A Bio-*logical* Approach to Catalysis in the Pharmaceutical Industry

Greg Mann* and Frédéric V. Stanger

Abstract: Enzymes have the potential to catalyse complex chemical reactions with unprecedented selectivity, under mild conditions in aqueous media. Accordingly, there is serious interest from the pharmaceutical industry to utilize enzymes as biocatalysts to produce medicines in an environmentally sustainable and economic manner. Prominent advances in the field of biotechnology have transformed this potential into a reality. Using modern protein engineering techniques, in a matter of months it is possible to evolve an enzyme, which fits the demands of a chemical process, or even to catalyse entirely novel chemistry. Consequently, biocatalysis is routinely applied throughout the pharmaceutical industry for a variety of applications, ranging from the manufacture of large volumes of high value blockbuster drugs to expanding the chemical space available for drug discovery.

Keywords: Biocatalysis · Enzymes · Green · Sustainable · Synthesis

Greg Mann received his PhD in Enzyme Biochemistry in 2016 from the University of St Andrews, Scotland, under the supervision of Prof. James H. Naismith, where he investigated enzymes catalysing the synthesis of diverse cyclic peptides. He then worked at the Swiss Institute of Technology, Lausanne as a postdoctoral researcher in the laboratory of Prof. Oliver Hantschel, where he conducted research in the field of therapeutic protein engineering, developing monobodies as potent biotherapeutics against notorious oncoproteins. In 2019, he joined Novartis Pharma AG, Basel as Molecular Biologist in the Biocatalysis team. In his current role, he uses advanced molecular biology techniques and high-throughput robotics to generate enzyme libraries for biochemical screening, for which he then conducts next generation sequencing experiments and analyses.

Frédéric V. Stanger received his PhD in Structural Biology in 2014 from the University of Basel. He then conducted postdoctoral research at the Biozentrum of the University of Basel. From 2016 to 2019, he received two postdoctoral fellowships from the Swiss National Science Foundation to conduct research at the Johns Hopkins University School of Medicine in Baltimore, USA. Since 2019, he joined Novartis Pharma AG in Basel where he holds the position of Bioinformatician/ Computational Biologist in the Biocatalysis team. His major focus is on using state-of-the-art computational tools for enzyme modelling, substrate docking and prediction of favourable mutations as well as automatization of the analysis of high-throughput screening data to support the need-for-speed in protein engineering.

1. Introduction - Biocatalysis in the Pharmaceutical Industry

Over millennia, nature has evolved enzymes to control and conduct chemical reactions with exquisite chemo-, regio- and enantioselectivity. In this regard, these biocatalysts rival and surpass even the best manmade catalysts. Furthermore, enzymes are nontoxic, constructed of readily available, renewable, inexpensive materials and are specially adapted to function in water and under mild conditions, making them highly attractive tools for industrial application.[1]

Since the dawn of history, humankind has harnessed the advantageous properties of enzymes, albeit indirectly, in food and beverage production, through the fermentative activity of microorganisms. However, enzymes themselves were not recognized as the active agent of these biologically mediated chemical transformations until the mid-19th century and isolated enzymes were not used until the early 20th century.[2] In the first half of the 20th century industrial applications of enzymes were limited; the food industry made use of polysaccharide hydrolysing enzymes,[3] while proteases and lipases were utilized in laundry detergent manufacture.[4] The development of molecular biology techniques in the 1970s paved the way for a new era; an era of biotechnology. It is now possible to wield DNA and thus produce proteins in the laboratory. This facilitated enzyme production, expanding the number of enzymes available to the chemical industry. However, the available wild-type enzymes suffered from an overall poor stability, low activity and limited substrate tolerance, which hindered their industrial applicability, particularly for the production of fine chemicals and pharmaceutically relevant compounds.[5] Over the last 50 years, the fledgling field of biotechnology has grown exponentially and it shows no signs of slowing down. Numerous improvements in this area, encompassing developments in molecular biology, structural biology, biochemistry and computational biology have contributed to heightened interest in the field of biocatalysis, as reflected by the growing number of publications and patents.[1,6] Chiefly, advanced protein-engineering techniques, utilizing both rational and stochastic approaches, have made it possible to manipulate the properties of enzymes. Scientists can alter the specificity of an enzyme towards a select substrate, improve its stability under challenging, unnatural conditions, which favours industrial processes, and are able to generate enzymes capable of novel chemistry. Essentially it is possible to evolve enzymes that are fit for a specific process, rather than trying to fit

Correspondence: Dr. G. Mann, E-mail: greg.mann@novartis.com
Chemical and Analytical Development, Novartis Pharma AG, CH-4056 Basel, Switzerland

the process to the available enzymes;[7] an important distinction that broadens the industrial scope of biocatalysis, especially in the pharmaceutical industry. The dogma that biocatalysis is resigned to a few selected, highly specialized transformations was lifted and a huge variety of different enzymes from different families and classes are presently being used to catalyse a growing number of reactions *en route* to the production of single enantiomer APIs.[8] Furthermore, the environmental need to reduce waste, use fewer toxic chemicals and focus on greener, more sustainable chemistry is paramount; a need for which enzymes are ideally suited to provide a solution.[1,9] Therefore, it is foreseen that biocatalysis will be at the forefront of the pharmaceutical industry for years to come.

This review article aims to highlight how biotechnological advancement – especially protein engineering – has been integral in realizing the potential of biocatalysis, and discuss the state-of-the-art developments in biocatalysis throughout the pharmaceutical industry.

2. Advances in Biotechnology have Empowered Industrial Biocatalysis

Numerous technological advances in recent years have contributed to the implementation of a greater number of biocatalytic synthetic methodologies within the pharmaceutical industry. These advances, primarily in the areas of molecular biology and bioinformatics have not only deepened scientists' understanding of the relationship between an enzyme's sequence, structure and its function, which is an important prerequisite for the rational design of bespoke catalysts but also provided tools to rapidly access novel enzyme sequences in a cost-efficient manner. The combination of these new technologies constitute an intricate and effective toolkit for the modern protein engineer and will be briefly discussed below.

Next-generation DNA sequencing technology enabled the sequencing of entire genomes of organisms from all domains of life and from different environments all over the world, providing a plethora of new enzyme sequences.[10] In parallel, improvements in bioinformatics allow scientists to harness this vast amount of information, to mine and analyse these data in the hunt for new biocatalysts.[11] Access to these newly identified genes has been simplified and expedited due to developments in DNA synthesis techniques.[12] Commercial DNA synthesis services offer synthetic DNA in an optimized format (codon optimization for different expression hosts, removal of restriction enzyme sites, *etc*.) quickly and at low cost. As such, it is seen as the preferred starting point when acquiring new enzymes, compared with the more hands-on approach of isolating DNA from exotic organisms and cloning the gene of interest from the obtained genomic DNA. Complementing the upsurge in the availability of sequence information, there has been a steady increase in the number of protein structures available; the RCSB Protein Data Bank (*http://www.rcsb.org*) contains over 150'000 protein structures to date. Given the inherent relationship between an enzyme's primary sequence of amino acids, its three-dimensional structure and its function, the accrual of sequence and structural information has empowered computational tools, which can be used to predict and model enzymatic reactions and guide scientists in their search for suitable enzymes.[13] Nevertheless, these predictions are not perfect and it is often necessary to screen large numbers of enzymes to identify candidates which are capable of the desired chemistry. Historically, such a task was laborious and time consuming. However, improvements in high-throughput robotics and automation of liquid handling, sample analytics and even data analysis permit the rapid and robust identification of the most promising enzyme candidates.[14] In some cases, the activity of these wild-type enzymes is sufficient to be utilized at scale in the chemical process. More often, the enzyme is challenged with substrates it

has never encountered in nature and under reaction conditions for which it has not evolved. In these cases enzyme performance can be optimized through protein engineering.

Once an enzyme exhibiting measurable activity for a desired reaction has been identified, it is then possible to alter the properties of that enzyme *via* protein engineering. The methodology aims to mimic natural evolution, whereby genetic mutations engender changes in the amino acid sequence, and these new amino acids can affect the functional properties of the enzyme. In contrast to nature, where these adaptations occur over millennia, scientists can accelerate the process by artificially installing mutations and subjecting these mutants to a predefined measure of fitness in the laboratory, in a process called directed evolution.[15] Those mutations, which confer an improvement compared with their parent, are themselves used as parents in further rounds of evolution. However, the strategy is not to create huge libraries of random mutations and then conduct an exhaustive search for the illusive needle in a haystack. Instead, by utilizing the aforementioned advances in molecular biology, computational biology and automated robotics, it is possible to expedite this process. Evolution can be streamlined through the computationally guided design of directed evolution libraries focused on the regions of the enzyme most likely to bring the desired change, thus reducing the screening burden of the campaign. Computational analysis of the protein sequence, structure and activity relationships of all the mutants screened in the library allow the identification of beneficial and deleterious mutations, which in turn inform the design of 'smart' libraries with a high probability of introducing beneficial mutations in subsequent rounds of evolution.[16] Following such an iterative cycle, beneficial mutations accumulate over generations until the goal of the evolutionary campaign is accomplished. In this way, in only a few rounds of evolution and a few months of laboratory work, it is possible to improve the activity as well as the chemo-, regio- and enantioselectivity of an enzyme for a reaction of interest. Furthermore, protein engineering is routinely employed to enhance the enzymes' tolerance to organic solvents, extremes of pH and elevated temperatures.

3. Engineered Enzymes Drive Drug Development Forward

The speed and robustness of modern protein engineering approaches has elevated the technology from an academic exercise to something that can bring real value to the pharmaceutical industry, in its quest to find selective, sustainable, and cost effective routes to produce complex molecules. This last decade has seen a number of high-profile success stories for the implementation of biocatalytic steps, more often than not enabled by protein engineering, in the synthesis of blockbuster drugs, for which some examples are given below.

