REVIEW SUMMARY

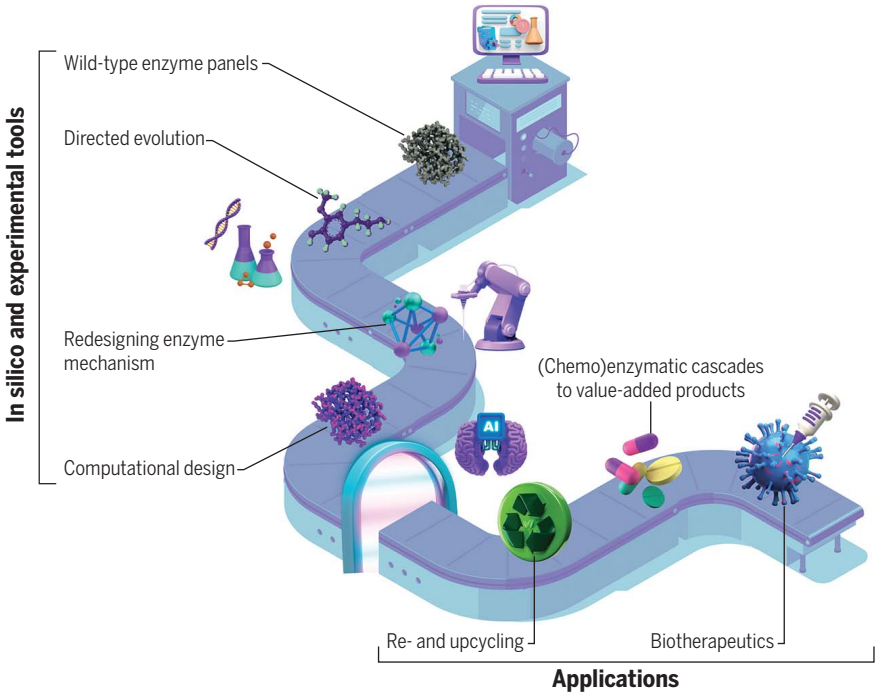
BIOCATALYSIS

From nature to industry: Harnessing enzymes for biocatalysis

R. Buller, S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore, U. T. Bornscheuer*

BACKGROUND: Biocatalysis is an approach to synthetic chemistry in which enzymes carry out chemical reactions. Historically, enzymes from natural sources have been used to break down oils and proteins in laundry detergents, produce semisynthetic antibiotics, and create simple chiral precursors for the pharmaceutical industry. In the past 5 years, the number of available protein sequences has increased by a staggering 20-fold, accelerating the discovery of enzymes with useful activities and properties. Directed evolution, the cornerstone of our ability to tailor enzymatic properties, is allowing researchers to tailor enzymes for the synthesis of complex molecules, the modification of biological therapeutics, and the breakdown of plastic waste. Machine learning–driven protein-structure prediction, coupled with advances in automation and high-throughput screening, is further advancing our ability to create enzymes with desired function. The ability to add non-biological catalytic elements to enzymes means that enzyme engineers no longer have to rely solely on natural catalytic machinery. Scientists have thus gained the capacity to redesign, reimagine, and repurpose enzymes. Illustrative applications include enzyme cascades to manufacture the antivirals islatravir and molnupiravir, biocatalysts that can generate and control radicals, and enzymes that exploit photocatalysis to affect stereocontrolled C–C couplings, hydroaminations, or Diels-Alder reactions.

ADVANCES: The past 5 years have witnessed a surge in the development of data-driven tools enabling the accelerated discovery, engineering, and deployment of enzymes for applications in chemistry, medicine, and food technology. By leveraging enzymes' ability to control the environment of a chemical reaction, scientists have developed successful platforms for the construction of complex small molecules from suitable starting materials or even CO_2. In addition to their use in (chemo)-enzymatic cascade reactions, precisely tailored biocatalysts can manufacture new therapeutic modalities, such as antisense oligonucleotide therapeutics, and bioconjugates. DNA synthesis, which still relies on phosphoramidite chemistry, is being reinvented with template-independent deoxynucleotidyl transferases to make the process faster and cheaper. Machine learning, already dominating protein-structure prediction and design, is finding applications in enzyme engineering, including improvement of functions such as enantioselectivity, activity, and stability.



The accelerating development of biocatalysis. In silico and experimental tools developed in the last decade allow the fast creation of tailored enzymes. In addition to the concise synthesis of complex small molecules in elegant (chemo)-enzymatic cascades, newly accessible applications include enzymatic DNA synthesis, the generation of therapeutic oligonucleotides, and up- and recycling of plastic waste. [Figure created with Canva]

OUTLOOK: In the decade ahead, biocatalysis research and applications will continue to profit from advances in data mining, machine learning, and DNA reading and writing. The combinatorial design of enzymes and the ease with which new enzyme variants can be experimentally generated lends itself to train data-intense machine-learning algorithms. Ideally, the sequence-function data of variants screened in a directed-evolution campaign could be used to predict which variants to evaluate next. Machine learning works best with clean and reproducible data, mandating the standardization and reporting of data and the further development of experimental techniques, including molecular biology methods, automation, and high-throughput screening assays.

The toolbox of drug discovery is expanding beyond traditional small molecules (molecular weights <500 g/mol) to include RNA therapeutics, protein degraders, cyclopeptides, antibody drug conjugates, and gene therapy. Consequently, the synthetic complexity of lead molecules and clinical candidates is increasing with a growing percentage of chiral and beyond-rule-of-5 molecules. Enzymatic synthesis will play a key role beyond its present impact in small-molecule drug discovery and development.

The discovery of new enzyme families and the rational design of new enzyme functions will expand the toolbox of available biocatalysts. Development and deployment of retrosynthetic tools containing enzymatic reactions will be an important step toward democratizing biocatalysis and making it available to the nonexpert. Fueled by these innovations, nature's catalysts will be profitably used to address current challenges, including the fight against diseases, provision of affordable clean energy, and reduction of industrial and consumer waste. ■

The list of author affiliations is available in the full article online.
*Corresponding author: Email: uwe.bornscheuer@uni-greifswald.de
Cite this article as R. Buller *et al.*, *Science* **382**, eadh8615 (2023). DOI: 10.1126/science.ahd8615

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adh8615

REVIEW

BIOCATALYSIS

From nature to industry: Harnessing enzymes for biocatalysis

R. Buller[1], S. Lutz[2], R. J. Kazlauskas[3], R. Snajdrova[4], J. C. Moore[5], U. T. Bornscheuer[6]*

Biocatalysis harnesses enzymes to make valuable products. This green technology is used in countless applications from bench scale to industrial production and allows practitioners to access complex organic molecules, often with fewer synthetic steps and reduced waste. The last decade has seen an explosion in the development of experimental and computational tools to tailor enzymatic properties, equipping enzyme engineers with the ability to create biocatalysts that perform reactions not present in nature. By using (chemo)-enzymatic synthesis routes or orchestrating intricate enzyme cascades, scientists can synthesize elaborate targets ranging from DNA and complex pharmaceuticals to starch made in vitro from CO_2-derived methanol. In addition, new chemistries have emerged through the combination of biocatalysis with transition metal catalysis, photocatalysis, and electrocatalysis. This review highlights recent key developments, identifies current limitations, and provides a future prospect for this rapidly developing technology.

To innovate in synthetic chemistry, academic and industrial scientists increasingly apply enzymes to make simple and complex functionalized (bio)molecules (*1*). Inspired by the precise control that enzymes can exert over reaction outcomes, chemists are using biocatalytic transformations that complement or even substitute for more traditional chemical routes (*2*). However, as synthetically relevant reactions rarely have a counterpart in nature, the challenge in biocatalysis is the identification of a suitable enzyme catalyst for the desired application. Early uses of biocatalysis relied on accessible wild-type enzymes used in food or other industries to produce laundry detergents, semisynthetic antibiotics, and simple chiral precursors for the pharmaceutical industry (*3*–*5*). Although such repurposing still occurs occasionally, most new applications require the discovery of enzymes with distinct reactivity or the engineering of existing enzymes to catalyze the desired reaction, accept the desired substrate, and be stable and active at the desired application conditions. The computational and experimental advances of the past few years have sped and simplified the tailoring of biocatalysts to access a rapidly growing set of chemical transformations. Today, strategies for the development of performant biocatalysts include screening natural diversity to discover a desired enzyme activity, engineering biocatalysts to alter the substrate range, redesigning mechanisms to create previously unknown reactivity, and computationally designing enzymes de novo (Fig. 1). These strategies now allow the field to be much bolder in choosing which reactions are attempted enzymatically. Biocatalysts are being redesigned, reimagined, and repurposed to grant access to the desired targets with great efficiency.

Here, we evaluate the recent experimental and computational developments that led to this boost in biocatalysis and its applications (Fig. 2). Highlighted experimental innovations include optimized strategies to read and write DNA, advanced tools for automation and high-throughput screening, and the ability to incorporate novel catalytic elements into enzymes, including noncanonical amino acids. Computational progress is reflected in improved access to suitable enzyme-encoding sequences, machine learning–based methods to accurately predict protein structures, and data-driven tools to guide enzyme engineering in the creation of information-enriched, functional variant libraries. These achievements have translated into a variety of applications, including (chemo-) enzymatic cascade reactions, new-to-nature chemistries, enzymatic plastic degradation, and the synthesis of biologics and therapeutics (Fig. 2).

Developing biocatalysts

There are several approaches that can be used to find or create an enzyme with the desired catalytic activity (Fig. 1) (*6*). First, the target activity could exist in nature but needs to be discovered. High-quality, low-cost sequencing of DNA has now revealed complete genomes of species from all over the globe (*7*, *8*), as well as DNA fragments from metagenome samples. The number of available protein sequences increased more than 20-fold in the last 5 years [2023, >2.4 billion (*9*, *10*); 2018, ~123 million (*11*)]. Searching these sequences reveals many putative enzymes, some with predictable activities and many that are of unknown function. Combining this sequence data with computational tools such as sequence similarity networks (*12*), phylogenetic analyses (*13*), and protein-structure prediction improves the precision of the search.

In a representative example, polyethylene terephthalate (PET) hydrolases (PETases) to degrade PET plastics were found through classical screening (*14*), bioinformatics searches (*15*), and sequencing of environmental samples (*16*). Alternatively, legacy enzyme collections of enzymes that proved useful in the past are an invaluable resource for identifying suitable biocatalysts or advanced starting points for engineering projects. Commercially available enzyme screening kits and private libraries of enzyme variants for frequently used enzymes such as lipases, esterases, ketoreductases, and transaminases can considerably shorten timelines toward yielding biocatalysts with the desired target activity. More recently, advances in protein structure prediction (see "Computational tools" section below) have further improved the search for suitable enzymes in sequence databases. For example, in a search for dehalogenases, a sequence similarity search identified 2905 putative target enzymes (*17*). Subsequent analysis of homology models—which considered the size of the active site, the presence of catalytic residues, and tunnels to access the active site—narrowed the search to 45 genes, 40 of which yielded catalytically active dehalogenases.

