An-Ping Zeng   *Editor*

Fundamentals and Application of New Bioproduction Systems

Springer

137
Advances in Biochemical
Engineering/Biotechnology

Series editor

T. Scheper, Hannover, Germany

For further volumes:
http://www.springer.com/series/10

Aims and Scope

This book series reviews current trends in modern biotechnology and biochemical engineering. Its aim is to cover all aspects of these interdisciplinary disciplines, where knowledge, methods and expertise are required from chemistry, biochemistry, microbiology, molecular biology, chemical engineering and computer science.

Volumes are organized topically and provide a comprehensive discussion of developments in the field over the past 3–5 years. The series also discusses new discoveries and applications. Special volumes are dedicated to selected topics which focus on new biotechnological products and new processes for their synthesis and purification.

In general, volumes are edited by well-known guest editors. The series editor and publisher will, however, always be pleased to receive suggestions and supplementary information. Manuscripts are accepted in English.

In references, Advances in Biochemical Engineering/Biotechnology is abbreviated as *Adv. Biochem. Engin./Biotechnol.* and cited as a journal.

An-Ping Zeng
Editor

Fundamentals and Application of New Bioproduction Systems

With contributions by

Inés Ardao · Lisa Katharina Blaß · Uwe T. Bornscheuer
Frank Breitling · Andreas K. Brödel
Yun-Chien Cheng · Edgar Doersam · Elmar Heinzle
Ee Taek Hwang · U. Jandt · Steffen Krauser · Stefan Kubick
Fanny C. Liu · Felix F. Loeffler · Bastian Muenster
Alexander Nesterov-Mueller · Robert B. Quast
F. Ralf Bischoff · Steffen Rupp · Rita Sachse
Marlitt Stech · Jakob Striffler · Tsz Kin Tam
Christian Weyler · C. You · A.-P. Zeng · An-Ping Zeng
Y.H.-P. Zhang · Y.-H.Percival Zhang · Zhiguang Zhu

∲ Springer

Editor
An-Ping Zeng
Institut für Bioprozess- und Biosystemtechnik
Technische Universität Hamburg-Harburg
Hamburg
Germany

Preface

Currently, large-scale production of chemicals, fuels, biomaterials, and pharmaceutical products from renewable materials draws on biological systems and processes primarily based on microbial, mammalian, and plant cells. With a few exceptions, enzymes have been mainly used as biocatalysts for biotransformations to produce specialties and intermediates. While these bioproduction systems have many important advantages compared to chemical production processes, indeed these have contributed greatly to the health and welfare of human beings and to the protection of our environment, they have several inherent limitations. First, the product yield is often limited (maximum about 50 % by weight of substrate used even for primary products such as ethanol) due to the underlying biochemical principle of coupled oxidation and reduction of the substrate: much of the substrate is released as CO_2. Second, microorganisms are evolutionarily optimized for their growth and maintenance, not for the formation of a special product which we desire to have. Despite the tremendous advances in biological fundamentals and bioengineering techniques, it is today still difficult to precisely and predictably manipulate a microorganism (e.g., by metabolic engineering or synthetic biology) for the purpose of bioproduction. In addition to the complex regulation mechanisms at different molecular levels, a major challenge and barrier is the cell membrane or cell wall which prevents the accumulation or secretion of desired products. Adding to the fact that cellular production processes take place almost always in water, we have normally a much diluted "aqueous soup" as the product from a typical bioproduction process, leading to high energy and investment costs for the product recovery. Furthermore, complex substrates such as hemi or lignocellulosic materials which represent the most abundant and cheap renewable substrates can still not be efficiently used in present bioproduction systems using single microorganisms. This is one of the obstacles for the realization of a bio-based economy with the aim to move away from a fossil-based one. There are therefore needs to look for alternatives or new concepts of bioproduction systems. One of the new and alternative bioproduction systems is the cell-free biosynthesis which could, in principle, overcome many of the major constraints of present bioproduction systems. In general, the de-coupling of biosynthesis reactions or pathways from the cellular machinery can allow for much higher product yield and thus reduce by-product or waste generation. The biosynthesis could be more easily manipulated and also carried out in nonaqueous solutions. Biosynthesis outside the

cells could provide completely new possibilities which are unthinkable with the present bioproduction systems.

In this volume, eight contributions are lined up to review and illustrate some of the fundamental issues, tools, promises, and applications of cell-free bioproduction systems. The article of Loeffler et al. describes the latest developments in peptide synthesis in an array format and presents a novel synthesis method using a microelectronic chip printer. This technology represents a so-called "printing bioproduction system" and allows the precise and massive production of artificial molecules or molecular building blocks. Bornscheuer illustrates the concept of protein engineering and its potential application for novel bioproduction systems. Jandt et al. gives insight into molecular strategies and modeling approaches for understanding and designing compartmentalization and metabolic channeling in multienzymatic systems for efficient biosynthesis, also in the context of in-vitro biosynthesis. The article of Ardao et al. addresses multienzymatic reaction systems from a more technologically oriented point of view, including process design aspects, strategies of co-immobilization of multienzymes in nano/micro-sized materials. It also highlights recent developments in applying multienzyme microreactors for biosynthesis. Rupp gives an overview on recent developments in design and modeling of molecular circuits with cell-free protein synthesis technology and for in vitro biosynthesis. Furthermore, he illustrates the need and perspectives of developing novel reactor concepts for cell-free biotechnology. The article of Stech et al. presents recent advances for cell-free synthesis of glycoproteins, membrane proteins, and fluorescently labeled proteins. These are difficult-to-express proteins and require organelles or microsomal vesicles as functional modules for translocation and post-translational modification. This work builds a bridge to synthetic and chemical biology. The article of Zhu et al. reviews the current status of cell-free technology and its applications in the production of electricity and bioenergy (hydrogen, alcohols, and jet fuels). They also address issues related to the high costs and low stability of enzymes and cofactors as well as compromised optimal conditions for enzymes from different sources. These are often argued as critical issues against cell-free biosystems as means for large-scale biomanufacturing. The last article of this volume by Krauser et al. presents a new bioproduction system as a middle-way between living cells and cell-free systems, namely tailored and permeabilized cells for multienzymatic biosynthesis. Practical examples are described to use this probably more cost-efficient production system for the synthesis of phosphorylated carbohydrates, sugar nucleotides, and polyketides.

The reader may have noticed that the majority of the authors of this volume are from Germany. Among other reasons, this may reflect the desire and efforts of the German government to develop the so-called "Next Generation Biotechnological Processes" in its research program "Strategy Process—Biotechnology 2020+." Most of the authors from Germany are involved in this long-term research program. Indeed, many of the new bioproduction systems presented in this volume may really need a long-term research and development to be made mature for large-scale biomanufacturing. This long-term endeavor is, in my opinion, both

timely and economically necessary in view of the limitations of present bioproduction systems and the large hope for a successful bioeconomy. It is also scientifically both highly interesting and challenging. Of course, I should mention that cell-free biosynthesis is only one type of the new bioproduction systems under development. The upcoming era of the synthetic bio(techno)logy will certainly result in new bioproduction systems with synthetic cells, prototype cells, self-reproducing organelles, or even vesicles, which may have unique properties and advantages. The substrate basis and energy supply systems may also completely change, i.e., from sugars move to CO_2, H_2O, light, and electricity. We are living in an exciting time for developing new and efficient bioproduction systems for the sake of human beings and our environment.

I cordially thank the authors for their excellent contributions and hope that this volume will inspire the future development of new bioproduction systems.

Hamburg, August 2013   An-Ping Zeng

Contents

Printing Peptide Arrays with a Complementary Metal Oxide Semiconductor Chip

Felix F. Loeffler, Yun-Chien Cheng, Bastian Muenster, Jakob Striffler, Fanny C. Liu, F. Ralf Bischoff, Edgar Doersam, Frank Breitling and Alexander Nesterov-Mueller

Abstract In this chapter, we discuss the state-of-the-art peptide array technologies, comparing the spot technique, lithographical methods, and microelectronic chip-based approaches. Based on this analysis, we describe a novel peptide array synthesis method with a microelectronic chip printer. By means of a complementary metal oxide semiconductor chip, charged bioparticles can be patterned on its surface. The bioparticles serve as vehicles to transfer molecule monomers to specific synthesis spots. Our chip offers 16,384 pixel electrodes on its surface with a spot-to-spot pitch of 100 μm. By switching the voltage of each pixel between 0 and 100 V separately, it is possible to generate arbitrary particle patterns for combinatorial molecule synthesis. Afterwards, the patterned chip surface serves as a printing head to transfer the particle pattern from its surface to a synthesis substrate. We conducted a series of proof-of-principle experiments to synthesize high-density peptide arrays. Our solid phase synthesis approach is based on the 9-fluorenylmethoxycarbonyl protection group strategy. After melting the particles, embedded monomers diffuse to the surface and participate in the coupling reaction to the surface. The method demonstrated herein can be easily extended to the synthesis of more complicated artificial molecules by using bioparticles with artificial molecular building blocks. The possibility of synthesizing artificial

F. F. Loeffler · B. Muenster · J. Striffler · F. C. Liu · F. Breitling · A. Nesterov-Mueller (✉)
Karlsruhe Institute of Technology, Institute of Microstructure Technology,
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
e-mail: Alexander.Nesterov-Mueller@KIT.edu

F. F. Loeffler
e-mail: Felix.Loeffler@KIT.edu

F. Breitling
e-mail: Frank.Breitling@KIT.edu

F. Ralf Bischoff
German Cancer Research Center, Chip Based Peptide Libraries, Im Neuenheimer Feld 580,
69120 Heidelberg, Germany

Y.-C. Cheng · E. Doersam
Institute of Printing Science and Technology, Magdalenenstr. 2, 64289 Darmstadt, Germany

peptides was also shown in an experiment in which we patterned biotin particles in a high-density array format. These results open the road to the development of peptide-based functional modules for diverse applications in biotechnology.

