PAPER

Spiers Memorial Lecture: Engineering biocatalysts

Donald Hilvert [ID]

Enzymes are being engineered to catalyze chemical reactions for many practical applications in chemistry and biotechnology. The approaches used are surveyed in this short review, emphasizing methods for accessing reactivities not expressed by native protein scaffolds. The successful generation of completely *de novo* enzymes that rival the rates and selectivities of their natural counterparts highlights the potential role that designer enzymes may play in the coming years in research, industry, and medicine. Some challenges that need to be addressed to realize this ambitious dream are considered together with possible solutions.

Introduction

For millennia, mankind has harnessed the power of enzymes for baking, brewing and sundry other purposes. Although such applications remain relevant, our understanding of what enzymes are and how they function has advanced substantially. Today, these agents serve as highly active and selective catalysts in many areas of biology and chemistry.

The properties of enzymes that make life possible—impressive catalytic efficiency and precision—lend themselves to applications far beyond their normal biological roles. Ready access to enzymes capable of copying, cutting, and splicing DNA enabled the biotechnological revolution, for example. Other enzymes are employed directly as therapeutic agents to treat diseases like leukemia (asparaginase[1]), severe combined immunodeficiency (adenosine deaminase[2]) and phenylketonuria (phenylalanine hydroxylase[3]). Biocatalysts have also been recognized as pivotal for a greener, more sustainable chemical industry.[4,5]

Although natural enzymes are remarkable catalysts, they are often poorly suited for non-natural applications due to challenges in production, substrate specificity, or stability. Robust methods are consequently needed to tailor their properties or, more ambitiously, create biocatalysts with entirely new activities. In this short review, common methods for tackling these issues are highlighted.

Laboratory of Organic Chemistry, ETH Zürich, 8093 Zürich, Switzerland. E-mail: hilvert@ethz.ch

Enzyme engineering and evolution

The success of an enzyme application hinges on identifying a suitable catalyst for the reaction of interest. Genomic databases are one source of natural enzymes; strain collections and metagenomic libraries are another.[4,5] For some applications, these proteins can be implemented directly. For others, engineering may be required. If structural and mechanistic information is available, rational approaches can successfully modify their properties. Often, a few mutations suffice to achieve the desired features.[6] In other instances, extensive remodeling of the enzyme may be necessary.

Approaches that mimic Darwinian evolution are well suited for demanding engineering challenges (Fig. 1).[7–11] In a directed evolution experiment, the gene for the enzyme of interest is randomized to create DNA libraries. Following *in vitro* or *in vivo* production of the encoded variants, those with improved properties are identified by screening or selection, and their respective genes are amplified for characterization or further diversification. As in natural evolution, beneficial mutations accumulate over iterative mutagenesis and screening/selection cycles, resulting in gradual catalyst optimization.

Over the last three decades, the rates and substrate preferences of many bio-catalysts have been engineered by directed evolution. This methodology is also useful for enhancing a protein's physical characteristics, including expression yield, thermal stability, and tolerance to organic solvents, which may be important for large-scale manufacturing processes. All that is typically needed for a successful evolutionary campaign is for the starting protein to exhibit at least low starting activity for the desired transformation.

If non-native reactivity is targeted, the ability of an enzyme to catalyze reactions unrelated to its normal biological function, also called catalytic
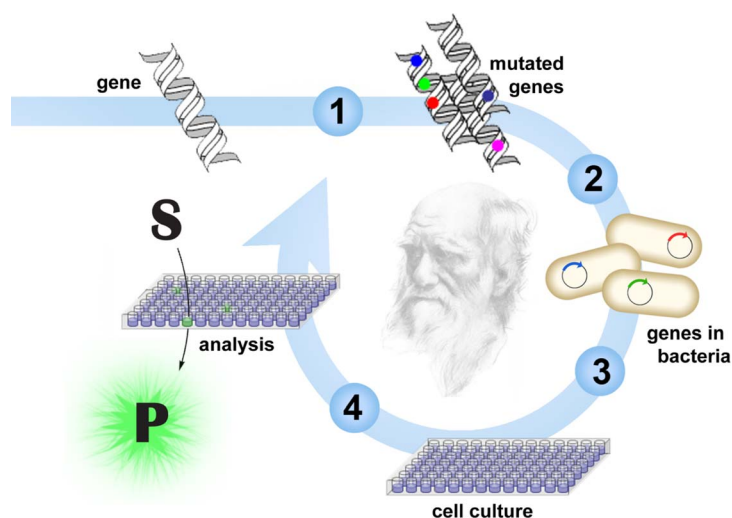
Fig. 1 Typical workflow for directed evolution. The process involves gene diversification (1), transformation of bacteria cells (2), growing up individual variants in microtiter plates (3), and screening the variants, for example with a fluorogenic substrate (4). Iterative rounds of mutagenesis and screening allow gradual optimization of catalyst properties.

This journal is © The Royal Society of Chemistry 2024

promiscuity,[12] may provide a solution. The catalysis of carbene and nitrene transfer reactions by cytochrome P450 and other heme proteins is a case in point.[13] These reactions proceed *via* metal–carbenoid or metal–nitrenoid intermediates mechanistically related to the native ferryl–oxo intermediates involved in oxygen insertion into unactivated C–H bonds. Diversification of this promiscuous activity by directed evolution has produced families of enzymes that catalyze cyclopropanations, intramolecular C–H aminations, intermolecular amination of benzylic C–H bonds, N–H insertions, aziridinations, and sulfimidations, usually with good total turnovers and excellent stereoselectivity. Biocatalytic routes to enantiopure organosilicon and organoboron compounds by heme-mediated carbon–silicon[14] and carbon–boron[15] bond formation are also notable in this context.

Common biological cofactors have been discovered to be another source of reactivity unexplored by natural evolution. For instance, the photoexcited states of NAD(P)H and FMN have been shown to facilitate electron transfer to substrates bound within enzyme active sites, enabling enantioselective radical dehalogenation reactions by NADPH-dependent ketoreductases[16] and asymmetric hydro-dehalogenations of α-bromoesters by flavin-dependent ene-reductases.[17] The mechanistic similarities between radical-relay reactions from synthetic chemistry and the native reactions of nonheme iron enzymes have similarly inspired efforts to enzymatically convert synthetic radical precursors into highly reactive species such as nitrogen-centered radicals that then undergo abiological reactions with high selectivity.[18,19] In another recent example, a pyridoxal 5′-phosphate (PLP)-dependent enzyme has been combined with photoredox catalysis to create a novel radical route to diverse noncanonical amino acids, including those containing multiple chiral centers, without the need for protecting groups.[20]