Sitagliptin (**1**) is the active pharmaceutical ingredient (API) in Januvia, a leading drug for the treatment of type 2 diabetes and the largest product by both volume and sales in Merck & Co/MSD's portfolio.[17] In the synthesis of sitagliptin, the central chiral amine of the molecule was installed *via* asymmetric hyodrogenation of an enamine using a rhodium-based chiral catalyst.[18] Merck & Co. sought to replace the rhodium-based catalyst, which catalysed the reaction with suboptimal stereoselectivity and led to rhodium traces in the product, requiring additional purification steps, with a transaminase biocatalyst (Fig. 1.). Native transaminases suffer from limited substrate promiscuity and none of the commercially available transaminases demonstrated measurable activity towards the bulky, prositagliptin ketone substrate **2**.[19] Following extensive protein engineering over the course of a year, a transaminase variant, containing 27 mutations with respect to the inactive wild-type enzyme was evolved, able to convert 200 g/l prositagliptin ketone to sitagliptin with an assay yield of 92% and at >99.95% ee using enzyme

loading of 3 wt%.[19] Crystallographic studies of the initial enzyme, ATA-117 and its variant G136F revealed that the introduced mutation modifies the conformation of a loop next to the active site, which results in a significantly modified-substrate binding site. The modification of the volume and pocket properties allows the accommodation of the large ketone substrate.[20] Importantly, compared with the rhodium-catalysed process, the biocatalytic process increased the overall yield of sitagliptin by up to 13%, improved productivity by 53% and reduced overall waste by 19%, as well as excluding the need for heavy metals and reducing manufacturing costs.[19]

in this synthesis, protein engineering was required to optimise their performance. The activity of the KRED and the GDH were improved by a factor of 7 and 13 respectively, while the performance of the HHDH was improved ~4,000-fold.[16,21b] This led to a highly efficient final process where starting material is converted to product with an isolated yield of >90%, with >98% chemical purity and >99.9% enantiomeric purity.[21b] In addition, this enzymatic route is much greener than the previous chemical syntheses, in accordance with the twelve principles of green chemistry.[21b,23]
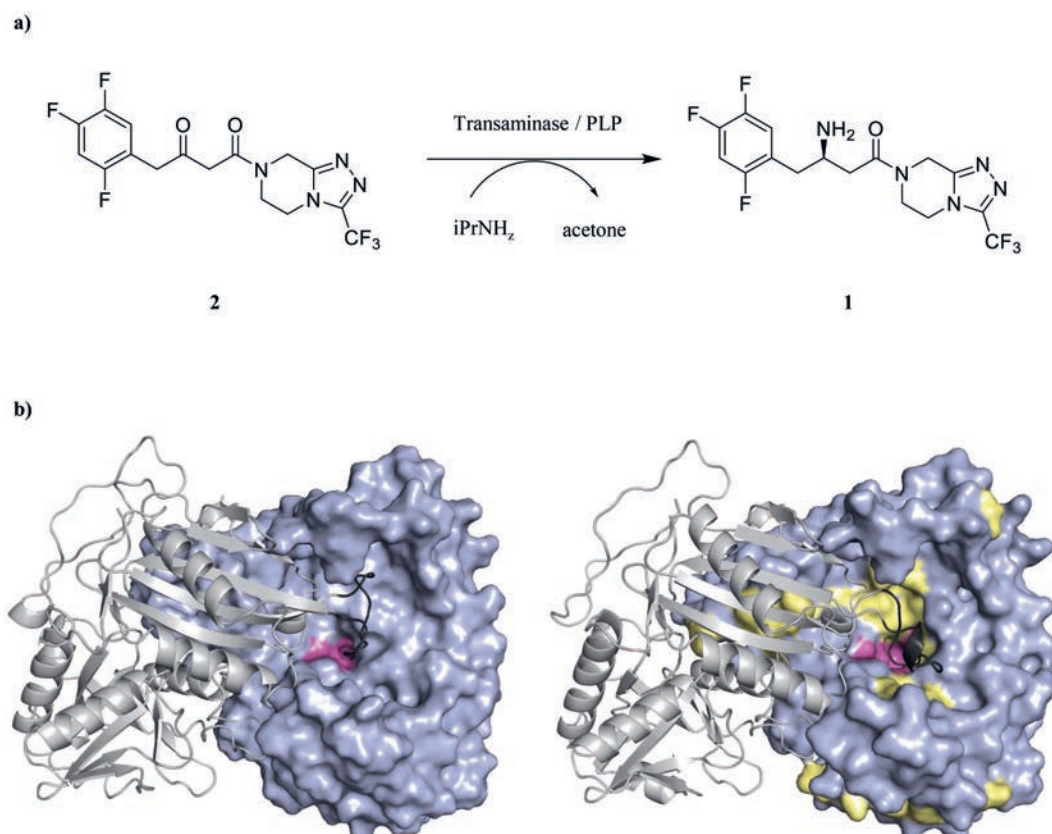

Fig. 1 a) Enzymatic transamination of prositagliptin ketone to form sitagliptin (**1**). b) The physiological homodimer of ATA-117 is displayed with one monomer shown as a grey cartoon and the other one as a blue surface representation. The cofactor (PLP) binding site is shown in magenta to highlight the enzyme active site, and the 27 positions mutated in ATA-117-RdII to yield the commercial target are highlighted in yellow. Mutations cluster at the active site, dimeric interface and surface-exposed patches. Noteworthy mutation G136F results in a conformational change of the loop highlighted in dark grey at the active site as described by the recent crystallographic study published by Guan *et al.*[20] ATA-117 homology model was built using the crystallographic structure 3wwh (>99% sequence identity to ATA-117) and ATA-117-Rd11 model was refined by extending its N-terminus based on the crystallographic structure 5fr9.

Atorvastatin (**3**) is the active ingredient in Pfizer's cholesterol lowering drug Lipitor, the most commercially successful drug of all time, generating peak sales of nearly $13 billion in 2007. Not surprisingly there has been significant interest in developing efficient, sustainable, cost-effective routes to synthesize atorvastatin, and multiple routes involving biocatalysis have been developed.[7,21] The focus of these methods has been on the synthesis of the chiral 3,5-dihydroxy acid side chain, a crucial element of atorvastatin and other statin-based drugs, which is essential for activity.[21a] One approach introduces the two stereogenic centres in one reaction by utilising deoxyribose-5-phosphate aldolase (DERA) to catalyse a one-pot tandem aldol reaction of simple, achiral building blocks *via* a six-membered lactol intermediate.[21a,22] An alternative synthetic route proceeds *via* ethyl (*R*)-4-cyano-3-hydroxybutyrate (**4**) (a.k.a. 'hydroxynitrile') as an intermediate, which is prepared in a two-step, three-enzyme process (Scheme 1).[21b] In the first step, a ketoreductase (KRED) converts ethyl 4-chloro-acetoacetate (**5**) to ethyl (*S*)-4-chloro-3-hydroxybutyrate (**6**), introducing one of the key stereocentres. The essential cofactor, NADPH is recycled by the addition of a second enzyme, glucose dehydrogenase (GDH) to the reaction mixture (Scheme 1). In a second step, the key hydroxynitrile intermediate **4** is generated using a halohydrin dehalogenase (HHDH). For all enzymes employed

In a more recent example, the synthesis of sacubitril (**7**), one of two APIs (the other being valsartan (**8**)) in Novartis's blockbuster drug sacubitril/valsartan used in the treatment of heart failure, also benefited from the implementation of biocatalysis (Scheme 2). In contrast to sitagliptin and atorvastatin, where biocatalytic steps were introduced to replace chemocatalytic steps within the existing synthetic routes, the synthetic route of **7** was completely redesigned with the specific aim to incorporate biocatalysis, to improve the manufacturing process and reduce the environmental
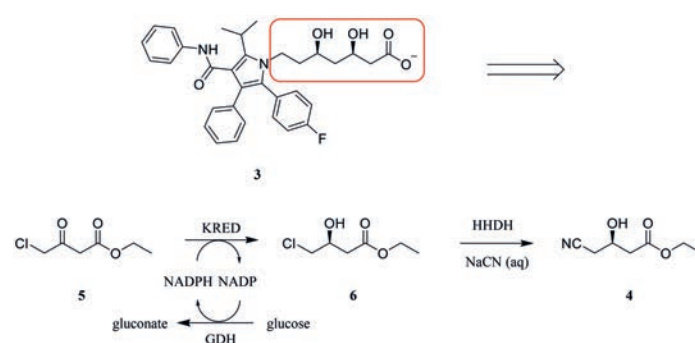
Scheme 1. Three-enzyme, two-step process for the synthesis of a key atorvastatin intermediate.

footprint in anticipation of large production volumes. The synthesis redesign focused on the production of key amino acid intermediate **9**. In this new route, the first chiral centre is established by asymmetric hydrogenation of a readily available itaconic acid monomethyl ester to form (*R*)-methyl succinic acid monomethyl ester. The second chiral centre was introduced using a transaminase to convert the γ-ketoacid **10** to the desired (2*S*,4*R*)-amino acid **9**.[24] After screening a panel of transaminases, one transaminase that could process trace amounts of the ketone substrate, was identified but delivering exclusively the undesired (*R*)-amine.[24a] Therefore, a comprehensive protein engineering campaign was required to enhance the performance of the enzyme to meet the synthesis demands. In the first round of evolution, a variant with a single point mutation was obtained and proved to be sufficient to completely reverse the selectivity of the enzyme to produce the desired (*S*)-amine and improve the activity 3-fold.[24a] A further ten rounds of evolution and an additional 25 mutations were required to improve the activity and the thermal stability of this enzyme. The final enzyme, active up to 65 °C and tolerant of up to 2 M isopropylamine, could convert more than 90% of the γ-ketoacid substrate to the (2*S*,4*R*)-amino acid product, at a substrate concentration of 75 g/L with an enzyme loading of 1 wt%.[24] The environmental impact of this new synthesis was ascertained through measurement of the total carbon dioxide release, and the new route, with solvent recycling employed reduced the total carbon dioxide release 3-fold, with respect to the original synthesis.[24b]
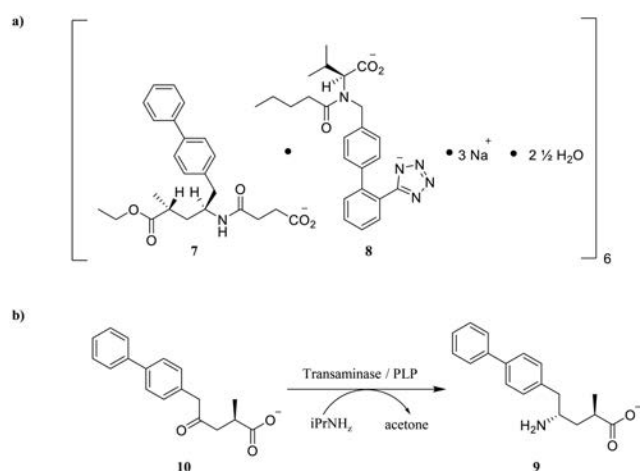

Scheme 2. a) Sacubitril/valsartan contains LCZ696, which is a crystalline supramolecular complex that combines the two APIs – sacubitril (**7**) and valsartan (**8**) in their respective anionic forms in a 1:1 molar ratio, together with sodium and water. b) An evolved transaminase was used to convert ketone **10** to key amino acid intermediate **9**.

