Advanced Synthesis & Catalysis

DOI: 10.1002/adsc.201500202

Designer Microorganisms for Optimized Redox Cascade Reactions – Challenges and Future Perspectives

Thomas Bayer,[a] Sofia Milker,[a] Thomas Wiesinger,[a] Florian Rudroff,[a,*] and Marko. D. Mihovilovic[a]

[a] Department of Applied Synthetic Chemistry, TU Wien, Getreidemarkt 9, 163 OC, 1060 Vienna, Austria
Fax: (+43)-1-58801-916-3618, phone: (+43)-1-58801-163-618; e-mail: florian.rudroff@tuwien.ac.at

Received: February 26, 2015; Revised: April 30, 2015; Published online: May 15, 2015

Abstract: An immense number of chemical reactions are carried out simultaneously in living cells. Nature's optimization approach encompasses the assembly of reactions in cascades and to embed them in finely tuned metabolic networks. With the vast progress in the field of biocatalysis, man-made cascades, especially redox cascades, have reached a degree of complexity that needs tools for improved control and optimization. Combined strategies from biocatalysis, metabolic engineering and synthetic biology lead to the establishment of artificial metabolic pathways with minimized interference with the cellular host environment. This review will focus on genetic and metabolic engineering tools for the assembly and introduction of *de novo* redox pathways into the host *Escherichia coli* and will present state of the art redox cascades performed by tailor-made microbial cell factories.

Keywords: biotransformations; knock-in/knock-out; metabolic engineering; redox biocatalysis; synthetic biology

1 Introduction

The concept of biocatalysis has transformed tremendously from its starting point by employing (isolated) enzymes in single-step reactions to today's realization of complex synthetic schemes by applying multi-enzymatic cascades.[1] Although the optimization of *in vitro* approaches was similar to classical organic synthesis, for example, by changing reaction parameters such as substrate and enzyme concentrations, temperature or pH, scientists soon had to face limitations. Many *in vitro* transformations consist of only a few enzymes and cannot compete with the complexity and efficiency of natural metabolic pathways.[2] The impracticality of cell-free systems starts with the prior preparation of biocatalysts which usually involves the expression, isolation, and purification of heterologous enzymes from host cells. It is a major drawback that many enzymes [e.g., oxidases, alcohol dehydrogenases (ADHs)] are dependent on cofactors which have to be either added in stoichiometric amounts or regenerated by enzymatic recycling systems.[2,3] Both render *in vitro* redox cascades often uneconomical and add complexity to a system that started out being simple. Consequently, scientists applied multi-step biotransformations preferably in living cells. Within such *in vivo* approaches there is no need to isolate cascade components, as enzymes are usually more stable in a cellular environment, and the host metabolism supplies and recycles cofactors.[2,3] Nonetheless, it is the complexity of whole cell systems that makes them superior to transformations carried out *in vitro* and, at the same time, complicates control in many aspects.

While nature had millions of years to optimize challenging reactions such as redox reactions and to embed them in complex and efficiently regulated met-

Thomas Bayer, born in 1985, graduated in molecular biology at the University of Vienna in 2012. Before joining the group of Prof. Marko D. Mihovilovic for his PhD on the design of microbial cell factories at the Vienna University of Technology (VUT), he already worked in industry with a strong focus on immunology. During a short-term stay in Prof. U. T. Bornscheuer's group at the Ernst-Moritz-Arndt University Greifswald, he further deepened his interest in biocatalysis and metabolic engineering.

Sofia Milker was born in St. Petersburg, Russia in 1987 and moved to Germany in 1996. She conducted her studies in biotechnology with specialization in biochemical engineering at the RWTH Aachen. After graduation she moved to Vienna and in January 2013 she started working on her PhD project in the group of Prof. Marko D. Mihovilovic, which deals with design and application of non-native enzyme cascades in *E. coli* in order to produce chiral target compounds.

Thomas Wiesinger was born in 1986 in Vöcklabruck, Austria and studied at the Technical University of Graz, where he completed his diploma in chemistry in 2013. He then moved to the Technical University of Vienna and is currently working on his PhD in the group of Prof. Marko D. Mihovilovic. His work involves the development of a non-native pathway in a living cell factory.

Marko D. Mihovilovic graduated in 1993 and received his PhD in organic chemistry from Vienna University of Technology (VUT) in 1996. Post-doc placements as an Erwin Schrödinger fellow followed at the University of New Brunswick (Canada; 1997) and the University of Florida (USA; 1998), in the fields of biocatalysis and molecular biology. Returning to VUT, he set up his own research group in 1999. He completed his habilitation in 2003 in bio-organic chemistry and was promoted to Associate Professor in 2004. In 2014 he was appointed as Full Professor for Bioorganic Synthetic Chemistry and he acts as Head of the Institute for Applied Synthetic Chemistry at VUT since 2013.

Florian Rudroff graduated in 2003 and received his PhD in 2007 from Vienna University of Technology (VUT) in bioorganic chemistry. He then joined the group of Prof. Uwe Sauer at ETH Zurich as an Erwin Schrödinger postdoctoral fellow and expanded his field of expertise from biocatalysis towards molecular systems biology. Since 2011 he is Senior PostDoc in the group of Prof. Marko D. Mihovilovic at Vienna University of Technology. His current research interests focus on the interplay between biocatalysis and metabolic engineering to design artificial metabolic pathways of naturally non-related enzymes; the approach is based on the concept of "retrosynthesis" and aims for the production of pharmaceuticals, fine- and platform chemicals.

abolic networks, scientists usually do not have that much time. With interdisciplinary concepts from metabolic engineering and synthetic biology, scientists in the field of biocatalysis try to mimic nature and aim for the establishment of redox cascades that are equally sophisticated, demanding and optimized on different molecular levels.

With the research progress made in the last decades, genetic and metabolic engineering tools for the modification and manipulation of the host *E. coli* are available that not only enable the design and application of entire *de novo* pathways but also their optimization by identifying bottlenecks which make these cascades less efficient in regard to product titers and host cell viability. Such phenomena are often associated with the term metabolic burden.[4] Bottlenecks can originate from limited permeability of host cells[5] or leaking of pathway intermediates, their accumulation

due to different kinetics or unbalanced production of heterologously expressed enzymes[6] or a depletion of redox cofactors.[3] Bottlenecks can also be caused by the metabolic background of the host including by-product formation from cascade intermediates catalyzed by natively expressed enzymes.[7] Additional limitations can be caused by diffusion of intermediates out of the cell or their individual toxicity. Several optimization strategies are available on all molecular levels and include, for example, the use of balanced promoter systems[8] (transcription), codon-optimized genes[9] (translation), and the spatial organization of pathways by scaffold proteins.[10] Recent advances in optimization target on improved cofactor availability and identification as well as removal of competing background reactions.[11]

This work will deal with the design, construction and optimization of redox cascades in the best studied host organism *Escherichia coli* (*E. coli*). Moreover, this review will focus on holistic approaches made to simultaneously both reduce the metabolic burden for the host and optimize redox cascades by multilevel manipulations through strategies and techniques from metabolic engineering, systems biology, synthetic biology and reaction engineering. Most recent genetic tools used to introduce *de novo* pathways in *E. coli*, one of the most commonly used hosts in metabolic engineering, will be compared. It will be depicted how rational gene knock-out (KO) leads to optimized redox cascades with minimized background. Moreover a short outlook on computational modelling and metabolic flux analysis will be given. Examples for tailor-made microbial cell factories and the cascade reactions they are able to perform will highlight the interdependency between redox biocatalysis and host metabolism.

2 Tools for Cellular Pathway Engineering

A well-characterized host for the heterologous expression of all *de novo* pathway components (usually proteins) represents the foundation of engineering new metabolic routes. The Gram-negative bacterium *E. coli* is most suitable because of its rapid growth at high density on cheap nutrients, the availability of a large number of different cloning vectors, and its well-understood genetics.[12] Starting in the late 1990s, the elucidation of whole genome sequences for various microbes [e.g., *E. coli*, *Bacillus subtilis* (*B. subtilis*) or *Saccharomyces cerevisiae* (*S. cerevisiae*)] went hand in hand with the fast development in the *omics* disciplines (transcriptomics, proteomics, metabolomics, fluxomics). This granted access to metabolic network information for the first time and led to the development of methods for gene KO and knock-in (KI) towards the construction of mutant libraries (e.g., the Keio collection[13]). Such a gain of knowledge also supported the emerging field of synthetic biology and enabled the remodelling of biological systems and manipulations on larger scales.[2]

The key players in biosynthetic pathways and metabolic networks are proteins, to be exact, enzymes. Immense numbers of different enzymes are simultaneously produced by the transcriptional and ribosomal machinery in living cells to ensure cellular functionality, growth, and reproduction.[2,12b] Enzymes perform sequential multi-step transformations with high demands on chemo-, regio-, and stereo-selectivity under the certain reaction conditions of the cellular environment (e.g., different compartments of the cell) which make them compatible with each other and enable the construction of artificial pathways consisting of naturally non-related enzymes.[1b] For both biocatalytic *in vitro* and *in vivo* applications, enzymes have to be supplied in sufficient and stoichiometric amounts. Early attempts were achieved by over-expressing enzymes from plasmids. This classic approach is still extensively used for (soluble) proteins based on increasing numbers of different vectors available and standardized guidelines for expression.[12b] The expression of many different (heterologous) pathway components though is far more elaborate and will be discussed in the following section together with improved genetic tools and new strategies for their introduction in *E. coli*.

2.1 The Challenges of Multiple Recombinant Protein Expression

The expression of each enzyme from a single plasmid is not feasible for the establishment of whole metabolic pathways as this approach involves separate rounds of molecular cloning which can be tricky and time-consuming.

Major obstacles are encountered after introduction into the host as the over-production of recombinant target proteins[14] and the replication of foreign DNA[15] utilize a significant amount of the host cell's resources.[4] The imposed metabolic load leads to a drain of resources from the host metabolism inducing stress responses (e.g., heat shock, starvation).[12b] This results in down-regulation of housekeeping genes involved in transcription, translation and amino acid biosynthesis,[16] reducing both growth rate[14,17] and cell viability.[18] Consequently, the flux through the *de novo* pathway can be strongly impaired which leads to low productivities.[19] In line with tuning chemical reactions, also stress responses can be minimized by the adaption of expression conditions (e.g., the use of different plasmids, changing inducer concentration, growth temperature, and medium composition). This simple approach may not suffice as re-

Table 1. Replicon incompatibility groups.

Plasmid(s)	ORI	Copy No.[a]	Incompatible Replicons
pUC	pMB1[b]	300–500	[c]
pGEM	pMB1[b]	300–500	[c]
pBluescript	ColE1[d] & F1[e]	300–500	[c]
pRSF	RSF1030	>100	RSF1030
pET (all)	pBR322 & F1[e]	~40	[c]
pETDuet-1	ColE1	~40	[c]
pCDF	CloDF13	20–40	CloDF13
pCOLA	ColA	20–40	ColA
pR6K	R6K[f]	15–20	[g]
pACYC	P15A	10–12	P15A
pACYCDuet-1	P15A	10–12	P15A
pRARE	P15A	10–12	P15A
pETcoco™	Mini-F/RK2	~40[h]	[g]
pSC101	pSC101	~5	[g]

[a] Actual copy numbers may vary depending on factors such as the type/size of insert, *E. coli* strain used and growth conditions. Numbers adapted from Sambrook et al.,[21b] Held et al.[22] or Sørensen and Mortensen.[12b]
[b] pMB1 derivative. This ORI is also related to pBR322 and ColE1 (i.e., they belong to the same incompatibility group).
[c] Incompatible replicons include pBR322, pMB1, ColE1 and their derivatives.
[d] ColE1 derivative.
[e] F1 is a phage-derived ORI that allows for the replication and packaging of ssDNA into phage particles.
[f] Requires *pir* gene for replication.[23]
[g] Incompatible replicons include F1 and Mini-F1/RK2, R6K, P15A, and pSC101.
[h] Amplifiable up to ~40 copies/cell.

sponses triggered by cells under stress conditions are complex and the metabolic burden bestowed upon the host by the introduction of many pathway elements is innately high. The flaws and faults of recombinant protein production in *E. coli* have been addressed in detail by different authors.[12,20] Recent advances aim for more efficient and convenient tools for multiple gene expression that allow the assembly of whole biosynthetic pathways and, simultaneously, reduce the metabolic burden.

The immediate solution to reduce the plasmid burden is the co-expression of two (or more) genes from a single plasmid. In such a way, a metabolic pathway could be expressed with improved stability from only a handful of plasmids when considering replicon incompatibility[21] (Table 1) and positive selection tools (e.g., antibiotic resistance markers) for host cells harboring all desired vector constructs.

Zhang et al. introduced a pathway for the synthesis of non-natural alcohols in *E. coli,* starting from L-threonine (**1**) and optimized it for the production of (*S*)-3-methyl-1-pentanol (**4**) (Figure 1). Their modular approach consisted of three plasmids encoding three synthetic operons (Figure 2A); two of which drive the

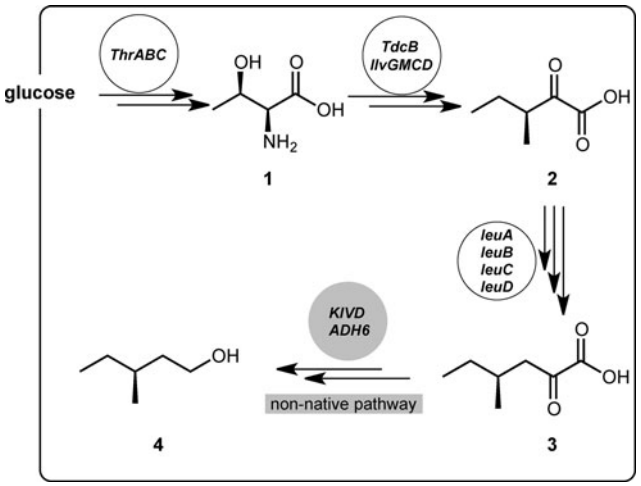

Figure 1. Synthetic pathways for non-natural alcohol production. Overexpression of *ThrABC*, *TdcB* and *IlvGMCD* drives carbon flux towards 2-keto-3-methyl-valerate (**2**), converted by *LeuABCD* to (*S*)-4-methyl-2-oxohexanoic acid (**3**), *KIVD* and *ADH6* comprise the non-native pathway leading to (*S*)-3-methyl-1-pentanol (**4**).

carbon flux towards the natural precursor 2-keto-3-methyl-valerate (**2**) and one comprising the *de novo* pathway.[24]

One popular approach for protein co-expression is the Duet expression system from Novagen. It combines different compatible origins of replication (ORIs; ColE1, CloDF13, P15A, ColA and RSF1030) and different antibiotic markers (ampicillin, streptomycin, chloramphenicol and kanamycin). Held and co-workers demonstrated the co-expression of up to eight proteins in one cell from four Duet vectors.[22,25] Within the context of metabolic pathway incorporation, scientists in the Prather group used the Duet system for the biosynthesis of (*R*)- and (*S*)-3-hydroxybutyrate,[26] butanol[27] and the non-natural alcohols pentanol[28] and 4-methyl-1-pentanol.[29] The last compound is formed *via* a ten-step *de novo* pathway with enzymes taken from nine different microorganisms. In the cases of non-natural alcohol production, enzymes are grouped in modules that allowed rational combination of different modules to elevate product titers.

Such modules or enzymatic toolboxes have to be assembled thoughtfully by conventional cloning. Despite the trendsetting examples of pathway engineering above, the Duet vector sets have drawbacks. Quick modular exchanges of a single or multiple pathway elements or changes in the pathway configuration (Figure 2) are not possible and gene expression is not tunable with only the strong isopropyl β-D-1-thiogalactopyranoside (IPTG)-inducible T7 promoter in place. But the establishment of artificial metabolic routes apparently needs modularity and fine tuning not only on different steps of the pathway but also on molecular levels (e.g., transcription, translation).[30]

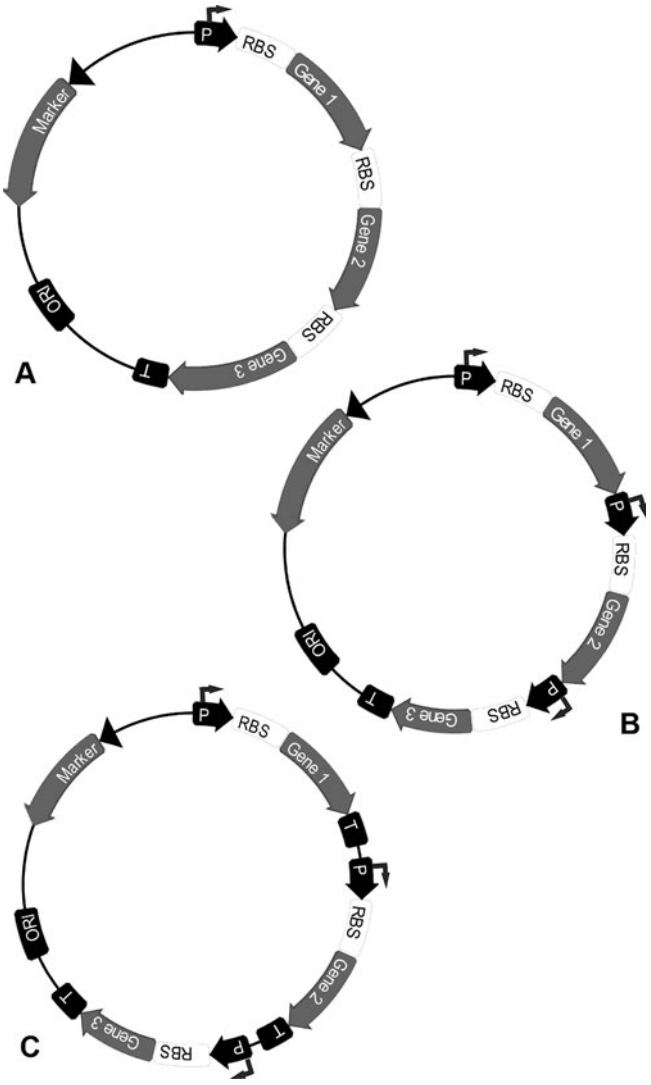

Figure 2. Pathway configurations. Plasmids contain an ORI for their maintenance in host cells and a marker gene (e.g., antibiotic resistance) for the selection of transformants. Pathways can be arranged on plasmids in different configurations. A) In operon forms, a gene cluster is under the control of one promoter (P) and one terminator (T). B) Pseudo-operons are terminated by a single terminator whereas the expression of each gene is controlled by individual promoters. C) In monocistronic configurations, multiple genes are arranged back-to-back. Each gene is controlled by a single promoter and terminator.

The production of all pathway enzymes needs to be coordinated, individual expressions must be tuned to balance the simultaneous expression of all pathway elements in order to reduce the metabolic burden on the host.[8a,19a,30b,31] Attempts to streamline pathway construction were suggested by Rebatchouk in 1996 within a general cloning strategy referred to as nucleic acid ordered assembly with directionality (NOMAD),[32] which was revived and extended to the BioBrick™ standard by Knight in 2003.[33] The BioBrick™ standard aims for the design and construction of genetic systems from standardized biological parts [e.g., promoter, ribosome binding site (RBS)].[34] It utilizes the isocaudomer pairs *XbaI* and *SpeI*. Isocaudomers are pairs of restriction enzymes that have slightly different recognition sequences but, upon cleavage, generate identical cohesive termini. Ligation results in a scar sequence that cannot be cleaved by either of the original restriction enzymes. BioBrick™ parts are assembled by iterative rounds of restriction enzyme digestion and ligation. The group of Keasling utilized the compatible cohesive ends produced by *BglII* and *BamHI* giving rise to different sets of BglBrick vectors.[35] Whereas the BioBrick™ scar sequence encodes an in-frame stop codon, the BglBrick scar sequence encodes Gly-Ser which is suitable for the production of fusion proteins.[31c,35] Other BioBrick™ standard compliant vectors featuring additional restriction sites[36] or regulatory elements[37] were also constructed.

Peng et al. developed a set of vectors compatible with BioBrick™ standards utilizing four isocaudamer pairs (*XbaI*, *SpeI*, *AvrII*, and *NheI*) which were already suggested by Shetty and co-workers.[34] Their ePathBrick vectors support the modular assembly of multi-component pathways in different configurations (Figure 2). In addition, transcriptional fine-tuning is enabled by incorporation of activator and repressor elements in the regulatory regions of the engineered vectors.[30b] The group impressively demonstrated the functionality of a seven-gene flavonoid pathway (~ 9 kb) assembled on one single ePathBrick vector in monocistronic configuration (Figure 2C).

With expression systems such as the ePathBrick vector sets, expression levels can be tuned and the number of plasmids can be dramatically reduced. The vector systems presented so far exclusively feature medium- and high-copy number plasmids (Table 1). It was described by Bailey that stress responses induced by plasmid maintenance are related to the copy number.[38] Silva et al. highlighted the physiological and metabolic alterations in *E. coli* due to the induced stress responses by high-copy number plasmids in their comprehensive review; however, this survey was limited to studies in context of recombinant high-yielding vector production for gene therapy applications and studies were not related to metabolic engineering.[18b] Low-copy number plasmids confer a low metabolic burden. They are more stably maintained and capable of replicating larger pieces of DNA due to their larger native size and the replication mechanism. This makes low-copy number plasmids excellent but disregarded alternatives to medium- or high-copy number plasmids for vector-based metabolic engineering.[8a,39]

The simplification of biology into parts and standardization for custom-assembly in a Lego brick-like

fashion made synthetic biology a powerful tool for metabolic engineering and artificial pathway construction. Nevertheless, the many rounds of molecular cloning and the step-by-step assembly of all pathway elements are time-consuming. Additional problems can occur if a target gene contains one or more recognition sites for restriction enzymes that are already in use. Evidently, other restriction enzymes can be used and codon degeneracy allows certain alterations in a nucleotide sequence, ditching undesired restriction sites without changing the amino acid sequence. But such compromises can be laborious and, as it can be with molecular cloning, neither promise efficiency nor success. Consequently, scientists developed cloning methods that compass iterative restriction enzyme digestion and ligation and are capable of assembling whole metabolic pathways (ideally) in one step.

