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

30/09/2022

THE FUTURENZYME PORTFOLIO OF 1000 ENZYME (RECOMBINANT/ NATIVE/BIOMIMETIC) MATERIAL, OBTAINED

DELIVERABLE NUMBER D4.2

MANUEL FERRER
CSIC
MARIE CURIE 2, 28049 MADRID, SPAIN

Document information sheet

Work package:	WP4, Small-scale enzyme production and characterisation.
Authors:	CSIC (Manuel Ferrer, Patricia Molina)
Document version:	1
Date:	30/09/2022
Starting date:	01/06/2021
Duration:	40 months
Lead beneficiary:	CSIC
Participant(s):	CSIC, UDUS, UHAM, CNR, IST-ID, Bangor, EUC, BioC-CheM
Dissemination Level:	Confidential, only for members of the consortium (including the Commission Services)
Type	Other
Due date (months)	16
Contact details:	Manuel Ferrer, mferrer@icp.csic.es; Patricia Molina, patricia.molina@icp.csic.es

Summary

The FuturEnzyme portfolio of 1000 enzyme (recombinant/ native/biomimetic) material, obtained

1. Scope of Deliverable

This deliverable consists in a set of protein samples of all 1000 enzymatic materials whose expression, preparation and production were undertaken by members of the consortium in a variety of vectors, hosts (heterologous or native) and cell-free systems. These materials were produced by, and exchanged/transferred to, partners for in deep characterization and analysis. To achieve this goal, appropriated SDS-PAGE analyses were performed to prove the expression levels and protein yields as well as the purity of the proteins after purification, the details of expression systems, production methods and purification methods.

2. Origin of the material

Along the already 16 months of project, different deliverables have been accomplished from which the present one nourishes. To be mentioned:

- D2.2: Set of 250,000 sequences pre-selected (November 2021)
  In this deliverable, a set of 250,000 sequences were pre-selected (D2.2, November 2021) were expected. This number was exceeded, as from a pool of 670 million sequences, a DIAMOND search disclosed 3.15 million sequences with interest for our project, which grouped in 457 different clusters as revealed by network analysis.
- D3.3: Set of 100 clones, 10 isolates, 10 enzymes shortlisted for sequencing, transfer to WP2 (March 2022)
  To accomplish this deliverable, we carried out bio-prospecting of the bio-resources from previous FP7, H2020 and EraNet funded projects, as well as internal sequence databases search. For this quest, the protocols detailed in deliverable "D3.2_Standard assays, analytics and calculations for monitoring enzymatic performance", were applied. Starting from 120 genomes from isolates, metagenomes from 47 microbial communities, 1200 microbial strains, 30 metagenome libraries and 500 enzymes, the outcome was 155 prospects (up to 163, noticing that some isolates present different activities, so most probably several enzymes) with interesting performances for FuturEnzyme, covering all the wanted classes of enzymes (except PluriZymes, to be designed by computational techniques).
- D2.3: Set of 1,000 enzymes selected using motif screens (May 2022)
  The aim of this deliverable was to select prospects for our objectives with bioinformatic tools from the 3.15 million sequences obtained in D2.2. A first filtering step comprised the following criteria: proper domain, catalytic residues, whether it was patented, and conservation (along with MSA). A set of 108 sequences were selected to be further analysed as follows: the 3D structure of the remaining sequences was modeled using AlphaFold 2.0, and substrates specified by the manufacturers were docked using Glide software (Schrödinger company) in the active site of these enzymes; the substrate positioning around the active site was further explored with PELE (Protein Energy Landscape Exploration) software from BSC.
- D2.4: Set of 180 enzymes for experimental focus (July 2022)
  From the previous deliverables, 205 enzymes and 54 microbes containing several enzymatic enzymes (which sums up to 304 candidates) were pointed out to be experimentally characterized. From D2.3 to the present D2.4, a number of candidates were added. Out of the 304 candidates, 258 enzymatic sequences are available, and the genomes of 13 microorganisms have been sequenced so far.

