



Cite this: *Green Chem.*, 2023, **25**, 2958

Received 8th February 2023,
Accepted 27th March 2023

DOI: 10.1039/d3ac00456b

rsc.li/greenchem

rsc.li/greenchem

Biocatalytic amide bond formation

Max Lubberink, † William Finnigan and Sabine L. Flitsch *

Amide bonds occur in over half of target compounds in medicinal chemistry patents, making amide bond formation the most commonly performed reaction in the pharmaceutical industry. The growing demand for green and sustainable catalytic methods for amide bond synthesis has driven the interest in biocatalytic methods. We have compiled the literature on this topic in a publicly available database (RetroBioCat) and use this database to discuss the most commonly used strategies. These include the use of low water systems for driving equilibria and kinetics in hydrolase-catalysed reactions towards amide bond formation. Alternatively, in aqueous medium, ATP-dependent enzymes have been shown to be particularly promising and can be coupled to ATP recycling systems.

Introduction

The amide bond plays a fundamental role in Nature and its prevalence in medicinal chemistry is exemplified by the fact that it occurs in over half of all compounds listed in medicinal chemistry patents and has been the most commonly performed reaction in the pharmaceutical industry for decades.^{1,2}

Accordingly, there is a wide choice for chemical methods of amide bond formations, which will not be reviewed here.³ However, many solvents and reagents commonly used for amide synthesis lack green credentials, and the development of sustainable methodologies is still of great interest.^{4–7} In 2018, a roundtable of pharmaceutical industry experts (GCIPR) identified amide bond formations as a top priority for the development of alternative more sustainable methods and biocatalysis was named as one of the most promising research fields to fulfil this need.⁸

Biocatalysis encompasses the use of enzymes for organic synthesis and has been used as a replacement for a wide variety of chemical processes for decades. Over recent years, enzymes have been increasingly utilized for the synthesis of more complex molecules such as pharmaceutical drugs, having advantages as catalysts with high chemo- and stereoselectivity.⁹ Limitations in reactivity and selectivity can be overcome by using large metagenomic databases as sources for new enzymes and by protein engineering.¹⁰ In particular, directed evolution, for which the 2018 Nobel Prize was awarded to Frances Arnold has transformed the efficiency and speed of development of industrial enzymes by orders of magnitude.¹¹ Other great benefits of using biocatalysis are its econ-

omic and environmental advantages over traditional chemical synthesis. Enzymes are biodegradable, non-toxic and produced from cheap renewable resources. Enzymatic reactions are often performed under aqueous conditions using mild reaction temperatures and frequently circumvent protection/deprotection strategies or functional group activation. Therefore, biocatalytic methods generally lead to less waste, are more cost-effective than conventional organic synthesis, and are therefore seen as potentially more sustainable.¹²

In Nature, the amide bond is ubiquitous and therefore, biosynthesis provides a wide variety of amide bond-forming strategies which can be harnessed for use in biocatalysis. This article focuses on cell-free applications of amide bond-forming enzymes for the synthesis of industrially relevant compounds.

To accompany this review, we have also curated data from across the literature for enzymatic amide bond formation into a database linked to the biocatalytic retrosynthesis tool RetroBioCat,¹³ allowing this information to be interactively explored and interrogated (available through retrobiocat.com). The database is continuously updated and is a community-based effort. As of February 2023, the database comprises 388 citations, 1831 enzymes, 4659 product molecules and 19 946 reactions, the detailed discussion of which is beyond this review. We have used the data collected in this database to analyse and discuss certain aspects of the amidation topic, such as the use of different organic solvents in lipase-catalysed amidation reactions and the range of amide bond-forming enzymes available for organic synthesis (Fig. 1).

Aminolysis of esters

Lipases

Lipases are the most studied enzymes for amide bond formation (Fig. 1). They are members of the family of serine

Manchester Institute of Biotechnology, Department of Chemistry, The University of Manchester, 131 Princess Street, Manchester M1 7DN, UK.

E-mail: Sabine.Flitsch@manchester.ac.uk

† Present address: Wageningen Food & Biobased Research, Wageningen, P. O. Box 17, 6700 AA, The Netherlands.



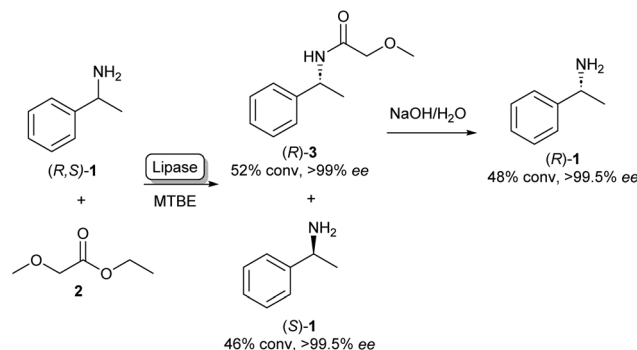
#	Enzyme Classes	Data in RetroBioCat
1	Lipases	886 unique product molecules
2	Penicillin G Acylases	63 unique product molecules
3	N-Acyltransferases	84 unique product molecules
4	ATP-dependent amide bond-forming enzymes	228 unique product molecules
5	Nitrile Hydratases	160 unique product molecules

Fig. 1 The four main enzymatic amidation strategies discussed in this review (top) and corresponding enzyme classes with their data availability in the RetroBioCat database (bottom).

hydrolases that naturally catalyse the hydrolysis of triglyceride ester bonds at the oil–water interface and are therefore naturally stable in hydrophobic conditions.¹⁴ Zaks *et al.* reported that lipases show increased thermostability and selectivity in anhydrous conditions and that transacylation reactions are favourable over hydrolysis under these conditions, thus promoting the aminolysis of esters leading to amides.^{15,16} This work is often seen as a landmark discovery, showing enzymes are more widely applicable in organic solvents than previously thought possible.¹² The enzymes used in these early studies were isolated from porcine pancreas (PPL)¹⁷ or *Candida cylindracea* (CCL),¹⁸ but over time the microbial *Candida antarctica* lipase B (CAL-B), in its free form or its many commercially available immobilized forms, has grown to be the most commonly used lipase for amidation reactions.^{19,20} Indeed, among the reactions listed in the RetroBioCat¹³ database of biocatalytic reactions, CAL-B is the most prevalent enzyme used for amidation reactions. CAL-B is a member of the serine hydrolase family with its three characteristic active site amino acids (aspartic acid, histidine and serine) forming the catalytic triad.¹⁹

Kinetic resolutions of amines

The enantiospecific properties of lipases have been extensively exploited for the synthesis of chiral amines through kinetic resolution. Lipase-catalysed kinetic resolution processes have been used on an industrial scale, for example by BASF as shown in Scheme 1: here, a lipase is used in organic solvent to



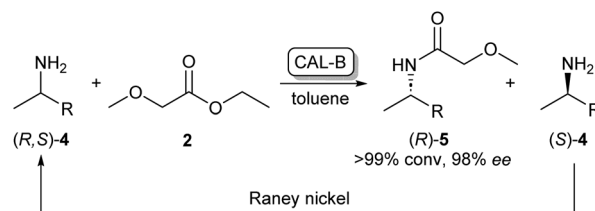
Scheme 1 Lipase catalysed kinetic resolution of amines: selective acylation of (*R,S*)-1 using lipase in MTBE (methyl *tert*-butyl ether) yielding (*S*)-1. Basic hydrolysis of amide 3 yields (*R*)-1.²¹

selectively acylate the *R*-enantiomer of racemic (*R,S*)-1 leaving enantiopure (*S*)-1 unreacted. After the separation of the reaction products, amide 3 can be hydrolysed under basic conditions yielding (*R*)-1 in excellent enantiopurity.^{21,22} As lipases are known for their exceptionally broad substrate scope, the variety of amines that can be resolved using this method is very diverse and can in some cases be done on multi-ton scales.²² A large amount of academic research has gone into the enantioselective acylation of amines following this principle, exploring the substrate scope and reaction conditions.