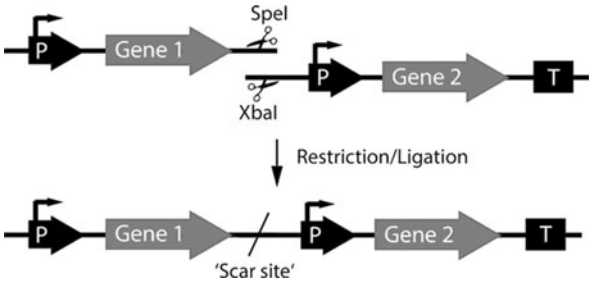

Figure 3. BioBrick™ assembly. Classical molecular cloning depends on overlapping DNA ends created by restriction enzymes that are linked by DNA ligases. Joining of the compatible ends from *SpeI* and *XbaI* digest results in a "scar sequence" that cannot be cut by the original enzymes. Therefore, *SpeI* and *XbaI* can be re-used in the next round of cloning to construct simple pathways, for example.

2.2 Advanced Cloning Techniques

In the last decade, major advances in DNA technologies and bioinformatics have dramatically reduced the costs in commercial DNA synthesis and sequencing.[40] The limiting technology for the construction of larger metabolic pathways with certain architecture and regulatory elements has long been its assembly. This issue seems to be overcome and there are excellent studies describing several DNA assembly methods in detail.[30a,41] In the following section, selected methods will be briefly discussed and applications in metabolic engineering will be presented.

To avoid scar sequences as generated during BioBrick™ (e.g., in-frame stop codon) or BglBrick assemblies (Figure 3) and to circumvent restriction sites that must not be utilized, various polymerase chain reaction (PCR)-based methods have been developed. These related methods depend on overlapping homologous DNA sequences and enable assembly of multiple DNA fragments in a standardized, seamless, and (almost) sequence-independent fashion.

A couple of ligation-independent PCR cloning techniques were applied to assemble linear biological parts that needed subsequent cloning steps for the construction of plasmids.[42] Others conveniently resulted in ready-to-transform plasmids bearing the desired inserts.[43] With the introduction of sequence- and ligation-independent cloning (SLIC), PCR cloning became interesting for the construction of pathways.[44] Since the original publication by Li et al., the SLIC protocol has been improved in various ways.[30a] One advanced SLIC protocol was introduced with FastCloning (FC). The insert and the target vector are separately amplified with overlapping ends (15–30 bases) by PCR. The mixture containing both the amplified vector and the insert is digested with *DpnI* to destroy the DNA templates used for PCR. The mixture is directly transformed into competent *E. coli* cells where DNA strand breaks are sealed by the bacterial DNA repair machinery (Figure 4).[45] Although FC is simple, sequence- and ligation-independent, it

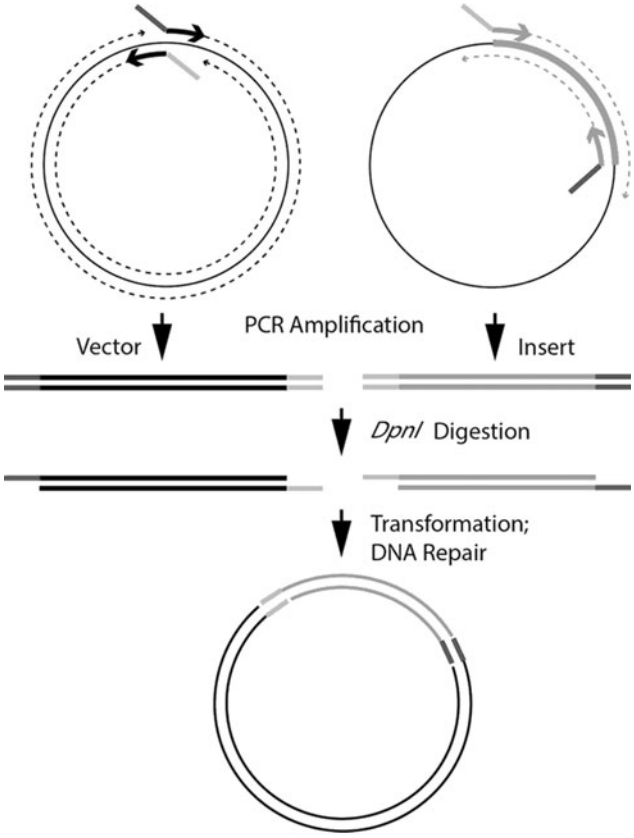

Figure 4. FastCloning (FC). FC is an improved SLIC protocol for the construction of plasmids. Vector and insert are separately amplified by PCR with overlapping primers. Both PCR products are mixed and digested with *DpnI*. The mixture can be directly transformed into competent cells where DNA gaps between the overlapping sequences are repaired *in vivo*.

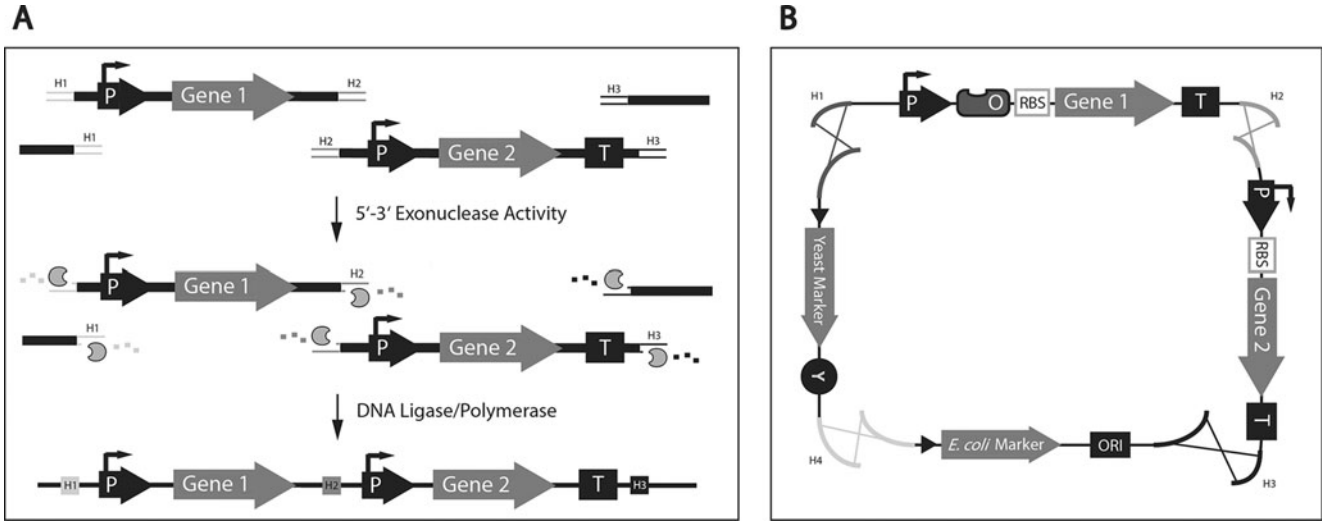

Figure 5. Large scale assembly methods. A) *In vitro* Gibson assembly utilizes the activity of 5′-3′exonuclease that chews back the ends of homologous DNA sequences (H1–H3) creating overlaps. Successively, a polymerase fills the gaps and a DNA ligase seamlessly joins the biological parts. B) Transformation-assissted recombination (TAR) cloning in yeast is also based on overlapping regions (H1–H4) between different biological parts that get efficiently assembled *via* homologous recombination *in vivo*.

has limitations. Many PCR-based overlap assembly approaches were applied to construct plasmids. But scaling up to assemble entire pathways can be difficult. Plasmids become less efficient at larger sizes. The innate error rate of PCRs and the GC content of nucleotide sequences also has to be taken into account when amplifying larger constructs.[30a] Additionally, complex mixtures only ligate with low efficiency in *in vivo* post-transformation.

The "Gibson" isothermal assembly is a very powerful *in vitro* technique omitting forbidden restriction sites and error-prone PCR amplification of long DNA sequences; it is ofted referred to as Gibson assembly (Figure 5A).[46]

It was successfully applied to assemble small circular genomes (~16–580 kb)[46,47] in one step by an enzyme cocktail of a high fidelity DNA polymerase, a T5 exonuclease, and a Taq DNA ligase. Besides its simplicity, short incubation times at one temperature and the parallel assembly of many DNA fragments with matching overlaps are most striking advantages. The elevated temperature (50 °C) can be beneficial to resolve stable secondary DNA structures that interfere with efficient assembly.[30a,46] As Gibson assembly uses three enzymes *in vitro*, it is more expensive than SLIC methods but the ligase may increase assembly efficiency when compared to ligations *in vivo*.

SLiCE (seamless ligation cloning extract) is another *in vitro* assembly method. Introduced by Zhang and co-workers, it utilizes an easy-to-produce bacterial cell extract.[48] Since different laboratory strains can be used as sources for SLiCE, it is a very cheap method for DNA assembly. SLiCE efficiency was improved by applying the PPY strain, an *E. coli* DH10B strain expressing a λ *Red* recombination system.[48]

Assembly methods like SLIC, Gibson assembly, and SLiCE are based on homologous DNA overlaps of various lengths depending on the technique used. This can be a major drawback if single stranded DNA (ssDNA) adopts stable secondary structures such as hairpins or stem loops (e.g., terminator sequences). As ssDNA is required to join the DNA fragments, such secondary structures can hinder assembly. Furthermore, identical homologous sequences must not be used repeatedly as this can result in constructs either not containing all desired DNA fragments or in the wrong configurations. Hence, the design of overlaps can be tedious and laborious especially for elaborate assembly mixtures containing many different biological parts. This renders such methods not truly sequence-independent.

As biocatalysis progressed from rather simple experiments *in vitro* to more and more complex systems *in vivo*, it was only a question of time for assembly methods to follow. Living cells have unmatched capabilities to repair DNA (e.g., double strand breaks; DSBs) by distinct processes involving homologous recombination.[49] Despite their successful demonstration of controlled assembly of large DNA fragments *in vitro*, Gibson et al. adapted a long known cloning protocol for DNA manipulations in *S. cerevisiae* termed transformation-associated recombination (TAR)[50] to assemble the entire circular genome of *Mycoplasma genetalium* (*M. genetalium*) in a single step.[51] The mechanism is, again, based on overlapping sequences that undergo homologous recombina-

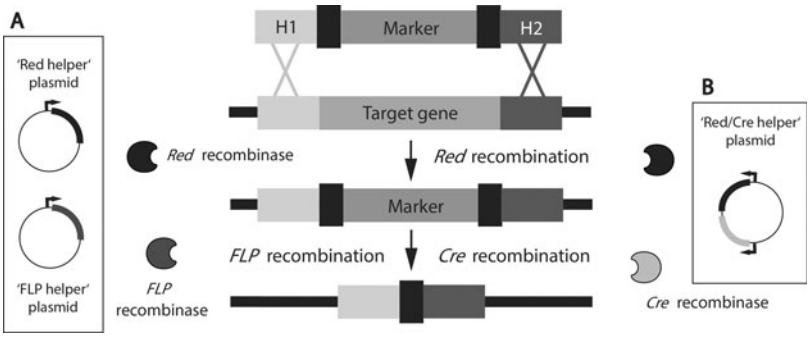

Figure 6. λ *Red* recombination. Target gene disruption is a two-step process. In the first round, λ *Red* recombinase facilitates recombination between homologous sequences (H1-H2) inserting a marker gene (e.g., antibiotic resistance) that is flanked by recognition sites (small black rectangles) for a second recombinase. After selection, the marker gene is excised by recognition of the flanking regions by the second recombinase leaving a "recognition scar sequence". A) The original protocol used two "helper" plasmids that had to be individually transformed for every recombination event. Black rectangles indicate FRT sequences recognized by the *FLP* recombinase that excises the marker gene. B) An improved protocol expressed the two individually inducible recombinases from a single plasmid omitting repeated transformation steps. Black rectangles indicate loxP sites recognized by the *Cre* recombinase. Marker excision leaves a loxP scar.

tion in yeast (Figure 5B). By including a yeast artificial chromosome (YAC) replicon and a suitable selection marker, the final construct can be maintained and easily propagated.

Due to both high fidelity and accuracy of enzymes involved in the homologous recombination and the tolerance of large YACs in *S. cerevisiae*, this method promises to be a reliable tool for pathway construction.[30a] Noteworthy, Shao et al. demonstrated a concurrently developed method called "DNA assembler" for pathway construction in yeast. To prove applicability, they assembled an eight-gene pathway combining D-xylose utilization and zeaxanthin biosynthetic pathway (~19 kb) onto one vector.[52] Furthermore, they constructed shuttle vectors encoding natural pathways for two polyketides, aureothin[53] and spectinabilin,[54] respectively, that were assembled by TAR in *S. cerevisiae* and heterologously expressed in *E. coli*.[55]

With the advanced DNA assembly tools available, single biological parts can now be more efficiently assembled into multi-part pathways even on a genomic scale. Although all of the described methods have their limitations, they form a solid foundation for further improvement of assembly efficiencies by exploiting novel enzymes[56] or polishing established protocols.[57] Future DNA assembly will be highly supported by computational tools to aid design and to lead to automated DNA assembly processes.[30a,58]

Imposed metabolic burden upon pathway introduction in the heterologous host has already been addressed in this review and elsewhere. Since the drain of resources originating from the synthesis of pathway elements is inevitable, the plasmid burden can be reduced by either applying metabolic engineering tools independent of plasmids (e.g., chemically inducible chromosomal evolution; ClChE[59]) or to ultimately integrate *de novo* pathways into the genome.

2.3 Genomic Integration Tools

Random modifications of bacterial genomes by chemical mutation,[60] UV irradiation[61] or transposon mutagenesis[62] have been used for decades. Introduction of undirected (unwanted) mutations into the genome represents a major drawback of these methods. The availability of complete genome information enabled targeted, homologous recombination-based DNA modifications. Such methods utilize the enzymatic activity of *RecA*[63] or certain phage-derived enzymes that facilitate homologous recombination.[41b,58a]

Currently, the λ *Red* recombination is commonly used in many labs and was successfully applied for metabolic pathway engineering.[58a,64] The original protocol was developed by Datsenko and Wanner to disrupt chromosomal genes in *E. coli* (Figure 6A).[65] [As a side note: This approach has been used to construct the Keio collection of single-gene knock-outs (KOs) in *E. coli*.[13]] The method depends on two "helper" plasmids encoding two different recombinases, λ *Red* and *FLP*, respectively. Insertion of an antibiotic resistance marker is facilitated by λ *Red* and homologous sequences targeting the gene to be knocked-out. The marker is additionally flanked by FRT (*FLP* recognition target) sequences. Colonies with the target gene disrupted are selected by their acquired antibiotic resistance and transformed with the second "helper" plasmid. The resistance gene is excised by the *FLP* recombinase. Both helper plasmids have temperature-sensitive replicons which can be easily cured.[65] Desired genes can be inserted into the genome at any loci *via* the flanking target homology regions. However, the efficiency of integration decreases for larger DNA fragments (~1.5–2.5 kb)[65,66] although the successful insertion of even larger fragments (>3 kb) was reported.[67]

Song and Lee modified the method described above and integrated two separately inducible recombinases, λ *Red* and *Cre*, on a single helper plasmid to perform gene KO and marker-excision, respectively (Figure 6B). With this method, the sequential deletion of four genes resulted in a fumaric acid over-producing *E. coli* strain without repeated rounds of helper plasmid transformations and curing.[68]

A more straightforward method employs the highly site-specific recombination machinery of the transposon Tn7. Insertion is readily efficient. The downside of this method is, although conserved in various bacterial species,[69] that insertion is strictly dependent on the occurrence of the attTn7 sequence in the given genome rendering this approach inflexible.[70]

Koma and co-workers combined the λ *Red* recombination, the *FLP*/FRT recombination, and P1 transduction[71] to insert multiple genes into target loci on the *E. coli* chromosome. They successfully incorporated essential genes of the shikimate pathway to accumulate aromatic amino acids. By integrating two heterologous decarboxylases from *Lactobacillus brevis* (*L. brevis*) and *Pseudomonas putida* (*P. putida*), Koma et al. were able to produce the aromatic compounds tyramine and phenethylamine, respectively, This clearly demonstrated that their method is a useful tool to integrate functional metabolic pathways in *E. coli*.[72] Furthermore, Koma et al. observed differences in the activity of single-copy insertions of a reporter gene *lacZ* (encoding β-galactosidase) depending on the insertion loci. Although the deleted genes were not essential, they might directly or indirectly influence the expression of other genes important to protein expression.[72] Another phenomenon termed context dependency states that many DNA elements are influenced by adjacent sequences or even distant ones.[30a,64b,67c,72] RBS, for example, perfectly illustrates context dependency. Although the core sequence of an RBS is only 6 bases long, it is absolutely required for translation initiation.[58d,73] RBS is always located before an open reading frame (ORF). In addition, flanking sequences (~50 bases) around the RBS modulate its efficiency.[30a,58d,74]

Sabri et al. also addressed context dependency and evaluated the expression levels of a reporter gene (*xynA*) for three non-essential loci (*arsB*, *lacZ*, and *rbsA-rbsR*) in two different *E. coli* strains in the course of their KIKO vector construction. Target genes can be conventionally cloned into a multiple cloning site (MCS). Insertion cassettes in the KIKO vectors are flanked by hairpin loops to isolate them from neighboring DNA elements at the insertion site. Genomic integration is mediated by λ *Red* recombinase *via* long homology arms (500 bases) to increase insertion efficiency.[44] Antibiotic resistance markers are removed by *FLP*/FRT recombination.

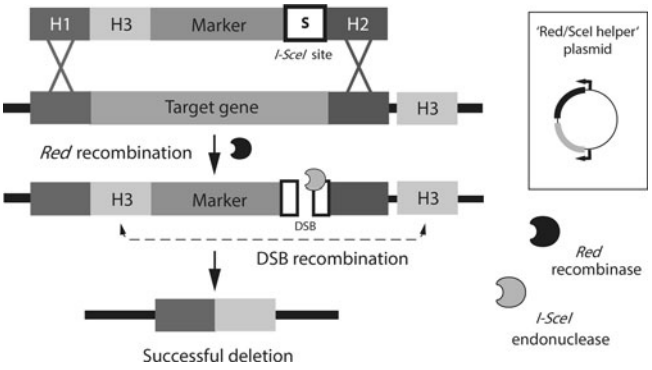

Figure 7. Efficient recombination by *I-SceI*-mediated double strand breakage. The *Red* recombinase and the *I-SceI* endonuclease are co-expressed from one plasmid. The endonuclease introduces a DSB at the *I-SceI* recognition site (S) that enhances recombination efficiency.

Both recombinases are encoded on the vector. Sabri et al. achieved the insertion of 5.4 kb at a single site making the KIKO vector set a very useful tool for the integration of multi-gene pathways in one go if flexibility of location is less important. Noteworthy, iterative rounds of KI resulted in unintended effects (e.g., deletions, rearrangements) as stated by Sabri and co-workers. These resulted from FRT scars remaining in the genome from previous integration events. Strains containing rearrangements or deletions have to be excluded by carefully analyzing the integration junctions *via* PCR.[67c]

Two-step techniques that combine the λ *Red* system with the yeast mitochondrial homing endonuclease *I-SceI* have been developed for introducing large DNA fragments onto the *E. coli* chromosome.[66,75] The recognition site of *I-SceI* is rather large (18 bases) for an endonuclease and is not found in the *E. coli* genome. Since DNA DSBs stimulate *in vivo* recombination,[76] the recombination efficiency can be greatly increased by using *I-SceI* sites at the target locus (Figure 7).

I-SceI sites were utilized within two studies: one prior to integration to improve cloning of large DNA fragments into the CRIM plasmid which bears attP sites that subsequently recombine with ϕ80-attB sites on the target chromosome.[75,77] This approach avoided the use of PCR cloning and allowed successful integration of an 8 kb fragment. The second method needs spadework as a "landing patch" has to be integrated first onto the chromosome. The λ *Red* recombinase and *I-SceI* facilitate anchoring of the desired integration cassette at the "landing site".[66] The utility of this method was demonstrated by the introduction of a 7 kb fragment at six different loci.[67c] Such two-step techniques provide convenient tools for location-independent insertion of very large DNA fragments onto the *E. coli* chromosome for applications in metabolic engineering and synthetic biology.

One very recent development revisited the utilization of transposons for genomic integration. The group of Nikel designed a set of mini-Tn5 delivery vectors termed pBAMD1-*x* that have a broad host range for Gram-negative bacteria (e.g., *P. putida*, *E. coli*), show good insertion efficiencies for large DNA fragments and can be re-used to target the same chromosome; however, this is achieved at the expense of directed site-specific insertion, as it can be encountered with transposable elements. Martínez-García et al. inserted the entire poly-(3-hydroxybutyrate) (PHB) biosynthetic pathway from *Cupriavidus necator* in the *E. coli* chromosome yielding up to 40% of PHB content of dry cell weight post-insertion.[78]

The aforementioned methods are capable of both disrupting target genes and, if homologous sequences are provided, inserting desired genes, pathway elements or whole metabolic pathways at the very same spot in the genome. Various groups took advantage of these techniques and incorporated heterologous pathways in the genome and simultaneously knocking-out genes that interfered with the *de novo* pathway. Thereby, the enzymatic host background was reduced and product titers were increased.