The implementation of biocatalysis within the synthesis of blockbuster drugs such as sitagliptin, atorvastatin and sacubitril illustrates the value that biocatalysis can bring to the pharmaceutical industry. However, for all three examples described above, a biocatalytic step was introduced after product launch. This requires refiling of the new process with regulatory agencies around the world, which is a time consuming and costly process.[17a] To deliver the highest impact, biocatalysis should be incorporated into the synthetic strategy as early in the drug development process as possible so that the drug can be manufactured with the most efficient chemistry by the time the product is ready for launch. Efficient protein engineering technology is paramount to achieve this goal and the importance of investing in such technology is reflected by the fact that three major, global pharmaceutical companies, (GSK, Merck and Co. and Novartis) have recently internalized a dedicated protein-engineering platform, pioneered

by Codexis, who are leaders in the protein-engineering field.[25] While advanced techniques in protein engineering are imperative, biocatalytic process development is also a key step. One should not wait for the perfect enzyme to be evolved before integrating it into the synthetic strategy of a developmental compound. Instead, process development should begin with a suboptimal biocatalyst, while the enzyme is evolved in parallel. Lessons learned from process development can guide the evolution campaign, while improved enzyme variants in turn can shape the chemical process; and so the iterative cycle continues until an optimum biocatalytic process is developed. Importantly, if a compound under clinical investigation is dropped from the portfolio then the investment in biocatalysis development is not wasted. The evolved enzymes can be used in future screening campaigns for new reactions, and the increased diversity generated over multiple evolution programs can increase the chances of rapidly finding a hit for a particular reaction.

In support of the idea that pharmaceutical companies should be considering biocatalysis as a Plan A, GSK have implemented an IRED-catalysed reductive amination into the synthesis of GSK2879552 (**11**), a lysine-specific demethylase-1 (LSD1) inhibitor, currently undergoing Phase II clinical trials for the treatment of acute leukaemia and small cell lung cancer.[26] To identify enzymes capable of the desired transformation, GSK screened their in-house panel of IREDs. Fortuitously, a number of enzymes were identified with good activity, generating the correct enantiomer with high enantiomeric excess.[26] Nevertheless, it was determined that protein engineering was required to develop a commercially viable process. In just three rounds of evolution, an enzyme variant displaying nearly a 40,000-fold improvement over the wild-type was identified, which was used to generate kilogram quantities of a key intermediate to GSK2879552 in 84% yield, at 99.9% purity and > 99.7% enantiomeric excess with a concomitant benefit in process mass intensity (Scheme 3).[26]
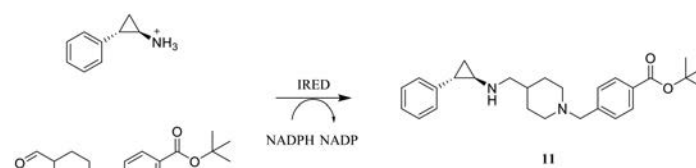
Scheme 3. IRED catalysed conjugation to form GSK2879552 (**11**).

Taking this one step further, scientists at Merck and Co. sought to introduce a biocatalysis step into the synthesis of MK-7246 (**12**), a CRTH2 (*c*hemo-attractant *r*eceptor expressed on *Th2* cells) antagonist for the potential treatment of respiratory disease, in early stage development, in preparation for the upscaling required for clinical trials.[27] In search for a transaminase that could selectively convert ketone **13** into the corresponding (*R*)-amine **14**, they screened a panel of (*R*)-selective transaminases that were previously evolved for the sitagliptin project. Gratifyingly, they identified an enzyme from this (*R*)-selective transaminase panel that could convert ketone **13** with a yield of 81%, giving **14** with 99% ee (Scheme 4).[27] This reemphasizes the advantage of investing in protein engineering technology, where the fruits of one project may accelerate the development of future projects. The transaminase step was reproducible at scale in the pilot plant (>100 Kg) leading to an isolated yield of 76%.[27]

The aforementioned examples highlight how the application of biocatalysis in the manufacture of pharmaceuticals can result in a streamlined synthesis, with measureable improvements towards environmental sustainability, laboratory safety and appreciable economic benefits. To maximise these benefits it may be wise to introduce multiple biocatalytic steps into a single synthesis when
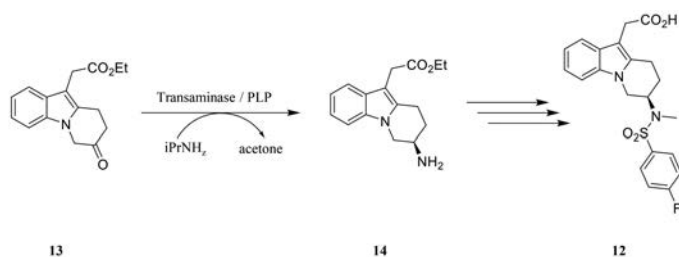
Scheme 4. A transaminase step converting ketone **13** to amine **14** was introduced en route to the synthesis of MK-7246 (**12**).

possible. Inspired by nature, where complex molecules are assembled *via* multi-enzyme cascades in one cell, there is interest in developing one-pot biocatalytic cascades *in vitro*. In such cascades, the selective nature of enzymes and their reactive compatibility potentially negate the need to isolate intermediates.[28] This has the advantage of reducing the amount of chemicals required to extract and purify the product from the reaction mixture, as well as the time and cost associated with establishing such methodology. Furthermore, for equilibrium-limited reactions, the addition of a second enzyme, which utilises product 1 as a substrate and converts it to product 2, can drive a reaction to completion, thus improving the overall yield of the synthesis.[28] Epitomising this concept, Merck and Co. have developed a biocatalytic cascade for the synthesis of islatravir (**15**), a nucleoside analogue currently under investigation for the treatment of HIV.[29] The biocatalytic cascade constructs islatravir from simple, achiral building blocks using nine enzymes in just three steps with an overall yield of 51% (Scheme 5).[29] Despite the need for extensive protein engineering of multiple enzymes to enable such a cascade, the advantages are clear: Firstly, the atom economy considerably surpasses that of earlier syntheses, which required between twelve and eighteen steps. Secondly, the entire synthesis occurs in aqueous media under mild conditions, without the need to isolate intermediates, supporting green chemistry initiatives and reducing the overall cost of production.[29]

These six examples demonstrate the potential of biocatalysis to shorten synthetic routes, catalyse chemical reactions with precise selectivity in a way that is more environmentally friendly and cost-effective compared with traditional chemical methods.

Consequently, biocatalysis forms an essential part of the drug development toolkit. As protein engineering technology continues to advance, enabling more and more biocatalysts, it would appear that the present utilisation of biocatalysis within the pharmaceutical industry is just the tip of the iceberg.

4. Biocatalysis Broadens the Possibilities of Drug Discovery

Beyond the role of biocatalysis in drug development and manufacture as described above, enzymes are frequently applied in multiple aspects of the drug discovery process. Enzymes are used in the generation of an array of chiral building blocks, which can be incorporated into the synthesis of new lead compounds. Such chiral building blocks include non-proteinogenic amino acids,[30] chiral amines[31] and enantiomerically pure inositol phosphates.[32] Furthermore, enzyme classes including cytochrome P450 monooxygenases, xanthine oxidoreductases, flavin monooxygenases, aldehyde oxidase and UDP-glucuronosyl transferases are utilised in the synthesis of drug metabolites.[33] The synthesis of these metabolites enable scientists to investigate the metabolic pathway, and the corresponding biological activity and toxicity of a drug candidate.[34] Moreover, the selective nature of enzymatic transformations make them attractive tools in medicinal chemistry and late stage functionalisation. In this regard, enzymes catalysing oxidation, halogenation, phosphorylation and *O*-methylation have gained the most attention, and with the aforementioned acceleration in the discovery and characterisation of novel enzymes, it is expected that medicinal chemists can make use of many more enzymes in the future.[35] These three aspects (synthesis of chiral building blocks, metabolism investigation and late stage functionalisation) are thoroughly covered in a review by Schwarz *et al.*, which is published in this issue of *Chimia*.[36] Instead, this review looks at how the ability of enzymes to conduct delicate chemistry in aqueous media can benefit the drug discovery process, focusing on examples in peptide drug discovery and DNA encoded chemical libraries, which will be described in the next paragraphs.

4.1 *Biocatalysis in Peptide Drug Discovery*

Peptide-based drugs exist in a chemical space between small molecules and biologics but display advantages of both classes of molecules. Like biologics, their larger size favours the disruption
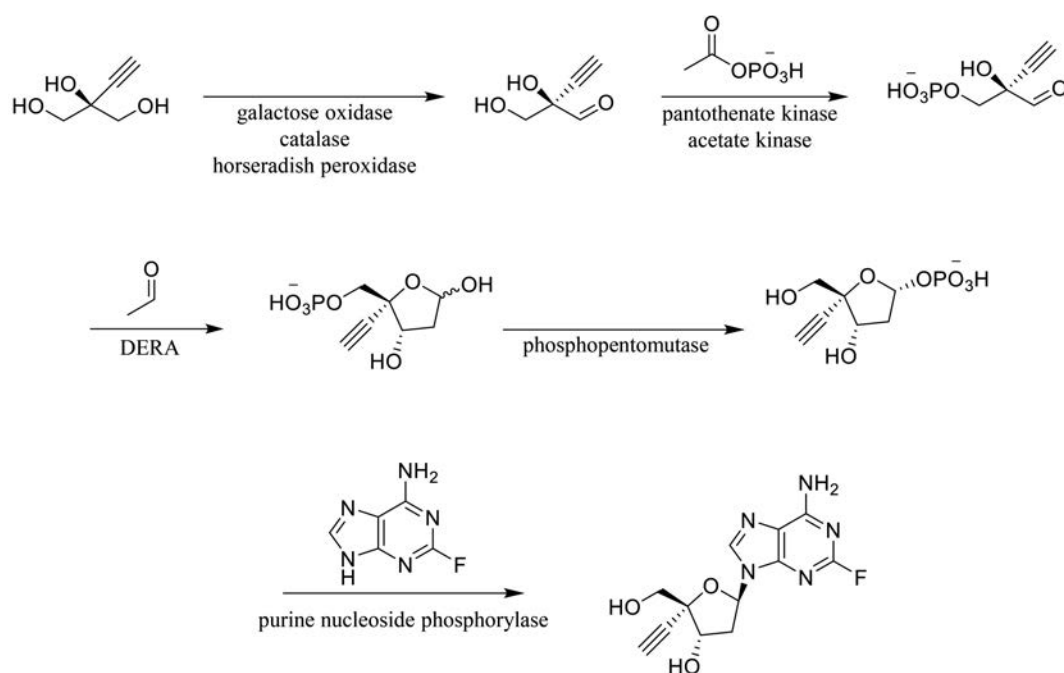
Scheme 5. Reaction scheme highlighting the biocatalytic cascade synthesis of islatravir.