As the kinetic resolution outlined in Scheme 1 inherently yields maximum conversion rates of 50% per enantiomer, there has been a significant interest in dynamic kinetic resolution (DKR) reactions. This method uses a racemisation agent, usually in the form of a metal or an organometallic catalyst, to convert the unreacted amine enantiomer back to the racemate, generating a cycle that allows these kinetic resolution reactions to reach full conversions to enantiopure amides (Scheme 2).^{23,24}

Use of non-aqueous solvents

The versatility of lipases for amide bond formation and other reactions has been reviewed extensively over the years.^{20,25} However, the need for organic solvents in lipase catalysed amidations raises the question if this is the most sustainable and safe synthesis method. Solvent selection guides such as published by CHEM21 have made efforts to rank commonly used



Scheme 2 CAL-B/RANEY® nickel catalysed dynamic kinetic resolution of 4 yielding enantiopure (*R*)-5. CAL-B selectively acylates (*R*)-4 while the unreacted (*S*)-4 is converted back to the racemate by the RANEY® nickel catalyst.²³



Overall, Fig. 2 illustrates that most solvents used in optimized lipase amide-forming reactions are not suitable as green solvents within the pharmaceutical industry and academia and pose substantial occupational health hazards. In this respect, moving towards amidation reactions under aqueous conditions is highly desirable, as water is ranked as the most recommended solvent in this selection guide. In some cases, lipases are able to catalyse amide bond formation in aqueous environments. In addition, a number of other enzyme classes have been shown to catalyse this reaction in water using a variety of mechanisms. These aqueous amidation strategies will be covered in the remaining part of this review.

Penicillin G acylases

A further group of hydrolase enzymes that have been exploited for amide bond formation are the penicillin G acylases (PGAs). PGAs natively catalyse the reversible hydrolysis of the amide bond in penicillin antibiotics (Scheme 3A). This characteristic

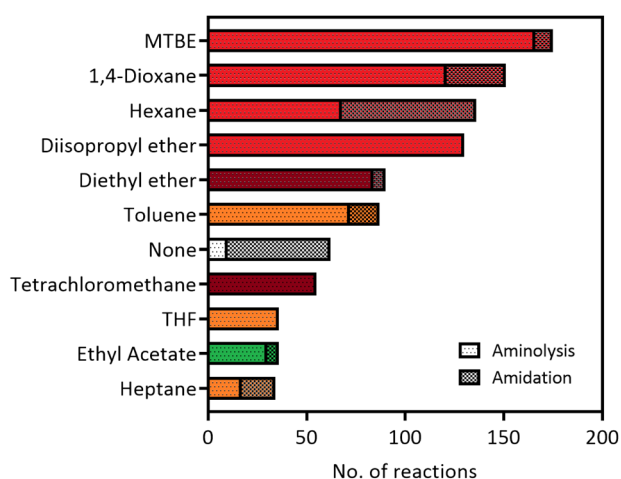
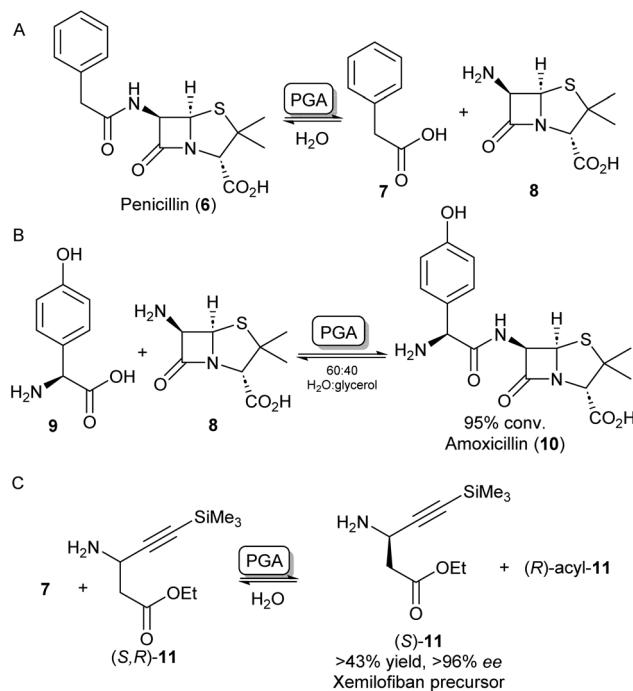


Fig. 2 The most used solvents for lipase catalysed amidation reactions in the RetroBioCat database. The individual solvents are split up between reaction types (aminolysis of esters and direct amidation of carboxylic acids) and are colour coded by their ranking using the ChEM21 solvent selection guide (dark red: highly hazardous, red: hazardous, yellow: problematic, green: recommended).⁵

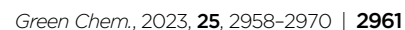


Scheme 3 Applications of penicillin G acylase (PGA). (A) PGA catalyses the hydrolysis of penicillin into **7** and **8**. (B) PGA can reinstall alternative acyl side chains such as **9** to synthesize analogous antibiotics like amoxicillin **10**.²⁶ (C) Racemic β -aminoester **11** can be deracemized on an industrial scale to obtain drug precursor (S)-**11** in high ee's.³⁴

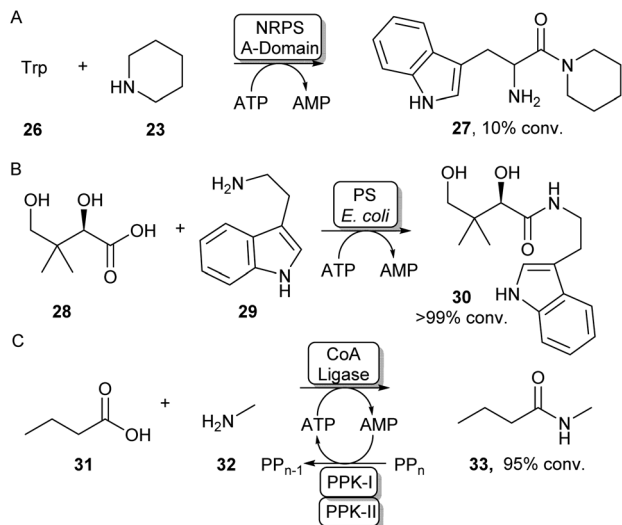
has been exploited for the synthesis of semi-synthetic penicillins and cephalosporins by hydrolysing the amide bond of penicillin-like structures and subsequent reinstallation of an alternative acyl side chain to the beta-lactam core structure (Scheme 3B).^{26–28} PGAs have been used on industrial scales for processes like these for decades.²⁹ Exploiting the substrate promiscuity, this enzyme class has also been utilized in the synthesis of dipeptides^{30,31} and the kinetic resolution of amines.^{32,33} For example, immobilized PGA was utilized for the resolution of the β -aminoester **11** using **7** as an acyl donor, resulting in an enantiopure precursor for the platelet aggregation inhibitor xemilofiban (*S*)-**11** on a 70 L scale (Scheme 3C).³⁴ Recently, PGA has been utilized by researchers at Merck for larger biomolecules. They report the selective installation of cleavable phenyl acetamide groups at either of three available primary amine positions of human insulin. Extensive directed evolution campaigns led to a toolbox of tailored PGA variants for the acylation or hydrolysis (*i.e.* protection/deprotection) for each of the targetable amino moieties. The selective modification of insulin led to homogeneous samples of bioconjugates that are highly valuable as biological probes or therapeutics.³⁵