The most common strategy to access a desired enzymatic function is repurposing of an existing enzyme that fits many, but not all, requirements for the new application. Often an existing enzyme catalyzes the desired reaction but does not accept the desired substrate or does not produce the desired product isomer. In this situation, protein engineering is required to shift the substrate or product preference of the enzyme. For example, the redesign of an aspartase enabled the hydroamination of α,β-unsaturated carboxylic acids to make β-amino acids (*18*). The target substrates contained a hydrophobic substituent at the position corresponding to the α-carboxylate of aspartic acid. The design replaced four mostly polar amino acids in the binding site with hydrophobic amino acids. In another case, triggered by the need for selective halogenation reactions, researchers expanded the substrate range of Fe(II)/α-ketoglutarate–dependent halogenases to include nonnatural substrates

[1]Competence Center for Biocatalysis, Institute of Chemistry and Biotechnology, Zurich University of Applied Sciences, 8820 Wädenswil, Switzerland. [2]Codexis Incorporated, Redwood City, CA 94063, USA. [3]Department of Biochemistry, Molecular Biology and Biophysics, Biotechnology Institute, University of Minnesota, Saint Paul, MN 55108, USA. [4]Novartis Institutes for BioMedical Research, Global Discovery Chemistry, 4056 Basel, Switzerland. [5]MRL, Merck & Co., Rahway, NJ 07065, USA. [6]Institute of Biochemistry, Dept. of Biotechnology and Enzyme Catalysis, Greifswald University, Greifswald, Germany.
*Corresponding author. Email: uwe.bornscheuer@uni-greifswald.de

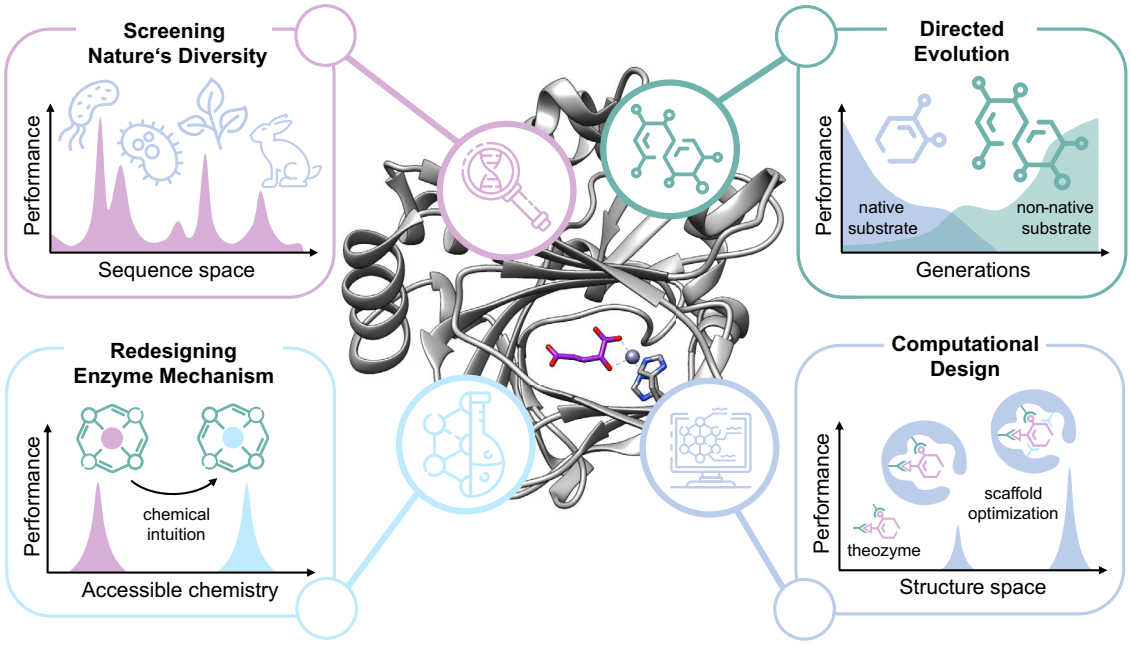


Fig. 1. Strategies for biocatalyst development. (Top left) Screening natural diversity from environmental samples or from enzyme libraries can aid in discovering the desired enzyme activity. (Top right) Directed evolution starts with an enzyme that has a small amount of performance on a desired reaction and tunes the enzyme to work well under desired reaction conditions. Substrate walking by means of directed evolution expands that capability by starting with an enzyme that is mechanistically competent to catalyze a desired reaction (but doesn't actually catalyze it) and engineers the enzyme to execute the reaction on the target substrate to make a desired product (i.e., alters its substrate range). (Bottom left) Redesigning the enzyme mechanism exploits chemical intuition to create new chemical functionality—for example, through cofactor repurposing and the addition of new prosthetic groups. (Bottom right) Computational design equips inert protein structures with enzymatic function. All images from stock.adobe.com, except where noted. Circled images, clockwise from bottom left: images 1 and 2, ylivdesign; image 3, Lifeking; image 4, ylivdesign. Images in boxes, clockwise from bottom left: box 1, by author; box 2, images 1 to 3 (left to right) by davooda, image 4 by SkyLine; box 3, Lifeking; box 4, by author.

(*19*–*21*), such as the macrolide soraphen A, a potent fungicide. By using algorithm-aided engineering, the apparent turnover number (k_{cat}) for soraphen was improved >100-fold, turning the halogenase into a suitable catalyst for aiding in structure-activity relationship studies in medicinal chemistry. The halogenase also accepts other anions, including azide, nitrate, and nitrite, enabling the creation of a wider range of products (*22*). In another example the selectivity of a protease was altered to degrade gluten peptides as a potential supplement for patients with celiac disease (*23*). The starting protease favored hydrolysis after a Pro-Arg/Lys sequence, whereas gluten contains many Pro-Gln-Gln/Leu sequences. Hydrolysis was desired between Gln and Gln/Leu residues. The researchers introduced eight substitutions, two in the new Gln site, one in the Gln/Leu site, and five more distant substitutions to stabilize the protein. The modified protease is currently undergoing clinical trials.

The concept of "substrate walking"—evolving an enzyme from transforming its natural substrate to accepting an industrially relevant substrates was rarely applied (*24*, *25*) a little over a decade ago. The evolution of a transaminase to manufacture sitagliptin (*26*) is an early example. Today, Pictet-Spenglerases that use benzaldehyde (*27*), flavin-dependent halogenases active on high–molecular weight indoles and carbazoles (*28*), opine dehydrogenases (reductive aminases) that use amines and ketones as opposed to amino- and ketoacids (*29*), and transketolases that use nonpolar aromatic substrates (*30*) have all been created through substrate walking.

In more difficult cases, no enzyme with the desired reactivity exists, but an enzyme with a mechanistically similar catalytic activity is known. Combining chemical reasoning with protein engineering can extend the natural catalytic activity to the desired activity (*6*, *31*). One example is the expansion of the catalytic repertoire of cytochrome P450 monooxygenases to catalyze carbene and nitrene reactions. Their oxidase mechanism involves an iron porphyrin oxo (Fe=O) intermediate, and it was reasoned that iron carbene [Fe=C(R1)(R2)] and iron nitrene (Fe=NR) species might form similarly given the appropriate reactive substrate (*32*, *33*). Initial experiments revealed inefficient reactions, but optimization by directed evolution and replacement of the proximal thiolate heme ligand with the more weakly donating serine—yielding the so-called P411 scaffolds—led to catalysts capable of new-to-nature chemistry, including cyclopropanation, cyclopropenation, Si–C bond formation, B–C bond formation, C–H insertion, and alkyl transfer (Fig. 3) (*34*), as well as aziridination, sulfide imidation, C–H amidation, and C–H amination (Fig. 3) (*35*). These examples are enzyme-catalyzed reactions in vitro, but this new-to-nature carbene-transfer chemistry was also used to extend biosynthesis in vivo. Engineered *Streptomyces* strains biosynthesized a carbene-transfer reagent, azaserine, as well as the acceptor styrene, to yield unnatural cyclopropanes (*36*). P450 monooxygenases can also be repurposed by following other strategies. For example, the P450 monooxygenase from *Labrenzia agreggata* was fashioned into a ketone synthase capable of the direct oxidation of internal arylalkenes to ketones by harnessing highly reactive carbocation species as key intermediates (*37*). To cite another example, researchers found that P450-based radical cyclases could be developed through the exploitation of a metalloredox strategy to catalyze stereoselective atom-transfer radical reactions yielding substituted γ-lactams (*38*) or arenes (*39*).

Changes in catalytic activity can also be more drastic; for example, a reductase was extended to a cyclase that forms C–C bonds through radical intermediates. Flavin-dependent "ene" reductases normally catalyze the reduction of electronically activated alkenes through the stepwise addition of H_2—the first step occurring through a hydride transfer (a two-electron transfer) from the flavin hydroquinone, followed by the second step, a proton transfer from a conserved tyrosine residue (Fig. 3). In other enzymes, flavins reduce substrates through two single-electron transfers, which creates radical intermediates. By replacing the activated alkene substrate of an ene reductase with an α-bromo ketone, which is a radical

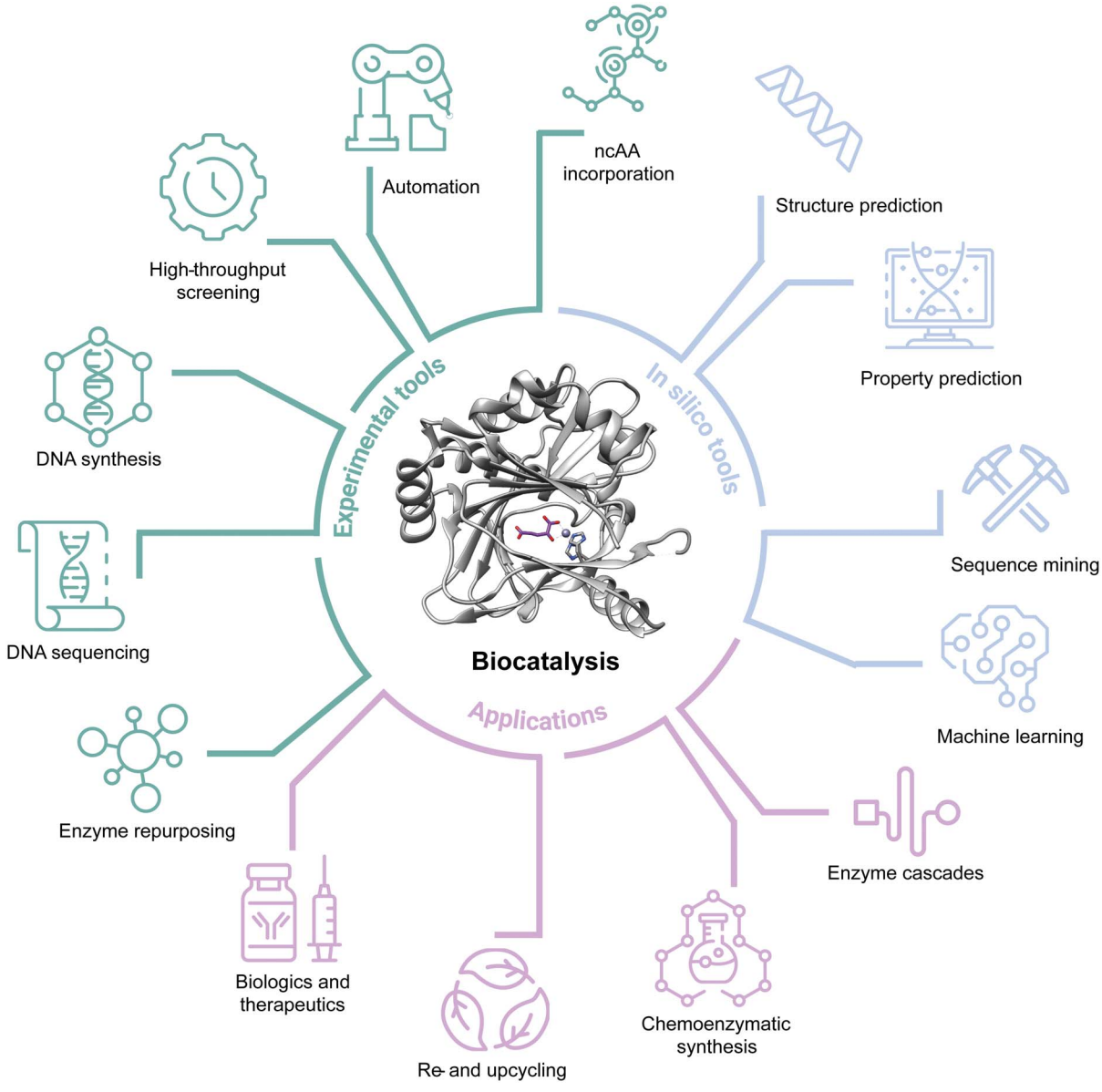


Fig. 2. Advances in experimental and computational tools have broadened the range of applications in biocatalysis. All images from stock.adobe.com, except where noted. Clockwise from bottom left: image 1 (Biologics and therapeutics), Sir.Vector; image 2, by author; images 3 and 4, ylivdesign; images 5 and 6, Artco; image 7, ylivdesign; image 8, muhamad; image 9, ylivdesign; image 10, davooda; image 11, Artco; image 12, RaulAlmu; image 13, ylivdesign; image 14, Skyline.

precursor, the ene reductase changes its mechanism and transfers a single electron. Rapid loss of bromide yields the corresponding α-ketonyl radical. In an appropriately selected substrate, this radical may cyclize, forming a C–C bond and abstracting a hydrogen atom from the flavin semiquinone (*40*). The surrounding active site guides the formation of new stereocenters in the product. Thus, the resulting reaction is an enantioselective reductive cyclization.