Biocatalysis for industry, medicine, and the circular economy

The ability to tailor enzyme properties reliably has made biocatalytic production of small-molecule drugs an increasingly attractive alternative to traditional synthetic chemistry. The creation of stereoselective biocatalysts to generate chiral amines, critical components of many pharmaceuticals, is illustrative. The successful redesign of a pyridoxal-dependent ω-transaminase for the industrial-scale production of sitagliptin, a blockbuster antidiabetic drug, was a watershed event in this regard (Fig. 2).[21] Not only was the enzymatic process superior to a previously established route involving a rhodium-based chiral catalyst, but the evolved enzyme displayed an unexpectedly broad substrate scope, rendering it an effective catalyst for the reductive amination of numerous prochiral ketones. Building on this success, many industrially relevant oxidoreductase enzymes for the synthesis of chiral amines, including monoamine oxidases (MAOs), amine dehydrogenases (AmDHs), imine reductases (IREDs) and reductive aminases (RedAms), have been engineered.[22,23] If product formation can be linked to a fluorescent readout, either directly or through a coupled reaction, ultrahigh-throughput microfluidic screening can dramatically accelerate the optimization process (Fig. 3).[24]
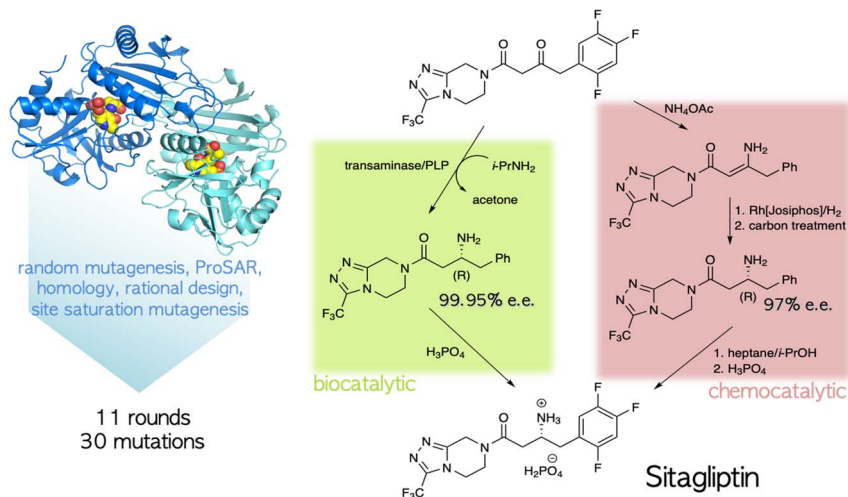
Fig. 2 Biocatalytic vs. chemocatalytic routes to the anti-diabetic medication sitagliptin.[21] The PLP-dependent transaminase was extensively mutagenized to meet the specifications of the industrial process.
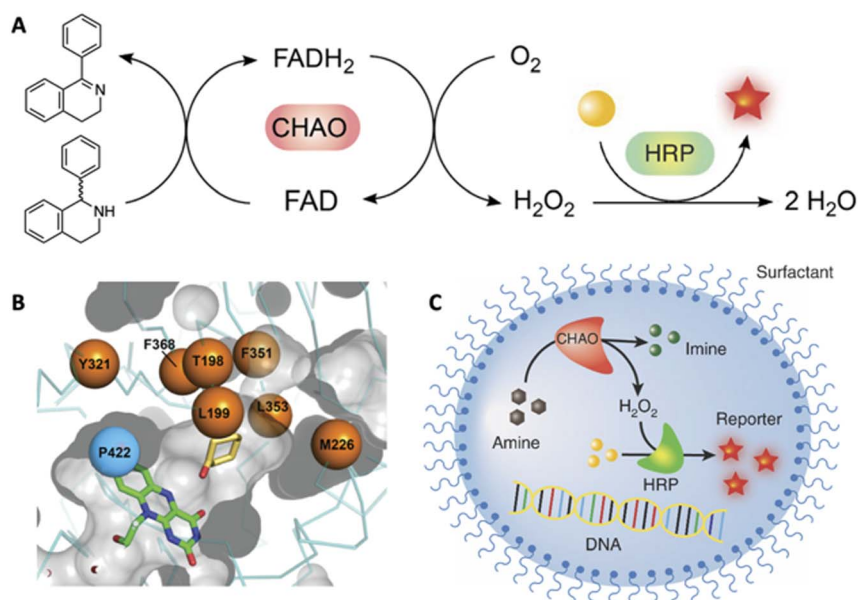

Fig. 3 High-throughput evolution of cyclohexylamine oxidase (CHAO) for the stereo-selective oxidation of 1-phenyl-1,2,3,4-tetrahydroisoquinoline, a precursor of the block-buster drug solifenacin. (A) The coupled assay used to monitor the reaction. (B) The CHAO active site showing the eight residues that were mutagenized to create a 1.7×10^6-membered library, which (C) was exhaustively screened by droplet-based microfluidics to identify a highly active and stereoselective variant.

Biocatalysis has much to offer the burgeoning field of synthetic biology as well. Efforts to expand the genetic code, for example, have relied on engineered aminoacyl-tRNA synthetases (aaRSs) that selectively aminoacylate suppressor tRNAs with more than 250 noncanonical amino acids.[25,26] Adding such enzymes to the translational machinery of bacterial and eukaryotic hosts has enabled the production of recombinant proteins containing unique chemical functionality, opening exciting avenues in protein biochemistry and cell biology. In another area, CRISPR nucleases, such as Cas9, have been engineered to improve RNA-guided genome editing technologies.[27] These efforts have afforded tools for highly specific base editing, regulation of gene expression, and genome imaging. Notably, phage-assisted continuous evolution enables the optimization of these enzymes on much faster timescales than conventional methods.

Global environmental concerns are also being addressed through biocatalysis. In this context, enzymatic degradation of plastic followed by polymerization of their constituent building blocks or valorization into other products has received considerable attention as a means of waste recycling. Natural hydrolytic enzymes that catalyze the breakdown of polyethylene terephthalate (PET) plastics have been evolved in the lab to improve both efficiency and thermostability for commercial applications. One variant, called HotPETase, rapidly depolymerizes semicrystalline PET and selectively decomposes the PET-containing components of laminated materials.[28] Microwave pre-treatment has been found to accelerate such enzymatic depolymerization reactions substantially, likely by making the polymer chains more accessible to the enzyme.[29] Recycling enzymatically depolymerized PET waste has also yielded PET polymers that exhibit the same properties as petrochemical PET,[30] supporting the feasibility of a circular carbon economy for these plastics.

Biocatalytic pathways, cascades, cells and systems

Integrating sequentially acting enzymes into synthetic reaction cascades, either *in vivo* or *in vitro*, is an effective strategy for producing complex, high-value chemicals under mild conditions.[31] Natural biosynthetic pathways have been successfully engineered to enhance the production of valuable natural products like the malaria drug artemisinin[32] (Fig. 4A) and, more recently, a plant-derived glycosylated triterpenoid used as a potent vaccine adjuvant.[33]

Biocatalysts from different organisms can even be combined to create hybrid pathways. The construction of an artificial CO_2 fixation cycle consisting of 17 enzymes from nine different organisms illustrates how complex such systems can be.[34,35] This pathway includes an efficient enoyl-CoA carboxylase/reductase for the central carboxylation step to circumvent limitations associated with natural RuBisCo-based systems. Its efficiency after optimization, which included engineering the rate-limiting enzyme, exceeded that of the Calvin cycle. Embedding such pathways in living organisms could conceivably yield whole-cell biocatalysts capable of capturing atmospheric CO_2 as a carbon feedstock.

In principle, engineered proteins can be used to construct completely artificial pathways. This was the case for a multistep enzymatic cascade developed for manufacturing the HIV drug islatravir (Fig. 4B).[36] Five of the nine component enzymes were engineered by directed evolution to recognize non-natural substrates. The final pathway, which eliminated purification steps, recycled
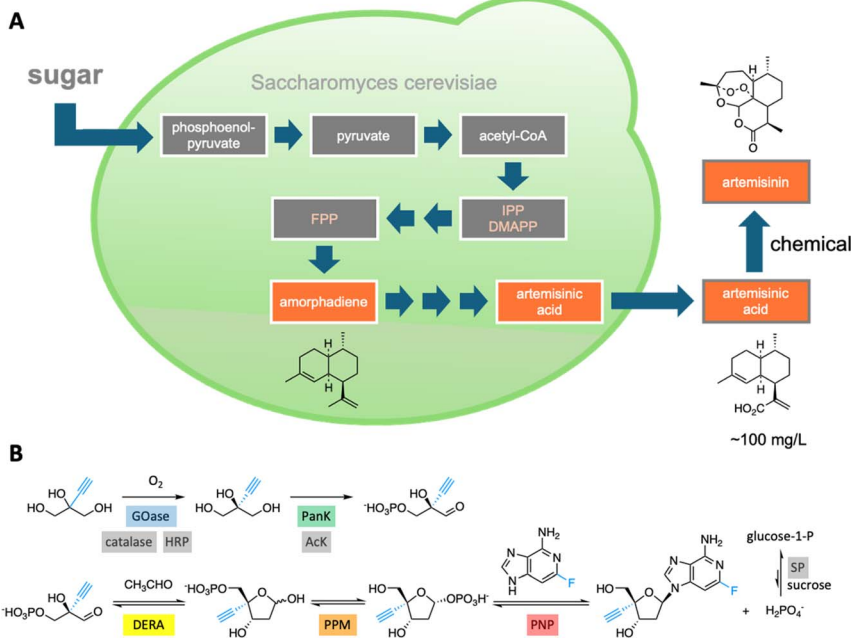
Fig. 4 Natural and artificial enzymatic cascades. (A) An engineered pathway in yeast cells for the production of artemisinic acid, a precursor of the malaria drug artemisinin.[32] (B) An *in vitro* enzymatic pathway to the HIV drug islatravir uses nine enzymes, five of which were engineered to accommodate the non-natural substrates (highlighted in color).[36]

expensive cofactors, and coupled energetically favorable and unfavorable reactions, was more efficient and substantially shorter than alternative chemical routes.