Baumgärtner et al. chromosomally integrated a 2′-fucosyllactose (2′-FL) pathway in *E. coli*. To prevent the intracellular degradation of L-fucose, the expression cassette was inserted into the region coding for two degrading enzymes, *fucI* and *fucK*. Additionally, bottlenecks for 2′-FL production were identified and overcome by providing another copy of the *futC* gene generating 2′FL from GDP-L-fucose and lactose.[79]

Kunjapur et al. constructed an *E. coli* strain termed the RARE (reduced aromatic aldehyde reduction) strain by deleting up to six genes with reported activity on their model substrate benzaldehyde. They used the engineered strain to heterologously produce vanillin and L-phenylacetylcarbinol upon expression of a carboxylic acid reductase (*CAR*) providing a model microbial cell factory for the production of valuable compounds such as aromatic aldehydes (see Figure 26).[11]

In a very recent review, Song, Lee and Lee highlighted genome engineering tools for bacterial strain development.[41b] They presented chromosomal integration tools mainly used in eukaryotes rather than bacteria.[80] Additionally, they compiled genomic integration mechanisms different from the ones described above but offering alternatives for genome editing in bacteria recalcitrant to traditional engineering tools (e.g., mobile group II intron-mediated genome engineering[81]).[41b]

Genomic integration tools are getting more versatile, reliable, and efficient. Entire pathways can be anchored in the genome turning simple strains into potent microbial cell factories. Whereas the metabolic load is reduced and pathway performance could be improved, for example, in terms of productivity,[19] context dependency in the chromosomal environment becomes a future challenge for metabolic engineering and synthetic biology.

3 Optimization of *de novo* Pathway Elements

In the case of a metabolic pathway, usually several genes have to be expressed each encoding a pathway enzyme. As pointed out earlier, the coordinated and balanced production of all pathway components is essential not only to reduce the metabolic burden from protein over-production but to optimize the flux through the metabolic pathway.[8a,19a,31c] As generally known, proteins are produced by cellular machineries *via* two interconnected processes: transcription and translation. Both can be finely modulated but also properties of pathway enzymes can be altered to meet the requirements of a synthetic pathway. The main regulatory element of transcription is certainly the promoter. Translation is strictly dependent on a functional RBS and influenced by the stability of mRNA.[30c] Recently, regulatory elements that can act in *cis*/*trans* were introduced for the regulation of gene expression in bacteria: small regulatory RNAs (sRNAs).[82] Also self-regulatory elements can be exploited as metabolic engineering tools (e.g., auto-inducers) but might resemble obstacles in biotechnology processes (e.g., feedback inhibition).[30c] This section will give an updated overview of optimization strategies for metabolic engineering on the levels of transcription and translation, as well as present examples for engineered enzymes involved in redox cascades to produce valuable compounds. Additionally, scaffolding has positively influenced the flux through *de novo* pathways towards desired products.

3.1 On the Levels of Transcription and Translation

Enzymes are the key players in metabolic networks and each biosynthetic pathway. On the DNA level, genes encode these enzymes. Besides the coding sequence, a gene can include many regulatory elements (Figure 8).

Manipulating regulatory elements (e.g., promoter, operator, RBS, terminator) will alter the levels of expression and is one main tool for pathway optimization.

Transcription is the first step of gene expression. The promoter DNA sequence is recognized by a group of proteins responsible for the initiation of transcription; it can also serve as very economical starting point for regulation. For the heterologous ex-

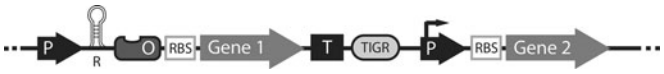

Figure 8. Possible elements of a protein-coding sequence. Gene 1 is under the control of multiple regulatory elements: a promoter (P), a riboswitch (R), an operator (O), and a terminator (T). Located between gene 1 and 2 is a tunable intergenic region (TIGR) that may influence downstream regulatory elements such as the RBS.

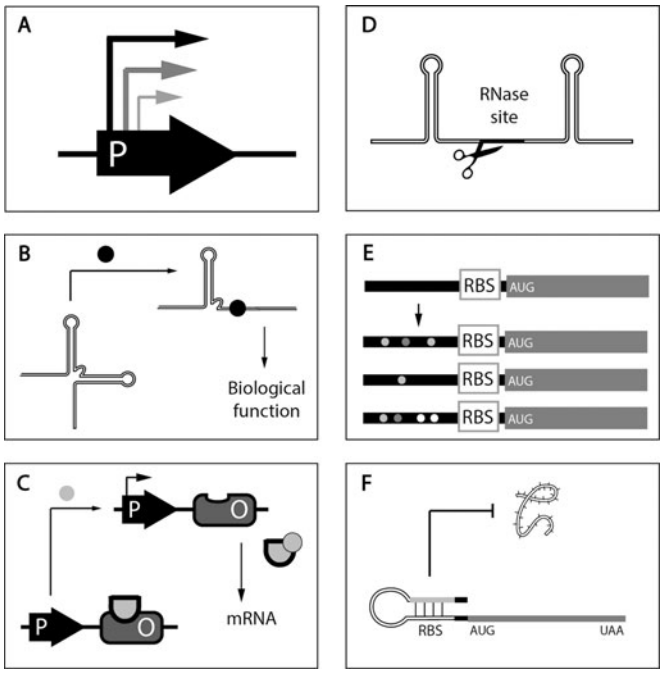

Figure 9. Regulatory mechanisms. A) Promoters of different strengths influence the transcription rate and, consequently, the amounts of protein produced. B) Riboswitches are structural mRNA elements that can bind small molecules. Conformational changes affect the biological function of the mRNA molecule. C) An operator region downstream of a promoter recruits inhibitory proteins that block a proceeding DNA polymerase. Small molecules (i.e., the inducers) bind to these proteins allowing transcription to continue. D) The introduction of RNase sites directly influences the half-lives of mRNA molecules and subsequent translation. E) Translation efficiency can be manipulated by sequence changes close to the RBS. F) Sequestration of the RBS can block translation.

pression of pathway enzymes, native promoters are often replaced by inducible promoters. Promoter strength has profound influence on the amount of enzyme produced and, thus, on the flux through the *de novo* pathway (Figure 9A).[8a] P_{lac}, P_{BAD}, and P_{tac} are frequently used strong, inducible promoters (Table 2). Makrides and Hannig gave an overview about promoters used for high-level expression of genes in *E. coli*.[20b] Brautaset et al. compared an expanded set of promotors in bacterial expression systems and their applications. This survey included promoters not only regulated by the presence of sugars (e.g., L-arabinose, L-rhamnose, and xylose) but also by organic compounds such as different alkanes, substituted benzenes or even peptides.[83] Recently, Balzer et al. re-examined commonly used promoter systems in *E. coli* (*XylS*/P_m, *LacI*/P_{T7lac}, *LacI*/P_{tac}, and *AraC*/P_{BAD}) and conducted comparative expression studies.[84] *LacI*/P_{T7lac} turned out as highly advantageous at a transcriptional level. Contradictory, on the level of translational, transcripts were often translated into vast amounts of inactive (insoluble) proteins.[84] This is long known for the popular pET system which also utilizes the *LacI*/P_{T7lac} promoter.[85] Gene expression is induced by IPTG, a lactose analogue that is not readily metabolized granting prolonged high expression levels. Promoters for metabolic engineering must show tight control to avoid undesired metabolic load from leaky expression which is the case with the T7 expression system.[8a,86] Tightly controlled systems are based on promoters like the arabinose-induced P_{BAD}[12b] or the more recently introduced cumate gene switch-based expression system[86b] and may be more suitable for metabolic engineering (Table 2).[86b]

Table 2. Seclected small-molecule inducible promoter systems.

Promoter System	Inducer	Features[a]
LacI/P_{T7Lac}	IPTG[b]	−/+
LacI/P_{tac}	IPTG[b]	−/+
AraC/P_{BAD}	L-arabinose	+/+
RhaR-RhaS/rha_{BAD}	L-rhamnose	+/+
P_{xylA}	xylose	+/−
XylS/P_m	[c]	+/+
XylR/P_u	[d]	+/+
CymR/P_{cum}	cumate	+/+
ChnR/P_b	cyclohexanone	+/+

[a] Promoter control/expression levels categorized in: leaky (−) and tight (+)/low (−) and high (+).
[b] IPTG as a lactose analogue.
[c] Multiple inducers such as substituted benzenes (e.g., 3-methylbenzoic acid).
[d] Multiple inducers like substituted benzenes but differently regulated than *XylS*/P_m.

Auto-responsive promoters are not induced by the addition of certain compounds, unlike conventional promoters, and represent promising alternatives. They respond to environmental stimuli (e.g., oxygen[87] or light[88]) or metabolites produced during the growth of the microorganism.[30c] Stress-response promoters belong to the latter.[89] A process in bacteria called "*quorum sensing*" coordinates the expression of genes and, consequently, the behavior of whole cell populations. It is triggered by specific signal molecules (i.e., auto-inducers) whose concentration is proportional to the cell density, for example.[90] Tsao and co-workers

successfully tested this system and heterologously expressed different model enzymes (e.g., *LacZ*).[91] As the expression of target genes only occurs at a certain cell density threshold, the optimization of induction initiation can be avoided eliminating continual monitoring of bacterial growth prior to induction.[30c] New promoter systems promise to be powerful tools for coordinated gene expression in metabolic engineering as well as in synthetic biology and engineering of known promoters systems further boosts diversity of tuning gene expression at a fundamental level.[92]

The repertoire of transcriptional control elements can be extended to riboswitches, operators and even intergenic regions (Figure 8). A riboswitch is part of the mRNA molecule that can directly bind a small target molecule which affects the activity of the very gene containing the riboswitch in a concentration-dependent manner (Figure 9B).[93]

The function of an operator region placed in front of a gene, operon or pseudo-operon (Figure 2) is extensively used in the form of the *lac* operator (e.g., in the pET expression system[85b]). The operator region can control the transcription of downstream genes by binding the repressor protein *LacI* which blocks the proceeding of the RNA polymerase (Figure 9C).[94] However, other regulatory elements take precedence over the use of operator sequences and literature examples related to metabolic engineering are rare.[30c]

Intergenic regions occur in genomes between coding sequences (Figure 8). The majority of the human genome is considered intergenic;[95] bacterial genomes are generally smaller and, therefore, have a different organization.[96] Despite the extensive existence of overlapping genes, intergenic regions also exist in organisms such as bacteria and fungi.[97]

Intergenic regions can direct the cleavage of transcripts by encoded RNase sites or stabilize mRNA secondary structures directly influencing translation (Figure 9D).[98] Pfleger et al. assessed the expression levels of multiple genes in operon form (Figure 2A) by using post-transcriptional control elements and tunable intergenic regions. Balancing resulted in a 7-fold productivity increase of a heterologous biosynthetic mevalonate pathway.[30c,99]

While promoters have been thoroughly investigated and used as regulatory element, transcriptional terminators apparently have not.[30c] The function of transcription terminators is simply to stop transcription which was reviewed by Henkin (Figure 8).[100] In addition to this primary function, transcriptional terminators were found to stabilize their own mRNA.[101] Engineering transcriptional terminators may therefore provide another interesting tool for regulating both transcription and translation. The longer the half-life of an mRNA molecule is, the more often it can be translated. Bernstein et al. studied mRNA decay and abundance in *E. coli* at single-gene resolution and stated that mRNA stability is related to both the number of transcripts and the function of a gene product.[102] Apart from stabilizing mRNA molecules, a more important strategy to control translation is to target the RBS (Figure 8). The recognition of the RBS by the ribosome is absolutely necessary for protein synthesis. The RBS [the Shine-Dalgarno (SD) sequence in bacteria] is a short sequence upstream of the start codon (AUG in bacteria; coding a methionine).[73,103] As already mentioned, the sequence around the RBS (~50 bases) modulates translation efficiency (Figure 9E).[30a,58d,74]

Furthermore, stable secondary structures of the RBS or close to it can drastically reduce translation rates or even prevent translation (Figure 9F).[104]

Apparently, variations of the RBS can reduce or increase translation in bacteria. Tools are available to design synthetic RBSs.[58d,105] Wang et al. invented a method called multiplex automated genome engineering (MAGE) which was applied to engineer a recombinant *E. coli* strain for lycopene production.[41b,106] Targeting the 20 genes responsible for lycopene synthesis, the RBS regions were modified through allelic replacements using oligo-nucleotides containing degenerated RBS sequences (DDRRRRRDDDD; D = A, G, T; R = A, G). High similarity between the replaced RBS region and the canonical SD sequence (TAAGGAGGT) gave rise to enhanced translation efficiencies.[30c,106]

The insufficient translations of several enzymes in a biosynthetic mevalonate-to-amorphadiene pathway lead to the accumulation of intermediates. Amorphadiene is a precursor for the anti-malarial drug artemisinin.[107] To overcome these bottlenecks, Nowroozi et al. applied a combinatorial approach to screen for suitable RBSs for different cascade enzymes. Upon testing various combinations and relating them to growth, protein levels, and accumulation of intermediates, this optimization improved the production of amorphadiene (**8**) about 5-fold. (Figure 10).[108]

Codon usage also influences the translation efficiency and the availability of a tRNA corresponding to its codon on the mRNA depends on the species. The use of codon-optimized genes is interesting in the context of heterologous pathway design when enzymes from different (especially higher) species are used (and produced in a lower host organism).[109]

Another RNA-based post-transcriptional/translational regulation strategy in bacteria is the use of riboregulators or sRNAs. sRNAs belong to a small subset in the group of non-coding RNAs in prokaryotes and eukaryotes and have various structural, regulatory and enzymatic functions.[110] Non-coding RNAs can form a portion of an mRNA molecule as it is the case with riboswitches. sRNAs, too, can respond to chemicals[111] and environmental signals (e.g., temperature[112]) and relay them to regulate gene expres-

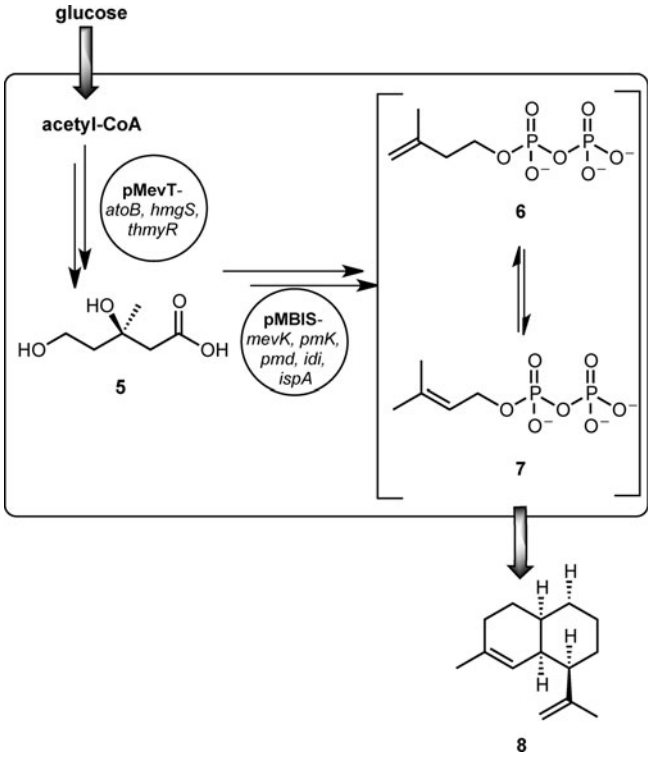

Figure 10. Key intermediates in the mevalonate-to-amorphadiene pathway. The upper and the lower mevalonate pathway, pMevT and pMBIS, respectively, produce the key intermediates isopentenyl pyrophosphate (IPP) (**6**) and dimethylallyl pyrophosphate (DMAPP) (**7**) in a synthetic amorphadiene pathway, converting mevalonic acid (**5**). Tuning of the RBSs overcame the issue of unbalanced enzyme production and the accumulation of (toxic) intermediates.

sion.[113] Translation can be either repressed or activated (Figure 11).[114]

Sets of riboregulators were designed and successfully applied[115] and Kang et al. used an artificially overexpressed sRNA (*RhyB*) for metabolic engineering in an arabinose-inducible expression system to accumulate succinate (and acetate) in *E. coli*.[116] Previously, they constructed a stress-induced sRNA-based system to produce polyhydroxyalkanoates.[117]

Recently, the group of Lee designed arrays of synthetic sRNAs to knock-down target genes and to increase the production of tyrosine and cadaverine in engineered *E. coli* strains.[82] In the case of the overproduction of tyrosine, Na et al. achieved titers ($2\,g\,L^{-1}$) with their sRNA approach as high as Juminaga et al. who used a plasmid-based expression system including promoter and operon engineering.[118]

The utilization of synthetic sRNAs offers future perspectives as alternatives for conventional gene KO strategies. KOs cannot be easily undone whereas sRNAs are easy to implement and gene knock-down usually is reversible. Furthermore, the introduction of sRNAs only conveys a minimal metabolic burden due to their natively small sizes. Moreover, sRNAs can be applied to simultaneously tune the expression levels of various target genes allowing gene-to-function studies of essential genes that cannot be deleted.[41b]

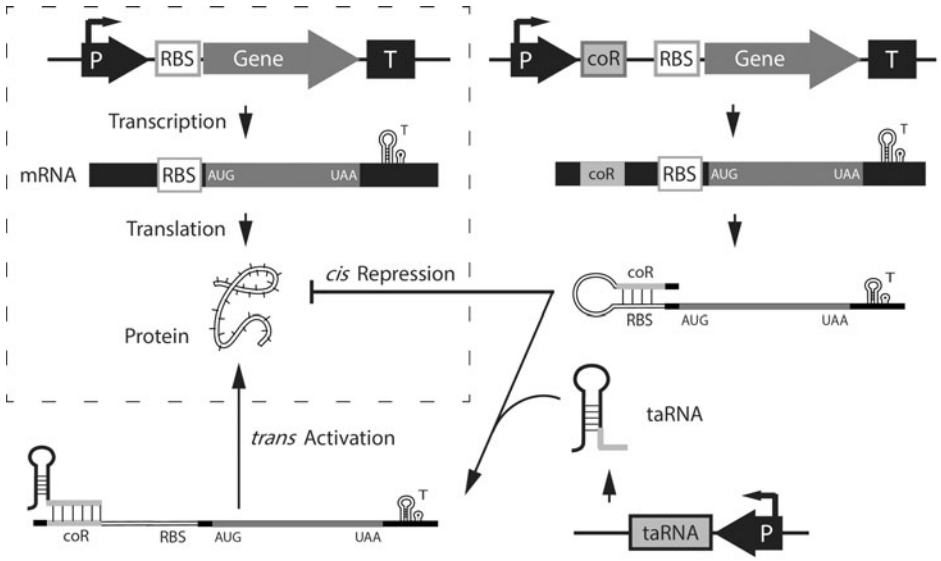

Figure 11. Post-transcriptional regulation of gene expression by a riboregulator. By including a DNA element coding for a sequence complementary to the RBS (coR) upstream of the RBS, translation can be blocked (*cis* repression). Expression of a short DNA molecule complementary to coR in *trans* (taRNA) resolves the secondary structure of the RBS and allows translation (*trans* activation).

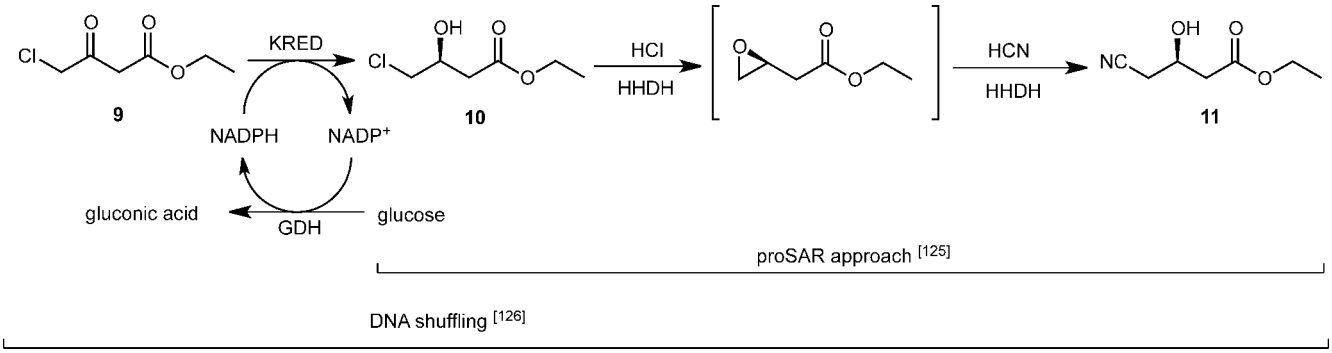

Figure 12. Improvement of the atorvastatin precursor synthesis by directed evolution. A combined approach of directed evolution and proSAR greatly increased the HHDH activity leading to the key precursor (**11**) in the atorvastatin synthesis. Another approach using a DNA shuffling methodology not only improved the HHDH activity but included a KRED and a GDH for cofactor recycling leading to a "green-by-design" set-up for industrial application.

3.2 Improving Enzyme Performance

The previous chapters dealt with sophisticated tools for pathway assembly and strategies to optimize introduced pathways on different molecular levels. In this part of the review we focus on the biocatalyst itself and the optimization tools that are available. Altering the sequence of regulatory elements can enhance or decrease the expression of associated genes and change the amounts of protein produced. Despite the many successful applications of enzymes in biocatalysis,[1b,40b,119] enzymes regularly failed to meet industrial process criteria (e.g., stereo- and regioselectivity, pH and thermostability, limited substrate scope).[119a] To overcome these shortcomings, the protein-coding sequence itself can be mutated which will directly influence the enzyme's properties. Two concepts have been extensively used: directed evolution[120] and rational design.[121]

Directed evolution essentially mimics the process of evolution as it takes place in nature but at a much higher pace.[122] Other than directed evolution, rational design largely depends on the availability of structure-function relationship of the target protein. Simple approaches are based on sequence homology comparisons and aim for the improvement of certain properties of the protein of interest (e.g., thermostability, solubility, organic solvent tolerance).[30c,123] Protein engineering by directed evolution involves iterative cycles of gene mutagenesis, expression, and selection of mutant enzymes. Commonly used mutagenesis techniques include error-prone PCR (epPCR), saturation mutagenesis, combinatorial active-site saturation test (CASTing) and DNA shuffling.[119a] Many examples from the literature highlight the potential of directed evolution and rational design in the contexts of biocatalysis and metabolic engineering and were recently reviewed by Bornscheuer et al.,[40b] Reetz,[119a] Otte and Hauer,[124] and others.