In other cases, radical precursors are less reactive and require light irradiation to initiate the single-electron transfer. In photobiocatalysis, a cofactor or amino acid within the protein active site is photoexcited to promote the electron or energy transfer required to convert starting materials to desired products. Only three natural enzymes follow this synthetic logic: fatty acid photodecarboxylase (*41*), DNA photolyase (*42*), and protochlorophyllide reductase (*43*). Whereas the latter two enzymes have less synthetic relevance, fatty-acid decarboxylases have been studied to catalyze the hydrodecarboxylation of fatty acids, a redox-neutral reaction that leads to alkanes, rendering them promising catalysts for biofuel production (*44*) and chemical building-block synthesis (*45*).

Apart from radical cyclizations of well-chosen organohalides (*46*), the photobiocatalysis approach similarly allows the alkylation of arenes (*47*), the asymmetric cross-electrophile coupling of alkyl halides and nitroalkanes (*48*) (Fig. 3), and hydroaminations through the generation of amidyl radicals (*49*). Nonenzymatic photoredox catalysts can also generate substrates for enzymes to create new reactions. Adding xanthene-based photocatalysts enabled an ene reductase to catalyze an enantioselective deacetoxylation (*50*). In another example, noncanonical amino acids were synthesized by using separate photocatalysis and enzymatic steps. The photocatalysis generates an alkyl radical from an alkyl trifluoroborate precursor. In the same solution, a modified tryptophan

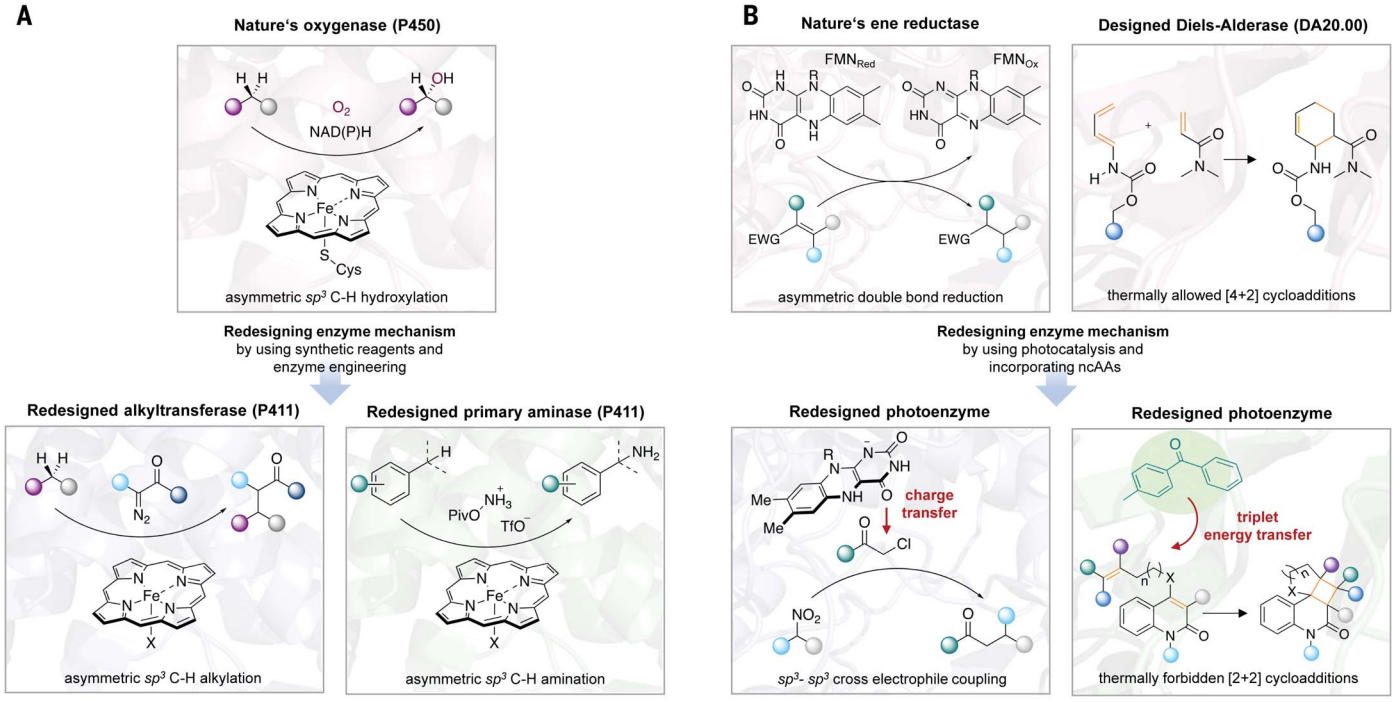


Fig. 3. Examples of redesigning the enzyme mechanism. (**A**) (Top) Oxygenation catalyzed by wild-type cytochrome P450 monooxygenase. The structural similarity of Fe oxenes and Fe nitrenes inspired the use of synthetic reagents to carry out new chemistries with engineered P450 enzymes. (Bottom left) These so-called P411 enzymes, which contain a serine residue as proximal ligand (denoted as X), were evolved to assemble C–C bonds through sp^3 C–H functionalization (*34*) and to aminate at benzylic and allylic (not shown) positions to yield (bottom right) enantioenriched, unprotected primary amines (*35*). Piv, pivaloyl; Tf, trifluoromethanesulfonyl. (**B**) (Top left) Asymmetric double-bond reduction catalyzed by an ene reductase through a hydride transfer (two-electron transfer) from flavinmononucleotide (FMN) followed by a proton transfer from a conserved tyrosine residue (not shown). By applying photocatalysis, the catalytic machinery of the ene reductase can be repurposed when radical precursors are used as substrates (as shown here, an α-chloro ketone). EWG, electron withdrawing group. (Top right) Thermally allowed [4+2] cycloaddition between 4-carboxybenzyl-*trans*-1,3-butadiene-1-carbamate and *N*,*N*-dimethylacrylamide catalyzed by a designed Diels-Alderase (*155*). (Bottom left) Irradiation initiates a single-electron transfer from the flavin hydroquinone cofactor and facilitates the formation of an α-ketonyl radical, which can engage in stereoselective sp^3–sp^3 cross electrophile couplings in the confines of the engineered enzyme scaffold (*48*). (Bottom right) Installation of the noncanonical amino acid 4-benzoylphenylalanine (green) by genetic-code expansion expands the reaction scope of the Diels-Alderase. Upon irradiation of the designer enzyme, the ncAA allows for triplet energy transfer to appropriately selected substrates, giving access to thermally forbidden reactions, such as intra- and bimolecular (not shown) [2+2] cycloadditions with high stereoselectivities (*156*). Structural illustrations are adapted from the PDB (PDB IDs: 2IJ2, cytochrome $P450_{BM3}$; 7TNB, ene reductase; and 3I1C, Diels-Alderase).

synthase catalyzes the dehydration of serine to form an enzyme-bound pyridoxyl-5′-phosphate aminoacrylate intermediate. Quenching of the radical and release of the intermediate yields a noncanonical amino acid. Adjusting the shape of the active site fine-tuned the enantioselectivity of all described reactions for up to three stereocenters (*51*).

Instead of modifying or extending existing enzymatic activities, researchers have also computationally designed new catalytic activities into a protein scaffold. This de novo design approach identifies the transition state of the desired transformation and then builds a binding site to stabilize it, reducing the problem of biocatalysis to one of molecular recognition. Although de novo enzyme design has created protein catalysts for model transformations, including proton-transfer, bimolecular-aldol, and Diels-Alder reactions (*52*). The initial activities of the designer catalysts were low, but directed evolution increased their catalytic activities. For example, directed evolution of a designed retroaldolase increased its catalytic activity from barely detectable to that of natural enzymes. The directed evolution introduced 31 substitutions, leading to a million-fold increase in activity (*53*). More recently, a deep learning–based approach generated large numbers of idealized protein structures and the sequences that encode it. The diverse scaffolds were employed to design an artificial luciferase. Three substitutions introduced by site-saturation mutagenesis yielded a 100-fold–higher photon flux than did the parent design (k_{cat}/K_m = 10^6 M^{-1} s^{-1}) (*54*) (where k_{cat}/K_m is the catalytic efficiency and K_m is the Michaelis constant). Although de novo design of enzymes with activities that rival their natural counterparts remains a major challenge at present, these studies have improved our understanding of how sequences fold into proteins and of how basic enzyme activity is created. De novo enzyme design will continue to grow as our understanding of the sequence-function relationship in enzymes improves, protein design methods mature, and computational power increases.

Experimental tools

Regardless of the strategy chosen to identify a suitable biocatalyst, laboratory work is always required to produce and, if needed, to tailor the target enzyme. The range of experimental tools available to protein engineers is consistently expanding, whether by lowering the cost of synthetic genes, speeding up individual directed-evolution cycles, or allowing for introduction of new catalytic elements to be added to proteins.

Biocatalyst development starts with a DNA construct encoding the enzyme of interest, which requires artificial DNA synthesis. Present DNA synthesis uses phosphoramidite chemistry developed 40 years ago (*55*); however, DNA synthesis performed with enzymes (*56*–*58*)

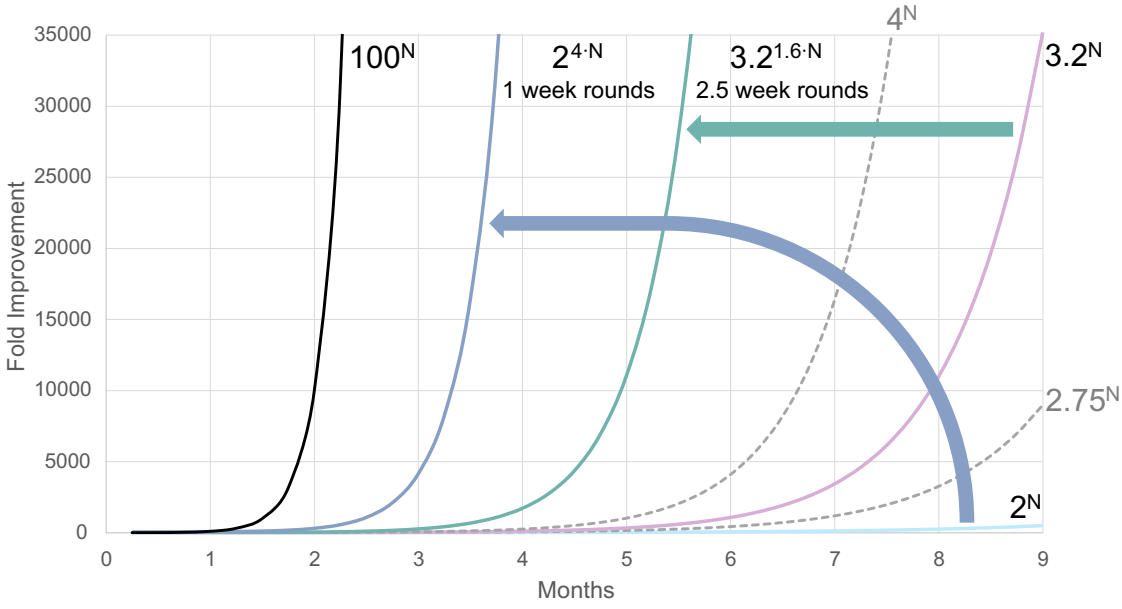


Fig. 4. Modeling of directed evolution methodologies. Recursive accumulation of single-point mutations, such as those generated by error-prone PCR methods, produces fitness improvements on the order of 2^N (light blue curve). Adding recombination techniques to evolution improves fitness accumulation to between 2.75^N (*26*) and 4^N (*122*, *152*) (dashed curves), averaging ~3.2^N (pink curve). Cell-free expression technologies increase the rate at which a cycle of evolution can be performed (green curve), and continuous-evolution techniques are estimated to accelerate accumulation of point mutations even more quickly (periwinkle curve). Microfluidics-based ultrahigh throughput screening has yielded 100-fold improvements per round (*76*). *N* indicates the number of evolution cycles. In this overview, it is assumed that one evolution cycle can be carried out per month.

yields higher-quality DNA and may be faster and cheaper, thus representing a biocatalytic solution in and of itself. Terminal deoxynucleotidyl transferases (TdTs) (Fig. 4) polymerize deoxynucleoside triphosphates (dNTPs) to the 3′-end of a DNA sequence in a template-independent fashion. Using dNTPs with a blocked 3′-hydroxyl stops TdT action after a single nucleotide incorporation, hence controlling an otherwise runaway polymerization. Native TdTs react slowly with blocked dNTPs, so improved TdTs were engineered for this application (*59*, *60*). TdT-based approaches have driven DNA oligonucleotide synthesis to (literally) new lengths of up to 1000 nucleotides, enabled by incorporation efficiencies of >99.6%, reviving the promise for rapid, single-run whole-gene synthesis. Milder aqueous reaction conditions also benefit DNA quality and greatly improve the overall process sustainability, opening up opportunities for gene editing and diagnostic applications. TdTs can also assemble short synthetic RNA fragments. Such sequences, containing modified nucleotides in addition to the standard RNA building blocks, hold substantial therapeutic potential as antisense oligonucleotides and small interfering RNAs (*61*–*63*).

