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 biotransformations 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.[4,5] 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 outperform 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.[14] Substrate channeling is
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. Some researchers thoroughly investigated substrate channeling in their systems, and provided evidence for presence or absence thereof. 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. In addition, we have included only a few studies on carbohydrate-active and biomass-degrading enzyme fusions, as these have also been reviewed elsewhere.

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 result in fused enzymes. Some known examples are tryptophan synthase, carbamoyl phosphatase synthase, and cytochrome P450 BM3. 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 glycol production and in another case through a rapid-evolution method that was used with selection for growth on guaiacol. 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. 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. 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. From this point on, many more enzyme fusions were produced and studied, of which some have been reviewed before. This Minireview highlights some recent developments in fusing enzymes for biocatalytic applications.

2.1. Baeyer–Villiger monoxygenase (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. 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 peroxoFLavin 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 NADPH to NADP+. Some typical examples of such enzymes are glucose dehydrogenase (GDH), formate dehydrogenase (FDH), and phosphite dehydrogenase (PTDH).

Pazmino and co-workers looked into the effects of fusing PTDH from Pseudomonas sp. to BVMOs so as to provide direct recycling of NADPH. 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 monoxygenase (CHMO), phenylacetone monoxygenase (PAMO), and cyclopentanone monoxygenase (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 $K_{\text{m}}$ for NADP+ and phosphite, whereas the BVMOs showed a small decrease in $K_{\text{m}}$ 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.

Still, for a large-scale reaction, the sacrificial substrate can constitute a significant cost contribution, as it also gives a by-
to oxidize the alcohol to a ketone to produce Tm

For Dehydrogenase fusion with a BVMO, in which the dehydrogenase

The fusions were evaluated based on their activi-

Inspired by the previous study on E. coli

The flavin reductase is often referred to as “styrene

As /C23

20

2019

2019

www.chembiochem.org

The authors linked the improved mass balance that

influence this ratio, possibly through slight conformational

Another study explored the same combination of enzyme

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. As the group had
demonstrated prior to this work, by combining these three en-
zymes, one can convert unsaturated cyclic alcohols to enanti-

product. An alternative to a sacrificial substrate for regenerat-
ing 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 sub-
strates to convert these in a cascade reaction to esters. 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 be-
tween 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 inter-
mediate. This would be an interesting advantage for such
fusion enzymes, and it is reminiscent of the substrate channel-
ing in natural fusion enzymes, for example, tryptophan syn-
thase.[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 monoxygenase (TmCHMO)
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 inac-
tive 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 oligo-
merization 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.[26] For
the other ADH and for TmCHMO 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 detri-
tmental to correct folding or oligomerization.

Another interesting exercise in producing fusions with

2.2. Flavin reductase (FR) fusions

Similarly to BVMOs, styrene monoxygenases (SMOs) use a re-
duced flavin cofactor (FADH₂) 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
monoxygenase 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 investi-
gating the effects of different linkers.

In 2016, the groups of Gassner and Tischler carefully investi-
gated the effect of fusing the reductase with the epoxidase
component.[15] The fusions were evaluated based on their activi-
ty 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-
This activity was higher than that of naturally fused StyA2B and the fusion enzymes addressed above.\(^\text{[5]}\) 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.\(^\text{[7]}\) 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*.\(^\text{[12]}\) 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\(^\text{[29]}\) and/or to improve electron-coupling efficiency.\(^\text{[28]}\) 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.\(^\text{[21]}\) 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.\(^\text{[32]}\)

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.\(^\text{[33]}\) 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.\(^\text{[12, 33]}\) A similar study was performed with a heterotrimeric P450 from *Pseudomonas putida*, in which it was fused to PTDH for self-sufficient hydroxylations.\(^\text{[29]}\)

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 aldol oxidase with a cytochrome P450 (Scheme 3A). The particular P450 monooxygenase from *Jeotgalicoccus* sp., named OleT,
\[\text{P450} + \text{O}_{2} + \text{NADPH} \rightarrow \text{hydroxylated product} + \text{NADP}^{+} + \text{H}^{+}\]
can act as an peroxgenase and primarily catalyzes the oxidative decarboxylation of fatty acids. In its mode as peroxxygenase, 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 peroxgenase. 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.\(^\text{[32]}\) The choice of aldol oxidase allowed the use of glycerol, which is very cheap and
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, thereby allowing for complete alcohol oxidations, it is necessary to recycle NAD(P)H [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, thereby allowing for complete alcohol oxidations, it is necessary to recycle NAD(P)H [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, thereby allowing for complete alcohol oxidations, it is necessary to recycle NAD(P)H [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, thereby allowing for complete alcohol oxidations, it is necessary to recycle NAD(P)H [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, thereby allowing for complete alcohol oxidations, it is necessary to recycle NAD(P)H [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.
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.\[23\] 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 $\alpha$-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,\[49,50\] 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.\[41\]

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.\[52,53\]

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-residue 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).\[51\] In a recent study of a natural FR–SMO fusion from Variorax 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 $K_{\text{m}}$ for FAD (from 33 to 1.8 $\mu$M) 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,\[29\] 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.\[56\]

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...
af lavodoxin 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.

Strong differences were also found in a β-glucanase–xylanase fusion with different linker lengths, with greatly improved activities for both enzyme activities.