15

of protein–protein interactions, which are typified by a large surface area and the lack of defined binding pockets, making them challenging to target with small molecules.[37] Furthermore, the extended interface of a peptide-target interaction, increases selectivity, reducing off-target toxicity and undesired side effects.[37a,38] However, unlike biologics, drugs derived from peptides can potentially to be taken orally sharing a major advantage with small molecule drugs;[39] although it is important to note that the peptide must often be significantly modified to be made more 'drug-like' to realize this potential. In particular, considerable alterations to the peptide scaffold are required to increase the lipophilicity of the molecule to facilitate passage across cell membranes to improve their absorption properties. Furthermore, modification is often required to enhance their stability against proteolytic degradation.[40] To decorate a synthetic peptide to such an extent, in a site-selective and non-destructive way has proved a challenge to synthetic chemists. As ever, nature is capable of such modification and offers a vast number of bioactive peptide-based natural products, many having been exploited as therapeutic agents. Prominent examples include the antibacterial agent vancomycin (16),[41] the immunosuppressant cyclosporine A (17),[42] and α-amanitin (18), which is being explored as a cytotoxic warhead as part of antibody drug conjugates (ADCs)[43] (Fig. 2).
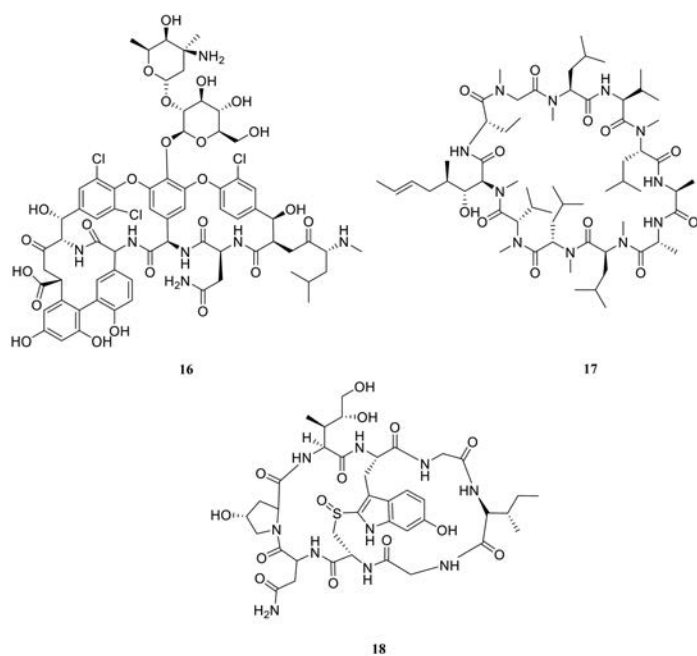

Fig. 2. Pharmaceutically relevant peptide-based natural products: vancomycin (16), cyclosporine A (17) and α-amanitin (18).

There are two biosynthetic routes to peptide-based natural products: Non-Ribosomal Peptide Synthesis (NRPS)[44] and Post Ribosomal Peptide Synthesis (PRPS).[45] The former involves huge modular protein complexes (>1 MDa), where each module, through its associated domains incorporates (and modifies if necessary) a specific amino acid into a growing peptide chain.[44] The immensity of these synthetases coupled with the observed intractability of the individual domains in isolation has hindered their applicability for the biotransformation of peptides *in vitro*. In contrast, PRPS pathways are more malleable. Here, ribosomally-synthesized and post-translationally modified peptides (RiPPs) are produced initially by the ribosome, which are post-translationally modified to become the final product (Fig. 3).

These post-translational modification (PTM) enzymes catalyse a variety of transformations including proteolysis, macrocyclisation, heterocyclisation, dehydration, methylation and prenylation, amongst others.[46] More often than not enzyme activity
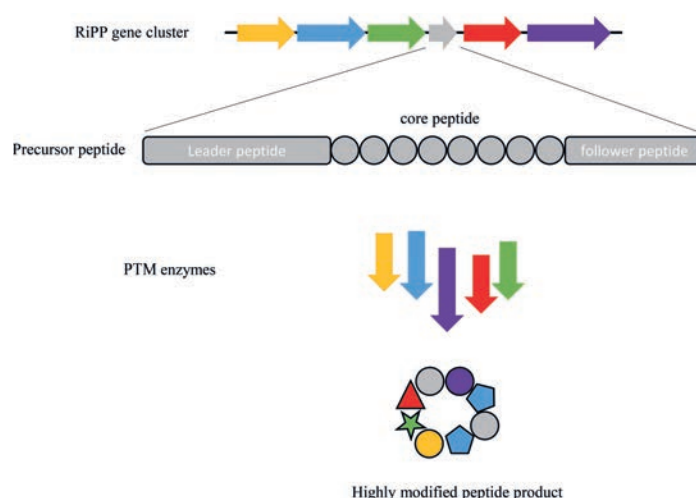

Fig. 3. Cartoon representation of a RiPP gene cluster encoding the precursor peptide (grey) and post-translational modification (PTM) enzymes. The precursor peptide is composed of a core peptide, which is flanked by recognition elements within the leader and follower peptides, which direct PTMs to modify the core peptide to generate the highly modified product.

can be attributed to an individual protein or even a single domain, which retains activity in isolation; these smaller units are more amenable to recombinant production and protein engineering pipelines. Furthermore, these enzymes have evolved to segregate substrate recognition from catalysis; they recognize short amino acid sequences outside of the peptide sequence that they modify. These recognition elements are ultimately disposed of as the peptide matures to the final product.[47] By including these recognition sequences onto a peptide of interest, it is possible to direct an enzymatic transformation to this peptide. In one approach, for an enzyme that catalyses the cyclodehydration reaction of cysteine residues to form thiazoline heterocycles, scientists fused the enzyme's recognition element to the enzyme itself, creating a constituently active enzyme, capable of acting on a variety of substrate peptides.[48] Other enzymes achieve catalytic promiscuity by recognising simple chemical motifs within the context of the larger peptide, irrespective of the surrounding amino acid residues.[46e,49] The relaxed substrate specificity offered by these enzymes give rise to remarkable biotechnological plasticity making them promising tools for the design of bespoke peptide drugs.

4.2 *Biocatalysis in DNA-encoded Libraries*

In the search for novel small molecule drugs, pharmaceutical companies frequently turn to high-throughput screening (HTS) as a method to identify active compounds from large chemical libraries. However, such campaigns are logistically challenging and are often limited to screening fewer than one million compounds, representing a tiny fraction of the possible chemical space.[50] Moreover, the identified hits often lack the potency and selectivity required to become a drug.[50] While it might be possible *via* medicinal chemistry to improve the properties of said hit, scientists are left wondering "what if it were possible to screen more compounds?" DNA-encoded libraries (DELs) may provide the answer to this question. DELs contain pools of lower molecular weight fragments, each one tagged with a unique sequence of DNA, encoding that fragment. Diversity of the library is expanded *via* split and pool synthesis, whereby the fragment pool is split and a second chemical fragment is connected to the first, then a second sequence of DNA, uniquely identifying that second fragment is ligated to the end of the first DNA molecule (Fig. 4). The DNA tags function as a barcode, tracking the synthesis of the compounds in the library. The library is then pooled and the cycles repeated until

a library of the desired size is created.[51] In general the library size can be defined by x to the power of y (x^y), where x is = the number of unique fragments and y is = the number of cycles used to create the library.[52] In this way the split and pool libraries can grow to contain huge numbers of compounds, with libraries containing greater than 800 million unique compounds reported.[52] Such libraries are too vast to be screened by even the best-equipped HTS labs. Instead, in DELs, 'hits' are identified *via* a selection process. The most common selection strategy involves immobilising a protein of interest (POI) onto a solid support, such as *e.g.* magnetic beads. The immobilised protein is incubated with the entire DEL in one test tube. Those compounds that bind the POI are 'captured' as the immobilised protein is isolated from library pool. Stringent washing procedures serve to remove compounds that bind non-specifically and enrich for compounds that bind the POI with the highest affinity. The DNA tags associated with the 'hits' are amplified and sequenced, and the sequence informs the researcher on the exact identity of the small molecule ligand.[51]

architecture containing multiple stereocentres, which requires highly selective chemistry. In a paper by Thomas *et al.,* the authors demonstrate that enzymes can be used to perform the site selective transglycosylation of sialic acid onto variety of mono- and disaccharides (Scheme 6). The conversions of these enzymatic reactions are modest, yielding 42–78% product but importantly the reactions were highly specific with no undesired side products observed.[54] In addition, the authors use galactose oxidase (GOase) to selectively oxidize the C(6) alcohol of a variety of sugars with excellent conversions (82–94%) (Scheme 6).[54] The resulting aldehyde was used as a handle for chemical reductive amination for the site-specific introduction of orthogonal functionality.

