Daniel Kracher, Robert Peharz

Algorithms in the Test Tube: How AI Accelerates Enzyme Engineering

Many chemical reactions that are essential for a sustainable, circular economy are either difficult to carry out or place a considerable burden on the environment. Traditionally, the chemical industry has relied on high heat, high pressure and often toxic solvents. Enzymes offer an attractive alternative because they catalyse reactions under mild conditions, are biodegradable and could replace many energy-intensive industrial processes. However, the enzyme toolbox available in nature is often limited and not sufficient for challenging biochemical reactions.

To overcome these limitations, scientists increasingly rely on computational methods to guide the modification and improvement of enzymes. In the TU Graz Lead Project DigiBioTech, Daniel Kracher (Institute of Molecular Biotechnology) and Robert Peharz (Institute of Machine Learning and Neural Computation) are joining forces to address one of the most difficult challenges in enzyme catalysis: the cleavage of carbon–halogen bonds.

These bonds are widespread in anthropogenic compounds and are often extremely stable. As a result, many of these substances are degraded only very slowly in nature, if at all. They therefore accumulate in soil and water and can threaten ecosystems and human health. Among the best-known examples are per- and polyfluoroalkyl substances (PFAS), often referred to as “forever chemicals”. Carbon–halogen bonds are also common in pharmaceuticals, agrochemicals and speciality materials. Their high stability and useful chemical properties make them valuable in many industrial and consumer products, but are also very difficult to remove once they enter the environment.

Daniel Kracher is an enzymologist who has spent his career investigating redox enzymes. These enzymes can overcome high-energy barriers and can react with molecules that are otherwise inert. In nature, these enzymes play important roles in many biological processes, for example, in the breakdown of complex organic matter. Kracher and his team aim to understand how these enzymes work at the molecular level. This knowledge can then be used to adapt existing enzymes or develop new ones that perform useful reactions under industrially relevant conditions. However, enzyme engineering is a complex and time-consuming process. Even very small changes in a protein structure can cause the whole system to fail – or to perform substantially better. These changes often have unpredictable effects on its activity, stability and selectivity. This is where computational methods become important. Instead of testing thousands or hundreds of thousands of random enzyme variations experimentally, which is time-consuming and expensive, researchers increasingly use data-driven models that guide the search for improved designs like a compass.

Robert Peharz develops machine learning methods for this purpose. His expertise lies in probabilistic machine learning, a branch of AI that explicitly takes uncertainty into account. Using experimental data generated in Kracher’s group, AI models developed in Peharz’s group can estimate an enzyme’s “fitness landscape”, meaning how changes in its sequence affect its performance.

In this project, the team investigates a rather unusual enzyme found in an unexpected place: a haemoglobin from the marine organism Amphitrite ornata. We usually think of haemoglobin as the protein that carries oxygen in the blood. However, this organism has to survive in an environment that is rich in naturally occurring toxic compounds, secreted by competing species. In order to survive, it has developed its most abundant protein into a specialized