Not all biocatalytic pathways involve discrete enzymes. The biosynthesis of many bioactive natural products is coordinated by gigantic multienzyme assemblies that function in an assembly-line fashion.[37] Polyketide synthases[38] (PKSs) and nonribosomal peptide synthetases[39] (NRPSs) consist of repeating modules of several enzymatic domains, strung together like beads on a string, which are responsible for building block selection, activation, and subsequent oligomerization. Additional structural diversification can be achieved by specialized tailoring domains, and, once assembly is complete, the product is typically released by a terminal thioesterase. Because of their modularity, PKSs and NRPSs are appealing targets for engineering.

Diverse strategies for engineering these biosynthetic assembly lines have been investigated.[40] In one approach, a recently developed yeast display system for screening millions of protein variants in a single experiment has been used to reprogram the substrate specificity of NRPS adenylation[41–43] and condensation[44] domains in high throughput (Fig. 5). The ability of the modified domains to process noncanonical building blocks such as β-amino acids, α-hydroxy acids, and fatty acids without compromising catalytic efficiency highlights its efficacy. Because such high-throughput strategies are likely general, they have the potential to facilitate the sustainable production of novel natural product analogs and other bioactive compounds.
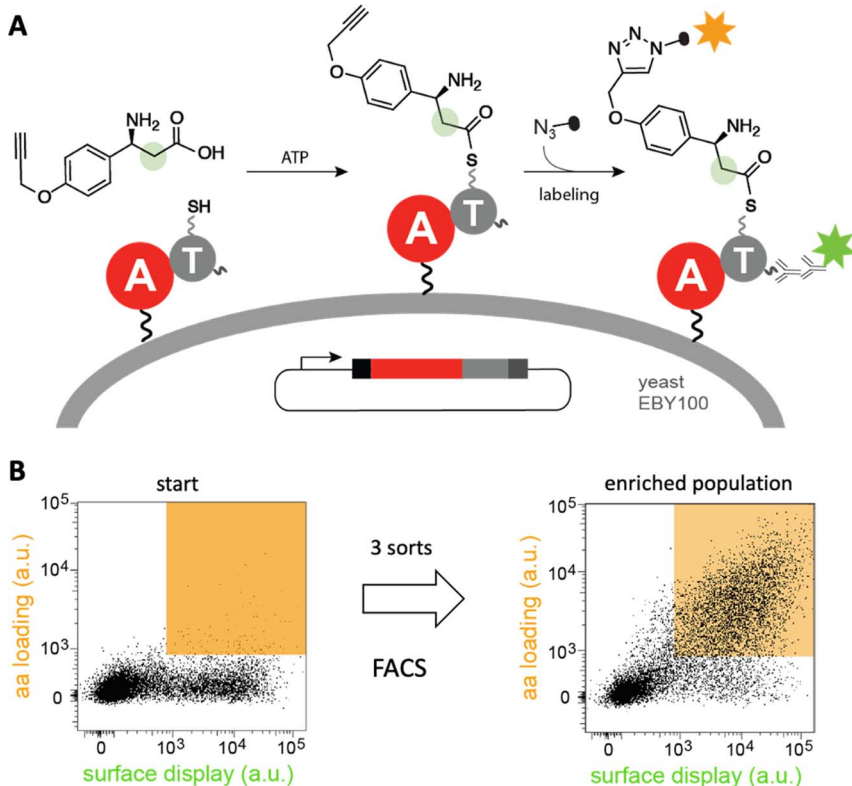
Fig. 5 High-throughput engineering of an NRPS A domain. (A) A library of adenylation (A)-thiolation (T) didomains displayed on the surface of yeast were challenged with a non-canonical β-amino acid. If successfully activated by the A domain, the substrate is transferred to the pantetheinyl cofactor of the T domain and can be detected by "click"-labeling with a fluorophore. (B) Analysis of a library of tyrocidine A variants by fluorescence activated cell sorting (FACS) shows that variants that recognize β-amino acid (orange field) become enriched after three sorts.[41]

Artificial, biomimetic, and hybrid enzymes

Conferring completely new activities on a protein scaffold is generally more difficult than augmenting weak promiscuous activity. One of the easiest ways to access novel chemistry is to incorporate reactive functional groups into a binding pocket.[45] For example, the protease subtilisin was converted into a peroxidase by chemically replacing its active-site serine with selenocysteine.[46]

Genetic code expansion has greatly extended this approach.[25,26] Recombinant incorporation of noncanonical amino acids into proteins has been used to capture transient reaction intermediates, probe mechanisms, tune metal ion reactivity, and map long-range proton and electron transfer pathways.[47] Catalysis of abiological reactions is also possible. Recent efforts in this direction include the creation of a photoenzyme that exploits triplet energy transfer to promote enantioselective [2 + 2] cycloadditions,[48] a *p*-aminophenylalanine-containing

enzyme that catalyzes condensation reactions of aldehydes with hydrazines and hydroxylamines *via* an iminium ion intermediate,[49] and another that employs a genetically encoded arylboronic acid for the kinetic resolution of hydroxyketones by oxime formation.[50] Notably, these reaction types have no equivalent in conventional homogeneous or enzyme catalysis. Moreover, because the responsible noncanonical amino acids are genetically encoded, the properties of the respective enzymes can be optimized by directed evolution.

Just as natural biochemical cofactors extend enzymatic capabilities, artificial prosthetic groups can also provide new-to-nature activities. This can be accomplished by replacing a natural cofactor like an iron heme with an abiological analog such as an iridium porphyrin. Myoglobin and P450 variants complexed with the iridium cofactor catalyze various carbene and nitrene insertion reactions,[51] including carbene insertion into unactivated C–H bonds.[52] Alternatively, cofactors can be attached either covalently or noncovalently to any protein to generate hybrid constructs, sometimes called semisynthetic enzymes,[45] that combine the inherent reactivity of the cofactor with the protein's molecular recognition properties. Binding biotinylated metal complexes to streptavidin, for example, is a rewarding route to artificial metalloenzymes for reactions like alkene hydrogenation, transfer hydrogenation of ketones, and olefin metathesis.[53] Although rate enhancements are typically modest even after evolutionary optimization, these systems often exhibit high enantioselectivity. In another approach, amino acids like (2,2′-bipyridin-5yl)alanine can be incorporated genetically into an appropriate scaffold and have been used to position metal ions within binding pockets. Artificial copper-dependent enzymes prepared in this way have been shown to catalyze Diels–Alder cycloadditions, olefin hydration reactions, and Friedel–Crafts alkylations.[54]

Computational enzyme design

State-of-the-art computational methods offer a powerful means of remodeling enzyme active sites for activities not found in nature.[55] Conceptually analogous to catalytic antibody technology—which uses synthetic haptens to elicit immunoglobulins with binding pockets configured to catalyze a chemical reaction[56,57]—computational design begins with an idealized model of a minimal active site, consisting of a quantum mechanically calculated structure of the rate-limiting transition state(s) of the target reaction plus amino acid side chains spatially arranged to stabilize it.[58] This ensemble is then docked into structurally characterized protein scaffolds *in silico*, linking the catalytic groups to the protein backbone, and repacking the active site to optimize interactions between the protein and the transition state. The resulting designs are ranked based on calculated energies, with the most promising candidates proceeding to experimental validation.

