGVYALVPYISGMYGRS
REEREAELPSLVECDGYFISCDLCAVFVEVYDPGTAHLTDPLAWPYHAAREDLVGLPPHVISVNEVDPLRDEGLAYYRKLVEA
GVEARSRVVPGACHAADMMFRKAAPDMYEATVQDIHDFVTSLHR

Synthesized sequence once introduced in the expression vector:

MTTTSNSTETGRRPGRLGDPDRCLRTDPRTDPRTVEALAPFGLDVNAAPAPIGPDAPREQQLEYAMGAEAAFEGVFAAL
MDGLDPVPGIERRTETISGPAGNEIKLYVHRPAGAVGPLPGIFHIHGGGMVILQAAGPVYVRFRDELAATGTVVVGVEYRN
GAGVLGPHPFPAGLHDCAVALDWVHARRAELGISTLTVAGESGGGNLTLATAIRAKREGRLDAIDGVYALVPYISGMYGRS
REEREAELPSLVECDGYFISCDLCAVFVEVYDPGTAHLTDPLAWPYHAAREDLVGLPPHVISVNEVDPLRDEGLAYYRKLVEA
GVEARSRVVPGACHAADMMFRKAAPDMYEATVQDIHDFVTSLHR

Comments:

- Rather high activity in washing liquor, high substrate promiscuity (60 out of 9 esters, see 10.1111/febs.15680) and moderate thermostability.
- Availability of experimental structural data.
- Full or high percentage residual activity in the presence of surfactants like in washing liquor (**Figure 10**).

Nominated for: as additive for Henkel detergent; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), although it may be not priority as it is not the most active lipase towards fatty standard stains.

Nagoya protocol compliance: The gene was recovered in 2015 from a sample taken at the site of Bayer AG in Germany and provided by Bayer in the frame of the H2020 project INMARE. It matches protein WP_073459097.1 that was isolated from a *Pseudonacardia thermophila* isolated before 1993.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >50% identity):

- Sequence 394 from patent US 8298799, 58.63% (accession AFX20682.1)
- Sequence 4 from patent US 9193961, 56.44% (accession AOE00123.1)

Candidate number: 8

Name: EstLip_TBEc304

Activity: Lipase

Partner: UDUS

Source: *Metagenome Slaughterhouse drain, Düren; Germany* (10.1038/srep27035)

Expression system: *Escherichia coli* (pET22b)

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MTDTPTPRPDVAAFLGFLNSLEGPKMHQVTPDVSRQMMLAMGPIAELPAGEVAVVRDLIIPGPAGGIPARLYDNRESRG
AGPVMVFFHGGGFVIGDLEVYHPYCAEVARQMDMPVISVDYRLAPEHPFPAASDDCEAATRWVAESPAELGLDVTGLVL
SGDSAGGNLTIVTAMALRDRPAKVPVILQHPIYPAVSSSNDWQSMRDFAEGHLLTAEGMAWFYDHYKPDNSDYRGSPLD
FDQTGMPPSLVVTAGLDPLRDQGAAYAEKLKASGVPTEYRNSEGTIHGYINLRKAIPSAQEDVTANLTTLRAMLDRILDAQ

Synthesized sequence once introduced in the expression vector:

MTDTPTPRPDVAAFLGFLNSLEGPKMHQVTPDVSRQMMLAMGPIAELPAGEVAVVRDLIIPGPAGGIPARLYDNRESRG
AGPVMVFFHGGGFVIGDLEVYHPYCAEVARQMDMPVISVDYRLAPEHPFPAASDDCEAATRWVAESPAELGLDVTGLVL
SGDSAGGNLTIVTAMALRDRPAKVPVILQHPIYPAVSSSNDWQSMRDFAEGHLLTAEGMAWFYDHYKPDNSDYRGSPLD
FDQTGMPPSLVVTAGLDPLRDQGAAYAEKLKASGVPTEYRNSEGTIHGYINLRKAIPSAQEDVTANLTTLRAMLDRILDAQ

Comments:

- Excellent expression yields.
- High activity at elevated temperatures (melting point beyond 60°C; **Figure 8**) and in washing liquor.
- High activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (**Figure 9**).
- Initial screenings indicate a high substrate promiscuity including the PET monomer BHET.

Nominated for: as additive for Henkel detergent; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), being the second most active from the set of lipases from UDUS partner.

Nagoya protocol compliance: The gene was retrieved from a sample that was isolated in slaughterhouse drain in Düren, Germany (50.80669915437404, 6.471127329492571) (in 2008, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >60% identity):

- Sequence 6 from patent US 8956838, 62.54% (accession AKW88674.1)
- Sequence 796 from patent US 8298799, 60.71% (accession AFX20883.1)

Candidate number: 9

Name: PEH_Paes_PE-H_Y250S
Activity: Lipase, PETase
Partner: UDUS
Source: Synthetic mutant, *Halopseudomonas aestusnigri*
Expression system: *Escherichia coli* (pET22b)
Expression level: Soluble (expression +)
Sequence (signal peptide underlined):
MPFNKKSVLALCGAGALLFSMSALANNPAPTDPGDSGGGSAYQRGPDPSVSFLEADRGQYSVRSSRVSSLVSGFGGGTIYY
PTGTTGTMGAVVVIPGFVSAESSIDWWGPKLASYGFVVMTIDTNTGFDQPPSRARQINNALDYLVSQNSRSSSPVRGMID
TNRLGVIGWSMGGGGTLRVASEGRIKAAIPLAPWDTTSYYASRSQAPTLIFACESDVIAPVLQHASPFYNSLPSSIDKAFVEI
NGGSHSCGNGGSIYNDVLSRFGVSWMKLHLDEDSRYKQFLCGPNHTSDSQISDYRGNCPYLE
Synthesized sequence once introduced in the expression vector:
NNPAPTDPGDSGGGSAYQRGPDPSVSFLEADRGQYSVRSSRVSSLVSGFGGGTIYYPTGTTGTMGAVVVIPGFVSAESSID
WWGPKLASYGFVVMTIDTNTGFDQPPSRARQINNALDYLVSQNSRSSSPVRGMIDTNRLGVIGWSMGGGGTLRVASEG
RIKAAIPLAPWDTTSYYASRSQAPTLIFACESDVIAPVLQHASPFYNSLPSSIDKAFVEINGGSHSCGNGGSIYNDVLSRFGVS
WMKLHLDEDSRYKQFLCGPNHTSDSQISDYRGNCPYLE
Comments:

- High specific activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (**Figure 9**).
- Full or high percentage residual activity in the presence of surfactants, i.e., washing liquor (**Figure 7**).
- Availability of experimental structural data.
- Activity on PET fabric.
- Activity in washing liquor.
- Activity on PET fabrics (**Figure 11**).
- Structural features of cutinases/PETase.

Nominated for: as additive for Henkel detergent and for end-of-life fabric recycling; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), although it may be not priority as it is not the most active lipase towards fatty standard stains, but it is the best candidate for end-of-life fabric recycling.

Nagoya protocol compliance: The original gene was retrieved from *Pseudomonas aestusnigri* (now *Halopseudomonas aestusnigri*) that was isolated before 2005 in Carnota, Galicia, Spain, so that before the Nagoya Protocol entered into force on 12 October 2014. Strain and sampling information were obtained from DSMZ.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >60% identity):

- Sequence 16 from patent US 9303264, 62.88% (accession APN37499.1)

Candidate number: 10

Name: PEH_Pbau_PE-H
Activity: Lipase, PETase
Partner: UDUS
Source: *Halopseudomonas bauzanensis*
Expression system: *Escherichia coli* (pET22b)
Expression level: Soluble (expression +)
Sequence (signal peptide underlined):

MINKNLSQSLLAMMAAGALLLSSSAFAVNPPTDGPTDPDQAYERGPDPSVAFLEAPTGPHSVRTSRVSGLVSGFGGGTIHY
PTGTTGTMAAIVVIPGFVSAESSIEWWGPKLASHGFVVMTIDTNTGFDQPPSRARQINNALDYLVSQNTSRTSPVNGMID
TERLGVIGWSMGGGGTLRVASEGRIKAAIPLAPWDTTRFRGVQAPTLIFACESDLIAPVRSHASPFYNQLPDDIDKAYVEIN
NGSHYCANGGGLNNDVLSRFGVSWMKRFLDNDTRYSQFLCGPNHESDRNISEYRGNCPY
Synthesized sequence once introduced in the expression vector:
VNPPTDGPTDPDQAYERGPDPSVAFLEAPTGPHSVRTSRVSGLVSGFGGGTIHYPTGTTGTMAAIVVIPGFVSAESSIEWW
GPKLASHGFVVMTIDTNTGFDQPPSRARQINNALDYLVSQNTSRTSPVNGMIDTERLGVIGWSMGGGGTLRVASEGRIKA
AIPLAPWDTTRFRGVQAPTLIFACESDLIAPVRSHASPFYNQLPDDIDKAYVEINNGSHYCANGGGLNNDVLSRFGVSWMK
RFLDNDTRYSQFLCGPNHESDRNISEYRGNCPY
Comments:

- High specific activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (**Figure 9**).
- Activity in washing liquor.
- Activity on PET fabrics (**Figure 11**).
- Structural features of cutinases/PETase.

Nominated for: as additive for Henkel detergent and for end-of-life fabric recycling; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), being the most active towards fatty standard stains from the UDUS lipase collection.

Nagoya protocol compliance: The gene was retrieved from *Pseudomonas bauzanensis* (now *Halopseudomonas bauzanensis*) that was isolated in March 2008 in Bozen, South Tyrol, Italy, so that before the Nagoya Protocol entered into force on 12 October 2014. Strain and sampling information were obtained from DSMZ.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >60% identity):

- Sequence 16 from patent US 9303264, 61.70% (accession APN37499.1).

Candidate number: 11

Name: PEH_Pform_PE-H
Activity: Lipase, PETase
Partner: UDUS
Source: *Halopseudomonas* sp., Compost, Germany
Expression system: *Escherichia coli* (pET22b)
Expression level: Soluble (expression +)
Sequence (signal peptide underlined):
MKAPAIKLTGIAMGMLLSASALAVNPGNPDPGTPPGTGFPAVSDFAASGPFSVTNSNVNTTCAVYRPRTLGENGLKHPIIL
WGNGTGGSPSTYGRLLEHWASHGFVVAAARTSNAGTGEEMISCLNYLITENGRSTGTFAGKLDVNRVATAGHSQGGGGS
IMAGQDSRVTVSAPFQPYTLGLGHDRSSQRNQSGPMFLMTGGSDTIASPTLNALPVYNNANVPVFWGELRGAGHFEPV
GNAGGYRGPSTAWLRYHLMEDQNAESTFYGSNCGLCSDRNWDVRRKGINLE
Synthesized sequence once introduced in the expression vector:
VNPGNPDPGTPPGTGFPAVSDFAASGPFSVTNSNVNTTCAVYRPRTLGENGLKHPIILWGNGTGGSPSTYGRLLEHWASH
GFVVAAARTSNAGTGEEMISCLNYLITENGRSTGTFAGKLDVNRVATAGHSQGGGGSIMAGQDSRVTVSAPFQPYTLGLG
HDRSSQRNQSGPMFLMTGGSDTIASPTLNALPVYNNANVPVFWGELRGAGHFEPVGNAGGYRGPSTAWLRYHLMEDQ
NAESTFYGSNCGLCSDRNWDVRRKGINLE
Comments:

- Activity in washing liquor.
- Activity on PET fabrics (**Figure 11**).
- Structural features of cutinases/PETase.

Nominated for: end-of-life fabric recycling.

Nagoya protocol compliance: The gene used here for expression was a synthetic construct encoding a protein that was retrieved from the yet unpublished genome sequence of *Pseudomonas* sp. (now *Halopseudomonas* sp.) that was isolated in March 2020 in Eschweiler, Germany, obtained. The sequence data was provided by Prof.- Nick Wierckx, IBG-1, Forschungszentrum Jülich.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >65% identity):

- Sequence 10 from patent US 8298799, 69.08% (accession AFX20490.1)
- Sequence 12 from patent US 8298799, 67.06% (accession AFX20491.1)
- Sequence 76 from patent US 8298799, 66.02% (accession AFX20523.1)

Candidate number: 12

Name: PEH_Poce_PE-H
Activity: Lipase, PETase
Partner: UDUS
Source: *Halopseudomonas oceani*
Expression system: *Escherichia coli* (pET22b)
Expression level: Soluble (expression +)
Sequence (signal peptide underlined):
MYPTTKTTHDSWRLHLMPFSKKGILAICGAGALLFSMSALANNPPPTDPGDGGGSSYQRGPDPSVSFLEADRGQYSVRSS
RVSGLVSGFGGGTIYYPTGTTGTMAAVVVIPGFVSAESSIDWWGPKLASYGFVVMTIDTNTGFDQPPSRARQINNALDYLI
SQNSRSTSPVRGMIDTNRLGVVGWSMGGGGTLRVASEGRIKAAIPLAPWDTSSYYASRAQAPTLIFACQADVIAPVFQHA
SPFYNSLPSNIDKAFVEINGGSHFCANGGSIYNDVLGRLGVSWMKLHLDEDNRYKQFLCGPNHTSDFQISGYRGNCPY
Synthesized sequence once introduced in the expression vector:
NNPPPTDPGDGGGSSYQRGPDPSVSFLEADRGQYSVRSSRVSGLVSGFGGGTIYYPTGTTGTMAAVVVIPGFVSAESSIDW
WGPKLASYGFVVMTIDTNTGFDQPPSRARQINNALDYLISQNSRSTSPVRGMIDTNRLGVVGWSMGGGGTLRVASEGRI
KAAIPLAPWDTSSYYASRAQAPTLIFACQADVIAPVFQHASPFYNSLPSNIDKAFVEINGGSHFCANGGSIYNDVLGRLGVS
WMKLHLDEDNRYKQFLCGPNHTSDFQISGYRGNCPY
Comments:

- High specific activity on fatty standard stains, particularly Lipstick, pink on polyester/cotton (P-S-16), Beef fat on cotton (C-S-61) (**Figure 9**).
- Activity in washing liquor.
- Activity on PET fabrics (**Figure 11**).
- Structural features of cutinases/PETase.