Keywords Functional particle deposition · High-throughput screening · Peptide microarray · Solid-phase synthesis

Abbreviations

Ac$_2$O	Acetic anhydride
CMOS	Complementary metal oxide semiconductor
DIPEA	Diisopropylethylamine
DMF	Dimethylformamide
Fmoc	9-Fluorenylmethoxycarbonyl
MMA	Methyl methacrylate
mol	Mole(s)
OPfp	Pentafluorophenyl
PEGMA	Poly(ethylene glycol) methacrylate
PTFE	Polytetrafluoroethylene
Boc	*tert*-butyloxycarbonyl
TFA	Trifluoroacetic acid

Contents

1 Introduction[1]

In the fields of genomics and transcriptomics, microarrays have revolutionized assay applications [4, 5]. Nevertheless, significant research effort is invested into technological improvements [6, 7]. The fabrication of molecule libraries in a microarray format on solid surfaces allows for massively parallel high-throughput assays with frugal consumption of analyte. Yet, it requires sophisticated technologies to facilitate such high-density arrays. Although the synthesis of DNA or RNA arrays only requires four different building blocks, highly versatile methods for combinatorial synthesis are required to generate peptides or polysaccharides [8, 9]. In the case of natural peptides, 20 different building blocks (amino acid monomers) need to be processed in a fully combinatorial manner, yielding the possibility to assemble any desired peptide sequence.

The Nobel Prize awarded solid phase peptide synthesis was invented more than 40 years ago, when Merrifield consecutively coupled amino acid monomers to a growing peptide chain immobilized on a solid support [10]. His approach allowed for incubating the growing oligomer chain with excess monomers, which drives the coupling reaction near completeness and, thereby, results in repetitive coupling yields that routinely exceed 95 % during peptide synthesis. Because the growing peptide products remain tethered to the support throughout the synthesis, all nonreacted monomers are simply washed away. These two basic principles are the prerequisite for the cost-efficient synthesis of peptides in the array format. In this chapter, we describe the state-of-the-art peptide array technologies and present a novel method using a microelectronic chip printer.

2 State-of-the-Art Peptide Array Technologies

With the invention of the phage display technique in 1985 [11], the generation of large protein libraries was possible. A few years later, the first antibody libraries became feasible due to the emergence of new methods, such as antibody phage display and recombinant antibodies [12, 13]. The necessity for an efficient characterization of large numbers of antibodies predetermined the development of peptide arrays in the early 1990s. Interestingly, the development of array technologies was not only stimulated by the search for antibody epitopes, but it was also driven by novel powerful methods in genome analysis. In this section, we divided peptide array technologies according to one characteristic trait: the number of coupling cycles. As you will see, this very trait seems to be crucial for success.

---

[1] Parts of this book chapter have already been published [1–3] and were reprinted with permission.

2.1 One-Coupling-Cycle-Per-Monomer Methods

In 1992, Stephen Fodor founded the spinoff company Affymetrix, based on the so-called lithographic method for the generation of peptide arrays [14, 15]. However, soon afterwards, Affymetrix completely shifted its focus from peptide to oligo-nucleotide arrays and became a world leader in oligonucleotide array production [16]. Meanwhile oligonucleotide arrays with thousands of oligonucleotides per cm^2 are commercially available (Affymetrix, USA), which are used for purposes such as detecting genome-wide transcription activity [17] or linking gene variants with diseases [18].

The reason for this peptide-to-oligonucleotide shift is quite obviously due to a peptide-specific drawback inherent to all lithographic synthesis methods [19]. During lithographic synthesis, light selectively removes a photolabile transient protecting group at the end of the growing oligomer chains, but only in those areas defined by the lithographic mask. The whole array then is incubated with a solution of chemically activated monomers that react with the deprotected oligomer. After the coupling reaction, unreacted monomers are washed away. This cycle adds only one monomer to the growing oligomer chains; therefore, it must be repeated until all the different monomers from one synthesis layer have been added to the array. Then, the whole process is repeated to elongate the growing oligomers by another monomer.

Peptides are composed of 20 different amino acids, whereas oligonucleotides comprise only four different bases. This fact leads to the peptide-specific large number of coupling cycles intrinsic in all lithographic syntheses. Although only 4×10 coupling cycles are needed to generate a 10-mer oligonucleotide array, the 20 different amino acid monomers demand 20×10 coupling cycles to synthesize a 10-mer peptide array. Such a large number of coupling cycles is expensive and usually accompanied by accumulation of an intolerable amount of unwanted side reactions, causing a very low yield.

Pellois et al. used a photo acid in combination with standard Boc protection instead of a photolabile protection group [17]. Photo acids are neutral precursor molecules that are transformed into an acid when illuminated by light. Thereby, a two-dimensional pattern of light with very small feature sizes is translated into a corresponding pattern of acidic versus neutral areas. This leads to selective and very efficient cleavage of standard Boc groups in those areas illuminated by light. However, this one-coupling-cycle-per-monomer technology requires physical barriers to confine the acid and prevent its diffusion to neighboring spots. This is possible using microfluidic µParaflo reaction devices (LC Sciences, Houston, Texas, USA), for example, which have a current yield of about 4,000 different peptides per chip [8].

Even more than 20 years since the introduction of the lithographic method, there are still efforts being made to improve its efficiency for high-density peptide array synthesis. In 2005, Li et al. reported a new chemical procedure for the in situ addition of photo-cleavable protection groups to a growing peptide chain [20].

Recently, Buus et al. combined this new chemistry with a digital micromirror device and pointed out the theoretical possibility of synthesizing 2 million individual peptide spots per array [21]. It remains to be seen if this variant lithographic synthesis method indeed yields high-density arrays that display peptides of sufficient length and in good yield.

2.2 One-Coupling-Cycle-Per-Layer Methods

Although the lithographic method shifted from peptide to oligonucleotide array production, the spot technique for the synthesis of peptide arrays appeared at around the same time, causing an opposite "oligonucleotide-to-peptide" development [19].

In 1989, Southern [22] introduced his novel concept for the combinatorial in situ synthesis of oligonucleotide arrays , adopting Ekins' array concept of spotting many different presynthesized molecules in the array format [23]. The basic idea behind this endeavor was to use a very limited number of chemically activated monomers (four different bases) to synthesize many different oligomers on neighboring places. Thereby, he came very close to the goal of cheap and high-density oligonucleotide arrays. A few years later, Frank [24] adopted this concept for the synthesis of peptide arrays: He parallelized the Merrifield synthesis by patterning 20 different activated amino acid derivatives as small droplets on a flat two-dimensional surface. There, the chemically activated monomers react with the solid support, with each droplet defining a small reaction sphere. Consecutively printed layers result in the parallel growth of many different peptides, whereas the number of different spots per area is only dependent on the achievable miniaturization of individual spots. Over the years, Ronald Frank's spot synthesis earned a reputation of reliability and wide applicability and thus still dominates the field.

High-density in situ synthesized peptide arrays with densities that exceed 25 peptides per cm^2, however, are difficult to obtain with the spot synthesis, mainly due to the difficulty of handling tiny droplets, which tend to evaporate or spread on the surface of the array. Interestingly, high-density peptide arrays manufactured with an inkjet printer were patented as early as 1994 [25]. Although oligonucleotide arrays are commercially available through Agilent's SurePrint technology [26], corresponding high-density peptide arrays have not been reported yet. This striking discrepancy might be due to the solvents required for peptide synthesis. These solvents are usually viscous, which makes it difficult to print them with inkjet or piezoelectric printers. The attempts of a world leader in printing technology—Hewlett-Packard (Palo Alto, California, USA)—in developing an inkjet printer for peptide arrays in 2002–2005 have also failed.[2]

---

[2] Stasiak, J.W., Hewlett-Packard, *Printable Nano Electronics*, COMS 2010, Aug 29th–Sept 2nd, Albuquerque, USA.

In analogy to the peptide laser printer, which is discussed elsewhere [19, 27], we describe in this book chapter the development and application of a novel microelectronic chip printer, which overcomes this peptide-specific drawback. In this particle-based method, only one coupling cycle per amino acid layer is required for the synthesis of peptide arrays. The chip technique [2, 28, 29] offers spot densities that approach those of lithographic methods. Because of the mentioned drawbacks of lithographic methods regarding the amount of coupling cycle steps and the coupling yield, we immobilized activated monomers in solid particles. Thus, we gained several orders of magnitude in chemical stability of the monomers compared to activated monomers in solvents [27, 28]. Addressing the monomers by patterning the chip surface with different particle types and subsequent printing (i.e. transferring) to a glass slide, we can elongate several thousand different biomolecule types layer-by-layer simultaneously. We melt the particle matrix to initiate the monomer coupling onto the glass slide, converting the particles into a gel-like phase and preserving the spatial pattern resolution. The integrity of these synthesized oligomers has been shown previously [27].

