nature catalysis

REVIEW ARTICLE

https://doi.org/10.1038/s41929-019-0385-5

Engineering new catalytic activities in enzymes

Kai Chen and Frances H. Arnold*

The efficiency, selectivity and sustainability benefits offered by enzymes are enticing chemists to consider biocatalytic transformations to complement or even supplant more traditional synthetic routes. Increasing demands for efficient and versatile synthetic methods, combined with powerful new discovery and engineering tools, has prompted innovations in biocatalysis, especially the development of new enzymes for precise transformations or 'molecular editing'. As a result, the past decade has witnessed an impressive expansion of the catalytic repertoire of enzymes to include new and useful transformations not known (or relevant) in the biological world. In this Review we illustrate various ways in which researchers have approached using the catalytic machineries of enzymes for new-to-nature transformations. These efforts have identified genetically encoded catalysts that can be tuned and diversified by engineering the protein sequence, particularly by directed evolution. Discovery and improvement of these new enzyme activities is opening a floodgate that connects the chemistry of the biological world to that invented by humans over the past 100 years.

Nature has evolved an astonishing array of enzymes to catalyse the chemical transformations that enable biological systems to eke out a living in diverse environments. Enzymes synthesize biological building blocks from available elemental resources, with which more enzymes go on to assemble new life, including essential biomolecules, complex natural products and macromolecular materials. Enzymes also break down these compounds into reusable fragments. While executing their biologically relevant functions, enzymes can, when needed, exert precise control over reaction outcomes. The unique ability of enzymes to do such 'molecular editing' has prompted addition of some of nature's catalysts to the organic synthesis toolbox[1–3].

Nature's repertoire of enzyme functions is striking — from photosynthesis to nitrogen fixation, water splitting to aliphatic carbon assembly, there are still no human-made catalysts that can compete with these fundamental processes of life. On the other hand, the biological world has not taken the same path as human-invented chemistry, and many valuable transformations invented by synthetic chemists have no known enzyme-catalysed counterparts. Among other reasons, nature does not use many of our favourite transformations because (1) the products are not useful to living systems, (2) the required reagents do not (stably) exist in nature or (3) the conditions that are needed to effect the reactions are not available.

To bridge nature's catalytic repertoire and the demands of synthetic chemistry, chemists and biologists have started to import human-invented chemistry into enzymes. One approach that researchers have tried is computational de novo enzyme design based on knowledge of the reaction transition-state structure[4,5]. However, given our limited understanding of how enzymes function at an atomic level and how a sequence encodes catalytic function in macromolecular design, an alternative avenue of engineering existing proteins has proven more successful, at least for now. Researchers are quickly unlocking new catalytic activities of existing enzymes simply by challenging and/or engineering them to work with non-natural reagents and in new environments. New activities can be released with relatively small modifications; for example, by introducing a different metal centre or changing a few amino acids in an active site (perhaps this is not surprising, as it is also nature's innovation strategy). Powerful molecular biology tools such as directed evolution can then tune and diversify these new functions to provide catalysts that bring the benefits of nature's biosynthetic machinery to chemical synthesis. This Review will cover these latter efforts to engineer new enzymes by starting from nature's designs (Fig. 1).

A pivotal feature of enzymes, their promiscuity with respect to the substrates they accept and even the reactions they catalyse, has played a central role in the discovery and development of new biocatalytic functions. Enzymes can often accept various substrates in addition to their native one(s); they can even catalyse different transformations when offered the right reagents and environments[6]. Such promiscuous activities may be left over from ancestral enzyme functions, or were never explored in the natural world and come simply as a result of having catalytic machinery that exhibits its hidden capabilities when the environment changes[7]. There is immense potential in nature's vast repertoire of contemporary enzymes for us to discover and make use of, just as nature has done for more than three billion years. Early examples with hydrolytic enzymes showed, for instance, that enzymes whose native function is amide or ester hydrolysis can also utilize their finely tuned networks of active-site residues for the hydrolysis of other bonds or even the formation of new bonds[8]. More recently, chemomimetic approaches developed by transferring human-invented chemistries to cofactor-dependent enzyme have significantly expanded the chemical space accessible to biocatalysis. Protein engineering that uses non-physiological reaction conditions and a combination of chemo- and biocatalysis has further unveiled the potential for chemical innovation in existing enzymes.

Rising demands for efficient, selective and versatile synthetic methods call for new enzymatic functions that may not be relevant in the biological world[9]. The challenges are daunting, requiring not only that enzymes take on new functions, but also that the newly developed biocatalysts exhibit activity and selectivity comparable to or better than current chemocatalytic methods or that they fill gaps in synthetic chemistry. In this Review we summarize the current status of non-natural biocatalysis and describe how protein engineering integrated with chemical rationalization enables innovations that expand the chemical space accessible to enzymes. The creation of abiological enzymatic functions represents a growing area of research that requires knowledge from different fields, including protein engineering, enzymology and synthetic chemistry. We hope this Review will help chemists and biologists recognize, explore and use enzymes for new chemistry.

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA. *e-mail: frances@cheme.caltech.edu

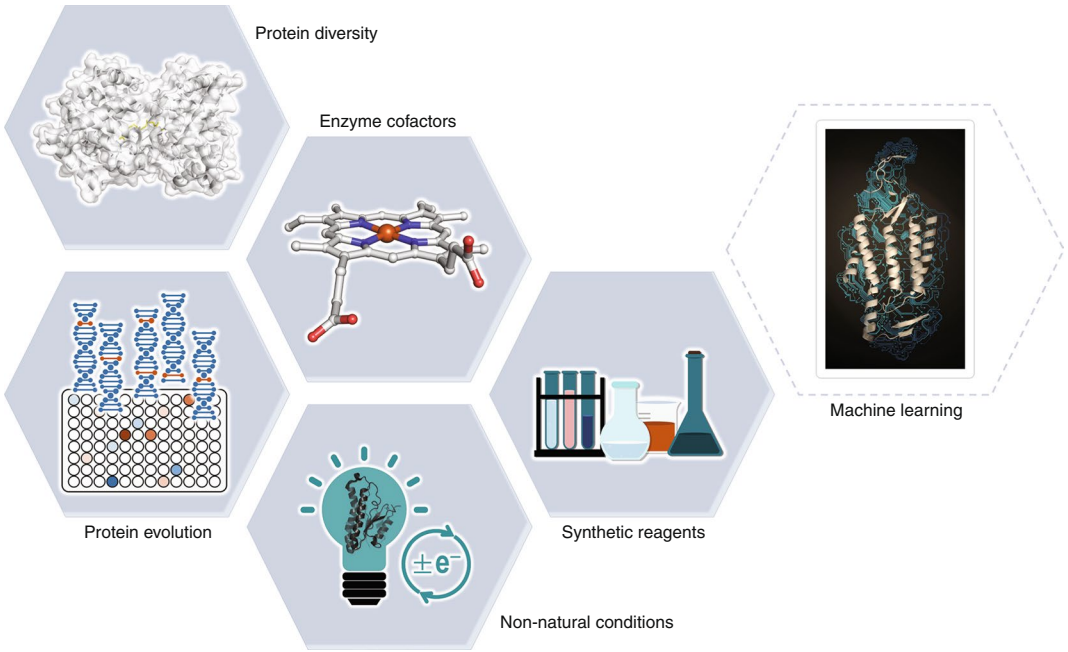

Fig. 1 | Strategies for the discovery of new enzyme functions. Discovery of new enzymes, engineering and diversification of proteins by directed evolution, exploitation of cofactors for new reactivities, and use of synthetic reagents and non-natural conditions have accelerated new enzyme activity development; other advances such as machine learning will expand the capabilities of biocatalysis by learning from growing gene/function databases.

Using established active sites for new enzymatic functions

Early studies of enzyme promiscuity illustrated how enzyme active sites can catalyse physiologically irrelevant chemical transformations[10]. Although different in the mechanism or path of bond formation and breakage, these non-native enzyme functions are typically enabled by the superior capacity of enzyme active sites to stabilize similar transition states and precisely control key intermediates, as exemplified by a large set of hydrolytic enzymes. Many hydrolases feature an oxyanion hole consisting of backbone amides or positively charged residues that stabilize the negative charge on a deprotonated oxygen or alkoxide in the transition state. Such structural properties allow the same hydrolases to catalyse diverse chemical reactions proceeding through oxyanionic intermediates, including aldol reactions, Michael additions, Mannich reactions and even peroxide-involved oxidative reactions[8]. For example, Berglund and co-workers reported in 2003 that *Candida antarctica* lipase B (CAL-B) catalyses aldol reactions between aliphatic aldehydes; they used quantum molecular modelling to illustrate the importance of the oxyanion hole in stabilizing the enolate intermediate (Fig. 2a)[11]. Another class of hydrolase, glycosidase, was intensively investigated and engineered for promiscuous activities of glycoside synthesis: by mutating key residues such as catalytic acid–base pairs, Withers and co-workers converted a glycosidase into a glycosynthase[12] or thioglycoligase[13] employing different α-glycosyl substrates and acceptor sugars.

The promiscuous functions of enzymes have been used for industrial production of valuable compounds. One representative example is halohydrin dehalogenase (HHDH), which in nature catalyses epoxide formation from the corresponding substituted chloro- or bromohydrins[14]. Structural studies on various HHDHs revealed several highly conserved catalytic residues in the active sites that specifically bind epoxides and halide anions. A key tyrosine residue can act as a catalytic base for hydroxyl group deprotonation or a catalytic acid for epoxide protonation, raising the possibility that it could catalyse the reverse reaction, epoxide ring opening. Acceptance of various nucleophiles, including azide, cyanide, nitrite, cyanate, thiocyanate and formate, and high enantioselectivity in the epoxide opening process render HHDHs desirable catalysts for synthetic purposes, especially for the preparation of enantioenriched β-substituted alcohol and epoxide products. Among the biocatalytic applications of HHDH is the asymmetric synthesis of ethyl (*R*)-4-cyano-3-hydroxybutyrate, a precursor of atorvastatin, as reported by scientists at Codexis (Fig. 2b)[15]. Directed evolution was used to enhance the activity of *Agrobacterium radiobacter* HHDH, enabling production of the precursor with >99.9% e.e. based on a substrate loading of 130 g l^{-1}.

