Project presentation

Project summary:

The aim of the thesis is to explore the importance of proteolysis in the adaptation of microorganisms to deep environments. Genomic data from LM2E and biochemical studies carried out at IBS (Grenoble) indicate that in environments with very little organic matter, certain families of intracellular peptidases form large enzyme complexes, are more diversified and function more efficiently by combining specific activity spectra (families M42 and M18 according to the MEROPS classification). The exploration of (meta)genomes from marine environments and the enzymatic and structural characterisation of candidate proteins will allow us to test this hypothesis. this project also has biotechnological potential that will be developed with LEMAR. Indeed, these enzymatic complexes could be used to generate hydrolysates with probiotic properties from biomass for use in the agri-food industry. Knowledge of the structural motifs associated with the oligomeric edifices specific to M42/M18 peptidases will make it possible to identify by bioinformatics analysis the genes of interest in the LM2E resource (meta)genomes. Particular attention will be paid to (meta)genomes from deep sediments. The corresponding proteins will be expressed in *E. coli* and purified at LM2E and IBS. Peptide libraries will be screened to identify, via the analysis of HPLC chromatographic profiles, the natural substrates of these enzymes. These substrates will be synthesized and used to identify, under extreme conditions of temperature, salts, pH and pressure, metallic co-factors and optimal enzymatic activation conditions. In parallel, the structure of these enzymes will be determined using an X-ray crystallography approach. This will enable i) a better understanding of the functioning, role and evolutionary history of these enzyme complexes in extremophiles ii) the design of engineering strategies to combine different active sites in the same macromolecular assembly and iii) testing, in partnership with LEMAR, the effect of these enzymes to functionalise hydrolysates from various biomass sources.

Detailed presentation of the project :

1 - Hypothesis and questions addressed, state of the art, identification of scientific bottlenecks

Analysis of the genomes of certain deep-sea and abyssal marine microorganisms reveals a marked enrichment in genes coding for putative peptidases. Most of these enzymes belong to the metallopeptidase families M18 and M42 (MEROPS classification). Structural characterisations of several types of M42 peptidases from hyperthermophilic archaea show that some of these enzymes are capable of self-assembling to form large macromolecular complexes (Appolaire *et al.*, 2015 ; figure 1).

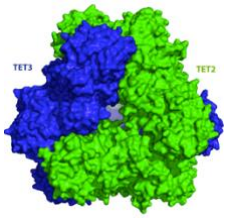


Figure 1. 500 kDa molecular assembly structure from the archaeon P. horikoshii. The complex is build up from 6 leucyl aminopeptidase subunits and 2 Lysyl aminopeptidases (In blue and green respectively) (Appolaire et al., 2014).

Biophysical and enzymatic experiments have deciphered the precise mechanisms that control their assembly and activation and have established the key role of various metallic co-factors in substrate assembly and specificity (Colombo *et al.*, 2016). In parallel, we have shown that these enzymes possess very narrow substrate specificities and that they assemble to form super

enzyme complexes with enhanced hydrolysis properties (Appolaire *et al.*, 2013). Finally, the study of optimal conditions of stability and activity with respect to temperature, pressure, salinity and metals also shows novel characteristics (Basbous *et al.*, 2018). These distinctive properties indicate that the diversification of M42/M18 peptidases corresponds to an environmental adaptation in microorganisms living in environments with a low organic matter content or in hyperthermophilic organisms that are required to survive when chemosynthesis conditions are no longer possible. The hypothesis is that, in these organisms, peptidases have diversified from ancestral genes or via horizontal transfers to optimise the cleavage activities of polypeptides, thus allowing more efficient degradation of peptides or an improvement in the function of intracellular proteolysis.

The aim of the thesis is to verify this hypothesis by carrying out an extended study of the functional and structural properties of different giant peptidases representative of the extremophilic biodiversity of deep marine environments. The main milestones are 1) identifying new peptidases via a bioinformatics strategy, 2) developing purification protocols, 3) identifying substrate peptides, the type of enzymatic activity and activation parameters and 4) determining the structure of these enzymes. This strategy has already been used successfully in the past (2 Phd, 5 new enzymes, 7 publications, 3 patents).

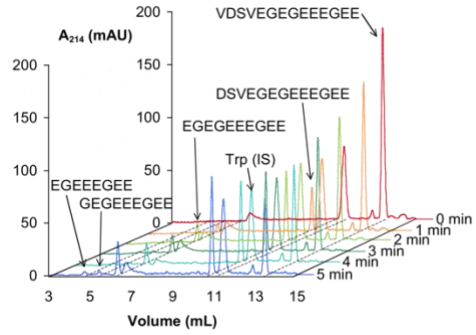
The detailed knowledge of the mechanisms of M42 peptidase assembly allows us to define consensus patterns that we use to identify genes coding for unknown giant peptidases in environmental (meta)genomes. Bioinformatics work extended to abyssal microbes, mainly sedimentary, will make it possible to specify the link between the multiplicity of peptidases and adaptation to extreme or particular metabolic conditions. Based on this work, a set of about ten enzymes will be selected on the basis of criteria of sequences, ecosystems of origin and of multiplicity within an organism, etc. Three enzymes of interest have already been targeted and cloned based on phylogenomic analysis in archaea, which already guarantees the availability of starting material for the PhD programme. The enzymes will be expressed in recombinant form and experimental work will be necessary to determine their substrates, activation conditions and stability. At the same time, an attempt will be made to determine the crystallographic structure in order to better understand their mode of action.

This thesis project also includes an important biotechnological component. Depending on their properties, certain enzymes could be patented for industrial applications and improved by molecular engineering processes, for example to produce multi-catalytic enzyme complexes. The new extremophilic peptidases will also be tested on the platforms of the LEMAR laboratory, which is developing innovative enzymatic hydrolysis processes on biomass with the aim of generating probiotic properties sought after by the aquaculture sector in the Brittany region.

2 - Methodological approach and techniques :

The identification of as yet unknown oligomeric peptidases in the (meta)genomes of bacteria and extremophilic archaea from deep marine environments will be carried out by bioinformatics analysis on the basis of specific motifs identified on the basis of comparative structural analyses. A phylogenetic study will also help us to elucidate the origin of assembly domains and the emergence of the oligomerisation process in these giant peptidases. We will thus determine the duplication events of the subunits as well as the horizontal losses and transfers that led to the development of multienzymatic systems with complementary substrate specificities, versus the maintenance of more "generalist" ancestral peptidases in other types of microorganisms. A set of about ten enzymes will be

selected on the basis of criteria of originality of sequences, original ecosystems or the number of versions in the genomes. The proteins of interest will be produced in *E.coli* from synthetic genes. Peptidases will be purified using protocols already tested at LM2E. The substrate specificity and mode of action of the new peptidases will be determined experimentally by an *ab initio* approach using synthetic peptide libraries, model hydrolysates and hydrolysates produced by LEMAR for food processing applications. The peptide profiles obtained after digestion will be analysed by reverse phase HPLC chromatography (C2/C18) (figure above). Within the more complex hydrolysates, the peptides of interest will be identified after size exclusion HPLC purification with the help of a LEMAR proteomics IR specialist for LC-MS-MS analysis. Substrate peptides will be synthesized by services provider (Montpellier) and will be used to determine the activation parameters of the new enzymes: metallic co-factors, optimal conditions of temperature, pH, pressure, salt, etc. The mode of action of the new enzymes (amino or carboxy-peptidase, endo-peptidases, di- or tri-peptidyl peptidases etc) will be deduced from the kinetic analyses of the reaction products. This work is based on spectrometry and fluorometry measurements which will be carried out under extreme conditions of temperature, pressure and pH. The oligomeric state of the peptidases will be characterised by biophysical studies (Sec-MALS, AUC, SAXS) and attempts will be made to determine the 3D structures of the new enzymes by X-ray crystallography. For this we will use protocols that have been proven at IBS. The determination of the structures will be done by experts from the ELMA group at IBS and the student will benefit from training in crystallogenesis and resolution of crystallographic structures. This work will allow to understand the mode of action of new enzymes and to initiate engineering work to design multicatalytic assemblies.



This thesis programme will take place :

- in the LM2E laboratory for bioinformatics analysis, selection of candidate proteases, expression of recombinant proteins and preliminary characterisation of preferential substrates (first 12-16 months).

- in the ELMA group of the IBS for the detailled enzymatic characterization on substrate banks and hydrolysates, the study of the peptide profiles obtained and the structural approach (last part of the PhD).

3 - Positioning and scientific environment in the regional, national and international context :

This project is in line with the institutional objectives of UBO and Ifremer and the site policy coordinated by IsBlue with regard to these objectives concerning the study of life forms that inhabit the ocean/lithosphere interface that give us the opportunity to better understand the development of life on earth, certain adaptive mechanisms, including those of adaptation to high temperature and pressure. Furthermore, it is also in line with the aim to amplify transfer actions within the biotech axis of the IUEM and the Ifremer innovation department and of course with the research and innovation strategy of the Brittany region in the framework of the development of biotechnologies and marine bioresources (DIS1). In terms of research themes, this programme also joins national and international initiatives to promote the study of the deep ocean, namely the Carnot MERS institute, the Ocean research priority programme which cites deep environments as priority ecosystems and the decade of

ocean sciences for sustainable development, particularly objective 14.

4 - Scientific and partnership context: general elements (ERC, CPER, ERDF, Breizhcop ...)

This project will be able to build on the infrastructure of the interreg NWE ALG-AD project coordinated by LEMAR, whose objective is to propose new treatments for nutrient-rich digestate with the aim of producing biomass that can be used for animal nutrition. These new treatments are partly based on the anaerobic digestion of food and agricultural waste, so that new proteases with original substrate specificities can be directly tested in order to assess their suitability for the valorisation scheme of this European project.

Appolaire, A., M. Colombo, H. Basbous, F. Gabel, E. Girard & B. Franzetti, (2015) TET peptidases: A family of tetrahedral complexes conserved in prokaryotes. *Biochimie*.

Appolaire, A., M.A. Dura, M. Ferruit, J.P. Andrieu, A. Godfroy, S. Gribaldo & B. Franzetti, (2014) The TET2 and TET3 aminopeptidases from Pyrococcus horikoshii form a hetero-subunit peptidasome with enhanced peptide destruction properties. *Mol Microbiol* **94**: 803-814.

Appolaire, A., E. Rosenbaum, M.A. Dura, M. Colombo, V. Marty, M.N. Savoye, A. Godfroy, G. Schoehn, E. Girard, F. Gabel & B. Franzetti, (2013) Pyrococcus horikoshii TET2 peptidase assembling process and associated functional regulation. *J Biol Chem* **288**: 22542-22554.

Basbous, H., A. Appolaire, E. Girard & B. Franzetti, (2018) Characterization of a glycyl-specific TET aminopeptidase complex from Pyrococcus horikoshii. *J Bacteriol*.

Colombo, M., E. Girard & B. Franzetti, (2016) Tuned by metals: the TET peptidase activity is controlled by 3 metal binding sites. *Scientific reports* **6**: 20876.