Many of these approaches were employed *en route* to a commercially relevant biocatalytic process for ethyl (*R*)-4-cyano-3-hydroxybutyrate (**11**); (Figure 12), a key intermediate in the synthesis of atorvastatin. Atorvastatin is a cholesterol-lowering drug sold under the brand name Lipitor® with a sales volume of >10 billion US $ in 2011. The key enzyme towards (**11**) is a halohydrin dehalogenase (HHDH) from *Agrobacterium radiobacter* (Figure 12). Fox et al. combined a recombination-based directed evolution approach and a strategy for statistical analysis of protein sequence activity relationships (proSAR). Their hybrid approach massively improved the cyanation capability of the HHDH under process conditions increasing productivity ~4 000-fold.[125]

Ma and co-workers successfully executed a "green-by-design" process yielding the same atorvastatin intermediate (**11**); (Figure 12). Utilizing a DNA shuffling technology,[126] they were able to improve the activity of the HHDH by >2,500-fold compared to the wild-type enzyme. In addition, they applied a ketoreductase (KRED) to reduce ethyl 4-chloroacetoacetate (**9**) to ethyl (*S*)-4-chloro-3-hydroxybutyrate (**10**) which was subsequently transformed by HHDH into (**11**). A glucose dehydrogenase (GDH) was applied to recycle the cofactor NADPH (Figure 12). Through several rounds of DNA shuffling, the GDH activity was improved 13-fold and the KRED activity by a factor of 7.[127]

By combining *in silico* design and various protein engineering technologies, Savile et al. equipped a transaminase lacking activity towards the prositagliptin ketone (**13**), the precursor for a type II diabetes drug,[128] with high activity for the ketone (Figure 13A). By additionally optimizing the transaminase towards the process parameters (i.e., tolerance of DMSO, acetone and *i*-PrNH_2 at elevated temperature), process productivity was increased by 53% applying the mutant transaminase.[129]

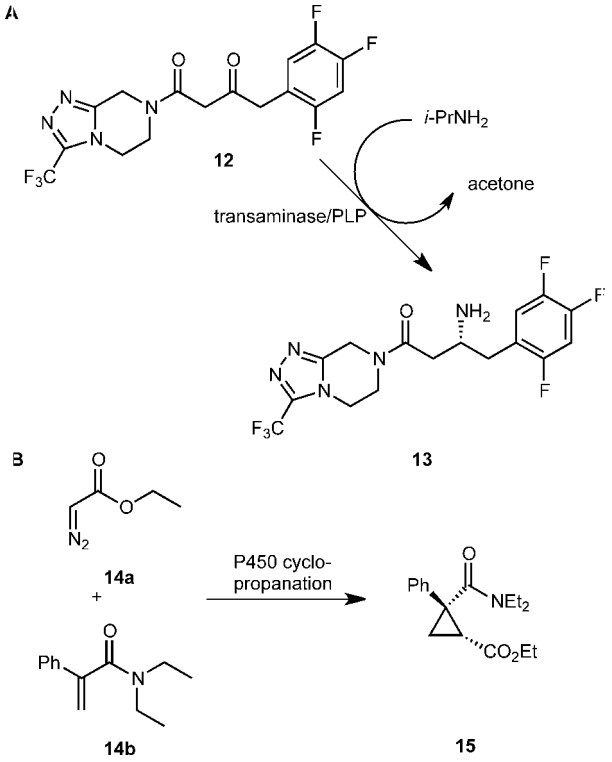

Figure 13. Directed evolution. New features for known enzymes. A) Protein engineering resulted in a transaminase accepting a sterically demanding substrate (**12**) and fitting process requirements necessary for the sitagliptin synthesis. B) A P450-BM3 mutant with increased cyclopropanation activity for the synthesis of levomilnacipran (**15**).

Agudo and Reetz devised a redox cascade comprised of two successive regioselective oxidations performed by a cytochrome P450 enzyme from *Bacillus megaterium* (*P450-BM3*), followed by stereoselective olefin reduction catalyzed by (*R*)- or (*S*)-selective mutants of the enoate reductase *YqjM*.[7a]

Both enzymes, *P450-BM3* and *YqjM*, were subjected to directed evolution to obtain P450-BM3 mutants with increased activity towards the substrate 1-cyclohexenecarboxylic acid methyl ester and *YqjM* mutants providing the proper stereoselectivity. This example is interesting in many regards and will be discussed in detail in the last chapter (see Figure 25).

The group of Arnold engineered the same wild-type *P450-BM3* by site-saturation mutagenesis towards enhanced cyclopropanation activity, a new function required for the enantioselective synthesis of levomilnacipran (**15**),[130] an antidepressant sold as Fetizma® (Figure 13B).

Zhang et al. also demonstrated the possible alteration of the *P450-BM3* scaffold to fine-tune the acceptance of other demanding substrates (**14a**, **14b**).[131]

Finally, another recent example of directed evolution utilizing saturation mutagenesis has been applied to reverse the enantioselectivity of a phenylacetone monooxygenase (PAMO) in the asymmetric sulfoxidation of prochiral thioethers. The four single point mutations synergistically turn the wild-type enzyme with (*S*)-preference for sulfoxide formation (90% *ee*) into the mutant PAMO with (*R*)-preference (95% *ee*).[132] These selected examples undeniably demonstrate the power of state-of-the-art enzyme engineering technologies and how far biocatalysis, especially redox biocatalysis, has come and evolved in the last decades.

3.3 Assembling the Cast: Scaffolding

In nature, scaffolding is used for the spatial organization[133] of metabolic pathways to control substrate channeling, cross-talk between interacting enzymes and to increase the efficiency of a multi-enzyme cascade. For *de novo* cascades, scaffolding provides an attractive strategy to control substrate flow/productivity and minimize intrinsic activity of the host organism. On the other hand, these structural elements are also heterologous to the host organism and can contribute to the metabolic burden. Preliminary design rules to control the number and orientation of enzymes in a spatially nanostructured scaffold system were clustered by Lin et al. in three aspects: (i) inter-enzyme distance, (ii) active site orientation, and (iii) multi-enzyme architecture.[134]

During the last years, various innovative examples to create biomolecular nanostructures have been reported. These structures can be formed by protein–protein interactions[135] as well as nucleic acids (DNA,[136] RNA[137]) or polymers (e.g., cell mimicking polymersomes[138]) (Figure 14).

The group of Keasling exemplified this approach by constructing a heterologous mevalonate pathway which contains hydroxyl-methylglutaryl-CoA synthase (*HMGS*) and hydroxyl-methylglutaryl-CoA reductase (*HMGR*) from *S. cerevisiae*) together with the endogeneous acetoactetyl-CoA transferase (*AtoB*) in *E. coli*.[10] These enzymes displayed different activities and the pathway flux was terminated at high cellular concentrations of the toxic intermediate HMG-CoA.[139] This bottleneck was circumvented by establishing three synthetic protein scaffolds from metazoan genomes: the GTPase binding domain *GBD*, the *SH3* domain and the PSD95/DlgA/Zo-1 (*PDZ*) which are tagged to the cognate peptide ligands of the pathway enzymes. In comparison to the non-spatially organized pathways, the scaffolded mevalonate pathway operated at a 77-fold higher product titer underscoring the potential of this particular pathway tuning strategy.

Another example was presented by Wilner et al.[140] using a self-assembly DNA scaffold of single-stranded nucleic acids, tethering three different nucleic aci-

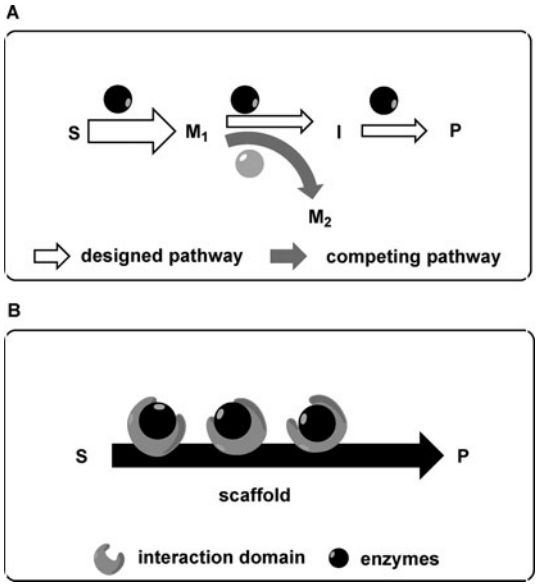

Figure 14. Illustration of a non-structured and a spatial organized enzyme cascade. A) Without metabolic channelling, substrate flux can be interrupted and potential toxic metabolites (M_1, M_2) generated. B) Scaffolds equipped with interaction domains which allow modular docking of enzymes to control product formation and balanced protein production.

functionalized enzymes and the cofactor NAD^+ equipped with the bi-functional cross-linker *N*-[(ε-maleimidocapropyloxy)sulfo-succinimide ester]-(Sulfo-EMCS). This regulation machinery (hexagon-type DNA scaffold) was necessary to integrate biocatalytic activity in their mini pathway, which was corroborated by control experiments without spatial organization. Another strategy to establish metabolic channeling elements in living *E. coli* cells was described by Sachdeva et al. using designed RNA scaffolds.[141] These RNA-binding domains (aptamers) were fused with two modified enzymes (N-termini), the acyl-ACP reductase (AAR) and an aldehyde deformylation oxygenase (ADO), for the synthesis of pentadecane. This *in vivo* channeling system improved production levels by 2.4-fold and intrinsic aldehyde reductase activity was inhibited. Furthermore, this strategy was used to enhance the substrate flux through the succinate pathway up to 88% and represents a useful toolkit for the optimizations of artificial pathways.

Recently, Chen and co-workers summarized various channeling strategies (e.g., cohesion/docking interaction modules, fusion enzymes, MAPK scaffolds) in their review[142] and discussed their applications for *in vivo* or *in vitro* multi-enzyme cascades co-expressed in different host organisms (Table 3).

In general, spatial organization of multi-enzyme pathways can be (highly) beneficial for improving product titers by adjusting stoichiometric ratios between pathway enzymes. Hence, substrate channeling can avoid the accumulation of toxic or unstable metabolites.[147]

Table 3. Scaffolds for *in vivo* cell factories.

Strategy	Interaction domain	Reference
chimeric adapters	protein	[143]
fusion enzymes	protein	[144]
GBD-PDZ-SH3 scaffold	protein	[10]
MAPK scaffold	protein	[145]
two-component systems	protein	[146]

4 Redox Cascades in Cellular Contexts

As already pointed out, redox cascades in whole-cell biocatalysts have to overcome several obstacles on different levels. In general, the tight interaction between the newly introduced pathway and the metabolism of the host on the genetic as well as on the metabolic level may result in increased probability of side reactions, competition for metabolites (e.g., redox cofactors), toxicity of pathway intermediates and, therefore, decreased overall productivity.[3] As another objective insufficient uptake of substrate and product release from the cell becomes an issue which can be addressed by reaction engineering. In the following chapter, such challenges will be discussed and methodological background as well as optimization strategies to overcome these obstacles will be outlined.

4.1 Productivity Enhancement by *in vivo* and *in silico* Strategies

Several proof-of-concept studies including novel biosynthetic routes have been established in the last decades. A major challenge at this stage of development is further progress beyond lab scales in order to demonstrate the potential for industrial processes. Low overall productivity represents a significant obstacle at present which needs to be overcome to enable an industrially relevant and profitable exploitation.

In general, there are two different strategies, a knowledge-based classical *in vivo* and a more holistic *in silico* approach. *In vivo* methods target specific side reactions by obvious gene KOs based on known metabolic network structures in order to rewire the flux through the synthetic pathway with the aim to improve overall productivity. The computational approach as second strategy requires a holistic view of the whole-cell biocatalyst as a system. To be able to do so, the microorganism is modelled as a set of reactions which correspond to the metabolic pathways within the cell.

The classical approach was established due to historical developments as the first microorganisms were used as wild-type strains and their natural ability for production of certain chemicals was exploited. Fermentative production of ethanol by the yeast *S. cerevisiae* represents the most prominent example.[148] Organic acids of the tricarbonic acid (TCA) cycle like citrate or succinate are mostly produced on the industrial scale by different microorganisms. For example, *Aspergillus niger* is used for citric acid, *Lactobacillus rhamnosus* for lactic acid, and *E. coli* for succinic acid production, respectively.[149] With increasing knowledge about the biochemistry and the network structures of different organisms, productivity improvements were mainly based on both rational KO/KI strategies to remove unwanted side reactions in combination with adjustment of target protein expression (e.g., engineered *E. coli* for shikimate production which led to the production of 14.6 $g L^{-1}$ shikimate with a yield of 0.29 $g g^{-1}$ glucose in a 7-L bioreactor[150]).

A lot of classical metabolic engineering approaches were performed for improved butanol production by *Clostridia* strains since this alcohol is an important bulk chemical. These improvements were summarized elsewhere.[151] Medium chain methyl ketones (MKs), another example for bulk chemicals of relevance to biofuel and flavor-and-fragrance industries were published lately by the group of Keasling. Titers of 3.4 $g L^{-1}$ were reported after a 45 h fermentation which, so far, has been the best for MKs. Modifications included balancing enzyme overexpression to increase fatty acid flux, consolidation of the pathway from two plasmids into one, codon optimization, and KOs for flux re-routing.[152]

In the field of biocatalysis, especially in redox processes, whole-cell transformations were extensively used in the past.[119a] The living organism provides a convenient and cheap cofactor recycling system which facilitates the whole process and reduces costs significantly. Overall, whole-cell redox biocatalysis is applied in many different research areas, for example, for the synthesis of various pharmaceutical compounds with different reductases or oxidases.[153] With the extension from single enzyme reactions to the construction of multi-enzyme mini-pathways, attention focused on the host itself, potential interfering side reactions, and metabolic bottlenecks. In contrast to many fermentative processes in the metabolic engineering area, the complexity of multi-enzyme biocatalysis is significantly lower because of the combination of metabolically unrelated enzymes. Nevertheless, very recently, Oberleitner et al. combined *in vitro* and *in vivo* experiments of an artificial mini-cascade consisting of three redox enzymes to improve the productivity by identifying a competing side reaction and deleting of the responsible gene in *E. coli*.[7b] The native enoate reductase *NemA* was removed from the host and different heterologously expressed EREDs were introduced with a distinct substrate profile and stereopreference. Additionally, the group of Reetz targeted the same background reaction (*NemA*) to establish another evolutionary non-related mini-pathway in *E. coli* (see Figure 25).[7a] In another study, *E. coli* was engineered to produce riboflavin (vitamin B_2) by introduction of two dehydrogenases from *Corynebacterium glutamicum* and deletion of competing reactions. This strain has the highest yield amongst all reported riboflavin production strains (2.7 $g L^{-1}$ in a shake flask culture) which is a 12-fold increase compared to the basic producer strain of riboflavin.[154] As mentioned before, the Prather group also enhanced the performance of a whole-cell biocatalyst by knocking-out different aldehyde reductases in their *CAR* expressing strain in order to produce vanillin from vanillate.[11]

On the other side, it became more and more common to investigate systems *in silico* beforehand and to conduct selected experiments based on computational simulations. This is a holistic while complex method which is only possible because computer-aided applications have drastically facilitated the design of experiments in many different research areas. From predictions of artificial biochemical production pathways with enzymes from databases[155] to kinetic enzymatic models,[156] computational approaches have been broadly applied in order to gain further knowledge and deeper understanding of biological systems. These methods enable researchers to perform optimizations of the whole system and to identify crosslinks and relationships, which are not obvious on first sight due to the complexity of the applied cellular system.

In this review, the methods of flux balance analysis (FBA) and metabolic flux analysis (MFA) will be addressed to show their possibilities with regard to metabolic optimization of microorganisms. In the last two decades researchers in the field of systems biology put tremendous efforts into understanding and describing (quantitatively) various microorganisms by computational models. Knowledge about a distinct metabolic flux distribution in production hosts with respect to productivity becomes more and more important especially in the area of biotechnology. Computational models were used for predictions and simulations to reduce unnecessary experimental work. Such a tool is based on genome scale metabolic models and is called flux balance analysis (FBA) (Figure 15).[157]

A genome-scale metabolic network reconstruction, either from online platforms like the Model SEED[158] or BIGG[159] or other metabolic model databases represents the initial requirements for this kind of analyses. These models represent stoichiometric reaction

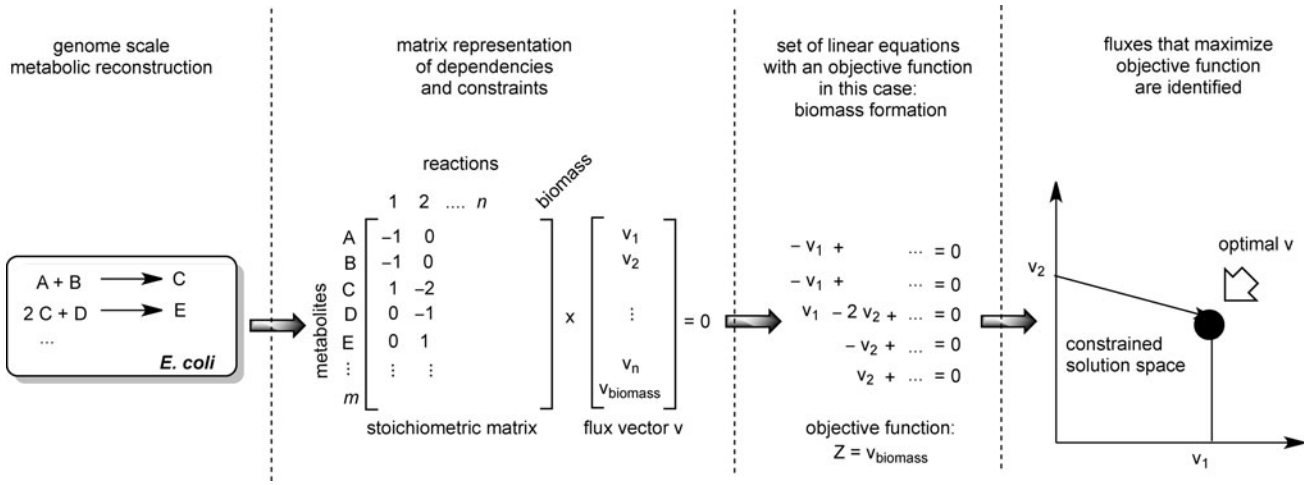

Figure 15. FBA analysis. The reactions within the metabolic pathways from a genome scale metabolic reconstruction are transformed into a matrix form and multiplied with a flux vector. The steady-state conditions set the change over time to zero. This set of linear equations determines the constraints and dependencies in the cell. If the objective function is maximized, defined solutions within the solution space can be identified.

equations of the reactions in the metabolism of the specific microorganism and can be either analyzed with online tools (http://www.theseed.org) or with the free available MATLAB COBRA toolbox.[160] A variety of online tools and toolboxes for FBA were compared and evaluated by Lakshmanan et al.[161] A simplified workflow is described in the following paragraph.

Different software tools create a stoichiometric matrix from all known reaction equations which usually also includes artificial reactions to account for the formation of biomass or transport reactions for substrates like glucose or oxygen. For example, the matrix is then multiplied with a flux vector that represents the fluxes through all the reactions in the metabolism. These fluxes are unknown variables which have to be identified during an FBA. For every flux analysis, constraints like substrate availability or oxygen uptake are set to minimize the solution space. These are fluxes which are already known. Other fluxes can be restricted to certain values (e.g., to only positive or negative values) if a reaction is unidirectional. Since FBA is mostly performed at steady state conditions, the change over time equals zero. A certain objective function is chosen which has to be optimized during the analysis.[162] Different objective functions are responsible for different results in regard to the prediction of flux distributions in the cell.[163] Simulating growing cells, the most frequently used objective function is the maximum formation of biomass.

FBA has been shown to be an effective tool for predictions of phenotypes of different KO mutants. Within an elaborate study for *S. cerevisiae* with a library of 4,658 mutants under five different environmental conditions the high predictive accuracy of the applied model was successfully demonstrated, which ranged between 96–98% for viable phenotypes and 73–80% for lethal phenotypes.[164] The accuracy of predictions enables such simulations of different growth conditions and constraints to perform KO studies for product maximization *in silico*. Therefore, this approach saves time or gives new ideas for KO strategies, which were not obvious in the first place.

Special focus is put on prevention of lethal KOs for the whole-cell biocatalyst, which can be identified with FBA. It is a general advantage of FBA that the concept is independent from kinetic data, although hybrid models are also investigated and applied by the biosystems community which partly consider dynamic attributes in FBA.[165]

The production of fumaric acid in *E. coli* represents a prominent example for the combination of classical metabolic engineering with FBA (Figure 16). Initially, the *iclR* gene for the isocitrate lyase repressor was knocked-out and TCA flux was redirected to the glyoxylate shunt. The most obvious genes, namely the three fumaric acid hydratases *fumA*, *fumB* and *fumC* were deleted in order to enhance fumaric acid production. Additionally, several other genes such as *aspA* were deleted after performing *in silico* flux responses. The resulting strain displayed a three-fold increased fumaric acid production compared to the classically modified strain.[166]

In another case study, two different *E. coli* KO mutants were predicted computationally for the improvement of L-phenylalanine production, constructed *in vivo* and showed high redundancies in the central carbon metabolism of *E. coli*.[167] This study enables a more precise FBA development and an easier improvement of the L-phenylalanine production in the future, since the L-phenylalanine yield remained the same compared to the reference strain without the

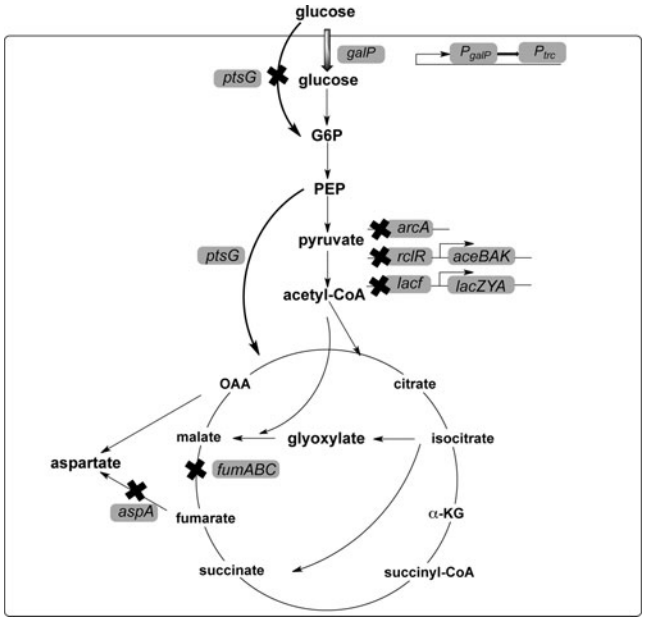

Figure 16. Knock-out (KO) strategies for fumaric acid production. KO of the *iclR* gene for the isocitrate lyase repressor and KO of three fumaric acid hydratases *fumA*, *fumB* and *fumC* enhanced the fumaric acid production. *In silico* predicted KO of *aspA* resulted in an additional increase of the fumaric acid yield.