Importantly the authors note that the DNA remained unmodified by the biotransformations.[54] It should be noted that the enzymes used in this study were not optimised for the purpose, and it should be possible to improve the conversion of the transformations through protein engineering. Nevertheless, this

Fig. 4. Schematic showing the split and pool synthesis procedure of DELs over two cycles. In cycle 1, small molecule building blocks are covalently attached to a unique piece of DNA. In cycle 2, a second small molecule fragment is connected to the first and a DNA corresponding with that fragment is ligated to the end of the initial DNA molecule, allowing the synthesis to be tracked.
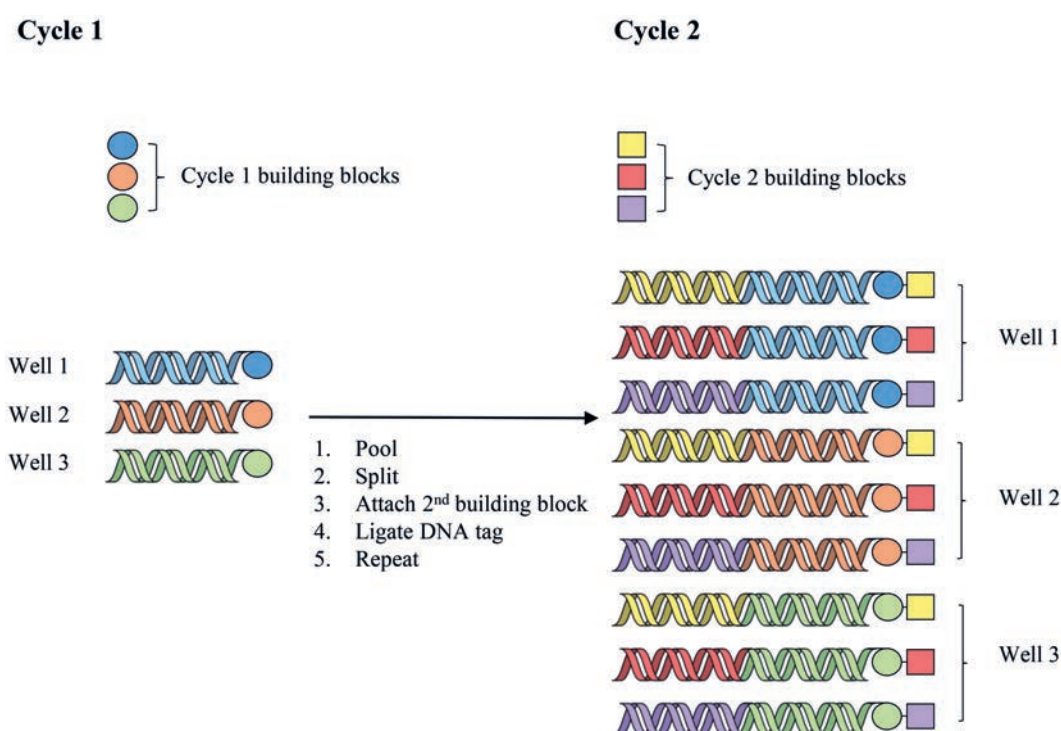
The main advantage of DEL technology is the ability to explore a chemical space containing a huge variety of structural diversity. However to create such a library the chemistry must be compatible with DNA. If the chemistry is destructive to DNA then the ability to identify compounds of interest is lost.[52] While the fragment scaffolds can be synthesised prior to attachment to DNA and so can be manufactured without restriction, many standard methods to assemble two fragments together asymmetrically may be at least partly destructive to DNA.[53] In general reactions which require harsh acidic or basic conditions, photoredox reactions and reactions involving radical intermediates cannot be employed.[52] One solution could be to combine DEL technology with biocatalysis; the ability of enzymes to conduct selective chemistry in aqueous media under mild conditions could expand the synthetic repertoire of DELs, while maintaining DNA integrity. In one example, scientists at GSK and the University of Manchester expanded the chemical space of their DELs by using enzymes to create a carbohydrate-based DEL.[54] Carbohydrates are highly functional molecules in biology but carbohydrate moieties are often underexplored in drug molecules because of their complex

study nicely highlights that enzymatic transformations can be employed in conjunction with chemical transformations to enhance the structural diversity obtainable in DELs. Given the power of DELs to accelerate the drug discovery process, and the plethora of enzymes available to the scientific community, it seems highly likely that more examples of biocatalysis in DELs will arrive in the coming years.
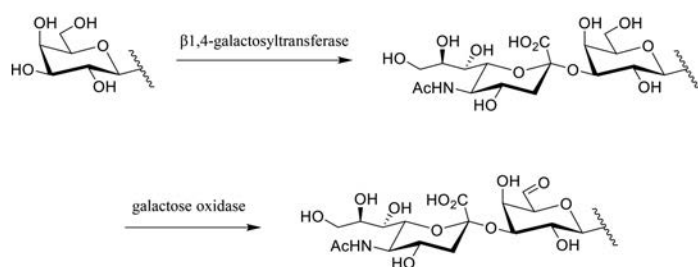
Scheme 6. Reaction scheme highlighting the use of biocatalysis in the generation of carbohydrate-based DELs.

5. Biocatalysis Beyond Small Molecule APIs

For most pharmaceutical companies small molecule drugs still constitute the major part of their portfolio, and this is expected to remain the case for years to come. Consequently, much of this review has focused on the use of biocatalysis in the manufacture of small molecule APIs. In addition to small molecule drugs, there is growing interest in biological-based drugs, including monoclonal antibodies (mAbs), antibody–drug conjugates (ADCs), therapeutic oligonucleotides, cell therapies and gene therapies. Typically, the production costs for biologics are high and development is often hindered by challenges in homogeneous sample preparation. In some cases, biocatalysis may provide a solution, and although in its infancy compared with the biocatalytic repertoire on offer for small molecule manufacture, examples highlighting the potential of enzymes in the preparation of ADCs and therapeutic oligonucleotides are discussed in the following sections.

5.1 Site-selective Conjugation in ADC Manufacture

The excitement surrounding ADCs as a therapeutic modality stems from the fact that they combine the targeted specificity of mAbs with the cytotoxic potency of small molecule drugs.[54] The benefit of this combination is exemplified by Kadcyla (trastuzumab-DM1), a conjugation of the human epidermal growth factor receptor 2 (HER2) targeting mAb trastuzumab and DM1, an antimitotic agent derived from a maytansinoid natural product, used in the treatment of breast cancer.[56] Trastuzumab-DM1 widened the therapeutic window of DM1 and exhibited increased efficacy compared with the unconjugated mAb.[56] The drug to antibody ratio (DAR) of ADCs is a key parameter in defining the therapeutic window: Too little drug and the ADC will lack potency; too high and the payload may perturb the properties of the antibody.[57] In general, a DAR between 2 and 4 is considered optimal.[55,57] Traditional coupling strategies rely on stochastic, unspecific chemical modification of surface exposed lysine and/or cysteine residues, which results in heterogeneous mixtures with a broad distribution of DARs. In contrast, enzyme based strategies have the power to create homogeneous ADCs with precise control over the location and distribution of the chemical warhead on the antibody scaffold, allowing scientists to better tune the therapeutic window (Fig. 5). Multiple strategies towards biocatalytic conjugation exist[57,58] and a few key approaches are highlighted in this review.

One of the most frequently investigated enzymes for this task is Sortase A (SrtA), a transpeptidase, which attaches surface proteins to the peptidoglycan cell wall of gram-positive bacteria. SrtA hydrolyses the peptide bond between a threonine and glycine of a C-terminal LPXTG (where X can be any amino acid) tag, creating an acyl-enzyme intermediate. Nucleophilic attack of this intermediate by the N-terminal amine of an incoming peptide regenerates a native peptide bond.[59] SrtA-mediated ligation is most effective when the incoming nucleophile is fronted by a di- or tri-glycine motif.[59] However, the transpeptidase reaction is limited by slow reaction kinetics and a reversible reaction (the glycine leaving group can function as a competing nucleophile) requiring a large excess of the desired nucleophile.[58]

A transpeptidase offering vastly improved reaction kinetics is butelase-1. Butelase-1 efficiently catalyses inter- and intramolecular peptide ligation with a catalytic efficiency up to 3'000 fold higher than SrtA, generating product with 95% conversion with enzyme loadings <1 wt%, requiring only the simple, three amino acid recognition element, NHV, making it an attractive enzyme for bioconjugation.[60] Despite the promising characteristics of butelase-1, application of this enzyme was hindered by limited availability as it needed to be isolated from plants. Recently, scientists achieved to produce this enzyme recombinantly.[61] OaAEP1, a related enzyme that was recently described offers a similar catalytic efficiency to butelase-1 and can be readily produced in *E. coli* and so arises as a promising alternative.[62]

Bacterial transglutaminases, which catalyse the γ-acyl-transfer of glutamine to the ε-amino group of lysine, are another promising class of enzyme with the potential for biocatalytic ADC manufacture.[57,58] Scientists at Rinat/Pfizer identified a 'glutamine tag or Q tag', which is defined by the amino acid sequence LLQGA and permits site-selective modification by a transglutaminase from *Streptomyces mobaraensis* (mTG).[63] This combination was used to introduce payloads to the C-terminus of the heavy chain, the C-terminus of the light chain and both the C-termini of the heavy and light chains of an antibody.[63]

In an alternative approach, formylglycine-generating enzymes (FGE) convert cysteine residues to formylglycine, introducing a unique aldehyde moiety, which can function as a bio-orthogonal handle for coupling reactions.[57] Introduction of the FGE recognition motif, CXPXR into different regions of the antibody sequence, allows the site-selective modification and thus drug conjugation at different locations on an antibody scaffold.[64]

The examples herewith demonstrate that biocatalysis presents multiple solutions for site-selective bioconjugation, and these approaches offer the same potential benefits in terms of environmental sustainability and cost-effective manufacturing as for the synthesis of small molecule APIs. However, these are not the only benefits of biocatalysis in ADCs preparation. The orthogonality of enzyme reactions makes it possible to conjugate multiple different payloads to the same mAb adding flexibility to ADCs design. Furthermore, by co-expressing the enzyme alongside the

Fig. 5. Cartoon highlighting different strategies to make ADCs. Nonspecific, chemical conjugation of a drug molecule to a mAb results in a heterogeneous population of ADCs with a broad distribution of DARs. In contrast, enzymes enable site selective conjugation with precise control of the DAR giving rise to homogenous population of ADCs, simplified downstream processing.
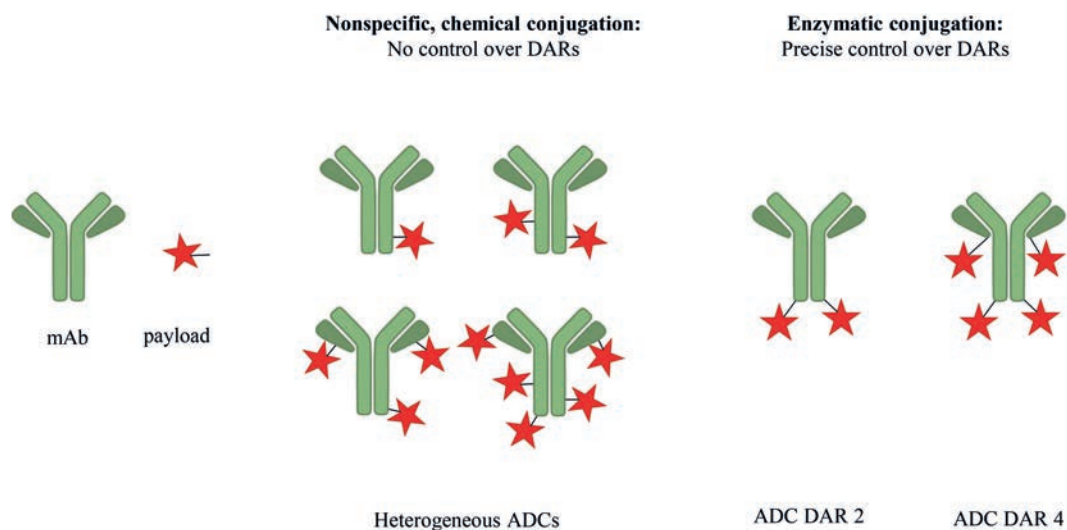
biologic or by adding the enzyme to a crude preparation of the biologic (*e.g.* cell lysate or secretion media) it should be possible to perform the conjugation prior to the isolation of the biologic, resulting in a conjugated biologic from a single purification, which saves time, effort and cost.