The promiscuity of terpene cyclases has also attracted attention for abiological chemistry[16]. Terpene cyclases typically use acid–base catalytic residues for cationic cyclization of polyenes. Squalene-hopene cyclase (SHC), natively responsible for polycyclization of squalene to pentacyclic hopene and hopanol, has been explored as a promiscuous Brønsted-acid biocatalyst to harness a plethora of non-natural reactions driven by protonation processes. Hauer and co-workers employed SHC to construct abiological carbocyclic skeletons by using different internal nucleophilic terminators in the cyclization process[17]. Reshaping the active-site structure of SHC also allowed non-native acidic isomerizations of β-pinene, monocyclization of geraniol and Prins reaction of citronellal (Fig. 2c)[18]. For instance, a single amino acid mutation I261A in SHC from *A. acidocaldarius* improved the activity by eleven folds for a Prins reaction of (*S*)-citronellal to an (−)-iso-isopulegol isomer product (>99% e.e. and >99% d.e.). The literature offers more examples of native enzymes displaying promiscuous activities. But this scenario for searching new enzymatic activities has the obvious limitation that the new reactivity must be quite closely related to the naturally established function.

Exploiting the catalytic potentials of organocofactors

Cofactor-dependent enzymes are of particular interest as many cofactors possess expansive catalytic potential for chemical transformations. This potential has been exploited in natural evolution to create families of enzymes whose functions cover two-electron and single-electron redox/non-redox chemistries. The protein scaffold can distinguish and direct the pathway through which a given reaction will proceed from among two or more possibilities. In this

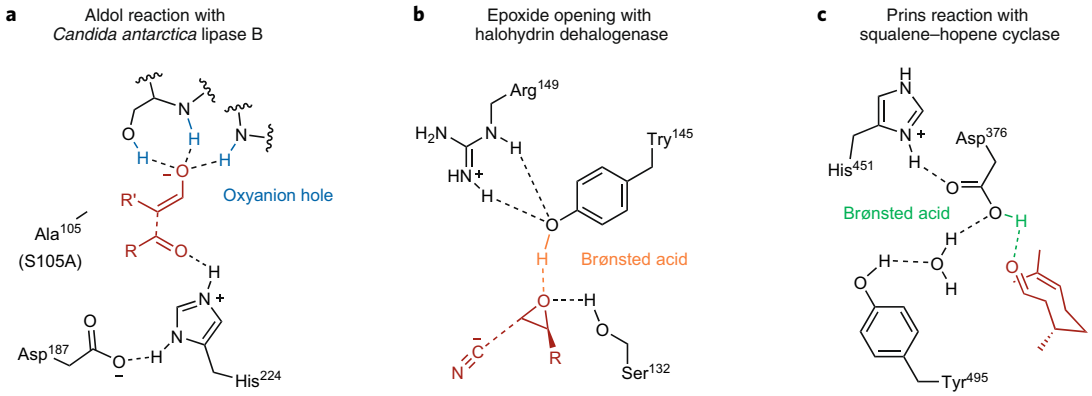

Fig. 2 | Promiscuous functions enabled by versatile active sites. a, Aldol reaction with *Candida antarctica* lipase B[11]. **b**, Epoxide opening with halohydrin dehalogenase[15]. **c**, Prins reaction with squalene–hopene cyclase[18].

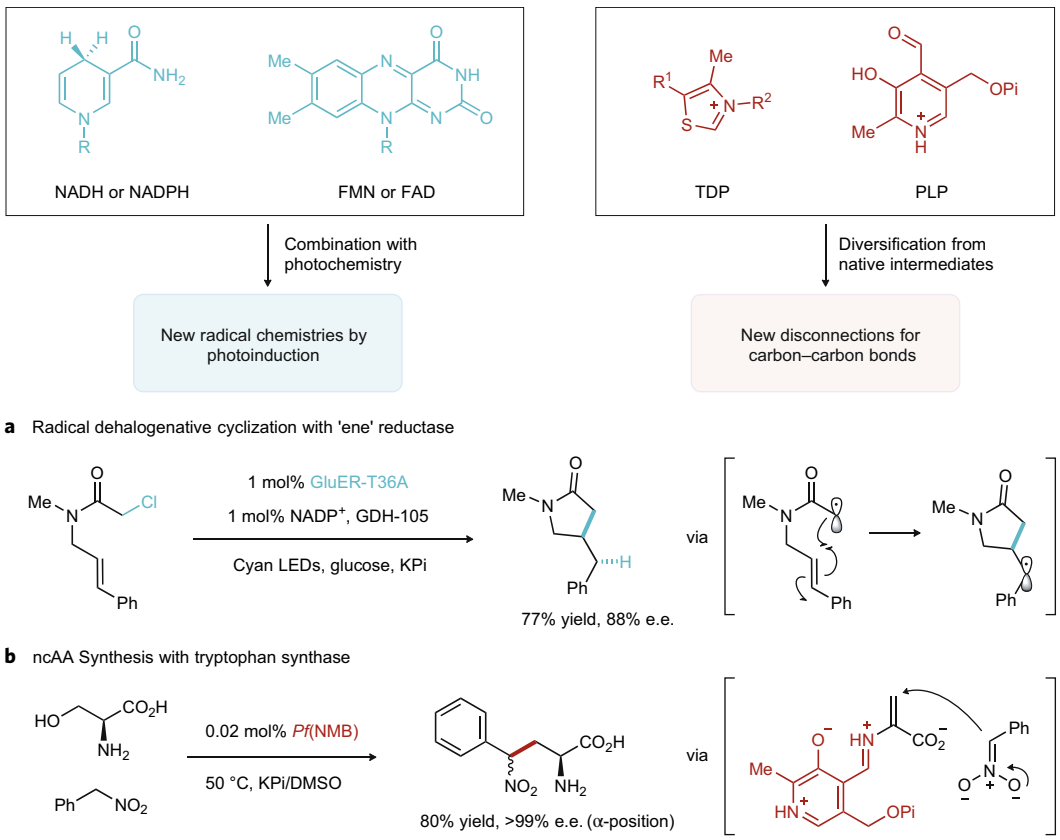

Fig. 3 | New chemistries with cofactor-dependent enzymes. a, Radical dehalogenative cyclization with ERED[27]. **b**, Non-canonical amino acid synthesis with engineered tryptophan synthase[45]. LEDs, light-emitting diodes.

section we will outline recent advances in realizing new chemistries with enzymes having organocofactors, including nicotinamide adenine dinucleotide and its phosphate form (NADH or NADPH), flavin nucleotides (FMN or FAD), thiamine diphosphate (TDP) and pyridoxal phosphate (PLP) (Fig. 3). Reaction design and protein engineering have both promoted the discovery of new functions for these enzymes.

NADH or NADPH act as biological reducing cofactors/co-substrates of numerous oxidoreductases. For example, keto-reductases (KREDs) reduce endogenous carbonyl compounds to alcohols. NADPH-dependent imine reductases (IREDs)[19] are also physiologically capable of reducing carbon–nitrogen double bonds. Turner and co-workers recently identified a subclass of IREDs that catalyse imine formation from ketones and amines prior to imine reduction and named them reductive aminases (RedAms)[20]. A RedAm from *Aspergillus oryzae*, an IRED homologue from a eukaryotic source, was discovered to accept a broad range of structurally diverse amines and ketones, which provides an attractive biocatalytic route to secondary and tertiary amines[21].

Although NAD(P)H is generally recognized as a hydride donor in catalysis, it was also found to be able to implement single-electron-transfer chemistry recently. Hyster and co-workers reported an abiological asymmetric radical dehalogenation of α-bromo-α-aryl/alkyl lactones with KREDs enabled by a photo-induced

electron-transfer strategy[22]. The NAD(P)H cofactor can be excited by blue light to a triplet state, which serves as a potent single-electron reductant. A single electron transfer forms a substrate radical anion that undergoes heterolytic cleavage of a C–Br bond to generate an alkyl radical. NAD(P)H$^{\bullet+}$ subsequently serves as a hydrogen atom donor to afford the dehalogenated lactone product. Several KREDs were found to catalyse dehalogenation of α-bromo γ/δ-lactones in good yields and enantioselectivities. Another system based on the combination of a NADPH-dependent double bond reductase (DBR) and the photocatalyst Rose bengal enables enantioselective radical deacetoxylation of α-acetoxyl ketones, where the enzyme plays a role in activating the substrate for electron transfer followed by deacetoxylation and hydrogen atom transfer from NADPH[23].

Flavin cofactor natively appears in the form of FMN or FAD and is exceptionally versatile in enzymes, mediating a plethora of oxidative and reductive activities[24]. Flavin can exist in multiple redox states, including flavin-N_5-oxide (FMN or FAD oxide), oxidized flavin (quinone, FMN or FAD), flavin semiquinone (FMNH$^{\bullet}$ or FADH$^{\bullet}$), reduced flavin (hydroquinone, $FMNH_2$ or $FADH_2$), as well as flavin-C_4-peroxide (FMNOOH or FADOOH) when engaging with molecular oxygen. The protein scaffold and reaction conditions determine which states of the cofactor are accessible, leading to the diverse redox chemistry of natural flavoenzymes.

Compared with two-electron processes catalysed by flavoenzymes, biological reactions with a single-electron mechanism involving the semiquinone state are rare. The Hyster laboratory discovered a promiscuous radical dehalogenation of α-bromo esters with flavin-dependent ene reductases (EREDs) by making use of the semiquinone state[25]. Mutation Y177F in ERED from *G. oxydans* abolishes the native ene-reduction function in the absence of a proton donor, but considerably improves this non-native function, which further supports the radical mechanism for dehalogenation. Coupled with photoredox catalysis, EREDs could also perform ketone reduction through a radical-mediated pathway[26]. Hyster and his team recently further expanded the capacity of EREDs to perform photo-induced radical cyclization to make various lactams (Fig. 3a)[27]. They reasoned that hydroquinone in the excited state can act as a single-electron reductant that is strong enough to activate α-chloro acetamide and generate the α-acetamide radical, which would cyclize to an *endo*-double bond and then abstract a hydrogen atom from semiquinone to form the desired lactam product. Lactam products with ring sizes that range from five to eight members were accessible via *endo*- or *exo*-cyclization processes in this system. A naturally occurring flavin-dependent photodecarboxylase from *Chlorella variabilis* NC64A was discovered recently by Beisson and his team to employ a semiquinone state for a light-induced radical decarboxylation of fatty acids to alkanes or alkenes[28,29]. These newly demonstrated photoenzymatic platforms have revealed previously unknown catalytic potentials of the flavin cofactor.

The thiamine-dependent enzymes offer another good example of how promiscuous catalytic functions can be exploited in nature and by chemists. The thiamine diphosphate cofactor comprises a thiazolium core, an aminopyrimidine group and a diphosphate moiety. The aminopyrimidine group acts as a key base for deprotonation of the C_2 position in the thiazolium ring, which leads to the formation of a nucleophilic thiazolium carbene and initiates all types of thiamine catalysis in nature[30]. The thiazolium carbene is a superior nucleophile for addition to carbonyl groups, resulting in an enaminol species, known as the Breslow intermediate, for a variety of nucleophilic reactions[31]. In such a way, thiamine enzymes can take electrophilic aldehydes or other carbonyl substrates, turn them into a nucleophilic form and further enable desired bond constructions.

Thiamine-dependent enzymes have been explored for abiological asymmetric C–C bond-forming reactions that take advantage of the nucleophilic feature of the Breslow intermediate. A cross-benzoin condensation between acetaldehyde (after decarboxylation of pyruvate) and benzaldehyde was achieved by Müller and co-workers using the cyclohexane-1,2-dione hydrolase (CDH) from *Azoarcus* sp. to produce chiral α-hydroxy ketone products in high enantioselectivity[32]. CDH was also engineered to accept ketones as electrophiles[33,34]. The cross-benzoin reaction between two aromatic aldehydes is particularly challenging due to the chemoselectivity problem that arises from homo-coupling and mixed cross-couplings[35]. However, benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* and a variant of benzoylformate decarboxylase (BFD) from *Pseudomonas putida* were found to catalyse this reaction, where the steric control from the *ortho*-substituted aldehyde as the electrophile substrate is key to high chemoselectivity[36]. Instead of aldehydes or ketones, α,β-unsaturated carbonyl substrates have also been investigated for this nucleophilic addition of the Breslow intermediate in a 1,4-conjugate manner (the Stetter reaction). A thiamine enzyme PigD from *Serratia marcescens,* catalyses acetaldehyde addition to α,β-unsaturated ketones with high enantioselectivity[37].

Pyridoxal phosphate is another highly versatile enzyme cofactor. Taking advantage of the aldimine intermediate formed through the condensation between the aldehyde group of PLP and the amino group from a substrate, PLP-dependent enzymes catalyse transamination, amino acid decarboxylation, deamination, racemization and more. *O*-Acetylserine sulfhydrylase (OASS) is a PLP-dependent enzyme that is used for cysteine biosynthesis. It forms a key aminoacrylate intermediate through the loss of an acetate from the aldimine between *O*-acetylserine and PLP, and nucleophilic addition of H_2S to the aminoacrylate gives L-cysteine[38]. Early work reported that a variety of heteroatom-based nucleophiles could also be used by OASS to synthesize non-canonical β-substituted alanine derivatives[39].

Similar to OASS in mechanism, tryptophan synthase catalyses the formation of tryptophan through addition of indole to the aminoacrylate electrophile formed with serine and PLP in its β-subunit (TrpB). Early work demonstrated that tryptophan synthase could accept some indole derivatives or other heterocyclic nucleophiles for the synthesis of tryptophan analogues[40]. Buller and co-workers engineered the β-subunit of tryptophan synthase from *Pyrococcus furiosus* (*Pf*TrpB) to serve as a stand-alone enzyme for non-canonical amino acid (ncAA) synthesis with different nucleophiles[41,42]. Further engineering of *Pf*TrpB expanded the scope of the serine electrophile to include threonine and other β-alkyl serine derivatives for production of β-substituted tryptophan analogues[43,44]. Recently, Romney and co-workers reported that an evolved P*f*TrpB variant, P*f*(NMB), can accommodate nitroalkane nucleophiles structurally distinct from the indole analogues, amines and thiols that were demonstrated previously (Fig. 3b)[45]. As the nitro group can serve as a handle for further modification, this biocatalytic strategy provides a convenient route to diverse ncAAs with aryl and alkyl side chains.

Compared with the diversity of small-molecule catalyst scaffolds invented by chemists, nature uses a more limited set of organocofactors for catalysis. However, the catalytic potential of these cofactors is still far from fully discovered or explored. Different protein structures or reaction conditions may completely alter the properties of the cofactors, a feature that natural evolution has exploited to create functionally diverse enzyme families. This chemical flexibility provides opportunities to use the diverse electronic and photo-chemical properties of cofactors to develop new reaction pathways that have not been explored by nature. Taking inspiration from the studies described here, we imagine that future efforts to the cofactors described here and others such as tetrahydrobiopterin (THB)[46], 4-methylideneimidazole-5-one (MIO)[47] and prenylated flavin (prFMN)[48], will lead to the discovery of yet more functionally diverse enzymes.

Taming metalloenzymes for non-native reactions

Natural enzymes also use metal ions or metal-based cofactors to implement diverse, challenging transformations, as exemplified by nitrogenase for nitrogen fixation. The versatility of transition metal electronic states and coordination modes lays the foundation for transition metal catalysis in chemistry. This versatility also provides opportunities to develop new chemistries starting from nature's vast collection of metalloproteins.

A given metalloenzyme family can encompass diverse functions, but usually their reactions proceed via a specific type of metallo-intermediate. For instance, iron(II)- and α-ketoglutarate-dependent (Fe/αKG) enzymes employ a high-valent iron–oxo (Fe^{IV}) intermediate for C–H hydroxylation, desaturation of aliphatic hydrocarbons, epoxidation of olefins, epimerization of sp^3-hybridized carbon centres and others. SyrB2, an Fe/αKG enzyme from the syringomycin biosynthetic pathway of *Pseudomonas syringae* B301D, is responsible for C–H halogenation of the side-chain methyl group of a threonine moiety tethered with its carrier protein SyrB1[49,50]. Mechanistically, a homolytic coupling between a halogen ligand of iron and a carbon-centred radical formed through a hydrogen atom abstraction process results in the carbon–halogen bond formation. This halogenation activity is thought to originate from the hydroxylation activity of Fe/αKG enzyme homologues. Based on further mechanistic study, Bollinger and co-workers discovered that incorporation of a non-oxygen ligand at the iron centre could lead to new enzymatic functions. Azide or nitrite anions can bind to the iron centre of SyrB2, thus proceeding through radical azidation and nitration with different amino acid-based substrates[51]. These reactions still require carrier protein-appended substrates and can only occur in modest yields under single-turnover conditions, but such activity provides an unprecedented enzymatic route to C–N bond formation with aliphatic C–H bonds. Future work on identifying other homologues and enzyme engineering may further expand the C–H functionalization chemistries of Fe/αKG enzymes and also allow these enzymes to accept diverse substrates for synthetic purposes[52].

Iron–oxo-mediated alkene epoxidation through oxo transfer to C–C double bonds is a well-established transformation with iron-based oxygenases, such as Fe/αKG or cytochrome P450 enzymes. Epoxidation typically follows a concerted [2+1] cycloaddition pathway with a low energy barrier. With a particular cytochrome P450, however, aldehydes were observed as side products of this epoxidation reaction[53]. Hammer et al. hypothesized that this promiscuous activity resulted from a stepwise pathway of radical addition of iron–oxo to the alkene substrate, single electron transfer and a subsequent 1,2-hydride migration to deliver the *anti*-Markovnikov oxidation product[54]. Hammer then engineered this P450 from the rhodobacterium *Labrenzia aggregata* ($P450_{LA1}$) to divert the iron–oxo intermediate into this *anti*-Markovnikov oxidation over the kinetically favoured epoxidation. Accumulation of mutations in the haem domain of $P450_{LA1}$ through directed evolution led to a variant (aMOx) that catalyses the *anti*-Markovnikov oxidation of styrene to phenylacetaldehyde with 3,800 TTN and 81% selectivity. By providing a chiral environment in the enzyme's active site for this stepwise oxo-transfer pathway, the first example of enantioselective *anti*-Markovnikov oxidation was demonstrated with a prochiral α-methylstyrene substrate, giving 82% e.e.

The above examples demonstrate the feasibility of hijacking key intermediates in the catalytic cycles of metalloenzymes into different reaction pathways to access new enzymatic activities. Alternatively, one could also enable new chemistries by introducing new reactive intermediates that mimic the native intermediates, structurally and functionally[55]. Metalloporphyrin complexes have been investigated for more than half a century as structural analogues to haem cofactors in proteins. Many metalloporphyrins were developed in an effort to mimic the oxo-transfer activities of cytochrome P450s through the formation of high-valent metal–oxo intermediates[56]. Meanwhile, intermediates that are analogous to metal–oxo species, typically metallocarbenes and metallonitrenes, could also be formed by the metalloporphyrins and used for carbene and nitrene transfer reactions[57,58]. For instance, a precursor such as a diazo compound can react with the transition metal centre (for example, iron, cobalt, rhodium, ruthenium, iridium, osmium and so on) in metalloporphyrin complexes to generate a metal–carbenoid intermediate that can undergo transfer to organic molecules[59,60]. Reasoning that translating the non-natural activities of metalloporphyrin catalysts to the corresponding haem-dependent enzymes would be a promising way to access new enzymatic activities, Coelho and co-workers described the first 'carbene transferase' enzyme in 2013 (Fig. 4)[61].

Cytochrome P450 from *Bacillus megaterium* ($P450_{BM3}$) catalyses alkene cyclopropanation via an abiological iron–carbenoid intermediate[61]. A diazo reagent, ethyl diazoacetate, reacts with the cytochrome P450 in the Fe^{II} state to yield an iron–carbenoid intermediate; subsequent carbene transfer to styrene substrates leads to the corresponding cyclopropane products. A cysteine-to-serine mutation at the haem-ligating residue furnished a new set of enzymes, designated P411s; this mutation increased the reduction potential of the ferric state of the iron centre, allowing endocellular reductant NADPH to reduce the ferric state and thus confer carbene transfer activity in vivo[62,63]. Histidine ligation was also structurally tolerated by the P450, where it accelerated cyclopropanation of acrylamide substrates[64]. Nature offers a diversity of haem proteins, and our group and the Fasan group have demonstrated that various small haem proteins, including protoglobin, nitric oxide dioxygenase[65] and myoglobin[66], could also be engineered to catalyse cyclopropanation reactions in high efficiency and stereoselectivity, even with electron-deficient and electron-neutral alkenes.

Chen and co-workers hypothesized that an enzyme could catalyse carbene transfer to alkynes to construct highly strained cyclopropenes and that a second carbene-transfer step would give even more strained bicyclo[1.1.0]butane structures[67]. Despite very few precedents of bicyclobutane formation using a carbene transfer strategy, an evolved P411 variant, P411-E10 V78F S438A, adopted this approach to synthesize bicyclobutanes through successive carbene addition to phenylacetylene substrates (Fig. 4a). With aliphatic alkyne substrates, two stereo-complementary P411 variants were obtained for the enantiodivergent synthesis of cyclopropenes. The identity of the amino acid residue at position 87 determines the stereochemistry.

Aziridination via nitrene transfer to alkenes can also be achieved with P411 haem proteins. Our group initially demonstrated aziridination of styrene-type substrates with engineered P411 variants using tosyl azide as the reactive nitrene precursor[68]. After this work, Ohnishi and co-workers disclosed that BezE, a cytochrome P450 in the biosynthetic pathway of benzastatin from *Streptomyces* sp. RI18, is responsible for an aziridination process via formation of an iron-nitrenoid intermediate with an *N*-acetoxy substrate and a subsequent nitrene transfer to a proximal double bond[69]. This is an excellent example of how non-natural biocatalysis can illuminate underexplored paths and lead to new discoveries in biological chemistry. In turn, the Ohnishi study inspired the Arnold group to look at hydroxylamine-type nitrene precursors, which can be more accessible and have fewer stability issues compared to the azide compounds. Cytochrome *c* (an electron-transfer protein) was recently shown to be capable of styrene aziridination by using *O*-pivaloyl hydroxylammonia triflate as the reagent to generate a putative unprotected nitrene species[70]. These aziridines are labile under aqueous conditions and undergo hydrolysis to afford unprotected chiral 1,2-amino alcohols.

In carbene- and nitrene-transfer chemistries, alkene cyclopropanation and aziridination are analogous to the native epoxidation activities of P450s; the abiological counterparts to C–H hydroxylation are C–H alkylation and C–H amination. Early in

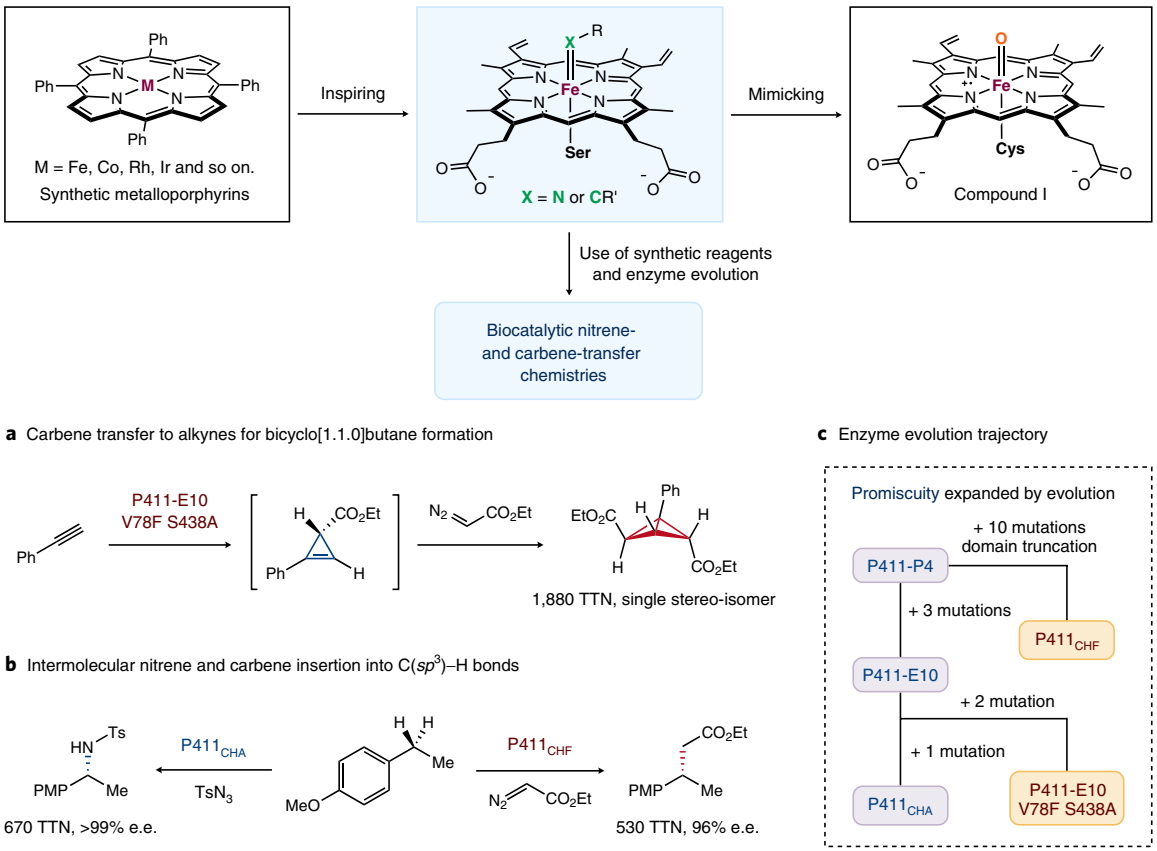

Fig. 4 | Chemomimetic carbene- and nitrene-transfer chemistries with engineered haem proteins. a, Carbene transfer to alkynes for bicyclo[1.1.0]butane formation[67]. **b**, Intermolecular carbene and nitrene insertion into C(sp^3)–H bonds[78,79]. **c**, Rapid expansion of the promiscuity of P450s to carbene and nitrene transfer (in **4a** and **4b**) by laboratory evolution.

the 1980s, Dawson and co-workers reported that a rabbit-liver cytochrome P450 catalysed a nitrene C–H insertion reaction using an abiological iminoiodinane substrate as nitrene precursor with very limited turnovers[71]. Inspired by this work, McIntosh et al. achieved intramolecular nitrene insertion into a proximal C(sp^3)–H bond using a P411 as a whole-cell catalyst, affording sultam products in decent yield and good enantioselectivity[72]; and nitrene insertion into primary, secondary and tertiary C(sp^3)–H bonds could all be achieved with engineered P411 variants for the efficient and selective synthesis of various diamine products[73]. Fasan also showed that $P450_{BM3}$[74] or myoglobin[75] variants are capable of such intramolecular C–H amination reactions. Furthermore, the evolved P450 variants can also function with azidoformates for intramolecular C–N bond formation, providing oxazolidinones as the amination products[76].

Engineering site-selective C–H amination is of great interest, since regioselectivity with small-molecule catalysis is usually dominated by the inherent properties of the C–H bonds. Hyster and co-workers demonstrated that the active site of P450 can be reshaped to facilitate C–H amination in a regiodivergent manner[77]. With a sulfonylazide bearing two types of C(sp^3)–H bonds that are geometrically accessible to amination, different P411 variants were evolved to direct intramolecular nitrene insertion to benzylic and homobenzylic C(sp^3)–H bonds, forming five- and six-membered sultam products selectively. Given the difference in bond dissociation energies of the C–H bonds, the distinct product outcomes show that properly engineering the active-site environment for catalysis can override the inherent reactivity of the C–H bonds and guide product formation along desired reaction pathways.

P411s were also engineered to catalyse intermolecular nitrene C–H insertion enantioselectively[78], a problem with few examples of solutions for small-molecule catalysts. Accumulation of active-site mutations in a P411 enzyme helped to precisely orient the substrate in the distal pocket and accelerate the desired C–H insertion. Using tosyl azide as the nitrene precursor, the evolved $P411_{CHA}$ enantioselectively aminated benzylic C–H bonds of alkyl benzene substrates. More importantly, this $P411_{CHA}$ variant provided a versatile platform for evolving biocatalysts for diverse carbene- or nitrene-transfer reactions inaccessible to chemical catalysis, such as bicyclobutane formation.

Carbene C–H insertion to install alkyl groups onto organic molecules is also feasible with cytochrome P411 variants[79,80]. A variant obtained in the evolutionary lineage of $P411_{CHA}$ displayed promiscuous activity for carbene insertion into C–H bonds and thus served as a parent for evolution of a powerful alkyl transferase, $P411_{CHF}$, which can target benzylic, allylic, propargylic and α-amino C–H bonds for carbene insertion. Interestingly, $P411_{CHF}$ alkylates a benzylic C–H bond in a substrate bearing a terminal alkene moiety, whereas an early P411 variant, P-I263F, only cyclopropanates the double bond in the same molecule. This demonstration of catalyst-controlled chemoselectivity once again speaks to the high tunability of enzyme catalysis, which provides a promising solution to long-standing selectivity challenges.

Analogous to heteroatom oxidation by cytochrome P450s, electrophilic nitrene or carbene intermediates formed with haem proteins (such as P411 or myoglobin variants) can be intercepted by sulfides to furnish sulfimides[81] and sulfonium ylides[82]. Moreover, with prochiral allylic sulfides, the corresponding allylic sulfimide

or sulfonium ylide products can further undergo [2,3]-sigmatropic rearrangement to yield chiral allylic amines[83] or chiral sulfides[81].

Another useful class of carbene-transfer reactions is X–H bond insertion (where X is a heteroatom, including N, S, Si, B, P and others). Our group showed that $P450_{BM3}$ variants can catalyse *N*-alkylation of aniline substrates via a formal carbene N–H insertion process[84]; the Fasan laboratory also described N–H insertion[85] and S–H insertion[86] reactions with engineered myoglobins. Chen et al. recently reported a P450-catalysed enantioselective S–H insertion reaction using a lactone-derived carbene that proceeds through a radical mechanism[87].

Carbene Si–H or B–H insertion reactions provide efficient routes for building C–Si and C–B bonds, which are not found in biological systems but are useful and important in human-made products. Haem proteins are capable of forming these bonds with much higher activities than reported for transition metal catalysts. Wild-type cytochrome *c* from *Rhodothermus marinus* (*Rma* cyt *c*) was discovered to catalyse carbene Si–H insertion using ethyl 2-diazopropanoate and phenyldimethylsilane as substrates with a modest turnover (44 TTN) but good enantioselectivity (97% e.e.)[88]. Introduction of three active-site mutations improved catalytic efficiency by over 30-fold. Evaluation of silane scope established that the evolved *Rma* cyt *c* is particularly selective for the desired silylation even with substrates bearing other functionalities that can participate in carbene-transfer chemistry. *Rma* cyt *c* was also engineered for enantioselective carbene B–H insertion reactions using *N*-heterocyclic carbene-stabilized boranes[89–91]. By modifying the active-site structure of *Rma* cyt *c*, a variety of structurally different carbenes can be accommodated for this B–H insertion reaction.

With the established carbene chemistries of haem proteins, Lewis was able to capture a carbene intermediate bound to the iron centre of an *Rma* cyt *c* mutant in a crystal structure[92]. The physical chemistry study of this iron–carbenoid species together with computational investigation provided insight into how protein structure enables the desired chemistries. The Hilvert laboratory also characterized a different type of carbene adduct with an N_δ-methylhistidine-ligated myoglobin variant, a Fe(III)–C–N(pyrrole)-bridged carbenoid, which could equilibrate with an end-on carbenoid isomer and thus participate in carbene-transfer reactions[93].

These newly discovered carbene- and nitrene-transfer activities of haem proteins expand nature's catalytic repertoire to include many transformations that are not biologically relevant but are highly useful for chemical synthesis. For most of the reactions described here, the free haem cofactor catalyses the reaction not at all or only with very poor efficiency, highlighting the contribution of the protein to enabling and facilitating these transformations. Although small-molecule catalysts have been developed for many of these chemistries, haem proteins stand out as competent catalysts with high catalytic efficiency and readily tunable stereo-/regio-/chemoselectivities. With the help of strategies to discover and improve new carbene and nitrene transferases, we foresee that haem protein biocatalysts will address more challenging problems in synthetic chemistry and will move to wider use at scale.

Developing new enzymes with artificial cofactors

The above examples show how small-molecule catalysis can inspire discovery of new enzyme functions. However, there are still many synthetically important reactions carried out with human-invented catalysts for which enzyme candidates have not yet been identified. Chemists have been trying to fill some of this large gap between classical catalysis and biocatalysis by incorporating catalytically competent artificial cofactors into proteins. Artificial metalloenzymes (ArMs), for example, can be traced back to the late 1970s when Wilson and Whitesides assembled an artificial metallohydrogenase for hydrogenation of α-acetamidoacrylic acid by introducing a biotin-tethered diphosphine–rhodium(I) complex to avidin[94]. Most early studies focused on proving that new enzymes with human-invented metallocatalysts could be made, and many functions performed by these ArMs, such as ester/amide hydrolysis, alcohol/olefin oxidation and ketone/imine/acrylate reduction, were already well-known for enzymes despite mechanistic differences. More recently, the development of ArMs has been greatly accelerated by advances in organometallic chemistry and protein engineering. A broad range of ArMs have been created for important transformations in synthetic chemistry. A recent Review from the Ward and Lewis groups comprehensively summarizes work in the field of ArMs[95]. Here we will focus only on representative ArMs with abiological functions.

Non-covalent binding of metallo-cofactors using specific protein ligands is a strategy used widely to construct ArMs. The biotin-(strept)avidin system is a typical example, as avidin and streptavidin (Sav) feature deep binding pockets for biotin and provide a chiral environment for the catalytic centre. Ward and co-workers applied this to the creation of palladium-ArMs for Suzuki cross-coupling[96] and allylic alkylation[97] and rhodium-ArMs for C–H activation/annulation reactions[98]. Based on their early work on ruthenium-ArMs for olefin metathesis[99], the team developed a system for selective assembly of artificial 'metathases' within the periplasm of *Escherichia coli* cells through the fusion of Sav with the signal peptide OmpA[100]. This in vivo construction strategy substantially expedited application of directed evolution, leading to an evolved metathase with higher activity than the free 2nd-generation Grubbs catalyst towards a ring-closing metathesis reaction. Recently, the Rovis lab also reported a monomeric streptavidin (mSav) Rh(III)-ArM for enantioselective C–H activation/annulation with acrylamide hydroxamate esters and styrenes for the synthesis of a variety of substituted δ-lactams (Fig. 5a)[101]. Artificial transfer hydrogenases (ATHs) based on iridium cofactors have been constructed with the biotin-Sav framework and engineering the protein has enabled enantioselective reduction of ketone, imine and quinoline substrates with high efficiency[102,103]. A similar non-covalent anchoring strategy based on the high binding affinity of an iron(III)-containing azotochelin complex with protein CeuE was also employed by Duhme-Klair for the assembly of an iridium-based ATHase; interestingly, the (dis)association of the cofactor could be tuned through the different oxidation states of the iron centre[104]. Furthermore, a recent study of an ArM that is created with an albumin binding protein demonstrated prodrug activation of an anticancer agent through ring-closing metathesis[105].

ArMs can also be assembled through covalent linkages of metallo-cofactors to the protein. The thiol group on cysteine is typically used as a handle for covalent assembly via nucleophilic substitution or conjugate addition, as demonstrated for the ruthenium-based olefin metathases reported by Hilvert[106] and Rh(I)-based olefin hydroformylase reported by Jarvis and Kamer[107]. However, selective ArM assembly with this thiol-linkage strategy is difficult when the protein has multiple accessible cysteine residues. Lewis and co-workers thus used a genetically encoded *p*-azidophenylalanine (pN_3Phe) for specific coupling with strained alkyne-modified metallo-cofactors through click chemistry[108]. They selected a prolyl oligopeptidase (POP) featuring a large internal cavity as the protein scaffold and constructed diRh(II)-ArMs through the cycloaddition between a strained alkyne pre-installed on a dirhodium catalyst and a (pN_3Phe) residue in the active site of POP[109]. This rapid and selective assembly system allowed them to develop a practical platform to evolve diRh(II)-ArMs for stereoselective cyclopropanation of styrenes with donor-acceptor diazos using random mutagenesis and screening (Fig. 5b)[110].

Binding of metals or metal complexes with coordinating residues represents an alternative strategy for constructing ArMs. Ueno, Watanabe and co-workers reported a 'Suzukiase', enabled by ligation of a $[Pd(allyl)Cl]_2$ complex in apo-ferritin[111]. Metallation

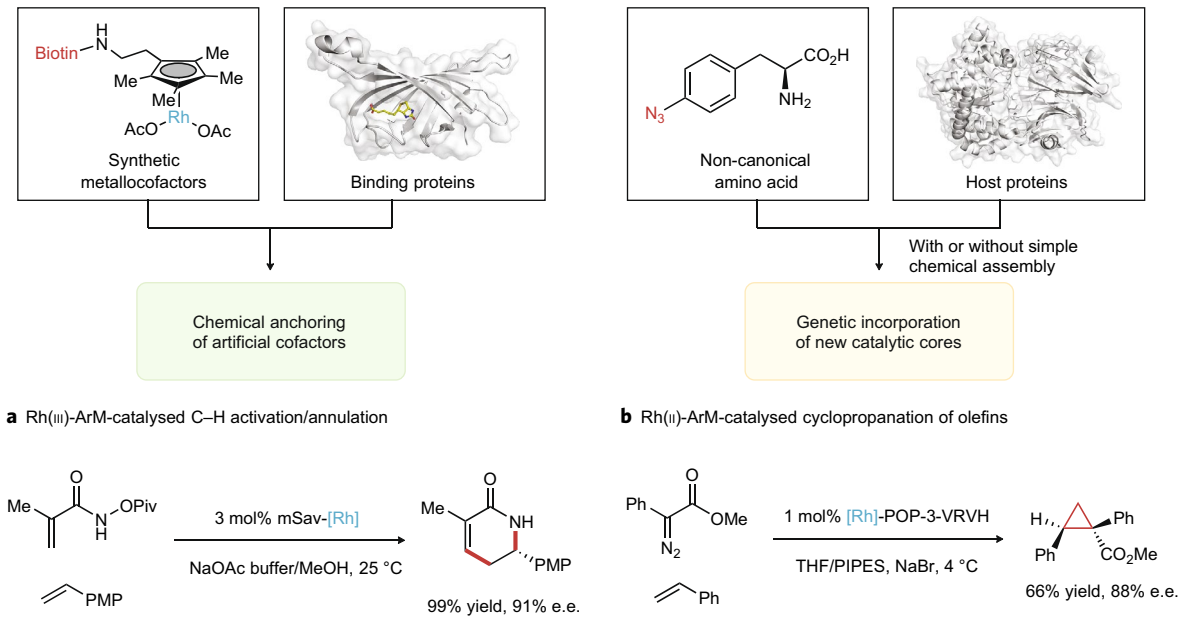

Fig. 5 | Different strategies for artificial enzyme construction. a, Artificial Rh(III)-enzyme-catalysed C–H activation/annulation[101]. **b**, Artificial diRh(II)-enzyme-catalysed olefin cyclopropanation[110]. mSav, monomeric streptavidin.

of native carbonic anhydrase with a [Rh(acac)$(CO)_2$] complex generated a hydroformylase, and further protein engineering helped to improve the overall enzyme activity and selectivity for the linear aldehyde product over the branched one[112]. The Roelfes group also utilized coordinating ncAAs to bind metal ions for new catalytic functions, as exemplified by the Cu(II)-ArMs for Friedel–Crafts alkylation of indoles[113] and hydration with α,β-unsaturated ketones[114]. A cofactor switch strategy has also been realized to generate ArMs. Hartwig, Clark and co-workers explored a set of iridium-substituted ArMs generated from myoglobin[115] and cytochrome P450[116] for carbene C–H insertion reactions. Engineered Ir-ArMs are also able to carry out cyclopropanation of unactivated or internal alkenes[117] and intramolecular nitrene C–H insertion[118]. It is interesting to note that most of these functions have been demonstrated with engineered iron-haem proteins that are fully genetically encoded.

Artificial enzymes with non-natural catalytic centres can be constituted in a fully genetically encoded scenario using ncAA incorporation. A wide range of ncAAs can now be genetically encoded and incorporated into protein scaffolds, which allows for the introduction of unnatural cofactors for new catalytic functions. A recent Review by Budisa summarizes the development of biocatalysts using this genetic strategy[119]. Most examples only employ ncAAs to tune the properties of enzyme active sites or natural cofactors and to enhance the native catalytic functions. Until now there were only a few examples of using catalytically functional ncAAs for new enzymatic activity, among which is the use of *p*-aminophenylalanine for catalytic condensation of carbonyls with hydrazines and hydroxylamines by the Roelfes' group[120,121]. Recently, the Green laboratory designed and evolved an artificial hydrolytic enzyme using an unconventional mechanism based on a non-canonical N_δ-methylhistidine (Me-His) as the catalytically functional residue[122].

The reaction design based on the selection of suitable catalytic scaffolds together with protein engineering has produced artificial enzymes for synthetically important chemical transformations. However, compared to natural enzymes, most artificial enzymes still exhibit low catalytic efficiency with limited turnovers and are usually not as versatile as the best small-molecule catalysts for the same type of reactions. Preparation of most ArMs still requires tedious processes, including chemical synthesis of specific metal cofactors, purification of the apoproteins followed by assembly steps and sometimes removal of excess cofactor from the system, which render tuning by directed evolution difficult. Developing versatile, amenable, readily evolvable systems is an important challenge for future research.

Conclusions

The third wave of biocatalysis, starting in the early 1990s and empowered by directed evolution and other methods, saw solutions to many practical problems in enzyme catalysis, including enzyme stability issues, limitations in substrate breadth, efficiency, selectivity and others[123]. These developments laid a solid foundation for widespread adoption of biocatalysis in pharmaceuticals, fine chemicals, agriculture, materials and more[124]. In a recent Perspective, Bornscheuer describes a fourth wave of biocatalysis[125] with discovery of new enzyme classes and development of non-natural activities as major new directions. These efforts lead not only to a broader appreciation of enzymes' capabilities but also fulfil the demand for new, sustainable methods in organic synthesis[126].

We predict that enzymes invented in the laboratory will become powerful complements and alternatives to synthetic catalysts. For example, engineered haem protein carbene and nitrene transferases, unknown less than ten years ago, are capable of catalysing diverse transformations that are also accessible with synthetic catalysts based on rhodium, iridium and other transition metals. However, directed evolution has enabled the enzymes to display orders of magnitude higher turnovers, using an earth-abundant iron centre; the enzymes also have selectivities that none of the small-molecule catalyst can offer. Perhaps most exciting, the enzymes can make molecules that are inaccessible to small-molecule catalysts (for example, pure stereoisomers of bicyclobutanes[67]). It is thrilling to realize that the ability of enzymes to control reaction intermediates and specifically accelerate a desired reaction can now be used to control the chemistry invented by us!

On the other hand, many of the new enzymes described here still exhibit low catalytic efficiencies, limited substrate ranges and moderate selectivities. It is reasonable to think that a natural enzyme

co-opted for new chemistry or an artificial enzyme assembled from a protein scaffold and a synthetic cofactor does not have an active site that is optimal for the entire mechanistic pathway of a new reaction. Directed evolution can step in to reorganize enzyme structures for non-natural catalysis, but this requires systems for mutagenesis and screening to identify beneficial changes in the protein sequence. We can anticipate that improved rational protein design as well as new protein engineering methods based on machine learning will help navigate the landscape of enzyme activities and protein sequence to guide further engineering with reduced experimental effort[127].

The field still has to grapple with the fact that the catalytic repertoire of enzymes is still quite restricted compared to synthetic methods. For example, organofluorine moieties are particularly important in medicinal chemistry, but until now only one class of enzyme, fluorinase, is known to catalyse a C–F bond-forming reaction with high substrate specificity[128]. And bimolecular cycloadditions, developed by synthetic chemists to build various ring structures in a modular way, are barely utilized by natural enzymes[129]. Furthermore, most newly identified enzymes are not robust enough for synthetic or industrial application. Compared to transaminases, for example, which natively catalyse C–N bond formation and are also used widely in the pharmaceutical industry[130], the recent enzymatic C–H amination strategy using nitrene transfer to a C–H bond provides a straightforward way to get to the same targets without needing a preinstalled carbonyl functionality. However, utility for synthetic purposes necessitates further improvements in the enzyme and reaction engineering; the enzymes also have to be made broadly available to users. In principle, genetically encoded catalysts are available to anyone with access to the sequence. In practice, however, few synthetic laboratories have the expertise and equipment to exploit them.

Overall, we see a bright future for enzymes in a world that needs clean, efficient catalysts. New activities will be discovered at an ever-faster pace as chemists look at enzymes with their goals in mind. The current challenges are feasible and worthy targets for creative problem solvers.

Received: 25 August 2019; Accepted: 21 October 2019;
Published online: 20 January 2020

References

1. Hönig, M., Sondermann, P., Carreira, E. M. & Turner, N. Enantioselective chemo- and biocatalysis: partners in retrosynthesis. *Angew. Chem. Int. Ed.* **56**, 8942–8973 (2017).
This Review covers many examples of designing retrosynthetic routes with enzymes.
2. de Souza, R. O. M. A., Miranda, L. S. M. & Bornscheuer, U. T. A retrosynthetic approach for biocatalysis in organic synthesis. *Chem. Eur. J.* **23**, 12040–12063 (2017).
3. Turner, N. J. & Humphrey, L. *Biocatalysis in Organic Synthesis: The Retrosynthesis Approach* (Royal Society of Chemistry, 2018).
4. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of *de novo* protein design. *Nature* **537**, 320–327 (2016).
5. Zanghellini, A. *de novo* computational enzyme design. *Curr. Opin. Chem. Biol.* **29**, 132–138 (2014).
6. Khersonsky, O. & Tawfik, D. S. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **79**, 471–505 (2010).
An in-depth Review of how to recognize and understand enzyme promiscuity.
7. Tawfik, D. S. Messy biology and the origins of evolutionary innovations. *Nat. Chem. Bio.* **6**, 692–696 (2010).
8. Busto, E., Gotor-Fernández, V. & Gotor, V. Hydrolases: catalytically promiscuous enzymes for non-conventional reactions in organic synthesis. *Chem. Soc. Rev.* **39**, 4504–4523 (2010).
9. Campos, K. R. et al. The importance of synthetic chemistry in the pharmaceutical industry. *Science* **363**, eaat0805 (2019).
10. Bornscheuer, U. T. & Kazlauskas, R. J. Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed.* **43**, 6032–6040 (2004).
11. Branneby, C. et al. Carbon–carbon bonds by hydrolytic enzymes. *J. Am. Chem. Soc.* **125**, 874–875 (2003).
12. Mayer, C. et al. The E358S mutant of *Agrobacterium* sp. β-glucosidase is greatly improved glycosynthase. *FEBS Lett.* **466**, 40–44 (2000).
13. Jahn, M. et al. Thioglycoligases: mutant glycosidases for thioglycoside synthesis. *Angew. Chem. Int. Ed.* **42**, 352–354 (2003).
14. Schallmey, A. & Schallmey, M. Recent advances on halohydrin dehalogenases—from enzyme identification to novel biocatalytic applications. *Appl. Microbiol. Biotechnol.* **100**, 7827–7839 (2016).
15. Fox, R. J. et al. Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344 (2007).
16. Syrén, P.-O. et al. Squalene-hopene cyclases—evolution, dynamics and catalytic scope. *Curr. Opin. Chem. Biol.* **41**, 73–82 (2016).
17. Seitz, M. et al. Synthesis of heterocyclic terpenoids by promiscuous squalene–hopene cyclases. *ChemBioChem* **14**, 436–439 (2013).
18. Hammer, S. C. et al. Squalene hopene cyclases are protonases for stereoselective Brønsted acid catalysis. *Nat. Chem. Bio.* **11**, 121–126 (2015).
19. Mangas-Sanchez, J. et al. Imine reductases (IREDs). *Curr. Opin. Chem. Biol.* **37**, 19–25 (2017).
20. Sharma, M. et al. A mechanism for reductive amination catalyzed by fungal reductive aminases. *ACS Catal.* **8**, 11534–11541 (2018).
21. Aleku, G. A. et al. A reductive aminase from *Aspergillus oryzae*. *Nat. Chem.* **9**, 961–969 (2017).
22. Emmanuel, M. A., Greenberg, N. R., Oblinsky, D. G. & Hyster, T. K. Accessing non-natural reactivity by irradiating nicotinamide-dependent enzymes with light. *Nature* **540**, 414–417 (2016).
This ground-breaking example of photo-induced enzyme promiscuity using the excited states of biological cofactors opens a whole new route to accessing non-natural enzyme activities.
23. Biegasiewicz, K. F. et al. Catalytic promiscuity enabled by photoredox catalysis in nicotinamide-dependent oxidoreductases. *Nat. Chem.* **10**, 770–775 (2018).
24. Dockrey, S. A. B. & Narayan, A. R. H. Flavin-dependent biocatalysts in synthesis. *Tetrahedron* **75**, 1115–1121 (2019).
25. Sandoval, B. A., Meichan, A. J. & Hyster, T. K. Enantioselective hydrogen atom transfer: discovery of catalytic promiscuity in flavin-dependent 'ene'-reductases. *J. Am. Chem. Soc.* **139**, 11313–11316 (2017).
26. Sandoval, B. A. et al. Photoenzymatic catalysis enables radical-mediated ketone reduction in ene-reductases. *Angew. Chem. Int. Ed.* **58**, 8714–8718 (2019).
27. Biegasiewicz, K. F. et al. Photoexcitation of flavoenzymes enables a stereoselective radical cyclization. *Science* **364**, 1166–1169 (2019).
28. Sorigué, D. et al. An algal photoenzyme converts fatty acids to hydrocarbons. *Science* **357**, 903–907 (2017).
29. Zhang, W. et al. Hydrocarbon synthesis via photoenzymatic decarboxylation of carboxylic acids. *J. Am. Chem. Soc.* **141**, 3116–3120 (2019).
30. Meyer, D., Neumann, P., Ficner, R. & Tittmann, K. Observation of a stable carbene at the active site of a thiamin enzyme. *Nat. Chem. Bio.* **9**, 488–490 (2013).
31. Breslow, R. On the mechanism of thiamine action. IV.[1] Evidence from studies on model systems. *J. Am. Chem. Soc.* **80**, 3719–3726 (1958).
32. Loschonsky, S. et al. Catalytic scope of the thiamine-dependent multifunctional enzyme cyclohexane-1,2-dione hydrolase. *ChemBioChem* **15**, 389–392 (2014).
33. Loschonsky, S. et al. Extended reaction scope of thiamine diphosphate dependent cyclohexane-1,2-dione hydrolase: From C–C bond cleavage to C–C bond ligation. *Angew. Chem. Int. Ed.* **53**, 14402–14406 (2014).
34. Bernacchia, G. et al. Enzymatic chemoselective aldehyde–ketone cross-couplings through the polarity reversal of methylacetoin. *Angew. Chem. Int. Ed.* **54**, 7171–7175 (2015).
35. Demir, A. S. et al. Enantioselective synthesis of α-hydroxy ketones via benzaldehyde lyase-catalyzed C–C bond formation reaction. *Adv. Synth. Catal.* **344**, 96–103 (2002).
36. Dünkelmann, P. et al. Development of a donor–acceptor concept for enzymatic cross-coupling reactions of aldehydes: the first asymmetric cross-benzoin condensation. *J. Am. Chem. Soc.* **124**, 12084–12085 (2002).
Excellent example of how enzyme catalysis can solve long-standing chemoselectivity problems in synthetic methodologies.
37. Dresen, C. et al. The enzymatic asymmetric conjugate umpolung reaction. *Angew. Chem. Int. Ed.* **49**, 6600–6603 (2010).
38. Tai, C.-H. & Cook, P. F. Pyridoxal 5′-phosphate-dependent α, β-elimination reactions: mechanism of *O*-acetylserine sulfhydrylase. *Acc. Chem. Res.* **34**, 49–59 (2001).
39. Maier, T. H. P. Semisynthetic production of unnatural l-α-amino acids by metabolic engineering of the cysteine-biosynthetic pathway. *Nat. Biotechnol.* **21**, 422–427 (2003).
40. Phillips, R. S. Synthetic applications of tryptophan synthase. *Tetrahedron.: Asymmetry* **15**, 2787–2792 (2004).
41. Buller, A. R. et al. Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. *Proc. Natl. Acad. Sci. USA* **112**, 14599–14604 (2015).

42. Romney, D. K. et al. Unlocking reactivity of TrpB: a general biocatalytic platform for synthesis of tryptophan analogues. *J. Am. Chem. Soc.* **139**, 10769–10776 (2017).
43. Boville, C. E. et al. Engineered biosynthesis of β-alkyl tryptophan analogs. *Angew. Chem. Int. Ed.* **57**, 14764–14768 (2018).
44. Herger, M. et al. Synthesis of β-branched tryptophan analogues using an engineered subunit of tryptophan synthase. *J. Am. Chem. Soc.* **138**, 8388–8391 (2016).
45. Romney, D. K., Sarai, N. S. & Arnold, F. H. Nitroalkanes as versatile nucleophiles for enzymatic synthesis of noncanonical amino acids. *ACS Catal.* **9**, 8726–8730 (2019).
46. Werner, E. R., Blau, N. & Thöny, B. Tetrahydrobiopterin: biochemistry and pathophysiology. *Biochem J.* **438**, 397–414 (2011).
47. Heberling, M. M., Wu, B., Bartsch, S. & Janssen, D. B. Priming ammonia lyases and aminomutases for industrial and therapeutic applications. *Curr. Opin. Chem. Biol.* **17**, 250–260 (2013).
48. Marshall, S. A., Payne, K. A. P. & Leys, D. The UbiX–UbiD system: the biosynthesis and use of prenylated flavin (prFMN). *Arch. Biochem. Biophys.* **632**, 209–221 (2017).
49. Vaillancourt, F. H., Yin, J. & Walsh, C. T. SyrB2 in *syringomycin E* biosynthesis is a nonheme Fe^{II} α-ketoglutarate- and O_2-dependent halogenase. *Proc. Natl. Acad. Sci. USA* **102**, 10111–10116 (2005).
50. Wong, S. D. et al. Elucidation of the iron(iv)–oxo intermediate in the non-haem iron halogenase SyrB2. *Nature* **499**, 320–323 (2013).
51. Matthews, M. L. et al. Direct nitration and azidation of aliphatic carbons by an iron-dependent halogenase. *Nat. Chem. Bio.* **10**, 209–215 (2014).
52. Neugebauer, M. E. et al. A family of radical halogenases for the engineering of amino-acid-based products. *Nat. Chem. Bio.* **10**, 209–215 (2014).
53. de Visser, S. P., Kumar, D. & Shaik, S. How do aldehyde side products occur during alkene epoxidation by cytochrome P450? Theory reveals a state-specific multi-state scenario where the high-spin component leads to all side products. *J. Inorg. Biochem.* **98**, 1183–1193 (2004).
54. Hammer, S. C. et al. Anti-Markovnikov alkene oxidation by metal-oxo-mediated enzyme catalysis. *Science* **358**, 215–218 (2017).
This study demonstrates how a new enzyme activity can be obtained by promoting an alternative reaction pathway.
55. Prier, C. K. & Arnold, F. H. Chemomimetic biocatalysis: exploiting the synthetic potential of cofactor-dependent enzymes to create new catalysts. *J. Am. Chem. Soc.* **137**, 13992–14006 (2015).
56. Che, C.-M. & Huang, J.-S. Metalloporphyrin-based oxidation systems: from biomimetic reactions to application in organic synthesis. *Chem. Commun.* 3996–4015 (2009).
57. Lu, H. & Zhang, X. P. Catalytic C–H functionalization by metalloporphyrins: recent developments and future directions. *Chem. Soc. Rev.* **40**, 1899–1909 (2011).
58. Che, C.-M., Lo, V. K.-Y., Zhou, C.-Y. & Huang, J.-S. Selective functionalisation of saturated C–H bonds with metalloporphyrin catalysts. *Chem. Soc. Rev.* **40**, 1950–1975 (2011).
59. Maxwell, J. L., Brown, K. C., Bartley, D. W. & Kodadek, T. Mechanism of the rhodium porphyrin-catalyzed cyclopropanation of alkenes. *Science* **256**, 1544–1547 (1992).
60. Wolf, J. R. et al. Shape and stereoselective cyclopropanation of alkenes catalyzed by iron porphyrins. *J. Am. Chem. Soc.* **117**, 9194–9199 (1995).
61. Coelho, P. S., Brustad, E. M., Kannan, A. & Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **339**, 307–310 (2013).
This report of a non-natural carbene-transfer reaction with an engineered haem enzyme laid the foundation for development of novel carbene and nitrene transferases.
62. Coelho, P. S. et al. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins *in vivo*. *Nat. Chem. Biol.* **9**, 485–487 (2013).
63. Brandenberg, O. F. et al. Stereoselective enzymatic synthesis of heteroatom-substituted cyclopropanes. *ACS Catal.* **8**, 2629–2634 (2018).
64. Wang, Z. J. et al. Improved cyclopropanation activity of histidine-ligated cytochrome P450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chem. Int. Ed.* **53**, 6810–6813 (2014).
65. Knight, A. M. et al. Diverse engineered heme proteins enable stereodivergent cyclopropanation of unactivated alkenes. *ACS Cent. Sci.* **4**, 372–377 (2018).
66. Bordeaux, M., Tyagi, V. & Fasan, R. Highly diastereoselective and enantioselective olefin cyclopropanation using engineered myoglobin-based catalysts. *Angew. Chem. Int. Ed.* **54**, 1744–1748 (2015).
67. Chen, K. et al. Enzymatic construction of highly strained carbocycles. *Science* **360**, 71–75 (2018).
68. Farwell, C. C. et al. Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450. *ACS Cent. Sci.* **1**, 89–93 (2015).
69. Tsutsumi, H. et al. Unprecedented cyclization catalyzed by a cytochrome P450 in benzastatin biosynthesis. *J. Am. Chem. Soc.* **140**, 6631–6639 (2018).
70. Cho, I. et al. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angew. Chem. Int. Ed.* **58**, 3138–3142 (2019).
71. Svastits, E., Dawson, J. H., Breslow, R. & Gellman, S. H. Functionalized nitrogen atom transfer catalyzed by cytochrome P-450. *J. Am. Chem. Soc.* **107**, 6427–6428 (1985).
72. McIntosh, J. A. et al. Enantioselective intramolecular C–H amination catalyzed by engineered cytochrome P450 enzymes *in vitro* and *in vivo*. *Angew. Chem. Int. Ed.* **52**, 9309–9312 (2013).
73. Yang, Y. et al. An enzymatic platform for the asymmetric amination of primary, secondary and tertiary C(sp^3)–H bonds. *Nat. Chem.* **11**, 987–993 (2019).
74. Singh, R., Bordeaux, M. & Fasan, R. P450-catalyzed intramolecular sp^3 C–H amination with arylsulfonyl azide substrates. *ACS Catal.* **4**, 546–552 (2014).
75. Bordeaux, M., Singh, R. & Fasan, R. Intramolecular C(sp^3)–H amination of arylsulfonyl azides with engineered and artificial myoglobin-based catalysts. *Bioorg. Med. Chem.* **22**, 5697–5704 (2014).
76. Singh, R., Kolev, J. N., Sutera, P. A. & Fasan, R. Enzymatic C(sp^3)–H amination: P450-catalyzed conversion of carbonazidates into oxazolidinones. *ACS Catal.* **5**, 1685–1691 (2015).
77. Hyster, T. K. et al. Enzyme-controlled nitrogen-atom transfer enables regiodivergent C–H amination. *J. Am. Chem. Soc.* **136**, 15505–15508 (2014).
78. Prier, C. K. et al. Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron–haem enzyme. *Nat. Chem.* **9**, 629–634 (2017).
79. Zhang, R. K. et al. Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp^3 C–H functionalization. *Nature* **565**, 67–72 (2019).
This work highlights the immense potential of natural metalloproteins in addressing important problems in synthetic chemistry, especially those considered to be challenging for earth-abundant transition metal catalysis.
80. Zhang, J., Huang, X., Zhang, R. K. & Arnold, F. H. Enantiodivergent α-amino C–H fluoroalkylation catalyzed by engineered cytochrome P450s. *J. Am. Chem. Soc.* **141**, 9798–9802 (2019).
81. Farwell, C. C. et al. Enantioselective imidation of sulfides via enzyme-catalyzed intermolecular nitrogen-atom transfer. *J. Am. Chem. Soc.* **136**, 8766–8771 (2014).
82. Tyagi, V. et al. Biocatalytic synthesis of allylic and allenyl sulfides through a myoglobin-catalyzed Doyle–Kirmse reaction. *Angew. Chem. Int. Ed.* **55**, 13562–13566 (2016).
83. Prier, C. K. et al. Asymmetric enzymatic synthesis of allylic amines: a sigmatropic rearrangement strategy. *Angew. Chem. Int. Ed.* **55**, 4711–4715 (2016).
84. Wang, Z. J., Peck, N. E., Renata, H. & Arnold, F. H. Cytochrome P450-catalyzed insertion of carbenoids into N–H bonds. *Chem. Sci.* **5**, 598–601 (2014).
85. Sreenilayam, G. & Fasan, R. Myoglobin-catalyzed intermolecular carbene N–H insertion with arylamine substrates. *Chem. Commun.* **51**, 1532–1534 (2015).
86. Tyagi, V., Bonn, R. B. & Fasan, R. Intermolecular carbene S–H insertion catalysed by engineered myoglobin-based catalysts. *Chem. Sci.* **6**, 2488–2494 (2015).
87. Chen, K. et al. Alternate heme ligation steers activity and selectivity in engineered cytochrome P450-catalyzed carbene-transfer reactions. *J. Am. Chem. Soc.* **140**, 16402–16407 (2018).
88. Kan, S. B. J., Lewis, R. D., Chen, K. & Arnold, F. H. Directed evolution of cytochrome *c* for carbon–silicon bond formation: bringing silicon to life. *Science* **354**, 1048–1051 (2016).
89. Kan, S. B. J. et al. Genetically programmed chiral organoborane synthesis. *Nature* **552**, 132–136 (2017).
90. Huang, X. et al. A biocatalytic platform for synthesis of chiral α-trifluoromethylated organoborons. *ACS Cent. Sci.* **5**, 270–276 (2019).
91. Chen, K. et al. Engineered cytochrome *c*-catalyzed lactone-carbene B–H insertion. *Synlett* **30**, 378–382 (2019).
92. Lewis, R. D. et al. Catalytic iron–carbene intermediate revealed in a cytochrome c carbene transferase. *Proc. Natl. Acad. Sci. USA* **115**, 7308–7313 (2018).
This report describes the first iron-carbene intermediate captured in a protein structure and is among the very few examples of a carbene intermediate bound to any type of iron complex.
93. Hayashi, T. et al. Capture and characterization of a reactive haem–carbenoid complex in an artificial metalloenzyme. *Nat. Catal.* **1**, 578–584 (2018).
94. Wilson, M. E. & Whitesides, G. M. Conversion of a protein to a homogeneous asymmetric hydrogenation catalyst by site-specific modification with a diphosphinerhodium(i) moiety. *J. Am. Chem. Soc.* **100**, 306–307 (1978).