Computational enzyme design has yielded catalysts for a handful of mechanistically distinct chemical reactions (Fig. 6). These include ester hydrolysis,[59] a unimolecular proton transfer reaction,[60–63] a stereoselective Diels–Alder cycloaddition,[64,65] and multistep aldol[66–68] and Morita–Baylis–Hillman[69,70] reactions. Although the initial computational designs exhibit modest efficiency, they could be readily optimized by directed evolution. The highest-performing catalysts exhibit activities that exceed those of comparable catalytic antibodies and, in the
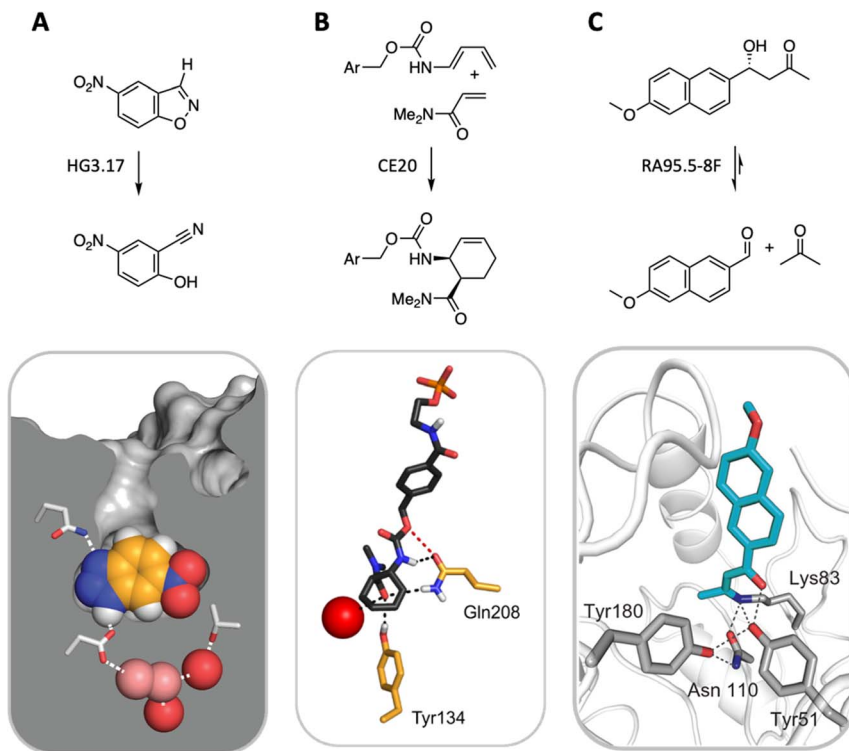
Fig. 6 Three representative examples of computationally designed enzymes that have been evolved to high activity and selectivity by iterative rounds of mutagenesis and screening. (A) A Kemp eliminase for proton transfer from carbon,[62] (B) a stereospecific Diels–Alderase,[65] and (C) a multi-step aldolase that exploits amine catalysis.[68]

best cases, even rival the speed and stereoselectivity of natural enzymes, achieving rate accelerations exceeding a billionfold and producing products as single stereoisomers on a preparative scale. Based on structural and mechanistic analyses of the "winners," the improvements achieved can be ascribed to substantial remodeling of the active site, installation of more complex arrays of functional groups, and/or minimizing unproductive states by changing protein conformation landscapes. Feeding these insights back into the design cycle may inform efforts to improve the starting computational design and thus minimize the need for time-consuming experimental optimization. Machine learning methods also offer exciting new computational tools for improving the properties of these systems.[71–75]

De novo enzymes

Natural proteins offer a diverse range of architectures for enzyme engineering. Nevertheless, their evolutionary histories and complex sequence–structure relationships may complicate efforts to implant completely new activities. Because they are evolutionarily naïve, simplified *de novo* scaffolds can potentially circumvent such issues.

Because the rules for designing helical bundles are relatively well understood, they were among the first *de novo* scaffolds to serve as scaffolds for both receptors and catalysts.[76] The assembly of heme-binding four-helix bundles that mimic the function of natural oxygen-binding proteins like myoglobin illustrates this approach.[77,78] Another four-helix bundle protein featuring a hydrophobic binding pocket adjacent to a di-iron center catalyzes phenol oxidation,[79] while a three-helix bundle, organized around a structural Hg(II) ion and a catalytic Zn(II) ion, functions as a carbonic anhydrase mimic.[80] A zinc-containing supramolecular protein assembly was found to exhibit promiscuous β-lactamase activity.[81] Hydrolytic activity was also installed in a thermostable α-helical barrel composed of seven helices arranged around a lumenal channel by introducing a Cys-His-Glu catalytic triad on the inward face of each helix.[82]

As an alternative to rational design, combinatorial libraries of binary-patterned four-helix bundles have proven to be a rich source of new chemical activities.[83] The genetic selection of an esterase that rescues an auxotrophic *Escherichia coli* strain lacking a key gene for iron uptake in an iron-limited medium is impressive in this regard.[84,85] This catalyst provides an enzymatic function that sustains the growth of a living organism using a sequence, structure, and mechanism that are dramatically different from those selected by nature to solve this specific biochemical challenge.

Recent progress toward computational design of atomically accurate proteins has greatly expanded the palette of *de novo* scaffolds, untouched by natural evolution, that can be utilized for enzyme design.[86] In one example, retro-aldolase activity was conferred on a *de novo* eight-stranded β-barrel protein by using computational methods to install a reactive amine in a hydrogen-bonded network of functional groups, thus jump-starting innovation.[87] However, the scaffold proved difficult to evolve, likely because its small size and extreme thermal stability hindered its ability to adapt to the different intermediates along the multistep reaction pathway. A closed α helical solenoid repeat protein, designed to bind heme in a reconfigurable pocket, proved more amenable to optimization and yielded a proficient peroxidase that reacts *via* a stable high-valent ferryl intermediate.[88] More recently, a *de novo* TIM barrel scaffold that tightly binds lanthanide ions in its central cavity[89] was found to activate molecules under mild conditions by cerium photoredox catalysis, catalyzing the radical C–C bond cleavage of aromatic and aliphatic 1,2-diols in aqueous solution, including lignin surrogates, upon irradiation with visible light.[90]

A computationally designed dimeric zinc-binding peptide called MID1 (metal interface design 1)[91,92] has proven to be an unexpectedly versatile template for catalyst design (Fig. 7). Mimicking the biogenesis of primitive proteins, the two peptides were fused to generate a globular protein, and one of the two zinc sites was removed by replacing metal binding residues with non-coordinating amino acids suggested by computation.[93] The single-chain version of the MID1 protein exhibited weak esterase activity, owing to an open coordination site on its zinc ion. This activity was improved >10 000 fold over ten rounds of directed evolution to afford an enzyme that cleaved a chiral ester substrate with remarkably high catalytic efficiency ($k_{cat} = 1.6$ s^{-1}; $k_{cat}/K_M = 960\,000$ M^{-1} s^{-1}) and stereospecificity ($E_{S/R} = 1000$). Indeed, the catalyst is orders of magnitude more efficient than other designed esterases and comparable to natural hydrolytic enzymes. A high-resolution structure of the optimized catalyst in complex with a transition state
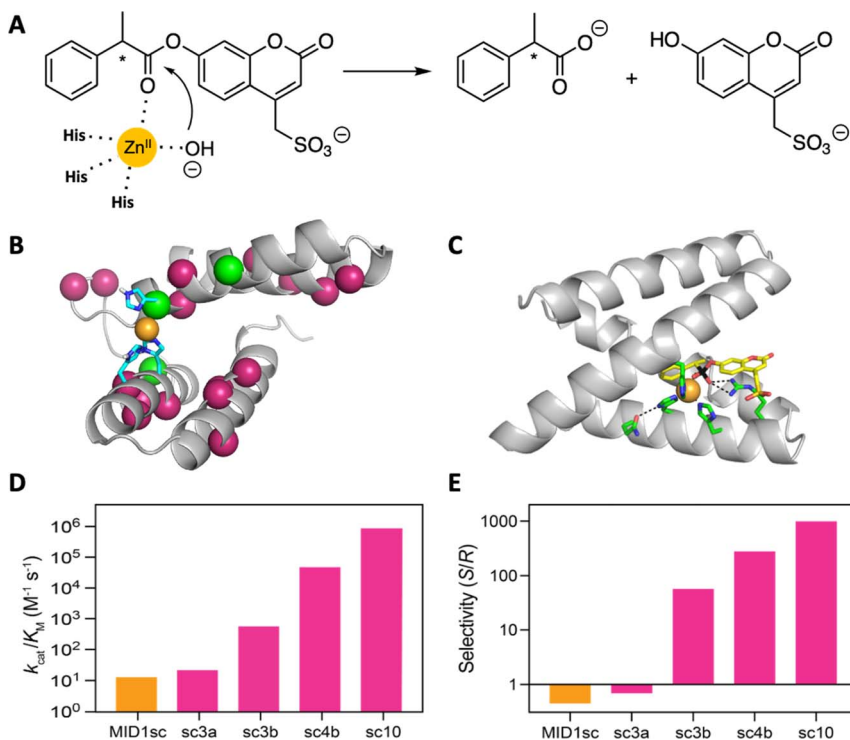
Fig. 7   Design and evolution of a highly active and enantiospecific metalloesterase.[93] (A) The computationally designed scaffold binds zinc and shows weak hydrolytic activity toward a coumarin ester of phenylpropionate. (B) Mutations (colored balls) introduced over 10 rounds of mutagenesis and screening are distributed throughout the protein scaffold. (C) A crystal structure of the optimized variant in complex with a transition state analog. (D) Activity increased by five orders of magnitude over the course of evolution, and (E) the stereospecificity first inverted and then increased in step with activity.

