Enzyme Fusions in Biocatalysis: Coupling Reactions by Pairing Enzymes

Friso S. Aalbers and Marco W. Fraaije*[a]

One approach to bringing enzymes together for multienzyme biocatalysis is genetic fusion. This enables the production of multifunctional enzymes that can be used for whole-cell bio-transformations or for in vitro (cascade) reactions. In some cases and in some aspects, such as expression and conversions, the fused enzymes outperform a combination of the individual enzymes. In contrast, some enzyme fusions are greatly compromised in activity and/or expression. In this Minireview, we give an overview of studies on fusions between two or more enzymes that were used for biocatalytic applications, with a focus on oxidative enzymes. Typically, the enzymes are paired to facilitate cofactor recycling or cosubstrate supply. In addition, different linker designs are briefly discussed. Although enzyme fusion is a promising tool for some biocatalytic applications, future studies could benefit from integrating the findings of previous studies in order to improve reliability and effectiveness.

1. Introduction

Finding new and efficient ways to synthesize specific compounds is a central endeavor for both academic and industrial chemists. One particular approach involves the combination of multiple enzymes that catalyze highly selective and sequential reactions in one pot.[1] This multienzyme biocatalysis approach is attracting a steady increase in interest.[1] Biocatalytic processes are nontoxic, can be operated under mild aqueous conditions, and can potentially be applied on an industrial scale.[2] The characterization of more and more enzymes over the years, with diverse substrate scopes, has enabled researchers to devise novel and elegant reaction combinations. In addition, tailoring and optimizing biocatalysts through enzyme engineering is becoming faster and more efficient.[3] The multienzyme strategy is deemed to be a critical stepping stone on the way to the large-scale industrial application of biocatalysis.

One advancing aspect of multienzyme biocatalysis is the approach of fusing enzymes.[1e,f] Primarily, enzyme fusion is a convenient way of producing two or more enzymes in one cell, thereby allowing either whole-cell biotransformations or co-purification. By combining enzymes in this fashion, a single multifunctional biocatalyst can be produced that can catalyze a cascade of reactions. Remarkably, in a number of studies, evidence was found that the covalently tethered enzymes outper-

form the classic combination of the separate enzymes in some features such as enhanced expression,[4,5] higher conversions of a cascade reaction,[4,6–9] improved stability,[9a,c,10] and improved catalytic activity.[9c,10–12] However, these promising observations are not consistent across studies, and generic methods to predictably design an effective bifunctional biocatalyst are still lacking. Moreover, some studies find a decrease in activity as a consequence of enzyme fusion, thus showing that fusing enzymes can be delicate.

Enzyme fusions can be made by genetically placing multiple genes together without a stop codon between them. Translation would then lead to a single polypeptide with multiple enzymes covalently fused. Alternatively, it is possible to create enzyme fusions post-translationally by using pairs of tags or scaffolds, of which there are several elegant examples,[13] although this type of approach is beyond the scope of this review. To ensure the proper folding and protein stability of fused enzymes, it is common to introduce a sequence coding for a peptide linker between the genes. This linker region can affect the orientation of the two enzymes. Another term often used for an enzyme fusion is a "chimeric enzyme", although this should concern the combination of particular parts/domains of different enzymes, rather than full-length enzymes, as the name suggests.

The relative improvement in catalytic efficiency of a two-step reaction with a fusion enzyme has been attributed to the phenomenon of substrate channeling: direct migration of the product of one enzyme to the active site of the second, which is faster than random diffusion through the bulk solution. A recent review on compartmentalized cascade reactions elaborated on this topic. The authors stated that proximity only provides an advantage if the intermediate concentration remains low, and that substrate channeling is only enabled when the fusion enzymes form larger clusters.[1e] Substrate channeling is

[a] F. S. Aalbers, Prof. Dr. M. W. Fraaije
Molecular Enzymology Group, University of Groningen
Nijenborgh 4, 9747AG Groningen (The Netherlands)
E-mail: m.w.fraaije@rug.nl

ⓘ The ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201800394.

quite beneficial for cascade reactions, as it increases the rate of the overall reaction, and decreases the lifetime of the intermediate; this can be important for reactive and/or labile intermediates.[14] Some researchers thoroughly investigated substrate channeling in their systems, and provided evidence for presence or absence thereof.[14, 15]

In this review, we focus on recent studies on enzyme fusions for biocatalytic applications, and compare the approaches and findings. In particular, studies concerning fusions of two enzymes and/or catalytic domains are selected. This excludes studies about enzymes fused to peptides or protein domains, such as affinity tags, carbohydrate-binding modules, expression-boosting domains, antibodies, and other catalytically inactive proteins, which are reviewed elsewhere.[16a] In addition, we have included only a few studies on carbohydrate-active and biomass-degrading enzyme fusions, as these have also been reviewed elsewhere.[16b]

First, we give an overview of recent studies in which enzyme fusions were created and applied for biocatalytic (cascade) reactions, in particular involving oxidative enzymes (Section 2). General linker design, in terms of flexibility, length and orientation, has been addressed to some extent in previous reviews; however, the degree to which linkers affect activity and expression has not been reviewed recently, to our knowledge. Based on the studies we discuss in Section 2, we provide an overview of the linker design of these fusions and the effects of different linkers (Section 3). Lastly, we point out the knowledge gaps and challenges, and we consider the advantages and disadvantages of the approach of fusion enzyme engineering (Section 4).

2. Enzyme Fusions in Biocatalysis

The first examples of fusion enzymes were found as natural fusions. Various organisms have evolved gene organizations that results in fused enzymes. Some known examples are tryptophan synthase,[17] carbamoyl phosphatase synthase,[18] and cytochrome P450 BM3.[19] That such enzyme fusions have evolved and are frequently encountered, suggests that enzyme fusion can provide an evolutionary advantage for some metabolic functions. In at least two studies, a genetic fusion has evolved: in one case as a result of selection for glycerol production[20a] and in another case through a rapid-evolution method that was used with selection for growth on guaiacol.[20b] Related to this, in one study, two enzymes were fused and produced in vivo to direct metabolic flux towards sesquiterpene production; this shows that enzyme fusion can affect the growth and phenotype of an organism.[20c] However, it is not universally beneficial, as natural fusions are fairly rare, considering the total number of operons that contain multiple genes. Likewise, studies on enzyme fusions found improvements with some combinations of enzymes, but not with all types of enzymes.

