Esterases Computational Study

Ruben Cañadas[#1], Gerard Santiago[#2], Victor Guallar[#*3]

[#]*Joint BSC-CRG-IRB Research Program in Computational Biology, Barcelona Supercomputing Center, 08034 Barcelona, Spain*

[1] ruben.canadas@bsc.es, [2] gerard.santiago@bsc.es

[*]*Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain*

[3] victor.guallar@bsc.es

***Keywords*— esterases, PELE, QM/MM, machine learning**

EXTENDED ABSTRACT

Esterases are a group of alpha/beta hydrolases enzymes that split an ester into an alcohol and a carboxylic acid. These types of enzymes have special interest in industry, being used in food processing, perfume industry, degradation of synthetic materials..

Their catalytic mechanism is well-known in the literature[1]. They have a catalytic triad typically formed by three amino acids: Serine, histidine and aspartic acid. In some cases we can find a glutamic acid instead of an aspartic acid. Many other hydrolases such as proteases use this mechanism, being on of the most conserved catalytic mechanism.

The histidine deprotonates the hydroxyl group of the serine. Then, this serine carries out a nucleophilic attack towards the ester. After that, a tetrahedral intermediate is formed followed by an electronic rearrangement, resulting in the formation of an alcohol. The next step is the entrance of a molecule of water. Thanks to this molecule, the second product of the reaction, a carboxylic acid, is formed.

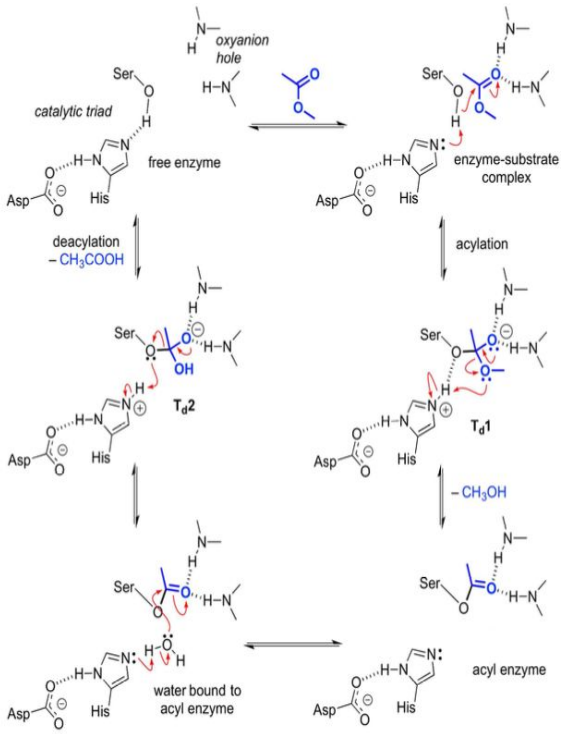

Figure 1. Canonical esterase mechanism for hydrolysis of methyl acetate, a typical ester

The Asp-His hydrogen bond is very important to maintain the catalytic triad in an active conformation. Using molecular dynamics, one can see that if this hydrogen bond is broken, this triad can adopt an inactive conformation [2]. Another important aspect of the catalytic mechanism is the oxyanion hole. This is formed by two backbone hydrogens which stabilize the negative charge of the oxygen when the tetrahedral intermediate is formed.

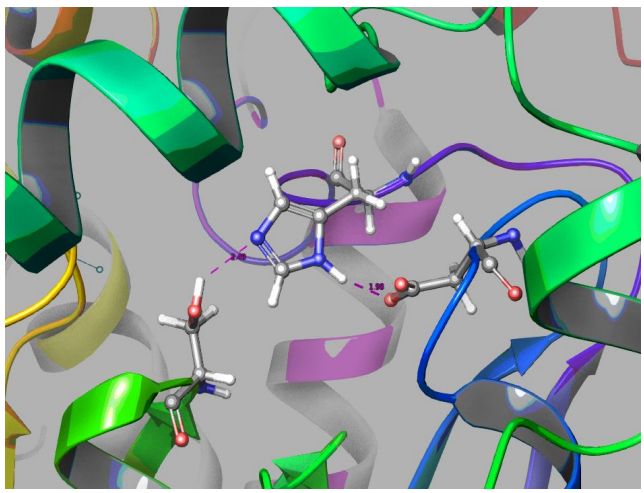

Figure 2. Typical esterase catalytic triad.

In our project, we are focusing on two different esterases called LAE6 and LAE5. LAE6 is very promiscuous since it presents activity towards a wide range of different substrates, whereas LAE5 presents activity towards few substrates. This promiscuity can be due to the location of the binding cavity: in LAE6 the binding cavity is located buried inside the protein, so the substrates can be retained longer, correlating with a higher reactive probability . LAE5 binding cavity is situated on the surface of the protein, in direct contact with the solvent, which, in our working hypothesis, is the main reason for its much less promiscuity.

We are using several computational techniques to extract important descriptors in order to construct a mathematical model to predict the activity of these two enzymes towards different substrates. We normally start performing a Monte-Carlo-Metropolis calculation using an in-house software named PELE (Protein Energy Landscape Calculation). This method uses a pseudorandom number

generator to perturb the ligand randomly. Then the protein structure is adapted (relaxed) to such perturbation using protein structure prediction techniques. Once the system has been relaxed, a Metropolis criterion is used to decide whether this structure will be accepted or not. The overall technique, along with the atomic force field and implicit solvent is often biased towards low energy conformations, so we can get stuck in some local minimum, preventing us from getting over high energy barriers and consequently visiting other local minima. Thus, we introduced a machine learning adaptive technique to improve sampling.
After running PELE simulations, we can extract the best poses according to catalytic distances and binding energy profiles. QM/MM optimizations might also be done to refine structures since PELE uses a minimization algorithm based on molecular mechanics. Once the structure is minimized , we are ready to extract several descriptors. For example catalytic distances, deprotonation angles, charge distribution, molecular weight… Other methods are being used, such as molecular dynamics or metadynamics to deduce more accurate descriptors.

All this descriptors will be used in the future, to create a model, using machine-learning techniques. Ideally, this model will be able to predict the activity of a wide range of substrates. Thus, a correlation between experimental activity and predicted activity needs to be found.

References

[1] A. Rauwerdink and R. J. Kazlauskas, "How the same core catalytic machinery catalyzes 17 different reactions: the serine-histidine-aspartate catalytic triad of alpha/beta hydrolase fold enzymes," ACS catal. 2015, 5, 6153-6176

[2] E. Y. Lau and T. C. Bruice, "Consequences of breaking the Asp-His hydrogen bond of the catalytic triad: effects on the structure and dynamics of the serine esterase cutinase," Biophysical Journal, Vol. 77, Pp. 85-98, July 1999.

Author biography

Ruben Cañadas is a bachelor's degree student at his fourth year who is currently carrying out his degree final project at the atomic and electronic protein modeling group at the BSC (Barcelona Supercomputing Center). He is focused on the study of two esterases with relevant importance for the industry.