The length of the in situ synthesized oligomer is only limited by the coupling yield as in standard combinatorial chemistry. Furthermore, this technique has the potential for wide application in any combinatorial chemistry approach, as long as the monomeric building blocks are compatible with the chemical steps. With low space requirements for the experimental setup, our method is easily extendable to up to 40 (e.g. D- and L- amino acids) or more monomeric building blocks, which allows for a high combinatorial freedom.

In the following, we show in a proof-of-principle peptide array fabrication that the parallel synthesis of 16,384 different peptides (10,000 spots/cm^2) onto the surface of a glass slide is in fact possible using our chip printer approach.

3 New Materials and Methods

3.1 Basic chip principle

The microelectronic complementary metal oxide semiconductor (CMOS) chip, shown in Fig. 1a and b, features a total of 16,384 (128 × 128) pixel electrodes on its surface with a spot-to-spot pitch of 100 μm and a density of 10,000 pixels per cm^2. For combinatorial particle deposition, the chip is programmed to exhibit a voltage pattern on its surface (Fig. 1, c1). Then, the chip is exposed to an aerosol of microparticles: Particles are only attracted and deposited onto those pixel electrodes that exhibit an electric field. The voltage pattern is thus transformed into the according particle pattern with high precision (Fig. 1, c2, 3) [30].

A key feature in this process is that particles will only reach the chip surface, where the local electric field of a switched-on electrode attracts particles. Once particles have contact with the surface, adhesion forces become so dominant that
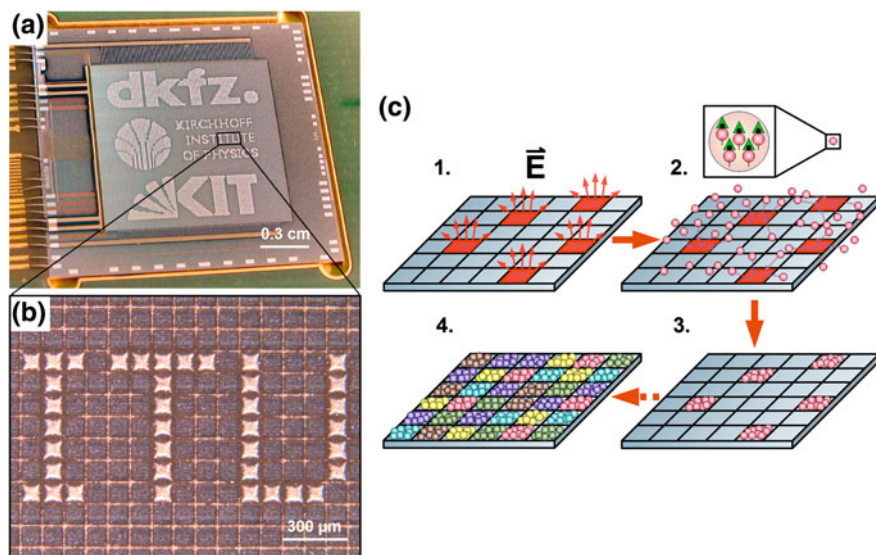
Fig. 1 Microelectronic complementary metal oxide semiconductor chip with deposition pattern. Chip (**a**), a magnified detail (**b**), and schematic deposition process (**c**): The chip is programmed to exhibit a voltage pattern on its surface (*1*). Then, the chip is brought into contact with an aerosol of microparticles and the particles are attracted by the resulting electric field and are deposited according to the desired pattern (*2, 3*). The magnification of a particle (*2*) shows that each particle type incorporates only one kind of monomeric building block (also see Fig. 9). After application of a first particle type, the chip is reprogrammed to a different voltage pattern. Due to the high adhesion forces of the particles, the already deposited pattern remains on the surface without distortion. Next, the particle deposition is repeated with different particle types consecutively (*4*), each type comprising a different monomer. These deposition steps can be repeated with hardly any cross-reactions until the final deposition layer is completed (from [2])

particles stick to the surface and conserve the deposition pattern. The magnification of a single particle (Fig. 1, c2) illustrates that each particle type incorporates only one kind of monomeric building block. After application of a first particle type, the chip is reprogrammed to a different voltage pattern. Due to the high adhesion forces of the particles, the already deposited pattern remains on the surface without distortion. Next, we repeat the particle deposition consecutively without overlap of the different particle types (Fig. 1, c4), each type comprising a different monomer. These deposition steps can be repeated with hardly any wrongly deposited particles until the final deposition layer is completed ([2] supplementary information). Although arbitrary voltage patterns were already achievable with microelectronic chips [29], until recently, we observed that the contamination rate in any other pattern (other than the chessboard pattern consisting of only two different peptides) was still so high that combinatorial synthesis was practically impossible. Regarding this problem, we have theoretically and numerically analyzed particle dynamics [30], leading to a crucial reduction of contamination by using new particles with very small diameters and an especially

designed deposition device. Exploiting those improvements in on-chip particle deposition, we are now able to combinatorially deposit arbitrary particle patterns and different particle types with high precision and flexibility. We showed that the consecutive deposition of 20 patterns with very low contamination is in fact possible ([2] supplementary information).

Higher spot densities of up to 40,000 pixels per cm^2 are feasible [28, 31]; however, for the sake of precision, quality, and robustness, we chose a lower spot density. The present chip is fabricated in a high-voltage CMOS process, which allows for applying either 0 or up to 100 V to each pixel electrode in any desired pattern.

3.2 Aerosol Generation and Deposition

To generate an aerosol for particle deposition, we have designed a computer-controlled aerosol generator (in cooperation with R.O.T GmbH, Celle, Germany) to ensure reproducible aerosol properties (Fig. 2a). It allows us to generate 20 different aerosol types, corresponding to the 20 proteinogenic amino acids. The process of aerosol generation is shown in Fig. 2b: Particles are siphoned out of a particle reservoir using the air ejector principle. Two different air inlets (dosage and transport air) not only provide for the desired siphoning effect, but can also be used to control the aerosol density and velocity.

The particles are ejected through a polytetrafluoroethylene (PTFE) tube, which leads to triboelectric particle charging. Afterwards, they are brought into contact with the chip at an outlet with an especially designed geometry (Fig. 3). A negatively charged sieve filters and selects the particle size and particle charge to ensure a high precision of the deposition pattern [30]. The surface-to-volume ratio of the particles is large due to the rather small size of the particles (about 4 μm in diameter). Therefore, minor surface friction suffices for the charging. Hence, it is unnecessary to implement an additional special setup for further triboelectric particle charging.

We used an aerosol generator consisting of 20 different reservoirs and outlets to deposit the different particle types, each containing a different amino acid monomer, consecutively onto the chip.

A schematic of the outlet geometry is shown in Fig. 3: The PTFE supply tube from the aerosol generator ends in the outlet, which is then expanding from a diameter of 0.8 cm to a diameter of 3 cm, reducing the flow velocity. Afterwards, the aerosol has to pass a sieve with a mesh size of 8–10 μm, thus restraining bigger particles, especially agglomerates, from reaching the chip surface. Finally, the particles come into the proximity of the chip surface for particle deposition.
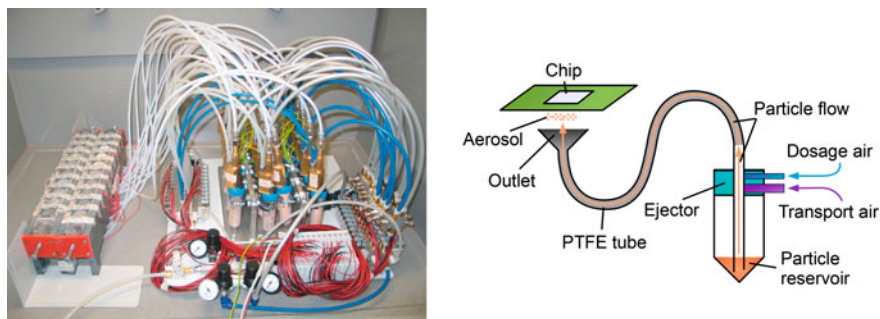
Fig. 2 Aerosol generator for 20 different particle types. Aerosol generator (*left*) and schematic of aerosol generation and deposition principle for one particle type (*right*): particles are siphoned out of a particle reservoir using the air ejector principle, two different air inlets (dosage and transport air) not only provide for the desired siphoning effect, but can also be used to control the aerosol density and velocity (for further details, see text) (from [2])
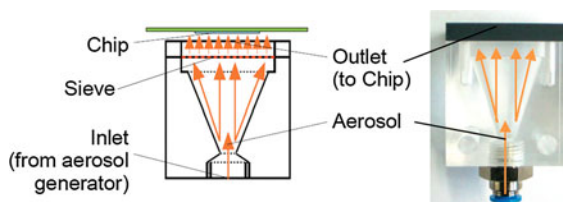
Fig. 3 Geometry of the aerosol outlet and the complementary metal oxide semiconductor chip position in the aerosol stream for particle deposition. Schematic construction drawing (*left*) and original outlet (*right*)

3.3 Chip Printer

In this chapter, we describe a variant procedure of our particle-based synthesis method that should add even more flexibility to the method, and, in addition, reduce costs. We constructed a CMOS "chip printer" that gathers particles on defined pixel areas and, afterwards, "prints" these amino acid particles in high resolution to a support that is suitable for peptide synthesis.