N-Acyltransferases

In some cases, the formation of esters as more reactive intermediates for biocatalytic amide bond formation is used in natural metabolism. For example, in the biosynthesis of



This journal is © The Royal Society of Chemistry 2023



Scheme 7 Utilising adenylating enzymes for amide bond formation. (A) The truncated A-domain from the NRPS TycA was able to adenylate tryptophan to a range of amines.⁵⁸ (B) PS from *E. coli* has a promiscuous amine scope allowing the synthesis of a range of vitamin B5 antimetabolites.⁵⁹ (C) Adenylation activity from a thermostable CoA ligase allowed the preparative synthesis of **33** using a two-enzyme ATP recycling system.⁶⁰

primary and secondary amines (Scheme 7A).⁵⁸ Although these A-domains are usually very specific for their carboxylic acid substrate, Zhu *et al.* have recently shown the 3-isocyanobutanoic acid adenylating enzyme SfaB to be able to couple a relatively broad range of aliphatic acids to a variety of primary and secondary amines.⁶¹ Adenylating enzymes involved in other natural product synthesis pathways have also been used for synthetic amide bond formation. For example, the standalone adenylating enzyme TamA, involved in the biosynthesis of tambjamine YP1, was used to couple a range of fatty acids to some α -amino acids and primary amines.⁶² Furthermore, pantothenate synthase (PS), which is involved in the biosynthesis of coenzyme A (CoA) was found to accept a range of amine substrates other than its native β -alanine substrate and has been exploited for the preparative synthesis of a range of pharmaceutically relevant amides like **30** (Scheme 7B).⁵⁹ In a similar fashion, the adenylating activity of CoA-ligases can also be used for amide bond formation. In their native function, these enzymes adenylate their carboxylic acid substrates for the subsequent coupling with CoA forming a CoA-thioester. For example, a thermostable CoA ligase isolated from *Geobacillus thermodenitrificans* has been used for the preparative synthesis of **33** using a two-enzyme ATP recycling system (Scheme 7C).⁶⁰

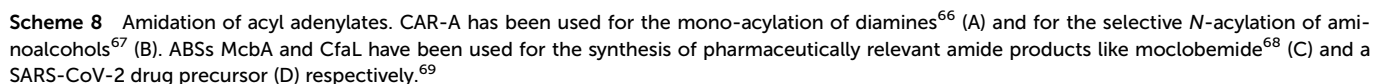
A different protein engineering strategy starts with Carboxylic acid reductases (CARs), which are multidomain proteins whose architecture resembles a single NRPS module: the A-domain activates a carboxylic acid substrate using ATP, which subsequently becomes covalently attached to the T-domain in the form of a thioester. In the final step of the

CAR-catalysed reaction, this thioester intermediate is then transported to the reductase (R) domain where it is reduced to the aldehyde using the cofactor NADPH.⁶³ CARs are known to have a very broad carboxylic acid scope due to their relatively large and hydrophobic active site pocket.⁶⁴ In order to use this system for amide bond formations, it was shown that by addition of an amine nucleophile, the acyl adenylate could be intercepted forming a variety of primary, secondary or tertiary amides.⁶⁵ Given that solely the adenylation activity of CAR was shown to be required for amidation, the CAR gene was truncated so that only the CAR A-domain would be expressed. It was found that the truncated CAR enzyme from *Mycobacterium marinum* (CAR_{mm}-A) had improved amidation activity compared to the full-length enzyme. After further optimisation by introducing a polyphosphate kinase from *Cytophaga hutchinsonii* (CHU) to regenerate ATP from AMP using cheap polyphosphate, this amidation method was used for the mono-acylation of symmetrical diamines. As solely the mono-acylated product was obtained using CAR-A, this method circumvented the protection/deprotection strategies used in traditional chemical methods towards these motifs. The pharmaceutical relevance was demonstrated by synthesising a range of drugs and drug precursors, for example, the antiparkinsonian agent lazabemide **36** was synthesized in a single biocatalytic step (Scheme 8A).⁶⁶

In further studies, the promiscuity of the CAR-A system was exploited for the amide coupling of amino alcohols to fatty acids showing a very high selectivity of amide bond formation over ester formation. This allowed the one-step synthesis of commercially relevant glucamide surfactants like **39** with advantages over lipase-catalysed syntheses where mixtures of amide and ester products are obtained (Scheme 8B).⁶⁷ It has been hypothesised that the acyl adenylate intermediate in this system might undergo nucleophilic attack independent from the enzyme, so, the *N* over *O*-selectivity observed in this reaction is likely due to the intrinsic reactivity of acyl phosphates in water.⁶⁷ This feature might also explain the exceptionally large amine scope of the CAR-A system.

Indeed, it is not uncommon in Nature to use the adenylation strategy for direct amidation with an amine nucleophile. ATP-dependent amide bond synthases (ABSs) have been identified to catalyse amide bond formation in secondary metabolite pathways by accepting an amine directly into the active site of the enzyme, not requiring the involvement of a thioester intermediate. For example, the ABS isolated from *Marinactinospora thermotolerans* (McbA) has been shown to couple its native β -carboline-derived acid substrate with a range of primary amines.⁷⁰ Structural studies of McbA revealed that the enzyme also displayed promiscuity for its carboxylic acid substrate and that the enzyme catalyses both half-reactions of the adenylation-amidation sequence.⁷¹ The enzyme was subsequently exploited for the synthesis of the antidepressant drug **42** in a preparative scale using a two-enzyme ATP recycling system (Scheme 8C).⁶⁸ Recently, another ABS, named coronafacic acid ligase (CfaL), was discovered to possess great selectivity towards aliphatic and aromatic acids. The enzyme was shown





is suitable for the synthesis of a wide variety of amide products, as long as the right combination of two complementary enzymes is being used. A precursor of the kinase inhibitor **49** was obtained in high yields on a preparative scale using a two-enzyme combination in a whole cell system (Scheme 9A).⁷³