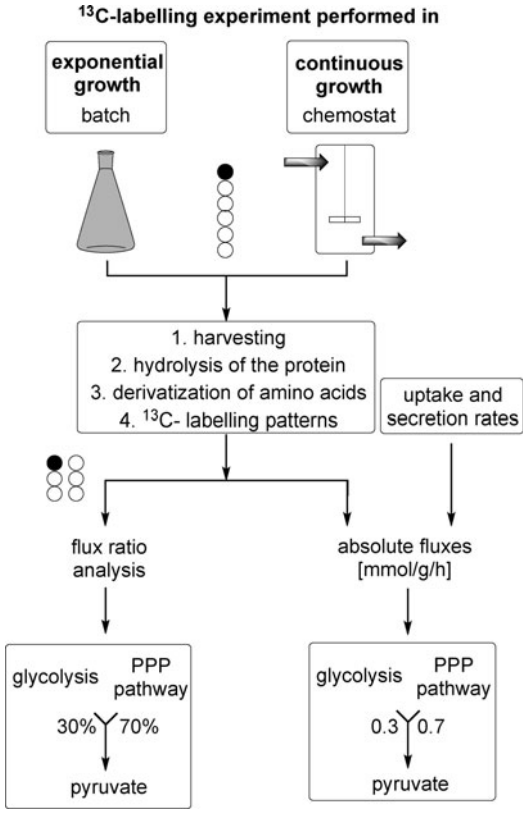

Figure 17. ^{13}C metabolic flux analysis workflow. The cultures are grown on labelled substrates and harvested during steady-state. The proteins are hydrolyzed and the amino acids derivatized for GC-MS measurements. The labelling patterns are recognized by specific software and result either in flux ratios or, with additional information about uptake and secretion rates, in absolute fluxes.

KOs. The reduced solution space makes future FBAs more precise to predict improvement strategies.

A triple KO in *E. coli* central carbon metabolism lead to a 7.4-fold increase (33.9 ± 1.2 C-mol%) in the production of 3-hydroxypropionic acid (3HP) from glycerol. This was achieved with a gene KO simulation prior to constructing the strain *in vivo*.[168]

Apart from FBA, the impact of cascades which are coupled to the carbon metabolism of a microorganism can be evaluated *via* ^{13}C metabolic flux analysis. This method is used to analyze the flux distributions of the central metabolism. Therefore, experiments with [1-^{13}C] and uniformly carbon [U-^{13}C]-labelled glucose as only carbon source are performed. Cultures are grown on labelled substrates and harvested during steady-state. Proteins are hydrolyzed and amino acids are derivatized for GC-MS measurements. Since every pathway of the central carbon metabolism shows a characteristic fragmentation pattern of the applied carbon source, it can be identified *via* differences in the ^{13}C-labelling patterns of amino acids. The analysis is usually performed *via* GC-MS and computational analysis has to be performed with a suitable software like FiatFlux,[169] 13CFLUX[170] or OpenFLUX.[171] The software applications for quantitative metabolic flux analysis[172] and general methods[173] were recently reviewed elsewhere. The resulting flux ratios between the different metabolic pathways can be used for calculation of absolute fluxes if uptake and secretion rates of substrate(s) and product(s) are provided (Figure 17).[174]

Within a recent study ^{13}C-MFA was used very efficiently to show that fluxes towards the pentose phosphate pathway (PPP) of *Corynebacterium glutamicum* were not altered after enhancement of cofactor ability by altering the coenzyme specificity of the native NAD^+-dependent glyceraldehyde 3-phosphate dehydrogenase to $NADP^+$. The desired lysine production could be improved by approximately 60%.[175] In another study recently published by the Antoniewicz group, a complementary parallel labelling experiments technique for metabolic flux analysis (COMPLETE-MFA) was established. In this study, 14 parallel labelling experiments with 8 differently ^{13}C-labelled glucose tracers were performed in *E. coli*. The best tracers for resolving the fluxes in the upper (glycolysis and PPP) and lower (TCA cycle and anaplerotic reactions) metabolism were identified. By doing so, it was shown that the best tracers for upper metabolism displayed a poor performance for lower metab-

olism and *vice versa*. This result should clearly be considered for the future design of experiments.[176]

Overall, *in silico* methods can facilitate the planning of *in vivo* experiments for product enhancement in whole cells and reduce the experimental workload, therefore, saving time, money and also offering new insights into all possible solutions.

Absolute quantification of metabolites in the host cell is also an important method as it provides the information on a possible maximal yield of the product. It also exposes possible bottlenecks within the metabolism, especially when it comes to cofactor recycling and cofactor availability.[177] Additionally, it is useful to correct FBA results since they mostly do not contain kinetic data which could be a (major) limitation in depicting real flux distributions in the cell. The quantification of intracellular metabolites is referred to as metabolomics. Several analytical techniques (GC-MS or LC-MS/MS) are known and depend on the properties and the origin of the metabolic compounds. Since the physical and chemical properties of metabolites vary from very polar (glycolytic intermediates) to strongly lipophilic (lipids, fatty acids) it has not been possible so far to cover the whole metabolome with only one analytical method. In the case of metabolic engineering and redox multi-step catalysis, investigations of the central carbon metabolism are most prominent. Methods tend to cover glycolysis, TCA cycle metabolites, amino acids, PPP intermediates and cofactors [e.g., $NAD(P)^+/H$, FAD]. The accuracy of these methods relies on the use of proper internal standards and is usually performed with ^{13}C-labelled cell extracts to obtain reproducible results and reduce matrix effects.[178] This is necessary because the intracellular compounds can be quite unstable and tend to decompose rapidly. In order to have a reference with comparable individual decomposition rates, the same compounds are applied as fully ^{13}C-labelled analogues. Hence, a fast sampling with rapid quenching or rapid extraction of metabolites is mandatory. Otherwise, the metabolic reactions will progress under non-experimental conditions, experimental error will increase and results might not be representative anymore. This is especially the case if the dynamic behavior of the metabolome is being investigated.[179] The difference between bacteria and yeast or fungi should be taken into account since the cell membranes of these species have different stabilities and the treatment method has to be adjusted in order to prevent metabolic leakage.[180] As bacterial outer membranes are less stable, filtration techniques are favored with a subsequent quenching/extraction step. In the case of more stable cell walls (e.g., yeast), samples are quenched with cold methanol, for example, and extracted afterwards. The quantification is usually performed *via* LC-MS/MS.[181] Metabolomics were recently used to identify the successful metabolic engineering of *E. coli* for triglyceride accumulation; within this study a high conservation of triglyceride composition was confirmed.[182] Another recent example for the success of metabolomics as a control strategy was disclosed by metabolic flux redirection towards a synthetic pathway with a metabolic toggle switch for changing the *E. coli* metabolism to isopropyl alcohol production. Metabolomics were used to follow the changes in the intracellular fluxes.[183]

The different computational approaches enable the construction and optimization of whole-cell biocatalysts *in silico* prior to their construction *in vivo* and reduce the number of experiments to determine the proper KO strategies. Furthermore, metabolomics is a powerful tool to both refine and control the simulated results and improve computational models.

4.2 Substrate Uptake and Product Release

Permeability of the cell membrane is very often an issue in whole-cell biotransformations. Since the outer membrane of a bacterium is a lipid bilayer, only lipophilic substances can pass the membrane *via* diffusion. This limits the substrate uptake for cascade reactions not only in terms of polarity but for non-polar substances also in terms of uptake velocity since diffusion is a passive and a slow process. Different strategies can be applied for improving already existing cascade reactions and for establishing new cascades not feasible due to hindered substrate uptake. These strategies will be discussed in the following section.

Treatment of bacterial cells with solvents and/or detergents represents the basic approach of improving cell wall permeability. This method is widely used in biocatalysis.[184] While the existence of straightforward preparation protocols is an obvious advantage of this approach, a lot of different procedures are described and identification of the suitable one for a certain biotransformation is often only achieved by simple trial and error. A lot of treatments are performed with EDTA and toluene, exemplified by the bioconversion of ethyl 4-chlorooxobutanoate (ECOB) to ethyl (*R*)-4-chloro-3-hydroxybutanoate (ECHB) by *E. coli* expressing the yeast reductase *YOL151W* and GDH for cofactor recycling purposes. In this example, 30 mM of ECOB could be fully converted to ECHB within 180 min. Without pretreatment, no conversion was observed.[185]

As the effects are non-specific it is advantageous that only limited knowledge is required prior to the treatment procedure. Additional incubation times and the addition of substances that might affect the metabolic processes of the microorganism clearly represent obstacles of such procedures.[186]

The genetic modulation of the cell enables a more elegant solution to the problem by the expression of

uptake systems for a specific kind of substrates. Some case studies along this line have been published recently employing *E. coli*. The expression and use of the alkane transporter *alkL* from *P. putida* GPO1 as uptake protein facilitated the rate-limiting step within a biooxidation cascade:[187] By co-expression of *alkL* as a transporter plug-in in addition to the product yielding cascade, specific yields could be improved by up to 100-fold for the biooxidation of $>C_{12}$ alkanes to fatty alcohols and acids. The importance of this system was the introduction of a tightly regulated expression vector since the overexpression of this transporter is a toxic burden for *E. coli*. With this modification, the alkane oxidation was further improved by approximately 10-fold. The same transporter system was also used in *E. coli* to produce (*S*)-styrene oxide in two-liquid-phase cultures with a space-time-yield of $2\,g\,L^{-1}\,h^{-1}$[188] which was more than a 4-fold improvement compared to previously published space-time-yields of around $0.45\,g\,L^{-1}\,h^{-1}$.[189]

Utilization of porins represents another possibility for higher cell permeability towards certain molecules. Porins are outer membrane channels in Gram-negative bacteria and are required for the influx of nutrients and the efflux of waste products. It is known that non-specific diffusion of hydrophilic solutes across the outer membrane occurs *via* porins. Therefore, porins can be used in biocatalysis for increasing the permeability for hydrophilic substances. The two major porins in *E. coli* are *OmpF* and *OmpC*. *OmpF* is slightly larger than *OmpC* and allows the diffusion of compounds such as antibiotics.[190] Within a recently published study it was outlined that high salt concentrations lead to a higher permeability of *OmpC*.[191] As for regulation of expression, porins are regulated by environmental factors (e.g., osmolarity, pH, temperature, nutrient concentrations). The *E. coli* porin *OmpF* is expressed at low temperatures. Another transporter, *PhoE* is expressed only under phosphate starvation which makes its expression inducible[190] and, consequently, applicable for biocatalytic purposes.

These different strategies can be applied to improve the performance of whole-cell biocatalysts and increase space-time-yields in biocatalytic processes.

4.3 Cofactor Recycling

The introduction of redox reactions into a host cell is inevitably connected to the additional consumption of redox cofactors. Since this shortage of redox cofactors affects growth of the host microorganism and production of target molecules, it is necessary to recycle the consumed cofactor molecules.[192] This can be achieved by knocking-out competing reactions which consume the same cofactors and/or by knocking-in reactions which recycle the required cofactors. An extended overview of the different strategies for cofactor recycling was published, recently.[193]

For NADPH recycling, most engineering approaches in aerobic cultivations focus on increasing the flux through the PPP. In recent years, strategies were expanded to the measurement of redox cofactor levels and the prediction of experimental outcomes with FBA[194] or increasing the yield by PPP cyclization. The confirmation of the cyclization was achieved *via* ^{13}C-MFA.[195]

Other problems arise with the anaerobic production of NADPH. In this case, the overexpression of PPP enzymes does not yield more NADPH. Therefore, other strategies for NADPH generation should be pursued. Modulation of the NADP transhydrogenase and NAD kinase activity in order to produce NADPH and NAD^+ from $NADP^+$ and NADH, respectively, represents a prominent example. This combined strategy worked for aerobic as well as for anaerobic production of NADPH.[196]

For NAD^+/NADH recycling, different mechanisms have to be addressed as depicted for NADPH recycling since it is produced and recycled in a different way. As an example, *E.coli* was applied for the CO_2 fixation *via* phosphoenol pyruvate carboxylase (PPC) for succinic acid production under anaerobic cultivation conditions. KO of competing pathways led to accumulation of pyruvic acid and limited the regeneration of NAD^+ from glycolytic NADH. Therefore, the regeneration of NAD^+ had to be achieved by the co-expression of nicotinic acid phosphoribosyl transferase (NAPRTase) and pyruvic acid carboxylase (PYC). NAPRTase is rate limiting in *E. coli* NAD(H) synthesis. PYC synthesizes oxaloacetate (OAA) from pyruvic acid. (Figure 18).[197]

These KIs led to a significant increase in cell mass and succinic acid production of $12.08\,g\,L^{-1}$ under anaerobic conditions.[197] This was not possible before since the glycolytic flux was blocked by the missing recycling of NAD^+.[198] Introduction of complete recycling systems into the whole-cell biocatalyst serves as a different method for an improved cofactor recycling. A recombinant *E. coli* equipped with the ADH (*R-ADH*) from *Lactobacillus kefir* (*L. kefir*) and a GDH (*TA-GDH*) from *Thermoplasma acidophilum* for cofactor recycling was applied for the reduction of α-halogenated ketones.[153b] Recently, an *E. coli* whole-cell approach was reported for a redox self-sufficient whole-cell catalyzed amination of alcohols. Thereby, an ADH from *Bacillus stearothermophilus* was used for substrate oxidation, a transaminase from *Vibrio fluvialis* was applied for aldehyde-intermediate amination and an alanine dehydrogenase from *B. subtilis* was expressed for recycling NADH.[199]

Connecting an enzyme to perform the desired reaction to a cofactor recycling enzyme *via* a linker-

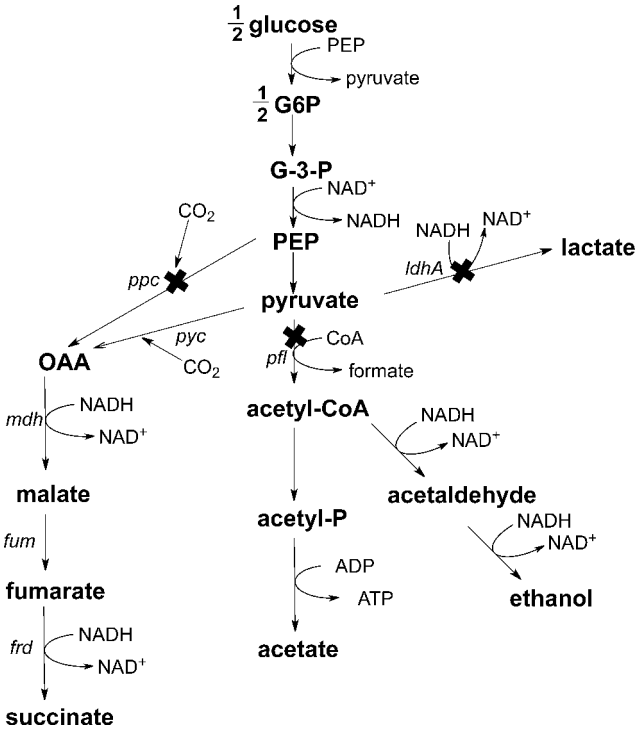

Figure 18. Possible NAD^+ recycling strategy. The regeneration of NAD^+ was achieved by co-expression of nicotinic acid phosphoribosyl transferase (NAPRTase) and pyruvate decarboxylase (PYC). NAPRTase is the rate limiting enzyme in the *E. coli* NAD(H) synthesis and PYC is an enzyme which synthesizes OAA from pyruvic acid.

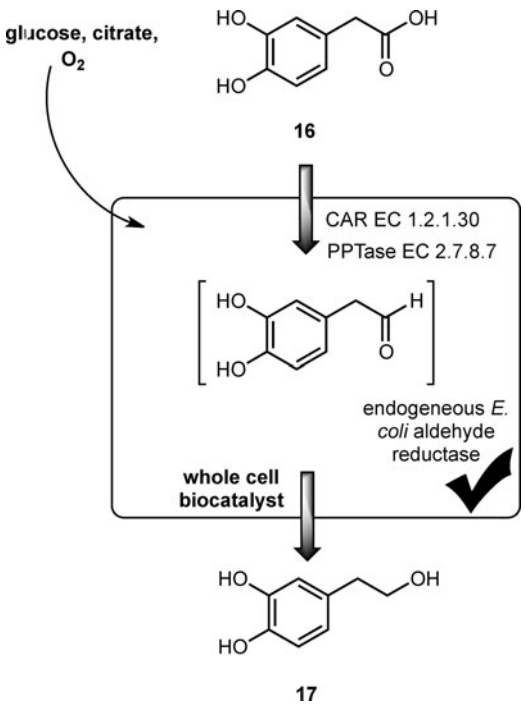

Figure 19. Advantageous intrinsic cell activity. Heterologous reduction of the substrate (**16**) by *CAR* and subsequent reduction by endogenous *E. coli* aldehyde reductases yielded the final product (**17**).

domain represents another elegant solution to the problem. This was demonstrated for several BVMOs which were linked to a thermostable variant of phosphite dehydrogenase.[200] This concept worked well in cell-free extracts since phosphite cannot diffuse trough the cell membrane. But the general concept of self-sustaining enzymes with an integrated recycling system could also be applied to whole-cell biocatalysis. With cofactor recycling as a final optimization step in a whole-cell biocatalytic process, the next chapter will focus on depicting successful examples of *in vivo* redox cascades.

5 Microbial Cell Factories Conducting Optimized Redox Cascades

Minimization of background reactions (e.g., targeted gene KO) and flux optimization (e.g., scaffolding) are important concepts of engineering microorganisms to enhance host productivity. To successfully establish artificial pathways in genetically tractable host microorganisms such as *E. coli*,[201] the host's metabolic (i.e., intrinsic) background as well as extrinsic factors have to be considered to minimize undesired side reactions.[202] There are three general strategies to improve selectivity and yields in particular and the general overall performance of microbial cell factories: (i) adaption of cellular properties such as genetics (e.g., gene KI and/or KO) and/or physiology (e.g., membrane permeability[185] and substrate uptake[203]); (ii) reaction engineering[204] (e.g., buffer, temperature, pH, use of biphasic systems), and (iii) catalyst design (e.g., protein engineering).[205] Optimization methods on chemical and molecular levels combined with the concept of biocatalytic retrosynthesis[206] will open doors to new high performance microorganisms for the efficient production of drugs and various complex natural compounds.[72,207]

In the first example, Winkler and co-workers presented an artificial enzyme cascade for the synthesis of 3-hydroxytyrosol (3-HT; **17**).[208] *CAR* from *Nocardia iowensis* (*N. iowensis*), which is dependent on the activity of an *E. coli* phosphopantetheinyl transferase, produced the aldehyde intermediate that is conveniently reduced to product 3-HT (**17**) by the activities of endogeneous *E. coli* aldehyde reductases (Figure 19).

A preparative biotransformation was performed in the presence of low citrate and glucose concentrations (10 mM) for improved cofactor regeneration. This enabled generation of 1 mM 3-HT (**17**) within 20 h starting from 10 mM 3,4-dihydroxyphenylacetic acid (DOPAC) (**16**).

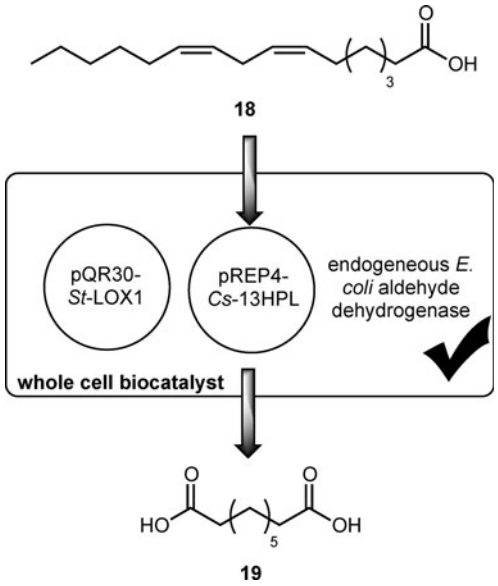

Figure 20. Optimized cascade for azelaic acid (**19**) production. Co-expression of cascade enzymes in *E. coli* from a dual expression system utilizing a high- and a low-copy number vector for balanced protein production and exploitation of endogeneous enzymatic host activity yielded the desired product (**19**).

Very recently, the group of Hauer presented a dual expression system for the synthesis of azelaic acid (**19**).[209] Their artificial pathway contained a lipoxygenase (St-LOX1) from *Solanum tuberosum* (*S. tuberosum*) and a hydroperoxide lyase from *Cucumis melo* (Cs-9/13HPL). Together with the intrinsic aldehyde dehydrogenase activity of the *E. coli* host, linoleic acid (**18**) was converted into the desired product (**19**) (Figure 20).

The previously reported *in vitro* approach[210] using cell-free extracts of *St-LOX1* and *Cs-13HPL* showed effects of unbalanced protein levels (*St-LOX1/Cs-13HPL* = 20:1) resulting in a low overall productivity. Co-expression of both enzymes in the same cell from different plasmids solved this obstacle. Since the flux through the cascade suffers from unbalanced protein production, the lipoxygenase *St-LOX1* was inserted into the high-copy number vector pQE30 and *Cs*-13HPL into the low-copy number vector pREP4. Especially the reduced production of *Cs-13HPL* from the low-copy plasmid was important for the optimization of the whole-cell catalyst.