Nominated for: as additive for Henkel detergent and for end-of-life fabric recycling; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), although it may be not priority as is not the best active lipase towards fatty standard stains.

Nagoya protocol compliance: The gene was retrieved from *Pseudomonas oceani* (now *Halopseudomonas oceani*) that was isolated on 17.04.2014 from international waters (Okinawa Trough), so that before the Nagoya Protocol entered into force on 12 October 2014. Strain and sampling information were obtained from DSMZ.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >60% identity):

- Sequence 16 from patent US 9303264, 62.50% (accession APN37499.1)

Comparative properties EstLip_Dim_#008, EstLip_Paes_TB035, EstLip_PtEst1, EstLip_TBEc304, PEH_Paes_PE-H_Y250S, PEH_Pbau_PE-H, PEH_Pform_PE-H, PEH_Poce_PE-H (candidates 5-12):

- Comparably high specific activity on fatty standard stains (PEH_Pbau_PE-H, EstLip_Dim_#008, PEH_Paes_PE-H_Y250S, EstLip_TBEc304, PEH_Poce_PE-H) (**Figure 5**).
- Melting points beyond 60°C (EstLip_Dim_#008, EstLip_Paes_TB035, EstLip_TBEc304) (**Figure 6**),
- Full or high percentage of residual activity in the presence of surfactants like in washing liquor (EstLip_Paes_TB035, EstLip_PtEst1, PEH_Paes_PE-H_Y250S) (**Figure 7**).
- Activity on PET fabric PEH_Pbau_PE-H, PEH_Poce_PE-H, PEH_Paes_PE-H_Y250S, PEH_Pform_PE-H.
- Availability of experimental structural data (EstLip_PtEst1, EstLip_Paes_TB035, PEH_Paes_PE-H_Y250S).
- PEH_Pbau_PE-H, PEH_Poce_PE-H, PEH_Paes_PE-H_Y250S, PEH_Pform_PE-H constitute a set of homologous enzymes from *Halopseudomonas* sp. They show the structural features of

cutinases/PETase. Consequently, they were suggested for further exploration because of their activity towards PET fabrics (**Figure 4**) and activity in washing liquor and remarkable activities in the experiments with standardized stains. Although sharing 75-90% sequence identities, they show different properties and constitute, therefore a small naturally evolved mutagenesis library. Notably, PEH_Paes_PE-H_Y250S is a variant of PEH_Paes_PE-H_Y250S with improved features, e.g. a massively increased activity towards PET. The crystal structures of this variant and the wild type are available (10.3389/fmicb.2020.00114). We plan to transfer this beneficial mutation to the other three proteins and compare the improved enzyme regarding PET hydrolysis, fatty acid release from fatty stains and stability. PEH_Pform_PE-H from a problable *Halopseudomonas formosensis* is a rather recent discovery; hence, some characteristics are still unknown. Nonetheless, this enzyme outperformed the homologous (wild type) enzymes in most so far investigated activities; thus we think it worth considering proceeding with these enzymes.

Candidate number: 13

Name: GEN0105

Activity: Lipase, PETase

Partner: Bangor

Source: metagenome of a mesophilic anaerobic digester, Evry, France (48.626 N 2.464 E)

Expression system: *Escherichia coli* De3 Lobstr(p15TVL)

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MSETSSASALPAYARIVVDKRAPFIRAILYLILRYVIKRSMKPDADILKLRAMQLRADQKYAHPAADAVMTPVDCDGVKAN
WITLPGARPERVIFYLHGGAWMFNFPRTYAAMLGRWARLLNARVLMVDYRLAPEHRYPAGANDCETAYRWLLAQGIDS
KQIVIGGDSAGGNLTLTTLLRLKSANQPLPACAVALSPFVDFTLSSPSMITNEKIDPMFTLEAMLGLRPHYLDPQDFLNVDAS
PIFGDFSGLPPIFFQSSNTEMLRDESVRAAARAHQHGVTVELELWQHLPHVFQALQKLPQADAALQSIVRFINSHTGWQA

Synthesized sequence once introduced in the expression vector:

MSETSSASALPAYARIVVDKRAPFIRAILYLILRYVIKRSMKPDADILKLRAMQLRADQKYAHPAADAVMTPVDCDGVKAN
WITLPGARPERVIFYLHGGAWMFNFPRTYAAMLGRWARLLNARVLMVDYRLAPEHRYPAGANDCETAYRWLLAQGIDS
KQIVIGGDSAGGNLTLTTLLRLKSANQPLPACAVALSPFVDFTLSSPSMITNEKIDPMFTLEAMLGLRPHYLDPQDFLNVDAS
PIFGDFSGLPPIFFQSSNTEMLRDESVRAAARAHQHGVTVELELWQHLPHVFQALQKLPQADAALQSIVRFINSHTGWQA

Comments:

- GEN0105 was selected based on the outstanding activity with ester bond plastic (3PET, PLA, PCL) (**Figure 12**). The enzyme was not found to be thermostable (data not shown).
- It was tested along with Kest3 on fatty standard stains (lipase activity present when screened on coconut oil). GEN0105 shows activity with Lipstick, pink on polyester/cotton (P-S-16), Butterfat on cotton (C-S-10), Fluid make-up on cotton (C-S-17), Pigment with oil on polyester/cotton (PC-09).
- Activities were comparable with Henkel enzyme mix, in the buffered washing liquor, though loosing activity significant in the not buffered conditions (**Figure 2**).
- Good candidate for increasing stability.

Nominated for: as additive for Henkel detergent; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), being the best active lipase towards fatty standard stains from the Bangor lipase collection.

Nagoya protocol compliance: GEN0105 was isolated from the metagenome of a mesophilic anaerobic digester, Evry, France (48.626 N 2.464 E). The sample was collected on 18.03.2002, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >60% identity):

- Sequence 918 from patent US 8298799, 61.69% (accession AFX20944.1)
- Sequence 1422 from patent US 10676751, 60.88% (accession QPT78057.1)

Candidate number: 14

Name: GEN0095

Activity: Cellulase

Partner: Bangor

Source: metagenome of a mesophilic anaerobic digester, Evry, France (48.626 N 2.464 E).

Expression system: *Escherichia coli* De3 Lobstr(p15TVL)

Expression level: Soluble (expression +)

Sequence (no signal peptide):

MNTIINDNISSFEILKQMKIGWNLGNTFDSNADWIESGNPVDYETSWGNPYTTQELISKINSAGFKTIRIPITWTNHFDLKTG
KINTEWIKRIKEVVDFAISINMYVIINIHHEKWHNPIYVNLDNASSILNKLWTQISLFFQSYDEHLLFEGLNEPRIVGSEHEWS
GGNLETHDVINKLNSTFIKTVRATGGNNSFRHLIVPTHAASATQNAIHGFEIPNDNKIIVSIHSYSPYNFALNPNGTTSWNAC
DPIDTFPIDDFINRLYENFISEGVPVIIGEFGAVNKNNLEERVSWAEYYVSRATEKNIVCIWWDNGLFDGDGENFGLIDRNNL
TWVYPEIVSSMMKSLKPKASINL

Synthesized sequence once introduced in the expression vector:

MNTIINDNISSFEILKQMKIGWNLGNTFDSNADWIESGNPVDYETSWGNPYTTQELISKINSAGFKTIRIPITWTNHFDLKTG
KINTEWIKRIKEVVDFAISINMYVIINIHHEKWHNPIYVNLDNASSILNKLWTQISLFFQSYDEHLLFEGLNEPRIVGSEHEWS
GGNLETHDVINKLNSTFIKTVRATGGNNSFRHLIVPTHAASATQNAIHGFEIPNDNKIIVSIHSYSPYNFALNPNGTTSWNAC
DPIDTFPIDDFINRLYENFISEGVPVIIGEFGAVNKNNLEERVSWAEYYVSRATEKNIVCIWWDNGLFDGDGENFGLIDRNNL
TWVYPEIVSSMMKSLKPKASINL

Comments:

- GEN0095 was selected based on activity with CMC and remarkable Topt (screen at 30, 50 and 80˚C) and neutral pH optimum (**Figure 13**).
- Enzyme showed comparable activity with commercially available cellulase from *Trichoderma viride* (C1794, Sigma) (**Figure 13**).

Nominated for: as additive for textile applications; possibly not priority target for WP6 (large scale production) and WP7 (pre-industrial validations) because cellulases are not priority enzymes.

Nagoya protocol compliance: GEN0095 was isolated from the metagenome of a mesophilic anaerobic digester, Evry, France (48.626 N 2.464 E). The sample was collected on 18.03.2002, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >52% identity):

- Sequence 3858 from patent US 10717768, 52.50% (accession QPV28741.1)
- Sequence 3851 from patent US 10717768, 52.36% (accession QPV28734.1)

Candidate number: 15

Name: VD_PL9

Activity: Extracellular exopolygalacturonate lyase (PL9) - Hyaluronidase

Partner: CNR, CSIC, BSC

Source: *Vibrio diabolicus* V4. Isolation source: shallow thermal white smokers (Fe^{2+} rich) of Basiluzzo Island (Chimney_Sagrada Familia, 38°39’25”N, 15°6’0”E). Isolation date 10 September 2021.

Expression system: *Escherichia coli* (pET-45b(+))

Expression level: Soluble (expression +)

Sequence (signal peptide underlined):

MKKHTLALCLAAILAPVAHAAEIKVEDLTWKAITFGQSTDMNFGSTILPEKVGVNQVTVNGEAVAAGKLASTFTIESRGGKL
ANSHEGLTFYYTELPTDVNFTLSADVVLEQLGPETGATPNRQEGAGLMVRDILGAERLVPQPEGHEEFPSASNMVMNLM
RSHTRTNDGMTNINASFREGVYQPWGTPGNRLSRVDYVEGVPYGTAETYRMTLTRTNDGFKVSYRNGEKFIEQAVKGAN
ANIVEMQNSDSQYVGFFASRNAKMTVSNVDLQLAAADTVDAPKYAVKQGELVFKIASSPRSATKEYPVQARANYSGEFEV
LHNDKVVAKQTVTAGDLFSQWLTLDSGANQMEVRFTAIDGPNKETQAHRYSVDVVSLPDPMTLYVAPNGSDKGNGSQA
QPLDLATAVELLPTGGTIILKDGDYQGMEIPLTASGSADKLKHLRAEGDNVRFVSELRHEANYWHYQGIEVAGAQFIVHGS
HNTFEKMVTHGAPDTGFVITSPENIGRALWASYNQVIESESYNNMDPSQINADGFAAKMRIGDGNTFIRCLSHHNIDDG
WDLFNKVEDGANGAVTILDSISFSNGRTLDVANKGGTIGNGFKLGGEGIPVPHVVKNSLSFNNNMDGFTDNFNPGALVLS

DNVSIDNKRFNYLFRKSPYSGEIEQGTFTNNRSYRFHVSSKYDDVINSAKSTGNALVENGTTYTSDGKAVDSKMLAPLKQAS
VIDTQQAIPGKQEAMQLKQLIH
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMAEIKVEDLTWKAITFGQSTDMNFGSTILPEKVGVNQVTVNGEAVAAGKLASTFTI
ESRGGKLANSHEGLTFYYTELPTDVNFTLSADVVLEQLGPETGATPNRQEGAGLMVRDILGAERLVPQPEGHEEFPSASNM
VMNLMRSHTRTNDGMTNINASFREGVYQPWGTPGNRLSRVDYVEGVPYGTAETYRMTLTRTNDGFKVSYRNGEKFIEQ
AVKGANANIVEMQNSDSQYVGFFASRNAKMTVSNVDLQLAAADTVDAPKYAVKQGELVFKIASSPRSATKEYPVQARAN
YSGEFEVLHNDKVVAKQTVTAGDLFSQWLTLDSGANQMEVRFTAIDGPNKETQAHRYSVDVVSLPDPMTLYVAPNGSDK
GNGSQAQPLDLATAVELLPTGGTIILKDGDYQGMEIPLTASGSADKLKHLRAEGDNVRFVSELRHEANYWHYQGIEVAGA
QFIVHGSHNTFEKMVTHGAPDTGFVITSPENIGRALWASYNQVIESESYNNMDPSQINADGFAAKMRIGDGNTFIRCLSH
HNIDDGWDLFNKVEDGANGAVTILDSISFSNGRTLDVANKGGTIGNGFKLGGEGIPVPHVVKNSLSFNNNMDGFTDNFN
PGALVLSDNVSIDNKRFNYLFRKSPYSGEIEQGTFTNNRSYRFHVSSKYDDVINSAKSTGNALVENGTTYTSDGKAVDSKML
APLKQASVIDTQQAIPGKQEAMQLKQLIH
Comments:

- The sequence contains a signal peptide.
- The sequence was analysed at BSC and results indicates that is a good hyaluronidase candidate to sequence .
- Capable of degrading hyaluronic acid: production of 1-2 kDa products (**Figure 14**).