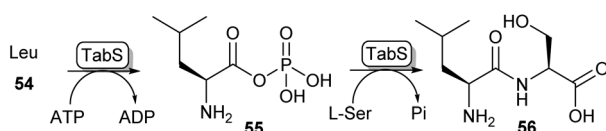
Recently, it was found that besides using an amine nucleophile to intercept the adenylate intermediate in the CAR-A reactions discussed previously, CAR-A was able to couple a wide variety of carboxylic acids to thiols, in particular CoA-SH, generating a range of CoA-thioesters and thus serving as a universal CoA ligase.⁷⁴ This general acyl-S-CoA recycling system can be used *in situ* with substoichiometric amounts of thiol and is an attractive alternative to chemical ligation strategies.^{45,74} Coupling the CAR-A mediated CoA-ester synthesis to a complementary NAT enzyme, the system was exploited to synthesize an exceptionally wide range of amide products, including a set of amides derived from a variety of anilines like **52**, which are usually regarded as more challenging substrates because of their low nucleophilicity and are not accessible through the direct CAR-A amidation method (Scheme 9B). Selected examples were performed on a preparative scale, demonstrating the value of this approach for synthesis. Further exploiting CAR-A as a universal CoA-thioester synthetase, the system was coupled to a lysine acetyltransferase to achieve acylation of histone-derived peptide H4-20 including non-natural bio-orthogonal groups amenable to further ‘click chemistry’.⁷⁴

Amidation of acyl phosphates

A second strategy toward amide bond formation prevalent in biosynthetic pathways involves adenylation of carboxylic acid substrates and subsequent conversion to a CoA-thioester by a CoA ligase. The CoA-thioester is then coupled to an amine by a further *N*-acyltransferase (NAT). Because this two-enzyme system for amide bond formation uses two cofactors (ATP and CoA), some research has gone into *in vitro* regeneration systems to make this strategy a viable method for synthetic amide bond formation. Mordhorst *et al.* designed a cascade consisting of a CoA ligase and a NAT coupled to a two-enzyme ATP recycling system. Using different combinations of ligases and NATs moderate to good conversions were observed to acetyl *para*-aminobenzoic acid and hippuric acid products.⁷²

Even though some NATs display moderately broad substrate scopes, CoA ligases are known to be very selective for their native carboxylic acid substrate.⁵⁴ This necessitates the selection of a two-enzyme combination with complementary substrate profiles to obtain the desired amide product. Using high-throughput screening of a library of CoA-ligases with a range of carboxylic acid substrates, and a second screening of a library of NATs with a combination of CoA-esters and amines, scientists at GSK have demonstrated that this system

Scheme 9 Amidations involving SCoA-esters. (A) CoA Ligase enzymes can adenylate and subsequently couple CoA to their carboxylic acid substrate. By selecting complementary CoA Ligases and NATs based on their substrate specificities, amides can be formed via an acyl-CoA intermediate.⁷³ (B) CAR-A has been shown to accept CoA as a nucleophile and, because of its wide substrate scope, is able to serve as a universal CoA ligase. Therefore, CAR-A can be utilized in a comparable strategy when coupled to a NAT, forming a variety of amides.⁷⁴



Scheme 10 ATP grasp enzymes form amide bonds via an acyl phosphate intermediate (55). The ATP grasp enzyme isolated from *Pseudomonas syringae* (TabS) has been shown to synthesise a variety of dipeptides such as 56.⁷⁶

they share between them, responsible for “grasping” or holding a molecule of ATP in the active site.¹⁹ The acid substrate **54** is reacted with ATP forming an acyl-phosphate intermediate (**55**) which subsequently undergoes nucleophilic attack by an amine leading to amide **56**. A great variety of ATP-grasp enzyme-catalysed amide bond-forming reactions have been described, involved in either primary or secondary metabolism.

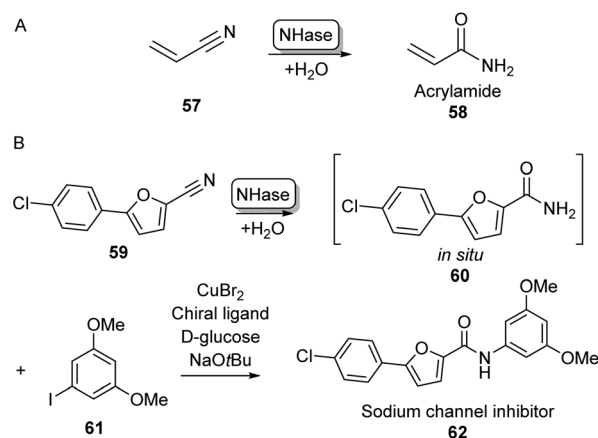
Carboxylic acids accepted by the ATP-grasp family range from small acids like formic acid to large proteins being accepted as substrates, while for the amine component, a range from ammonium ions to the amine of a biotin prosthetic group of a carrier protein have been described as substrates. Even though the property of this enzyme class has been described as highly promising for use in synthetic chemistry,¹⁹ they have not been implemented as such, most likely due to the high specificity for both the acid and the amine substrates. To date, ATP-grasp enzymes have mainly been used for the synthesis of di- or oligopeptides where some promiscuity has been observed.⁷⁵

For example, ywE from *Bacillus subtilis* was shown to accept a handful of proteinogenic amino acids as the carboxylic acid donor, and accept up to 15 amino acids as the amine donor.⁷⁷ The L-amino acid ligase TabS from *Pseudomonas syringae* has shown the broadest substrate promiscuity thus far, being able to accept all 20 proteinogenic L-amino acids and β -alanine as either an acid or an amine substrate (Scheme 10).⁷⁶

Nitrile hydration

Nitrile hydratases

Although amine acylation is the most common synthetic route toward amides, primary amides can also be obtained by the hydration of nitrile compounds. Chemical nitrile hydration methods rely on harsh reaction conditions such as high temperatures (200–300 °C) and high pressure usually obtaining low yields without chemo- or regioselectivity. Therefore, the cobalt- or iron-dependent enzyme class of nitrile hydratase (NHase) has been used in several industrial-scale processes for decades as a more eco-friendly alternative.⁷⁸ In fact, NHase catalysed industrial-scale synthesis of acrylamide (**58**) from acrylonitrile **57** is the first example of the use of biocatalysis for the production of commodity chemicals and the first example of the use of biotechnology in the petrochemical industry (Scheme 11A).⁷⁹ NHase has also been employed as an industrial catalyst for the manufacturing of nicotinamide and 5-cyanovaleramide.⁸⁰ Much research has gone into the expression,⁸¹ structure,⁸² mechanism,⁸³ stability,⁸⁴ selectivity⁸⁵ and sub-



Scheme 11 Nitrile Hydratase (NH) catalysed amide bond formation. (A) NHase catalysed acrylamide synthesis. (B) Chemoenzymatic strategy for the expansion of amide scopes using copper catalysis.⁸⁷

strate scope⁸⁶ of NHases and has been extensively reviewed elsewhere.^{78–80}

In a notable recent example, NHases have been used in combination with copper catalysis to extend the product scope of this amidation method beyond primary amides. Using two complementary NHase subtypes (from *Rhodospseudomonas palustris* and *Rhodococcus erythropolis* respectively) in whole *E. coli* cells to perform nitrile hydration, the obtained primary amide was then coupled *in situ* to iodobenzene derivatives like **61** using CuBr₂ and a diamine ligand in water, obtaining a broad range of aliphatic and aromatic amide products such as **62** on gram scales using substrate loadings in the molar range (Scheme 11B).⁸⁷