Despite the advances in enzymatic DNA synthesis, most laboratories still rely on traditional molecular biology strategies to create large, randomized libraries through methods such as error-prone polymerase chain reaction (PCR) or the construction of a few specific variants through site-directed mutagenesis by using a chemically synthesized gene as template. Building libraries of predefined enzyme variants remains expensive and challenging with these traditional approaches. "Oligo-pools,"—which consist of a few hundred up to 1000 distinct polynucleotides with a length of about 300 base pairs—are a less-expensive alternative solution (0.00001 to 0.001 USD per nucleotide, depending on length, scale, platform, or vendor) (*64*). Despite disadvantages such as truncated DNA molecules and high error rates (*65*), the oligo-pool option may be more cost-effective than the degenerate or reduced codon-coverage primers typically used for library construction and may allow more flexible library design (*66*).

The cornerstone of protein engineering (*1*, *3*, *67*, *68*), the Nobel prize–winning strategy of directed evolution, relies heavily on such randomized gene libraries. When directed evolution appeared three decades ago, error-prone PCR was the driver of sequence variability, resulting in round-over-round accumulation of single mutations and an approximated two-fold improvement of enzyme performance per round of evolution (*69*) (Fig. 4, light blue curve). The contemporaneous addition of recombination techniques upgraded the rate of improvement to between 2.75 and 4 per evolution cycle (Fig. 4, dashed curves) by incorporating multiple mutations per round, reflecting the state of the art over the past two decades. Computational and molecular biology efforts have improved mutation prediction and accelerated enzyme-variant creation so that the screening burden can be reduced, but not to the point that screening cycles or improvements per cycle have been greatly altered from these historical norms (*70*). Additionally, protein engineers continue to increase the complexity of attempted evolution campaigns. Together, these developments point to an increasing bottleneck in biocatalysis: higher demand for protein-engineering resources and a lack of improvement in the way evolution experiments are performed. Technical innovations are on the horizon, however. Cell-free protein-expression technologies provide all the necessary components to transcribe and translate many DNA sequences into functional proteins, avoiding the time-consuming steps of cell transformation, growth, and induction (*71*). Using this technology to produce enzyme libraries for testing in evolution experiments has the potential to accelerate the time to complete a round of evolution by ~1.6-fold (Fig. 4, green curve). Turning to the biological world, continuous-evolution strategies tie the success of a desired enzyme activity to the growth rate of a producing microbe and simultaneously provide the microbe with a mechanism of introducing mutation into the DNA coding for the desired protein only (*72*, *73*). Evolution of the desired protein then occurs in continuous culture, which circumvents the same laboratory manipulations required for cell-free protein expression and in addition avoids the man-made construction of new variant libraries, leading to an eightfold increase in the rate of evolution performance compared with error-prone PCR [and a five-fold improvement compared with current state of the art (Fig. 4, periwinkle curve)]. For example, mRNA display libraries (*74*) have generated new proteins, new types of catalytic activity, and binding proteins with very high affinity for their targets. Similarly, microfluidic technologies have demonstrated the ability to

screen large library sizes, which positively changes the nature of directed-evolution experiments (Fig. 4, black curve) (*75*). Encapsulating cells into picoliter-sized droplets, along with a lysis reagent and an analytical readout, has allowed scientists to assay thousands of samples per second, which is on par with the rates achieved by fluorescence-activated cell sorting (FACS) techniques. Because millions of samples can be analyzed, libraries with multiple mutations per sequence can be screened deeply enough to uncover rare events, reaching 100-fold improvements per evolution round or greater (*76*). Although lack of commercialized hardware, relatively high implementation costs, and technical complexity has so far limited the use of such tools to expert practitioners, these strategies show promise for managing the longer evolution timelines required for increasingly complex targets.

Other emerging tools that include useful in protein engineering include cryo–electron microscopy (cryo-EM), which has been profitably employed to guide improvement of a nitrilase (*77*), and microcrystal electron diffraction (MicroED), which supported the mechanistic investigation of carbene transfer in a designed protein (*78*). Both methods enable gathering of mechanistic insights and structural data complementary to x-ray crystallography, which nevertheless remains an important tool for generating high-resolution structures, especially of smaller enzymes.

Complementing the biocatalysts obtained by enzyme-engineering campaigns, some applications require researchers to equip proteins with new-to-nature functionalities. This is achieved through the incorporation of noncanonical amino acids (ncAAs) by using either amber stop-codon suppression or promiscuous aminoacyl tRNA synthetases (aaRSs), which expands the repertoire of available catalytic elements beyond those encoded in the 20 proteinogenic amino acids (*79*). ncAAs can be harnessed to tune enzyme properties, reveal mechanisms in complex catalytic machineries, and create enzymes with functions not found in nature.

An illustrative example of tuning enzyme properties by incorporating ncAAs is the introduction of N_δ-methylhistidine (NMH) as proximal ligand into the heme protein ascorbate peroxidase, which led to substantial increases in turnover number without compromising catalytic efficiency (*80*). Similarly, introducing NMH modulates the reactivity of compound II in cytochrome *c* peroxidase (*81*, *82*) and enhances the promiscuous peroxidase (*83*) and cyclopropanation (*84*, *85*) activities of myoglobin. Most recently, this noncanonical nucleophile was introduced into a designer enzyme created for the Morita-Baylis-Hillman reaction. During evolution, NMH substantially altered the evolutionary trajectory, yielding an order of magnitude– more-active variant compared with previously engineered enzymes that did not include the new catalytic entity (*86*).

The installation of tyrosine analogs is a valuable tool for mechanistic investigations of enzymes involving radical intermediates, in which such residues may play a catalytic role. Incorporating halogenated tyrosine analogs at key positions in the active site of *Escherichia coli* ribonucleotide reductase (*87*) or verrucologen synthase (*88*), an Fe(II)/α-ketoglutarate–dependent dioxygenase, enabled the use of analytical tools such as electron magnetic resonance, giving key insights into the reaction mechanism.

Installation of a photosensitizer through the use of amber stop-codon suppression created an enzyme that catalyzes a photoinduced [2+2]-cycloaddition. Upon irradiation of the unnatural amino acid 4-benzoylphenylalanine—which was installed in the active site of a de novo designed Diels-Alderase—triplet energy transfer initiated thermally forbidden [2+2]-cycloadditions. Directed evolution to fine-tune the enzyme yielded a catalyst with high enantioselectivity for both intra- and intermolecular reactions (*89*) (Fig. 3).

Computational tools

Machine learning–based structure-prediction tools promise to limit the excessive screening or selection necessary to analyze the huge libraries typically generated during directed evolution. The advanced statistics system AlphaFold2, developed by DeepMind, predicts protein structures from amino acid sequence with much greater accuracy than did previous methods. AlphaFold2 uses deep neural networks to predict inter-residue distances from amino acid sequences (*90*). Its predictions rely on the structural data in the Protein Data Bank (PDB) (www.wwpdb.org), which contains the three-dimensional structures of 200,000 proteins and nucleic acids. The neural networks assume that substructures that appear frequently in the database are more stable than absent or rarely occurring substructures. At present, a database of predicted structures (https://alphafold.com/) contains over 200 million entries and is continuously growing. Similarly useful alternatives are RoseTTAFold (*91*) and ESMFold (*92*), created by Meta Platforms.

Structure prediction has the potential to reveal the arrangement of active-site residues of the target enzyme, guiding enzyme optimization (*93*). However, enzymatic function also requires substrate binding and product release and often relies on cofactors or metal ions. Tools such as AlphaFill (*94*) and DiffDock (*95*) project missing organic molecules and metal ions into the protein pockets, providing a model that may more accurately reflect the desired chemistry. To capture motion, AlphaFold2 can approximate conformational heterogeneity by, for example, reducing the depth of the multiple sequence alignments that serve as input for the algorithm (*96*), as demonstrated when researchers used AlphaFold2 to create models of alternate conformations of a tryptophan synthase (*97*).

The deep neural networks trained to predict native protein structures from their amino acid sequences can be inverted to design new proteins (*98*–*101*). The first step is to predict the structure of a random amino acid sequence, which yields parts of a structure with confident predictions and other parts with uncertain predictions. Repeated modification of the amino acid sequence of the uncertain regions eventually leads to a sequence whose entire structure is predicted confidently. Testing these predictions often yields stable proteins with structures close to those predicted. These neural network approaches to protein design are faster than physics-based methods such as Rosetta Match (*102*), Rosetta Design (*103*), and Rosetta Ligand (*104*) because they do not attempt to optimize interactions such as side-chain packing. However, the neural network approaches lack the physical transparency of methods such as Rosetta. Adding diffusion models trained on proteins [similar to Stable Diffusion, used to generate images (*105*), or ChatGPT, used to generate text] to protein design expands the number and variety of designs generated, thus increasing the chances of success (*106*). Furthermore, language models can generate functional protein sequences across diverse families (*107*).

Although the design of protein structures is improving, design of proteins with catalytic function remains challenging. Design typically yields inefficient enzymes, which then require extensive optimization by directed evolution (see "Experimental tools" section above). This might be because efficient catalysis requires more precise positioning of catalytic groups than is presently achievable with the algorithms. Another reason for this problem is an incomplete understanding of the structural features, including motions, that are needed for catalysis. Classical protein-structure determination with x-ray crystallography uses transition-state analogs or suicide substrates to reveal how a substrate binds to the active site of an enzyme, to learn where binding pockets are located, and to gather information about protein motions that contribute to catalysis. For instance, the structural basis for the higher catalytic activity of lipases at an oil-water interface remained a puzzle despite several available x-ray structures of lipases, which showed a buried catalytic site. Brzozowski *et al.* solved this puzzle in 1991 with a lipase structure that showed a dramatic movement of a helical lid covering the catalytic site (*108*). A more recent puzzle was a monoamine oxidase, whose x-ray structure also showed a closed conformation that could not explain the effects of mutations on catalytic activity and substrate scope. In

this case, a computational approach using long-timescale molecular dynamics identified partially and fully open conformations that could account for the changes in catalysis (*109*). Current structure-prediction tools do not include protein dynamics in most cases and likewise do not include information about disulfide bonds and posttranslational modifications (which are often not covered in protein sequences in databases) such as glycosylations and acetylations, mandating the use of complementary computational tools and experiments.

Connecting protein structures to protein functions such as reactivity or selectivity is an advancing field in machine learning. Both neural networks based on protein structure (*110*) and those based on sequences using contrastive learning have increased their reliability to predict protein function (*111*). In many cases, the application of machine learning to catalytic properties of proteins is limited by the availability of reliable experimental data to train the neural network; recently, the EnzymeML database was established to address this issue (*112*). In special cases in which the properties of tens of thousands of protein variants have been measured, machine learning can predict variants with improved binding properties (*113*, *114*). The use of data from computational modeling predictions has also aided machine-learning approaches to design more-selective enzymes (*115*, *116*).