3. Methodology

The starting point of D4.2 has been the materials (isolates, enzymes) pre-selected and listed in D2.3 and D3.3. The analysis performed in the frame of the present deliverable revealed the list of 304 priority target to focus on listed in the deliverable D2.4. The partners implemented some strategies, that include the synthesis of codon-optimized genes in vectors (supporting His_6 fusion) and cloning strategies, as well as expression systems, validated for recombinant expression, available at the partners' laboratories (**Table 1**).

Table 1. List of expression vectors and hosts to be used in FuturEnzyme for enzyme expression.

Expression host	Host features	Expression vectors
Escherichia coli	Well-established expression host for non-food applications; used for molecular cloning and standard expression; Gram-negative; several strains available adapted to particular applications, e.g. BL21(DE3), BL22, BL21(DE3) Codon Plus-RIL, DH5α, Rosetta2, Rosetta-gami 2, JM109, TOP10, MC1061, T7-Shuffle, etc.	pET, pDrive, pCF-duet, pET-Duet, pCR-XL TOPO, pBBR, pVLT, pEBP, pASK, pEST, pPro, pBX, P15TV-L, pRhok, pRhot, pCEV-G4-Km, pMAL-p4x, pQE-80L series, etc.
Pseudomonas putida	Produces numerous co-factors; Gram-negative; versatile metabolism; high tolerance against many antibiotics and organic solvents	pBBR, pVLT, pEBP, pBBR22b, pBNTMCS, pBT'T, pJT'T, pJNT, and TREX/yTREX series
Bacillus subtilis	Well-established industrial expression host; Food-grade production host; Gram-positive; highly efficient secretion systems; examples: WB800N (No/low proteases)	pHT01, pBSMul, pEBP, pBX, and pBE-S series
Bacillus megaterium	Food grade production host; large-capacity versatile production (intracellular or extracellular); Gram positive; examples: WH320, YYBm1 (low protease background)	pHT01, pBSMul, pBX, pEBP, and pHIS1525 series
Burkholderia glumae	GC rich β-proteobacterium; production strain for an industrially applied lipase	pBBR, pVLT, and pBBR22b series
Rhodobacter capsulatus	Facultative anaerobic phototrophic Gram-negative bacterium; suitable for expression of O_2-sensitive and membrane proteins; intracytoplasmic membrane system; produces variety cofactors and possesses non-toxic LPS	pRhok, pRhon, and pRhot series
Haloferax volcanii	Obligate halophilic euryarchaeon suitable for intra-cytoplasmic and extracellular extremozymes	pTA941, pTA963, and pTA1228 series
Saccharomyces cerevisiae, Pichia pastoris	Highly efficient production and secretion systems	pPICZα, and pIB4 series
Thermus thermophilus	Derivatives of HB27 strain; Gram-negative; Growth between 55-80 ºC; GRAS; different (an)aerobic strains available; high transformation efficiency	pMK, pMH, pWUR, and pMS series
Aspergillus	Recommended for fungal laccases and peroxidases	*In house* vectors
Streptomyces	Recommended for fungal laccases and peroxidases	*In house* vectors

CSIC partner contributed to enzyme expression and production (see list in D3.3) using the following:

1. *E.coli* hosts for gene synthesis, cloning and expression : BL21, BL21 (DE3), Arctic Express (DE3)RIL, and MC1061.
2. Vectors for cloning and expression: pET-45b(+), pET-46 Ek/LIC, p15TV, pCC1FOS, pCC2FOS, pET21a, pET22b, pBXCH and pBXNH3. In many cases, the sequences of target enzymes were synthesized by GenScript Biotech (EG Rijswijk, Netherlands), which were codon-optimized to maximize expression in *E. coli*. The genes were flanked by BamHI and HindIII (stop codon) restriction sites and inserted in a pET-45b(+) expression vector with an ampicillin selection marker (GenScript, US), which was further introduced into *E. coli* BL21(DE3).
3. Cultivation media: LB (Luria-Bertani), TB (Terrific Broth) with 0.5% glycerol, SOC Broth medium (for the recovery of transformants of *E.coli*).
4. Cultivation volume: 0.1-5 L of LB or TB with appropriate antibiotics.