Lactamisation

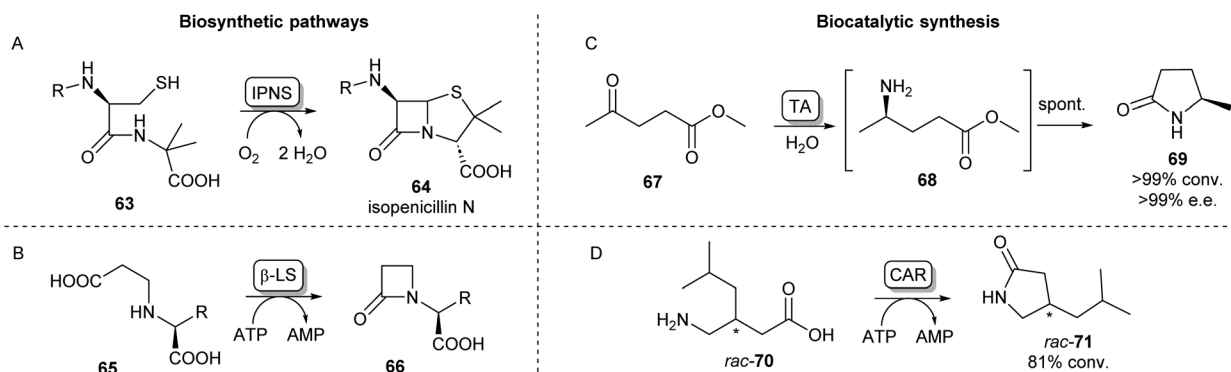
Cyclic amides (lactams) are found as motifs in many natural products and APIs and are important precursors in the synthesis of commercial polymers.⁸⁸ A well-known example from nature is the four-membered β -lactam ring present in the most important class of currently approved antibiotics, first observed in penicillin.⁸⁹ The biosynthetic pathway of the well-studied penicillin N (**64**) proceeds *via* a tripeptide precursor produced by an NRPS. This peptide is cyclized by isopenicillin N synthase (IPNS), a non-heme iron protein making use of a ferrous ion and molecular oxygen, yielding the β -lactam ring in **64** (Scheme 12A).^{90,91} An alternative pathway toward β -lactams has been found in the biosynthesis of clavulanic acid, an inhibitor of β -lactamases which is used to overcome antibiotic resistance. Here, an ATP-dependent β -lactam synthetase (β -LS) catalyses the intramolecular ring closure between a terminal carboxylic acid and a secondary amine (Scheme 12B).⁹⁰ β -LS has been engineered using rational design to accept methyl and ethyl substituted substrates yielding clavulanic derivatives.⁹² Carbanepam synthase (CarA) also utilizes an ATP-dependent mechanism to produce a lactam

ring but is besides four-membered rings also able to form 5- and 6-membered lactam rings.⁹³ This enzyme has subsequently been used in a two-enzyme cascade to prepare a range of bicyclic β -lactams in excellent diastereomeric ratios.⁹⁴ (BAS) was exploited to form 5-membered lactam rings *via* an intramolecular Dieckmann condensation. Coupling malonyl-CoA with aminoacyl-CoA esters a range of unnatural tetramic acid derivatives were obtained with one example showing anti-leukemic activity.⁹⁵

In a different lactamisation strategy, the wide substrate promiscuity of the polyketide synthase benzalacetone synthase.

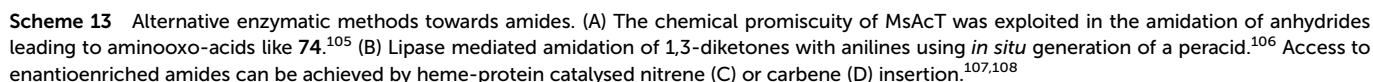
Synthetically, lactams can be formed using an intramolecular cyclisation reaction of amino esters or amino acids catalysed by lipases.⁴⁶ PPL has been shown to catalyse the lactamisation of a range of amino esters in *tert*-amyl alcohol,⁹⁶ whereas CAL-B has been used for the cyclisation of 4-, 5-, 6-, and 8-aminooctanoic acid to their corresponding lactam products.⁹⁷ In a scaled-up example, Willis *et al.* have utilized CAL-B for the lactamisation of a fluorinated diester in aqueous conditions on a 10 g scale.⁹⁸

An alternative biocatalytic method towards lactam formation is *via* the reductive amination of keto- or aldehyde esters, the products of which have been shown to spontaneously cyclise to the corresponding lactams. This was first shown using the (*R*)-selective transaminase ATA-117 synthesising (*R*)-4-phenylpyrrolidin-2-one in excellent yields.⁹⁹ Using similar lactamisation strategies and thorough reaction optimisations, these reactions have been applied on multi-gram scales showing excellent stereoselectivities.^{100,101} The versatility of this lactamisation strategy has been shown by Mourelle-Insua *et al.* forming a wide range of lactams from γ - and δ -keto esters by using a set of commercially available transaminases (Scheme 12C).⁸⁸ Using a different reductive amination strategy, an NADPH-dependent reductive aminase has been shown to catalyse the coupling of a keto ester with propargylamine, which then formed the (*R*)-lactam products in high yields and



Scheme 12 Lactam forming enzymes. Naturally occurring biosynthetic enzymes forming a beta-lactam ring (left). (A) The non-heme iron protein IPNS catalyzes the cyclisation of **63** forming isopenicillin N (**64**) and (B) β -LS catalyzes the ring closure of **65** forming **66**, an intermediate in the clavulanic acid biosynthetic pathway.⁹⁰ Enzymatic methods towards synthetic lactams (right). (C) Enantioselective TAs catalyse reductive amination of ketoesters like **67** forming intermediate **68** which undergoes spontaneous cyclisation forming lactams like **69**.⁸⁸ (D) Using CAR catalysed amidation a variety of amino acids have been converted into lactams.¹⁰³





Using both wild-type and mutants of CAR from *Segniliparus rugosus* various substituted lactams of different ring sizes were synthesised with some examples showing stereoselectivity (Scheme 12D).¹⁰³ This lactamisation activity of CAR has also been applied in a one-pot biocatalytic cascade forming ϵ -caprolactam from cyclohexanol.¹⁰⁴

to catalyse the cyclopropanation of a variety of acrylamide and acrylate olefins¹¹⁰ allowing the synthesis of a levomilnacipran precursor¹¹¹ or amide-bearing alkenes in high diastereo- and enantioselectivities.¹¹² A similar intramolecular cyclopropanation strategy was also shown to be catalysed by myoglobin mutants leading to enantioenriched lactams (Scheme 13D).¹⁰⁸

Conclusion

As traditional chemical methods for amide bond formation most often rely on hazardous reagents and coupling agents, scientists have been looking to develop more sustainable and greener alternatives for this very common chemical reaction for decades. Biocatalysis has played a big role in this regard. In fact, the discovery of the ability of lipases to catalyse amide bond formation has helped to kick-start the very field of biocatalysis, possessing excellent stability and substrate scopes. However, with the increasing impact of climate change, the desirability for moving away from fossil fuels and to more sustainable chemistry is becoming more and more important. As outlined in Fig. 2, lipase catalysed amidations often rely on hazardous solvents. Aqueous amidation methods are therefore highly sought after, and some highly promising enzymatic technologies have been described in this review: NATs like MsAcT and its mutants have proven very successful for the acetylation of a range of amines and even thiols in water, whereas ATP-dependent acyl adenylate forming enzymes like McbA, CfaL and CAR-A have shown to be able to couple a very diverse set of carboxylic acids and amines under aqueous conditions, forming a wide range of molecules relevant to pharmaceutical and fine chemical industries. RetroBioCat seeks to cover all papers in this area and researchers are encouraged to help populate the database as new methodologies are published. Publicly available data curation tools allow new reactions, enzymes, or papers to be suggested by anyone. Furthermore, a comprehensive data curation portal is available for the addition of enzyme sequences and activity data, with database additions to the database attributed to each user.