analog revealed the origins of these properties, highlighting an unexpected change in the coordination environment of the metal ion and the acquisition of an oxyanion-stabilizing arginine residue in the active site.[93]

Although conventional plate assays failed to yield further improvements in esterase activity, recent unpublished work has demonstrated that the enzyme's efficiency can be increased by at least another order of magnitude (Oliver Allemann). This advance was achieved by exploiting an ultrahigh-throughput fluorescence-activated droplet sorting assay to screen libraries containing >10^6 members. Because screening could be carried out at very high substrate concentrations ([S] $\gg K_M$), the improvement was largely manifest as a higher turnover number.

Strikingly, this small metalloprotein can also catalyze an abiological hetero-Diels–Alder reaction.[94] In this case, modest starting activity was obtained through computational design and then improved by five orders of magnitude over multiple rounds of mutagenesis and screening. Detailed kinetic analysis showed that the optimized enzyme is highly stereoselective and >100 times more proficient than other known Diels–Alderases, including designed catalysts and natural

enzymes involved in the biosynthesis of polyketide natural products. Its high activity can be ascribed to the use of a zinc cation, absent in the other catalysts, as a Lewis acid to activate the dienophile for reaction.

Simultaneous optimization of structure and function in a naïve scaffold not only illustrates how the first enzymes might have emerged from simpler precursors, but the high activities achieved also attest to the value of harnessing Lewis acids for new catalytic activities. Indeed, this helical bundle scaffold is remarkably malleable. It promiscuously catalyzes the reduction of azachalcones by Hantzsch esters, bimolecular Michael additions, and zinc-dependent retro-aldol reactions (Yusuke Ota, unpublished). Although starting activities are generally low, they can be increased substantially by multiple rounds of mutagenesis and screening. Changing the metal ion could conceivably provide a rich source of additional activities.

Perspectives

Laboratory evolution and computational methods have become invaluable tools for engineering biocatalysts with customized activities and specificities. Not only can they adapt existing enzymes for new tasks, but they have also been key to realizing Emil Fischer's 120-year-old dream[95] of generating *de novo* enzymes that are as efficient and selective as their counterparts in nature. Being able to accelerate abiological reactions in this way has the potential to greatly expand the toolkit of biocatalysts available for chemical synthesis and other applications.

Nonetheless, many challenges persist. Producing new enzymes for energetically demanding transformations is still difficult, reflecting the complexity of the interconnected hydrogen-bonded networks needed to preorganize active site residues in catalytically competent orientations and to facilitate concerted interactions between side chains during catalysis. Even when design works, success rates and initial activities are low, pointing to inadequacies in the design tools that will require rectification.

The advent of deep learning technologies for predicting protein structure from sequence[96] and "hallucinating" atomically accurate protein scaffolds[97–99] will undoubtedly help address these issues. By learning how sequence maps to fitness, such methods will navigate vastly larger regions of sequence space than will ever be possible experimentally. The recent design of a *de novo* luciferase highlights the enormous promise of these approaches.[100] To capture the dynamic properties of proteins, which is currently ignored in most design protocols, classical molecular dynamics (MD) and quantum mechanics/molecular mechanics (QM/MM) simulations may augment these efforts.[101,102]

Although time-consuming and costly, laboratory evolution can be expected to play a key role in enzyme design for the foreseeable future. As computational methods improve, though, its function should shift from fixing flawed designs to fine-tuning their properties. Advances in gene synthesis and next-generation sequencing have the potential to expedite the implementation and analysis of "smart" libraries,[103,104] while (ultra)high-throughput screening methods[105] can facilitate the search of large populations of variants and thus increase the likelihood of discovering synergistic combinations of rare but beneficial mutations. Recent initiatives to develop "self-driving laboratories" that explore protein fitness landscapes by integrating machine learning, decision-making, protein

design, and experimental validation provide additional opportunities to automate and accelerate biocatalyst optimization.[106,107]

Protein design and evolution have made amazing progress in the last decade, signaling an exciting future for enzyme engineering in the years to come. Being able to design enzymes at will for any reaction you fancy will benefit efforts to address countless real-world challenges, paving the way to a more sustainable and prosperous future.

Data availability

No primary research results, software or code were included, and no new data were generated or analyzed for this review.

Author contributions

DH wrote this review article.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 Z. Wang, Q. Xie, H. Zhou, M. Zhang, J. Shen and D. Ju, Amino acid degrading enzymes and autophagy in cancer therapy, *Front. Pharmacol*, 2021, **11**, 582–587.

2 H. M. Tartibi, M. S. Hershfield and S. L. Bahna, A 24-year enzyme replacement therapy in an adenosine-deaminase-deficient patient, *Pediatrics*, 2016, **137**, e20152169.

3 W. Kim, H. Erlandsen, S. Surendran, R. C. Stevens, A. Gamez, K. Michols-Matalon, S. K. Tyring and R. Matalon, Trends in enzyme therapy for phenylketonuria, *Mol. Ther.*, 2004, **10**, 220–224.

4 E. L. Bell, W. Finnigan, S. P. France, A. P. Green, M. A. Hayes, L. J. Hepworth, S. L. Lovelock, H. Niikura, S. Osuna, E. Romero, K. S. Ryan, N. J. Turner and S. L. Flitsch, Biocatalysis, *Nat. Rev. Methods Primers*, 2021, **1**, 46.

5 R. Buller, S. Lutz, R. J. Kazlauskas, R. Snajdrova, J. C. Moore and U. T. Bornscheuer, From nature to industry: Harnessing enzymes for biocatalysis, *Science*, 2023, **382**, eadh8615.

6 M. D. Toscano, K. J. Woycechowsky and D. Hilvert, Minimalist active-site redesign: teaching old enzymes new tricks, *Angew. Chem., Int. Ed.*, 2007, **46**, 3212–3236.

7 F. H. Arnold, Design by directed evolution, *Acc. Chem. Res.*, 1998, **31**, 125–131.

8 M. T. Reetz, Directed evolution of enantioselective enzymes as catalysts for organic synthesis, *Adv. Catal.*, 2006, **49**, 1–69.

9 M. S. Packer and D. R. Liu, Methods for the directed evolution of proteins, *Nat. Rev. Genet.*, 2015, **16**, 379–394.

10 C. Zeymer and D. Hilvert, Directed Evolution of Protein Catalysts, *Annu. Rev. Biochem.*, 2018, **87**, 131–157.

11 Y. Wang, P. Xue, M. Cao, T. Yu, S. T. Lane and H. Zhao, Directed Evolution: Methodologies and Applications, *Chem. Rev.*, 2021, **121**, 12384–12444.

12 R. B. Leveson-Gower, C. Mayer and G. Roelfes, The importance of catalytic promiscuity for enzyme design and evolution, *Nat. Rev. Chem.*, 2019, **3**, 687–705.