A further challenge in enzyme design is predicting the effects of distant amino acid substitutions. Experiments reveal that such residues influence catalysis, making their prediction

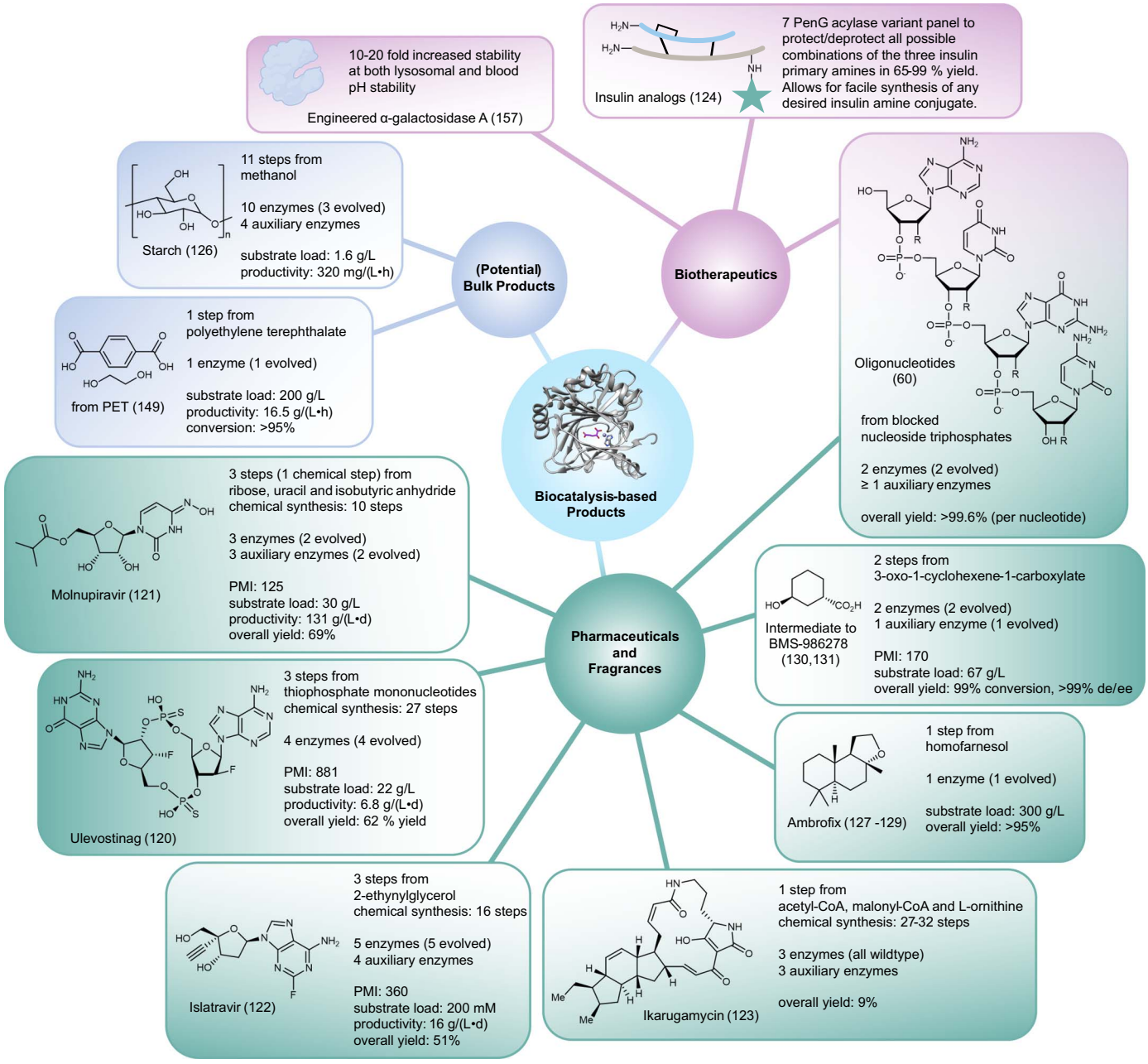


Fig. 5. Selected examples of recent biocatalysis-based products. Examples shown include biotherapeutics [engineered α-galactosidase A (*157*), insulin analogs (*124*), and oligonucleotides synthesized from TdTs (*60*)]; potential bulk products [starch synthesized from CO_2 (*126*) and plastic bottles recycled from PET (*149*)]; pharmaceuticals [molnupiravir (*121*), ulevostinag (*120*), islatravir (*122*), ikarugamycin (*123*), and an intermediate to BMS-986278 (*130*, *131*)]; and the fragrance Ambrofix (*127*–*129*).

essential to the design of efficient enzymes. These residues are too far away to directly interact with the substrate. They may act through a domino-like effect as the protein flexes and moves to alter positioning and flexibility of the catalytic residues and the substrate(s) within the active site, similarly to allostery. One promising computational approach to capture the effect of relevant distant residues is the shortest path map (*117*). Starting from a molecular dynamics simulation, this approach identifies residues that move together and are connected to the catalytic residues. When this strategy was applied, distant substitutions abolished the allosteric activation of tryptophan synthase B, activating it permanently (*118*).

Applications

Combining several engineered enzymes into a cascade creates new biochemical pathways (*119*). Recent examples (Fig. 5) include the syntheses of pharmaceuticals—such as the cyclic dinucleotide ulevostinag (*120*), molnupiravir (*121*), islatravir (*122*), and ikarugamycin (*123*)—and bioconjugates to make novel insulins (*124*), as well as artificial sweeteners. CO_2-fixation pathways capable of yielding compounds through the CETCH cycle (*125*), and even starch synthesis from CO_2-derived methanol (*126*), are further examples of complex engineered pathways.

In the fragrance industry, (−)-Ambrox, which has an ambery and woody odor, is one of the most widely used biodegradable fragrance ingredients. The previous synthesis of (−)-Ambrox (marketed as Ambrofix) was a multistep route from the diterpene sclareol isolated from clary sage. A new one-step process from biosynthetically manufactured homofarnesol uses an engineered squalene-hopene cyclase (Fig. 5) (*127*–*129*).

Another example of an enzyme cascade is a one-pot, two-enzyme reduction of allylic ketone 3-oxocyclohexene-1-carboxylate to the corresponding saturated alcohol reported in early 2023 by Bristol Myers Squibb (*130*) and Codexis (*131*). The ene reductase (ERED) and ketoreductase (KRED) process showed high enantio- and chemoselectivity. Maximum product yields benefited from the compatibility of the cofactor regeneration system and avoidance of side reactions. Separately, substrate inhibition and enzyme robustness under process conditions were improved to deliver a scalable enzyme cascade while greatly reducing the number of synthesis steps, improving overall yield, and lowering process mass intensity (PMI)—an indicator of environmental impact—from a value of 2017 for the initial chemical route to a PMI of only 170.

The development of novel enzyme cascades has been simplified by the availability of retrosynthesis tools—the disconnection approach used by organic chemists for decades—to plan synthetic routes (*132*, *133*). Programs [e.g., RetroBioCat (*134*)] identify possible routes for the multistep biocatalysis reactions while keeping in mind aspects such as commercial availability, cofactor requirements, and solvent tolerance (*135*–*138*). Machine learning can also assist with this purpose (*139*). Similarly, metabolic pathway planning predicts routes to complex natural products and also simpler molecules such as 1,4-butanediol (*140*).

Developing drugs for more difficult-to-drug targets leads to molecules with increasing numbers of chiral centers, beyond-rule-of-5 scaffolds (*141*), and bioconjugates (*142*) [e.g., radiotherapeutics, antisense oligonucleotide therapeutics, and antibody-drug conjugates (*143*)]. To address these synthetic (and later manufacturing) challenges requires finding enzymatic tools to produce these complex scaffolds. For example, it was reported that selective acylation of three amino groups of insulin—either of the two amino termini or at an internal lysine residue—was achieved by using engineered penicillin G acylase. These enzyme variants enabled installation of a cleavable phenylacetamide protecting group in a programmable manner, while leaving one or more amino groups unprotected for subsequent chemical modification. The high enzymatic regioselectivity improved the overall purity and yield of the insulin conjugates (*124*) (Fig. 5).

Despite the fact that chemical synthesis methodologies for modified oligonucleotides are well established, large-scale manufacturing in a sustainable and economically feasible manner remains challenging (*144*). Errors in the chemical coupling efficiency accumulate as the length of the oligonucleotide increases. Several companies and academic laboratories have developed technology to enzymatically assemble chemically modified oligonucleotides (*62*, *145*). The approach uses the well-known catalytic activity of RNA ligases to form an adenosine triphosphate (ATP)–dependent covalent bond between the 3′-OH and 5′-PO_4 termini of two oligoribonucleotides to form one larger, continuous strand (*146*). A recent method describes the use of such an RNA ligase to synthesize a chemically modified RNA starting from short (≤9-nucleotide) oligonucleotide fragments (*147*) with 40 to 80% conversion.

An orthogonal enzymatic approach to oligonucleotides (*62*) used self-priming (hairpin-like) templates added in catalytic amounts. DNA polymerase amplifies the complementary sequence in the presence of nucleoside triphosphate (NTP) building blocks. A specific endonuclease cleaves the newly synthesized chain and then releases the template for the next catalytic cycle. The synthesis uses unprotected building blocks in water, without large amounts of acetonitrile (as is typically the case in solid-phase synthesis), hence addressing a major sustainability challenge. Whether the unprotected NTP will be possible to source on sufficiently large scales with acceptable timelines and cost to fully embed this concept on a manufacturing scale remains to be seen. However, given the success of enzymatic synthesis of unnatural mono- and dinucleosides (e.g., ulevostinag, Fig. 5), biocatalytic production can be considered feasible.

Enzymes can also be used to address plastic pollution. Since the 1950s, almost 9 billion tons of plastics have been manufactured, and waste plastics are a major environmental concern worldwide. Considerable efforts have been made to degrade and recycle commodity plastics with biocatalytic approaches (*148*)—especially for polyesters, polyamides, and polyurethanes—because key chemical bonds can be hydrolyzed to yield the corresponding monomers, which can be used to make new virgin polymer. The most advanced recycling process has been reported for PET. Rational design methods created a quadruple mutant of the leaf-branch compost cutinase (LCC) that is efficient enough to establish a robust process currently implemented on an industrial scale (*149*) (Fig. 5). In a more recent example, a PETase from *Ideonella sakaiensis* was engineered by using machine learning (*150*). Present research efforts focus on polyamide and polyurethane hydrolyzing enzymes; the first candidate enzymes were recently identified in a metagenome library (*151*).

Conclusions

Advanced tools developed in the past few years have sped up protein engineering to the point where enzymes are an equal counterpart to conventional organic synthesis catalysts, resulting in an upsurge of biocatalysis applications in pharmaceutical manufacture (*152*) and many other areas. Future enzyme engineering also needs to speed up discovery-design-test cycles to maintain momentum and expand synthetic contributions to new enzyme classes. Likewise, the combination of enzymes with other (catalytic) synthetic chemistry methods such as transition metal catalysis, photocatalysis, and electrocatalysis (*119*, *153*, *154*) will be required to address humankind's challenges, such as combating climate change, degrading plastic waste, transitioning to renewable energy, and developing new therapies for medical treatment. Repurposing of enzyme mechanisms to broaden the repertoire of biocatalytic reactions represents a new concept for creating desired activities in enzyme catalysts. Other exciting strategies en route to novel enzymatic reactivities are de novo enzyme design (*54*) and computer "hallucination" (*98*).

Eleven years ago, the third wave of biocatalysis catapulted enzyme technology from "designing a biocatalytic process around the limitations of the enzyme to engineering the enzyme to fit the specifications of the process" (*1*). Now, the technology has taken yet another leap: The enzyme engineer no longer has to

depend on the enzyme's native catalytic prowess to define the accessible chemistry but can instead dream up bold new-to-nature reactivities and bring these to biocatalytic reality.