Furthermore, using a carbene transfer reaction with ethyl diazoacetate, this same class of enzymes was shown to be able

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

ML thanks CoEBio3 for funding for a studentship.

References

- 1 D. G. Brown and J. Boström, *J. Med. Chem.*, 2016, **59**, 4443–4458.
- 2 E. Massolo, M. Pirola and M. Benaglia, *Eur. J. Org. Chem.*, 2020, **2020**, 4641–4651.
- 3 R. M. de Figueiredo, J.-S. Suppo and J.-M. Campagne, *Chem. Rev.*, 2016, **116**, 12029–12122.
- 4 J. R. Dunetz, J. Magano and G. A. Weisenburger, *Org. Process Res. Dev.*, 2016, **20**, 140–177.
- 5 D. Prat, A. Wells, J. Hayler, H. Sneddon, C. R. McElroy, S. Abou-Shehadeh and P. J. Dunn, *Green Chem.*, 2016, **18**, 288–296.
- 6 S. Hazra, F. Gallou and S. Handa, *ACS Sustainable Chem. Eng.*, 2022, **10**, 5299–5306.
- 7 D. Procopio, C. Siciliano, S. Trombino, D. E. Dumitrescu, F. Suciú and M. L. di Gioia, *Org. Biomol. Chem.*, 2022, **20**, 1137–1149.
- 8 M. C. Bryan, P. J. Dunn, D. Entwistle, F. Gallou, S. G. Koenig, J. D. Hayler, M. R. Hickey, S. Hughes, M. E. Kopach, G. Moine, P. Richardson, F. Roschangar, A. Steven and F. J. Weiberth, *Green Chem.*, 2018, **20**, 5082–5103.
- 9 R. J. Young, S. L. Flitsch, M. Grigalunas, P. D. Leeson, R. J. Quinn, N. J. Turner and H. Waldmann, *JACS Au*, 2022, **2**, 2400–2416.
- 10 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, *Nat. Rev. Methods Primers*, 2021, **1**, 46.
- 11 F. H. Arnold and F. H. Arnold– Nobel Lecture, <https://www.nobelprize.org/prizes/chemistry/2018/arnold/lecture/>, (accessed 23 November 2022).
- 12 R. A. Sheldon and J. M. Woodley, *Chem. Rev.*, 2018, **118**, 801–838.
- 13 W. Finnigan, L. J. Hepworth, S. L. Flitsch and N. J. Turner, *Nat. Catal.*, 2021, **4**, 98–104.
- 14 M. T. Reetz, *Curr. Opin. Chem. Biol.*, 2002, **6**, 145–150.
- 15 A. Zaks and A. M. Klivanov, *Science*, 1984, **224**, 1249–1251.
- 16 A. Zaks and A. M. Klivanov, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 3192–3196.
- 17 A. L. Margolin and A. M. Klivanov, *J. Am. Chem. Soc.*, 1987, **109**, 3802–3804.
- 18 V. Gotor, R. Brieva and F. Rebolledo, *Tetrahedron Lett.*, 1988, **29**, 6973–6974.
- 19 A. Goswami and S. G. Van Lanen, *Mol. BioSyst.*, 2015, **11**, 338–353.
- 20 R. N. Lima, C. S. dos Anjos, E. V. M. Orozco and A. L. M. Porto, *Mol. Catal.*, 2019, **466**, 75–105.
- 21 F. Balkenhohl, K. Dittrich, B. Hauer and W. Ladner, *J. Prakt. Chem.*, 1997, **339**, 381–384.
- 22 M. Breuer, K. Dittrich, T. Habicher, B. Hauer, M. Keßler, R. Stürmer and T. Zelinski, *Angew. Chem., Int. Ed.*, 2004, **43**, 788–824.
- 23 A. N. Parvulescu, P. A. Jacobs and D. E. De Vos, *Adv. Synth. Catal.*, 2008, **350**, 113–121.
- 24 R. N. Lima, C. S. dos Anjos, E. V. M. Orozco and A. L. M. Porto, *Mol. Catal.*, 2019, **466**, 75–105.
- 25 V. Gotor, *Bioorg. Med. Chem.*, 1999, **7**, 2189–2197.
- 26 X. Pan, L. Wang, J. Ye, S. Qin and B. He, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 1749–1758.
- 27 X.-L. Zhang, M.-H. Zong and N. Li, *Bioresour. Bioprocess.*, 2016, **3**, 49.
- 28 C. Fuganti, C. M. Rosell, R. Rigoni, S. Servi, A. Tagliani and M. Terreni, *Biotechnol. Lett.*, 1992, **14**, 543–546.
- 29 B. M. Dorr and D. E. Fuerst, *Curr. Opin. Chem. Biol.*, 2018, **43**, 127–133.
- 30 P. C. Pereira, I. W. C. E. Arends and R. A. Sheldon, *Tetrahedron Lett.*, 2014, **55**, 4991–4993.
- 31 M. Koreishi, K. Tani, Y. Ise, H. Imanaka, K. Imamura and K. Nakanishi, *Biosci., Biotechnol., Biochem.*, 2007, **71**, 1582–1586.
- 32 L. M. Van Langen, N. H. P. Oosthoek, D. T. Guranda, F. Van Rantwijk, V. K. Švedas and R. A. Sheldon, *Tetrahedron: Asymmetry*, 2000, **11**, 4593–4600.
- 33 X. Gong, E. Su, P. Wang and D. Wei, *Tetrahedron Lett.*, 2011, **52**, 5398–5402.
- 34 B. H. Landis, P. B. Mullins, K. E. Mullins and P. T. Wang, *Org. Process Res. Dev.*, 2002, **6**, 539–546.
- 35 A. Fryszkowska, C. An, O. Alvizo, G. Banerjee, K. A. Canada, Y. Cao, D. DeMong, P. N. Devine, D. Duan, D. M. Elgart, I. Farasat, D. R. Gauthier, E. N. Guidry, X. Jia, J. Kong, N. Kruse, K. W. Lexa, A. A. Makarov, B. F. Mann, E. M. Milczek, V. Mitchell, J. Nazor, C. Neri, R. K. Orr, P. Orth, E. M. Phillips, J. N. Riggins, W. A. Schafer, S. M. Silverman, C. A. Strulson, N. Subramanian, R. Voladri, H. Yang, J. Yang, X. Yi, X. Zhang and W. Zhong, *Science*, 2022, **376**, 1321–1327.
- 36 M. Funabashi, Z. Yang, K. Nonaka, M. Hosobuchi, Y. Fujita, T. Shibata, X. Chi and S. G. Van Lanen, *Nat. Chem. Biol.*, 2010, **6**, 581–586.
- 37 X. Liu, Y. Jin, W. Cai, K. D. Green, A. Goswami, S. Garneau-Tsodikova, K. Nonaka, S. Baba, M. Funabashi, Z. Yang and S. G. Van Lanen, *Org. Biomol. Chem.*, 2016, **14**, 3956–3962.
- 38 N. G. Schmidt, T. Pavkov-Keller, N. Richter, B. Wilschi, K. Gruber and W. Kroutil, *Angew. Chem., Int. Ed.*, 2017, **56**, 7615–7619.
- 39 A. Zadło-Dobrowolska, N. G. Schmidt and W. Kroutil, *Chem. Commun.*, 2018, **54**, 3387–3390.