13 Y. Yang and F. H. Arnold, Navigating the Unnatural Reaction Space: Directed Evolution of Heme Proteins for Selective Carbene and Nitrene Transfer, *Acc. Chem. Res.*, 2021, **54**, 1209–1225.

14 S. B. Kan, R. D. Lewis, K. Chen and F. H. Arnold, Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life, *Science*, 2016, **354**, 1048–1051.

15 S. B. J. Kan, X. Huang, Y. Gumulya, K. Chen and F. H. Arnold, Genetically programmed chiral organoborane synthesis, *Nature*, 2017, **552**, 132–136.

16 M. A. Emmanuel, N. R. Greenberg, D. G. Oblinsky and T. K. Hyster, Accessing non-natural reactivity by irradiating nicotinamide-dependent enzymes with light, *Nature*, 2016, **540**, 414–417.

17 K. F. Biegasiewicz, S. J. Cooper, X. Gao, D. G. Oblinsky, J. H. Kim, S. E. Garfinkle, L. A. Joyce, B. A. Sandoval, G. D. Scholes and T. K. Hyster, Photoexcitation of flavoenzymes enables a stereoselective radical cyclization, *Science*, 2019, **364**, 1166–1169.

18 J. Rui, Q. Zhao, A. J. Huls, J. Soler, J. C. Paris, Z. Chen, V. Reshetnikov, Y. Yang, Y. Guo, M. Garcia-Borràs and X. Huang, Directed evolution of nonheme iron enzymes to access abiological radical-relay C(sp3)-H azidation, *Science*, 2022, **376**, 869–874.

19 Q. Zhou, M. Chin, Y. Fu, P. Liu and Y. Yang, Stereodivergent atom-transfer radical cyclization by engineered cytochromes P450, *Science*, 2021, **374**, 1612–1616.

20 T. C. Wang, B. K. Mai, Z. Zhang, Z. Bo, J. Li, P. Liu and Y. Yang, Stereoselective amino acid synthesis by photobiocatalytic oxidative coupling, *Nature*, 2024, **629**, 98–104.

21 C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman and G. J. Hughes, Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture, *Science*, 2010, **329**, 305–309.

22 A. P. Green and N. J. Turner, Biocatalytic retrosynthesis: Redesigning synthetic routes to high-value chemicals, *Perspect. Sci.*, 2016, **9**, 42–48.

23 G. Grogan, Synthesis of chiral amines using redox biocatalysis, *Curr. Opin. Chem. Biol.*, 2018, **43**, 15–22.

24 A. Debon, M. Pott, R. Obexer, A. P. Green, L. Friedrich, A. Griffiths and D. Hilvert, Ultrahigh-throughput screening enables efficient single-round oxidase remodelling, *Nat. Catal.*, 2019, **2**, 740–747.

25 P. Schultz, Expanding the genetic code, *Protein Sci.*, 2023, **32**, e4488.

26 J. W. Chin, Expanding and reprogramming the genetic code, *Nature*, 2017, **550**, 53–60.

27 A. V. Anzalone, L. W. Koblan and D. R. Liu, Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors, *Nat. Biotechnol.*, 2020, **38**, 824–844.

28 E. L. Bell, R. Smithson, S. Kilbride, J. Foster, F. J. Hardy, S. Ramachandran, A. A. Tedstone, S. J. Haigh, A. A. Garforth, P. J. R. Day, C. Levy, M. P. Shaver and A. P. Green, Directed evolution of an efficient and thermostable PET depolymerase, *Nat. Catal.*, 2022, **5**, 673–681.

29 B. Guo, X. Lopez-Lorenzo, Y. Fang, E. Bäckström, A. J. Capezza, S. R. Vanga, I. Furó, M. Hakkarainen and P. O. Syrén, Fast depolymerization of PET bottle mediated by microwave pre-treatment and an engineered PETase, *ChemSusChem*, 2023, **16**, e202301237.

30 V. Tournier, C. M. Topham, A. Gilles, B. David, C. Folgoas, E. Moya-Leclair, E. Kamionka, M. L. Desrousseaux, H. Texier, S. Gavalda, M. Cot, E. Guémard, M. Dalibey, J. Nomme, G. Cioci, S. Barbe, M. Chateau, I. André, S. Duquesne and A. Marty, An engineered PET depolymerase to break down and recycle plastic bottles, *Nature*, 2020, **580**, 216–219.

31 S. P. France, L. J. Hepworth, N. J. Turner and S. L. Flitch, Constructing biocatalytic cascades: In vitro and in vivo approaches to de novo multi-enzyme pathways, *ACS Catal.*, 2017, **7**, 710–724.

32 C. J. Paddon and J. D. Keasling, Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development, *Nat. Rev. Microbiol.*, 2014, **12**, 355–367.

33 L. B. B. Martin, S. Kikuchi, M. Rejzek, C. Owen, J. Reed, A. Orme, R. C. Misra, A. El-Demerdash, L. Hill, H. Hodgson, Y. Liu, J. D. Keasling, R. A. Field, A. W. Truman and A. Osbourn, Complete biosynthesis of the potent vaccine adjuvant QS-21, *Nat. Chem. Biol.*, 2024, **20**, 493–502.

34 T. J. Erb, P. R. Jones and A. Bar-Even, Synthetic metabolism: metabolic engineering meets enzyme design, *Curr. Opin. Chem. Biol.*, 2017, **37**, 56–62.

35 H. Schulz-Mirbach, B. Dronsella, H. He and T. J. Erb, Creating new-to-nature carbon fixation: A guide, *Metab. Eng.*, 2024, **82**, 12–28.

36 M. A. Huffman, A. Fryszkowska, O. Alvizo, M. Borra-Garske, K. R. Campos, K. A. Canada, P. N. Devine, D. Duan, J. H. Forstater, S. T. Grosser, H. M. Halsey, G. J. Hughes, J. Jo, L. A. Joyce, J. N. Kolev, J. Liang, K. M. Maloney, B. F. Mann, N. M. Marshall, M. McLaughlin, J. C. Moore, G. S. Murphy, C. C. Nawrat, J. Nazor, S. Novick, N. R. Patel, A. Rodriguez-Granillo, S. A. Robaire, E. C. Sherer, M. D. Truppo, A. M. Whittaker, D. Verma, L. Xiao, Y. Xu and H. Yang, Design of an in vitro biocatalytic cascade for the manufacture of islatravir, *Science*, 2019, **366**, 1255–1259.

37 D. E. Cane, C. T. Walsh and C. Khosla, Harnessing the biosynthetic code: combinations, permutations, and mutations, *Science*, 1998, **282**, 63–68.

38 A. Nivina, K. P. Yuet, J. Hsu and C. Khosla, Evolution and diversity of assembly-line polyketide synthases, *Chem. Rev.*, 2019, **119**, 12524–12547.

39 R. D. Süssmuth and A. Mainz, Nonribosomal peptide synthesis-principles and prospects, *Angew. Chem., Int. Ed.*, 2017, **56**, 3770–3821.

40 M. Winn, J. K. Fyans, Y. Zhuo and J. Micklefield, Recent advances in engineering nonribosomal peptide assembly lines, *Nat. Prod. Rep.*, 2016, **33**, 317–347.

41 D. L. Niquille, D. A. Hansen, T. Mori, D. Fercher, H. Kries and D. Hilvert, Nonribosomal biosynthesis of backbone-modified peptides, *Nat. Chem.*, 2018, **10**, 282–287.

42 A. Camus, G. Truong, P. R. E. Mittl, G. Markert and D. Hilvert, Reprogramming nonribosomal peptide synthetases for site-specific insertion of α-hydroxy acids, *J. Am. Chem. Soc.*, 2022, **144**, 17567–17575.

43 A. Camus, M. Gantz and D. Hilvert, High-throughput engineering of nonribosomal extension modules, *ACS Chem. Biol.*, 2023, **18**, 2516–2523.

44 I. B. Folger, N. F. Frota, A. Pistofidis, D. L. Niquille, D. A. Hansen, T. M. Schmeing and D. Hilvert, High-throughput reprogramming of an NRPS condensation domain, *Nat. Chem. Biol.*, 2024, **20**, 761–769.