The CMOS chip is used to generate the particle pattern (Fig. 4a, b), as mentioned before. Now, the chip is positioned parallel, about 40 μm apart from a glass slide with an electrode located beneath it (Fig. 4c). An electric potential between the chip and the electrode results in a force that transfers the charged particles from the chip to the glass slide (Fig. 4d). Thus, this technique allows for the transfer of a particle pattern to a glass slide, making the chip reusable.

Afterwards, the glass slide is heated and the particle matrix melts and forms distinct gel-like reaction spheres. Similar to Merrifield's solid-phase peptide synthesis, the particle matrix releases the monomers, which now couple to the
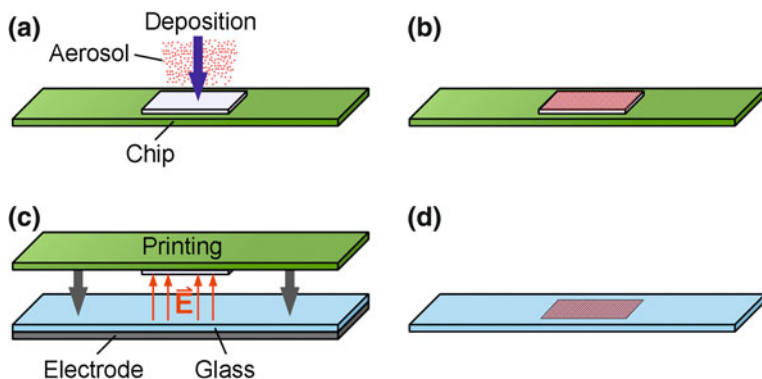
Fig. 4 Complementary metal oxide semiconductor (CMOS) chip printer principle. The scheme shows the particle deposition (**a**, **b**) as in the aforementioned CMOS chip deposition technique. The chip is then positioned parallel to a glass slide with an electrode beneath (**c**). An electric potential between the chip and the electrode results in a force that transfers the particles from the chip to the slide (**d**)

surface-bound amino groups, without losing the spatial resolution of the deposition pattern. Then, the glass slide is allowed to cool to room temperature, which causes the reaction spheres to solidify, freezing the coupling reaction. Finally, the chip is subjected to several washing steps, removing excess monomers and the particle matrix. However, the different monomers remain coupled to the surface. The complete process cycle is repeated layer by layer until the terminal molecule length is reached (see Sect. 3.4 for further information on the coupling and synthesis cycle).

To closely and parallel align the chip surface with the substrate where the synthesis of peptides is conducted after the printing of particles, we mounted the chip on a print head (Fig. 5). It is adjustable in *x, y,* and *z* directions with a precision of ± 0.1 μm and ± 5 μrad.

The chip, mounted on the print head, is moved in *x* direction to pick up particles on selected pixels that were switched to 100 V. Then, the chip is moved to the substrate position and the particles are printed by a strong electrical field (1,000 V/cm) to a glass slide. Because of the required repetitive removal of the synthesis slide for chemical processing, a very high reproducibility of the glass slide position is necessary. Therefore, a special slide holder was devised (Fig. 6). The synthesis surface (top of the slide) is fixed in the *z* direction by stable metal pillars on the top of the slide and a screw from the bottom fixes this position. The position in *x* and *y* direction is defined by three poles, which allow repetitive manual positioning with a tolerance below 5 μm.

The process of parallelization and distance control of the chip to the glass slide in *z* direction is achieved with the following procedure: Before printing, a slide with a conductive surface is fixed in the slide holder. Then, the chip is easily moved into the correct *x* and *y* position and, afterwards, gradually moved in direction of the conductive slide (*z* direction).
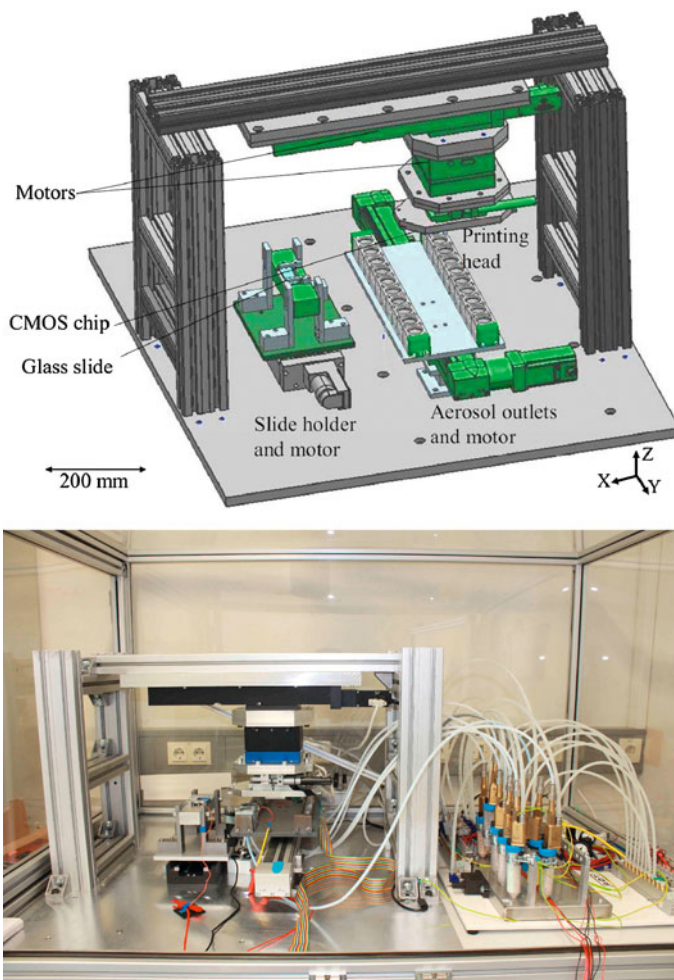
Fig. 5 The complementary metal oxide semiconductor chip printer. The design drawing (*top*) and the assembled experimental system (*bottom*), together with the connected aerosol generator, are shown. The chip is mounted on a print head and can be moved in *x, y,* and *z* directions with a precision of ±0.1 μm and ±5 μrad

In addition to the pixels, the chip surface also exhibits electric contact sensors, which can sense a current on contact of a conductor (Fig. 7). These sensors are now used for fine adjustment and tilting of the printing head.

The printing head is equipped with a tilt stage, which can be tilted in very small angles (±5 μrad). This is used to tilt and move the chip until an electric circuit is established between the conductive slide and the contact sensors. Only when all three contact sensors have established a circuit is the chip successfully parallelized.
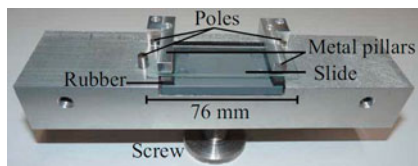
Fig. 6 Synthesis slide holder. The synthesis surface (*top of the slide*) is fixed in the *z* direction by stable metal pillars on the *top* of the slide and a screw from the *bottom* fixes this position. The position in *x* and *y* direction is defined by three poles, which allow repetitive manual positioning with a tolerance below 5 μm
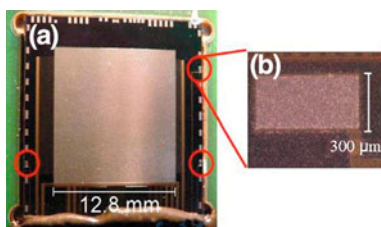
Fig. 7 Complementary metal oxide semiconductor chip surface with highlighted contact sensors. The sensors are used for fine adjustment and tilting of the printing head. Only when all three contact sensors have established a circuit is the chip successfully parallelized

3.4 Particle-Based Synthesis

The particles [27] used in the process are composed of 84.5 % w/w polymer matrix, 10 % w/w of 9-fluorenylmethoxycarbonyl (Fmoc)-protected amino acid monomer preactivated as pentafluorophenyl (OPfp) ester, 4.5 % w/w pyrazolone orange, and 1.0 % w/w of Fe(naphthol)$_2$ complex [32], acting as a charge control agent. The compound is homogenized and afterwards air milled. Finally, 0.05 % w/w silica nanoparticles are added to improve flow properties and to prevent agglomeration.

We have produced 20 different particle types, each containing a different amino acid monomer. The obtained mean charge per mass ratio of the different particle types was about −3 mC/kg [33–35].

The combinatorial synthesis of an array is shown in Fig. 8. First, particles are deposited onto the functionalized substrate exhibiting the reactive amino groups (black dots) (Fig. 8a). Each particle type imbeds a different amino acid type, as illustrated by the differently colored spheres.

When the particle matrix (in gray) is melted, the OPfp preactivated amino acids diffuse and couple to the amino groups on the surface (Fig. 8b). Afterwards, excess particles (gray matrix) and uncoupled monomers are washed away (Fig. 8c), using a mixture of dimethylformamide, acetic anhydride, and diisopropylethylamine
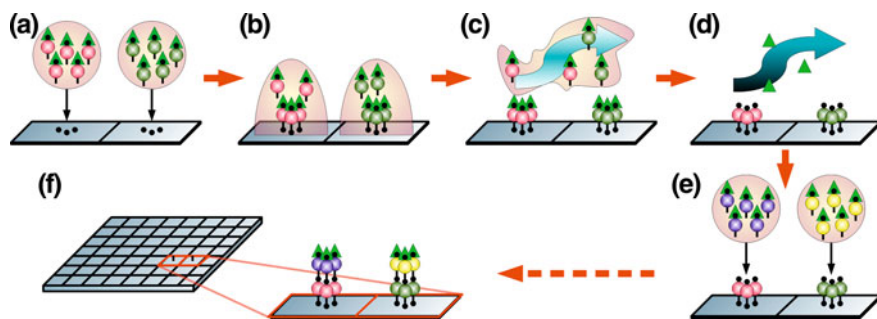
Fig. 8 Combinatorial synthesis of a peptide array. First, the particles are transferred to a substrate (**a**). Each particle type contains a different (amino acid) building block, represented as *colored spheres*. Afterwards, the particles are melted and the monomeric building blocks couple to the surface (**b**). Then, the melted matrix and uncoupled monomers are washed away (**c**) and finally, the protective groups (*green triangles*) are removed (**d**). These steps are repeated with different particle patterns (**e, f**), until the desired array is completed (from [2])

(DMF/Ac$_2$O/DIPEA), which at the same time acetylates (i.e. caps) unreacted amino groups. Finally, the Fmoc-protecting groups (green triangles) are removed from the coupled monomers (Fig. 8d), using a solution of 20 % piperidine in DMF. Hence, new amino groups are presented on the surface for the next coupling cycle. This procedure is repeated with layers of different particle patterns (Fig. 8e, f), until the desired combinatorial array is completed. In the end, side-chain protecting groups are cleaved with trifluoroacetic acid (TFA) (for more details, also see [2] supplementary information).

The glass slide surface is modified with a 40–50 nm thin copolymer film of poly(ethylene glycol) methacrylate (PEGMA) and methyl methacrylate (MMA) with mole fractions of 10 % PEGMA and 90 % MMA [36, 37]. The coating provides for the functional hydroxyl groups, which were subsequently modified with three β-alanine residues according to standard protocols to render the amino groups required for the synthesis [38]. Measurement of the starting density of amino groups upon Fmoc cleavage resulted in about 1 nmol/cm^2.

Using this technique, it is easily possible to generate a variety of biologically or chemically different molecules, given that the building blocks are compatible with the synthesis reaction. The maximum achievable peptide length depends on the amino acid coupling yield, which is about 90 % [27]. Because the signal-to-noise ratio (peptides to truncated peptides) might become increasingly relevant, we require a total peptide yield of at least 10 %. Thus, in our case, the maximum peptide length is about 20 amino acids, with a density of 0.1 nmol/cm^2; the rest of the peptides are truncated.

3.5 Bioparticle Production

The resolution of peptide arrays produced with particle-based procedures such as the CMOS chip or the laser printer is, from a practical point of view, only dependent on mechanical aspects—for example, the size of the used pixel electrodes in the CMOS chip or the adjustment mechanisms in the laser printer. However, if we theoretically neglect these equipment-specific factors, we reach another restricting parameter: the particle size. Hence, the array resolution and precision is also dependent on the particle size and shape.

Previous investigations and comparisons to commercially available toner particles showed that particle types with a median diameter of 10 μm lead to the best results in laser printing. For the CMOS chip technique, a smaller particle diameter is required. Particles of desired size distributions can be produced by grinding and mechanical winnowing. This procedure leads to particle fractions with a tunable median diameter between 1 and 10 μm. Bigger particle fractions are used for the laser printer, whereas finer fractions are applied in the chip approach.

Therefore, to produce high-resolution peptide arrays, it is important to manufacture particle batches with median diameters between 1 and 5 μm. Such particle size distributions could enable the synthesis of arrays with 1 million peptide spots per cm^2.

Experimental particle batches were produced using a milling procedure. The raw particle material was pulverized in a universal mill (M20, IKA, Staufen, Germany) and grinded in an air jet mill (50AS, Hosokawa Alpine, Augsburg, Germany). By fine-tuning the milling parameters, it is possible to achieve the latter particle size distributions (median diameter in the range of 1–5 μm), even without winnowing.

A different approach to create microparticles containing amino acids is the use of a spray drying device (Mini-Spraydryer B-290 in combination with the Inert Loop B-295 for drying of organic solvents; Büchi, Flawil, Switzerland).

The main advantage of this method is the saving of time. Especially for the production of novel experimental particle types containing building blocks, which are yet unverified for synthesis, this method is preferred. Another advantage is the particle shape: In comparison to the randomly shaped particles received by the milling approach (Fig. 9, left), the spray-drying device leads to round, almost uniformly shaped particles (Fig. 9, right).

The crumbly appearance of the particles' surface originates from hydrophobic silica nanoparticles, which were added to act as a dispersive medium in order to prevent agglomeration and improve the flow behavior of the microparticles.

The volume of the organic solvent used to dilute the particle material for the spray-drying process is an important parameter to control the size of the resulting particles. The less concentrated the solution, the smaller the particles will be. When the solution is too concentrated, the particles will become bigger; at a certain point, they even lose their shape. In Fig. 10, two different examples are shown: The chaotic particle shape and wide particle-size distribution caused by
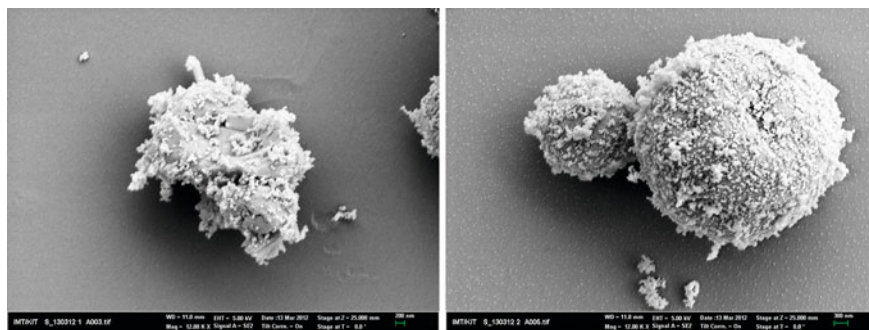
Fig. 9 Comparison of milled and spray-dried particles. The random shape of milled particles (*left*) is compared to the round shape (*right*) of spray-dried particles in a scanning electron microscope
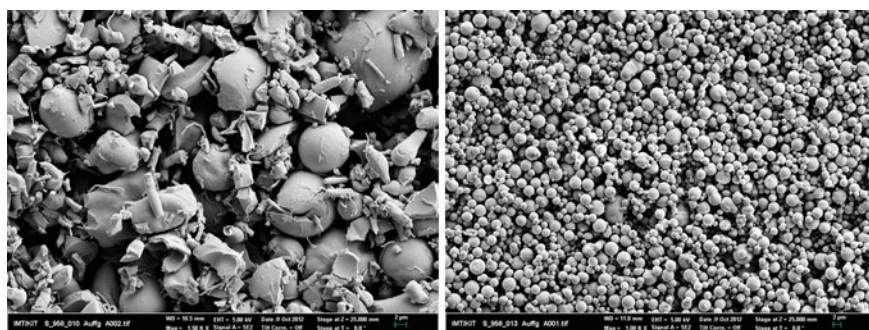
Fig. 10 Different spray-drying parameters. Scanning electron microscope pictures show the chaotic particle shape and wide particle-size distribution caused by spray drying a concentrated solution (*left*). In comparison, note the very small and almost perfectly *round* shaped particles (*right*) created by using high volumes of solvent for the manufacturing process
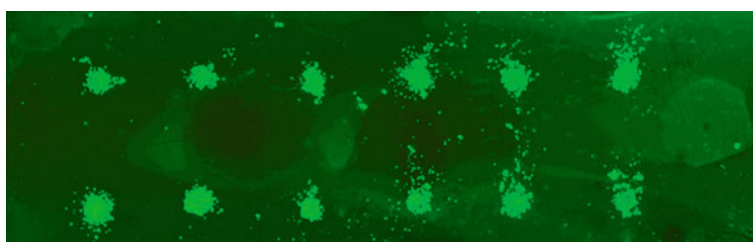
Fig. 11 Proof of Biotin coupling using novel Biotin particles. The coupling was verified by fluorescent staining with Streptavidin Alexa Fluor 546 conjugate
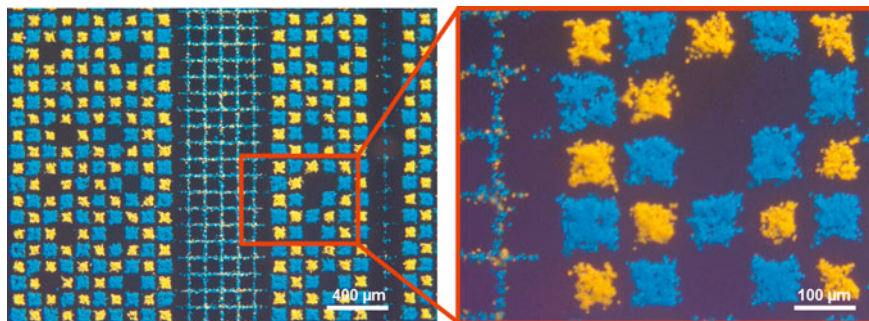
Fig. 12 Proof of printing reproducibility. Xerox (*blue* and *yellow*) proof of reproducibility of the printing process. We first printed a *blue* pattern onto the glass slide, then un- and remounted the glass slide manually. Now, a complementary pattern of *yellow* particles was deposited on the chip and then printed onto the glass slide. The pattern is clearly visible and the *blue* and *yellow* spots do not overlap

spray drying a very concentrated solution of particle material (Fig. 10, left) and, in comparison, the very small and almost perfectly round-shaped particles (Fig. 10, right) formed by using a higher dilution for the spray-drying process.

