Engineering a Haloacid-dehalogenase: a step closer to understand the most diverse enzyme family

By combination of computer analyses and molecular biology techniques, the active site of a haloacid dehalogenase (HAD) enzyme was unravelled. The HAD superfamily is comprised of the majority of regulatory enzymes in all organisms illustrating the significance of this superfamily.

Enzymes are a group of proteins that help chemical reactions in all living organisms to occur. The food we consume, for example, can only be converted into energy by a set of chemical reactions catalysed by dozens of enzymes. Enzymes in our body that play a role in our metabolism are largely controlled by chemical compounds that contain phosphate. Addition or breakdown of a phosphate to those compounds, for instance, decides whether our body needs to digest the food we eat into energy or to store it as fat. The breakdown of a phosphate from those compounds is also performed by a group of enzymes called phosphatases. Enzymes from HAD superfamily are the most numerous members of the phosphatases. Thus, knowledge of this superfamily is important to better understand our metabolism.

In this study, a HAD enzyme from *Thermatoga neopolitana* (HADTn) was used as a model. HADTn is a particularly interesting enzyme because it cleaves phosphate from fructose-6-phosphate (F6P) more efficiently than glucose-6-phosphate (G6P). Other HAD enzymes from *Escherichia coli*, the most studied organisms, do otherwise. Previous reports show that all enzymes from the HAD superfamily contain four regions that are important for the activity. Those four regions are conserved in the HAD enzymes found in all organisms. One study, however, proposes the presence of the one additional region in the HAD superfamily that has a role in the recognition of their substrates, e.g. F6P, G6P, etc. The significance of a region in the enzymes can be tested by substituting amino acids (the building blocks of all enzymes) in that region with different amino acids. Different activities will be observed if important amino acids are replaced.

Computer modelling was used in this work to determine the proposed fifth region in HADTn. Simulations by computer were also performed to identify other amino acids that might have a significant role in HADTn. A major obstacle to the computer analyses was a lack of previous studies reporting how the substrates interact with HADTn while several possibilities of interactions were generated from the analyses. To overcome the limitations of the computer predictions, a larger number of amino acids had to be studied. This approach then led to the second challenge, which was to develop a robust screening method to screen huge numbers of the variants produced. After having dealt with those two hurdles, screenings of hundreds of variants were carried out. Results indicated that substitutions of the amino acids located in the fifth region of HADTn increased the specificity toward F6P. These findings demonstrated the substantial role of the fifth region in HAD superfamily.

Identification of the fifth region is a pivotal step to utilise HAD enzymes for metabolism studies. With these results, scientists can simply alter the fifth region of a HAD enzyme to favour a reaction of a phosphate-containing-compound over others thus controlling the flows of the metabolisms. Manipulation of the metabolism's flows is a powerful approach to study the organisms. Cancer cells, for example, are also studied using this approach.