- 40 H. Müller, S. P. Godehard, G. J. Palm, L. Berndt, C. P. S. Badenhorst, A. Becker, M. Lammers and U. T. Bornscheuer, *Angew. Chem., Int. Ed.*, 2021, **60**, 2013–2017.
- 41 I. Mathews, M. Soltis, M. Saldajeno, G. Ganshaw, R. Sala, W. Weyler, M. A. Cervin, G. Whited and R. Bott, *Biochemistry*, 2007, **46**, 8969–8979.
- 42 H. Land, P. Hendil-Forsell, M. Martinelle and P. Berglund, *Catal. Sci. Technol.*, 2016, **6**, 2897–2900.
- 43 M. L. Contente, A. Pinto, F. Molinari and F. Paradisi, *Adv. Synth. Catal.*, 2018, **360**, 4814–4819.
- 44 M. L. Contente, S. Farris, L. Tamborini, F. Molinari and F. Paradisi, *Green Chem.*, 2019, **21**, 3263–3266.
- 45 M. L. Contente, D. Roura-Padrosa, F. Molinari and F. Paradisi, *Nat. Catal.*, 2020, **3**, 1020–1026.
- 46 M. R. Petchey and G. Grogan, *Adv. Synth. Catal.*, 2019, **361**, 3895–3914.
- 47 A. Baldessari and C. P. Mangone, *J. Mol. Catal. B: Enzym.*, 2001, **11**, 335–341.
- 48 P. G. Quintana, G. García Liñares, S. N. Chanquia, R. M. Gorjod, M. L. Kotler and A. Baldessari, *Eur. J. Org. Chem.*, 2016, **2016**, 518–528.
- 49 R. Irimescu and K. Kato, *Tetrahedron Lett.*, 2004, **45**, 523–525.
- 50 D. Manova, F. Gallier, L. Tak-Tak, L. Yotava and N. Lubin-Germain, *Tetrahedron Lett.*, 2018, **59**, 2086–2090.
- 51 S. Zeng, J. Liu, S. Anankanbil, M. Chen, Z. Guo, J. P. Adams, R. Snajdrova and Z. Li, *ACS Catal.*, 2018, **8**, 8856–8865.
- 52 J. Pitzer, K. Steiner, C. Schmid, V. K. Schein, C. Prause, C. Kniely, M. Reif, M. Geier, E. Pietrich, T. Reiter, P. Selig, C. Stückler, P. Pöchlauer, G. Steinkellner, K. Gruber, H. Schwab, A. Glieder and W. Kroutil, *Green Chem.*, 2022, **24**, 5171–5180.
- 53 M. Winn, J. K. Fyans, Y. Zhuo and J. Micklefield, *Nat. Prod. Rep.*, 2016, **33**, 317–347.
- 54 M. Winn, S. M. Richardson, D. J. Campopiano and J. Micklefield, *Curr. Opin. Chem. Biol.*, 2020, **55**, 77–85.
- 55 J. Pitzer and K. Steiner, *J. Biotechnol.*, 2016, **235**, 32–46.
- 56 R. D. Süßmuth and A. Mainz, *Angew. Chem., Int. Ed.*, 2017, **56**, 3770–3821.
- 57 T. Abe, Y. Hashimoto, S. Sugimoto, K. Kobayashi, T. Kumano and M. Kobayashi, *J. Antibiot.*, 2017, **70**, 435–442.
- 58 R. Hara, K. Hirai, S. Suzuki and K. Kino, *Sci. Rep.*, 2018, **8**, 1–8.
- 59 M. Z. Abidin, T. Saravanan, E. Strauss and G. J. Poelarends, *Org. Biomol. Chem.*, 2021, **19**, 4515–4519.
- 60 C. M. Lelièvre, M. Balandras, J. L. Petit, C. Vergne-Vaxelaire and A. Zaparucha, *ChemCatChem*, 2020, **12**, 1184–1189.
- 61 M. Zhu, L. Wang and J. He, *Angew. Chem., Int. Ed.*, 2021, **60**, 2030–2035.
- 62 P. M. Marchetti, S. Richardson, N. M. K. Mohamed Abdelmegid and D. J. Campopiano, *MedChemComm*, 2019, **10**, 1192–1196, DOI: [10.1039/C9MD00063A](https://doi.org/10.1039/C9MD00063A).
- 63 D. Gahloth, M. S. Dunstan, D. Quaglia, E. Klumbys, M. P. Lockhart-Cairns, A. M. Hill, S. R. Derrington, N. S. Scrutton, N. J. Turner and D. Leys, *Nat. Chem. Biol.*, 2017, **13**, 975–981.
- 64 M. Winkler and J. G. Ling, *ChemCatChem*, 2022, **14**(19), e202200441, DOI: [10.1002/cctc.202200441](https://doi.org/10.1002/cctc.202200441).
- 65 A. J. L. Wood, N. J. Weise, J. D. Frampton, M. S. Dunstan, M. A. Hollas, S. R. Derrington, R. C. Lloyd, D. Quaglia, F. Parmeggiani, D. Leys, N. J. Turner and S. L. Flitsch, *Angew. Chem., Int. Ed.*, 2017, **56**, 14498–14501.
- 66 M. Lubberink, C. Schnepel, J. Citoler, S. R. Derrington, W. Finnigan, M. A. Hayes, N. J. Turner and S. L. Flitsch, *ACS Catal.*, 2020, **10**, 10005–10009.
- 67 M. Lubberink, W. Finnigan, C. Schnepel, C. R. Baldwin, N. J. Turner and S. L. Flitsch, *Angew. Chem., Int. Ed.*, 2022, **61**, e202205054.
- 68 M. R. Petchey, B. Rowlinson, R. C. Lloyd, I. J. S. Fairlamb and G. Grogan, *ACS Catal.*, 2020, **10**, 4659–4663.
- 69 M. Winn, M. Rowlinson, F. Wang, L. Bering, D. Francis, C. Levy and J. Micklefield, *Nature*, 2021, **593**, 391–398.
- 70 C. Ji, Q. Chen, Q. Li, H. Huang, Y. Song, J. Ma and J. Ju, *Tetrahedron Lett.*, 2014, **55**, 4901–4904.
- 71 M. Petchey, A. Cuetos, B. Rowlinson, S. Dannevald, A. Frese, P. W. Sutton, S. Lovelock, R. C. Lloyd, I. J. S. Fairlamb and G. Grogan, *Angew. Chem., Int. Ed.*, 2018, 11584–11588.
- 72 S. Mordhorst, A. Maurer, D. Popadić, J. Brech and J. N. Andexer, *ChemCatChem*, 2017, **9**, 4164–4168.
- 73 H. K. Philpott, P. J. Thomas, D. Tew, D. E. Fuerst and S. L. Lovelock, *Green Chem.*, 2018, **20**, 3426–3431.
- 74 C. Schnepel, L. R. Pérez, Y. Yu, A. Angelastro, R. S. Heath, M. Lubberink, F. Falcioni, K. Mulholland, M. A. Hayes, N. J. Turner and S. L. Flitsch, *Nat. Catal.*, 2023, **6**, 89–99, DOI: [10.1038/s41929-022-00889-x](https://doi.org/10.1038/s41929-022-00889-x).
- 75 Y. Ogasawara and T. Dai, *Chem. – Eur. J.*, 2017, **23**, 10714–10724.
- 76 T. Arai, Y. Arimura, S. Ishikura and K. Kino, *Appl. Environ. Microbiol.*, 2013, **79**, 5023–5029.
- 77 K. Tabata, H. Ikeda and S. Hashimoto, *J. Bacteriol.*, 2005, **187**, 5195–5202.
- 78 S. Prasad and T. C. Bhalla, *Biotechnol. Adv.*, 2010, **28**, 725–741.
- 79 M. Kobayashi and S. Shimizu, *Nat. Biotechnol.*, 1998, **16**, 733–736.
- 80 Z. Cheng, Y. Xia and Z. Zhou, *Front. Bioeng. Biotechnol.*, 2020, **8**, 1–18.
- 81 Y. Tian, J. Chen, H. Yu and Z. Shen, *J. Microbiol. Biotechnol.*, 2016, **26**, 337–346.
- 82 S. Hourai, M. Miki, Y. Takashima, S. Mitsuda and K. Yanagi, *Biochem. Biophys. Res. Commun.*, 2003, **312**, 340–345.
- 83 Y. Yamanaka, Y. Kato, K. Hashimoto, K. Iida, K. Nagasawa, H. Nakayama, N. Dohmae, K. Noguchi, T. Noguchi, M. Yohda and M. Odaka, *Angew. Chem., Int. Ed.*, 2015, **54**, 10763–10767.