5.2 *Enzymatic Assembly of Oligonucleotide Therapeutics*

Oligonucleotide therapeutics is an umbrella term for synthetic polymers of nucleic acids encompassing antisense oligonucleotides (ASOs), small interfering RNA (siRNAs), microRNAs (miRNAs) and gapmers amongst others.[65] These state-of-the-art molecules modulate disease pathways with unrivalled precision by targeting cellular DNA and RNA. Early generations of synthetic oligonucleotides suffered from poor pharmacokinetics and pharmacodynamics, but these properties have been significantly improved, thanks in part to advances in chemical modification strategies of synthetic nucleotides.[66] Some of the chemically modified nucleotides routinely used in therapeutic oligonucleotides can be seen in Fig. 6. 2'-O-MOE, the last example displayed in Fig. 6a, was the focus of some of Oreste Ghisalba's biocatalytic work. 2'-O-methoxyethyl ribonucleosides were synthesised using a short chemo-enzymatic strategy.[66] Shortly, adenosine triphosphate (ATP) is cyclised to cyclic adenosine monophosphate (cAMP) using adenylate cyclase. A chemical alkylation results in the O-methoxylation of cAMP to yield 2'-O-cAMP. Marais and Ghisalba identified a phosphodiesterase from *Serratia marcescens* that is able to perform the decyclisation step of 2'-O-cAMP to 2'-O-methoxyethyl adenosine 5'-monophosphate (2'-O-MOE AMP).[67] The phosphate moiety of the modified nucleotide can then be cleaved off using a phosphatase to form the modified nucleoside (2'-O-methoxyethyl adenosine). This chemo-enzymatic synthesis leads to the formation of the building block 2'-O-MOE used in amongst others, antisense oliogonucleotides. Exemplars of oligonucleotide therapeutics are nusinersen (Spinraza), an antisense oligonucleotide approved for the treatment of spinal muscular atrophy[68] and givosiran (Givlaari), a GalNAc-modified siRNA, which has been approved for the treatment of acute hepatic porphyria (Fig. 6).[69]

The encouraging potential of these molecules necessitates the need for the development of an economical and sustainable synthesis. The present synthesis is divided into three steps: 1) A nucleotide is fixed to a solid support; 2) nucleotides are added sequentially in a cycle of chemical steps, first a coupling step and then the growing chain is deprotected ready for the next cycle; 3) the final oligonucleotide is then released from the solid support.[70]
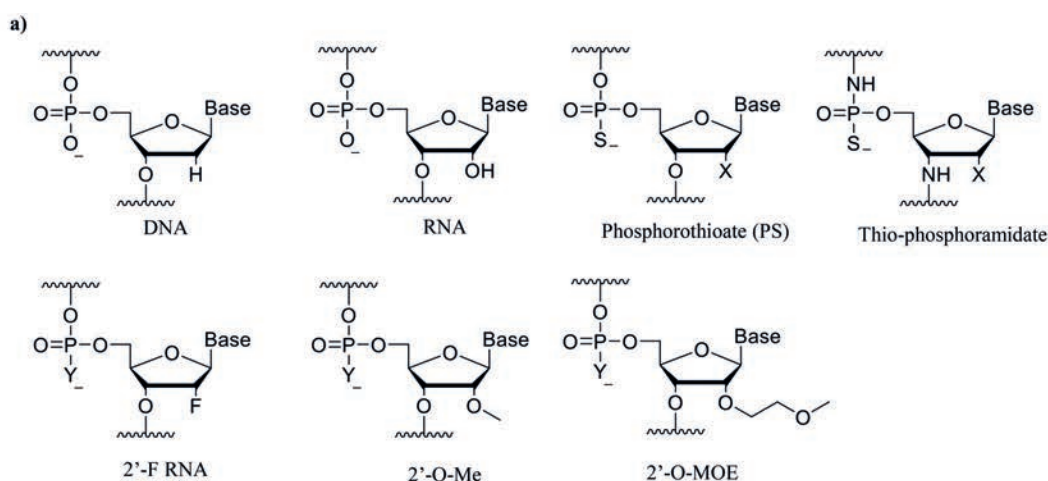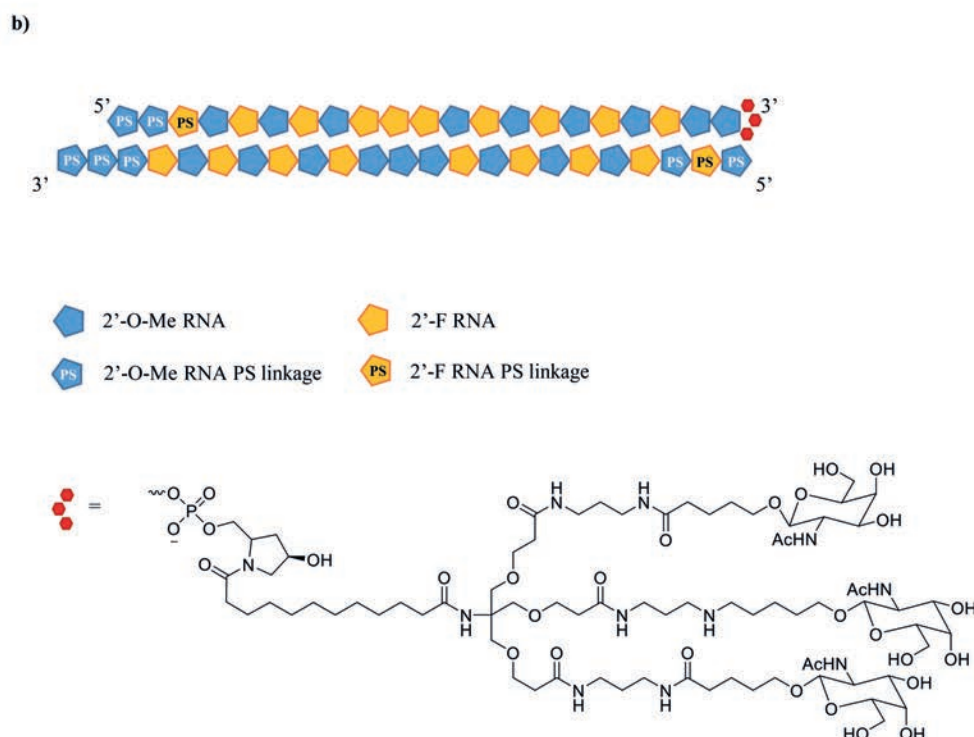

Fig. 6. a) Chemical structures of chemically modified nucleotide analogues, routinely employed in the synthesis of therapeutic oligonucleotides where X = H, OH, F, 2'-O-Me, 2'-O-MOE and Y = O, S.
b) Cartoon showing the structure of givosiran.

This synthesis method is well established but suffers from several limitations, particularly when large quantities need to be synthesised, as is necessary for oligonucleotide therapeutics. Errors in the coupling reaction, while small (typically <1%) accumulate as the length of the oligonucleotide increases. For example, assuming the yield of each coupling reaction is 99% then the overall yield for a 20mer would be 83%. Consequently, the crude oligonucleotide that is released from the resin requires purification *via* column chromatography, which is a costly and time-consuming process involving large volumes of solvent.

In routine molecular biology, a wide range of enzymes is used to manipulate DNA, albeit usually at a small scale. It makes sense therefore that an enzymatic synthesis of oligonucleotides is feasible. DNA polymerase may seem like an obvious candidate; however, it requires a primer of approximately 15 complimentary nucleotides to initiate polymerisation, which makes it unsuitable in the synthesis of therapeutic oligonucleotides, which typically consist of 20–25 nucleotides. Furthermore, the polymerisation reaction occurs in a pool of all the nucleotides that are to be incorporated and while DNA polymerase can align the appropriate nucleotide *via* Watson-Crick base pairing, it is unable to discriminate between deoxyribose and ribose modifications, which are common features of many therapeutic oligonucleotides. As an alternative, scientists at GSK disclosed a patented synthesis of therapeutic oligonucleotides using DNA ligase.[70] Their process starts from short oligonucleotide fragments, which can be synthesised accurately and with high yields. The fragments are assembled in the correct order using a biotinylated template strand of DNA and the fragments are ligated together by T4 DNA ligase. By immobilising the template strand hybridised to the ligated product using streptavidin-coated beads, it is possible to wash away impurities. A simple change in temperature releases the ligated product from the template strand providing an easy separation of the product.[70] Crucially, the invention describes the use of an evolved DNA ligase, which accepts chemically modified nucleotides as substrates, either side of the ligation junction,[70] making it a suitable biocatalyst for the synthesis of a variety of complex therapeutic oligonucleotides.

6. Concluding Remarks

Biocatalysis in the pharmaceutical industry has progressed enormously in the last 10–15 years, transforming from a fringe technology with limited application to one which is routinely applied in the synthesis of APIs, replacing established chemical methods. This growth in the application of biocatalysis mirrors the advancement of biotechnology, particularly in the field of protein engineering, which enables the development of biocatalysts that are optimised for the desired manufacturing process. Importantly, modern protein engineering techniques combining both directed evolution and computationally guided rational design are robust and rapid enough to be compatible with the time pressure of drug development. Further improvements in protein engineering technology will likely engender increased use of biocatalysis throughout the pharmaceutical industry.