Using this technique, it is possible to embed activated nonproteinogenic monomer building blocks into the particles to make them available for combinatorial particle-based array synthesis. In this context, we produced an experimental particle batch containing activated Biotin-OPfp ester. In a simple coupling experiment on a PEGMA-grafted glass slide functionalized with beta-Alanine, we could prove the integrity of the incorporated active ester by subsequent fluorescent staining with Streptavidin Alexa Fluor 546 conjugate (Fig. 11).

4 Results

4.1 Proof of Reproducibility and Flexibility

The first particle printing results with the CMOS chip were obtained with two commercial laser printer particle types from Xerox (blue and yellow). We tried to prove the reproducibility of the printing process, even when removing and, afterwards, remounting the glass slide in the slide holder. Therefore, we first printed a blue pattern onto the glass slide, then un- and remounted the glass slide manually. Then, a complementary pattern of yellow particles was deposited on the chip and afterwards printed onto the glass slide at the former position. The result of this experiment is shown in Fig. 12: The pattern is clearly visible and the blue and yellow spots do not overlap. No bias between the two deposited patterns is visible. Thus, we can expect a good reproducibility of deposition for several printing layers.
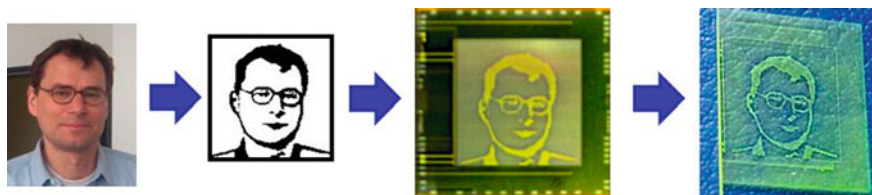
Fig. 13 Rapid and flexible particle printing approach. An image is easily processed into a pattern of 128 × 128 pixels. The chip is used to develop the corresponding voltage pattern into a particle pattern, which is then printed onto a glass slide

In another approach, we wanted to show the rapid production and flexibility of our approach (Fig. 13). Arbitrary images are easily processed into black-and-white (i.e. on and off) digital patterns of 128 × 128 pixels. Then, the CMOS chip can be used to develop the corresponding voltage pattern into a particle pattern, which is afterwards printed onto a slide. The process from particle pattern development until printing only requires about 20 s, whereas the precise movement of the chip from the aerosol outlet to the glass slide requires most of this time.

4.2 Proof of Amino Acid Coupling

In a subsequent experiment, we deposited an amino acid particle pattern on the chip (Fig. 14a) and printed it onto a functionalized glass surface (Fig. 14b). Finally, we melted the particles (Fig. 14c) to induce the coupling reaction to free amino groups on the functionalized glass slide, blocked nonreacted free amino groups with a large excess of acetic anhydride, removed the Fmoc-protecting groups, and stained the surface coupled amino acids with bromophenol blue (Fig. 14d).

The experiment shows that the printing pattern has a sufficient particle density for a successful coupling of amino acids, without causing uncoupled regions.

4.3 Proof-of-Principle Peptide Array Synthesis

After the successful basic printing and coupling experiments, we conducted a proof-of-principle peptide array synthesis. Therefore, we designed two complementary patterns and developed them in a two-step synthesis resulting in two different peptides.

The schematic in Fig. 15 (left) shows the dipeptides, which were synthesized with the CMOS chip printer: a dipeptide consisting of two consecutive glycines and a synthetic peptide consisting of alanine and biotin. Thus, in the first layer, we
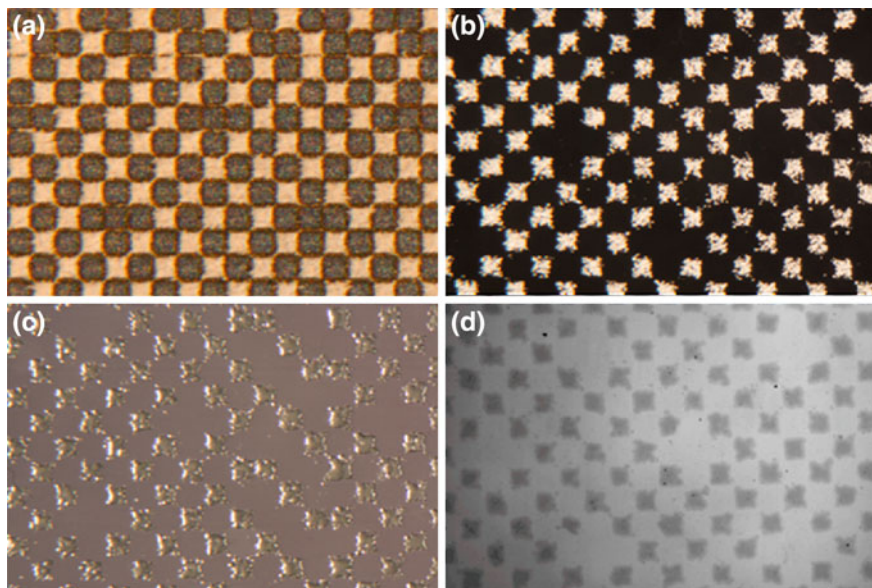
Fig. 14 Proof-of-principle amino acid printing and coupling. Lysine-containing particles were deposited onto the CMOS chip (**a**) and transferred onto a glass slide (**b**). In (**c**), the pattern is shown after the melting/coupling process. The final bromophenol blue staining of the amino groups is shown in (**d**)
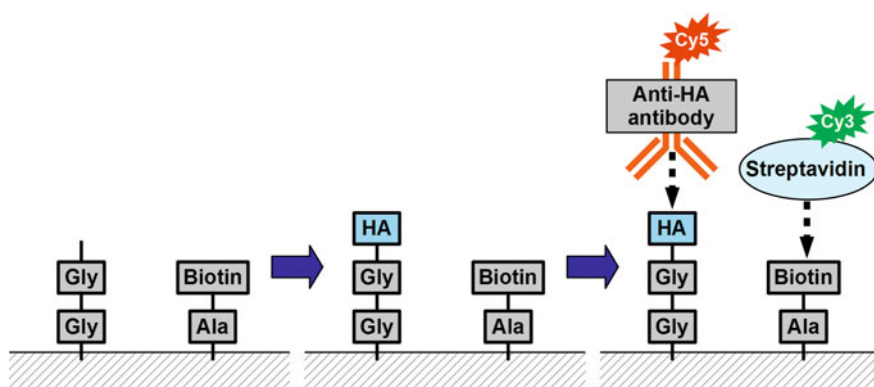
Fig. 15 Peptide synthesis of two different dipeptides and subsequent fluorescent labeling. The complementary patterns are composed of a dipeptide consisting of two consecutive glycines and a synthetic peptide consisting of alanine and biotin (*left*). In the next step (*middle*), a HA-tag peptide was coupled from solution to the free N-terminus of glycine. Finally, the fluorescent staining is performed (*right*) by incubating the array with an anti-HA antibody conjugated with a Cy5 fluorophore and a streptavidin conjugated with a Cy3 fluorophore
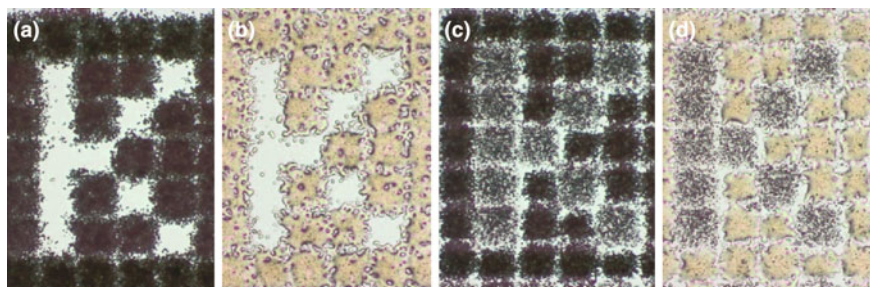
Fig. 16 Detail of the printed complementary pattern of glycine and alanine in the first synthesis layer on the glass slide. The printing of glycine with subsequent melting is shown (**a**, **b**). In a second experiment, the printing of glycine and afterwards alanine with subsequent melting is shown (**c**, **d**)

Fig. 17 Fluorescent staining results of the proof-of-principle peptide array. The green staining corresponds to the alanine-biotin spots, labeled with streptavidin conjugated with a Cy3 fluorophore (*top*). The red staining corresponds to the di-glycine spots, which were elongated with an HA peptide and labeled with anti-HA antibody conjugated with a Cy5 fluorophore (*bottom*)
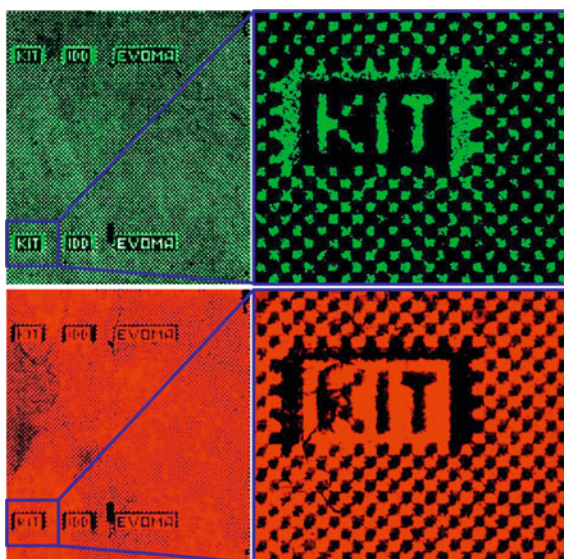
printed and coupled two complementary patterns of glycine and alanine; in the second layer, we used glycine and biotin.