Nominated for: degrading hyaluronic acid, and thus relevant for cosmetic applications; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), as the host organism is highly active towards hyaluronic acid.

Nagoya protocol compliance: VD_PL9 was isolated from the shallow thermal white smokers (Fe^{2+} rich) of Basiluzzo Island (Chimney_Sagrada Familia, 38°39’25”N, 15°6’0”E). The sample was collected on 10.09.2021. Italy is not a Party to the Nagoya Protocol.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown those with >88% identity):

- Sequence 7 from patent US 9527906, 88.89% (accession ATJ64490.1)

Candidate number: 16

Name: Hyal_Vibrio alginolyticus (PL22) - VD_PL22
Activity: Intracellular pectin lyase - Hyaluronidase
Partner: CNR, CSIC, BSC
Source: *Vibrio diabolicus* (PL22). Isolation source: shallow thermal white smokers (Fe^{2+} rich) of Basiluzzo Island (Chimney_Sagrada Familia, 38°39’25”N, 15°6’0”E). Isolation date 10 September 2021.
Expression system: *Escherichia coli* (pET-45b(+))
Expression level: Soluble (expression +)
Sequence (no signal peptide):
MAKGDVITLNFETFVDSDTQVKVTRLTPTDVICHRNYFYQKCFTQDGKKLLFAGDFDGNRNYYLLNLETQQAVQLTEGKGD
NTFGGFISTDERAFFYVKNELNLMKVDLETLEEQVIYTVDEEWKGYGTWVANSDCTKLVGIEILKRDWQPLTSWEKFAEFY
HTNPTCRLIKVDIETGELEVIHQDTAWLGHPIYRPFDDSTVGFCHEGPHDLVDARMWLVNEDGSNVRKIKEHAEGESCTHE
FWIPDGSAMAYVSYFKGQTDRVIYKANPETLENEEVMVMPPCSHLMSNFDGSLMVGDGCDAPVDVADADSYNIENDPF
LYVLNTKAKSAQKLCKHSTSWDVLDGDRQITHPHPSFTPNDDGVLFTSDFEGVPAIYIADVPESYKH
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMAKGDVITLNFETFVDSDTQVKVTRLTPTDVICHRNYFYQKCFTQDGKKLLFAGDF
DGNRNYYLLNLETQQAVQLTEGKGDNTFGGFISTDERAFFYVKNELNLMKVDLETLEEQVIYTVDEEWKGYGTWVANSDC
TKLVGIEILKRDWQPLTSWEKFAEFYHTNPTCRLIKVDIETGELEVIHQDTAWLGHPIYRPFDDSTVGFCHEGPHDLVDARM
WLVNEDGSNVRKIKEHAEGESCTHEFWIPDGSAMAYVSYFKGQTDRVIYKANPETLENEEVMVMPPCSHLMSNFDGSLM
VGDGCDAPVDVADADSYNIENDPFLYVLNTKAKSAQKLCKHSTSWDVLDGDRQITHPHPSFTPNDDGVLFTSDFEGVPAIY
IADVPESYKH
Comments:

- The sequence was analysed at BSC and results indicates that is a good hyaluronidase candidate to sequence.

- Capable of degrading hyaluronic acid: production of 1-2 kDa products (**Figure 14**).

Nominated for: degrading hyaluronic acid, and thus relevant for cosmetic applications; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), as the host organism is highly active towards hyaluronic acid.

Nagoya protocol compliance: Hyal_Vibrio alginolyticus (PL22) - VD_PL was isolated from the shallow thermal white smokers (Fe^{2+} rich) of Basiluzzo Island (Chimney_Sagrada Familia, 38°39'25"N, 15°6'0"E). The sample was collected on 10.09.2021. Italy is not a Party to the Nagoya Protocol.

Patent match information (BLASTP against the Patented Protein Sequences Database, only shown the first 3 entries of those with >69% identity):

- Sequence 64 from patent US 9771566, 69.23% (accession AVB00750.1)
- Sequence 116 from patent US 9771566, 69.23% (accession AVB00802.1)
- Sequence 120 from patent US 9771566, 69.23% (accession AVB00806.1)

Candidate number: 17

Name: VA_PL9 (HA22_PG2209_02210 PL9)

Activity: Hyaluronidase

Partner: CNR, CSIC

Source: *Vibrio alginolyticus* #23. Isolation source: Lake of Ganzirri, Sicily, Italy (38°15'39.95"N 15°37'01.9"E)

Expression system: *Escherichia coli* (pET-45b(+))

Expression level: Soluble (expression +)

Sequence (signal peptide underlined):

MLIVAIVASLSISLIGPVTAAETSTNVTFVESDITANTTWTPDGGPYRIVQDVQVEPGATLTIRPGTEVQLAERITLSVDGSLYA
NGTADRPVTVSQTDGAAANRRWASIRYNGTDTSRLTVRNTTLEGGTSGITVDSSDGAIRVVDSTLRNFATAGLSVSDTAAT
PSITIERSAFRLIDGHAIRATPSMGTTGDVSLTASPSGRDLRSEHTLSLDAGVGVAFDTIELRYLSDGSVGNVESASLKRIGLDR
NRNGSVEQSFGSLVTDVSSTDGRLRISLSRPVKIPSDGRLLVEYEDAVNPSTRGIYPVGVDLRKESISQLSDGVRASFVVGNVT
SPIDRSTSVNTRANRLTVRGSSFNQVGGNGIFVAADTVRRLRLFDNRISETDGSGIAVRAARSEMDLSNNDISADDAGVQID
ARSQTSMAAYGNRIHDAQTGIRIHQSGTQVFRAGEISIRKNTLTDNVGHGIGIDARTTKLRFHLTDNTIRDNGRDGVNIQT
WLTRRSSVDGNQITDNGDDGFALAGAAVSNLTLAHNEVVNNSGTGMGVQTRATARKVTIHNNTVRDGTGHGVTVRSDL
LIHQVDVRENRLANNAGSGLLVASPITHRSNLSVVHNTIAANSYGVIVRGAVETAIRENDIVFNTNAFAEPVPVSDVYLGTG
VYVAEGSSGVIVNQAQAKIPLEELVANPEMTEELTVAQLWDDTVVVLRTDGESETRPAEASSLLIRQVSGTTPTGVGIQKSR
GSVQNHRIVNNSVYGQHRGLVVDMAPLINVNTTARIIVDPIRTVHAESNYWGTESGPYHSSILPAGDGNSVVTERGWVDF
TPFRTATSGPRYERPTTRIEAPATATAKSQLRLSGANSTSNHSPVGRYHFVVNGTAQPAQSSPVLNVTMPNQSLEVRLSVE
NRLGIDSNNASTATIDTAEPVAMEQPAQSKTTSSNGPLTSVDAIDSGGLSAGLGSIWGLLGGVLYLGALVLGGHGMVLTLQ
NRSTPVSGRRIQGLAVAGVLVWVVGGLVSAGPLVSIGIVAVVSWGGLTAVAYVLATRG

Synthesized sequence once introduced in the expression vector:

MAHHHHHHVGTGSNDDDDKSPDPMAETSTNVTFVESDITANTTWTPDGGPYRIVQDVQVEPGATLTIRPGTEVQLAERI
TLSVDGSLYANGTADRPVTVSQTDGAAANRRWASIRYNGTDTSRLTVRNTTLEGGTSGITVDSSDGAIRVVDSTLRNFATA
GLSVSDTAATPSITIERSAFRLIDGHAIRATPSMGTTGDVSLTASPSGRDLRSEHTLSLDAGVGVAFDTIELRYLSDGSVGNVE
SASLKRIGLDRNRNGSVEQSFGSLVTDVSSTDGRLRISLSRPVKIPSDGRLLVEYEDAVNPSTRGIYPVGVDLRKESISQLSDGV
RASFVVGNVTSPIDRSTSVNTRANRLTVRGSSFNQVGGNGIFVAADTVRRLRLFDNRISETDGSGIAVRAARSEMDLSNNDI
SADDAGVQIDARSQTSMAAYGNRIHDAQTGIRIHQSGTQVFRAGEISIRKNTLTDNVGHGIGIDARTTKLRFHLTDNTIRDN
GRDGVNIQTWLTRRSSVDGNQITDNGDDGFALAGAAVSNLTLAHNEVVNNSGTGMGVQTRATARKVTIHNNTVRDGT
GHGVTVRSDLLIHQVDVRENRLANNAGSGLLVASPITHRSNLSVVHNTIAANSYGVIVRGAVETAIRENDIVFNTNAFAEPV
PVSDVYLGTGVYVAEGSSGVIVNQAQAKIPLEELVANPEMTEELTVAQLWDDTVVVLRTDGESETRPAEASSLLIRQVSGTT
PTGVGIQKSRGSVQNHRIVNNSVYGQHRGLVVDMAPLINVNTTARIIVDPIRTVHAESNYWGTESGPYHSSILPAGDGNSV
VTERGWVDFTPFRTATSGPRYERPTTRIEAPATATAKSQLRLSGANSTSNHSPVGRYHFVVNGTAQPAQSSPVLNVTMPN
QSLEVRLSVENRLGIDSNNASTATIDTAEPVAMEQPAQSKTTSSNGPLTSVDAIDSGGLSAGLGSIWGLLGGVLYLGALVLG
GHGMVLTLQNRSTPVSGRRIQGLAVAGVLVWVVGGLVSAGPLVSIGIVAVVSWGGLTAVAYVLATRG

Comments:

- The sequence contains a signal peptide.
- Capable of degrading hyaluronic acid: production of 1-2 kDa products (**Figure 14**).

Nominated for: degrading hyaluronic acid, and thus relevant for cosmetic applications; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), as the host organism is highly active towards hyaluronic acid.

Nagoya protocol compliance: VA_PL9 (HA22_PG2209_02210 PL9) was isolated from the Lake of Ganzirri, Sicily, Italy (38°15'39.95"N 15°37'01.9"E). The sample was collected on 10.09.2021. Italy is not a Party to the Nagoya Protocol.

Patent match information (BLASTP against the Patented Protein Sequences Database):

- Sequence 16074 from patent US 6833447, 32.38% (accession AAW98511.1)

Candidate number: 18

Name: Hyal_HRDSV_2334 (ID 3)
Activity: Hyaluronidase
Partner: CNR, CSIC, BSC
Source: *Halorhabdus* sp. SivX81
Expression system: *Escherichia coli* (pET-45b(+))
Expression level: Soluble (low expression)
Sequence (no signal peptide):
MCSALDGGSSPTDTADSATPTDGSDAPTSPEAESTEPNGDSWDLSWSDEFDGEEIDEAVWNFETGNGHPDVPGWGNEE
PQYYQRENAWLEDDHLVIEAREEHVEDDYGEYDYTSSRLNTQGAVATKYGRVDVRARLPRGQGIWPRIWMVGSDVAFA
GWPDCGEIDILEFRGDKPETVRGFAQGPGYTGADRLTDSYTVEEGALTDSTHVFSIRWQSDRIEWYVDETQYHTVSRETVE
DAGNEWVFDTPMYFVLAFACGGDPAGYPDETTPFPQRLEVDYVRHYERV
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMCSALDGGSSPTDTADSATPTDGSDAPTSPEAESTEPNGDSWDLSWSDEFDGEEI
DEAVWNFETGNGHPDVPGWGNEEPQYYQRENAWLEDDHLVIEAREEHVEDDYGEYDYTSSRLNTQGAVATKYGRVDV
RARLPRGQGIWPRIWMVGSDVAFAGWPDCGEIDILEFRGDKPETVRGFAQGPGYTGADRLTDSYTVEEGALTDSTHVFSI
RWQSDRIEWYVDETQYHTVSRETVEDAGNEWVFDTPMYFVLAFACGGDPAGYPDETTPFPQRLEVDYVRHYERV
Comments: Capable of degrading hyaluronic acid: production of 1-2 kDa products (**Figure 14**).
Nominated for: degrading hyaluronic acid, and thus relevant for cosmetic applications; possibly nominated for WP5 (genetic and supramolecular engineering), WP6 (large scale production) and WP7 (pre-industrial validations), as the host organism is highly active towards hyaluronic acid.