45 E. T. Kaiser and D. S. Lawrence, Chemical mutation of enzyme active sites, *Science*, 1984, **226**, 505–511.

46 Z. P. Wu and D. Hilvert, Selenosubtilisin as a glutathione peroxidase mimic, *J. Am. Chem. Soc.*, 1990, **112**, 5647–5648.

47 M. Tinzl and D. Hilvert, Trapping transient protein species by genetic code expansion, *ChemBioChem*, 2021, **22**, 92–99.

48 J. S. Trimble, R. Crawshaw, F. J. Hardy, C. W. Levy, M. J. B. Brown, D. E. Fuerst, D. J. Heyes, R. Obexer and A. P. Green, A designed photoenzyme for enantioselective [2+2] cycloadditions, *Nature*, 2022, **611**, 709–714.

49 C. Mayer, C. Dulson, E. Reddem, A. W. H. Thunnissen and G. Roelfes, Directed evolution of a designer enzyme featuring an unnatural catalytic amino acid, *Angew. Chem., Int. Ed.*, 2019, **58**, 2083–2087.

50 L. Longwitz, R. B. Leveson-Gower, H. J. Rozeboom, A. W. H. Thunnissen and G. Roelfes, Boron catalysis in a designer enzyme, *Nature*, 2024, **629**, 824–829.

51 H. M. Key, P. Dydio, D. S. Clark and J. F. Hartwig, Abiological catalysis by artificial haem proteins containing noble metals in place of iron, *Nature*, 2016, **534**, 534–537.

52 P. Dydio, H. M. Key, H. Hayashi, D. S. Clark and J. F. Hartwig, Chemoselective, enzymatic C-H bond amination catalyzed by a cytochrome P450 containing an Ir(Me)-PIX cofactor, *J. Am. Chem. Soc.*, 2017, **139**, 1750–1753.

53 H. J. Davis and T. R. Ward, Artificial metalloenzymes: Challenges and opportunities, *ACS Cent. Sci.*, 2019, **5**, 1120–1136.

54 G. Roelfes, LmrR: A privileged scaffold for artificial metalloenzymes, *Acc. Chem. Res.*, 2019, **52**, 545–556.

55 S. L. Lovelock, R. Crawshaw, S. Basler, C. Levy, D. Baker, D. Hilvert and A. P. Green, The road to fully programmable protein catalysis, *Nature*, 2022, **606**, 49–58.

56 R. A. Lerner, S. J. Benkovic and P. G. Schultz, At the crossroads of chemistry and immunology: catalytic antibodies, *Science*, 1991, **252**, 659–667.

57 D. Hilvert, Critical analysis of antibody catalysis, *Annu. Rev. Biochem.*, 2000, **69**, 751–793.

58 F. Richter, A. Leaver-Fay, S. D. Khare, S. Bjelic and D. Baker, De novo enzyme design using Rosetta3, *PLoS One*, 2011, **6**, e19230.

59 F. Richter, R. Blomberg, S. D. Khare, G. Kiss, A. P. Kuzin, A. J. Smith, J. Gallaher, Z. Pianowski, R. C. Helgeson, A. Grjasnow, R. Xiao, J. Seetharaman, M. Su, S. Vorobiev, S. Lew, F. Forouhar, G. J. Kornhaber, J. F. Hunt, G. T. Montelione, L. Tong, K. N. Houk, D. Hilvert and D. Baker,

Computational design of catalytic dyads and oxyanion holes for ester hydrolysis, *J. Am. Chem. Soc.*, 2012, **134**, 16197–16206.

60 D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik and D. Baker, Kemp elimination catalysts by computational enzyme design, *Nature*, 2008, **453**, 190–195.

61 H. K. Privett, G. Kiss, T. M. Lee, R. Blomberg, R. A. Chica, L. M. Thomas, D. Hilvert, K. N. Houk and S. L. Mayo, Iterative approach to computational enzyme design, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 3790–3795.

62 R. Blomberg, H. Kries, D. M. Pinkas, P. R. Mittl, M. G. Grütter, H. K. Privett, S. L. Mayo and D. Hilvert, Precision is essential for efficient catalysis in an evolved Kemp eliminase, *Nature*, 2013, **503**, 418–421.

63 R. Otten, R. A. P. Pádua, H. A. Bunzel, V. Nguyen, W. Pitsawong, M. Patterson, S. Sui, S. L. Perry, A. E. Cohen, D. Hilvert and D. Kern, How directed evolution reshapes the energy landscape in an enzyme to boost catalysis, *Science*, 2020, **370**, 1442–1446.

64 J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael and D. Baker, Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction, *Science*, 2010, **329**, 309–313.

65 N. Preiswerk, T. Beck, J. D. Schulz, P. Milovník, C. Mayer, J. B. Siegel, D. Baker and D. Hilvert, Impact of scaffold rigidity on the design and evolution of an artificial Diels-Alderase, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 8013–8018.

66 L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Röthlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas 3rd, D. Hilvert, K. N. Houk, B. L. Stoddard and D. Baker, De novo computational design of retro-aldol enzymes, *Science*, 2008, **319**, 1387–1391.

67 L. Giger, S. Caner, R. Obexer, P. Kast, D. Baker, N. Ban and D. Hilvert, Evolution of a designed retro-aldolase leads to complete active site remodeling, *Nat. Chem. Biol.*, 2013, **9**, 494–498.

68 R. Obexer, A. Godina, X. Garrabou, P. R. Mittl, D. Baker, A. D. Griffiths and D. Hilvert, Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase, *Nat. Chem.*, 2017, **9**, 50–56.

69 S. Bjelic, L. G. Nivón, N. Çelebi-Ölçüm, G. Kiss, C. F. Rosewall, H. M. Lovick, E. L. Ingalls, J. L. Gallaher, J. Seetharaman, S. Lew, G. T. Montelione, J. F. Hunt, F. E. Michael, K. N. Houk and D. Baker, Computational design of enone-binding proteins with catalytic activity for the Morita-Baylis-Hillman reaction, *ACS Chem. Biol.*, 2013, **8**, 749–757.

70 R. Crawshaw, A. E. Crossley, L. Johannissen, A. J. Burke, S. Hay, C. Levy, D. Baker, S. L. Lovelock and A. P. Green, Engineering an efficient and enantioselective enzyme for the Morita-Baylis-Hillman reaction, *Nat. Chem.*, 2022, **14**, 313–320.

71 M. Gantz, S. V. Mathis, F. E. H. Nintzel, P. Lio and F. Hollfelder, On synergy between ultrahigh throughput screening and machine learning in biocatalyst engineering, *Faraday Discuss.*, 2024, DOI: **10.1039/D4FD00065J**.

72 J. Yang, F. Z. Li and F. H. Arnold, Opportunities and Challenges for Machine Learning-Assisted Enzyme Engineering, *ACS Cent. Sci.*, 2024, **10**, 226–241.

73 C. R. Freschlin, S. A. Fahlberg and P. A. Romero, Machine learning to navigate fitness landscapes for protein engineering, *Curr. Opin. Biotechnol.*, 2022, **75**, 102713.

74 R. Feehan, D. Montezano and J. S. G. Slusky, Machine learning for enzyme engineering, selection and design, *Protein Eng., Des. Sel.*, 2021, **34**, gzab019.

75 S. Gelman, S. A. Fahlberg, P. Heinzelman, P. A. Romero and A. Gitter, Neural networks to learn protein sequence-function relationships from deep mutational scanning data, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2104878118.

76 R. B. Hill, D. P. Raleigh, A. Lombardi and W. F. DeGrado, De novo design of helical bundles as models for understanding protein folding and function, *Acc. Chem. Res.*, 2000, **33**, 745–754.

77 T. A. Farid, G. Kodali, L. A. Solomon, B. R. Lichtenstein, M. M. Sheehan, B. A. Fry, C. Bialas, N. M. Ennist, J. A. Siedlecki, Z. Zhao, M. A. Stetz, K. G. Valentine, J. L. R. Anderson, A. J. Wand, B. M. Discher, C. C. Moser and P. L. Dutton, Elementary tetrahelical protein design for diverse oxidoreductase functions, *Nat. Chem. Biol.*, 2013, **9**, 826–833.