5. Protein yields: 0.1-100 mg/L. In all cases, the soluble N-terminal histidine (His) tagged protein was produced and purified at 4 °C after binding to a Ni-NTA His-Bind resin.

Bangor contributed to enzyme expression and production (see list in D3.3) using the following:

1. *E.coli* hosts for cloning and expression : BL21, BL21 (DE3), Arctic Express (DE3)RIL, DH5α, EPI300™-T1R , Stellar, NovaBlueGiga Singles, MC1061, Lobstr.
2. Vectors for cloning and expression: pET-46 Ek/LIC, p15TV-L, pCC1FOS/pCC2FOS, pTA1392, pCDFDuet, pCR-XL-TOPO,pET22b, pBAD-derivatives, pBXNH3.
3. Cultivation media: LB (Luria-Bertani), TB (Terrific Broth) with 0.5% glycerol, SOC Broth medium (for the recovery of transformants of *E.coli*).
4. Cultivation volume: 1 L of LB or TB with appropriate antibiotics.
5. Protein yields: 0.1-20 mg/L. In all cases, the soluble N-terminal histidine (His) tagged protein was produced and purified at 4 °C after binding to a Ni-NTA His-Bind resin.

UDUS contributed to enzyme expression and production (see list in D3.3) using the following:

1. *E. coli* hosts for cloning and expression: BL21 (DE3), Arctic Express (DE3), Lobster.
2. *Bacillus* for cloning and expression: *Bacillus subtilis* Marburg 168.
3. Vectors for cloning and expression: pET22b, pET21a, pCR-XL-TOPO, pBSMul1(SPBox). Enzymes were subcloned from bacterial genomes or metagenomic library clones.
4. Cultivation media: LB (Luria-Bertani), Lindstrom Autoinduction media.
5. Cultivation volume: 10 mL - 1 L of autoinduction medium with appropriate antibiotics, for *E. coli* cultivations usually 0.5 L.
6. Protein yields: 0.1-40 mg/L. In all cases, the soluble N-terminal-histidine (His) tagged protein was produced and purified after binding to a Ni-NTA His-Bind resin.

UHAM contributed to enzyme expression and production (see list in D3.3) using the following:

1. *E. coli* hosts for cloning and expression: BL21 (DE3), T7 SHuffle, DH5α, Rosetta gami 2 (DE3), and T7 express lysY/Iq.
2. Vectors for cloning and expression: pET21a(+), pET28a, pMAL-c2x and pDrive.
3. Cultivation media: LB (Luria-Bertani), TB (Terrific broth) and autoinduction medium (AI) ZYM.
4. Cultivation volume: 1 mL – 2L depending on the assay.
5. Protein yields: 0.1-10 mg/L. In all cases, the soluble N-terminal-histidine (His) tagged protein was produced and purified after binding to a Ni-NTA His-Bind resin.

CNR contributed to producing enzymatic materials from isolates (see list in D3.3), as follows:
CNR cultivated five types of prokaryotes as native enzymatic materials. They include:

1. Freshwater psychrophilic and psychrotolerant bacteria;
2. Marine gamma-proteobacteria;
3. Extremely halophilic haloarchaea;
4. Extremely alkali(halo)philic archaea;
5. Moderately halophilic natronobacteria.

For the preparation of enzyme material and all enzymatic tests (esterases, lipases, proteinases and hyaluronidases), freshwater psychrophilic and psychrotolerant bacteria were grown initially in 50 mL of Rich Enigma Medium (REM). Oligomineral water *Mangiatorella* was used as the basis for REM. Yeast extract (1g L^{-1}), Luria Broth (1g L^{-1}), and Na-pyruvate (1g L^{-1}) were added and the liquid medium was further buffered with 5 g L^{-1} of TAPSO {3-[N-tris(hydroxymethyl) methylamino]-2-hydroxypropanesulfonic acid} (pH 8.7). The formation of colonies was monitored in solidified REM obtained by adding of agar (15 g L^{-1}). Two different temperatures of cultivation were used: 2 °C and 20 °C.