Nagoya protocol compliance: Hyal_HRDSV_2334 (ID 3) was isolated from lakes Razvel and Dunino, Sol-Iletsk region, Russia (51.151°N, 55.003°E). The sample was collected in June 2019. Russia is not a Party to the Nagoya Protocol.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown those with >51% identity):

- Sequence 7 from patent US 9527906, 100.00% (accession ATJ64490.1)
- Sequence 5 from patent US 9493751, 51.22% (accession ATJ14040.1)

Candidate number: 19

Name: *Vibrio diabolicus*_V4 and *Vibrio alginolyticus*_#23 (hosts positive for hyaluronidases)
Partner: CNR
Source: *Vibrio diabolicus*_V4 (Basiluzzo submarine hydrothermal field, Panarea Island, Aeolian Archipelago, South Tyrrhenian Sea, 38°40.315'N; 15°07.846'E) and *Vibrio alginolyticus*_#23 (Lake of Ganzirri, Sicily, Italy (38°15'39.95"N 15°37'01.9"E), genomes available
Expression system: native hosts *Vibrio diabolicus*_V4 and *Vibrio alginolyticus*_#23
Expression level: Soluble (expression +)
Sequence: to be identified, strains' genome available
Comments:

- When cultivated in the presence of hyaluronic acid, degradation products of 1-2 kDa products are observed (**Figure 14**).
- These two candidates are two native hosts that are herein selected as candidates either for scale up their fermentations (WP3 or WP6) and for optimization of the cultivation conditions to produce native

hyaluronidades with which to degrade hyaluronic acid. Alternatively, their genomes are being analysed for hyaluronidases.

Nominated for: degrading hyaluronic acid, and thus relevant for cosmetic applications; possibly nominated for WP6 (large scale production) to optimize fermentation conditions to produce native hyaluronidases produced by these two strains.

Nagoya protocol compliance: Vibrio diabolicus_V4 and Vibrio alginolyticus_#23 (hosts positive for hyaluronidases) was isolated from the shallow thermal white smokers (Fe^{2+} rich) of Basiluzzo Island (Chimney_Sagrada Familia, 38°39'25"N, 15°6'0"E) and the Lake of Ganzirri, Sicily, Italy (38°15'39.95"N 15°37'01.9"E). The sample was collected on 10.09.2021 Italy is not a Party to the Nagoya Protocol.

Candidate number: 20

Name: FE_EH37
Activity: Esterase
Partner: CSIC
Source: Milazzo harbor (Sicily, Italy; 38°12'30.10''N, 15°15'34.89''E), that generally suffers chronic petroleum pollution because of intensive maritime traffic and its limited hydrodynamic regimen and restricted area.
Expression system: *Escherichia coli* (Ek/LIC 46)
Expression level: Soluble (expression +)
Sequence (no signal peptide):
MSGDNPFDPELYKDAAVSAETRALNTALIDLLETSDDNWDIGVEEARARRDRGEGPFPAVPKSPRARTIQIPGKGGDIALRII
APETPKGVYLHFHGGGWVFGSADGQDPMLERISDTTGLVCVSVEYRLAPEHPYPAGPDDCESAALWLVENAKREFGTDLL
TIGGESAGGHLAAVTLLRMRDRHGFTGFAGANLVFGAFDLRWTPSARSYGNDRYLILRTLDLEKFDACFLPENVDRADPDI
SPLMANLHDMPPALFTVGTDDALLDDSLFMHARWAAAGNEAELAVYPGGAHGFVAFPGALAASAVQRMDAFLKRFTD
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMSGDNPFDPELYKDAAVSAETRALNTALIDLLETSDDNWDIGVEEARARRDRGEG
PFPAVPKSPRARTIQIPGKGGDIALRIIAPETPKGVYLHFHGGGWVFGSADGQDPMLERISDTTGLVCVSVEYRLAPEHPYP
AGPDDCESAALWLVENAKREFGTDLLTIGGESAGGHLAAVTLLRMRDRHGFTGFAGANLVFGAFDLRWTPSARSYGNDR
YLILRTLDLEKFDACFLPENVDRADPDISPLMANLHDMPPALFTVGTDDALLDDSLFMHARWAAAGNEAELAVYPGGAHG
FVAFPGALAASAVQRMDAFLKRFTD
Comments:

- Availability of experimental structural data.
- Availability of a large number of single mutants.

Nominated for: computational and experimental investigation of thermal behaviour, to allow establishing predictive tools of thermostability (**Figure 15**).

Nagoya protocol compliance: The enzyme was retrieved from microbial communities inhabiting seawater samples at the Milazzo harbor (Sicily, Italy; 38°12'30.10''N, 15°15'34.89''E), that generally suffers chronic petroleum pollution because of intensive maritime traffic and its limited hydrodynamic regimen and restricted area. The sample was collected April 2011, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown those with >49% identity):

- Sequence 12316 from patent US 9719064, 49.09% (accession AUF66202.1)

Candidate number: 21

Name: FE_Lip5
Activity: Lipase
Partner: CSIC, BSC
Source: MarRef - Marine Metagenomics Database
Expression system*: Escherichia coli* (pET-45b(+))
Expression level: Soluble (expression +)
Sequence (no signal peptide):

MTVSALDHRVTGYQLDHAYWLGKAAKLAYSGEEEIRAETARWGFDRFRFLHVVRDLPVPLDDTQAYLAASDHMIILAFRG
TEITQIKDWLTDATTPVAPGPADRGLVHLGFDQALATVLPLVCQGIKELRTNDQSIWLTGHSLGGALAMLAAATLYFEDPN
LTPDGVYTFGQPRTCDPRLAHAYDEALEGRTFRFVNNNDIVPQLPPEPVFRHVKAARYFDRTGALHEQLSLWGGIADKVG
GHTDELLLPGSDALKDHPMDRYLENIEKNL
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMTVSALDHRVTGYQLDHAYWLGKAAKLAYSGEEEIRAETARWGFDRFRFLHVVR
DLPVPLDDTQAYLAASDHMIILAFRGTEITQIKDWLTDATTPVAPGPADRGLVHLGFDQALATVLPLVCQGIKELRTNDQSI
WLTGHSLGGALAMLAAATLYFEDPNLTPDGVYTFGQPRTCDPRLAHAYDEALEGRTFRFVNNNDIVPQLPPEPVFRHVKA
ARYFDRTGALHEQLSLWGGIADKVGGHTDELLLPGSDALKDHPMDRYLENIEKNL
Biochemical features:

- FE_Lip5 is most active towards glyceryl tributyrate and glyceryl trioctanoate with low or no activity towards medium and long chain triglycerides.
- FE_Lip5 is most active at 45˚C, retaining about 25-35% of the maximum activity at 25-30˚C, and pH 6.5.

Comments: this lipase a priory may not be of interest for detergent applications, but it was selected for enhancing the hydrolytic activity of this lipase towards larger triglycerides through lid domain engineering. In brief, the engineering of the lid domain of lipase from the *Actinoalloteichus genus* (NCBI Accession Number: WP_075743487.1) to switch from the hydrolysis of small-chain triglycerides to medium-chain ones through site-directed mutagenesis (W89M/L60F) and to long-chain ones through lid swapping (with the lid domain of *Rhizopus oryzae* lipase). The datasets already generated confirmed the "lid domain substitution" as efficient approach to engineer lipases.

Nagoya protocol compliance: The sequence was identified by homology screen in the MarRef - Marine Metagenomics Database, a manually curated marine microbial reference genome database that contains completely sequenced genomes (https://mmp2.sfb.uit.no/marref/). This protein is similar to one isoltaed from a strain isolated from the seawater in Trondheimsfjord (Norway), front of Seongsan-ri, Jeju Island, South Korea (E-value of 2,53E-11 with WP_075743487.1_MMP03339750 MULTISPECIES: lipase [Actinoalloteichus] [mmp_id=MMP03339750] [mmp_db=marref]]). Sample was collected in 2005 (https://www.ebi.ac.uk/biosamples/samples/SAMN03339750), so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown those with >43% identity):

- Sequence 34862 from patent US 10696975, 43.52% (accession QPU44530.1)

Candidate number: 22

Name: TR_2E_2
Activity: PluriZyme (transaminase/esterase)
Partner: CSIC, BSC, UDUS
Source: The beach acidic pool on Vulcano Island (Aeolian archipelago; 38.4162° N, 14.9603° E)
Expression system: *Escherichia coli* (pET-45b(+))
Expression level: Soluble (expression +)
Sequence (no signal peptide):
MSQSQRSTADWQRLDAAHHLHPFTDYGELNTKGSRIITRAEGCYLWDSDGNQILDGMAGLWCVNIGYGRKELAEVAYR
QMQELPYYNNFFQCSHPPAIELSRLLSEVTPKHMNHVFFTGSGSDSNDTILRMVRYYWKLLGKPYKKVVISRENAYHGSTV
AGASLSGMKAMHSHGDLPIPGIEHIEQPYHFGRAPDMDPAEFGRQAAQALERKIDEIGECNVAAFIAEPIQGAGGVIIPPD
SYWPEIKRICAERDILLIVDEVITGFGRLGTWFGSQYYDLQPDLMPIAKGLSSGYMPIGGVMVSDRVAKVVIEEGGEFFHGY
TYSGHPVAAAVAAENIRIMRDEGIIERAGAEIAPYLQARWRELGEHPLVGEARGVGMVAALELVKSKQPLERFEEPGKVGS
LCRDLSVKNGLVMRAVGGTMIISPPLVLSREQVDELIDKARRTLDETHKAIGGA
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMSQSQRSTADWQRLDAAHHLHPFTDYGELNTKGSRIITRAEGCYLWDSDGNQIL
DGMAGLWCVNIGYGRKELAEVAYRQMQELPYYNNFFQCSHPPAIELSRLLSEVTPKHMNHVFFTGSGSDSNDTILRMVRY
YWKLLGKPYKKVVISRENAYHGSTVAGASLSGMKAMHSHGDLPIPGIEHIEQPYHFGRAPDMDPAEFGRQAAQALERKID

EIGECNVAAFIAEPIQGAGGVIIPPDSYWPEIKRICAERDILLIVDEVITGFGRLGTWFGSQYYDLQPDLMPIAKGLSSGYMPI
GGVMVSDRVAKVVIEEGGEFFHGYTYSGHPVAAAVAAENIRIMRDEGIIERAGAEIAPYLQARWRELGEHPLVGEARGVG
MVAALELVKSKQPLERFEEPGKVGSLCRDLSVKNGLVMRAVGGTMIISPPLVLSREQVDELIDKARRTLDETHKAIGGA
Biochemical features:

- The original enzyme TR_2 is an omega-amine transaminase.
- Availability of experimental structural data.

Comments: this omega-amine transaminase was selected for WP5 engineering as a target to design PluriZymes, with two different active sites. In brief, by applying Protein Energy Landscape Exploration (PELE) software, 20 mutants were generated potentially incorporating catalytic triads supporting ester hydrolysis. After gene synthesis and production and experimental validation, one variant named TR_2E_2 was an efficient PluriZyme with both transaminase and esterase active sites. The datasets already generated confirmed that the development of PluriZymes seems to be a robust procedure, with a nearly 90% success rate, and therefore the technology can be now used to target priority lipases, hyaluronidades and oxidoreductases relevant to the project.

Nagoya protocol compliance: The enzyme was isolated from the metagenome of microbial communities inhabiting the beach acidic pool on Vulcano Island (Aeolian archipelago; 38.4162° N, 14.9603° E). The sediments were collected at the Levante Bay (Vulcano Island, Aeolian archipelago, Italy) on 02.10.2012, so that before the Nagoya Protocol entered into force on 12 October 2014. Italy is not a Party to the Nagoya Protocol.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown those with >66% identity):

- Sequence 11896 from patent US 8343764, 66.74% (accession AGD14440.1)
- Sequence 29725 from patent US 8343764, 66.44% (accession AGD32269.1)
- Sequence 27768 from patent US 7517684, 66.30% (accession ACQ35254.1)

Candidate number: 23

Name: EH_{1AB1C}
Activity: PluriZyme (esterase/protease)
Partner: CSIC, BSC
Source: Evaporite karstic lake Arreo in Spain (42°46'N, 2°59'W; altitude, 655 m)
Expression system: *Escherichia coli* (pET-45b(+))
Expression level: Soluble (expression +)
Sequence (no signal peptide):
MLLPETRNLLDLMDAATRGGRPGCDTLPHAVGRKAVDKMSEDGEADPPEVAEVANGGFAGPASEIRFRRYRPLGEAAGL
LPTLIYYHGGGFVIGNIETHDSTCRRLANKSRCQVISIDYRLAPEHPFPAPIDDGIAAFRHIRDNAESFGADAARLAVGGDSAG
GAMAAVVCQACRDAGETGPAFQMLIYPATDSSRESASRVAFAEGYFLSKAHMDWFWEAYVPEDTDLTDLRLSPLLATDF
TGLPPAFVLTAGYDPLRDEGRAYADRLIEAGIKTTYVNYPGTIHGFFSLTRFLSQGLKANDEAAAVMGAHFGT
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPMLLPETRNLLDLMDAATRGGRPGCDTLPHAVGRKAVDKMSEDGEADPPEVAEV
ANGGFAGPASEIRFRRYRPLGEAAGLLPTLIYYHGGGFVIGNIETHDSTCRRLANKSRCQVISIDYRLAPEHPFPAPIDDGIAAF
RHIRDNAESFGADAARLAVGGDSAGGAMAAVVCQACRDAGETGPAFQMLIYPATDSSRESASRVAFAEGYFLSKAHMD
WFWEAYVPEDTDLTDLRLSPLLATDFTGLPPAFVLTAGYDPLRDEGRAYADRLIEAGIKTTYVNYPGTIHGFFSLTRFLSQGLK
ANDEAAAVMGAHFGT
Biochemical features:

- The original enzyme EH_{1AB1} is a PluriZyme with two active sites supporting ester hydrolysis, one native, and one artificially introduced.
- Availability of experimental structural data.