Inspired by nature, various researchers began exploring artificial fusions of two enzymes, and compared these bifunctional enzymes to the separate enzymes. One of the first papers describes a fusion of a histidinol dehydrogenase with an amino-

transferase.[21] Another important early study was done on fusing β-galactosidase with a galactose dehydrogenase; the fused enzyme showed improvement in the two-step conversion from lactose to galactone.[22] From this point on, many more enzyme fusions were produced and studied, of which some have been reviewed before.[16, 23] This Minireview highlights some recent developments in fusing enzymes for biocatalytic applications.

2.1. Baeyer–Villiger monooxygenase (BVMO) fusions

BVMOs are an interesting class of biocatalysts as they can catalyze Baeyer–Villiger reactions under mild conditions, transforming ketones to esters or lactones, often in a highly regio- and enantioselective manner. Interestingly, genome-sequence analysis has revealed that some natural fusions exist, such as BVMOs fused to an esterase that can hydrolyze a lactone or ester product from the Baeyer–Villiger reaction.[24] BVMOs contain a flavin cofactor (FAD) as prosthetic group. For catalysis, the FAD cofactor is first reduced by NADPH; this enables the reduced FAD to react with dioxygen and form the oxygenating peroxyflavin intermediate. The reliance on NADPH is one of the greatest challenges when applying BVMOs as biocatalysts. Because it is an expensive cofactor, efficient recycling is needed to make any BVMO-based biocatalytic process feasible. A typical approach to recycling is to use a cofactor-recycling enzyme that can oxidize a sacrificial substrate to transform $NADP^+$ to NADPH. Some typical examples of such enzymes are glucose dehydrogenase (GDH), formate dehydrogenase (FDH), and phosphite dehydrogenase (PTDH).

Pazmiño and co-workers looked into the effects of fusing PTDH from *Pseudonomas* sp. to BVMOs so as to provide direct recycling of NADPH.[25] It was found that PTDH–BVMO fusions enabled good expression of such bifunctional biocatalysts. By adding phosphite and only minor amounts of $NADP^+$, the bifunctional enzymes could catalyze continuous Baeyer–Villiger oxidations (Scheme 1). Three distinct BVMOs were initially selected for probing the fusion approach, with varying substrate scope and enantioselectivity: cyclohexanone monooxygenase (CHMO), phenylacetone monooxygenase (PAMO), and cyclopentanone monooxygenase (CPMO). The bifunctional biocatalysts were found to largely retain the catalytic properties of the original enzymes, such as regio- and stereoselectivity, and kinetic parameters. For PTDH, there was a small increase in K_M for $NADP^+$ and phosphite, whereas the BVMOs showed a small decrease in K_M for the ketone model substrate. With either whole cells or extracts from cells expressing the fusion enzymes, complete conversion was demonstrated for a large set of substrates. Conveniently, these conversions could even be performed without addition of the expensive cofactor $NADP^+$, as the cell extracts contained sufficient nicotinamide cofactor. In a follow-up study, the PTDH fusion partner was optimized for its stability and expression in *Escherichia coli*, and the developed expression vector was shown to be successful for producing a wide range of PTDH–BVMOs.[26]

Still, for a large-scale reaction, the sacrificial substrate can constitute a significant cost contribution, as it also gives a by-
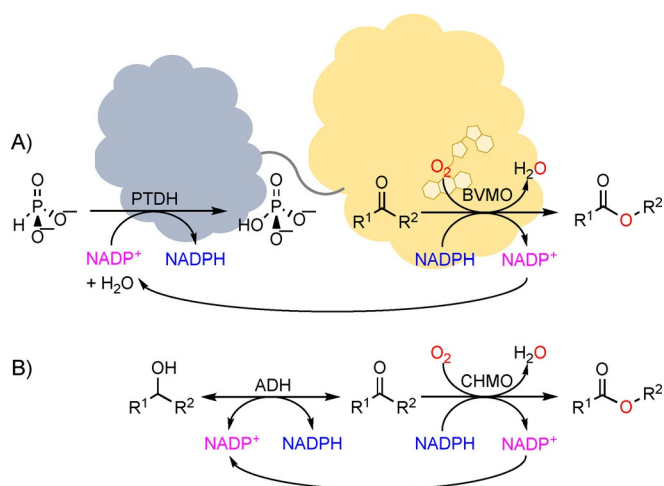
Scheme 1. Dehydrogenase fusion with a BVMO, in which the dehydrogenase provides NADPH that is needed for the BVMO reaction. A) A BVMO can also be paired with PTDH, which uses phosphite as sacrificial substrate to regenerate NADPH.[25,26] B) Fusion of alcohol dehydrogenase (ADH) with a BVMO for a self-sufficient cascade reaction from an alcohol substrate to an ester or lactone.[4,27]

product. An alternative to a sacrificial substrate for regenerating NADPH is a cascade reaction starting from the alcohol rather than the ketone (Scheme 1). An alcohol dehydrogenase can use NADP$^+$ to oxidize the alcohol to a ketone to produce NADPH, and the ketone and NADPH can then be used by the BVMO. Recently, alcohol dehydrogenase (ADH)–BVMO fusions were designed and expressed in *E. coli*, and the whole cells were used in reactions with different hydroxy fatty acid substrates to convert these in a cascade reaction to esters.[18] As comparison, cells co-expressing the separate ADH and BVMO were included. Relative to the latter cells, cells expressing the ADH–BVMO fusion with a glycine-rich linker showed greater cell activity, and higher conversions for all substrates tested. The authors ascribe this improvement primarily to the higher expression of the fusion. Another notable observation was the difference in ratio of normal to abnormal ester product between different linkers, which suggested that linker design can influence this ratio, possibly through slight conformational changes. The authors linked the improved mass balance that was found for the fusions to co-localization of the active sites, by considering the hydrophobicity of the substrate and intermediate. This would be an interesting advantage for such fusion enzymes, and it is reminiscent of the substrate channeling in natural fusion enzymes, for example, tryptophan synthase.[17]