For the preparation of enzyme material and all enzymatic tests (esterases, lipases, proteinases and hyaluronidases), marine gamma-proteobacteria were grown in an artificial seawater mineral salts medium (ONR7a) based on the ionic composition of seawater (*Int J Syst Bacteriol* 45: 116–123). This medium contained all of the major cations and anions that are present at concentrations greater than 1 mg L^{-1} in seawater. Nitrogen was provided in the form of NH_4Cl, and phosphorous was provided in the form of Na_2HPO_4. The ONR7a medium contained (per liter of distilled or deionized water) 22.79 g of NaCl, 11.18 g of $MgCl_2 \times 6H_2O$, 3.98 g of Na_2SO_4, 1.46 g of $CaCl_2 \times 2H_2O$, 1.3 g of TAPSO, 0.72 g of KCl, 0.27 g of NH_4Cl, 89 mg of $Na_2HPO_4 \times 7H_2O$, 83 mg of NaBr, 31 mg of $NaHCO_3$, 27 mg of H_3BO_3, 24 mg of $SrCl_2 \times 6H_2O$, 2.6 mg of NaF, and 2.0 mg of $FeCl_2 \times 4H_2O$. To prevent precipitation of ONR7a during autoclaving, three separate solutions were prepared and then mixed together after autoclaving when the solutions had cooled to at least 50 °C; one solution contained NaCl, Na_2SO_4, KCl, NaBr, $NaHCO_4$, H_3BO_3, NaF, NH_4Cl, Na_2HPO_4, and TAPSO (pH adjusted to 7.6 with NaOH), the second solution contained $MgCl_2$, $CaCl_2$, and $SrCl_2$ (divalent cation salts), and the third solution contained $FeCl_2$. For solid media, Bacto Agar (Difco) (15.0 g L^{-1}) or agarose (Sigma) (12.0 g L^{-1}) was added to the first solution. Lactate (5 g L^{-1}) and/or pyruvate (3 g L^{-1}) were used as carbon and energy sources.

For the preparation of enzyme material and all enzymatic tests (esterases, lipases, proteinases and hyaluronidases), extremely halophilic haloarchaea were grown in an '*Laguna Chitin*' mineral medium (LC medium). The LC medium was prepared on the basis of ONR7a medium (*Int J Syst Bacteriol* 45: 116–123), modified by the addition of (final concentration, g L^{-1}): 200 NaCl; 30 $MgCl_2$; 0.5 NH_4Cl and 5 HEPES (4-[2-hydroxyethyl]-1-piperazineethanesulfonic acid). The pH was adjusted to 7.0 by addition of 1 M KOH. D-(+)-Cellobiose (Sigma-Aldrich, catalog number C7252) was added at the final concentration 5 g L^{-1} to serve as growth substrate. After sterilization, the medium was supplemented with 1 ml L^{-1} acidic trace metal solution, 1 ml L^{-1} vitamin mix (*Arch Mikrobiol* 55, 245–256) and 50 mg L^{-1} yeast extract. The cultivation was performed at 40 °C in tightly closed 120 mL glass serum bottle with intensive shaking (100 rpm) in the dark.

For the preparation of enzyme material and all enzymatic tests (esterases, lipases, proteinases and hyaluronidases), extremely alkali(halo)philic haloarchaea (natronoarchaea) two mineral basic media (4M total Na^+) were used for enrichments and cultivation: a NaCl-base medium with a final pH 7.0 and a Na_2CO_3-base medium with a final pH 10.0. The former was similar as such for the enrichment and cultivation of neutrophilic extreme halophiles (LC medium), while for the natronoarchaea cultivation a 1:1 mix of the NaCl- and Na2CO3-base media (final pH 9.7) was used. The NaCl-base medium contained (g L^{-1}): NaCl, 240; KCl, 5; K_2HPO_4, 2.5; NH_4Cl, 0.5. The sodium carbonate- base medium contained (g L^{-1}): Na_2CO_3 190; $NaHCO_3$ 30; NaCl 16; KCl, 5.0; and K_2HPO_4 1.0. After autoclave sterilization, both base media were supplemented with 1 mM $MgCl_2$, 1 ml L^{-1} of acidic trace metal solution and 1 ml L^{-1} vitamin mix (*Arch Mikrobiol* 55, 245–256), 1 ml L^{-1} of alkaline Se/W solution (Pfennig &Lippert, 1966), and 100 mg L^{-1} of yeast extract.

For the preparation of enzyme material and all hyaluronidase enzymatic tests, moderately halophilic natronobacteria were enriched and, after isolation, grown for 7–28 days at 28 °C on sodium carbonate medium (SCM) containing 2 M total Na^+ with a pH 9.5 and included the following (g L^{-1}): Na_2CO_3 64, $NaHCO_3$ 48, NaCl 18, K_2HPO_4 1. The pH was adjusted to 9.5 by titration with 5M HCl. After sterilization at 120 ˚C for 30 min the medium was supplemented with 4 mM NH_4Cl, 1 mM $MgSO_4$, 50 mg L^{-1} of yeast extract, 1 mL each of vitamin mix and trace metal solution (*Arch Mikrobiol* 55, 245–256). High molecular weight Hyaluronic Acid (2 g L^{-1}) was added as the only source of carbon and energy.

All enzymatic tests, related with screening for esterase and lipase activities, were performed following FuturEnzyme protocols and methodology (assays). Substrates used were coconut, palm and olive oils. For protease activities calcium caseinate assay was used along with skim milk agar assay. Most active and performant isolates are ready to be cultivated upon request of FuturEnzyme partners. Hyaluronatolytic bacteria (*Vibrio* and *Paracoccus*), halo- (*Halomicrobium*, *Haloferax*, *Haloarcula*, *Halomicroarcula*, *Halorhabdus*, etc.) and natrono(halo)archaea (*Natrarchaeobius* sp. and *Natranaeroarchaeum aerophilum*)

were grown in liquid media (50 mL) and both the biomass and supernatant were collected for further testing of enzymatic efficiency and quality (size of obtained oligohyaluronates) of depolymerization.

IST contributed to producing enzymatic materials from isolates (see list in D3.3), as follows:
IST cultivated marine isolates on 100 mL erlenmeyers containing 20 mL of media, at 30 ˚C and 150 rpm. For high throughput screening tests, the isolates were grown on 24-deep well plates containing 4 mL of medium per well. The following media were used:

1. Marine broth (from Condalab).
2. Marine broth supplemented with olive oil (olive oil from Sigma-Aldrich).
3. Marine broth supplemented with hyaluronic acid (HA*; Hyacare from Evonik).
4. Marine broth supplemented with hyaluronic acid 50 (HA50; Hyacare 50 from Evonik).
5. Mineral medium supplemented with HA as carbon source.
6. Mineral medium supplemented with HA50 as carbon source.
7. Fang media.
8. Fang media without carbon source and with hyaluronic acid (HA as sole C source).
9. Fang media without carbon source and with hyaluronic acid 50 (HA50 as sole C source).

**HA, Hyaluronic Acid*

Cells were harvested during late exponential phase, washed with buffer, and used as whole cells in the biotransformation assays. Around 20-30 mg of cells were produced per cultivation in the erlenmeyers. Protocols described in deliverable D3.2 were used to assess lipase/esterase and hyaluronidase activities. Olive oil was used to favor the production of lipases/esterases, whilst HA and HA50 were used to favor the production of hyaluronidases.

Eucodis contributed to establishing expression systems, as follows:
Partner Eucodis improved current expression systems to be able to produce enzymes in the desired quantities for the downstream partners, through an approach comprising 2 parts: 1) Expression in *P. pastoris* through the design of integration plasmids for faster cloning, the establishment of secretion signal/pro-peptide library, and the fermentation optimization with improved plasmids; 2) Expression in *Corynebacterium glutamicum* through the design of integration plasmids for stable integration into genome, design and test of secretion signal peptide library, antibiotics-free expression for food/cosmetics grade enzymes. The improved *P. pastoris* protocols are already available and have been confirmed with current lipases. Compared to the established standard α-mating factor signal peptide, the new peptides doubled the secretion of selected lipases, which will contribute to make the future production of the consortium enzymes as efficient as possible and thereby enabling and supporting industrial applications. The improvement of the current *C. glutamicum* expression system is in progress and is expected to become available for the consortium enzymes in the next few months. Both expression systems, *P. pastoris* and *C. glutamicum* are GRAS hosts and well established for industrial production of enzymes, but also for food-grade production, which may become an important factor for the use of the consortium enzymes in cosmetics applications.