Comments: this PluriZyme (EH_{1AB1}) was selected for WP5 engineering as a target to design PluriZymes with two different active sites supporting different chemistry. In brief, by applying Protein Energy Landscape Exploration (PELE) software a mutant was generated potentially incorporating a catalytic dyad supporting proteolytic activity. After gene synthesis and production and experimental validation, the variant named

EH_{1AB1C} was an efficient PluriZyme with both protease and esterase active sites. The datasets already generated confirmed that the development of PluriZymes seems to be a robust procedure, with a nearly 90% success rate, and therefore the technology can be now used to target priority lipases, hyaluronidades and oxidoreductases relevant to the project.

Nagoya protocol compliance: The enzyme was isolated from the metagenome of microbial communities inhabiting the evaporite karstic lake Arreo in Spain (42°46'N, 2°59'W; altitude, 655 m). Sediment sampling was carried out in February 2007, so that before the Nagoya Protocol entered into force on 12 October 2014; all documentation (including pictures) that shows where the samples were taken can be made available upon request.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown those with >50% identity):

- Sequence 748 from patent US 8298799, 64.19% (accession AFX20859.1)
- Sequence 498 from patent US 8298799, 50.00% (accession AFX20734.1)

Candidate number: 24

Name: X11_mut1
Activity: PluriZyme (xylanase/esterase)
Partner: CSIC, BSC
Source: *Pseudothermotoga thermarum*
Expression system: *Escherichia coli* (pQE80L)
Expression level: Soluble (expression +)
Sequence (no signal peptide):
QIPSLKDVFLQDFKIGVALPVRVFSNSMDVELITKHFNSMTAENEMKPESILRRDASGKIYYDFTVADRYIEFAQKHGMVVR
GHTLVWHSQTPEWFFKDEKGNLLSREAMIERMREYIHTVVGRYRGKVYAWDVVNEAVDENQPDGLRRSLWYQVIGPDY
IELAFKFAHEADPDALLFYNDYNEFFPKKRDIIFKLVKEMREKGVPIHGIGMQQHLTLADNVGWIDIAIQKFKTISGIQIHITEL
DVSVYKSRSPSIIYQTPPSEVLHEQAEFYRKLFEIYRKHTDVITNVTFWGLKDDYSWLRFFFGRRNDWPLLFDENYQPKPAF
WSVIESVS
Synthesized sequence once introduced in the expression vector:
MAHHHHHHVGTGSNDDDDKSPDPQIPSLKDVFLQDFKIGVALPVRVFSNSMDVELITKHFNSMTAENEMKPESILRRDAS
GKIYYDFTVADRYIEFAQKHGMVVRGHTLVWHSQTPEWFFKDEKGNLLSREAMIERMREYIHTVVGRYRGKVYAWDVVN
EAVDENQPDGLRRSLWYQVIGPDYIELAFKFAHEADPDALLFYNDYNEFFPKKRDIIFKLVKEMREKGVPIHGIGMQQHLTL
ADNVGWIDIAIQKFKTISGIQIHITELDVSVYKSRSPSIIYQTPPSEVLHEQAEFYRKLFEIYRKHTDVITNVTFWGLKDDYSWL
RFFFGRRNDWPLLFDENYQPKPAFWSVIESVS
Biochemical features:

- The original enzyme X11 is a xylanase supporting xylan hydrolysis at temperatures as high as 80-90˚C.
- Availability of experimental structural data.

Comments: the xylanase X11 was selected for WP5 engineering as a target to design PluriZymes with two different active sites supporting different chemistry. In brief, by applying Protein Energy Landscape Exploration (PELE) software, a mutant was generated potentially incorporating a catalytic triad supporting feruloyl esterase activity. After gene synthesis and production and experimental validation, the variant named X11_mut was an efficient PluriZyme with both protease and esterase active sites. The datasets already generated confirmed that the X11_mut PluriZyme has an improved capacity to degrade xylan because a synergetic capacity to degrade different xylan components. These data confirmed that the development of PluriZymes seems to be a robust procedure, with a nearly 90% success rate, and therefore, the technology can be now used to target priority lipases, hyaluronidades and oxidoreductases relevant to the project.

Nagoya protocol compliance: The enzyme was isolated from the genome of *Pseudothermotoga thermarum*, from the NCBI database. Its genome was sequences by the US DOE Joint Genome Institute (JGI-PGF), and the sample (continental solfataric spring, Lac Abbe) from which the strain was isolated (BioSample: SAMN02232059), located at Djibouti (11.162815 N 41.780548 E) was taken before July 2013, so that before the Nagoya Protocol entered into force on 12 October 2014.

Patent match information (BLASTP against the Patented Protein Sequences Database,only shown the first three entries of those with >64% identity):

- Sequence 86 from patent US 7504120, 64.78% (accession ACP87319.1)
- Sequence 466 from patent US 8043839, 64.67% (accession AER76128.1)
- Sequence 36 from patent US 7504120, 64.67% (accession ACP87294.1)

4. Conclusion: summary nominated enzymes and isolates

According to the information provided above, until month 18 a set of 22 enzyme candidates and two isolates (*Vibrio diabolicus* and *Vibrio alginolyticus*), were selected as candidates for WP5-WP7. Their information is summarized in **Table 2**. About priorization, below the first priority candidates:

- FE_Lip9 (lipase and PET hydrolase activity) for detergent and textile applications because is capable to degrade all the stained fabrics tested and spinning oils in Schoeller's textiles; it is also stable in washing liquour.
- FE_ID9 (lipase activity) for detergent applications because is capable to degrade all the stained fabrics tested; it is also stable in washing liquour.
- Paes-PE-H Y250S (PET esterase activity, washing liquor stability, stain activity) for detergent and textile applications; structural data is available.
- Dim008 aka AZ Lipase (lipase activity on stains, thermal stability) for textile applications and, with some informed mutagenesis to increase detergent resistance, also for detergent application.
- GEN0105 (lipase) for detergent applications because is capable to degrade stained fabrics comparable with Henkel enzyme mix, in the buffered washing liquor, and spinning oils in Schoeller's textiles; it is also stable in washing liquour.
- Strains *Vibrio diabolicus* V4 and *Vibrio alginolyticus* #23 for cosmetic applications, as when cultivating in the presence of hyaluronic acid, degradation product with the size of interest are observed.

Second priority:

- Paes_TB035 because of its remarkable stabilities; structural data is available.
- PtEst1 because it combines relatively high activities on stains with rather high activity in washing liquor; structural data is available.
- FE_Polur1, for detergent applications because is capable to degrade stained fabrics; it is also stable in washing liquour.

Table 2. Summary of enzymes and isolates selected as best candidates for WP5-WP7.

ID	Name	Partner	Priority	Signal Peptide	Nagoya protocol	Homology (%) patent database	Application
1	Kest3 (lipase)	Bangor	No	No	Information provided	33.87	Detergent
2	FE_Lip9 (lipase)	CSIC	Yes (1st)	Yes	Information provided	99.45	Detergent, textile
3	FE_ID9 (lipase)	CSIC	Yes (1st)	No	Information provided	100	Detergent
4	FE_polur1 (lipase)	CSIC	Yes (2nd)	No	Information provided	97.3	Detergent
5	EstLip_Dim_#008 (lipase)	UDUS	Yes (1st)	No	Information provided	100	Detergent
6	EstLip_Paes_TB035 (lipase)	UDUS	Yes (2nd)	Yes	Information provided	41.98	Detergent
7	EstLip_PtEst1 (lipase)	UDUS	Yes (2nd)	No	Information provided	58.63	Detergent
8	EstLip_TBEc304 (lipase)	UDUS	No	No	Information provided	62.54	Detergent
9	PEH_Paes_PE-H_Y250S (PETase)	UDUS	Yes (1st)	Yes	Information provided	62.88	Detergent, textile
10	PEH_Pbau_PE-H (Lipase, PETase)	UDUS	No	Yes	Information provided	61.70	Detergent, textile
11	PEH_Pform_PE-H (Lipase, PETase)	UDUS	No	Yes	Information provided	69.08	Textile
12	PEH_Poce_PE-H (Lipase, PETase)	UDUS	No	Yes	Information provided	62.5	Detergent, textile
13	GEN0105 (Lipase, PETase)	Bangor	No	No	Information provided	61.69	Detergent
14	GEN0095 (cellulase)	Bangor	No	No	Information provided	52.5	Textile
15	VD_PL9 (hyaluronidase)	CNR	Yes	Yes	Information provided	88.89	Cosmetic
16	VD_PL22 (hyaluronidase)	CNR	No	No	Information provided	69.23	Cosmetic
17	VA_PL9 (hyaluronidase)	CNR	No	Yes	Information provided	32.38	Cosmetic
18	Hyal_HRDSV_2334 (hyaluronidase)	CNR	No	No	Information provided	100	Cosmetic
19	*V. diabolicus* V4; *V. alginolyticus* #23 (hyaluronidase)	CNR	Yes	-	Information provided	-	Cosmetic
20	FE_EH37 (esterase)	CSIC	No	No	Information provided	49.09	Predictive tools
21	FE_Lip5 (lipase)	CSIC	No	No	Information provided	43.52	Detergent
22	TR_2E_2 (PluriZyme)	CSIC	No	No	Information provided	66.74	All
23	EH_{1AB1C} (PluriZyme)	CSIC	No	No	Information provided	64.19	All
24	X11_mut1 (PluriZyme)	CSIC	No	No	Information provided	64.78	All

5. Figures

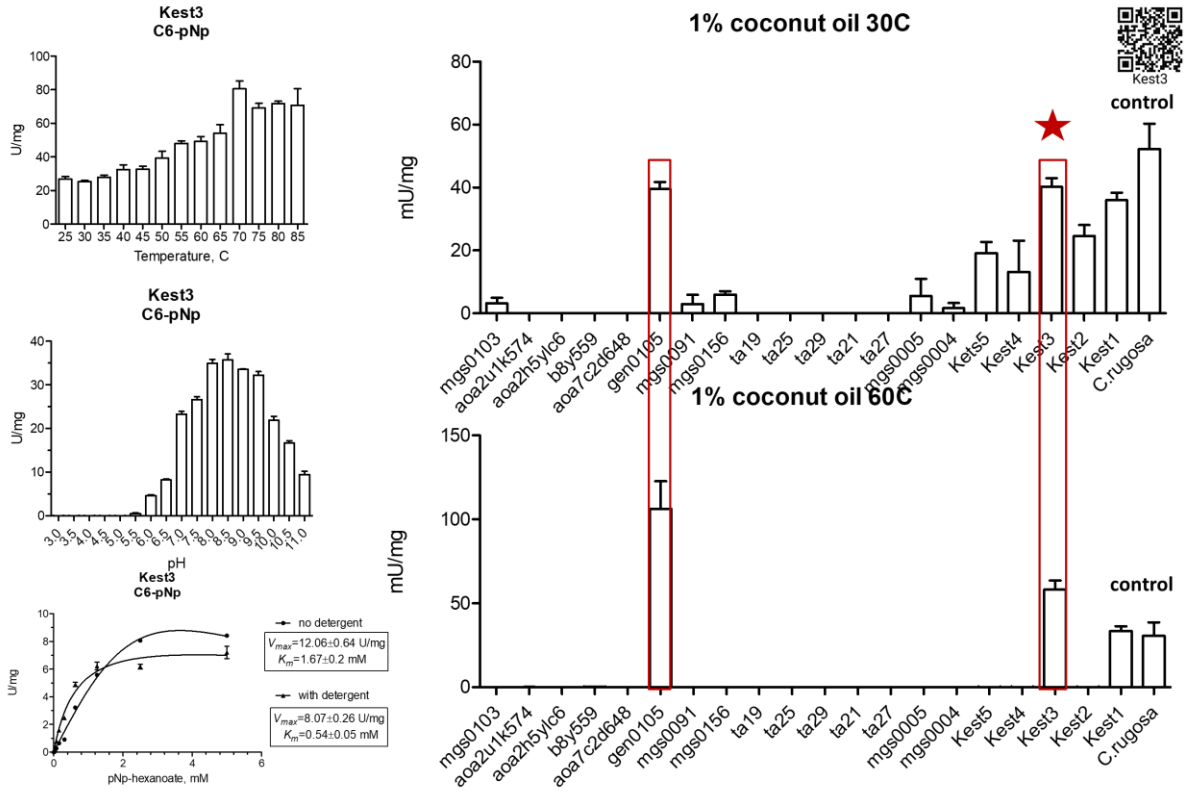


Figure 1. Characteristics of Kest3 lipase (candidate 1). Gen0105 (candidate 13) is also shown.

