# Beyond Trial-and-Error: Artificial Intelligence for PHA Depolymerase Engineering

**Laura I. de Eugenio**, José Daniel Jiménez Santos-García and M. Auxiliadora Prieto
CIB-MS (CSIC), C/Ramiro de Maeztu, 9, Madrid, Spain

lidem@cib.csic.es

PHAs are biobased, biodegradable, and biocompatible polyesters composed of (*R*)-3-hydroxyacids (HAs) of variable chain length, typically classified as short-chain-length (scl) or medium-chain-length (mcl) polymers. Their chemical diversity allows tuning of material properties across a wide range, enabling PHA-based materials to exhibit thermoplastic or elastomeric behavior depending on monomer composition [1]. In addition, PHAs occupy a distinctive position among bioplastics because they can be completely synthesized and degraded by microorganisms, making them attractive candidates for genuinely circular material schemes. However, PHAs should not be considered materials intended simply to disappear through degradation. Regulatory frameworks for packaging increasingly promote controlled recovery pathways rather than uncontrolled environmental breakdown [2]. Moreover, PHA depolymerization can generate enantiomerically pure HAs that are valuable chemical intermediates and building blocks for biotechnological and chemical applications [3]. The enzymes responsible for PHA degradation are collectively known as PHA depolymerases, which can be classified as intracellular or extracellular depending on their biological roles, either in mobilizing intracellular storage polymers or in degrading environmental PHA. In addition, these enzymes display specificity toward scl or mcl-PHA substrates [4]. The apparent enzymatic performance of PHA depolymerases depends strongly on substrate preparation and assay format, and it is influenced by material properties such as crystallinity, crystal organization, surface morphology, and surface accessibility [4–6]. In this context, enzymatic depolymerization of PHA emerges as a particularly attractive strategy to recover value while maintaining circularity. Yet, the practical development of enzymatic recycling routes for PHA-based materials remains limited because many of the biochemical and physicochemical determinants governing depolymerase performance are still not fully understood.

In this work, we compare classical protein engineering strategies with emerging artificial intelligence (AI)-driven approaches to evaluate their potential for depolymerase optimization. Classical protein engineering approaches, such as structure-guided and random mutagenesis, enable the discovery of improved variants both in predicted functional regions and at unexpected positions, but they typically require labor-intensive experimental screening. Within this framework, site-directed mutagenesis studies on PHA depolymerases have enabled the identification of catalytic residues in several enzymes [7–8], and have revealed the functional importance of residues located at the interface between lid and core domains in intracellular mcl-PHA depolymerase from *Pseudomonas putida* KT2440, where substitutions can severely impair enzymatic activity. Random mutagenesis approaches applied to this enzyme have generated variants displaying increased activity toward soluble model substrates such as *p*-nitrophenyl esters, as well as toward polymeric substrates. Mapping of these mutations has identified positions contributing to improved activity and their structural implications have been evaluated through molecular dynamics (MD) simulations to assess their impact on local flexibility and conformational behavior [9]. In parallel, evolution-guided computational strategies, including ancestral sequence reconstruction, offer an additional route to explore sequence space beyond extant variants and recover enzymes with enhanced robustness or altered substrate preferences [10,11]. For example, ancestral reconstruction applied to the extracellular mcl-PHA depolymerase from *Pseudomonas solani* GK13 species has enabled the generation of variants displaying improved activity and thermostability. However, ancestral reconstruction remains constrained by the availability and diversity of evolutionary information and primarily samples sequence solutions that have already been explored by nature, leaving large regions of functional sequence space inaccessible.

In this scenario, artificial intelligence (AI) and machine learning (ML) offer the possibility of exploring sequence-function relationships beyond evolutionary constraints, enabling the prioritization of variants that would be difficult to identify through conventional experimental strategies. In other enzymatic systems, ML-guided strategies have shown that beneficial variants can be identified more efficiently and with reduced screening effort, and in polymer-degrading hydrolases this logic has already translated into successful engineering outcomes, most notably in PET hydrolases improved through machine learning-assisted design [12]. Importantly, PHA depolymerases present an additional layer of complexity, as their substrates naturally occur in multiple physical forms, from intracellular granules to extracellular particles and highly processed materials. This distinctive feature differentiates PHA depolymerases from many other polymer-degrading systems and introduces additional challenges when translating enzyme optimization strategies into practical applications. Against this background, the limited understanding of how enzyme variants behave across different material forms creates practical challenges for identifying and optimizing depolymerases capable of performing efficiently across both native PHA diversity and processed PHA materials. As a result, depolymerase performance cannot be understood solely from enzyme sequence and structure: it is also shaped by the physical state of the substrate. Therefore,

extending ML-driven protein engineering strategies to PHA depolymerases requires models that do not treat the enzyme as the only informative input, but instead integrate enzyme descriptors with descriptors of the material context. Such frameworks could combine sequence information, structural models, mutagenesis outcomes, and selected molecular dynamics (MD)-derived features with experimentally defined properties of the tested PHA substrate, such as monomer composition, crystallinity, morphology, preparation route, and surface state.

We therefore propose moving from enzyme-centered prediction to enzyme–substrate-state prediction, explicitly incorporating the physical state of PHA materials into AI-guided depolymerase engineering. Rather than replacing experimentation, AI could make experimental work more targeted, more interpretable, and more scalable by embedding the material state of the substrate directly into the engineering loop. This shift may be essential if depolymerases are to move from biological polymer turnover toward robust enzymatic recycling of real PHA-based plastic products.

References

[1] Park H, He H, Yan X, Liu X, Scrutton NS, Chen G-Q. Biotechnol Adv, 71 (2024) 108320.
[2] European Parliament and Council. Regulation (EU) 2025/40 on packaging and packaging waste. Official Journal of the European Union (2025).
[3] de Roo G, Kellerhals MB, Ren Q, Witholt B, Kessler B. Biotechnol Bioeng, 77 (2002) 717–722.
[4] Jendrossek D, Handrick R. Annu Rev Microbiol, 56 (2002) 403–432.
[5] Jendrossek D. Appl Microbiol Biotechnol, 74 (2007) 1186–1196.
[6] Iwata T, Doi Y, Nakayama S, Sasatsuki H, Teramachi S. Int J Biol Macromol, 25 (1999) 169–176.
[7] Braaz R, Handrick R, Jendrossek D. FEMS Microbiol Lett, 224 (2003) 107–112.
[8] García-Hidalgo J, Hormigo D, Arroyo M, de la Mata I. PLoS One, 8 (2013) e71699.
[9] de Eugenio LI, Jiménez JD, Ramos E, Serrano-Aguirre L, Sanz JM, Prieto MA. Appl Microbiol Biotechnol, 109 (2025) 215.
[10] Thornton JW. Nat Rev Genet, 5 (2004) 366–375.
[11] Gumulya Y, Gillam EMJ. Biochem J, 474 (2017) 1–19.
[12] Lu H, Diaz DJ, Czarnecki NJ, et al. Nature, 604 (2022) 662–667.