REFERENCES AND NOTES

1. U. T. Bornscheuer *et al*., Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012). doi: 10.1038/nature11117; pmid: 22575958
2. K. Chen, F. H. Arnold, Engineering new catalytic activities in enzymes. *Nat. Catal.* **3**, 203–213 (2020). doi: 10.1038/s41929-019-0385-5
3. E. L. Bell *et al*., Biocatalysis. *Nat. Rev. Methods Primers* **1**, 46 (2021). doi: 10.1038/s43586-021-00044-z
4. K. Buchholz, V. Kasche, U. T. Bornscheuer, *Biocatalysts and Enzyme Technology*. (Wiley-VCH, ed. 2, 2012).
5. A. S. Bommarius, B. R. Riebel, *Biocatalysis*. (Wiley-VCH, 2004).
6. C. Klaus, S. C. Hammer, New catalytic reactions by enzyme engineering. *Trends Chem.* **4**, 363–366 (2022). doi: 10.1016/j.trechm.2022.03.002
7. L. R. Thompson *et al*., A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017). doi: 10.1038/nature24621; pmid: 29088705
8. H. A. Lewin *et al*., Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4325–4333 (2018). doi: 10.1073/pnas.1720115115; pmid: 29686065
9. The Uniprot Consortium, UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2020).
10. L. Richardson *et al*., MGnify: The microbiome sequence data analysis resource in 2023. *Nucleic Acids Res.* **51** (D1), D753–D759 (2023). doi: 10.1093/nar/gkac1080; pmid: 36477304
11. M. Steinegger, J. Söding, Clustering huge protein sequence sets in linear time. *Nat. Commun.* **9**, 2542 (2018). doi: 10.1038/s41467-018-04964-5; pmid: 29959318
12. J. A. Gerlt *et al*., Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta. - Prot. Proteom.* **1854**, 1019–1037 (2015). doi: 10.1016/j.bbapap.2015.04.015
13. D. M. Emms, S. Kelly, SHOOT: Phylogenetic gene search and ortholog inference. *Genome Biol.* **23**, 85 (2022). doi: 10.1186/s13059-022-02652-8; pmid: 35346327
14. R.-J. Müller, H. Schrader, J. Profe, K. Dresler, W.-D. Deckwer, Enzymatic Degradation of Poly(ethylene terephthalate): Rapid Hydrolyse using a Hydrolase from *T. fusca*. *Macromol. Rapid Commun.* **26**, 1400–1405 (2005). doi: 10.1002/marc.200500410
15. E. Erickson *et al*., Sourcing thermotolerant poly(ethylene terephthalate) hydrolase scaffolds from natural diversity. *Nat. Commun.* **13**, 7850 (2022). doi: 10.1038/s41467-022-35237-x; pmid: 36543766
16. S. Sulaiman *et al*., Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach. *Appl. Environ. Microbiol.* **78**, 1556–1562 (2012). doi: 10.1128/AEM.06725-11; pmid: 22194294
17. M. Vasina *et al*., Advanced database mining of efficient haloalkane dehalogenases by sequence and structure bioinformatics and microfluidics. *Chem Catal.* **2**, 2704–2725 (2022). doi: 10.1016/j.checat.2022.09.011
18. R. Li *et al*., Computational redesign of enzymes for regio- and enantioselective hydroamination. *Nat. Chem. Biol.* **14**, 664–670 (2018). doi: 10.1038/s41589-018-0053-0; pmid: 29785057
19. T. Hayashi *et al*., Evolved Aliphatic Halogenases Enable Regiocomplementary C-H Functionalization of a Pharmaceutically Relevant Compound. *Angew. Chem. Int. Ed.* **58**, 18535–18539 (2019). doi: 10.1002/anie.201907245; pmid: 31589798
20. J. Büchler *et al*., Algorithm-aided engineering of aliphatic halogenase WelO5* for the asymmetric late-stage functionalization of soraphens. *Nat. Commun.* **13**, 371 (2022). doi: 10.1038/s41467-022-27999-1; pmid: 35042883
21. S. Duewel *et al*., Directed Evolution of an Fe ^{II} -Dependent Halogenase for Asymmetric C(sp ^{3})–H Chlorination. *ACS Catal.* **10**, 1272–1277 (2020). doi: 10.1021/acscatal.9b04691
22. A. Papadopoulou, F. Meyer, R. M. Buller, Engineering Fe(II)/α-Ketoglutarate-Dependent Halogenases and Desaturases. *Biochemistry* **62**, 229–240 (2023). doi: 10.1021/acs.biochem.2c00115; pmid: 35446547
23. C. Wolf *et al*., Engineering of Kuma030: A Gliadin Peptidase That Rapidly Degrades Immunogenic Gliadin Peptides in Gastric Conditions. *J. Am. Chem. Soc.* **137**, 13106–13113 (2015). doi: 10.1021/jacs.5b08325; pmid: 26374198
24. J. M. Smith, Natural selection and the concept of a protein space. *Nature* **225**, 563–564 (1970). doi: 10.1038/225563a0; pmid: 5411867
25. M. Kaltenbach, N. Tokuriki, Dynamics and constraints of enzyme evolution. *J. Exp. Zool. B Mol. Dev. Evol.* **322**, 468–487 (2014). doi: 10.1002/jez.b.22562; pmid: 24522979
26. C. K. Savile *et al*., Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 305–309 (2010). doi: 10.1126/science.1188934; pmid: 20558668
27. E. Eger *et al*., Asymmetric Biocatalytic Synthesis of 1-Aryltetrahydro-β-carbolines Enabled by "Substrate Walking". *Chemistry* **26**, 16281–16285 (2020). doi: 10.1002/chem.202004449; pmid: 33017078
28. J. T. Payne, C. B. Poor, J. C. Lewis, Directed evolution of RebH for site-selective halogenation of large biologically active molecules. *Angew. Chem. Int. Ed.* **54**, 4226–4230 (2015). doi: 10.1002/anie.201411901; pmid: 25678465
29. O. Alvizo, M. A. Mayo, J. C. Moore, US Patent US10370648B2. (2014).
30. P. E. Affaticati *et al*., Structural Analysis of an Evolved Transketolase Reveals Divergent Binding Modes. *Sci. Rep.* **6**, 35716 (2016). doi: 10.1038/srep35716; pmid: 27767080
31. H. Renata, Z. J. Wang, F. H. Arnold, Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **54**, 3351–3367 (2015). doi: 10.1002/anie.201409470; pmid: 25649694
32. O. F. Brandenberg, R. Fasan, F. H. Arnold, Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotechnol.* **47**, 102–111 (2017). doi: 10.1016/j.copbio.2017.06.005; pmid: 28711855
33. Y. Yang, F. H. Arnold, Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer. *Acc. Chem. Res.* **54**, 1209–1225 (2021). doi: 10.1021/acs.accounts.0c00591; pmid: 33491448
34. R. K. Zhang *et al*., Enzymatic assembly of carbon-carbon bonds via iron-catalysed sp ^{3} C-H functionalization. *Nature* **565**, 67–72 (2019). doi: 10.1038/s41586-018-0808-5; pmid: 30568304
35. Z.-J. Jia, S. Gao, F. H. Arnold, Enzymatic Primary Amination of Benzylic and Allylic C(sp ^{3})-H Bonds. *J. Am. Chem. Soc.* **142**, 10279–10283 (2020). doi: 10.1021/jacs.0c03428; pmid: 32450692
36. J. Huang *et al*., Complete integration of carbene-transfer chemistry into biosynthesis. *Nature* **617**, 403–408 (2023). doi: 10.1038/s41586-023-06027-2; pmid: 37138074
37. S. Gergel *et al*., Engineered cytochrome P450 for direct arylalkene-to-ketone oxidation via highly reactive carbocation intermediates. *Nat. Catal.* **6**, 606–617 (2023). doi: 10.1038/s41929-023-00979-4
38. Q. Zhou, M. Chin, Y. Fu, P. Liu, Y. Yang, Stereodivergent atom-transfer radical cyclization by engineered cytochromes P450. *Science* **374**, 1612–1616 (2021). doi: 10.1126/science.abk1603; pmid: 34941416
39. W. Fu *et al*., Enzyme-controlled stereoselective radical cyclization to arenes enabled by metalloredox biocatalysis. *Nat. Catal.* **6**, 628–636 (2023). doi: 10.1038/s41929-023-00986-5
40. H. Fu *et al*., Ground-State Electron Transfer as an Initiation Mechanism for Biocatalytic C-C Bond Forming Reactions. *J. Am. Chem. Soc.* **143**, 9622–9629 (2021). doi: 10.1021/jacs.1c04334; pmid: 34114803
41. T. M. Hedison, D. J. Heyes, N. S. Scrutton, Making molecules with photodecarboxylases: A great start or a false dawn? *Curr. Res. Chem. Biol.* **2**, 100017 (2022). doi: 10.1016/j.crchbi.2021.100017
42. A. Sancar, Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev.* **103**, 2203–2237 (2003). doi: 10.1021/cr0204348; pmid: 12797829
43. P. Vedalankar, B. C. Tripathy, Evolution of light-independent protochlorophyllide oxidoreductase. *Protoplasma* **256**, 293–312 (2019). doi: 10.1007/s00709-018-1317-y; pmid: 30291443
44. M. A. Emmanuel *et al*., Photobiocatalytic Strategies for Organic Synthesis. *Chem. Rev.* **123**, 5459–5520 (2023). doi: 10.1021/acs.chemrev.2c00767; pmid: 37115521
45. W. Zhang *et al*., Photobiocatalytic synthesis of chiral secondary fatty alcohols from renewable unsaturated fatty acids. *Nat. Commun.* **11**, 2258 (2020). doi: 10.1038/s41467-020-16099-7; pmid: 32382158
46. K. F. Biegasiewicz *et al*., Photoexcitation of flavoenzymes enables a stereoselective radical cyclization. *Science* **364**, 1166–1169 (2019). doi: 10.1126/science.aaw1143; pmid: 31221855
47. C. G. Page *et al*., Regioselective Radical Alkylation of Arenes Using Evolved Photoenzymes. *J. Am. Chem. Soc.* **145**, 11866–11874 (2023). doi: 10.1021/jacs.3c03607; pmid: 37199445
48. H. Fu *et al*., An asymmetric sp ^{3} -sp ^{3} cross-electrophile coupling using 'ene'-reductases. *Nature* **610**, 302–307 (2022). doi: 10.1038/s41586-022-05167-1; pmid: 35952713
49. Y. Ye *et al*., Using enzymes to tame nitrogen-centred radicals for enantioselective hydroamination. *Nat. Chem.* **15**, 206–212 (2023). doi: 10.1038/s41557-022-01083-z; pmid: 36376390
50. K. F. Biegasiewicz, S. J. Cooper, M. A. Emmanuel, D. C. Miller, T. K. Hyster, Catalytic promiscuity enabled by photoredox catalysis in nicotinamide-dependent oxidoreductases. *Nat. Chem.* **10**, 770–775 (2018). doi: 10.1038/s41557-018-0059-y; pmid: 29892028
51. L. Cheng *et al*., Stereoselective amino acid synthesis by synergistic photoredox-pyridoxal radical biocatalysis. *Science* **381**, 444–451 (2023). doi: 10.1126/science.adg2420; pmid: 37499030
52. S. L. Lovelock *et al*., The road to fully programmable protein catalysis. *Nature* **606**, 49–58 (2022). doi: 10.1038/s41586-022-04456-z; pmid: 35650353
53. R. Obexer *et al*., Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat. Chem.* **9**, 50–56 (2017). doi: 10.1038/nchem.2596; pmid: 27995916
54. A. H.-W. Yeh *et al*., De novo design of luciferases using deep learning. *Nature* **614**, 774–780 (2023). doi: 10.1038/s41586-023-05696-3; pmid: 36813896
55. S. L. Beaucage, M. H. Caruthers, Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett.* **22**, 1859–1862 (1981). doi: 10.1016/S0040-4039(01)90461-7
56. A. Hoose, R. Vellacott, M. Storch, P. S. Freemont, M. G. Ryadnov, DNA synthesis technologies to close the gene writing gap. *Nat. Rev. Chem.* **7**, 144–161 (2023). doi: 10.1038/s41570-022-00456-9; pmid: 36714378
57. D. Verardo *et al*., Multiplex enzymatic synthesis of DNA with single-base resolution. *Sci. Adv.* **9**, eadi0263 (2023). doi: 10.1126/sciadv.adi0263; pmid: 37418522
58. S. Palluk *et al*., De novo DNA synthesis using polymerase-nucleotide conjugates. *Nat. Biotechnol.* **36**, 645–650 (2018). doi: 10.1038/nbt.4173; pmid: 29912208
59. I. Sarac, M. Hollenstein, Terminal Deoxynucleotidyl Transferase in the Synthesis and Modification of Nucleic Acids. *ChemBioChem* **20**, 860–871 (2019). doi: 10.1002/cbic.201800658; pmid: 30451377
60. C. Ching *et al*, Codexis Inc., World Intellectual Property Organization Patent Application WO 2023/069865, US Patent Application PCT/US2022/078071 (2023).
61. M. Friedrich, A. Aigner, Therapeutic siRNA: State-of-the-Art and Future Perspectives. *BioDrugs* **36**, 549–571 (2022). doi: 10.1007/s40259-022-00549-3; pmid: 35997897
62. E. R. Moody, R. Obexer, F. Nickl, R. Spiess, S. L. Lovelock, An enzyme cascade enables production of therapeutic oligonucleotides in a single operation. *Science* **380**, 1150–1154 (2023). doi: 10.1126/science.add5892; pmid: 37319201
63. D. J. Wiegand *et al*., A platform for controlled template-independent enzymatic synthesis of RNA oligonucleotides and therapeutics. bioRxiv 2023.06.2029.547106 [Preprint] (2023). doi: 10.1101/2023.06.29.547106
64. S. Kosuri, G. M. Church, Large-scale de novo DNA synthesis: Technologies and applications. *Nat. Methods* **11**, 499–507 (2014). doi: 10.1038/nmeth.2918; pmid: 24781323
65. B. P. Kuiper, R. C. Prins, S. Billerbeck, Oligo Pools as an Affordable Source of Synthetic DNA for Cost-Effective Library Construction in Protein- and Metabolic Pathway Engineering.