The benefits of incorporating biocatalysis are numerous: 1) Enzymes offer precise control over the enantioselectivity and diastereoselectivity of a reaction. 2) Enzymes can simplify the manufacturing process by shortening synthetic routes and minimising the number of purification steps. 3) Biocatalysis confers environmental sustainability to the manufacturing process by replacing stoichiometric reactions with catalytic reactions, reducing organic solvent waste and removal of rare and toxic metals. 4) Enzymes operate under mild conditions, reducing safety hazards associated with many chemical approaches. Furthermore, this preference for aqueous media and mild reactions make them suitable as catalysts of transformations on delicate, water-soluble scaffolds, such as peptides and DNA tagged molecules in DELs. 5) Enzymes are capable of novel chemistry allowing exploration of chemical space that was previously not accessible to synthetic chemists.

These benefits have been recognised by the pharmaceutical industry and there is a consensus to consider biocatalysis, whenever possible, as the primary means of transformation, rather than as an afterthought. Moreover, there is much interest in the development of enzymatic cascade reactions, which have the potential to maximise the impact of biocatalysis. Of course, this commitment to redefine chemical synthesis does require investment in biotechnology; however, the burgeoning evidence coming from the pharmaceutical industry is that this investment is easily recompensed. After all, the application of enzyme catalysts in the synthesis of intricate pharmaceutical compounds makes bio-*logical* sense!

[1]   R. A. Sheldon, D. Brady, *ChemSusChem* **2019**, *12*, 2859, DOI: 10.1002/cssc.201900351.

[2]   J. J. M. Leisola, O. Pastinen, O. Turunen, H. E. Shoemaker, in 'Medical and Health Sciences', Vol. 2, Eds. M. A. Osmo, O. P. Hanninen, B. P. Mansourian, A. Wojtezak, S. M. Mahfouz, H. Majewski, E. Elisabetsky, N. L. Etkin, R. Kirby, T. G. Downing, M. I. El Gohary, EOLSS Publications Oxford, UK, **2010**, p. 1.

[3]   P. Fernandes, *Enzyme Res.* **2010**, *2010*, 862537, DOI: 10.4061/2010/862537.

[4]   H. S. Olsen, P. Falholt, *J. Surfact. Deterg.* **1998**, *1*, DOI: 10.1007/s11743-998-0058-7.

[5]   C. Schmidt-Dannert, F. H. Arnold, *Trends Biotechnol.* **1999**, *17*, 135, DOI: 10.1016/s0167-7799(98)01283-9.

[6]   P. D. de Gonzalo, A. R. Alcantara, *Catalysts* **2019**, *9*, DOI: ARTN 80210.3390/catal9100802.

[7]   U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185, DOI: 10.1038/nature11117.

[8]   R. N. Patel, *Bioorg. Med. Chem.* **2018**, *26*, 1252, DOI: 10.1016/j.bmc.2017.05.023.

[9]   R. Wohlgemuth, *Curr. Opin. Biotechnol.* **2010**, *21*, 713, DOI: 10.1016/j.copbio.2010.09.016.

[10] L. Ufarte, G. Potocki-Veronese, E. Laville, *Front. Microbiol.* **2015**, *6*, 563, DOI: 10.3389/fmicb.2015.00563.

[11] E. Ladoukakis, F. N. Kolisis, A. A. Chatziioannou, *Front. Cell Dev. Biol.* **2014**, *2*, 70, DOI: 10.3389/fcell.2014.00070.

[12] J. Liang, Y. Z. Luo, H. M. Zhao, *Wires Syst. Biol. Med.* **2011**, *3*, 7, DOI: 10.1002/wsbm.104.

[13] a) M. P. Jacobson, C. Kalyanaraman, S. W. Zhao, B. X. Tian, *Trends Biochem. Sci.* **2014**, *39*, 363, DOI: 10.1016/j.tibs.2014.05.006; b) R. Lonsdale, J. N. Harvey, A. J. Mulholland, *Chem. Soc. Rev.* **2012**, *41*, 3025, DOI: 10.1039/c2cs15297e.

[14] M. Dorr, M. P. C. Fibinger, D. Last, S. Schmidt, J. Santos-Aberturas, D. Bottcher, A. Hummel, C. Vickers, M. Voss, U. T. Bornscheuer, *Biotechnol. Bioeng.* **2016**, *113*, 1421, DOI: 10.1002/bit.25925.

[15] J. D. Bloom, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2009**, *106 Suppl 1*, 9995, DOI: 10.1073/pnas.0901522106.

[16] R. J. Fox, S. C. Davis, E. C. Mundorff, L. M. Newman, V. Gavrilovic, S. K. Ma, L. M. Chung, C. Ching, S. Tam, S. Muley, J. Grate, J. Gruber, J. C. Whitman, R. A. Sheldon, G. W. Huisman, *Nat. Biotechnol.* **2007**, *25*, 338, DOI: 10.1038/nbt1286.

[17] a) M. D. Truppo, *ACS Med. Chem. Lett.* **2017**, *8*, 476, DOI: 10.1021/acsmedchemlett.7b00114; b) A. A. Desai, *Angew. Chem. Int. Ed.* **2011**, *50*, 1974, DOI: 10.1002/anie.201007051.

[18] K. B. Hansen, Y. Hsiao, F. Xu, N. Rivera, A. Clausen, M. Kubryk, S. Krska, T. Rosner, B. Simmons, J. Balsells, N. Ikemoto, Y. Sun, F. Spindler, C. Malan, E. J. Grabowski, J. D. Armstrong, 3rd, *J. Am. Chem. Soc.* **2009**, *131*, 8798, DOI: 10.1021/ja902462q.

[19] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305, DOI: 10.1126/science.1188934.

[20] L. J. Guan, J. Ohtsuka, M. Okai, T. Miyakawa, T. Mase, Y. Zhi, F. Hou, N. Ito, A. Iwasaki, Y. Yasohara, M. Tanokura, *Sci. Rep.* **2015**, *5*, 10753, DOI: 10.1038/srep10753.

[21] a) W. A. Greenberg, A. Varvak, S. R. Hanson, K. Wong, H. J. Huang, P. Chen, M. J. Burk, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5788, DOI: 10.1073/pnas.0307563101; b) S. K. Ma, J. Gruber, C. Davis, L. Newman, D. Gray, A. Wang, J. Grate, G. W. Huisman, R. A. Sheldon, *Green Chem.* **2010**, *12*, 81, DOI: 10.1039/b919115c.

[22] H. J. M. Gijsen, C. H. Wong, *J. Am. Chem. Soc.* **1994**, *116*, 8422, DOI: DOI 10.1021/ja00097a082.

[23] P. Anastas, N. Eghbali, *Chem. Soc. Rev*. **2010**, *39*, 301, DOI: 10.1039/b918763b.

[24] a) S. Novak, Personal Communication, Codexis, Inc; b) F. Kleinbeck, Personal Communication, Novartis Pharma AG.

[25] a) Codexis Inc., 'Codexis Announces CodeEvolver Thechnology Transfer and License Agreement with Global Pharmaceutical Leader', Codexis Press Release 15/05/**2019**, DOI: https://ir.codexis.com/news-releases/news-release-details/codexis-announces-codeevolverr-technology-transfer-and-license; b) Codexis Inc., 'Codexis Completes CodeEvolver Technology Transfer to GSK; Earns $7.5 Million Milestone Payment', Codexis Press Release 02/05/**2016**; c) Codexis Inc., 'Codexis Completes Transfer of CodeEvolver Technology to Merck', Codexis Press Release 05/10/**2016**, DOI: https://ir.codexis.com/news-releases/news-release-details/codexis-completes-transfer-codeevolverr-technology-merck.

[26] M. Schober, C. MacDermaid, A. A. Ollis, S. Chang, D. Khan, J. Hosford, J. Latham, L. A. F. Ihnken, M. J. B. Brown, D. Fuerst, M. J. Sanganee, G. D. Roiban, *Nat. Catal.* **2019**, *2*, 909, DOI: 10.1038/s41929-019-0341-4.

[27] C. Molinaro, P. G. Bulger, E. E. Lee, B. Kosjek, S. Lau, D. Gauvreau, M. E. Howard, D. J. Wallace, P. D. O'Shea, *J. Org. Chem.* **2012**, *77*, 2299, DOI: 10.1021/jo202620r.

[28] E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv. Synth. Catal.* **2011**, *353*, 2239, DOI: 10.1002/adsc.201100256.

[29] M. A. Huffman, A. Fryszkowska, O. Alvizo, M. Borra-Garske, K. R. Campos, K. A. Canada, P. N. Devine, D. Duan, J. H. Forstater, S. T. Grosser, H. M. Halsey, G. J. Hughes, J. Jo, L. A. Joyce, J. N. Kolev, J. Liang, K. M. Maloney, B. F. Mann, N. M. Marshall, M. McLaughlin, J. C. Moore, G. S. Murphy, C. C. Nawrat, J. Nazor, S. Novick, N. R. Patel, A. Rodriguez-Granillo, S. A. Robaire, E. C. Sherer, M. D. Truppo, A. M. Whittaker, D. Verma, L. Xiao, Y. Xu, H. Yang, *Science* **2019**, *366*, 1255, DOI: 10.1126/science.aay8484.

[30] K. Laumen, O. Ghisalba, K. Auer, *Biosci. Biotechnol. Biochem.* **2001**, *65*, 1977, DOI: 10.1271/bbb.65.1977.

[31] a) D. Ghislieri, N. J. Turner, *Top. Catal.* **2014**, *57*, 284, DOI: 10.1007/s11244-013-0184-1; b) E. E. Ferrandi, D. Monti, *World J. Microbiol. Biotechnol.* **2017**, *34*, 13, DOI: 10.1007/s11274-017-2395-2.

[32] C. You, T. Shi, Y. Li, P. Han, X. Zhou, Y. P. Zhang, *Biotechnol. Bioeng.* **2017**, *114*, 1855, DOI: 10.1002/bit.26314.

[33] R. D. Schmid, V. B. Urlacher, 'Modern Biooxidation: Enzymes, Reactions and Applications', Wiley VCH Verlag GmbH & Co. KGaA, **2007**.

[34] R. Gasser, in 'Molecular and Applied Aspects of Oxidative Drug Metabolizing Enzymes', Vol. 303, Eds. E. Arinç, J. B. Schenkman, Springer, Boston, MA, **1999**, pp 183.

[35] N. C. Goodwin, J. P. Morrison, D. E. Fuerst, T. Hadi, *ACS Med. Chem. Lett.* **2019**, *10*, 1363, DOI: 10.1021/acsmedchemlett.9b00410.

[36] J. Schwarz, K. Rosenthal, R. Snajdrova, M. Kittelmann, S. Lütz, *Chimia* **2020**, *74*, 368.