78 C. T. Armstrong, D. W. Watkins and J. L. Anderson, Constructing manmade enzymes for oxygen activation, *Dalton Trans.*, 2013, **42**, 3136–3150.

79 M. Faiella, C. Andreozzi, R. T. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado and A. Lombardi, An artificial di-iron oxo-protein with phenol oxidase activity, *Nat. Chem. Biol.*, 2009, **5**, 882–884.

80 M. L. Zastrow, A. F. Peacock, J. A. Stuckey and V. L. Pecoraro, Hydrolytic catalysis and structural stabilization in a designed metalloprotein, *Nat. Chem.*, 2012, **4**, 118–123.

81 W. J. Song and F. A. Tezcan, A designed supramolecular protein assembly with in vivo enzymatic activity, *Science*, 2014, **346**, 1525–1528.

82 A. J. Burton, A. R. Thomson, W. M. Dawson, R. L. Brady and D. N. Woolfson, Installing hydrolytic activity into a completely de novo protein framework, *Nat. Chem.*, 2016, **8**, 837–844.

83 M. H. Hecht, A. Das, A. Go, L. H. Bradley and Y. Wei, De novo proteins from designed combinatorial libraries, *Protein Sci.*, 2004, **13**, 1711–1723.

84 M. A. Fisher, K. L. McKinley, L. H. Bradley, S. R. Viola and M. H. Hecht, De novo designed proteins from a library of artificial sequences function in *Escherichia coli* and enable cell growth, *PLoS One*, 2011, **6**, e15364.

85 A. E. Donnelly, G. S. Murphy, K. M. Digianantonio and M. H. Hecht, A de novo enzyme catalyzes a life-sustaining reaction in *Escherichia coli*, *Nat. Chem. Biol.*, 2018, **14**, 253–255.

86 P. S. Huang, S. E. Boyken and D. Baker, The coming of age of de novo protein design, *Nature*, 2016, **537**, 320–327.

87 Y. Kipnis, A. O. Chaib, A. A. Vorobieva, G. Cai, G. Reggiano, B. Basanta, E. Kumar, P. R. E. Mittl, D. Hilvert and D. Baker, Design and optimization of enzymatic activity in a de novo β-barrel scaffold, *Protein Sci.*, 2022, **31**, e4405.

88 I. Kalvet, M. Ortmayer, J. Zhao, R. Crawshaw, N. M. Ennist, C. Levy, A. Roy, A. P. Green and D. Baker, Design of heme enzymes with a tunable substrate binding pocket adjacent to an open metal coordination site, *J. Am. Chem. Soc.*, 2023, **145**, 14307–14315.

89 S. J. Caldwell, I. C. Haydon, N. Piperidou, P. S. Huang, M. J. Bick, H. S. Sjöström, D. Hilvert, D. Baker and C. Zeymer, Tight and specific

lanthanide binding in a de novo TIM barrel with a large internal cavity designed by symmetric domain fusion, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 30362–30369.

90 A. S. Klein, F. Leiss-Maier, R. Mühlhofer, B. Boesen, G. Mustafa, H. Kugler, *et al.*, A de novo metalloenzyme for cerium photoredox catalysis, *ChemRxiv*, 2024, preprint, DOI: **10.26434/chemrxiv-2024-6g4px**.

91 B. S. Der, M. Machius, M. J. Miley, J. L. Mills, T. Szyperski and B. Kuhlman, Metal-mediated affinity and orientation specificity in a computationally designed protein homodimer, *J. Am. Chem. Soc.*, 2012, **134**, 375–385.

92 B. S. Der, D. R. Edwards and B. Kuhlman, Catalysis by a de novo zinc-mediated protein interface: implications for natural enzyme evolution and rational enzyme engineering, *Biochemistry*, 2012, **51**, 3933–3940.

93 S. Studer, D. A. Hansen, Z. L. Pianowski, P. R. E. Mittl, A. Debon, S. L. Guffy, B. S. Der, B. Kuhlman and D. Hilvert, Evolution of a highly active and enantiospecific metalloenzyme from short peptides, *Science*, 2018, **362**, 1285–1288.

94 S. Basler, S. Studer, Y. Zou, T. Mori, Y. Ota, A. Camus, H. A. Bunzel, R. C. Helgeson, K. N. Houk, G. Jiménez-Osés and D. Hilvert, Efficient Lewis acid catalysis of an abiological reaction in a de novo protein scaffold, *Nat. Chem.*, 2021, **13**, 231–235.

95 E. Fischer, Nobel Prize Outreach AB, 2022, **https://www.nobelprize.org/prizes/chemistry/1902/fischer/lecture/**.

96 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C. C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, O. Bertolli, A. Bridgland, A. Cherepanov, M. Congreve, A. I. Cowen-Rivers, A. Cowie, M. Figurnov, F. B. Fuchs, H. Gladman, R. Jain, Y. A. Khan, C. M. R. Low, K. Perlin, A. Potapenko, P. Savy, S. Singh, A. Stecula, A. Thillaisundaram, C. Tong, S. Yakneen, E. D. Zhong, M. Zielinski, A. Žídek, V. Bapst, P. Kohli, M. Jaderberg, D. Hassabis and J. M. Jumper, Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, 2024, **630**, 493–500.

97 I. Anishchenko, S. J. Pellock, T. M. Chidyausiku, T. A. Ramelot, S. Ovchinnikov, J. Hao, K. Bafna, C. Norn, A. Kang, A. K. Bera, F. DiMaio, L. Carter, C. M. Chow, G. T. Montelione and D. Baker, De novo protein design by deep network hallucination, *Nature*, 2021, **600**, 547–552.

98 J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N. P. King and D. Baker, Robust deep learning-based protein sequence design using ProteinMPNN, *Science*, 2022, **378**, 49–56.

99 J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, N. Hanikel, S. J. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. V. Torres, A. Lauko, V. De Bortoli, E. Mathieu, S. Ovchinnikov, R. Barzilay, T. S. Jaakkola, F. DiMaio, M. Baek and D. Baker, De novo design of protein structure and function with RFdiffusion, *Nature*, 2023, **620**, 1089–1100.

100 A. H. Yeh, C. Norn, Y. Kipnis, D. Tischer, S. J. Pellock, D. Evans, P. Ma, G. R. Lee, J. Z. Zhang, I. Anishchenko, B. Coventry, L. Cao, J. Dauparas, S. Halabiya, M. DeWitt, L. Carter, K. N. Houk and D. Baker, De novo design of luciferases using deep learning, *Nature*, 2023, **614**, 774–780.

101 S. Osuna, G. Jimenez-Oses, E. L. Noey and K. N. Houk, Molecular dynamics explorations of active site structure in designed and evolved enzymes, *Acc. Chem. Res.*, 2015, **48**, 1080–1089.

102 M. C. Childers and V. Daggett, Insights from molecular dynamics simulations for computational protein design, *Mol. Syst. Des. Eng.*, 2017, **2**, 9–33.

103 S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz and J. P. Acevedo, Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis, *ACS Synth. Biol.*, 2013, **2**, 83–92.

104 D. Patsch, T. Schwander, M. Voss, D. Schaub, S. Hüppi, M. Eichenberger, D. Stockinger, L. Schelbert, D. Giger, F. Peccati, G. Jiménez-Osés, M. Mutný, A. Krause, U. T. Bornscheuer, D. Hilvert and R. M. Buller, Enriching productive mutational paths accelerates enzyme evolution, *Nat. Chem. Biol.*, 2024, in press.

105 H. A. Bunzel, X. Garrabou, M. Pott and D. Hilvert, Speeding up enzyme discovery and engineering with ultrahigh-throughput methods, *Curr. Opin. Struct. Biol.*, 2018, **48**, 149–156.

106 J. T. Rapp, B. J. Bremer and P. A. Romero, Self-driving laboratories to autonomously navigate the protein fitness landscape, *Nat. Chem. Eng.*, 2024, **1**, 97–107.

107 E. Orsi, L. Schada von Borzyskowski, S. Noack, P. I. Nikel and S. N. Lindner, Automated in vivo enzyme engineering accelerates biocatalyst optimization, *Nat. Commun.*, 2024, **15**, 3447.