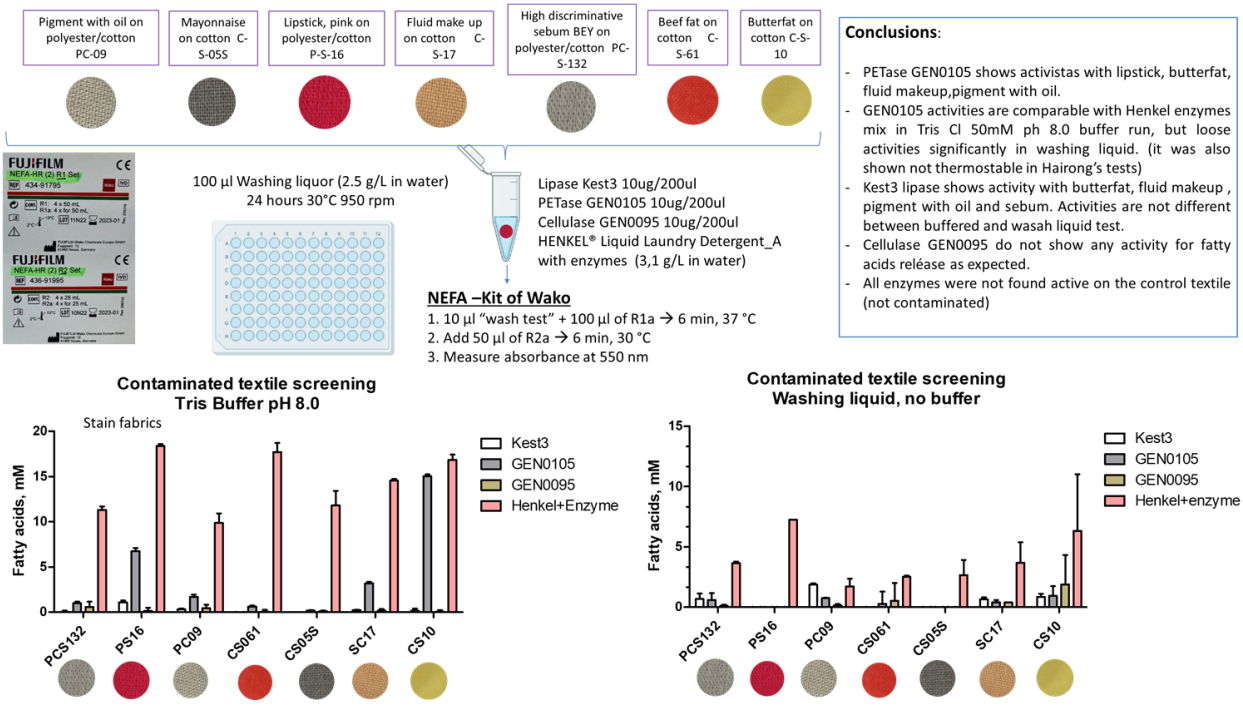


Figure 2. Capacity of Kest3 (candidate 1), GEN0105 (candidate 13) and GEN0095 (candidate 14) lipases to degrade stained swatches from CFT (Center For Testmaterials), compared to enzyme-containing Henkel liquid detergent.

Triglyceride	U/g		
	FE_Lip9	FE_Polur1	FE_ID9
Glyceryl tripropionate (TriC3)	471.0	169.6	12.0
Glyceryl tributyrate (TriC4)	965.1	348.3	100.4
Glyceryl trioctanoate (TriC8)	1720.1	392.1	1604.1
Glyceryl tridecanoate (TriC10)	886.8	156.0	174.7
Glyceryl tridodecanoate (TriC12)	314.2	117.9	4.0
Glyceryl trimyristate (TriC14)	111.4	33.7	0.4
Coconut oil	582.7	254.0	6.0
Palm oil	144.5	30.5	4.0
Olive oil	73.8	27.3	0.5

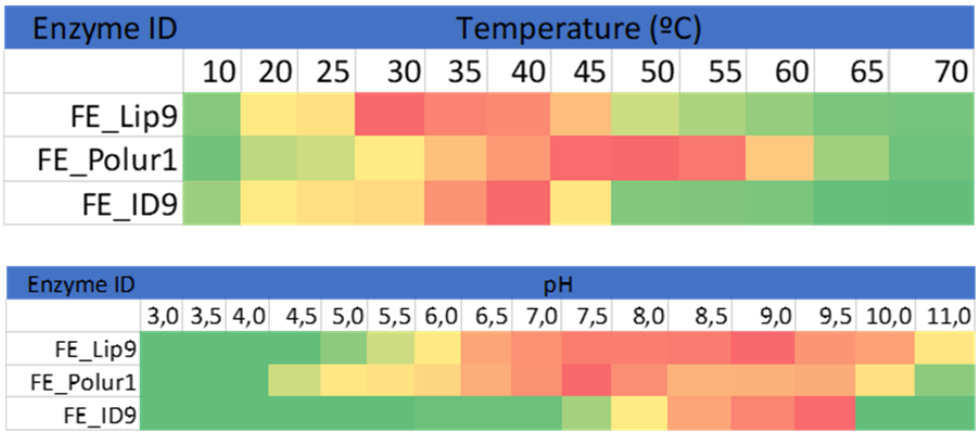


Figure 3. Characteristics of FE_Lip9 (candidate 2), FE_Polur (candidate 4) and FE_ID9 lipases (candidate 3).

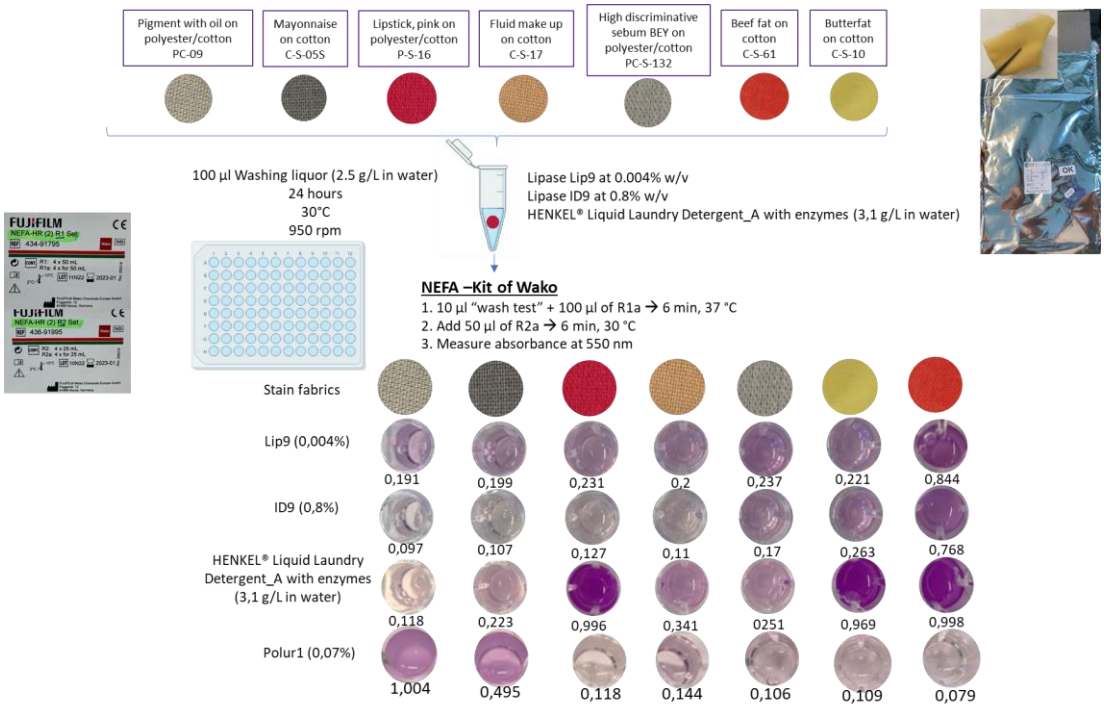


Figure 4. Capacity of FE_Lip9 (candidate 2), FE_Polur (candidate 4) and FE_ID9 lipases (candidate 3) lipases to degrade stained swatches from CFT (Center For Testmaterials), compared to enzyme-containing Henkel liquid detergent.

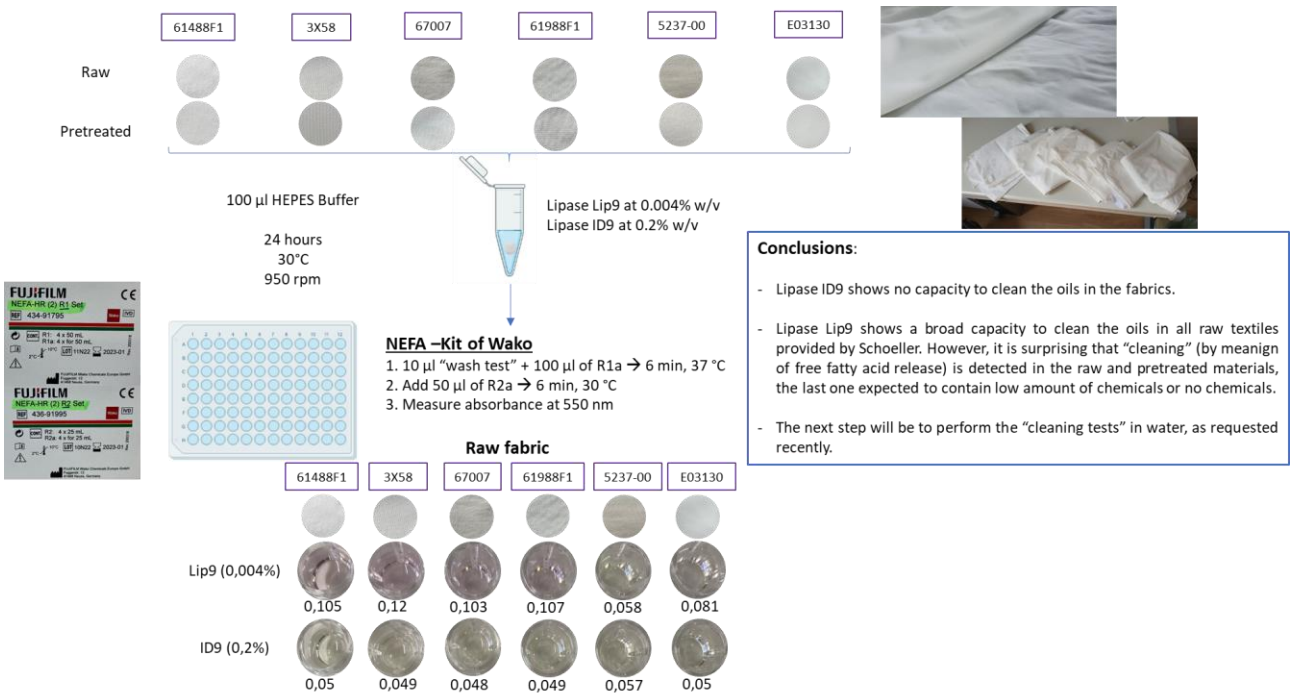


Figure 5. Capacity of FE_Lip9 (candidate 2) and FE_ID9 (candidate 3) lipases to clean the oils in all raw Scholler's textiles (61488F1, 3X58, 67007, 61988F1, 5237-00 and E03130).

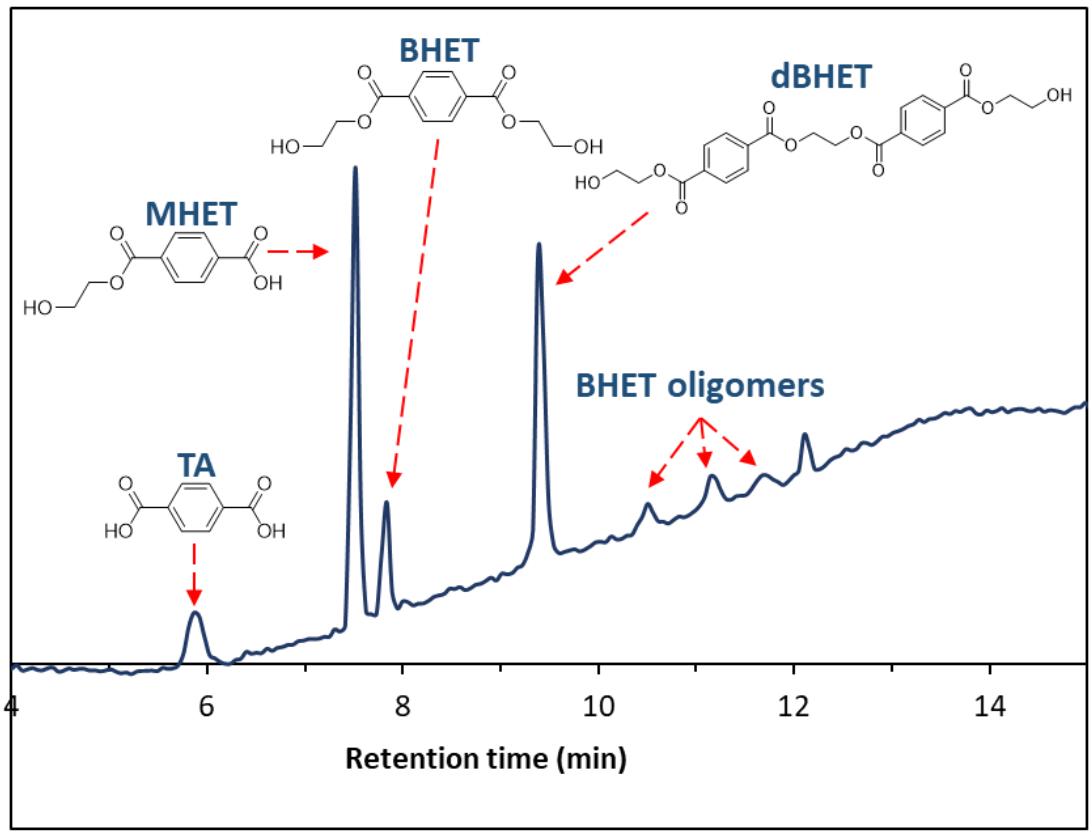


Figure 6. Capacity of FE_Lip9 (candidate 2) to degrade PET.

