**Unveiling the molecular determinants responsible for NAD(P)(H) cofactor specificity using enzyme structural information**

**Tiago Resende**[1], **Isabel Rocha**[1], **Cláudio M. Soares**[2]

[1] *Centro de Engenharia Biológica, Universidade do Minho*
[2] *ITQB NOVA, Instituto de Tecnologia Química e Biológica Antonio Xavier, Universidade Nova de Lisboa*

tiagofcresende@gmail.com

In this post-genomic era, gene homology annotations have become the foundation of systems biology. However, errors spread easily when functional annotation is not performed carefully due to overly unconstrained homology metrics [1,2].

As metabolic model reconstructions become a relevant tool for performing fundamental studies and bioprocess design, the impact of accurate enzymatic function assignments becomes evident [3].

The uncertainty of the usage of NADP(H) or NAD(H) as co-factors [4], even in well performed annotations, has a major impact in metabolic engineering applications, severely affecting Genome-scale metabolic model reconstruction due to the potential insertion of misleading reactions.

In this work, we unveiled the molecular determinants for cofactor specificity, using enzyme structural information. In order to do so, we created a representative dataset of all enzymes present in the PDB with NAD(P)(H) as cofactors and measured the occurrence of every aminoacid residue at a distance of 3.5 Angstrom[5] of all cofactor atoms. This allowed us to create a matrix with the total number of aminoacid residues at interacting distances from all cofactor atoms, for all structures.

After analyzing the matrix, we identified the residues that had a significantly higher number of contacts with the cofactors, and in which atoms this occurred.

With the intent of applying these findings in the unveiling of cofactor specificity for enzymes that are not structurally characterized experimentally we successfully replicated these findings using machine learning algorithms.

In future work we will apply the results and machine learning models obtained in order to assign cofactor specificity to enzymes with uncertain cofactor specificity.

These results may represent an important development in systems biology by allowing the reduction of annotation errors and the implementation

of erroneous or redundant reactions in GEM models, improving the overall performance of metabolic simulations.

1Tian W, et al (2003) How Well is Enzyme Function Conserved as a Function of Pairwise Sequence Identity? J.Mol.Biol. 333,863 882.

2Dobson PD, et al (2005) Predicting Enzyme Class From Protein Structure Without Alignments. J. Mol. Biol. (2005) 345,187 199

3Ferrari L, et al (2012) EnzML: multi-label prediction of enzyme classes using InterPro signatures. BMC Bioinformatics 2012 13:61.

4King ZA, et al (2013) Optimizing Cofactor Specificity of Oxidoreductase Enzymes for the Generation of Microbial Production Strains - OptSwap. Industrial Biotechnology.2013,9(4): 236-246.

5Jeffrey, George A.; An introduction to hydrogen bonding, Oxford University Pres, 1997.