- 84 J. Liu, H. Yu and Z. Shen, *J. Mol. Graphics Modell.*, 2008, **27**, 529–535.
- 85 M.-X. Wang, *Acc. Chem. Res.*, 2015, **48**, 602–611.
- 86 A. R. Mashweu, V. P. Chhiba-Govindjee, M. L. Bode and D. Brady, *Molecules*, 2020, **25**, 1–15.
- 87 L. Bering, E. J. Craven, S. A. Sowerby Thomas, S. A. Shepherd and J. Micklefield, *Nat. Commun.*, 2022, **13**, 1–10.
- 88 Á. Mourelle-Insua, L. A. Zampieri, I. Lavandera and V. Gotor-Fernández, *Adv. Synth. Catal.*, 2018, **360**, 686–695.
- 89 L. M. Lima, B. N. M. da Silva, G. Barbosa and E. J. Barreiro, *Eur. J. Med. Chem.*, 2020, **208**, 112829.
- 90 B. O. Bachmann, R. Li and C. A. Townsend, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 9082–9086.
- 91 P. L. Roach, I. J. Clifton, C. M. H. Hensgens, N. Shibata, C. J. Schofield, J. Hajdu and J. E. Baldwin, *Nature*, 1997, **387**, 827–830.
- 92 J. W. Labonte, F. Kudo, M. F. Freeman, M. L. Raber and C. A. Townsend, *MedChemComm*, 2012, **3**, 960–966.
- 93 B. Gerratana, A. Stapon and C. A. Townsend, *Biochemistry*, 2003, **42**, 7836–7847.
- 94 R. B. Hamed, J. R. Gomez-Castellanos, L. Henry, S. Warhaut, T. D. W. Claridge and C. J. Schofield, *Commun. Chem.*, 2019, **2**, 1–10.
- 95 T. Wakimoto, T. Mori, H. Morita and I. Abe, *J. Am. Chem. Soc.*, 2011, **133**, 4746–4749.
- 96 A. L. Gutman, E. Meyer, X. Yue and C. Abell, *Tetrahedron Lett.*, 1992, **33**, 3943–3946.
- 97 E. Stavila and K. Loos, *Tetrahedron Lett.*, 2013, **54**, 370–372.
- 98 N. J. Willis, C. A. Fisher, C. M. Alder, A. Harsanyi, L. Shukla, J. P. Adams and G. Sandford, *Green Chem.*, 2016, **18**, 1313–1318.
- 99 D. Koszelewski, D. Clay, K. Faber and W. Kroutil, *J. Mol. Catal. B: Enzym.*, 2009, **60**, 191–194.
- 100 C. K. Chung, P. G. Bulger, B. Kosjek, K. M. Belyk, N. Rivera, M. E. Scott, G. R. Humphrey, J. Limanto, D. C. Bachert and K. M. Emerson, *Org. Process Res. Dev.*, 2014, **18**, 215–227.
- 101 M. D. Truppo, J. David Rozzell and N. J. Turner, *Org. Process Res. Dev.*, 2010, **14**, 234–237.
- 102 G. A. Aleku, S. P. France, H. Man, J. Mangas-Sanchez, S. L. Montgomery, M. Sharma, F. Leipold, S. Hussain, G. Grogan and N. J. Turner, *Nat. Chem.*, 2017, **9**, 961–969.
- 103 Z. Qin, X. Zhang, X. Sang, W. Zhang, G. Qu and Z. Sun, *Green Synth. Catal.*, 2022, **3**(3), 294–297, DOI: [10.1016/j.gresc.2022.05.009](https://doi.org/10.1016/j.gresc.2022.05.009).
- 104 S. Sarak, T. P. Khobragade, H. Jeon, A. D. Pagar, P. Giri, S. Lee and H. Yun, *Green Chem.*, 2021, **23**, 9447–9453.
- 105 M. S. Christodoulou, M. L. Contente, S. Dallavalle and A. Pinto, *Green Chem.*, 2022, **24**, 4432–4436.
- 106 L. Zhang, F. Li, C. Wang, L. Zheng, Z. Wang, R. Zhao and L. Wang, *Catalysts*, 2017, **7**, 115.
- 107 S. V. Athavale, S. Gao, Z. Liu, S. C. Mallojjala, J. S. Hirschi and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2021, **60**, 24864–24869.
- 108 X. Ren, A. L. Chandgude and R. Fasan, *ACS Catal.*, 2020, **10**, 2308–2313.
- 109 Y. Yang and F. H. Arnold, *Acc. Chem. Res.*, 2021, **54**, 1209–1225.
- 110 H. Renata, Z. J. Wang, R. Z. Kitto and F. H. Arnold, *Catal. Sci. Technol.*, 2014, **4**, 3640–3643.
- 111 Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho and F. H. Arnold, *Angew. Chem., Int. Ed.*, 2014, **53**, 6810–6813.
- 112 O. F. Brandenburg, C. K. Prier, K. Chen, A. M. Knight, Z. Wu and F. H. Arnold, *ACS Catal.*, 2018, **8**, 2629–2634.