BioC-CheM contributed to establishing expression systems, as follows:
BioC-CheM Solutions dedicated effort to finding solutions for the production of selected enzymes and for the quantitative delivery of enzymes. The flowsheet of the tasks performed by BioC-CheM Solutions is schematically reported in **Figure 1**.

Figure 1. Schematic representation of BioC-CheM Solutions

Production of proteins from recombinant *P. pastoris* strains were performed by use of the BMGY medium with glycerol feeding (for strain designed to have constitutive expression) and methanol feeding (for strains designed to have an expression induced by methanol). Media LB or MHB were used for the production of proteins from recombinant *B. subtilis* and *E. coli*. Inducers were added based on the specific requirement of the expression system. Based on the levels of expression, cultivation of *P. pastoris, B. subtilis* and *E. coli* was performed in 20 L (Working Volume) Bioengineering or in 200 L (Working Volume) Sartorius Biostat fermenters. The yield of the process was mainly estimated based on the activity of the enzymes. However, the range in grams of pure proteins was in the range 5-30 grams for 20 L batches and 50-500 grams for 200 L batches.

For those "native" strains for which cloning was un-convenient (problems with the activity of the enzymes upon cloning in different hosts and/or sufficient expression already available in the native hosts), the selection of media was performed with the proprietary database BCSMedDat. The basics principles of the database are summarized in **Figure 2**.

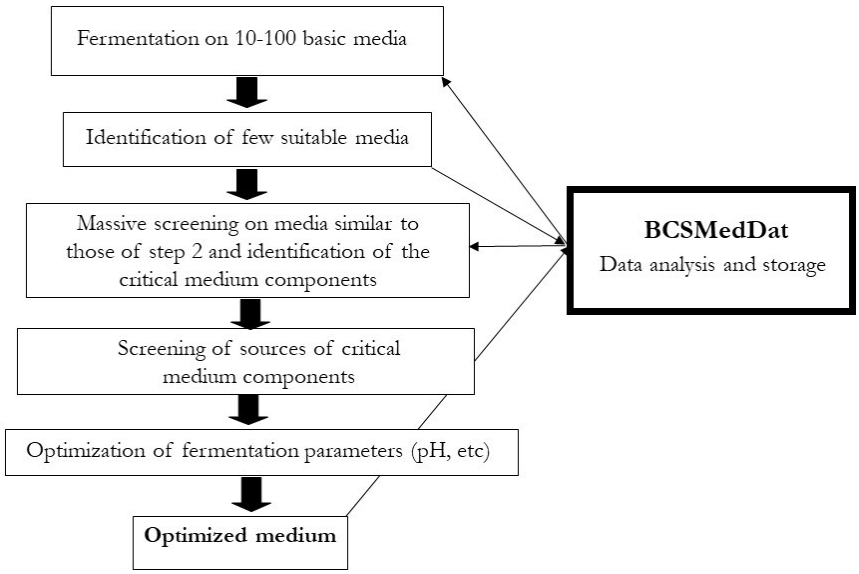

Figure 2. Schematic representation of the principles of the Database BCSMedDat

To date, media BCS340 (main carbon source glucose, glycerol and soybean oil, main nitrogen source soybean meal and peptone) and BCS321 (main carbon source glycerol, main nitrogen sources: peptone, corn steep liquor, soybean meal) were used for the production. The evaluation of the performance of the production process is in progress with best selected targets.

FHNW contributed to establishing expression systems, as follows:
FHNW developed and produced a novel class of artificial proteolytic enzymes exploiting the large capacity to produce esterases/lipases within the consortium to construct biomimetic proteolytic systems that can be produced in large quantities. It is based on a meticulous biochemical modification of the catalytic site of an esterase/lipase by a synthetic catalytic inhibitor (**Figure 3**). The method has been implemented using a bacterial transpeptidase as scaffold. A set of 7 esterase inhibitors has been synthetized and tested. Moreover, in this sense, esterase inhibition by mimicking Cys-protease is in progress (see details in deliverable D4.4); the objective is to produce an analogue, which can be accommodated in the active site of esterases. This is done via the design of peptides detailed in D4.4.