ChemBioChem **23**, e202100507 (2022). doi: 10.1002/cbic.202100507; pmid: 34817110
66. D. Patsch, M. Eichenberger, M. Voss, U. T. Bornscheuer, R. M. Buller, LibGENiE - A bioinformatic pipeline for the design of information-enriched enzyme libraries. *Comput. Struct. Biotechnol. J.* **21**, 4488–4496 (2023). doi: 10.1016/j.csbj.2023.09.013; pmid: 37736300
67. S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem. Int. Ed.* **60**, 88–119 (2021). doi: 10.1002/anie.202006648; pmid: 32558088
68. D. Yi *et al*., Recent trends in biocatalysis. *Chem. Soc. Rev.* **50**, 8003–8049 (2021). doi: 10.1039/D0CS01575J; pmid: 34142684
69. J. C. Moore, F. H. Arnold, Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. *Nat. Biotechnol.* **14**, 458–467 (1996). doi: 10.1038/nbt0496-458; pmid: 9630920
70. J. C. Moore *et al*., "Site and Mutation"-Specific Predictions Enable Minimal Directed Evolution Libraries. *ACS Synth. Biol.* **7**, 1730–1741 (2018). doi: 10.1021/acssynbio.7b00359; pmid: 29782150
71. A. D. Silverman, A. S. Karim, M. C. Jewett, Cell-free gene expression: An expanded repertoire of applications. *Nat. Rev. Genet.* **21**, 151–170 (2020). doi: 10.1038/s41576-019-0186-3; pmid: 31780816
72. R. S. Molina *et al*., In vivo hypermutation and continuous evolution. *Nat. Rev. Methods Primers* **2**, 36 (2022). doi: 10.1038/s43586-022-00119-5; pmid: 37073402
73. X. Yi, J. Khey, R. J. Kazlauskas, M. Travisano, Plasmid hypermutation using a targeted artificial DNA replisome. *Sci. Adv.* **7**, eabg8712 (2021). doi: 10.1126/sciadv.abg8712; pmid: 34272238
74. B. Seelig, J. W. Szostak, Selection and evolution of enzymes from a partially randomized non-catalytic scaffold. *Nature* **448**, 828–831 (2007). doi: 10.1038/nature06032; pmid: 17700701
75. M. Gantz, S. Neun, E. J. Medcalf, L. D. van Vliet, F. Hollfelder, Ultrahigh-Throughput Enzyme Engineering and Discovery in *In Vitro* Compartments. *Chem. Rev.* **123**, 5571–5611 (2023). doi: 10.1021/acs.chemrev.2c00910; pmid: 37126602
76. A. Debon *et al*., Ultrahigh-throughput screening enables efficient single-round oxidase remodelling. *Nat. Catal.* **2**, 740–747 (2019). doi: 10.1038/s41929-019-0340-5
77. A. E. Mulelu, A. M. Kirykowicz, J. D. Woodward, Cryo-EM and directed evolution reveal how *Arabidopsis* nitrilase specificity is influenced by its quaternary structure. *Commun. Biol.* **2**, 260 (2019). doi: 10.1038/s42003-019-0505-4; pmid: 31341959
78. E. Danelius, N. J. Porter, J. Unge, F. H. Arnold, T. Gonen, MicroED Structure of a Protoglobin Reactive Carbene Intermediate. *J. Am. Chem. Soc.* **145**, 7159–7165 (2023). doi: 10.1021/jacs.2c12004; pmid: 36948184
79. F. Agostini *et al*., Biocatalysis with Unnatural Amino Acids: Enzymology Meets Xenobiology. *Angew. Chem. Int. Ed.* **56**, 9680–9703 (2017). doi: 10.1002/anie.201610129; pmid: 28085996
80. A. P. Green, T. Hayashi, P. R. Mittl, D. Hilvert, A Chemically Programmed Proximal Ligand Enhances the Catalytic Properties of a Heme Enzyme. *J. Am. Chem. Soc.* **138**, 11344–11352 (2016). doi: 10.1021/jacs.6b07029; pmid: 27500802
81. M. Ortmayer *et al*., Rewiring the "Push-Pull" Catalytic Machinery of a Heme Enzyme Using an Expanded Genetic Code. *ACS Catal.* **10**, 2735–2746 (2020). doi: 10.1021/acscatal.9b05129; pmid: 32550044
82. M. Ortmayer *et al*., A Noncanonical Tryptophan Analogue Reveals an Active Site Hydrogen Bond Controlling Ferryl Reactivity in a Heme Peroxidase. *JACS Au* **1**, 913–918 (2021). doi: 10.1021/jacsau.1c00145; pmid: 34337604
83. M. Pott *et al*., A Noncanonical Proximal Heme Ligand Affords an Efficient Peroxidase in a Globin Fold. *J. Am. Chem. Soc.* **140**, 1535–1543 (2018). doi: 10.1021/jacs.7b12621; pmid: 29309143
84. D. M. Carminati, R. Fasan, Stereoselective Cyclopropanation of Electron-Deficient Olefins with a Cofactor Redesigned Carbene Transferase Featuring Radical Reactivity. *ACS Catal.* **9**, 9683–9697 (2019). doi: 10.1021/acscatal.9b02272; pmid: 32257582
85. T. Hayashi *et al*., Capture and characterization of a reactive haem–carbenoid complex in an artificial metalloenzyme. *Nat. Catal.* **1**, 578–584 (2018). doi: 10.1038/s41929-018-0105-6
86. R. Crawshaw *et al*., Engineering an efficient and enantioselective enzyme for the Morita-Baylis-Hillman reaction. *Nat. Chem.* **14**, 313–320 (2022). doi: 10.1038/s41557-021-00833-9; pmid: 34916595
87. G. Kang, A. T. Taguchi, J. Stubbe, C. L. Drennan, Structure of a trapped radical transfer pathway within a ribonucleotide reductase holocomplex. *Science* **368**, 424–427 (2020). doi: 10.1126/science.aba6794; pmid: 32217749
88. C.-Y. Lin *et al*., Use of Noncanonical Tyrosine Analogues to Probe Control of Radical Intermediates during Endoperoxide Installation by Verruculogen Synthase (FtmOx1). *ACS Catal.* **12**, 6968–6979 (2022). doi: 10.1021/acscatal.2c01037; pmid: 37744570
89. J. S. Trimble *et al*., A designed photoenzyme for enantioselective [2+2] cycloadditions. *Nature* **611**, 709–714 (2022). doi: 10.1038/s41586-022-05335-3; pmid: 36130727
90. J. Jumper *et al*., Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). doi: 10.1038/s41586-021-03819-2; pmid: 34265844
91. M. Baek *et al*., Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021). doi: 10.1126/science.abj8754; pmid: 34282049
92. Z. Lin *et al*., Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023). doi: 10.1126/science.ade2574; pmid: 36927031
93. G. Casadevall, C. Duran, S. Osuna, AlphaFold2 and Deep Learning for Elucidating Enzyme Conformational Flexibility and Its Application for Design. *JACS Au* **3**, 1554–1562 (2023). doi: 10.1021/jacsau.3c00188; pmid: 37388680
94. M. L. Hekkelman, I. de Vries, R. P. Joosten, A. Perrakis, AlphaFill: Enriching AlphaFold models with ligands and cofactors. *Nat. Methods* **20**, 205–213 (2023). doi: 10.1038/s41592-022-01685-y; pmid: 36424442
95. G. Corso, H. Stärk, B. Jing, R. Barzilay, T. Jaakkola, DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. arXiv:2210.01776v2 [q-bio.BM] (2023).
96. R. A. Stein, H. S. Mchaourab, SPEACH_AF: Sampling protein ensembles and conformational heterogeneity with Alphafold2. *PLOS Comput. Biol.* **18**, e1010483 (2022). doi: 10.1371/journal.pcbi.1010483; pmid: 35994486
97. G. Casadevall, C. Duran, M. Estévez-Gay, S. Osuna, Estimating conformational heterogeneity of tryptophan synthase with a template-based Alphafold2 approach. *Protein Sci.* **31**, e4426 (2022). doi: 10.1002/pro.4426; pmid: 36173176
98. I. Anishchenko *et al*., De novo protein design by deep network hallucination. *Nature* **600**, 547–552 (2021). doi: 10.1038/s41586-021-04184-w; pmid: 34853475
99. J. Dauparas *et al*., Robust deep learning-based protein sequence design using ProteinMPNN. *Science* **378**, 49–56 (2022). doi: 10.1126/science.add2187; pmid: 36108050
100. J. Wang *et al*., Scaffolding protein functional sites using deep learning. *Science* **377**, 387–394 (2022). doi: 10.1126/science.abn2100; pmid: 35862514
101. B. I. M. Wicky *et al*., Hallucinating symmetric protein assemblies. *Science* **378**, 56–61 (2022). doi: 10.1126/science.add1964; pmid: 36108048
102. A. Zanghellini *et al*., New algorithms and an in silico benchmark for computational enzyme design. *Protein Sci.* **15**, 2785–2794 (2006). doi: 10.1110/ps.062353106; pmid: 17132862
103. B. Kuhlman *et al*., Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, 1364–1368 (2003). doi: 10.1126/science.1089427; pmid: 14631033
104. J. Meiler, D. Baker, ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility. *Proteins* **65**, 538–548 (2006). doi: 10.1002/prot.21086; pmid: 16972285
105. R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-Resolution Image Synthesis with Latent Diffusion Models. arXiv:2112.10752v2 [cs.CV] (2021).
106. J. L. Watson *et al*., Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models. *bioRxiv*2022.2012.2009.519842 [Preprint] (2022). https://doi.org/10.1101/2022.12.09.519842.
107. A. Madani *et al*., Large language models generate functional protein sequences across diverse families. *Nat. Biotechnol.* **41**, 1099–1106 (2023). doi: 10.1038/s41587-022-01618-2; pmid: 36702895
108. A. M. Brzozowski *et al*., A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature* **351**, 491–494 (1991). doi: 10.1038/351491a0; pmid: 2046751
109. C. Curado-Carballada, F. Feixas, J. Iglesias-Fernández, S. Osuna, Hidden Conformations in *Aspergillus niger* Monoamine Oxidase are Key for Catalytic Efficiency. *Angew. Chem. Int. Ed.* **58**, 3097–3101 (2019). doi: 10.1002/anie.201812532; pmid: 30600584
110. V. Gligorijević *et al*., Structure-based protein function prediction using graph convolutional networks. *Nat. Commun.* **12**, 3168 (2021). doi: 10.1038/s41467-021-23303-9; pmid: 34039967
111. T. Yu *et al*., Enzyme function prediction using contrastive learning. *Science* **379**, 1358–1363 (2023). doi: 10.1126/science.adf2465; pmid: 36996195
112. S. Lauterbach *et al*., EnzymeML: Seamless data flow and modeling of enzymatic data. *Nat. Methods* **20**, 400–402 (2023). doi: 10.1038/s41592-022-01763-1; pmid: 36759590
113. S. Gelman, S. A. Fahlberg, P. Heinzelman, P. A. Romero, A. Gitter, Neural networks to learn protein sequence-function relationships from deep mutational scanning data. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104878118 (2021). doi: 10.1073/pnas.2104878118; pmid: 34815338
114. K. K. Yang, Z. Wu, F. H. Arnold, Machine-learning-guided directed evolution for protein engineering. *Nat. Methods* **16**, 687–694 (2019). doi: 10.1038/s41592-019-0496-6; pmid: 31308553
115. Z. Wu, S. B. J. Kan, R. D. Lewis, B. J. Wittmann, F. H. Arnold, Machine learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8852–8858 (2019). doi: 10.1073/pnas.1901979116; pmid: 30979809
116. E. J. Ma *et al*., Machine-Directed Evolution of an Imine Reductase for Activity and Stereoselectivity. *ACS Catal.* **11**, 12433–12445 (2021). doi: 10.1021/acscatal.1c02786
117. S. Osuna, The challenge of predicting distal active site mutations in computational enzyme design. *WIREs Comput. Mol. Sci.* **11**, e1502 (2021). doi: 10.1002/wcms.1502
118. M. A. Maria-Solano, T. Kinateder, J. Iglesias-Fernández, R. Sterner, S. Osuna, *In Silico* Identification and Experimental Validation of Distal Activity-Enhancing Mutations in Tryptophan Synthase. *ACS Catal.* **11**, 13733–13743 (2021). doi: 10.1021/acscatal.1c03950; pmid: 34777912
119. F. Rudroff *et al*., Opportunities and challenges for combining chemo- and biocatalysis. *Nat. Catal.* **1**, 12–22 (2018). doi: 10.1038/s41929-017-0010-4
120. J. A. McIntosh *et al*., A kinase-cGAS cascade to synthesize a therapeutic STING activator. *Nature* **603**, 439–444 (2022). doi: 10.1038/s41586-022-04422-9; pmid: 35296845
121. J. A. McIntosh *et al*., Engineered Ribosyl-1-Kinase Enables Concise Synthesis of Molnupiravir, an Antiviral for COVID-19. *ACS Cent. Sci.* **7**, 1980–1985 (2021). doi: 10.1021/acscentsci.1c00608; pmid: 34963891
122. M. A. Huffman *et al*., Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019). doi: 10.1126/science.aay8484; pmid: 31806816
123. C. Greunke, A. Glöckle, J. Antosch, T. A. M. Gulder, Biocatalytic Total Synthesis of Ikarugamycin. *Angew. Chem. Int. Ed.* **56**, 4351–4355 (2017). doi: 10.1002/anie.201611063; pmid: 28128495
124. A. Fryszkowska *et al*., A chemoenzymatic strategy for site-selective functionalization of native peptides and proteins. *Science* **376**, 1321–1327 (2022). doi: 10.1126/science.abn2009; pmid: 35709255
125. T. Schwander, L. Schada von Borzyskowski, S. Burgener, N. S. Cortina, T. J. Erb, A synthetic pathway for the fixation of carbon dioxide in vitro. *Science* **354**, 900–904 (2016). doi: 10.1126/science.aah5237; pmid: 27856910
126. T. Cai *et al*., Cell-free chemoenzymatic starch synthesis from carbon dioxide. *Science* **373**, 1523–1527 (2021). doi: 10.1126/science.abh4049; pmid: 34554807
127. E. Eichhorn, B. Schilling, D. Wahler, L. Fourage, E. Locher, Givaudan SA, World Intellectual Property Organization Patent WO 2016/170099 (2015).
128. E. Eichhorn, C. Baumgartner, M. Biermann, Biotechnology – A Tool to Transform Givaudan's Fragrance Ingredients Palette. *Chimia* **77**, 384–389 (2023). doi: 10.2533/chimia.2023.384
129. E. Eichhorn, C. Ullmann, Givaudan SA, World Intellectual Property Organization Patent Application WO 2021/110848 (2019).
130. S. McKenna *et al*., Bristol Myers Squibb, World Intellectual Property Organization Patent Application WO 2023/178061 (2023).