Another study explored the same combination of enzyme classes to facilitate a cascade reaction from cyclohexanol to caprolactone. Three different ADH enzymes were selected, and a thermostable cyclohexanone monooxygenase (*Tm*CHMO) was chosen as BVMO.[27] Inspired by the previous study on ADH–BVMO fusions, a glycine-rich linker was chosen, and all fusions were made in both orientations. Notably, the two short-chain dehydrogenase (SDR) ADHs were found to be inactive as N-terminal fusions (ADH–CHMO) and were also found

to be a monomer, rather than a dimer/tetramer. It is likely that the fusion at the C terminus of SDRs interferes with the oligomerization of the enzymes, which is necessary for activity. Other studies found a similar effect with a C-terminal fusion of other enzymes to an SDR, and with a C-terminal His tag.[6] For the other ADH and for *Tm*CHMO no substantial change in activity was observed upon fusion. The most active fusion was used in a small-scale conversion of 200 mM cyclohexanol, which reached 99% conversion, whereas the separate enzymes reached only 41% conversion. This study showed again that it is difficult to predict whether a protein can be used as a fusion partner. Even if inspection of crystal structures reveal accessible N and C termini, protein–protein interactions could be detrimental to correct folding or oligomerization.

Another interesting exercise in producing fusions with CHMO was recently reported by Peters et al. They attempted to bring together three enzymes: an ADH, an ene-reductase (ER), and a CHMO in one fusion protein.[6] As the group had demonstrated prior to this work, by combining these three enzymes, one can convert unsaturated cyclic alcohols to enantiopure chiral lactone products. Starting from an alcohol, the ADH would bring partial redox balancing, as the ER and CHMO both need the reduced nicotinamide cofactor. Their strategy was first to combine the ER and CHMO, and the authors found that this successfully yielded a bifunctional enzyme. Unfortunately, they found that any fusion including the studied ADHs was not expressed. Still, the cascade reaction catalyzed by the ER–CHMO fusion outperformed the same reaction with separate enzymes.

2.2. Flavin reductase (FR) fusions

Similarly to BVMOs, styrene monooxygenases (SMOs) use a reduced flavin cofactor (FADH$_2$) to activate oxygen, even though they catalyze a different reaction. SMOs bind styrene in such a way that the activated oxygen reacts with the unsaturated bond of styrene to form one enantiomer of styrene oxide (Scheme 2). However, unlike BVMOs, this class of enzymes does not bind FAD tightly, and cannot bind an electron donor like NAD(P)H to reduce FAD. Therefore, SMOs need to receive the reduced flavin from another enzyme: NADH-dependent flavin reductase (FR), which, in many genomes, is either in the same operon (i.e., co-expressed with the SMO) or fused to the SMO.[28] The flavin reductase is often referred to as "styrene monooxygenase reductase" or StyB, and the oxygenase that catalyzes the oxidation is denoted by StyA. Recently, several groups have been artificially fusing FRs to SMOs, and investigating the effects of different linkers.

In 2016, the groups of Gassner and Tischler carefully investigated the effect of fusing the reductase with the epoxidase component.[5] The fusions were evaluated based on their activity on indole, and derivatives thereof, to produce various dyes. Foremost, the authors argue that having the two enzymes fused together is convenient for engineering, expression, and performing oxidations. An important improvement that was observed, compared to the use of separate enzymes, was a higher coupling efficiency, which is the degree to which re-
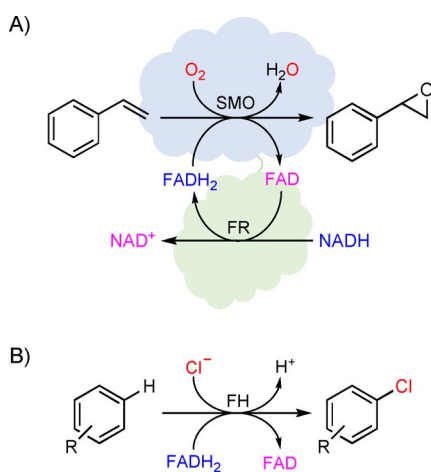
Scheme 2. Concept of fusing a flavin reductase to an $FADH_2$-dependent enzyme. Another enzyme (not shown) facilitates regeneration of NAD^+ to NADH (such as formate dehydrogenase). A) Fusion of SMO (StyA) with FR (StyB) for efficient coupling of a reduced flavin for styrene oxidation.[5, 11] B) Flavin halogenase can also be fused with a FR to couple the two reactions.[7]

duced FAD leads to oxidation versus the formation of hydrogen peroxide (referred to as uncoupling). Through fusion, the enzymes are in a strict 1:1 ratio, which prevents there being excess reductase that could lead to excessive uncoupling. In addition, expression of the reductase domain individually resulted in inclusion bodies, whereas expression of the fusion of the epoxidase and reductase gave soluble and active bifunctional enzymes. However, the epoxidation activity was significantly decreased compared to that of the separate enzymes. In contrast, another group that studied the same enzymes as a fusion found it had a similar activity to the combination of separate StyA/StyB in a 1:1 ratio.[11] This activity was higher than that of naturally fused StyA2B and the fusion enzymes addressed above.[5] The authors suspect that this difference was due to the glycine-rich flexible linker in their SMO; previous SMO fusions had had shorter and more-rigid linkers. Both studies emphasize that the fusion of the two enzymes is advantageous in terms of expression, purification, and applying the enzymes.

Another class of enzymes that relies on an FR for reduced flavin is that of flavin-dependent halogenases (FHs). This extraordinary class of enzymes can catalyze regioselective halogenations of aromatic substrates (Scheme 2). Similar to the FR–SMO fusions addressed above, fusions of various FRs with FHs were studied, with a particular emphasis on whole-cell applications.[7] A decrease in activity was observed in vitro, although when the fusions were expressed for whole-cell biotransformations, higher product titers were reached compared to those in cells co-expressing the separate enzymes. As with the studies above, the authors indicated that the fusion approach simplifies the studying and engineering of the system.

From these examples it seems that, for the combination of a flavin reductase with an enzyme that requires reduced flavin (SMO, FH), enzyme fusion provides notable advantages. This is also a type of fusion that is encountered in several genomes, such as natural FR–SMO fusions and reductase–P450 fusions. In this sense, it is quite important to see that there are examples in which it does not work that well in order to identify the important factors and prerequisites for creating functional fusions, such as linker design (which will be discussed in Section 3).