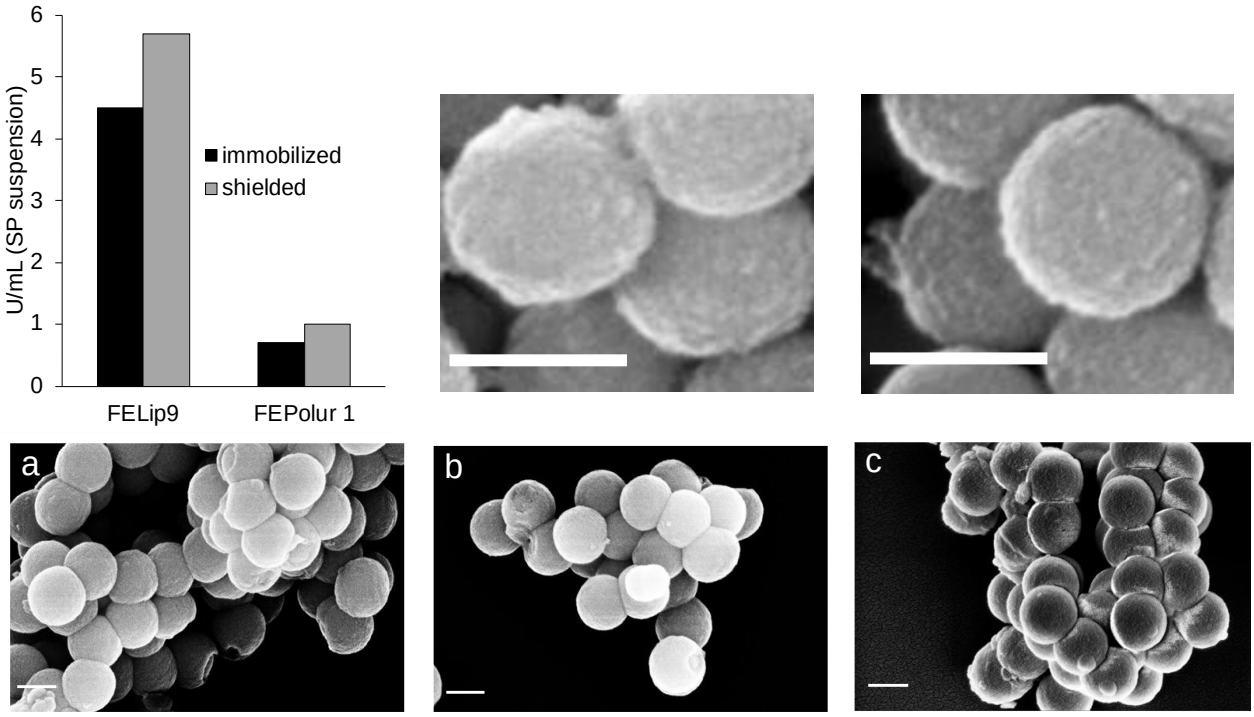


Figure 7. Immobilization-guided supramolecular engineering of FE_Lip9 (candidate 2) and FE_Polur1 (candidate 4). These lipases were immobilized on amino modified silica particles (SP, diameter 230 nm), and shielded with an organosilica layer. Glutaraldehyde was used as homo-bifunctional crosslinker. The lower images show SEM micrographs of lipase-based nanobiocatalysts shielded with an organosilica layer of controlled thickness. All scale bars represent 200 nm.

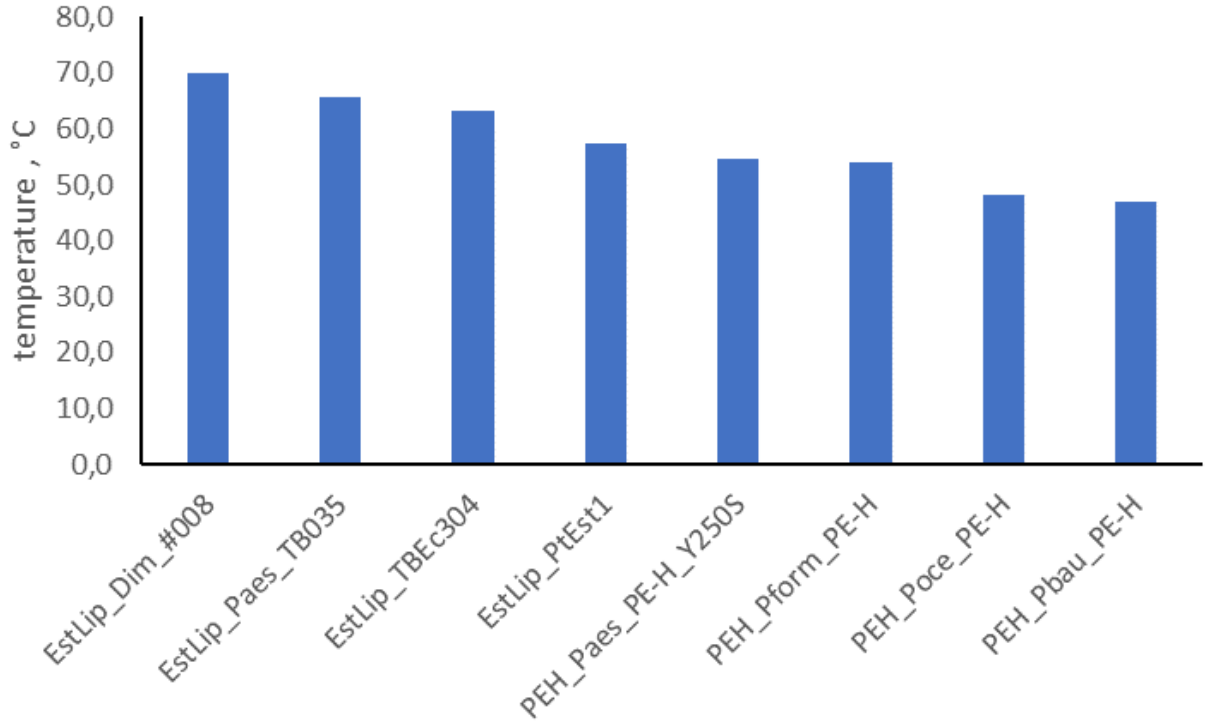


Figure 8. Tm and structural melting points of the selected enzymes EstLip_Dim_#008 (candidate 5), EstLip_Paes_TB035 (candidate 6), EstLip_PtEst1 (candidate 7), EstLip_TBEc304 (candidate 8), PEH_Paes_PE-H_Y250S (candidate 9), PEH_Pbau_PE-H (candidate 10), PEH_Pform_PE-H (candidate 11), PEH_Poce_PE-H (candidate 12).

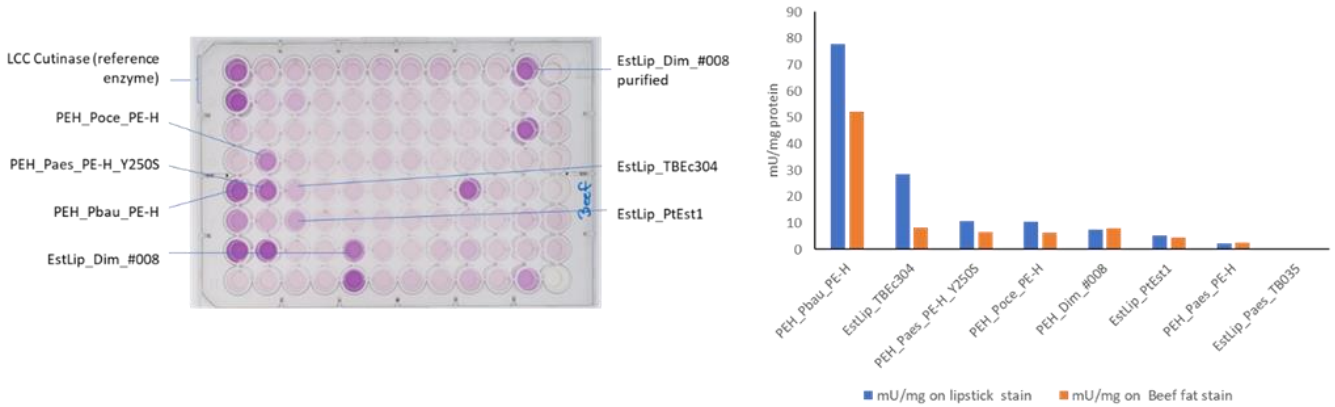


Figure 9. Activity of the selected esterases upon incubation with standard stained fabric material, determined by measuring the released fatty acids. Left: screening plate assaying the fatty acid concentration after overnight incubation with beef fat-stained fabric. The darker the violet, the higher the fatty acid concentration. Right: approximated specific activity of purified enzymes [1 U= 1 µmol fatty acid released per minute, determined with a serial dilution with oleic acid) after 2 h.

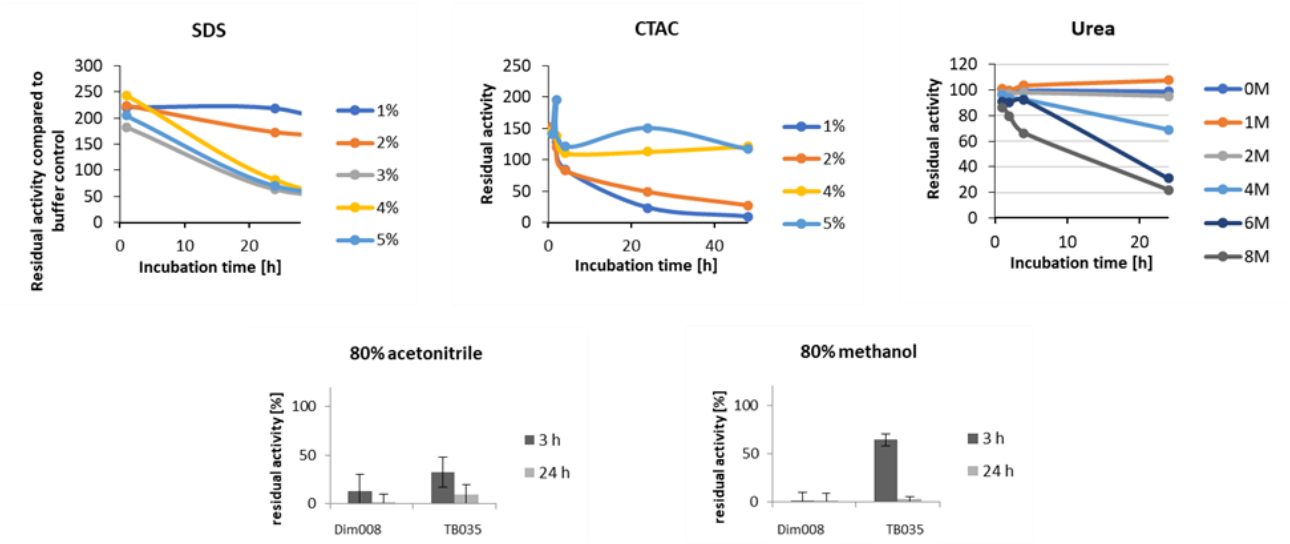


Figure 10. Resistance of EstLip_Paes_TB035 (candidate 6) towards different challenging chemicals.