Additionally, host strain selection[211] (*E. coli versus S. cerevisiae*) and tuning of reaction conditions such as pH adjustment or the application of a biphasic system [5% (*v*/*v*) cylcohexane in water] increased the overall productivity (29 mg L^{-1} in 8 h). As an outlook, the authors pointed out that reduction of intrinsic enzyme activity by gene KO, for example, may be beneficial and such studies are currently ongoing in Hauer's lab.

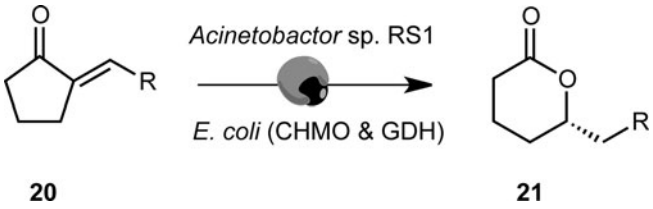

Figure 21. Mixed cell culture approach for minimized background reaction activity.

A simple but very effective approach to further exploit intrinsic host activity was presented by the group of Li by utilizing a mixed two-cell culture approach. The endogeneous *RS1* reductase from *Acinetobacter* sp. catalyzed the enantioselective reduction of the C=C double bond (**20**). The subsequent selective oxidation was performed by a cyclohexanone monooxygenase (CHMO) from *Acinetobacter calcoaceticus* expressed in *E. coli*. The GDH from *B. subtilis* was co-expressed for cofactor regeneration. The multienzyme cascade produced enantiopure δ-lactones (**21**) (Figure 21).[212]

Noteworthy, this mixed culture biotransformation was carried out in a sequential manner to minimize the intrinsic activity of *E. coli* to the stereochemical properties of the C=C bond. Substituted lactones (**21**) were produced in good yields (41–56%) and high optical purities (98% *ee*).

Another optimized whole-cell biocatalyst designed by the group of Li was published in 2015. For this artificial pathway, an epoxide hydroxylase (SpEH from *Sphingomonas* sp. HXN-200), an ADH (BDHA from *B. subtilis*) and an NADH oxidase (NOX from *L. brevis* DSM 20054) were co-expressed in a single organism for the enantioselective synthesis of α-hydroxy ketones (**23**) starting from *meso* or racemic epoxides (**22**).[213] After SpEH-mediated epoxide hydrolysis, NAD^+-dependent dehydrogenase oxidizes the diol to the ketone (Figure 22). Interestingly, two different en-

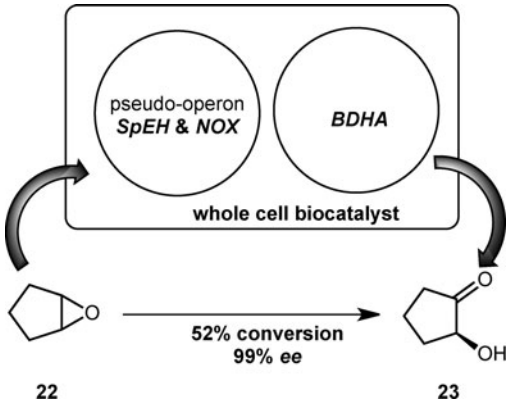

Figure 22. Whole-cell catalyst for the synthesis of α-hydroxy ketones (**23**) starting from *meso* or racemic epoxides (**22**) with optimized cofactor recycling system constructed on a pETduet and a pET28a vector.

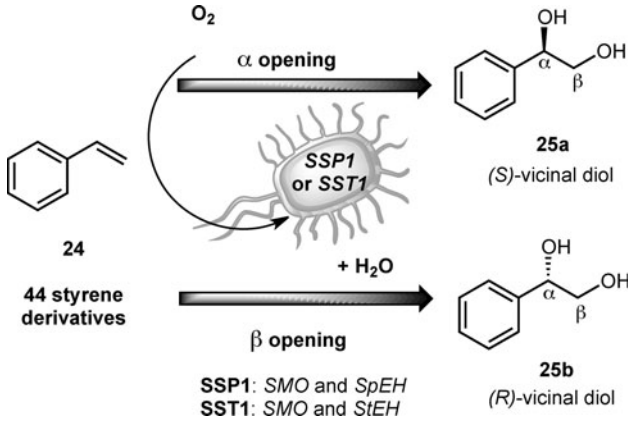

Figure 23. Microbial cell factory for the generation of various diol compounds (**25a**, **25b**) starting from styrene derivatives (**24**). (32 of 44 substrates >50% $yield_{analyt}$; 30 of 44 substrates >90% *ee*).

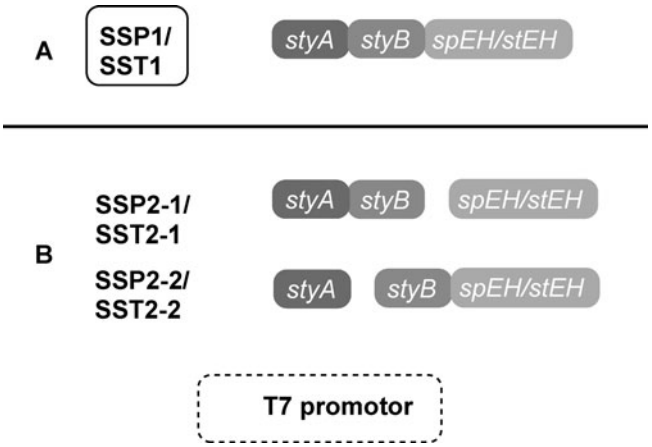

Figure 24. Genetic configurations of expression cassettes. Operon (A) and pseudo-operon (B) co-expressed in *E. coli* containing the two coding sequences of *SMO* (*styA* and *styB*) and *spEH* or *stEH*, respectively.

zymes were compared for efficient cofactor recycling: a lactate dehydrogenase (LDH) from *B. subtilis* and the aforementioned NOX from *L. brevis*. LDH requires pyruvate and NOX consumes O_2 for NAD^+ regeneration. *In vitro* experiments gave a total turnover number (TTN) of 5,500 to 26,000 for LDH for the NAD^+ regeneration and a practical TTN of 6,200 to 14,000 for NOX.

Different from the earlier established system using LDH,[214] the whole-cell catalyst was constructed by the co-transformation of pETDuet-1/*SpEH*-NOX and pET28a/*BDHA*. As both plasmids share the same ORI, plasmid maintenance exclusively relied on the selection by two different antibiotics. This is in contrast to the concept of plasmid incompatibility and might result in inconsistent expression patterns in different cell populations.[8a]

Besides the presented process engineering strategies, also improvements on molecular and genetic levels were published in recent years. Li and co-workers published one of the first multi-enzyme cascades optimized on a genetic level. Thereby, they established a biocatalytic opponent to the well-established Sharpless dihydroxylation.[215] In their two-step reaction sequence co-expressed in *E. coli*, a styrene monooxygenase from *Pseudomonas* sp. and two enantiocomplementary epoxide hydrolases from *Sphingomonas* sp. HXN-200 and *S. tuberosum*, respectively, are used for the enantioselective dihydroxylation of terminal and internal olefins (**24**) (Figure 23). Special emphasis is put on direct access to all enantiomeric products (**25a**, **25b**) only by changing the regioselectivity of a single enzyme-mediated reaction step (*SpEH* or *StEH*).

To balance expression levels of all genes of interest, three different expression cassettes of *styA*, s*tyB* and *spEH* were constructed on the common plasmid pRSFDuet-1 (Figure 24).

Competent *E. coli* cells were individually transformed with one of the constructs. The pathway was either arranged in an operon form (SSP1/SST1: one common T7 promotor; Figure 24A) or in two distinct pseudo-operon configurations (SSP2-1/SSP2-2 and SST2-1/SST2-2: individual T7 promoters; Figure 24B).

Best results were generated with the *E. coli* strain harboring SSP1 or SST1 (operon form) under basic conditions (pH 8.0) to avoid unselective self-hydrolysis of the epoxide intermediate.

Resting cell biotransformations showed good conversions. For example, (*S*)-1-phenylethane-1,2-diol (**25**) was produced with 86% yield and an excellent enantioselectivity of 97% *ee* after 5 h. The strains containing plasmids encoding the pathway as pseudo-operons also produced the target molecule, however, in lower concentrations. This might result from suboptimal expression levels.

Finally, three non-native pathways optimized on a genetic level are presented to demonstrate the degree of efficiency by artificial cascades already achieved (Figure 25 and Figure 26).

The groups of Mihovilovic and Bornscheuer (Figure 25A)[7b] and the group of Reetz (Figure 25B)[7a] published two independent redox cascades for the stereoselective synthesis of lactones (**28**) or cyclohexanones (**31**). Both cascade designs combine the concept of biocatalytic retrosynthesis with metabolic engineering tools and in both cases the same endogeneous *E. coli* ene-reductase, the known *N*-ethylmaleimide reductase A (*NemA*), had to be deleted in order to minimize undesired by-product formation. The directed gene deletion prevented non-selective C=C bond reductions on the activated alkenes (**27**, **30**) (Figure 25).

The work of Oberleitner et al.[7b] is a representative demonstration on efficient artificial pathways with prospective applications in the chemical industry in

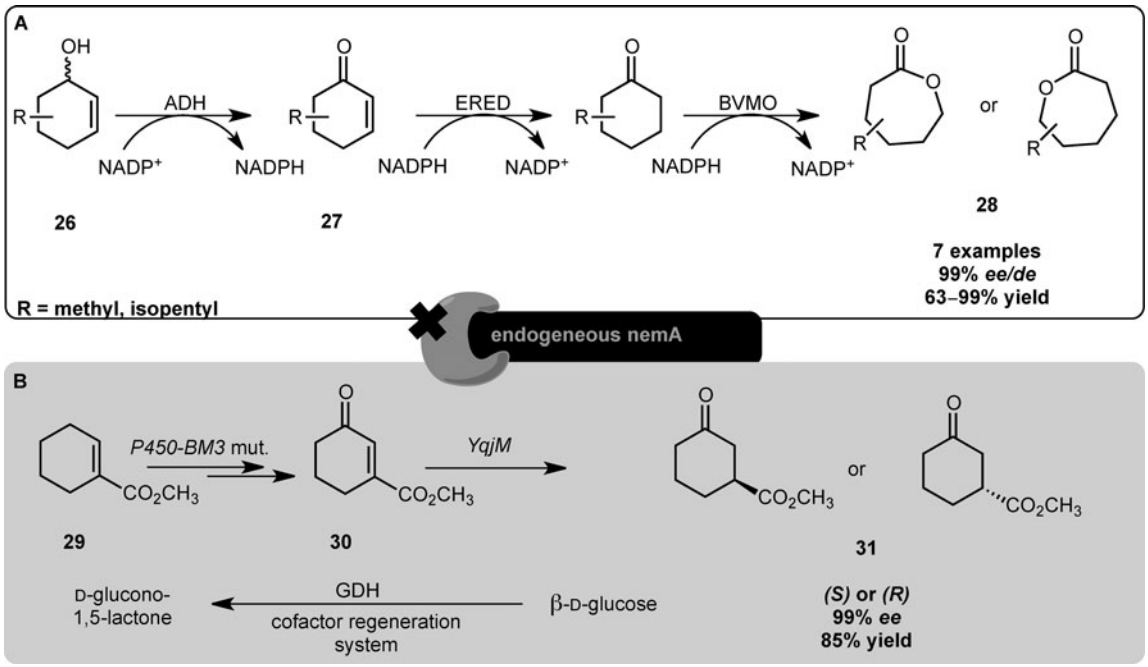

Figure 25. "Designer cells" with minimized intrinsic activity by gene KO strategies. (A) BVMO containing cascade for the stereoselective synthesis of various lactones (**28**), (B) P450-BM3 mut. and YqjM containing artificial pathway for the generation of substituted cyclohexanones (**31**).

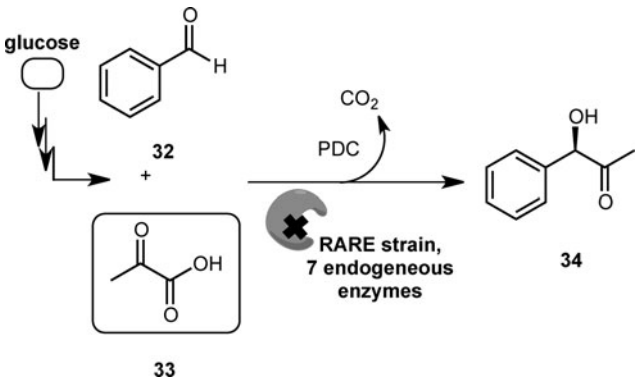

Figure 26. *In vivo* cascade for the synthesis of L-phenylacetyl carbinol (**34**), expressed in the optimized RARE strain.

the future (Figure 25A). Their expression system utilizes a two pET plasmid system for the co-expression of all three enzymes in one cell and plasmid maintenance is achieved by different antibiotics. While Agudo and Reetz relied on the λ *Red*/ET system[216] for the deletion of the *nemA* gene, Oberleitner et al. took advantage of the TargeTron™ gene KO system which is based on group II intron-mediated gene KO.[41b]

In a preparative biotransformation using resting cells of their Δ*nemA* strain and 100 mg substrate (**26**) [e.g., 3-methylcyclohexen-1-ol or (*1S,5S*)-carveol], they were able to isolate both products (**28**) after classical extractive work-up with very good overall yields of 55 and 60%, respectively. In addition, modularity and flexibility of this microbial cell factory were illustrated by desymmetrizations, kinetic resolutions and regio-divergent biotransformations within the BVMO-mediated step based on different cyclohexenol derivatives (**26**).

Remarkably, Agudo and Reetz used enzymes engineered by directed evolution and tested three different approaches to establish their redox cascade (Figure 25B) using techniques described in previous chapters and paragraphs: (i) a two-cell system in which one *E. coli* strain harbored the *P450-BM3*- and another the *YqjM*-coding plasmid, respectively; (ii) a single-strain approach co-expressing the two enzymes from two plasmids (pRSF/*P450-BM3* and pACYC/*YqjM*); and (iii) the use of an engineered *E. coli* strain with chromosomally integrated *YqjM* mutants while *P450- BM3* remained on a plasmid. The three approaches differ in the levels of metabolic load on the host with (ii) imposing the highest burden with two plasmids co-maintained in a single cell. Integration of *YqjM* was achieved by λ *Red* recombination[216] elegantly knocking-out the endogeneous *nemA* gene[217] coding for the competing ene-reductase in the same step.

Under process conditions including NAPH cofactor recycling by glucose/GDH, all three devices gave 40–60% yields with enantiomerically pure (*R*)- and (*S*)-cyclohexan-3-onecarboxylic acid methyl ester (99% *ee*) (**31**), respectively. In comparison, the *in vitro* approach with a cell-free extract from (ii) only resulted in $<26\%$ yield.[7a] The enzymatic cascade im-

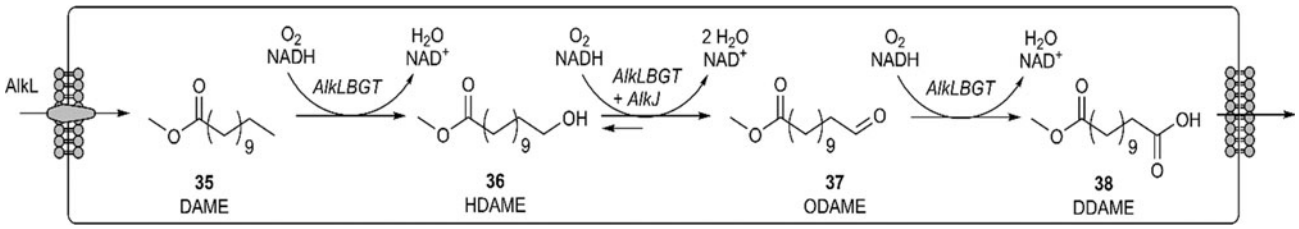

Figure 27. "Showcase" of an *in vivo* cascade using different optimization (e.g., substrate uptake, 2 phasic system) strategies to enhance the host productivity. [DAME: dodecanedioic acid methyl ester (**35**); HDAME: 12-hydroxydodecanoic acid methyl ester (**36**); ODAME: 12-oxododecanoic acid methyl ester (**37**); DDAME: dodecanedioic acid monomethyl ester (**38**)].

pressively combines a whole array of techniques and surely can be further optimized by coordinating the production of *P450-BM3* and *YqjM* as seen in approach (i) by the sequential addition of the two *E. coli* strains. In approach (i), yields were significantly lower without a time lag (in other words, when *P450-BM3* and *YqjM* were both present from the very beginning). Coordinated gene expression could be achieved by different inducible promoters or other regulatory elements. This study also adds temporal aspects to the complexity chart for cascade reactions for future considerations.

Prather and co-workers reported targeted gene deletions to minimize aromatic aldehyde reductase activity in *E. coli*.[11] Their so-called RARE strain contains nine gene deletions, of which six are responsible for unwanted aldehyde reductase activity (*dkgB*, *yeaE, yqhD*, *dkgA*, *yahK*, *yjgB*,) as well as an endonuclease (*endA)* and a recombinase (*recA*) which were deleted to increase host productivity and to improve plasmid stability, respectively. First, the engineered strain was investigated to produce vanillin from vanillate using a *CAR* from *N. iowensis*, without side products (or more precisely without detection of vanillyl alcohol, which is the main product of intrinsic aldehyde reductase activity). With their optimized *E. coli* strain, they presented a high performance whole-cell catalyst for the synthesis of the chiral pharmaceutical intermediate L-phenylacetyl carbinol (**34**) by C–C bond formation between benzaldehyde (**32**) and the glycolytic product pyruvate (**33**) (Figure 26).

Expression of a recombinant mutant pyruvate decarboxylase (PDC) and co-expression of previously reported *CAR* in order to avoid the reduction of benzoate to benzyl alcohol have increased the substrate conversion by 10-fold compared to the *E. coli* wild-type.

Within 24 h their optimized "designer cell" converted 5 mM benzaldehyde to the desired product without any side product formation.

Recently, the group of Bühler presented an efficient biocatalyst for the synthesis of dodecanedioic acid monomethyl ester (**38**) (DDAME), a potent building block for polymer chemistry.[203] The heterologous pathway contains the alkane monooxygenase *AlkBGT* and the ADH *AlkJ,* both from *P. putida,* and converts the dodecanedioic acid methyl ester (**35**) (DAME), a renewable feedstock chemical, into DDAME (**38**) (Figure 27). This optimized pathway perfectly illustrates the efficiency of reaction and catalyst engineering in order to enhance host productivity. Outer membrane protein *AlkL* was introduced to overcome limitation of substrate uptake and substrate flux through the cascade, as outlined above. In order to increase their biocatalyst stability, the benefit of a carrier solvent serving as both a substrate reservoir and a product sink was demonstrated. Finally, best results were obtained with an organic phase composition of 75% bis-(2-ethylhexyl)phthalate (BEHP) and 25% DAME (**35**). Furthermore, the heterologous expression of *AlkJ* shifted the reaction equilibrium from the primary alcohol (**36**) HDAME to the thermodynamically unfavored aldehyde moiety (**37**) (ODAME) which was successfully transformed to the target molecule DDAME (**35**).

In particular, the presented cell factories of Prather and Bühler highlighted how beneficial catalyst and/or reaction engineering in combination with optimization strategies on genetic and molecular levels can be for productivity enhancement.

6 Outlook

Recent trends in whole-cell-mediated redox biocatalysis progressed from single-step transformations towards multi-enzyme cascades. An increasing demand for the production of chemicals, pharmaceuticals, fuels and valuable materials requires future high performance cell factories. Combination of diverse strategies from various disciplines like metabolic engineering, systems biology and biocatalysis enable optimization of bio-based processes on different levels. Essentially, genome engineering techniques, proper gene expression control, protein engineering for enhanced catalyst performance and process engineering concepts are mandatory in performing whole-cell-mediated redox biocatalysis as well as metabolic engineering.

In this paper we have reviewed the most prominent concepts, recently developed tools and strategies in the context of multi-step redox biocatalysis. Thereby we tried to close the gap between different scientific disciplines in order to obtain the best performing biocatalytic system. Hence the use of computer-aided tools and the more holistic view on the cell as a system become more and more important, avoiding tedious trial and error experimental work and ultimately reducing development costs for industrially relevant processes. Future trends go into the direction of developing "designer cells" for either different enzyme classes (e.g., ERED production host, by elimination of intrinsic *nemA* and *fadH* activity in *E. coli*) or a specific substance class (e.g., RARE strain for the production of aldehydes). Overall, an increasing number of commercially available and easy-to-use tools for the genetic manipulation and computational simulation of a variety of production hosts (e.g., *E. coli*, *S. cerevisiae*) open new perspectives in the design and application of bio-based whole-cell-mediated redox processes.

Acknowledgements

We thank the FWF (grant no. I723-N17, P24483-B20) and the COST action systems biocatalysis WG2 for financial support.