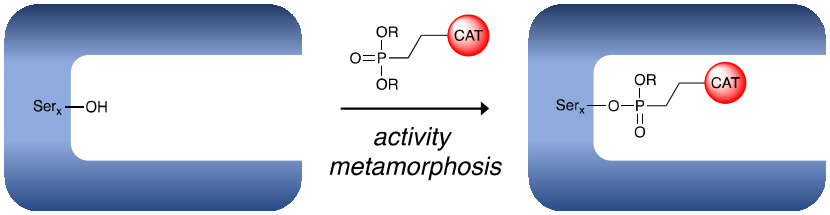

Figure 3. Schematic production of proteases from esterases

BSC & CSIC contributed to establishing expression systems for artificial proteases, as follows:
Proteases are abundant in prokaryotic genomes, about 10 per genome, but their recovery encounters expression problems, as only 1% can be produced at high levels; this value differs from that of similarly abundant esterases (1-15 per genome), 50% of which can be expressed at good levels. Partners BSC & CSIC designed a computational protocol by which catalytically efficient artificial protease that can be easily produced from esterases. The idea is based on locating in a native enzyme, through the Protein Energy Landscape Exploration (PELE) software, existing binding pockets where a target oligopeptide substrate could be accommodated and turning them into catalytic active sites by introducing all the residues needed for catalysis, particularly catalytic dyads, such as Cys-His.

PELE was used to find the non-catalytic peptide binding sites and check if catalytic poses can be reached in the functionalized variant. PELE is a Monte Carlo (MC) based algorithm coupled with protein structure prediction methods. The heuristic MC approach begins with the sampling of different microstates by initially applying small perturbations (translations and rotations) on the ligand. Then, the flexibility of the protein is taken into account by applying normal modes through the Anisotropic Network Model (ANM) approach. Once the system has been perturbed, side chains of the residues near the ligand are sampled with a library of rotamers to avoid steric clashes. Finally, a truncated Newton minimization with the OPLS2005 force field is performed, and the new microstate is accepted or rejected according to the Metropolis criterion. The Variable Dielectric Generalized Born Non-Polar (VDGBNP) implicit solvent was applied to mimic the influence of waters around the protein. The ΔΔG(mutWT) of stability in the experimentally tested variants was calculated using the module of thermodynamic stability from HotSpot Wizard, which uses FoldX to repair possible problems in the protein structure and Rosetta to perform the energy minimization and ΔΔG calculation (according to protocol 3 from Rosetta).

This protocol has been applied to one of the most active and well-produced esteasre obtained in the project, the esterase EH_{1AB1}. We proceed by preparing multiple dipeptides, namely, AH, AQ, DI, EA, FF, KA, LA, LL, NV, PF, QQ, RG, SW, TM, YN and YY, and performed a local PELE exploration for each of them. By doing so, we modeled the propensity to form catalytically active positions between each peptide bond and succeed in finding presumptive catalytic dyad (Cys24 and His214), that was further validated. By using this protocol, it is possible to design artificial proteases, starting from any native esterase, with good production yields (10-100 mg per liter of culture) compared to difficult to express natural proteases.

4. Results

Cell cultures, derived from enzyme-producing strains, have been obtained following optimization of the cultivation and expression conditions appropriated for each host and plasmid using semi-automatic devices (e.g. 96 and 48-well parrallel microfermentation devices and stirred Mini-Bioreactor Systems 6 x 0.25 L, etc.), or classical flask cultivation approaches, etc. Results of the expression tests (soluble protein, poor solubility, or lack of expression), for examples by analysing SDS-PAGE gels (see QR code in **Figure 4**) helped us to define which enzyme was produced in soluble form, to further perform activity tests, as well as to define whether alternative expression systems are needed. In addition, the enzymes to focus on, listed in deliverable D2.4, were retrieved.

Figure 4. QR code, relative to D4.2. Representative SDS-PAGE gels for key enzymes produced