[37] a) J. M. Mason, *Future Med. Chem.* **2010**, *2*, 1813, DOI: 10.4155/fmc.10.259; b) M. Gao, K. Cheng, H. Yin, *Biopolymers* **2015**, *104*, 310, DOI: 10.1002/bip.22625.

[38] D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.* **2013**, *81*, 136, DOI: 10.1111/cbdd.12055.

[39] a) A. T. Bockus, C. M. McEwen, R. S. Lokey, *Curr. Top. Med Chem.* **2013**, *13*, 821, DOI: 10.2174/1568026611313070005; b) W. M. Hewitt, S. S. Leung, C. R. Pye, A. R. Ponkey, M. Bednarek, M. P. Jacobson, R. S. Lokey, *J. Am. Chem. Soc.* **2015**, *137*, 715, DOI: 10.1021/ja508766b.

[40] L. Otvos, Jr., J. D. Wade, *Front Chem.* **2014**, *2*, 62, DOI: 10.3389/fchem.2014.00062.

[41] D. P. Levine, *Clin. Infect. Dis.* **2006**, *42*, S5, DOI: Doi 10.1086/491709.

[42] S. Ho, N. Clipstone, L. Timmermann, J. Northrop, I. Graef, D. Fiorentino, J. Nourse, G. R. Crabtree, *Clin. Immunol. Immunopathol.* **1996**, *80*, S40, DOI: 10.1006/clin.1996.0140.

[43] G. Moldenhauer, A. V. Salnikov, S. Luttgau, I. Herr, J. Anderl, H. Faulstich, *J. Natl. Cancer I* **2012**, *104*, 622, DOI: 10.1093/jnci/djs140.

[44] R. D. Sussmuth, A. Mainz, *Angew. Chem. Int. Ed.* **2017**, *56*, 3770, DOI: 10.1002/anie.201609079.

[45] W. van der Donk, *Biopolymers* **2013**, *100*, 229.

[46] a) G. Mann, L. Huo, S. Adam, B. Nardone, J. Vendome, N. J. Westwood, R. Muller, J. Koehnke, *Chembiochem* **2016**, *17*, 2286, DOI: 10.1002/cbic.201600406; b) J. Koehnke, A. Bent, W. E. Houssen, D. Zollman, F. Morawitz, S. Shirran, J. Vendome, A. F. Nneoyiegbe, L. Trembleau, C. H. Botting, M. C. M. Smith, M. Jaspars, J. H. Naismith, *Nat. Struct. Mol. Biol.* **2012**, *19*, 767, DOI: 10.1038/nsmb.2340; c) C. M. Czekster, H. Ludewig, S. A. McMahon, J. H. Naismith, *Nat. Commun.* **2017**, *8*, 1045, DOI: 10.1038/s41467-017-00862-4; d) H. Ludewig, C. M. Czekster, E. Oueis, E. S. Munday, M. Arshad, S. A. Synowsky, A. F. Bent, J. H. Naismith, *ACS Chem. Biol.* **2018**, *13*, 801, DOI: 10.1021/acschembio.8b00050; e) J. A. McIntosh, M. S. Donia, S. K. Nair, E. W. Schmidt, *J. Am. Chem. Soc.* **2011**, *133*, 13698, DOI: 10.1021/ja205458h; f) A. Parajuli, D. H. Kwak, L. Dalponte, N. Leikoski, T. Galica, U. Umeobika, L. Trembleau, A. Bent, K. Sivonen,

M. Wahlsten, H. Wang, E. Rizzi, G. De Bellis, J. Naismith, M. Jaspars, X. Liu, W. Houssen, D. P. Fewer, *Angew. Chem. Int. Ed.* **2016**, *55*, 3596, DOI: 10.1002/anie.201509920; g) H. Song, N. S. van der Velden, S. L. Shiran, P. Bleiziffer, C. Zach, R. Sieber, A. S. Imani, F. Krausbeck, M. Aebi, M. F. Freeman, S. Riniker, M. Kunzler, J. H. Naismith, *Sci. Adv.* **2018**, *4*, eaat2720, DOI: 10.1126/sciadv.aat2720; h) B. J. Burkhart, C. J. Schwalen, G. Mann, J. H. Naismith, D. A. Mitchell, *Chem. Rev.* **2017**, *117*, 5389, DOI: 10.1021/acs.chemrev.6b00623.

[47] D. Sardar, E. W. Schmidt, *Curr. Opin. Chem. Biol.* **2016**, *31*, 15, DOI: 10.1016/j.cbpa.2015.11.016.

[48] a) J. Koehnke, G. Mann, A. F. Bent, H. Ludewig, S. Shirran, C. Botting, T. Lebl, W. Houssen, M. Jaspars, J. H. Naismith, *Nat. Chem. Biol.* **2015**, *11*, 558, DOI: 10.1038/nchembio.1841; b) G. Mann, University of St Andrews, *http://hdl.handle.net/10023/10826*, **2017**.

[49] a) A. F. Bent, G. Mann, W. E. Houssen, V. Mykhaylyk, R. Duman, L. Thomas, M. Jaspars, A. Wagner, J. H. Naismith, *Acta Crystallogr. D Struct. Biol.* **2016**, *72*, 1174, DOI: 10.1107/S2059798316015850; b) S. Gao, Y. Ge, A. F. Bent, U. Schwarz-Linek, J. H. Naismith, *Biochemistry* **2018**, *57*, 5996, DOI: 10.1021/acs.biochem.8b00835.

[50] N. Favalli, G. Bassi, J. Scheuermann, D. Neri, *Febs Lett.* **2018**, *592*, 2168, DOI: 10.1002/1873-3468.13068.

[51] V. Kunig, M. Potowski, A. Gohla, A. Brunschweiger, *Biol. Chem.* **2018**, *399*, 691, DOI: 10.1515/hsz-2018-0119.

[52] P. Dickson, T. Kodadek, *Org. Biomol. Chem.* **2019**, *17*, 4676, DOI: 10.1039/c9ob00581a.

[53] M. L. Malone, B. M. Paegel, *ACS Comb. Sci.* **2016**, *18*, 182, DOI: 10.1021/acscombsci.5b00198.

[54] B. Thomas, X. Lu, W. R. Birmingham, K. Huang, P. Both, J. E. Reyes Martinez, R. J. Young, C. P. Davie, S. L. Flitsch, *Chembiochem* **2017**, *18*, 858, DOI: 10.1002/cbic.201600678.

[55] N. Diamantis, U. Banerji, *Br. J. Cancer* **2016**, *114*, 362, DOI: 10.1038/bjc.2015.435.

[56] T. T. Junttila, G. M. Li, K. Parsons, G. L. Phillips, M. X. Sliwkowski, *Breast Cancer Res. Tr.* **2011**, *128*, 347, DOI: 10.1007/s10549-010-1090-x.

[57] G. Falck, K. M. Muller, *Antibodies (Basel)* **2018**, *7*, DOI: 10.3390/antib7010004.

[58] T. Kline, A. R. Steiner, K. Penta, A. K. Sato, T. J. Hallam, G. Yin, *Pharm. Res-Dordr.* **2015**, *32*, 3480, DOI: 10.1007/s11095-014-1596-8.

[59] D. A. Levary, R. Parthasarathy, E. T. Boder, M. E. Ackerman, *PLoS One* **2011**, *6*, e18342, DOI: 10.1371/journal.pone.0018342.

[60] G. K. T. Nguyen, S. J. Wang, Y. B. Qiu, X. Hemu, Y. L. Lian, J. P. Tam, *Nat. Chem. Biol.* **2014**, *10*, 732, DOI: 10.1038/Nchembio.1586.

[61] a) A. M. James, J. Haywood, J. Leroux, K. Ignasiak, A. G. Elliott, J. W. Schmidberger, M. F. Fisher, S. G. Nonis, R. Fenske, C. S. Bond, J. S. Mylne, *Plant J.* **2019**, *98*, 988, DOI: 10.1111/tpj.14293; b) N. Pi, M. Gao, X. Y. Cheng, H. L. Liu, Z. K. Kuang, Z. X. Yang, J. Yang, B. L. Zhang, Y. Chen, S. Liu, Y. Q. Huang, Z. D. Su, *Biochemistry* **2019**, *58*, 3005, DOI: 10.1021/acs.biochem.9b00263.

[62] R. Yang, Y. H. Wong, G. K. T. Nguyen, J. P. Tam, J. Lescar, B. Wu, *J. Am. Chem. Soc.* **2017**, *139*, 5351, DOI: 10.1021/jacs.6b12637.

[63] S. E. Farias, P. Strop, K. Delaria, M. G. Casas, M. Dorywalska, D. L. Shelton, J. Pons, A. Rajpal, *Bioconjug. Chem.* **2014**, *25*, 240, DOI: 10.1021/bc4003794.

[64] a) P. M. Drake, A. E. Albers, J. Baker, S. Banas, R. M. Barfield, A. S. Bhat, G. W. de Hart, A. W. Garofalo, P. Holder, L. C. Jones, R. Kudirka, J. McFarland, W. Zmolek, D. Rabuka, *Bioconjug. Chem.* **2014**, *25*, 1331, DOI: 10.1021/bc500189z; b) B. C. B. Huang, Y. C. Kim, S. Banas, R. M. Barfield, P. M. Drake, I. Rupniewski, W. E. Haskins, D. Rabuka, *Mabs-Austin* **2018**, *10*, 1182, DOI: 10.1080/19420862.2018.1512327.

[65] W. Yin, M. Rogge, *Clin. Transl. Sci.* **2019**, *12*, 98, DOI: 10.1111/cts.12624.

[66] C. I. E. Smith, R. Zain, *Annu. Rev. Pharmacol.* **2019**, *59*, 605, DOI: 10.1146/annurev-pharmtox-010818-021050.

[67] G. Marais, O. Ghisalba, *Appl. Microbiol. Biotechnol.* **2005**, *66*, 512, DOI: 10.1007/s00253-004-1718-z.

[68] E. E. Neil, E. K. Bisaccia, *J. Pediatr. Pharmacol. Ther.* **2019**, *24*, 194, DOI: 10.5863/1551-6776-24.3.194.

[69] L. J. Scott, *Drugs* **2020**, *80*, 335, DOI: 10.1007/s40265-020-01269-0.

[70] A. Crameri, M. L. Hill, S. L. Lovelock, M. Schober, D. G. Tew, P. J. Thomas, WO Patent Appl. No. WO2018011067A3, **2018**.