2.3. Heme-containing enzymes: Cytochrome P450 and peroxidase fusions

In biocatalysis, one extensively studied cytochrome P450 monooxygenase is the natural fusion P450 BM3 from *Bacillus megaterium*.[22] It can catalyze enantio- and regioselective hydroxylations, and it consists of two catalytic domains: the CPR, which is a NADPH-dependent FAD- and FMN-containing reductase domain, and a monooxygenase heme domain. In some studies, various artificial reductase–monooxygenase fusions were made, either to improve stability[29] and/or to improve electron-coupling efficiency.[30] Sabbadin et al. developed a cloning vector that contains the gene coding for the P450 reductase domain RhF-Red, which enables seamless ligation-independent cloning of any P450 heme domain fused to RhF-Red.[31] This vector was used in a study in which 23 genes coding for heme domains from *Rhodococcus jostii* RHA1 were fused to RhF-Red so as to screen the P450 fusions for activity on various pharmaceutical compounds.[32]

Like the afore-mentioned monooxygenases, P450 monooxygenases rely on reduced NAD(P)H for catalysis. In order to establish a biocatalytic process using this class of enzyme, an NAD(P)H-recycling system is needed. As with the BVMOs mentioned above, Beyer and co-workers fused PTDH to P450 BM3 and studied this multifunctional self-sufficient enzyme.[33] Remarkably, the PTDH–P450 fused enzyme showed a slight improvement in activity and in several other aspects performed better than the combination of the two separate enzymes. Even though the fusion is rather large (155 kDa), expression is not compromised. The authors demonstrated the utility of PTDH–P450 by converting several pharmaceutical compounds to hydroxylated drug metabolites.[12,33] A similar study was performed with a heterotrimeric P450 from *Pseudonomas putida*, in which it was fused to PTDH for self-sufficient hydroxylations.[34]

Enzyme fusion can be a great way to generate a cofactor or cosubstrate for a particular enzyme reaction. An example of cosubstrate production is achieved by the fusion of alditol oxidase with a cytochrome P450 (Scheme 3 A). The particular P450 monooxygenase from *Jeotgalicoccus* sp., named OleT$_{JE}$, can act as an peroxygenase and primarily catalyzes the oxidative decarboxylation of fatty acids. In its mode as peroxygenase, the P450 merely requires hydrogen peroxide to perform oxygenations. In order to prevent peroxide-induced enzyme inactivation, it is attractive to have in situ formation of hydrogen peroxide when employing a peroxygenase. By tethering a peroxide-dependent enzyme to an oxidase (generating hydrogen peroxide) as a fusion, reactions can be catalyzed with the hydrogen peroxide produced in situ.[8] The choice of alditol oxidase allowed the use of glycerol, which is very cheap and

A)

B)
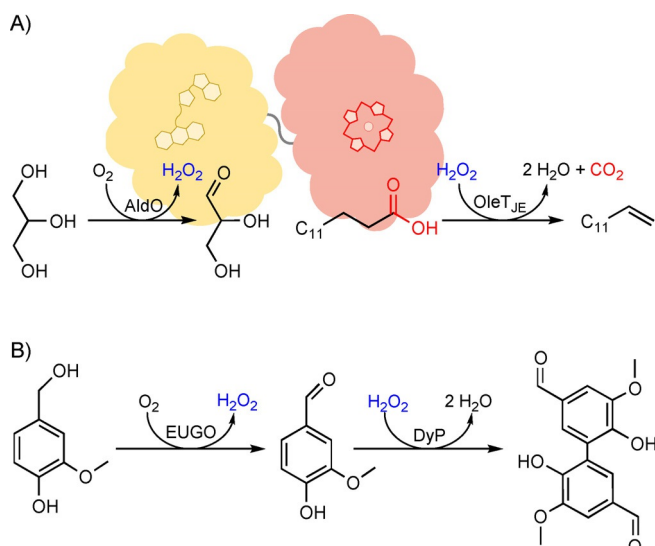
Scheme 3. Oxidase–peroxygenase and oxidase–peroxidase fusions. A) By combining alditol oxidase (AldO) with OleT$_{JE}$, the decarboxylation of myristic acid can be catalyzed with the peroxide formed in situ, at the cost of glycerol.[8] B) Vanillyl alcohol can be converted by eugenol oxidase (EUGO) into vanillin and hydrogen peroxide. The peroxidase DyP can then use the peroxide to create a radical form of vanillin, which can react with a second vanillin radical to form divanillin.[36]

abundant, as sacrificial substrate. The fusion enzyme was able to decarboxylate myristic acid to produce tridec-1-ene with the help of glycerol (Scheme 3 A). To examine the effect of the fusion, the authors designed a linker that can be cleaved with a specific protease, and then directly compared the fusion cleaved and uncleaved. They discovered that the uncleaved fusion enzyme had a higher conversion of myristic acid, thus suggesting that there is a beneficial proximity effect. Interestingly, fusion of OleT$_{JE}$ with the flavin reductase from BM3 produced a self-sufficient decarboxylase, yet with higher activity and broader substrate scope.[35] It relies on oxygen and NADPH rather than hydrogen peroxide, similar to the PTDH–P450 BM3 mentioned above.[33] This study demonstrated that the activity of OleT$_{JE}$ P450 can be altered by different types of electron-donor-producing fusion partners.

An oxidase, which produces hydrogen peroxide, and an enzyme that uses hydrogen peroxide as substrate would form a perfect match. A recent study explored fusions of oxidases with a dye-decolorizing peroxidase (DyP), which is a heme-containing enzyme that uses hydrogen peroxide to oxidize phenolic compounds and dyes (Scheme 3 B).[36] Four oxidases were selected and fused, without a linker, to the bacterial peroxidase SviDyP: alditol oxidase (HotAldO), 5-hydroxymethylfurfural oxidase (HMFO), chitooligosaccharide oxidase (ChitO), and eugenol oxidase (EugO). All four fusions could be expressed and purified with a high degree of heme and flavin incorporation. The oxidase components of the fusions lost some activity; this could be due to the absence of a suitable linker. One challenging aspect of this combination was that most oxidases have a pH optimum around 7–8, whereas DyPs are most active around pH 4–5. Nevertheless, the authors were able to demonstrate successfully several coupled reactions in which the fu-

sions functioned as biosensors or as bifunctional biocatalysts for performing cascade reactions, in which both the hydrogen peroxide and the product from the oxidase were used by the peroxidase.