References

[1] a) A. Liese, K. Seelbach, C. Wandrey, *Industrial Biotransformations*, 2nd edn., John Wiley & Sons, **2006**; b) E. Ricca, B. Brucher, J. H. Schrittwieser, *Adv. Synth. Catal.* **2011**, *353*, 2239–2262.
[2] F. Lopez-Gallego, C. Schmidt-Dannert, *Curr. Opin. Chem. Biol.* **2010**, *14*, 174–183.
[3] L. M. Blank, B. E. Ebert, K. Bühler, B. Bühler, *Antioxid. Redox Signaling* **2010**, *13*, 349–394.
[4] B. R. Glick, *Biotechnol. Adv.* **1995**, *13*, 247–261.
[5] R. R. Chen, *Appl. Microbiol. Biotechnol.* **2007**, *74*, 730–738.
[6] T. Schweder, H. Y. Lin, B. Jurgen, A. Breitenstein, S. Riemschneider, V. Khalameyzer, A. Gupta, K. Buttner, P. Neubauer, *Appl. Microbiol. Biotechnol.* **2002**, *58*, 330–337.
[7] a) R. Agudo, M. T. Reetz, *Chem. Commun.* **2013**, *49*, 10914–10916; b) N. Oberleitner, C. Peters, J. Muschiol, M. Kadow, S. Saß, T. Bayer, P. Schaaf, N. Iqbal, F. Rudroff, M. D. Mihovilovic, U. T. Bornscheuer, *ChemCatChem* **2013**, *5*, 3524–3528.
[8] a) J. D. Keasling, *Trends Biotechnol.* **1999**, *17*, 452–460; b) G. J. Balzer, C. Thakker, G. N. Bennett, K.-Y. San, *Metab. Eng.* **2013**, *20*, 1–8.
[9] M. C. Y. Chang, R. A. Eachus, W. Trieu, D.-K. Ro, J. D. Keasling, *Nat. Chem. Biol.* **2007**, *3*, 274–277.
[10] J. E. Dueber, G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. J. Prather, J. D. Keasling, *Nat. Biotechnol.* **2009**, *27*, 753–759.
[11] A. M. Kunjapur, Y. Tarasova, K. L. Prather, *J. Am. Chem. Soc.* **2014**, *136*, 11644–11654.
[12] a) F. Baneyx, *Curr. Opin. Biotechnol.* **1999**, *10*, 411–421; b) H. P. Sorensen, K. K. Mortensen, *J. Biotechnol.* **2005**, *115*, 113–128.
[13] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, H. Mori, *Mol. Syst. Biol.* **2006**, *2*, DOI: 10.1038/msb41000.
[14] F. Hoffmann, U. Rinas, *Adv. Biochem. Eng. Biotechnol.* **2004**, *89*, 73–92.
[15] W. Bentley, D. Kompala, *Biotechnol. Lett.* **1990**, *12*, 329–334.
[16] a) ref.[14]; b) D.-E. Chang, D. J. Smalley, T. Conway, *Mol. Microbiol.* **2002**, *45*, 289–306.
[17] a) C. G. Kurland, H. Dong, *Mol. Microbiol.* **1996**, *21*, 1–4; b) H. Dong, L. Nilsson, C. G. Kurland, *J. Bacteriol.* **1995**, *177*, 1497–1504.
[18] a) J. C. Diaz Ricci, M. E. Hernández, *Crit. Rev. Biotechnol.* **2000**, *20*, 79–108; b) F. Silva, J. A. Queiroz, F. C. Domingues, *Biotechnol. Adv.* **2012**, *30*, 691–708.
[19] a) P. K. Ajikumar, W.-H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, G. Stephanopoulos, *Science* **2010**, *330*, 70–74; b) L. Z. Yuan, P. E. Rouvière, R. A. LaRossa, W. Suh, *Metab. Eng.* **2006**, *8*, 79–90.
[20] a) S. Gräslund, P. Nordlund, J. Weigelt, J. Bray, O. Gileadi, S. Knapp, U. Oppermann, C. Arrowsmith, R. Hui, J. Ming, *Nat. Methods* **2008**, *5*, 135–146; b) G. Hannig, S. C. Makrides, *Trends Biotechnol.* **1998**, *16*, 54–60; c) P. Jonasson, S. Liljeqvist, P.-A. K. Nygren, S. Ståhl, *Biotechnol. Appl. Biochem.* **2002**, *35*, 91–105; d) B. Gasser, M. Saloheimo, U. Rinas, M. Dragosits, E. Rodriguez-Carmona, K. Baumann, M. Giuliani, E. Parrilli, P. Branduardi, C. Lang, D. Porro, P. Ferrer, M. Tutino, D. Mattanovich, A. Villaverde, *Microb. Cell Fact.* **2008**, *7*, 11.
[21] a) R. P. Novick, *Microbiol. Rev.* **1987**, *51*, 381–395; b) 21b J. Sambrook, E. F. Fritsch, T. Maniatis, *Biol. Unserer Zeit* **1989**, *20*, 285.
[22] D. Held, K. Yaeger, R. Novy, *inNovations* **2003**, *18*, 4–6.
[23] A. Shafferman, D. R. Helinski, *J. Biol. Chem.* **1983**, *258*, 4083–4090.
[24] K. Zhang, M. R. Sawaya, D. S. Eisenberg, J. C. Liao, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20653–20658.
[25] N. H. Tolia, L. Joshua-Tor, *Nat. Methods* **2006**, *3*, 55–64.
[26] H.-C. Tseng, C. H. Martin, D. R. Nielsen, K. L. J. Prather, *Appl. Environ. Microbiol.* **2009**, *75*, 3137–3145.
[27] D. R. Nielsen, E. Leonard, S.-H. Yoon, H.-C. Tseng, C. Yuan, K. L. J. Prather, *Metab. Eng.* **2009**, *11*, 262–273.
[28] H.-C. Tseng, K. L. J. Prather, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17925–17930.
[29] M. J. Sheppard, A. M. Kunjapur, S. J. Wenck, K. L. J. Prather, *Nat. Commun.* **2014**, *5*, 5031.

[30] a) T. Ellis, T. Adie, G. S. Baldwin, *Integr. Biol.* **2011**, *3*, 109–118; b) P. Xu, A. Vansiri, N. Bhan, M. A. Koffas, *ACS Synth. Biol.* **2012**, *1*, 256–266; c) T. W. Wang, X. Y. Ma, G. C. Du, J. Chen, *Mol. Biotechnol.* **2012**, *52*, 300–308.
[31] a) D. J. Pitera, C. J. Paddon, J. D. Newman, J. D. Keasling, *Metab. Eng.* **2007**, *9*, 193–207; b) V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotechnol.* **2003**, *21*, 796–802; c) T. S. Lee, R. A. Krupa, F. Zhang, M. Hajimorad, W. J. Holtz, N. Prasad, S. K. Lee, J. D. Keasling, *J. Biol. Eng.* **2011**, *5*, 12.
[32] D. Rebatchouk, N. Daraselia, J. O. Narita, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10891–10896.
[33] T. Knight, http://web.mit.edu/synbio/release/docs/biobricks.pdf, **2003**.
[34] R. P. Shetty, D. Endy, T. F. Knight Jr, *J. Biol. Eng.* **2008**, *2*, 5.
[35] J. C. Anderson, J. E. Dueber, M. Leguia, G. C. Wu, J. A. Goler, A. P. Arkin, J. D. Keasling, *J. Biol. Eng.* **2010**, *4*, 1.
[36] J. E. Norville, R. Derda, S. Gupta, K. A. Drinkwater, A. M. Belcher, A. E. Leschziner, T. F. Knight Jr, *J. Biol. Eng.* **2010**, *4*, 17.
[37] J. Vick, E. Johnson, S. Choudhary, S. Bloch, F. Lopez-Gallego, P. Srivastava, I. Tikh, G. Wawrzyn, C. Schmidt-Dannert, *Appl. Microbiol. Biotechnol.* **2011**, *92*, 1275–1286.
[38] J. E. Bailey, *Adv. Biochem. Eng./Biotechnol.* **1993**, *48*, 29–52.
[39] K. L. Jones, S. W. Kim, J. D. Keasling, *Metab. Eng.* **2000**, *2*, 328–338.
[40] a) R. Carlson, *Nat. Biotechnol.* **2009**, *27*, 1091–1094; b) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185–194.
[41] a) M. G. Montague, C. Lartigue, S. Vashee, *Curr. Opin. Biotechnol.* **2012**, *23*, 659–665; b) C. W. Song, J. Lee, S. Y. Lee, *Biotechnol. J.* **2015**, *10*, 56–68.
[42] a) C. Aslanidis, P. J. de Jong, *Nucleic Acids Res.* **1990**, *18*, 6069–6074; b) C. Li, R. M. Evans, *Nucleic Acids Res.* **1997**, *25*, 4165–4166.
[43] a) J. Quan, J. Tian, *PLoS ONE* **2009**, *4*, e6441; b) A. V. Bryksin, I. Matsumura, *BioTechniques* **2010**, *48*, 463–465.
[44] M. Z. Li, S. J. Elledge, *Nat. Methods* **2007**, *4*, 251–256.
[45] C. Li, A. Wen, B. Shen, J. Lu, Y. Huang, Y. Chang, *BMC Biotechnol.* **2011**, *11*, 92.
[46] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat. Methods* **2009**, *6*, 343–345.
[47] D. G. Gibson, H. O. Smith, C. A. Hutchison, J. C. Venter, C. Merryman, *Nat. Methods* **2010**, *7*, 901–903.
[48] Y. Zhang, U. Werling, W. Edelmann, *Nucleic Acids Res.* **2012**, *40*, e55.
[49] S. C. Kowalczykowski, D. A. Dixon, A. K. Eggleston, S. D. Lauder, W. M. Rehrauer, *Microbiol. Rev.* **1994**, *58*, 401–465.
[50] V. Larionov, N. Kouprina, J. Graves, X. N. Chen, J. R. Korenberg, M. A. Resnick, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 491–496.
[51] a) D. G. Gibson, G. A. Benders, K. C. Axelrod, J. Zaveri, M. A. Algire, M. Moodie, M. G. Montague, J. C. Venter, H. O. Smith, C. A. Hutchison, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20404–20409; b) G. A. Benders, V. N. Noskov, E. A. Denisova, C. Lartigue, D. G. Gibson, N. Assad-Garcia, R.-Y. Chuang, W. Carrera, M. Moodie, M. A. Algire, Q. Phan, N. Alperovich, S. Vashee, C. Merryman, J. C. Venter, H. O. Smith, J. I. Glass, C. A. Hutchison, *Nucleic Acids Res.* **2010**, *38*, 2558–2569.
[52] Z. Shao, H. Zhao, H. Zhao, *Nucleic Acids Res.* **2009**, *37*, e16.
[53] N. Traitcheva, H. Jenke-Kodama, J. He, E. Dittmann, C. Hertweck, *ChemBioChem* **2007**, *8*, 1841–1849.
[54] Y. S. Choi, T. W. Johannes, M. Simurdiak, Z. Shao, H. Lu, H. Zhao, *Mol. BioSyst.* **2010**, *6*, 336–338.
[55] a) Z. Shao, Y. Luo, H. Zhao, *Mol. BioSyst.* **2011**, *7*, 1056–1059; b) Z. Shao, H. Zhao, in: *Methods Enzymol.*, Vol. 517, (Ed.: A. H. David), Academic Press, **2012**, pp 203–224.
[56] S. D. Colloms, C. A. Merrick, F. J. Olorunniji, W. M. Stark, M. C. M. Smith, A. Osbourn, J. D. Keasling, S. J. Rosser, *Nucleic Acids Res.* **2013**, *42*, e23.
[57] N. G. A. Kuijpers, S. Chroumpi, T. Vos, D. Solis-Escalante, L. Bosman, J. T. Pronk, J.-M. Daran, P. Daran-Lapujade, *FEMS Yeast Res.* **2013**, *13*, 769–781.
[58] a) T. Fehér, V. Burland, G. Pósfai, *J. Biotechnol.* **2012**, *160*, 72–79; b) J. T. MacDonald, C. Barnes, R. I. Kitney, P. S. Freemont, G.-B. V. Stan, *Integr. Biol.* **2011**, *3*, 97–108; c) D. Densmore, T. H.-C. Hsiau, J. T. Kittleson, W. DeLoache, C. Batten, J. C. Anderson, *Nucleic Acids Res.* **2010**, *38*, 2607–2616; d) H. M. Salis, E. A. Mirsky, C. A. Voigt, *Nat. Biotechnol.* **2009**, *27*, 946–950; e) Y. Cai, B. Hartnett, C. Gustafsson, J. Peccoud, *Bioinformatics* **2007**, *23*, 2760–2767.
[59] a) K. E. J. Tyo, P. K. Ajikumar, G. Stephanopoulos, *Nat. Biotechnol.* **2009**, *27*, 760–765; b) A. H. Kachroo, M. Jayaram, P. A. Rowley, *Nat. Biotechnol.* **2009**, *27*, 729–731.
[60] V. C. Dunkel, E. Zeiger, D. Brusick, E. McCoy, D. McGregor, K. Mortelmans, H. S. Rosenkranz, V. F. Simmon, *Environ. Mutagen.* **1985**, *7*, 1–19.
[61] E. M. Witkin, *Bacteriol. Rev.* **1976**, *40*, 869–907.
[62] R. Simon, U. Priefer, A. Puhler, *Nat. Biotechnol.* **1983**, *1*, 784–791.
[63] C. M. Radding, J. Flory, A. Wu, R. Kahn, C. DasGupta, D. Gonda, M. Bianchi, S. S. Tsang, *Cold Spring Harbor Symp. Quant. Biol.* **1983**, *47*, 821–828.
[64] a) H. Alper, K. Miyaoku, G. Stephanopoulos, *Nat. Biotechnol.* **2005**, *23*, 612–616; b) D. Koma, H. Yamanaka, K. Moriyoshi, T. Ohmoto, K. Sakai, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 815–829.
[65] K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640–6645.
[66] T. E. Kuhlman, E. C. Cox, *Nucleic Acids Res.* **2010**, *38*, e92–e92.
[67] a) M. Orford, M. Nefedov, J. Vadolas, F. Zaibak, R. Williamson, P. A. Ioannou, *Nucleic Acids Res.* **2000**, *28*, e84; b) B. J. Yu, K. H. Kang, J. H. Lee, B. H. Sung, M. S. Kim, S. C. Kim, *Nucleic Acids Res.* **2008**, *36*, e84; c) S. Sabri, J. A. Steen, M. Bongers, L. K. Nielsen, C. E. Vickers, *Microb. Cell Fact.* **2013**, *12*, 60.

[68] C. W. Song, S. Y. Lee, *Biotechnol. J.* **2013**, *8*, 776–784.
[69] P. N. Kuduvalli, R. Mitra, N. L. Craig, *Nucleic Acids Res.* **2005**, *33*, 857–863.
[70] G. McKenzie, N. Craig, *BMC Microbiol.* **2006**, *6*, 39.
[71] L. C. Thomason, N. Costantino, D. L. Court, in: *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., **2001**.
[72] D. Koma, H. Yamanaka, K. Moriyoshi, T. Ohmoto, K. Sakai, *Appl. Environ. Microbiol.* **2012**, *78*, 6203–6216.
[73] B. S. Laursen, H. P. Sørensen, K. K. Mortensen, H. U. Sperling-Petersen, *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101–123.
[74] H. Chen, M. Bjerknes, R. Kumar, E. Jay, *Nucleic Acids Res.* **1994**, *22*, 4953–4957.
[75] A. A. Ublinskaya, V. V. Samsonov, S. V. Mashko, N. V. Stoynova, *J. Microbiol. Methods* **2012**, *89*, 167–173.
[76] a) M. Z. Li, S. J. Elledge, *Nat. Genet.* **2005**, *37*, 311–319; b) A. Kuzminov, *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 751–813.
[77] N. Minaeva, E. Gak, D. Zimenkov, A. Skorokhodova, I. Biryukova, S. Mashko, *BMC Biotechnol.* **2008**, *8*, 63.
[78] E. Martínez-García, T. Aparicio, V. de Lorenzo, P. I. Nikel, *Front. Bioeng. Biotechnol.* **2014**, *2*, 46.
[79] F. Baumgärtner, L. Seitz, G. A. Sprenger, C. Albermann, *Microb. Cell Fact.* **2013**, *12*, 40.
[80] T. Gaj, C. A. Gersbach, C. F. Barbas III, *Trends Biotechnol.* **2013**, *31*, 397–405.
[81] a) J. Zhong, M. Karberg, A. M. Lambowitz, *Nucleic Acids Res.* **2003**, *31*, 1656–1664; b) J. Perutka, W. Wang, D. Goerlitz, A. M. Lambowitz, *J. Mol. Biol.* **2004**, *336*, 421–439; c) J. Yao, J. Zhong, A. M. Lambowitz, *Nucleic Acids Res.* **2005**, *33*, 3351–3362.
[82] D. Na, S. M. Yoo, H. Chung, H. Park, J. H. Park, S. Y. Lee, *Nat. Biotechnol.* **2013**, *31*, 170–174.
[83] T. Brautaset, R. Lale, S. Valla, *Microb. Biotechnol.* **2009**, *2*, 15–30.
[84] S. Balzer, V. Kucharova, J. Megerle, R. Lale, T. Brautaset, S. Valla, *Microb. Cell Fact.* **2013**, *12*, 26.
[85] a) J. W. Dubendorff, F. W. Studier, *J. Mol. Biol.* **1991**, *219*, 61–68; b) J. W. Dubendorf, F. W. Studier, *J. Mol. Biol.* **1991**, *219*, 45–59.
[86] a) M. Giacalone, A. Gentile, B. Lovitt, N. Berkley, C. Gunderson, M. Surber, *BioTechniques* **2006**, *40*, 355–364; b) Y. J. Choi, L. Morel, T. Le François, D. Bourque, L. Bourget, D. Groleau, B. Massie, C. B. Míguez, *Appl. Environ. Microbiol.* **2010**, *76*, 5058–5066.
[87] X.-X. Wei, Z.-Y. Shi, M.-Q. Yuan, G.-Q. Chen, *Appl. Microbiol. Biotechnol.* **2009**, *82*, 703–712.
[88] a) S. Shimizu-Sato, E. Huq, J. M. Tepperman, P. H. Quail, *Nat. Biotechnol.* **2002**, *20*, 1041–1044; b) D. Camsund, P. Lindblad, A. Jaramillo, *Biotechnol. J.* **2011**, *6*, 826–836; c) U. Krauss, T. Drepper, K.-E. Jaeger, *Chem. Eur. J.* **2011**, *17*, 2552–2560.
[89] R. H. Dahl, F. Zhang, J. Alonso-Gutierrez, E. Baidoo, T. S. Batth, A. M. Redding-Johanson, C. J. Petzold, A. Mukhopadhyay, T. S. Lee, P. D. Adams, J. D. Keasling, *Nat. Biotechnol.* **2013**, *31*, 1039–1046.
[90] a) C. M. Waters, B. L. Bassler, *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 319–346; b) W.-L. Ng, B. L. Bassler, *Annu. Rev. Genet.* **2009**, *43*, 197–222; c) R. Popat, D. M. Cornforth, L. McNally, S. P. Brown, *J. R. Soc. Interface* **2015**, *12*, DOI: 10.1098/rsif.2014.0882.
[91] C.-Y. Tsao, S. Hooshangi, H.-C. Wu, J. J. Valdes, W. E. Bentley, *Metab. Eng.* **2010**, *12*, 291–297.
[92] a) H. Alper, C. Fischer, E. Nevoigt, G. Stephanopoulos, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12678–12683; b) J. R. Anthony, L. C. Anthony, F. Nowroozi, G. Kwon, J. D. Newman, J. D. Keasling, *Metab. Eng.* **2009**, *11*, 13–19; c) M. De Mey, J. Maertens, S. Boogmans, W. K. Soetaert, E. J. Vandamme, R. Cunin, M. R. Foulquie-Moreno, *BMC Biotechnol.* **2010**, *10*, 26; d) J. Blazeck, R. Garg, B. Reed, H. S. Alper, *Biotechnol. Bioeng.* **2012**, *109*, 2884–2895.
[93] a) A. G. Vitreschak, D. A. Rodionov, A. A. Mironov, M. S. Gelfand, *Trends Genet.* **2004**, *20*, 44–50; b) B. J. Tucker, R. R. Breaker, *Curr. Opin. Struct. Biol.* **2005**, *15*, 342–348; c) D. M. Mishler, S. Topp, C. M. K. Reynoso, J. P. Gallivan, *Curr. Opin. Biotechnol.* **2010**, *21*, 653–656; d) A. D. Garst, A. L. Edwards, R. T. Batey, *Cold Spring Harbor Perspect. Biol.* **2011**, *3*, a003533.
[94] F. Jacob, J. Monod, *J. Mol. Biol.* **1961**, *3*, 318–356.
[95] J.-M. Claverie, *Science* **2001**, *291*, 1255–1257.
[96] E. P. C. Rocha, *Annu. Rev. Genet.* **2008**, *42*, 211–233.
[97] S. Normark, S. Bergstrom, T. Edlund, T. Grundstrom, B. Jaurin, F. P. Lindberg, O. Olsson, *Annu. Rev. Genet.* **1983**, *17*, 499–525.
[98] a) C. D. Smolke, T. A. Carrier, J. D. Keasling, *Appl. Environ. Microbiol.* **2000**, *66*, 5399–5405; b) C. D. Smolke, J. D. Keasling, *Biotechnol. Bioeng.* **2002**, *80*, 762–776; c) L. Berg, R. Lale, I. Bakke, N. Burroughs, S. Valla, *Microb. Biotechnol.* **2009**, *2*, 379–389.
[99] B. F. Pfleger, D. J. Pitera, C. D. Smolke, J. D. Keasling, *Nat. Biotechnol.* **2006**, *24*, 1027–1032.
[100] T. M. Henkin, *Annu. Rev. Genet.* **1996**, *30*, 35–57.
[101] H. Aiba, A. Hanamura, H. Yamano, *J. Biol. Chem.* **1991**, *266*, 1721–1727.
[102] J. A. Bernstein, A. B. Khodursky, P.-H. Lin, S. Lin-Chao, S. N. Cohen, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9697–9702.
[103] a) M. Kozak, *Gene* **1999**, *234*, 187–208; b) M. Kozak, *Nucleic Acids Res.* **2005**, *33*, 6593–6602.
[104] M. H. de Smit, J. van Duin, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 7668–7672.
[105] a) D. Na, D. Lee, *Bioinformatics* **2010**, *26*, 2633–2634; b) D. Na, S. Lee, D. Lee, *BMC Syst. Biol.* **2010**, *4*, 71.
[106] H. H. Wang, F. J. Isaacs, P. A. Carr, Z. Z. Sun, G. Xu, C. R. Forest, G. M. Church, *Nature* **2009**, *460*, 894–898.
[107] a) P. J. Westfall, D. J. Pitera, J. R. Lenihan, D. Eng, F. X. Woolard, R. Regentin, T. Horning, H. Tsuruta, D. J. Melis, A. Owens, S. Fickes, D. Diola, K. R. Benjamin, J. D. Keasling, M. D. Leavell, D. J. McPhee, N. S. Renninger, J. D. Newman, C. J. Paddon, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E111–118; b) K. Zhou, R. Zou, C. Zhang, G. Stephanopoulos, H.-P. Too, *Biotechnol. Bioeng.* **2013**, *110*, 2556–2561.
[108] F. F. Nowroozi, E. E. K. Baidoo, S. Ermakov, A. M. Redding-Johanson, T. S. Batth, C. J. Petzold, J. D. Keasling, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 1567–1581.
[109] a) B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, C. Khosla, *Science* **2001**, *291*, 1790–1792; b) G. Cannarozzi, N. N. Schraudolph, M. Faty, P. von Rohr, M. T. Friberg, A. C. Roth, P. Gonnet, G. Gonnet, Y.