131. C. Ching D. Elgart *et al*., Codexis, World Intellectual Property Organization Patent Application WO 2023/150568 (2023).
132. E. J. Corey, Robert Robinson Lecture. Retrosynthetic thinking—Essentials and examples. *Chem. Soc. Rev.* **17**, 111–133 (1988). doi: 10.1039/CS9881700111
133. S. Warren, P. Wyatt, *Organic Synthesis: The Disconnection Approach* (Wiley-Blackwell, ed. 2, 2008).
134. W. Finnigan, L. J. Hepworth, S. L. Flitsch, N. J. Turner, RetroBioCat as a computer-aided synthesis planning tool for biocatalytic reactions and cascades. *Nat. Catal.* **4**, 98–104 (2021). doi: 10.1038/s41929-020-00556-z; pmid: 33604511
135. R. O. M. A. de Souza, L. S. M. Miranda, U. T. Bornscheuer, A Retrosynthesis Approach for Biocatalysis in Organic Synthesis. *Chemistry* **23**, 12040–12063 (2017). doi: 10.1002/chem.201702235; pmid: 28514518
136. M. Hönig, P. Sondermann, N. J. Turner, E. M. Carreira, Enantioselective Chemo- and Biocatalysis: Partners in Retrosynthesis. *Angew. Chem. Int. Ed.* **56**, 8942–8973 (2017). doi: 10.1002/anie.201612462; pmid: 28407390
137. N. J. Turner, L. Humphreys, *Biocatalysis in Organic Synthesis: The Retrosynthesis Approach* (Royal Society of Chemistry, 2018).
138. N. J. Turner, E. O'Reilly, Biocatalytic retrosynthesis. *Nat. Chem. Biol.* **9**, 285–288 (2013). doi: 10.1038/nchembio.1235; pmid: 23594772
139. T. Yu *et al*., Machine learning-enabled retrobiosynthesis of molecules. *Nat. Catal.* **6**, 137–151 (2023). doi: 10.1038/s41929-022-00909-w
140. H. Yim *et al*., Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **7**, 445–452 (2011). doi: 10.1038/nchembio.580; pmid: 21602812
141. C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **46**, 3–26 (2001). doi: 10.1016/S0169-409X(00)00129-0; pmid: 11259830
142. A. Debon, E. Siirola, R. Snajdrova, Enzymatic Bioconjugation: A Perspective from the Pharmaceutical Industry. *JACS Au* **3**, 1267–1283 (2023). doi: 10.1021/jacsau.2c00617; pmid: 37234110
143. A. Ali *et al*., European Medicinal Chemistry Leaders in Industry (EMCL) - On the Status and Future of Medicinal Chemistry Research in Europe. *ChemMedChem* **18**, e202300127 (2023). doi: 10.1002/cmdc.202300127; pmid: 37276375
144. B. I. Andrews *et al*., Sustainability Challenges and Opportunities in Oligonucleotide Manufacturing. *J. Org. Chem.* **86**, 49–61 (2021). doi: 10.1021/acs.joc.0c02291; pmid: 33253568
145. A. Crameri *et al*., Glaxo Smith Kline, World Intellectual Property Organization Patent WO 2018/011067 A2 (2018).
146. J. Nandakumar, S. Shuman, C. D. Lima, RNA ligase structures reveal the basis for RNA specificity and conformational changes that drive ligation forward. *Cell* **127**, 71–84 (2006). doi: 10.1016/j.cell.2006.08.038; pmid: 17018278
147. G. Mann *et al*., Biocatalytic assembly of chemically modified oligonucleotides. *Tetrahedron Lett.* **93**, 153696 (2022). doi: 10.1016/j.tetlet.2022.153696
148. V. Tournier *et al*., Enzymes' Power for Plastics Degradation. *Chem. Rev.* **123**, 5612–5701 (2023). doi: 10.1021/acs.chemrev.2c00644; pmid: 36916764
149. V. Tournier *et al*., An engineered PET depolymerase to break down and recycle plastic bottles. *Nature* **580**, 216–219 (2020). doi: 10.1038/s41586-020-2149-4; pmid: 32269349
150. H. Lu *et al*., Machine learning-aided engineering of hydrolases for PET depolymerization. *Nature* **604**, 662–667 (2022). doi: 10.1038/s41586-022-04599-z; pmid: 35478237
151. Y. Branson *et al*., Urethanases for the Enzymatic Hydrolysis of Low Molecular Weight Carbamates and the Recycling of Polyurethanes. *Angew. Chem. Int. Ed.* **62**, e202216220 (2023). doi: 10.1002/anie.202216220; pmid: 36591907
152. D. L. Hughes, Biocatalysis in Drug Development—Highlights of the Recent Patent Literature. *Org. Process Res. Dev.* **22**, 1063–1080 (2018). doi: 10.1021/acs.oprd.8b00232
153. S. González-Granda, J. Albarrán-Velo, I. Lavandera, V. Gotor-Fernández, Expanding the Synthetic Toolbox through Metal-Enzyme Cascade Reactions. *Chem. Rev.* **123**, 5297–5346 (2023). doi: 10.1021/acs.chemrev.2c00454; pmid: 36626572
154. H. Gröger, F. Gallou, B. H. Lipshutz, Where Chemocatalysis Meets Biocatalysis: In Water. *Chem. Rev.* **123**, 5262–5296 (2023). doi: 10.1021/acs.chemrev.2c00416; pmid: 36473100
155. J. B. Siegel *et al*., Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* **329**, 309–313 (2010). doi: 10.1126/science.1190239; pmid: 20647463
156. A. Crossley *et al*., A non-canonical nucleophile unlocks a new mechanistic pathway in a designed enzyme. *Research Square*, [Preprint] (2023). doi: https://doi.org/10.21203/rs.3.rs-2922796/v1
157. W. C. Hallows *et al*., Optimizing human α-galactosidase for treatment of Fabry disease. *Sci. Rep.* **13**, 4748 (2023). doi: 10.1038/s41598-023-31777-4; pmid: 36959353

ACKNOWLEDGMENTS

We thank T. Schwander and F. Özgen for help with figure preparation. **Funding:** R.J.K. thanks the NSF (CBET-2039039 and CBET-1930825), and U.T.B. thanks the German Research Foundation, DFG (NFDI4CAT and Bo1862/25-1). R.M.B. thanks NCCR Catalysis, a National Centre of Competence in Research funded by the Swiss National Science Foundation (Grant no. 180544). **Competing interests:** U.T.B. is a shareholder and head of the Supervisory Board of Enzymicals AG (Greifswald, Germany), head of the Advisory Board of the Toulouse White Biotechnology center (Toulouse, France), and a member of the Advisory Board of Carbios (Paris, France). R.S. is an employee of Novartis Pharma AG (Basel, Switzerland). S.L. is an employee of Codexis. (Redwood, CA, USA). J.C.M. is an employee of Merck Sharp & Dohme, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA, and is an Advisory Board member of EU Horizon's EVOdrops program. R.K. has served as a consultant to Sasya, a biotechnology company. R.K. is a co-inventor on a patent related to work cited (*73*). R.B. is a member of the Research Oversight Commitee Genome Canada Project, "Optimizing a microbial platform to break down and valorize waste plastic," a board member of the Swiss Academy of Natural Sciences - Chemistry Platform, a board member of the working group "Biotransformations" (VAAM/DECHEMA), a Steering Board member of the National Competence Center of Research "Catalysis," and Swiss Delegate of the European Section of Applied Biocatalysis (ESAB) Scientific Commitee.

Submitted 14 July 2023; accepted 17 October 2023
10.1126/science.adh8615

From nature to industry: Harnessing enzymes for biocatalysis

R. Buller, S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore, and U. T. Bornscheuer

Science **382** (6673), eadh8615. DOI: 10.1126/science.adh8615

Editor's summary
As nature's catalysts, enzymes are adaptable, specific, and interoperable. There have been important advances in the identification of transformations catalyzed by enzymes, and new methods have been developed for identifying, engineering, or creating de novo suitable enzymes for a wide range of applications. In a review, Buller *et al*. collect this recent progress in biocatalysis research and look ahead toward increasing the use of computational tools, accelerating design test cycles, and expanding chemistry beyond what is familiar in nature. Engineered enzymes provide a green chemistry solution to the synthesis of complex organic molecules and can be used in the degradation of waste plastic and chemicals. —Michael A. Funk