2.4. Alcohol dehydrogenase and transaminase fusions

Some other conceptually intriguing fusions include a couple of examples with an ADH. As in the examples mentioned in the section on BVMOs, ADHs can perform alcohol oxidations by using NAD(P)$^+$. However, this reaction is thermodynamically less favorable than the reverse reaction. In order to use ADHs for complete alcohol oxidations, it is necessary to recycle NAD(P)$^+$.[37] To enable efficient recycling, an ADH can be combined with an NAD(P)H-oxidase (NOX). Such FAD-containing enzymes can oxidize NAD(P)H and produce water or hydrogen peroxide as by-product. One study describes a fusion between a glycerol dehydrogenase (GlyDH) and a NOX to produce a self-sufficient enzyme that converts glycerol to dihydroxyacetone.[38] However, the activity of the enzymes as fusion was significantly decreased. This was possibly due to the absence of a linker, and/or the orientation of GlyDH-NOX, which could interfere with the quaternary structure of the GlyDH.

The asymmetric reduction of ketones to chiral alcohols can also be catalyzed by alcohol dehydrogenases. This reverse reaction relies on NAD(P)H, like the reactions catalyzed by BVMOs and P450s. A number of studies explored fusions of ADHs with NAD(P)H-regenerating enzymes. Three studies fused formate dehydrogenase with a specific ADH,[9] and in one study glucose dehydrogenase was used.[39] Even though fusing such dehydrogenases with one another can be more challenging, owing to the typical quaternary structures of these enzymes, these studies reported successful examples of active enzyme fusions with alcohol dehydrogenases.

Alcohol dehydrogenases can also serve in cascade reactions with transaminases, as the latter can act on a ketone substrate, which can be produced by an ADH. Fusions of an ADH with an aminotransferase (AT) were designed and used to catalyze a cascade reaction to produce amines from alcohols (Scheme 4).[40] Taking into account the quaternary structures of both enzymes, and to make space for proper folding, three
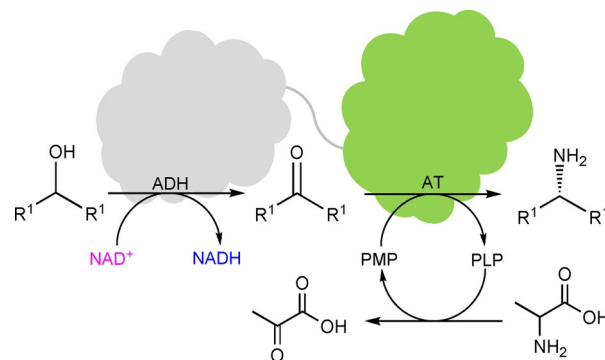
Scheme 4. Fusing an alcohol dehydrogenase with an AT enables a cascade reaction starting from an alcohol to produce a chiral amine, at the expense of NAD$^+$ and alanine, as well as NADH and pyruvate.[40]

linkers of different length (20, 40, and 60 amino acids) were used. The linkers consisted of PAS sequences (combinations of proline, alanine, and serine), which form disordered, uncharged, and highly soluble spacers.[41] Each of the linkers had a particular effect: PAS60 resulted in the highest soluble expression, PAS40 retained most activity, and PAS20 had higher conversion for the coupled reaction, with about twofold higher conversion compared to the separate enzymes. By measuring the initial rate of the second step with different concentrations of the first substrate, it was shown that the fused system displays a beneficial proximity effect. However, only low conversions were found, as the NADH was not recycled to NAD$^+$ in this system. A similar, yet redox-neutral, combination would be the fusion of an ADH with an amine dehydrogenase, using alcohols and ammonia to produce amines.[42]

3. Fusion Design and Linker Design

When fusing two genes to produce a fusion enzyme, there are at least two things to consider: in which order to place the genes and which linker to choose. In some cases, the effect of the arrangement is not drastic, such as for the PTDH–BVMO fusions, for which both enzymes were also active in the BVMO–PTDH orientation.[25] However, other enzymes are inactive and/or fold improperly when another protein is attached to their N or C terminus, as was observed for SDR enzymes with C-terminal fusions.[27] Similar problems were observed when an improper linker was chosen, or in the absence of a linker. In other words, linker design can be critical for the generation of an active fusion enzyme.

Broadly speaking, there are two aspects of the linker to consider: the composition (in terms of flexibility/rigidity and hydrophilicity/hydrophobicity) and the length. Linkers that contain predominantly glycine are flexible, whereas a prevalence of α-helix-associated amino acids, such as alanine and lysine, forms a rigid tether.[43] To investigate the nature of this dichotomy in linkers, Li and co-workers performed simulations and constructed a linker library consisting of various amounts of flexible and rigid parts.[44] By measuring variants from this library by using FRET, which is a perfect indicator of proximity, their simulations could be validated. The work provides guidance on controlling the flexibility of a linker. To study flexibility, another group investigated linkers with various amounts of glycine and of varying length.[45] Their findings showed that length and a low percentage of glycine in the linker correlated with lower FRET frequency. Even though these two studies used protein fusions rather than enzyme fusions, the knowledge of linker design should be generally applicable.

An extensive review on loops and linkers elaborates on linker design in great detail and emphasizes how linkers are not merely "connectors", but have an important effect on the microenvironment and orientation of the fusion.[46] It turns out that the orientation between the two enzymes, as a consequence of the linker, can have a significant impact on the efficiency of the reaction,[47] although, in practice, this is difficult to control.[48] Some studies point out that by performing computational simulations and/or using linker databases, a more

consistent and less arbitrary design can be established,[46,49] as exemplified by the studies mentioned above,[44,45] even though up to now this has hardly been integrated with enzyme fusions. In one study, simulations were used to decide the gene order for a fusion enzyme of formate dehydrogenase with leucine dehydrogenase.[9c]

Although studies on linker design help to narrow down the selection of a linker, there is no clear consensus on both composition and length. It seems to be case dependent, as there is variation in the N and C termini of the enzymes, in dynamics (i.e., conformational movements), and in the tertiary and quaternary structures. The fusions discussed in this review are listed in Table 1, alongside their respective linkers. Strikingly, there is quite some variation in the composition and length of linkers. A commonly used and relatively successful linker is a glycine-rich linker, which forms a disordered loop that, in theory, provides freedom for folding and other conformational movements. The latter in particular might be an underestimated factor because, for some enzymes including BVMOs, conformational changes play an important role in catalysis.[50,51]