95. Schwizer, F. et al. Artificial metalloenzymes: reaction scope and optimization strategies. *Chem. Rev.* **118**, 142–231 (2018).
96. Chatterjee, A. et al. An enantioselective artificial Suzukiase based on the biotin–streptavidin technology. *Chem. Sci.* **7**, 673–677 (2016).
97. Pierron, J. et al. Artificial metalloenzymes for asymmetric allylic alkylation on the basis of the biotin–avidin technology. *Angew. Chem. Int. Ed.* **47**, 701–705 (2008).
98. Hyster, T. K., Knörr, L., Ward, T. R. & Rovis, T. Biotinylated Rh(III) complexes in engineered streptavidin for accelerated asymmetric C–H activation. *Science* **338**, 500–503 (2012).
99. Lo, C. et al. Artificial metalloenzymes for olefin metathesis based on the biotin-(strept)avidin technology. *Chem. Commun.* **47**, 12065–12067 (2011).
100. Jeschek, M. et al. Directed evolution of artificial metalloenzymes for *in vivo* metathesis. *Nature* **537**, 661–665 (2016).
This study describes a platform for *in vivo* artificial metalloenzyme assembly and is one of the rare examples of directed evolution of ArMs.
101. Hassan, I. S. et al. Asymmetric δ-lactam synthesis with a monomeric streptavidin artificial metalloenzyme. *J. Am. Chem. Soc.* **141**, 4815–4819 (2019).
102. Dürrenberger, M. et al. Artificial transfer hydrogenases for the enantioselective reduction of cyclic imines. *Angew. Chem. Int. Ed.* **50**, 3026–3029 (2011).
103. Wu, S. et al. Breaking symmetry: engineering single-chain dimeric streptavidin as host for artificial metalloenzymes. *J. Am. Chem. Soc.* 15869–15878 (2019).
104. Raines, D. J. et al. Redox-switchable siderophore anchor enables reversible artificial metalloenzyme assembly. *Nat. Catal.* **1**, 680–688 (2018).
105. Eda, S. et al. Biocompatibility and therapeutic potential of glycosylated albumin artificial metalloenzymes. *Nat. Catal.* **2**, 780–792 (2019).
106. Mayer, C., Gillingham, D. G., Ward, T. R. & Hilvert, D. An artificial metalloenzyme for olefin metathesis. *Chem. Commun.* **47**, 12068–12070 (2011).
107. Jarvis, A. G. et al. Enzyme activity by design: an artificial rhodium hydroformylase for linear aldehydes. *Angew. Chem. Int. Ed.* **56**, 13596–13600 (2017).
108. Yang, H., Srivastava, P., Zhang, C. & Lewis, J. C. A general method for artificial metalloenzyme formation through strain-promoted azide–alkyne cycloaddition. *ChemBioChem* **15**, 223–227 (2014).
109. Srivastava, P., Yang, H., Ellis-Guardiola, K. & Lewis, J. C. Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nat. Commun.* **6**, 7789 (2015).
110. Yang, H. et al. Evolving artificial metalloenzymes via random mutagenesis. *Nat. Chem.* **10**, 318–324 (2018).
111. Abe, S. et al. Control of the coordination structure of organometallic palladium complexes in an apo-ferritin cage. *J. Am. Chem. Soc.* **130**, 10512–10514 (2008).
112. Jing, Q. & Kazlauskas, R. J. Regioselective hydroformylation of styrene using rhodium-substituted carbonic anhydrase. *ChemCatChem* **2**, 953–957 (2010).
113. Drienovská, I., Rioz-Martínez, A., Draksharapu, A. & Roelfes, G. Novel artificial metalloenzymes by *in vivo* incorporation of metal-binding unnatural amino acids. *Chem. Sci.* **6**, 770–776 (2015).
114. Drienovská, I. et al. Design of an enantioselective artificial metallohydratase enzyme containing an unnatural metal-binding amino acid. *Chem. Sci.* **8**, 7228–7235 (2017).
115. Key, H. M., Dydio, P., Clark, D. S. & Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **534**, 534–537 (2016).
116. Dydio, P. et al. An artificial metalloenzyme with the kinetics of native enzymes. *Science* **354**, 102–106 (2016).
117. Key, H. M. et al. Beyond iron: iridium-containing P450 enzymes for selective cyclopropanations of structurally diverse alkenes. *ACS Cent. Sci.* **3**, 302–308 (2017).
118. Dydio, P. et al. Chemoselective, enzymatic C–H bond amination catalyzed by a cytochrome P450 containing an Ir(Me)–PIX cofactor. *J. Am. Chem. Soc.* **139**, 1750–1753 (2017).
119. Agostini, F. et al. Biocatalysis with unnatural amino acids: enzymology meets xenobiology. *Angew. Chem. Int. Ed.* **56**, 9680–9703 (2017).
120. Drienovská, I., Mayer, C., Dulson, C. & Roelfes, G. A designer enzyme for hydrazone and oxime formation featuring an unnatural catalytic aniline residue. *Nat. Chem.* **10**, 946–952 (2018).
121. Mayer, C. et al. Directed evolution of a designer enzyme featuring an unnatural catalytic amino acid. *Angew. Chem. Int. Ed.* **58**, 2083–2087 (2019).
122. Burke, A. J. et al. Design and evolution of an enzyme with a non-canonical organocatalytic mechanism. *Nature* **570**, 219–223 (2019).
This artificial enzyme with a novel mechanism from organocatalysis will inspire further development of new enzymes with catalytically functional non-canonical amino acids.
123. Bornscheuer, U. T. et al. Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).
124. Hughes, G. & Lewis, J. C. Introduction: biocatalysis in industry. *Chem. Rev.* **118**, 1–3 (2018).
125. Bornscheuer, U. T. The fourth wave of biocatalysis is approaching. *Philos. Trans. R. Soc. A* **376**, 20170063 (2017).
126. Bornscheuer, U. T. Biocatalysis: successfully crossing boundaries. *Angew. Chem. Int. Ed.* **55**, 4372–4373 (2016).
127. Wu., Z. et al. Machine-learning-assisted directed protein evolution with combinatorial libraries. *Proc. Natl. Acad. Sci. USA* **116**, 8852–8858 (2019).
Machine learning-guided directed evolution of a haem protein for a non-native function facilitated exploration of sequence space.
128. O'Hagan, D. & Deng, H. Enzymatic fluorination and biotechnological developments of the fluorinase. *Chem. Rev.* **115**, 634–649 (2015).
129. Jeon, B., Wang, S.-A., Ruszczycky, M. W. & Liu, H. Natural [4+2]-cyclases. *Chem. Rev.* **117**, 5367–5388 (2017).
130. Kelly, S. A. et al. Application of ω-transaminases in the pharmaceutical industry. *Chem. Rev.* **118**, 349–367 (2018).

Acknowledgements

We thank D. J. Wackelin and Y. Yang (Caltech) for helpful discussions and comments on the manuscript. This work was supported by NSF Division of Molecular and Cellular Biosciences grant no. MCB-1513007, the US Army Research Office Institute for Collaborative Biotechnologies (cooperative agreement no. W911NF-19-2-0026) and the US Army Research Office Institute for Collaborative Biotechnologies (contract no. W911NF-19-D-0001). K.C. thanks the Resnick Sustainability Institute at Caltech for fellowship support. The content of this paper does not necessarily reflect the position or the policy of the funding agencies, and no official endorsement should be inferred.

Author contributions

All authors participated in designing and writing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to F.H.A.