Laboratoire de Biochimie Théorique
Institut de **B**iologie **P**hysico-**C**himique
13, rue Pierre et Marie Curie
75005 PARIS

S E M I N A I R E

Karine Bastard

Laboratoire d'Analyses Bioinformatiques pour la Génomique et le Métabolisme, Genoscope, Institut de biologie François-Jacob, Direction de la Recherche Fondamentale, CEA

« Structural modeling and classification of active sites for guiding enzyme functional annotation »

The rate of enzyme functional characterization by experiments lags far behind the rate of gene sequence discovery, leading to an accumulation of proteins with no known function. Moreover, in public databases, function is extrapolated from a small number of proteins to all homologous members of a family resulting in 60% of superfamilies being mis-annotated [1]. Our institute has developed an integrated strategy based on *in-silico* prediction of enzymatic activities and *in-vitro* screening of enzymes for the discovery of various activities involved in microbial metabolism. As part of this strategy, we developed a structural bioinformatics method, called ASMC, for Active Sites Modeling and Clustering, which classifies proteins of a family into iso-functional sub-families and identifies functional amino acids responsible of specific enzymatic activities [2]. Experiments based on ASMC led to the unearthing of 14 potential new enzymatic activities for a family of unknown function, DUF849, and to the description of 3D-patterns for further annotation of sequences [3]. ASMC was also used to classify two phylogenetically unrelated protein families, MetX and MetA, for which we detected numerous mis-annotations in public databases. We re-examined nearly 10 000 MetA and MetX proteins using homology modeling and corrected the function for about 60% of them [4]. Our results show that the functional diversity within a protein family may be largely underestimated.

[1] Schnoes, A. M., Brown, S. D., Dodevski, I. & Babbitt, P. C. Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. PLoS Comput Biol 5, e1000605 (2009).
[2] de Melo-Minardi RC, Bastard K, Artiguenave F. Identification of subfamily-specific sites based on active sites modeling and clustering. Bioinformatics. 2010. 26(24):3075-82.
[3] Bastard K, Smith AA, Vergne-Vaxelaire C, Perret A, Zaparucha A, De Melo-Minardi R, Mariage A, Boutard M, Debard A, Lechaplais C, Pelle C, Pellouin V, Perchat N, Petit JL, Kreimeyer A, Medigue C, Weissenbach J, Artiguenave F, De Berardinis V, Vallenet D, Salanoubat M. Revealing the hidden functional diversity of an enzyme family. Nat Chem Biol. 2014. 10(1):42-9.
[4] Bastard K, Perret A, Mariage A, Bessonnet T, Pinet-Turpault A, Petit JL, Darii E, Bazire P, Vergne-Vaxelaire C, Brewee C, Debard A, Pellouin V, Besnard-Gonnet M, Artiguenave F, Médigue C, Vallenet D, Danchin A, Zaparucha A, Weissenbach J, Salanoubat M, de Berardinis V. Parallel evolution of non-homologous isofunctional enzymes in methionine biosynthesis. Nat Chem Biol. 2017. 13(8):858-866.

Jeudi 16 novembre 2017
14h00

Salle de conférence