In the next step (Fig. 15, middle), a HA-tag peptide (glycine-glycine-glycine-tyrosine-proline-tyrosine-aspartic acid-valine-proline-aspartic acid-tyrosine-alanine-glycine-glycine-glycine) was coupled from solution to the free N-terminus of glycine; the biotin does not offer a reactive group for this coupling reaction. Finally, the fluorescent staining is performed (Fig. 15, right) by incubating the array with an anti-HA antibody conjugated with Cy5 fluorophore (647 nm) and a streptavidin conjugated with Cy3 fluorophore (546 nm).

In Fig. 16, a detail of the printing pattern of the first synthesis layer is shown: First, the glycine pattern with subsequent melting is shown (Fig. 16a, b). In a second experiment, the printing of glycine and afterwards alanine with subsequent melting was performed (Fig. 16c, d).

The fluorescent staining results of this proof-of-principle experiment are shown in Fig. 17.

The staining of the green pattern (Fig. 17 top) is homogenous with a few missing spots owed to a few incorrectly deposited particles. The staining of the red pattern (Fig. 17 bottom) is rather inhomogenous. We assume that this is a result of the manual staining procedure, which in this case might have led to contaminations and inaccurate washing and, thus, to the inhomogenous staining of the pattern.

5 Summary and Outlook

We have developed a particle-based method for the synthesis of molecular arrays with a CMOS chip printer. The CMOS chip is used to manipulate charged bioparticles, which serve as vehicles to transfer monomers to synthesis spots. The chip is the basic element of our method and fulfills the following xerographic functions: The latent image of electrical fields is generated on its surface. Then, this image is developed by attracting charged particles with powered electrodes on the chip. Afterwards, the chip surface serves as a grounded electrode to transfer particle patterns to the synthesis substrate.

The CMOS chip has 16,384 pixel electrodes on its surface with a spot-to-spot pitch of 100 μm and, thus, exhibits a density of 10,000 pixels per cm^2. It can switch voltage between 0 and 100 V separately for each electrode, which allows for arbitrary particle pattern deposition in combinatorial synthesis. We use more than 20 different very small particle types with diameters less than 5 μm. Hence, a special aerosol generator was developed, which integrates many different channels for the different kinds of particles. In comparison to other particle patterning techniques [39, 40], our method offers a compact approach and a high combinatorial freedom.

In this chapter, we described two methods of bioparticle production: milling and spray drying. Polymer particles with embedded amino acid derivates are produced with these methods and tested in proof-of-principle experiments for application in high-density peptide array synthesis.

The solid phase synthesis is based on the Fmoc-protecting group strategy. After melting the particles, embedded monomers diffuse to the surface and participate in the coupling reaction. Thus, the particle patterns are developed into a molecule array. In addition, we showed the possibility of synthesizing artificial peptides with patterning biotin particles in order to couple biotin derivates in high-density array format.

Further improvements of the CMOS chip printer will comprise the optimization of surface coating and particle quality reproducibility, which are the critical

players in biofunctional high-precision particle deposition. However, future developments should be also devised in the scope of further automation: Robotics used in "wet" chemical coupling steps will improve the robustness of the procedure for routine applications. The complete process will be associated with automatic quality control, as already described for our particle patterning [41]. Successful tackling of all these process steps will finally allow for an automated routine application of this research tool.

The fabrication of molecule libraries in a microarray format with our method allows for massively parallel high-throughput assays with frugal consumption of analyte. Our method offers an unprecedented spot density for in situ combinatorial arrays; the coupling yield is comparable to standard combinatorial chemistry. We integrated the artificial amino acid building block biotin into the synthesis; hence, our method should be easily extendable to D-amino acids, which would make it valuable for synthetic peptide drug discovery. We propose that this method is applicable not only in the medical and biological fields of peptide synthesis, but it also has the potential for a wide application in general combinatorial approaches, such as peptidomimetics, which can be used in chemical screenings for catalysts or for new organic electronics. Today's CMOS technique is mature and advanced, providing cost-efficient microchips and thus making this process feasible for biomolecule array production. By means of further miniaturization, the spot count and density can be easily increased [31], which opens the road to the development of peptide-based functional modules for diverse applications.

References

1. Löffler F (2012) Development of particle-based high-density peptide arrays for application in antibody assays. Ph.D., thesis, University of Heidelberg
2. Loeffler F, Schirwitz C, Wagner J, Koenig K, Maerkle F, Torralba G, Hausmann M, Bischoff FR, Nesterov-Mueller A, Breitling F (2012) Biomolecule arrays using functional combinatorial particle patterning on microchips. Adv Funct Mater 22(12):2503–2508. doi:10.1002/adfm.201103103
3. Cheng Y-C (2012) CMOS-chip based printing system for combinatorial synthesis. Ph.D., thesis, Technical University of Darmstadt
4. Carr PA, Church GM (2009) Genome engineering. Nat Biotechnol 27(12):1151–1162. doi:10.1038/nbt.1590
5. Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet 7(3):200–210. doi:10.1038/nrg1809
6. Tian JD, Quan JY, Saaem I, Tang N, Ma SM, Negre N, Gong H, White KP (2011) Parallel on-chip gene synthesis and application to optimization of protein expression. Nat Biotechnol 29(5):449–452. doi:10.1038/nbt.1847
7. Kosuri S, Eroshenko N, LeProust EM, Super M, Way J, Li JB, Church GM (2010) Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. Nat Biotechnol 28(12):1295–1299. doi:10.1038/nbt.1716
8. Katz C, Levy-Beladev L, Rotem-Bamberger S, Rito T, Rudiger SGD, Friedler A (2011) Studying protein–protein interactions using peptide arrays. Chem Soc Rev 40(5):2131–2145. doi:10.1039/C0cs00029a

9. Seeberger PH, Haase WC (2000) Solid-phase oligosaccharide synthesis and combinatorial carbohydrate libraries. Chem Rev 100(12):4349–4393. doi:10.1021/Cr9903104

10. Merrifield RB (1965) Automated synthesis of peptides. Science 150(3693):178–185

11. Smith GP (1985) Filamentous fusion phage—novel expression vectors that display cloned antigens on the virion surface. Science 228(4705):1315–1317. doi:10.1126/science.4001944

12. Mccafferty J, Griffiths AD, Winter G, Chiswell DJ (1990) Phage antibodies—filamentous phage displaying antibody variable domains. Nature 348(6301):552–554. doi: 10.1038/348552a0

13. Dübel S, Breitling F (1999) Recombinant antibodies, English edn. Wiley, New York

14. Fodor SPA, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D (1991) Light-directed spatially addressable parallel chemical synthesis. Science 251(4995):767–773

15. Pellois JP, Zhou XC, Srivannavit O, Zhou TC, Gulari E, Gao XL (2002) Individually addressable parallel peptide synthesis on microchips. Nat Biotechnol 20(9):922–926. doi: 10.1038/Nbt723

16. Lipshutz RJ, Fodor SPA, Gingeras TR, Lockhart DJ (1999) High density synthetic oligonucleotide arrays. Nat Genet 21:20–24. doi:10.1038/4447

17. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SPA, Gingeras TR (2002) Large-scale transcriptional activity in chromosomes 21 and 22. Science 296(5569):916–919. doi:10.1126/science.1068597

18. Matsuzaki H, Dong SL, Loi H, Di XJ, Liu GY, Hubbell E, Law J, Berntsen T, Chadha M, Hui H, Yang GR, Kennedy GC, Webster TA, Cawley S, Walsh PS, Jones KW, Fodor SPA, Mei R (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. Nat Meth 1(2):109–111. doi:10.1038/Nmeth718

19. Breitling F, Loffler F, Schirwitz C, Cheng YC, Markle F, Konig K, Felgenhauer T, Dorsam E, Bischoff FR, Nesterov-Muller A (2011) Alternative setups for automated peptide synthesis. Mini-Rev Org Chem 8(2):121–131

20. Li SW, Marthandan N, Bowerman D, Garner HR, Kodadek T (2005) Photolithographic synthesis of cyclic peptide arrays using a differential deprotection strategy. Chem Commun 5:581–583. doi:10.1039/B415578e

21. Buus S, Rockberg J, Forsstrom B, Nilsson P, Uhlen M, Schafer-Nielsen C (2012) High-resolution mapping of linear antibody epitopes using ultrahigh-density peptide microarrays. Mol Cell Proteomics 11(12):1790–1800. doi:10.1074/mcp.M112.020800