Some studies designed and produced fusions with different linkers, and proceed with the most suitable linker based on the initial results. For instance, with the FR–SMO fusions, four different linkers were tested, from which two linkers led to insoluble inclusion bodies. For the two linkers that did show soluble expression (StyAL1B, StyAL2B, Table 1), a decrease in activity compared to the separate enzymes was observed.[5] However, when another group fused the same enzymes (Fus–SMO, Table 1) with a 30-resdue glycine-rich linker, the level of activity was unaffected, and the expression level increased (40 mg L^{-1} compared to 12–15 mg L^{-1} for the StyAL1B and StyAL2B).[11] In a recent study of a natural FR–SMO fusion from *Variovorax paradoxus* EPS, the original linker (AREAV) was replaced with six distinct linkers.[52] For each variant, the expression was lower and some catalytic properties were altered. Remarkably, one linker (AAAAA) displayed a lower K_M for FAD (from 33 to 1.8 μM) and higher monooxygenase activity.

Whereas some studies show different consequences for different linkers, in others the linker hardly makes any difference. In the second study on PTDH–BVMO fusions,[26] a PTDH–PAMO fusion was produced with four different linkers (Table 1). The authors found no difference in specific activity or thermostability between the linkers, and concluded that, for these two enzymes in particular, the termini of the enzymes were accommodating, in terms of flexibility. However, in the study on the ER–CHMO fusions, in which the same linkers were used, the tryptophan linker performed significantly worse in conversions.[6]

Although some of the studies listed looked at linkers of different length, there are only a few studies that rigorously investigated the effect of linker length. In one of these studies, a reductase was fused to CYP153 from *Marinobacter aquaeolei*, and the linker length was varied between 11 and 32 residues.[30] Three variants were found to have improved activity, of which one (with a linker of 18 residues) also showed improved stability and coupling efficiency, although expression was decreased. In another study, in which a P450 was fused to

Table 1. List of enzyme fusions.

Fusion name(s)	Enzyme types	Linker	Utility of fusion	Refs
TmCHMO-ADHA TmCHMO-ADHMi TbADH-TmCHMO	ADH, BVMO	glycine-rich linker (13): SSGGSGGSGGSAG	cascade reaction cyclic alcohol to lactone	[27]
ADH-Gly-BVMO ADH-FOM-BVMO ADH-BVMO	ADH, BVMO	L1: (12) SGGSGGSGGSAG L2: (30) SASNCLIGLFLNDQE LKKKAKVYDKIAKDV L3: (0)	cascade reaction alcohol to ester	[4]
pCRE2-BVMO	PTDH, BVMO	L1: (6) SRSAAG L2: (13) SSGGSGGSGGSAG L3: (12) SSATGSATGSAG L4: (1) W	NADPH-recycling system	[25], [26]
pCRE2-P450-BM3	PTDH, P450	Short linker (6): SRSAAG	NADPH-recycling system	[12], [33]
StyAL1B StyAL2B StyAL3B StyAL4B	styrene monooxygenase (StyA), flavin reductase (StyB)	L1: (3) WYH L2: (6) WYHHHH L3: (6) WYHAAA L4: (3) WYH (−3 aa from StyB)	electron transfer for epoxidation of styrene	[5]
Fus-SMO	styrene monooxygenase (StyA), flavin reductase (StyB)	flexible linker (30): ASGGGGSGGGGSGGG GSGGGGSGGGGSGAS	electron transfer for epoxidation of styrene	[11]
FH-FR	flavin-dependent halogenase, flavin reductase	L1: (10) PSPSTDQSPS L2: (16) VLHRHQPVTIGEPAAR L3: (22) VLHRHQPVSPIHSRTIG EPAAR	electron transfer for halogenation	[7]
ADH-AT	ADH, aminotransferase	L1: PAS linker: (20) ASPAAPAPA SPAAPAPSAPA L2: (40) PAS ×2 L3: (60) PAS ×3	cascade reaction alcohol to amine, stabilization through linker	[40]
OleT$_{JE}$-AldO	P450, alcohol oxidase	L1: (18) GSGLEVLFQGPGSG GGGS L2: (45) A(EAAAK)$_4$-LEA-(EAAAK)$_4$A	hydrogen peroxide supply for decarboxylation reaction, cleavable linker (underlined)	[8]
P-EugO P-HFMO P-ChitO P-HotAldO	peroxidase, alcohol oxidase	none	hydrogen peroxide supply for peroxidase, cascade reactions, biosensor	[36]
XenB-CHMO	ene reductase, BVMO	L1: (13) SSGGSGGSGGSAG L2: (12) SSATGSATGSAG L3: (1) W	cascade reaction unsaturated cyclic alcohols to chiral lactones	[6]
GlyDH-NADH oxidase	ADH, NOX	none	recycling of NAD+ for glycerol oxidation	[38]
PTS-FPPS	synthases	none	metabolic flux redirection towards sesquiterpene production	[20b]

a flavodoxin shuttle protein, a library of linkers was designed, ranging from one to 16 amino acids. The variants were compared based on whole-cell conversion, and the optimal linker was found to contain ten amino acids.[53] Strong differences were also found in a β-glucanase–xylanase fusion with different linker lengths, with greatly improved activities for both enzyme activities.[10] Based on these three studies it seems that, in addition to linker composition, the linker length can have a pronounced effect on the properties of the enzymes. It would be interesting to see whether this is also the case for other enzyme systems than the ones studied so far. To that end, a recent study reported a straightforward PCR method to generate linker libraries for any fusion construct.[54]

4. Conclusion and Outlook

In the emerging platform of multienzyme biocatalysis, enzyme fusion can be a useful tool for the simplification and optimiza-tion of a multienzyme system, in particular for enzyme production. In addition, the expression of fused enzymes can enable cascade reactions in vivo, which in some cases outperform the cascade reactions with coexpressed enzymes. In particular, the tool seems well suited to cascade systems that rely on co-factors or cosubstrates, such as NAD(P)H, FADH$_2$, and H$_2$O$_2$. Although in some cases the production of a bifunctional enzyme is primarily a matter of convenience, in other cases the pairing of two enzymes provides an advantage in terms of expression, catalytic activity, and stability.

From the overview of linkers from this set of biocatalytic fusion studies (Table 1), it seems that some of the studies chose no linker or a generally safe variant (the glycine-rich linker), and others explored different linkers. Even though linker design up to now has been treated in a somewhat arbitrary fashion, as signified by this set of studies, the choice of linker(s) in future studies could benefit from the collections of linkers in databases and computational simu-

lations, as well as from the ground work that has been covered here.

This review has primarily covered enzyme fusions for (oxidative) biocatalytic reactions, though there have also been extensive developments in the field of carbohydrate-active enzyme fusions.[16b] From there and from the other developing areas of enzyme fusion described in this review, inspiration can be drawn for novel investigations and applications.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] a) J. M. Sperl, V. Sieber, *ACS Catal.* **2018**, *8*, 2385–2396; b) J. H. Schrittwieser, S. Velikogne, W. Kroutil, *Catal. Sci. Technol.* **2018**, *8*, 205; c) J. H. Schrittwieser, S. Velikogne, M. Hall, W. Kroutil, *Chem. Rev.* **2018**, *118*, 270–348; d) S. Wu, Z. Li, *ChemCatChem* **2018**, *10*, 2164–2178; e) K. S. Rabe, J. Müller, M. Skoupi, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2017**, *56*, 13574–13589; *Angew. Chem.* **2017**, *129*, 13760–13777; f) M. B. Quin, K. K. Wallin, G. Zhang, C. Schmidt-Dannert, *Org. Biomol. Chem.* **2017**, *15*, 4260–4271; g) S. P. France, L. J. Hepworth, N. J. Turner, S. L. Flitsch, *ACS Catal.* **2017**, *7*, 710–724.

[2] R. A. Sheldon, J. M. Woodley, *Chem. Rev.* **2018**, *118*, 801–838.

[3] a) R. A. Sheldon, P. C. Pereira, *Chem. Soc. Rev.* **2017**, *46*, 2678–2691; b) G. Li, J. Wang, M. T. Reetz, *Bioorg. Med. Chem.* **2018**, *26*, 1241–1251; c) J. L. Porter, R. A. Rusli, D. L. Ollis, *ChemBioChem* **2016**, *17*, 197–203.

[4] E. Y. Jeon, A. H. Baek, U. T. Bornscheuer, J. B. Park, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 6267–6275.

[5] T. Heine, K. Tucker, N. Okonkwo, B. Assefa, C. Conrad, A. Scholtissek, M. Schlömann, G. Gassner, D. Tischler, *Appl. Biochem. Biotechnol.* **2017**, *181*, 1590–1610.

[6] C. Peters, F. Rudroff, M. D. Mihovilovic, U. T. Bornscheuer, *Biol. Chem.* **2017**, *398*, 31–37.

[7] M. C. Andorfer, K. D. Belsare, A. M. Girlich, J. C. Lewis, *ChemBioChem* **2017**, *18*, 2099–2103.

[8] S. Matthews, K. L. Tee, N. J. Rattray, K. J. McLean, D. Leys, D. A. Parker, R. T. Blankley, A. W. Munro, *FEBS Lett.* **2017**, *591*, 737–750.

[9] a) B. Li, Y. Li, D. Bai, X. Zhang, H. Yang, J. Wang, G. Liu, J. Yue, Y. Ling, D. Zhou, et al., *Sci. Rep.* **2015**, *4*, 6750; b) I. Sührer, M. Haslbeck, K. Castiglione, *Process Biochem.* **2014**, *49*, 1527–1532; c) Y. Zhang, Y. Wang, S. Wang, B. Fang, *Eng. Life Sci.* **2017**, *17*, 989–996.

[10] P. Lu, M. G. Feng, *Appl. Microbiol. Biotechnol.* **2008**, *79*, 579–587.

[11] M. L. Corrado, T. Knaus, F. G. Mutti, *ChemBioChem* **2018**, *19*, 679–686.

[12] N. Beyer, J. K. Kulig, M. W. Fraaije, M. A. Hayes, D. B. Janssen, *ChemBioChem* **2018**, *19*, 326–337.

[13] a) H. Hirakawa, T. Nagamune in *Methods in Molecular Biology, vol. 978: Enzyme Engineering: Methods and Protocols* (Eds.: J. C. Samuelson), Humana, Totowa, **2013**, pp. 149–163; b) S. Schulz, D. Schumacher, D.

[14] I. Wheeldon, S. D. Minteer, S. Banta, S. C. Barton, P. Atanassov, M. Sigman, *Nat. Chem.* **2016**, *8*, 299–309.

[15] C. You, Y.-H. P. Zhang, *ACS Synth. Biol.* **2014**, *3*, 380–386.

[16] a) H. Yang, L. Liu, F. Xu, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8273–8281; b) S. Elleuche, *Appl. Microbiol. Biotechnol.* **2015**, *99*, 1545–1556.

[17] C. Yanofsky, *Bacteriol. Rev.* **1960**, *24*, 221.

[18] H. Nyunoya, K. E. Broglie, C. J. Lusty, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 2244–2246.

[19] L. O. Narhi, A. J. Fulco, *J. Biol. Chem.* **1986**, *261*, 7160–7169.

[20] a) I. Meynial Salles, N. Forchhammer, C. Croux, L. Girbal, P. Soucaille, *Metab. Eng.* **2007**, *9*, 152–159; b) M. Tumen-Velasquez, C. W. Johnson, A. Ahmed, G. Dominick, E. M. Fulk, P. Khanna, S. A. Lee, A. L. Schmidt, J. G. Linger, M. A. Eiteman, G. T. Beckham, E. L. Neidle, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 7105–7110; c) L. Albertsen, Y. Chen, L. S. Bach, S. Rattleff, J. Maury, S. Brix, J. Nielsen, U. H. Mortensen, *Appl. Environ. Microbiol.* **2011**, *77*, 1033–1104.

[21] J. Yourno, T. Kohno, J. R. Roth, *Nature* **1970**, *228*, 820–824.

[22] L. Bülow, P. Ljungcrantz, K. Mosbach, *Nat. Biotechnol.* **1985**, *3*, 821–823.

[23] R. J. Conrado, J. D. Varner, M. P. DeLisa, *Curr. Opin. Biotechnol.* **2008**, *19*, 492–499.

[24] J. B. van Beilen, F. Mourlane, M. A. Seeger, J. Kovac, Z. Li, T. H. M. Smits, U. Fritsche, B. Witholt, *Environ. Microbiol.* **2003**, *5*, 174–182.

[25] D. E. Torres Pazmiño, R. Snajdrova, B. J. Baas, M. Ghobrial, M. D. Mihovilovic, M. W. Fraaije, *Angew. Chem. Int. Ed.* **2008**, *47*, 2275–2278; *Angew. Chem.* **2008**, *120*, 2307–2310.

[26] D. E. Torres Pazmiño, A. Riebel, J. de Lange, F. Rudroff, M. D. Mihovilovic, M. W. Fraaije, *ChemBioChem* **2009**, *10*, 2595–2598.

[27] F. S. Aalbers, M. W. Fraaije, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 7557–7565.

[28] E. W. van Hellemond, D. B. Janssen, M. W. Fraaije, *Appl. Environ. Microbiol.* **2007**, *73*, 5832–5839.

[29] S. Eiben, H. Bartelmäs, V. B. Urlacher, *Appl. Microbiol. Biotechnol.* **2007**, *75*, 1055–1061.

[30] S. M. Hoffmann, M. J. Weissenborn, Ł. Gricman, S. Notonier, J. Pleiss, B. Hauer, *ChemCatChem* **2016**, *8*, 1591–1597.

[31] a) F. Sabbadin, R. Hyde, A. Robin, E. M. Hilgarth, S. Delenne, S. Flitsch, N. Turner, G. Grogan, N. C. Bruce, *ChemBioChem* **2010**, *11*, 987–994; b) F. Sabbadin, G. Grogan, N. C. Bruce in *Methods in Molecular Biology: vol. 987: Cytochrome P450 Protocols* (Eds.: I. R. Phillips, E. A. Shephard, P. R. Ortiz de Montellano), Humana, Totowa, **2013**, pp. 239–249.

[32] J. K. Kulig, C. Spandolf, R. Hyde, A. C. Ruzzini, L. D. Eltis, G. Grönberg, M. A. Hayes, G. Grogan, *Bioorg. Med. Chem.* **2015**, *23*, 5603–5609.

[33] N. Beyer, J. K. Kulig, A. Bartsch, M. A. Hayes, D. B. Janssen, M. W. Fraaije, *Appl. Microbiol. Biotechnol.* **2017**, *101*, 2319–2331.

[34] H. Watanabe, H. Hirakawa, T. Nagamune, *ChemCatChem* **2013**, *5*, 3835–3840.

[35] C. Lu, F. Shen, S. Wang, Y. Wang, J. Liu, W. J. Bai, X. Wang, *ACS Catal.* **2018**, *8*, 5794–5798.

[36] D. I. Colpa, N. Lončar, M. Schmidt, M. W. Fraaije, *ChemBioChem* **2017**, *18*, 2226–2230.

[37] G. Rehn, A. T. Pedersen, J. M. Woodley, *J. Mol. Catal. B* **2016**, *134*, 331–339.

[38] F. B. Fang, W. Jiang, Q. Zhou, S. Wang, *PLoS One* **2015**, *10*, e0128412.

[39] T. Sun, B. Li, Y. Nie, D. Wang, Y. Xu, *Bioresources Bioprocess.* **2017**, *4*, 21.

[40] A. Lerchner, M. Daake, A. Jarasch, A. Skerra, *Protein Eng. Des. Sel.* **2016**, *29*, 557–562.

[41] M. Schlapschy, U. Binder, C. Borger, I. Theobald, K. Wachinger, S. Kisling, D. Haller, A. Skerra, *Protein Eng. Des. Sel.* **2013**, *26*, 489–501.

[42] F. G. Mutti, T. Knaus, N. S. Scrutton, M. Breuer, N. J. Turner, *Science* **2015**, *349*, 1525–1529.

[43] X. Chen, J. L. Zaro, W. C. Shen, *Adv. Drug Delivery Rev.* **2013**, *65*, 1357–1369.

[44] G. Li, Z. Huang, C. Zhang, B.-J. Dong, R.-H. Guo, H.-W. Yue, L.-T. Yan, X.-H. Xing, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 215–225.

[45] M. van Rosmalen, M. Krom, M. Merkx, *Biochemistry* **2017**, *56*, 6565–6574.

[46] E. Papaleo, G. Saladino, M. Lambrughi, K. Lindorff-Larsen, F. L. Gervasio, R. Nussinov, *Chem. Rev.* **2016**, *116*, 6391–6423.

[47] S. I. Lim, B. Yang, Y. Jung, J. Cha, J. Cho, E. S. Choi, Y. H. Kim, I. Kwon, *Sci. Rep.* **2016**, *6*, 39587.

[48] Y. T. Lai, L. Jiang, W. Chen, T. O. Yeates, *Protein Eng. Des. Sel.* **2015**, *28*, 491–500.

[49] K. Yu, C. Liu, B. G. Kim, D. Y. Lee, *Biotechnol. Adv.* **2015**, *33*, 155–164.

[50] R. Orru, H. M. Dudek, C. Martinoli, D. E. Torres Pazmiño, A. Royant, M. Weik, M. W. Fraaije, A. Mattevi, *J. Biol. Chem.* **2011**, *286*, 29284–29291.

[51] B. J. Yachnin, T. Sprules, M. B. McEvoy, P. C. K. Lau, A. M. Berghuis, *J. Am. Chem. Soc.* **2012**, *134*, 7788–7795.

[52] D. Tischler, R. Schwabe, L. Siegel, K. Joffroy, S. Kaschabek, A. Scholtissek, T. Heine, *Molecules* **2018**, *23*, 809.

[53] K. D. Belsare, A. J. Ruff, R. Martinez, A. V. Shivange, H. Mundhada, D. Holtmann, J. Schrader, U. Schwaneberg, *Biotechniques* **2014**, *57*, 13–20.

[54] J. L. Norris, R. M. Hughes, *FEBS Open Bio* **2018**, *8*, 1029–1042.