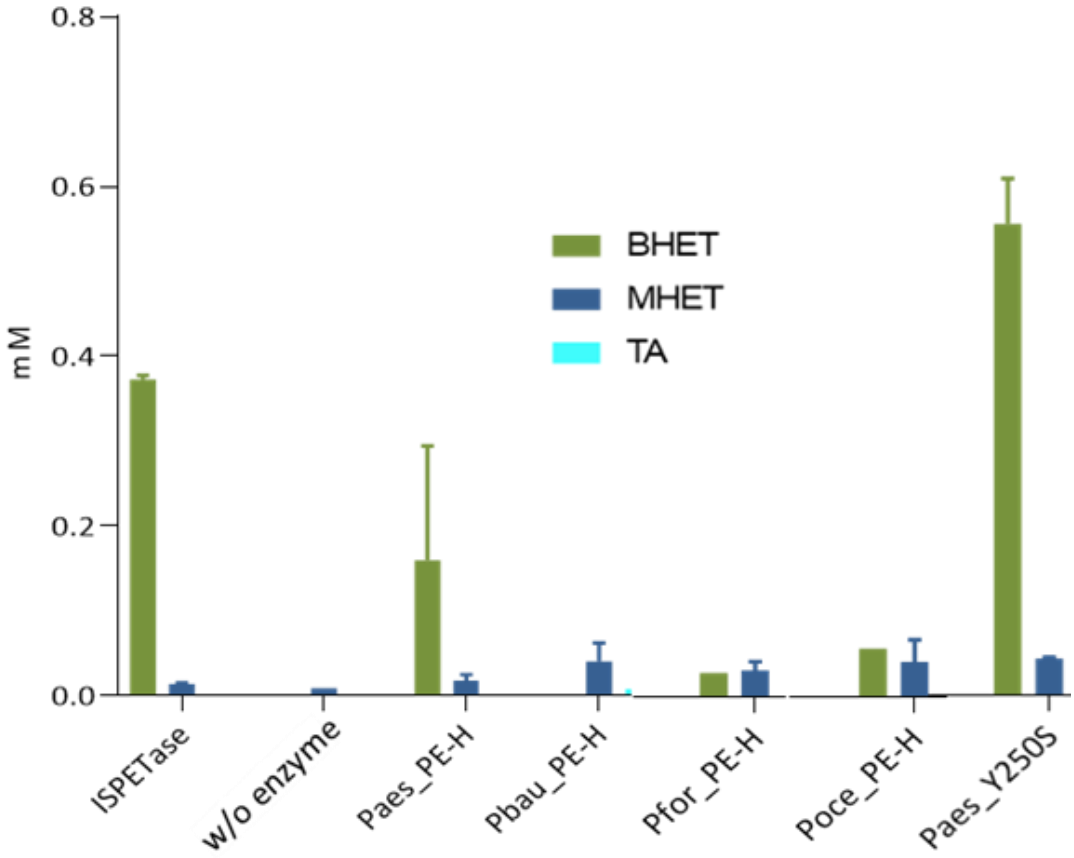


Figure 11. Release of PET degradation products from pretreated polyester fabric. ISPETase: PETase from *Ideonella sakaiensis* was used as a positive control.

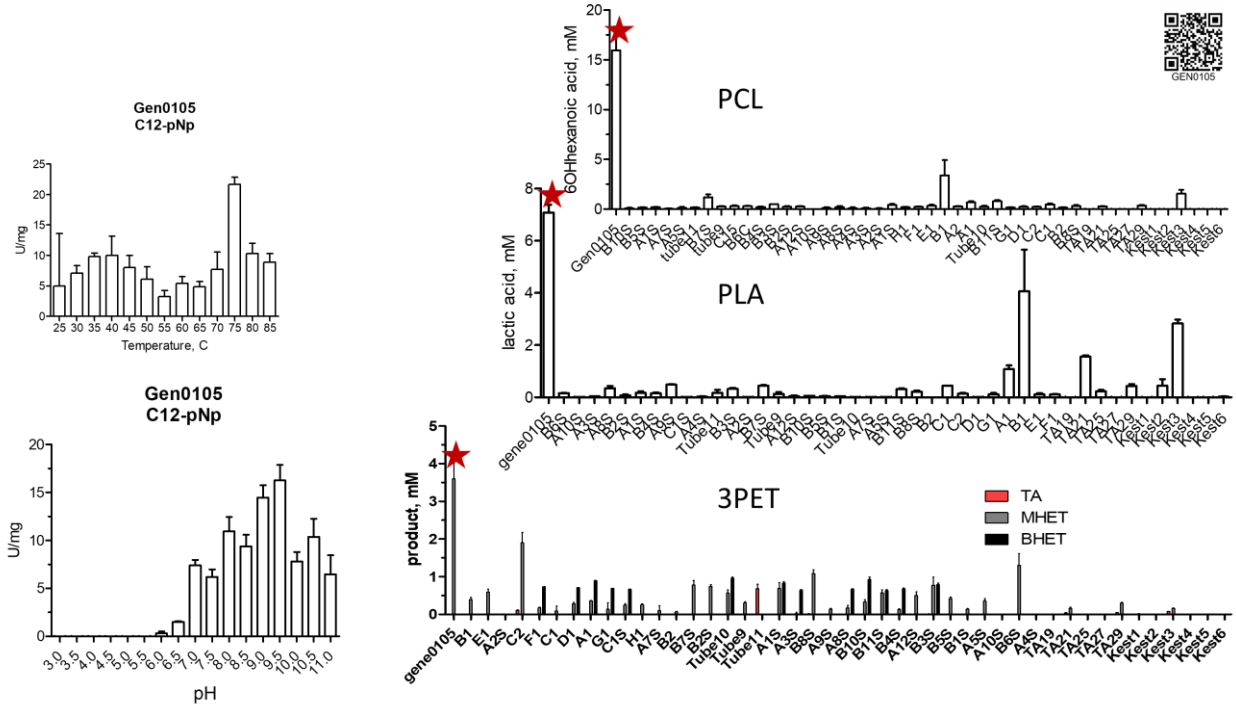


Figure 12. Characteristics of Gen0105 (candidate 13).

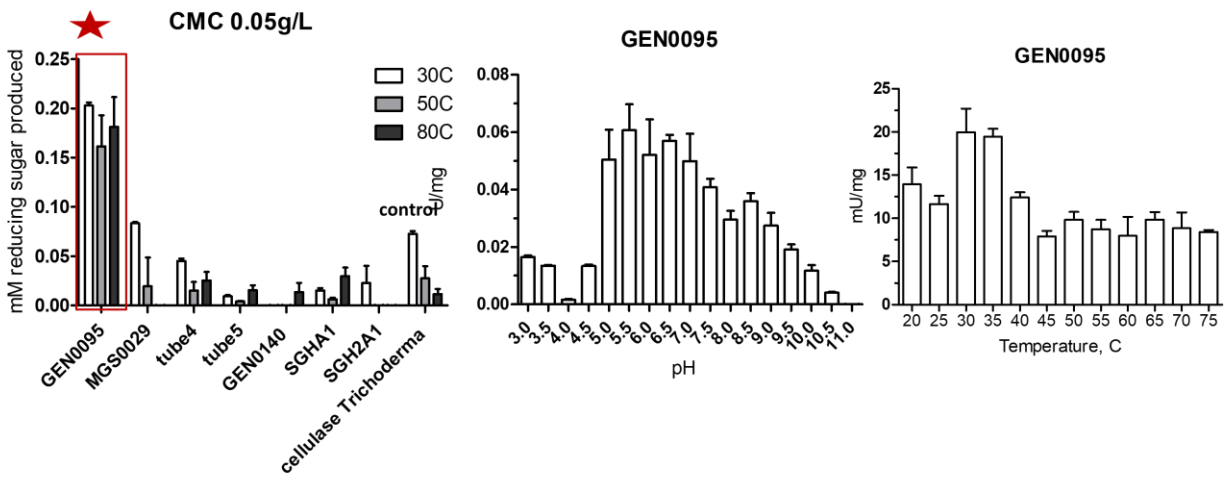


Figure 13. Characteristics of GEN0095 (candidate 14). PCL: PLA: CMC: carboxymethyl cellulose.

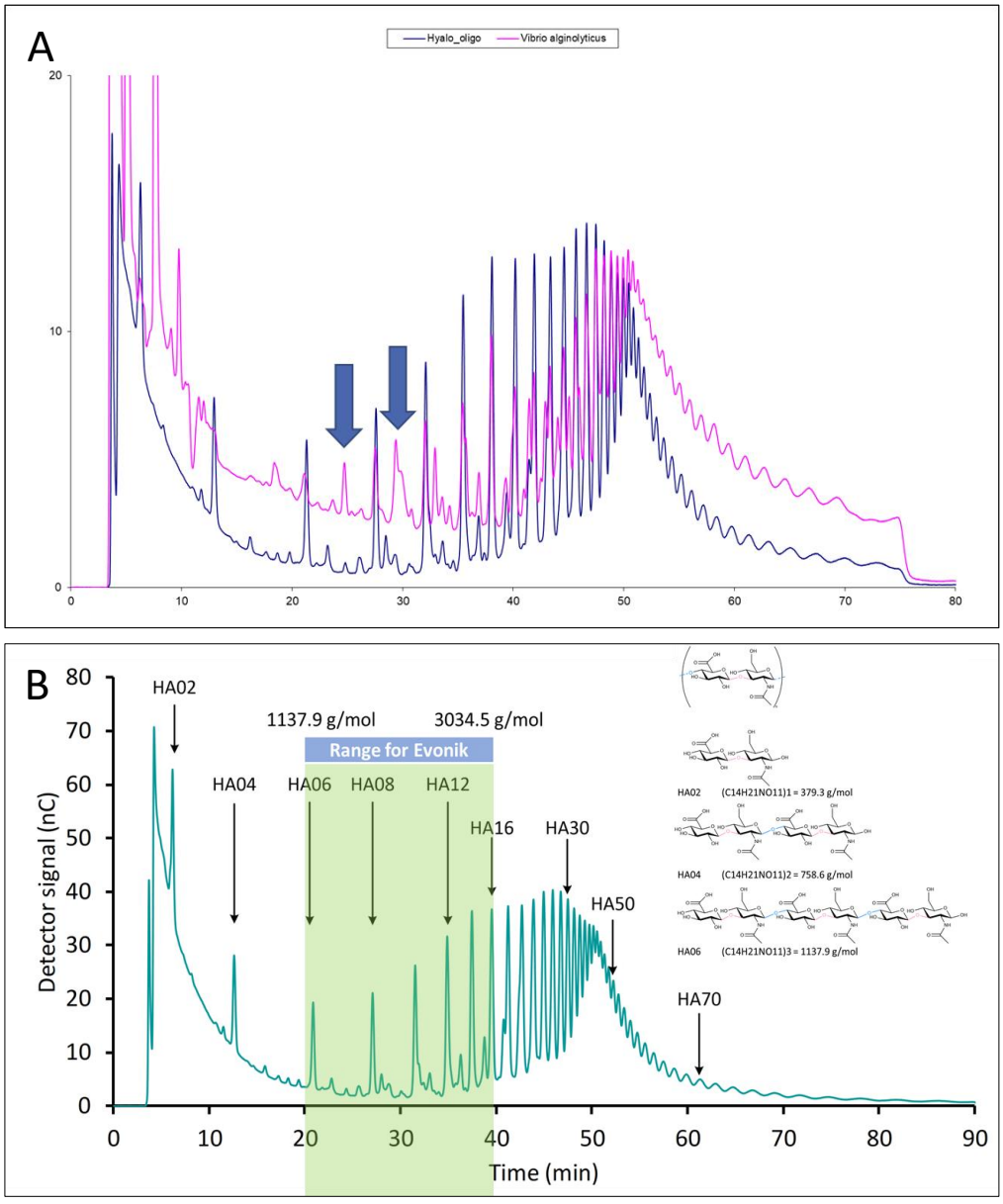
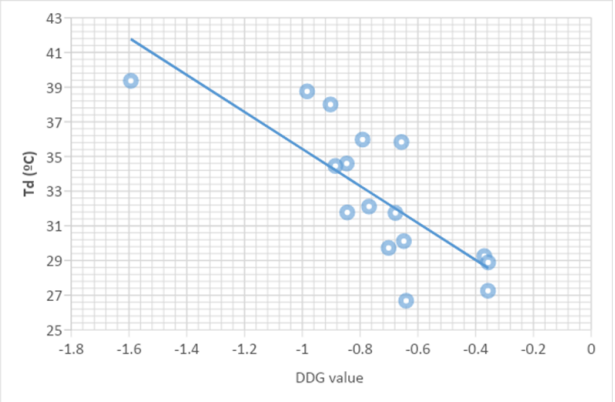


Figure 14. Degradation of hyaluronic acid by hyaluronidases of Vibrio strains (A) to produce degradation products of M_r 1-2 kDa (B).

- SCOT is a Random Forest based Machine Learning metapredictor that combines the estimations of 8 already published protein stability predictors and a molecular filter to produce a more reliable result

- Predictors: MAESTRO, CUPSAT, AUTOMUTE-SVM and AUTOMUTE-TR, FOLDX, INPS3D, MUPRO and I-MUTANT



Correlation between DDG value and Td

Figure 15. SCOT: Stability Consensus Metapredictor, using experimental data of EH37 mutants (candidate 13). DDG: delta delta G. Td: denaturing temperature.

6. Annex

Table 1. Detailed information on expression, activity and stability of the available enzymes. Shown are, among other datasets: 1) Enzymatic activity; 2) Name of the candidate; 3) Screen method; 4) Expression host; 5) Expression level; 6) Amino acid sequence or genome sequencing status; 7) Origin; 8) Details of stability features including denaturing temperature (Td), detergent stability; 9) Details of activity features, including substrate profile, optimal temperature and pH, etc.; 10) Sequence homology. Document available under the designation *D5.1_Annex 1_The shortlist of at least 18 enzymes nominated for engineering* at the FuturEnzyme web intranet through the following QR code (password needed), in the section *Shared data, Datasets*:

31