Barral, *Cell* **2010**, *141*, 355–367; c) K. Fredrick, M. Ibba, *Cell* **2010**, *141*, 227–229; d) T. Tuller, A. Carmi, K. Vestsigian, S. Navon, Y. Dorfan, J. Zaborske, T. Pan, O. Dahan, I. Furman, Y. Pilpel, *Cell* **2010**, *141*, 344–354.

[110] a) S. R. Eddy, *Nat. Rev. Genet.* **2001**, *2*, 919–929; b) R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **1993**, *75*, 843–854; c) E. G. H. Wagner, R. W. Simons, *Annu. Rev. Microbiol.* **1994**, *48*, 713–742; d) P. Stougaard, S. Molin, K. Nordström, *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 6008–6012.

[111] A. S. Mironov, I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov, E. Nudler, *Cell* **2002**, *111*, 747–756.

[112] M. Morita, M. Kanemori, H. Yanagi, T. Yura, *J. Bacteriol.* **1999**, *181*, 401–410.

[113] C.-H. Hoe, C. A. Raabe, T. S. Rozhdestvensky, T.-H. Tang, *Int. J. Med. Microbiol.* **2013**, *303*, 217–229.

[114] a) N. Majdalani, D. Hernandez, S. Gottesman, *Mol. Microbiol.* **2002**, *46*, 813–826; b) S. Gottesman, *Trends Genet.* **2005**, *21*, 399–404; c) T. Soper, P. Mandin, N. Majdalani, S. Gottesman, S. A. Woodson, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9602–9607.

[115] a) F. J. Isaacs, D. J. Dwyer, C. Ding, D. D. Pervouchine, C. R. Cantor, J. J. Collins, *Nat. Biotechnol.* **2004**, *22*, 841–847; b) F. J. Isaacs, J. J. Collins, *Nat. Biotechol.* **2005**, *23*, 306–307; c) R. Lutz, H. Bujard, *Nucleic Acids Res.* **1997**, *25*, 1203–1210; d) G. Rodrigo, T. E. Landrain, A. Jaramillo, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15271–15276.

[116] Z. Kang, X. Wang, Y. Li, Q. Wang, Q. Qi, *Biotechnol. Lett.* **2012**, *34*, 527–531.

[117] Z. Kang, Q. Wang, H. Zhang, Q. Qi, *Appl. Microbiol. Biotechnol.* **2008**, *79*, 203–208.

[118] D. Juminaga, E. E. K. Baidoo, A. M. Redding-Johanson, T. S. Batth, H. Burd, A. Mukhopadhyay, C. J. Petzold, J. D. Keasling, *Appl. Environ. Microbiol.* **2012**, *78*, 89–98.

[119] a) M. T. Reetz, *J. Am. Chem. Soc.* **2013**, *135*, 12480–12496; b) J. Muschiol, C. Peters, N. Oberleitner, M. D. Mihovilovic, U. T. Bornscheuer, F. Rudroff, *Chem. Commun.* **2015**, *51*, 5798–5811.

[120] a) C. Jäckel, P. Kast, D. Hilvert, *Annu. Rev. Biophys.* **2008**, *37*, 153–173; b) C. A. Tracewell, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2009**, *13*, 3–9.

[121] a) G. L. Butterfoss, B. Kuhlman, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 49–65; b) A. M. Poole, R. Ranganathan, *Curr. Opin. Struct. Biol.* **2006**, *16*, 508–513; c) S.-g. Kang, J. G. Saven, *Curr. Opin. Chem. Biol.* **2007**, *11*, 329–334.

[122] T.-W. Wang, H. Zhu, X.-Y. Ma, T. Zhang, Y.-S. Ma, D.-Z. Wei, *Mol. Biotechnol.* **2006**, *34*, 55–68.

[123] A. Davidson, in: *Protein Design*, Vol. 340, (Eds.: R. Guerois, M. de La Paz), Humana Press, **2006**, pp 171–181.

[124] K. B. Otte, B. Hauer, *Curr. Opin. Biotechnol.* **2015**, *35*, 16–22.

[125] 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*, 338–344.

[126] a) W. P. C. Stemmer, *Nature* **1994**, *370*, 389–391; b) J. E. Ness, M. Welch, L. Giver, M. Bueno, J. R. Cherry, T. V. Borchert, W. P. C. Stemmer, J. Minshull, *Nat. Biotechnol.* **1999**, *17*, 893–896.

[127] 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*, 81–86.

[128] B. Gallwitz, *Vasc. Health Risk Manage.* **2007**, *3*, 203–210.

[129] 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*, 305–309.

[130] J. Alliot, E. Gravel, F. Pillon, D.-A. Buisson, M. Nicolas, E. Doris, *Chem. Commun.* **2012**, *48*, 8111–8113.

[131] Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho, F. H. Arnold, *Angew. Chem.* **2014**, *126*, 6928–6931; *Angew. Chem. Int. Ed.* **2014**, *53*, 6810–6813.

[132] Z.-G. Zhang, R. Lonsdale, J. Sanchis, M. T. Reetz, *J. Am. Chem. Soc.* **2014**, *136*, 17262–17272.

[133] a) R. J. Conrado, J. D. Varner, M. P. DeLisa, *Curr. Opin. Biotechnol.* **2008**, *19*, 492–499; b) S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. H. Narayan, K. Hakansson, D. H. Sherman, J. L. Smith, G. Skiniotis, *Nature* **2014**, *510*, 512–517.

[134] J.-L. Lin, L. Palomec, I. Wheeldon, *ACS Catal.* **2013**, *4*, 505–511.

[135] a) W. R. Whitaker, J. E. Dueber, in: *Methods Enzymol.* Vol. 497, (Ed.: V. Chris), Academic Press, **2011**, pp 447–468; b) H. Hirakawa, T. Haga, T. Nagamune, *Top. Catal.* **2012**, *55*, 1124–1137.

[136] a) J. Fu, M. Liu, Y. Liu, N. W. Woodbury, H. Yan, *J. Am. Chem. Soc.* **2012**, *134*, 5516–5519; b) R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnšek, N. Tomšič, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Benčina, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala, M. P. DeLisa, *Nucleic Acids Res.* **2012**, *40*, 1879–1889.

[137] a) C. J. Delebecque, P. A. Silver, A. B. Lindner, *Nat. Protoc.* **2012**, *7*, 1797–1807; b) C. J. Delebecque, A. B. Lindner, P. A. Silver, F. A. Aldaye, *Science* **2011**, *333*, 470–474.

[138] a) A. J. Dzieciol, S. Mann, *Chem. Soc. Rev.* **2012**, *41*, 79–85; b) R. J. R. W. Peters, M. Marguet, S. Marais, M. W. Fraaije, J. C. M. van Hest, S. Lecommandoux, *Angew. Chem.* **2014**, *126*, 150–154; *Angew. Chem. Int. Ed.* **2014**, *53*, 146–150; c) M. Marguet, C. Bonduelle, S. Lecommandoux, *Chem. Soc. Rev.* **2013**, *42*, 512–529.

[139] L. Kizer, D. J. Pitera, B. F. Pfleger, J. D. Keasling, *Appl. Environ. Microbiol.* **2008**, *74*, 3229–3241.

[140] O. I. Wilner, Y. Weizmann, R. Gill, O. Lioubashevski, R. Freeman, I. Willner, *Nat. Nanotechnol.* **2009**, *4*, 249–254.

[141] G. Sachdeva, A. Garg, D. Godding, J. C. Way, P. A. Silver, *Nucleic Acids Res.* **2014**, *42*, 9493–9503.

[142] R. Chen, Q. Chen, H. Kim, K.-H. Siu, Q. Sun, S.-L. Tsai, W. Chen, *Curr. Opin. Biotechnol.* **2014**, *28*, 59–68.

[143] a) P. L. Howard, M. C. Chia, S. Del Rizzo, F.-F. Liu, T. Pawson, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11267–11272; b) Y.-M. Lim, S. Wong, G. Lau, O. N. Witte, J.

Colicelli, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12233–12238; c) G. Barnea, W. Strapps, G. Herrada, Y. Berman, J. Ong, B. Kloss, R. Axel, K. J. Lee, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 64–69.
[144] L. Albertsen, Y. Chen, L. S. Bach, S. Rattleff, J. Maury, S. Brix, J. Nielsen, U. H. Mortensen, *Appl. Environ. Microbiol.* **2011**, *77*, 1033–1040.
[145] a) C. J. Bashor, N. C. Helman, S. Yan, W. A. Lim, *Science* **2008**, *319*, 1539–1543; b) S.-H. Park, A. Zarrinpar, W. A. Lim, *Science* **2003**, *299*, 1061–1064; c) P. Wei, W. W. Wong, J. S. Park, E. E. Corcoran, S. G. Peisajovich, J. J. Onuffer, A. Weiss, W. A. Lim, *Nature* **2012**, *488*, 384–388.
[146] W. R. Whitaker, S. A. Davis, A. P. Arkin, J. E. Dueber, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 18090–18095.
[147] A. M. Weeks, M. C. Chang, *Biochemistry* **2011**, *50*, 5404–5418.
[148] K. M. Dombek, L. O. Ingram, *Appl. Environ. Microbiol.* **1987**, *53*, 1286–1291.
[149] M. Sauer, D. Porro, D. Mattanovich, P. Branduardi, *Trends Biotechnol.* **2008**, *26*, 100–108.
[150] X. Chen, M. Li, L. Zhou, W. Shen, G. Algasan, Y. Fan, Z. Wang, *Bioresour. Technol.* **2014**, *166*, 64–71.
[151] S. Y. Lee, J. H. Park, S. H. Jang, L. K. Nielsen, J. Kim, K. S. Jung, *Biotechnol. Bioeng.* **2008**, *101*, 209–228.
[152] E.-B. Goh, E. E. K. Baidoo, H. Burd, T. S. Lee, J. D. Keasling, H. R. Beller, *Metab. Eng.* **2014**, *26*, 67–76.
[153] a) G. W. Huisman, S. J. Collier, *Curr. Opin. Chem. Biol.* **2013**, *17*, 284–292; b) A. Berkessel, C. Rollmann, F. Chamouleau, S. Labs, O. May, H. Gröger, *Adv. Synth. Catal.* **2007**, *349*, 2697–2704.
[154] Z. Q. Lin, Z. B. Xu, Y. F. Li, Z. W. Wang, T. Chen, X. M. Zhao, *Microb. Cell Fact.* **2014**, *13*; 104 DOI: 10.1168/s12934-014-5.
[155] M. A. Campodonico, B. A. Andrews, J. A. Asenjo, B. O. Palsson, A. M. Feist, *Metab. Eng.* **2014**, *25*, 140–158.
[156] K. Tummler, T. Lubitz, M. Schelker, E. Klipp, *FEBS J.* **2014**, *281*, 549–571.
[157] J. S. Edwards, R. U. Ibarra, B. O. Palsson, *Nat. Biotechnol.* **2001**, *19*, 125–130.
[158] C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay, R. L. Stevens, *Nat. Biotechnol.* **2010**, *28*, 977–922.
[159] J. Schellenberger, J. O. Park, T. M. Conrad, B. O. Palsson, *BMC Bioinf.* **2010**, *11*; 213 DOI: 10.1168/1471-2105-11-213.
[160] J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke, B. O. Palsson, *Nat. Protoc.* **2011**, *6*, 1290–1307.
[161] M. Lakshmanan, G. Koh, B. K. S. Chung, D. Y. Lee, *Briefings Bioinf.* **2014**, *15*, 108–122.
[162] J. D. Orth, I. Thiele, B. O. Palsson, *Nat. Biotechnol.* **2010**, *28*, 245–248.
[163] R. Schuetz, L. Kuepfer, U. Sauer, *Mol. Syst. Biol.* **2007**, *3*, 119.
[164] L. Kuepfer, U. Sauer, L. M. Blank, *Genome Res.* **2005**, *15*, 1421–1430.
[165] K. Hoffner, S. M. Harwood, P. I. Barton, *Biotechnol. Bioeng.* **2013**, *110*, 792–802.
[166] C. W. Song, D. I. Kim, S. Choi, J. W. Jang, S. Y. Lee, *Biotechnol. Bioeng.* **2013**, *110*, 2025–2034.
[167] M. Weiner, J. Tröndle, C. Albermann, G. A. Sprenger, D. Weuster-Botz, *Biotechnol. Bioeng.* **2014**, *111*, 1406–1416.
[168] K. Tokuyama, S. Ohno, K. Yoshikawa, T. Hirasawa, S. Tanaka, C. Furusawa, H. Shimizu, *Microb. Cell Fact.* **2014**, *13*; 64 DOI: 10.1168/1475-2859-13-64.
[169] N. Zamboni, E. Fischer, U. Sauer, *BMC Bioinf.* **2005**, *6*; 209 DOI: 10.1168/1471-2105-6-209.
[170] W. Wiechert, M. Möllney, S. Petersen, A. A. de Graaf, *Metab. Eng.* **2001**, *3*, 265–283.
[171] L.-E. Quek, C. Wittmann, L. Nielsen, J. Kromer, *Microb. Cell Fact.* **2009**, *8*, 25.
[172] T. Dandekar, A. Fieselmann, S. Majeed, Z. Ahmed, *Briefings Bioinf.* **2014**, *15*, 91–107.
[173] M. Antoniewicz, *J. Ind. Microbiol. Biotechnol.* **2015**, 1–9.
[174] a) E. Fischer, N. Zamboni, U. Sauer, *Anal. Biochem.* **2004**, *325*, 308–316; b) E. Fischer, U. Sauer, *Eur. J. Biochem.* **2003**, *270*, 880–891; c) N. Zamboni, S.-M. Fendt, M. Ruhl, U. Sauer, *Nat. Protoc.* **2009**, *4*, 878–892.
[175] R. R. Bommareddy, Z. Chen, S. Rappert, A.-P. Zeng, *Metab. Eng.* **2014**, *25,* 30–37.
[176] S. B. Crown, C. P. Long, M. R. Antoniewicz, *Metab. Eng.* **2015**, *28*, 151–158.
[177] K. Schroer, B. Zelic, M. Oldiges, S. Lütz, *Biotechnol. Bioeng.* **2009**, *104*, 251–260.
[178] L. Wu, M. R. Mashego, J. C. van Dam, A. M. Proell, J. L. Vinke, C. Ras, W. A. van Winden, W. M. van Gulik, J. J. Heijnen, *Anal. Biochem.* **2005**, *336*, 164–171.
[179] W. M. van Gulik, *Curr. Opin. Biotechnol.* **2010**, *21*, 27–34.
[180] A. B. Canelas, C. Ras, A. ten Pierick, J. C. van Dam, J. J. Heijnen, W. M. Van Gulik, *Metabolomics* **2008**, *4*, 226–239.
[181] H. Link, J. M. Buescher, U. Sauer, in: *Methods in Microbiology*, Vol. 39, (Eds.: H. Colin, W. Anil), Academic Press, **2012**, pp 127–150.
[182] J. Rucker, J. Paul, B. A. Pfeifer, K. Lee, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 2753–2759.
[183] Y. Soma, K. Tsuruno, M. Wada, A. Yokota, T. Hanai, *Metab. Eng.* **2014**, *23*, 175–184.
[184] a) Y.-C. He, F. Liu, D.-P. Zhang, S. Gao, Z.-Q. Li, Z.-C. Tao, C.-L. Ma, *Appl. Biochem. Biotechnol.* **2014**, 1–12; b) S. Gough, M. Deshpande, M. Scher, J. P. N. Rosazza, *Biotechnol. Lett.* **2001**, *23*, 1535–1537; c) D. Gobinath, S. Prapulla, *J. Food Sci. Technol.* **2014**, 1–7.
[185] H. J. Park, J. Jung, H. Choi, K. N. Uhm, H. K. Kim, *J. Microbiol. Biotechnol.* **2010**, *20*, 1300–1306.
[186] G. Thedei, D. P. S. Leitao, M. Bolean, T. P. Paulino, A. C. C. Spadaro, P. Ciancaglini, *Braz. J. Med. Biol. Res.* **2008**, *41*, 1047–1053.
[187] C. Grant, D. Deszcz, Y.-C. Wei, R. J. Martinez-Torres, P. Morris, T. Folliard, R. Sreenivasan, J. Ward, P. Dalby, J. M. Woodley, F. Baganz, *Sci. Rep.* **2014**, *4*.
[188] S. Panke, A. Meyer, C. M. Huber, B. Witholt, M. G. Wubbolts, *Appl. Environ. Microbiol.* **1999**, *65*, 2324–2332.
[189] M. G. Wubbolts, J. Hoven, B. Melgert, B. Witholt, *Enzyme Microb. Technol.* **1994**, *16*, 887–894.

[190] H. Nikaido, *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 593–656.
[191] S. Kojima, H. Nikaido, *J. Biol. Chem.* **2014**, *289*, 26464–26473.
[192] Z. Xiao, C. Lv, C. Gao, J. Qin, C. Ma, Z. Liu, P. Liu, L. Li, P. Xu, *PLoS ONE* **2010**, *5*, 8860.
[193] Y. Wang, K. Y. San, G. N. Bennett, *Curr. Opin. Biotechnol.* **2013**, *24*, 994–999.
[194] J. W. Chin, P. C. Cirino, *Biotechnol. Prog.* **2011**, *27*, 333–341.
[195] S. Siedler, S. Bringer, L. M. Blank, M. Bott, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1459–1467.
[196] A. Shi, X. Zhu, J. Lu, X. Zhang, Y. Ma, *Metab. Eng.* **2013**, *16*, 1–10.
[197] R. Liu, L. Liang, M. Wu, K. Chen, M. Jiang, J. Ma, P. Wei, P. Ouyang, *Biochem. Eng. J.* **2013**, *79*, 77–83.
[198] A. Singh, M. D. Lynch, R. T. Gill, *Metab. Eng.* **2009**, *11*, 347–354.
[199] S. Klatte, V. F. Wendisch, *Bioorg. Med. Chem.* **2014**, *22*, 5578–5585.
[200] D. E. Torres Pazmiño, A. Riebel, J. de Lange, F. Rudroff, M. D. Mihovilovic, M. W. Fraaije, *ChemBioChem* **2009**, *10*, 2595–2598.
[201] J.-W. Song, E.-Y. Jeon, D.-H. Song, H.-Y. Jang, U. T. Bornscheuer, D.-K. Oh, J.-B. Park, *Angew. Chem.* **2013**, *125*, 2594–2597; *Angew. Chem. Int. Ed.* **2013**, *52*, 2534–2537.
[202] M. Schrewe, N. Ladkau, B. Bühler, A. Schmid, *Adv. Synth. Catal.* **2013**, *355*, 1693–1697.
[203] M. Schrewe, M. K. Julsing, K. Lange, E. Czarnotta, A. Schmid, B. Bühler, *Biotechnol. Bioeng.* **2014**, *111*, 1820–1830.
[204] A. S. Bommarius, B. R. Riebel, in: *Biocatalysis*, Wiley-VCH, **2005**, pp 91–134.
[205] a) B. M. Nestl, B. Hauer, *Nat. Chem. Biol.* **2013**, *9*, 470–471; b) F. H. Arnold, *Nature* **2001**, *409*, 253–257.
[206] N. J. Turner, E. O'Reilly, *Nat. Chem. Biol.* **2013**, *9*, 285–288.
[207] a) J. H. Shin, H. U. Kim, D. I. Kim, S. Y. Lee, *Biotechnol. Adv.* **2013**, *31*, 925–935; b) J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi, S. Y. Lee, *Nat. Chem. Biol.* **2012**, *8*, 536–546; c) G. M. Rodriguez, S. Atsumi, *Microb. Cell Fact.* **2012**, *11*, 90.
[208] K. Napora-Wijata, K. Robins, A. Osorio-Lozada, M. Winkler, *ChemCatChem* **2014**, *6*, 1089–1095.
[209] K. B. Otte, J. Kittelberger, M. Kirtz, B. M. Nestl, B. Hauer, *ChemCatChem* **2014**, *6*, 1003–1009.
[210] K. B. Otte, M. Kirtz, B. M. Nestl, B. Hauer, *ChemSusChem* **2013**, *6*, 2149–2156.
[211] D. Kuhn, B. Bühler, A. Schmid, *J. Ind. Microbiol. Biotechnol.* **2012**, *39*, 1125–1133.
[212] J. Liu, Z. Li, *ACS Catal.* **2013**, *3*, 908–911.
[213] J. Zhang, S. Wu, J. Wu, Z. Li, *ACS Catal.* **2014**, *4*, 51–58.
[214] J. Zhang, T. Xu, Z. Li, *Adv. Synth. Catal.* **2013**, *355*, 3147–3153.
[215] S. Wu, Y. Chen, Y. Xu, A. Li, Q. Xu, A. Glieder, Z. Li, *ACS Catal.* **2013**, *4*, 409–420.
[216] Y. Zhang, F. Buchholz, J. P. P. Muyrers, A. F. Stewart, *Nat. Genet.* **1998**, *20*, 123–128.
[217] K. Miura, Y. Tomioka, H. Suzuki, M. Yonezawa, T. Hishinuma, M. Mizugaki, *Biol. Pharm. Bull.* **1997**, *20*, 110–112.