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 enzymes 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.

4. Conclusion and Outlook

In the emerging platform of multienzyme biocatalysis, enzyme fusion can be a useful tool for the simplification and optimization 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 cofactors or cosubstrates, such as NAD(P)H, FADH₂, and H₂O₂. 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-

### Table 1. List of enzyme fusions.

<table>
<thead>
<tr>
<th>Fusion name(s)</th>
<th>Enzyme types</th>
<th>Linker Utility of fusion</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TmCHMO-ADHA</td>
<td>ADH, BVMO</td>
<td>glycine-rich linker (13): SSGGSGGGSGGSAG to lactone</td>
<td>[27]</td>
</tr>
<tr>
<td>TmCHMO-ADHMi</td>
<td>ADH, BVMO</td>
<td>L1: (12) SGGSGGGSGGSAG to lactone</td>
<td>[4]</td>
</tr>
<tr>
<td>ADH-Gly-BVMO</td>
<td>ADH, BVMO</td>
<td>L2: (30) SASNCGLFLNDQELKKKAKVYDKIADV to lactone</td>
<td>[4]</td>
</tr>
<tr>
<td>ADH-BVMO</td>
<td>PTDH, BVMO</td>
<td>L3: (0) SGGGSGGGSGGSAG to lactone</td>
<td>[4]</td>
</tr>
<tr>
<td>pCRE2-BVMO</td>
<td>PTDH, BVMO</td>
<td>L4: (1) W SGGGSGGGSGGSAG to lactone</td>
<td>[4]</td>
</tr>
<tr>
<td>pCRE2-P450-BM3</td>
<td>PTDH, P450</td>
<td>Short linker (6): SRRSAG to lactone</td>
<td>[12], [33]</td>
</tr>
<tr>
<td>StyAL1B</td>
<td>flavin reductase (StyB)</td>
<td>L1: (3) WYH to lactone</td>
<td>[7]</td>
</tr>
<tr>
<td>StyAL2B</td>
<td>flavin reductase (StyB)</td>
<td>L1: (6) WYHHH to lactone</td>
<td>[7]</td>
</tr>
<tr>
<td>StyAL3B</td>
<td>flavin reductase (StyB)</td>
<td>L2: (6) WYHAAA to lactone</td>
<td>[7]</td>
</tr>
<tr>
<td>StyAL4B</td>
<td>flavin reductase (StyB)</td>
<td>L3: (6) WYH (–3 aa from StyB) to lactone</td>
<td>[7]</td>
</tr>
<tr>
<td>Fus-SMO</td>
<td>flavin reductase (StyB)</td>
<td>L4: (4) WYH (–3 aa from StyB) to lactone</td>
<td>[7]</td>
</tr>
<tr>
<td>FH-FR</td>
<td>flavin-dependent halogenase, flavin reductase (StyB)</td>
<td>L1: (10) PSPSTDQSPS to lactone</td>
<td>[11]</td>
</tr>
<tr>
<td>ADH-AT</td>
<td>ADH, aminotransferase</td>
<td>L2: (60) PAS x = 3 to lactone</td>
<td>[11]</td>
</tr>
<tr>
<td>OleTₐ-AldO</td>
<td>P450, alcohol oxidase</td>
<td>L1: (18) GGGSLEVLFQQPGSGGG to lactone</td>
<td>[8]</td>
</tr>
<tr>
<td>P-EugO</td>
<td>peroxidase, alcohol oxidase</td>
<td>L1: (20) ASPAAPAPA to lactone</td>
<td>[36]</td>
</tr>
<tr>
<td>P-HFMO</td>
<td>none to lactone</td>
<td>[36]</td>
<td></td>
</tr>
<tr>
<td>P-ChitO</td>
<td>none to lactone</td>
<td>[36]</td>
<td></td>
</tr>
<tr>
<td>P-HotAldO</td>
<td>none to lactone</td>
<td>[36]</td>
<td></td>
</tr>
<tr>
<td>XenB-ChMO</td>
<td>ene reductase, BVMO</td>
<td>L1: (13) SSGGSGGGSGGSAG to lactone</td>
<td>[6]</td>
</tr>
<tr>
<td>GlyDH-NADH oxidation</td>
<td>ADH, NOX</td>
<td>L2: (12) SSATGSAATGSA to lactone</td>
<td>[38]</td>
</tr>
<tr>
<td>PTS-FPPS</td>
<td>synthases none to lactone</td>
<td>[38]</td>
<td></td>
</tr>
</tbody>
</table>

lations, as well as from the ground work that has been covered before. This review has primarily covered enzyme fusions for (oxida-
tive) biocatalytic reactions, though there have also been exten-
sive developments in the field of carbohydrate-active enzyme 
fusions.

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

This work has received funding from the European Union Project 
ROBOX (grant agreement no. 635734) under the EU’s Horizon 
2020 Programme Research and Innovation actions H2020-LEIT 
BIO-2014-1. The views and opinions expressed in this article are 
only those of the authors, and do not necessarily reflect those 
of the European Union Research Agency. The European Union is 
not liable for any use that may be made of the information con-
tained herein.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: biocatalysis · cascade reactions · enzyme fusion · enzymes · linkers


© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim