

INNOV'SAR: A NEW APPROACH FOR PROTEIN ENGINEERING AND SCREENING

Rudy PANDJAITAN, PEACCEL, Protein Engineering ACCEerator, n°6 square Albin Cachot, France.
rudy.pandjaitan@peaccel.com

Nicolas Fontaine, PEACCEL, Protein Engineering ACCEerator, n°6 square Albin Cachot, France.
Matthieu Ng Fuk Chong, PEACCEL, Protein Engineering ACCEerator, n°6 square Albin Cachot, France.
Frederic Cadet, PEACCEL, Protein Engineering ACCEerator, n°6 square Albin Cachot, France.

Key Words: Enzyme engineering, Biocatalyst, Chemical Manufacturing, Artificial Intelligence

We present a strategy that combines wet-lab experimentation and computational protein design for engineering polypeptide chains. The protein sequences were numerically coded and then processed using Fourier Transform (FT). Fourier coefficients were used to calculate the energy spectra called "protein spectrum". We use the protein spectrum to model the biological activity/fitness of protein from sequence data. We assume that the protein fitness (catalytic efficacy, thermostability, binding affinity, aggregation, stability...) is not purely local, but globally distributed over the linear sequence of the protein. Our patented method does not require protein 3D structure information and find patterns that correlate with changes in protein activity (or fitness) upon amino acids residue substitutions. A minimal wet lab data set sampled from mutation libraries (single or multiple points mutations) were used as learning data sets in heuristic approaches that were applied to build predictive models. We show the performance of the approach on designed libraries for different examples¹ and discuss how our approach can tackle epistatic phenomena². We can screen up to 1 billion (10^9) protein variants in a very short time.

1. F. Cadet, N. Fontaine, I. Vetrivel, M. Ng Fuk Chong, O. Savriama, X. Cadet, P. Charton (2018), Application of Fourier transform and proteochemometrics principles to protein engineering, *BMC Bioinformatics* (2018) 19:382, doi.org/10.1186/s12859-018-2407-8, <https://rdcu.be/9qjp>

2. F. Cadet, N. Fontaine, G. Li, J. Sanchis, M. Ng Fuk Chong, R. Pandjaitan, I. Vetrivel¹, B. Offmann & M. T. Reetz (2018), A machine learning approach for reliable prediction of amino acid interactions and its application in the directed evolution of enantioselective enzymes, *Scientific Reports* 8(1), 16757, doi: 10.1038/s41598-018-35033-y. www.nature.com/articles/s41598-018-35033-y