22. Southern E (2007) Method and apparatus for analysing polynucleotide sequences. EP Patent 0,373,203

23. Ekins RP (1989) Multi-analyte immunoassay. J Pharm Biomed 7(2):155–168. doi: 10.1016/0731-7085(89)80079-2

24. Frank R (1992) Spot-synthesis—an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. Tetrahedron 48(42):9217–9232

25. Nishioka GM (1995) Generation of combinatorial libraries. US Patent 5(449):754

26. www.chem.agilent.com/Scripts/Generic.ASP?lPage=557

27. Stadler V, Felgenhauer T, Beyer M, Fernandez S, Leibe K, Guttler S, Groning M, Konig K, Torralba G, Hausmann M, Lindenstruth V, Nesterov A, Block I, Pipkorn R, Poustka A, Bischoff FR, Breitling F (2008) Combinatorial synthesis of peptide arrays with a laser printer. Angew Chem Int Edit 47(37):7132–7135. doi:10.1002/anie.200801616

28. Beyer M, Nesterov A, Block I, Konig K, Felgenhauer T, Fernandez S, Leibe K, Torralba G, Hausmann M, Trunk U, Lindenstruth V, Bischoff FR, Stadler V, Breitling F (2007) Combinatorial synthesis of peptide arrays onto a microchip. Science 318(5858):1888. doi: 10.1126/science.1149751

29. Koenig K, Block I, Nesterov A, Torralba G, Fernandez S, Felgenhauer T, Leibe K, Schirwitz C, Loeffler F, Painke F, Wagner J, Trunk U, Bischoff FR, Breitling F, Stadler V, Hausmann M, Lindenstruth V (2010) Programmable high voltage CMOS chips for particle-based high-density combinatorial peptide synthesis. Sens Actuator B-Chem 147(2):418–427. doi: 10.1016/j.snb.2009.12.039

30. Loeffler F, Wagner J, Koenig K, Maerkle F, Fernandez S, Schirwitz C, Torralba G, Hausmann M, Lindenstruth V, Bischoff FR, Breitling F, Nesterov A (2011) High-precision combinatorial deposition of micro particle patterns on a microelectronic chip. Aerosol Sci Tech 45(1):65–74. doi:10.1080/02786826.2010.517814
31. Wagner J, Koenig K, Förtsch T, Loeffler F, Fernandez S, Felgenhauer T, Schirwitz C, Painke F, Torralba G, Lindenstruth V, Bischoff FR, Stadler V, Breitling F, Hausmann M, Nesterov-Müller A (2011) Microparticle transfer onto pixel electrodes of 45 μm pitch on HV-CMOS chips—simulation and experiment. Sens Actuator A-Phys 172(2):533–545
32. Kawagishi Y, Ishida Y, Ishikawa K (1983) Metal complexes for use in developers for electrostatic images, charge control function. United States Patent 4404271
33. Nesterov A, Loffler F, Konig K, Trunk U, Leibe K, Felgenhauer T, Stadler V, Bischoff R, Breitling F, Lindenstruth V, Hausmann M (2007) Noncontact charge measurement of moving microparticles contacting dielectric surfaces. Rev Sci Instrum 78(7):075111. doi:10.1063/1.2756629
34. Nesterov A, Loffler F, Konig K, Trunk U, Leibe K, Felgenhauer T, Bischoff FR, Breitling F, Lindenstruth V, Stadler V, Hausmann M (2007) Measurement of triboelectric charging of moving micro particles by means of an inductive cylindrical probe. J Phys D Appl Phys 40(19):6115–6120. doi:10.1088/0022-3727/40/19/053
35. Nesterov A, Loffler F, Cheng YC, Torralba G, Konig K, Hausmann M, Lindenstruth V, Stadler V, Bischoff FR, Breitling F (2010) Characterization of triboelectrically charged particles deposited on dielectric surfaces. J Phys D Appl Phys 43(16):0022–3727. doi:10.1088/0022-3727/43/16/165301
36. Beyer M, Felgenhauer T, Bischoff FR, Breitling F, Stadler V (2006) A novel glass slide-based peptide array support with high functionality resisting non-specific protein adsorption. Biomaterials 27(18):3505–3514. doi:10.1016/j.biomaterials.2006.01.046
37. Stadler V, Kirmse R, Beyer M, Breitling F, Ludwig T, Bischoff FR (2008) PEGMA/MMA copolymer graftings: Generation, protein resistance, and a hydrophobic domain. Langmuir 24(15):8151–8157. doi:10.1021/La800772m
38. Nesterov A, Dörsam E, Cheng Y-C, Schirwitz C, Märkle F, Löffler F, König K, Stadler V, Bischoff R, Breitling F (2010) Peptide arrays with a chip. In: Uttamchandani, Mahesh Yao, Shao Q (ed) Small molecule microarrays: methods and protocols, vol 669. Humana Press, New York, pp 109–124
39. Jacobs HO, Whitesides GM (2001) Submicrometer patterning of charge in thin-film electrets. Science 291(5509):1763–1766
40. Bieri NR, Chung J, Haferl SE, Poulikakos D, Grigoropoulos CP (2003) Microstructuring by printing and laser curing of nanoparticle solutions. Appl Phys Lett 82(20):3529–3531. doi:10.1063/1.1575502
41. Wagner J, Loffler F, Konig K, Fernandez S, Nesterov-Muller A, Breitling F, Bischoff FR, Stadler V, Hausmann M, Lindenstruth V (2010) Quality analysis of selective microparticle deposition on electrically programmable surfaces. Rev Sci Instrum 81(7):073703. doi:10.1063/1.3456986

Protein Engineering as a Tool for the Development of Novel Bioproduction Systems

Uwe T. Bornscheuer

Abstract Protein engineering is the most important tool to overcome the limitations of natural enzymes to facilitate their efficient use in biocatalysis, but also in biotransformations using engineered whole cell systems. In the past decade numerous novel concepts have been developed to create mutant libraries more efficiently, taking advantage of improved molecular biology methods and better expression systems for the target proteins. Most important are sophisticated designs of mutant libraries guided by bioinformatic tools. These achievements were complemented with the development of advanced high-throughput screening or selection methods using, for instance, fluorescence-activated cell sorting (FACS) or microfluidic systems. The latter are prime examples useful to establish novel bioproduction systems as cell-free protein biosynthesis allows protein expression and subsequent sorting of desired enzyme variants with these methods.

Abbreviations

BLAST	Basic local alignment search tool
CAST	Combinatorial active site-test
E	Enantioselectivity
FACS	Fluorescence-activated cell sorting
FADS	Fluorescence-activated droplet sorting
HRP	Horseradish peroxidase
ISM	Iterative saturation mutagenesis
IVC	In vitro compartmentalization
IVTT	In vitro transcription and translation

U. T. Bornscheuer (✉)
Department of Biotechnology and Enzyme Catalysis, Institute of Biochemistry,
Greifswald University, Felix-Hausdorff-Str. 4,
17487 Greifswald, Germany
e-mail: uwe.bornscheuer@uni-greifswald.de

PCR   Polymerase chain reaction
PFE   *Pseudomonas fluorescens* esterase
QM/MM   Quantum mechanical/molecular mechanics
WT   Wildtype

Contents

1 Introduction

For more than a century enzymes have found numerous applications as biocata-
lysts for synthetic organic chemistry due to their high stereoselectivity required for
the creation of chiral compounds. They are also widely used for processes where
mild reaction conditions are important, such as in the conversion of labile com-
pounds, in protection/deprotection steps or where chemo- and regioselectivity are
needed, in laundry applications, or food processing. Also the pretreatment of
renewable resources for the degradation of lignocellulose to make biofuels
recently became important.

Traditionally, only wildtype enzymes for single-step biocatalysis or microor-
ganisms for whole cell biotransformations were used, but the advent of recom-
binant DNA technology enabled the cloning and functional expression of the
enzyme of interest. Nowadays, complete pathways in microorganisms can be
altered and this area is called metabolic engineering and/or synthetic biology.

Natural (wildtype) enzymes have evolved in nature to solve a distinct task in the
host organism, which usually means that they have a specific substrate range and
selectivity and are active under reaction conditions determined by the host envi-
ronment. Substrate concentrations are usually very low (in the µM range) and pH-
and temperature regions are rather narrow. Intracellular enzymes are often highly
specific and intensively regulated by, for example, feedback inhibition whereas
extracellular enzymes typically are more stable and show a broader substrate scope
as their main task is the degradation of nutrients for the host (i.e., lipases split
triglycerides to provide fatty acids, amylases degrade carbohydrate polymers to
provide sugars).

Efficient application of enzymes in biocatalysis and biotransformations on an industrial scale requires substantial improvements of wildtype enzymes to make them robust. Typical goals of protein engineering are therefore improved substrate specificity (narrow or broader substrate scope depending on the application, altered regio- and stereoselectivity, increased reaction rate, elimination of substrate/product inhibition), and better tolerance to reaction conditions (pH, heat, solvents, reactive components, high substrate concentrations in the molar range). In addition, protein engineering can help to improve efficient protein manufacturing by improving expression and folding of active protein, to simplify purification and to achieve better storage stability. Furthermore, protein engineering is an important tool to investigate evolutionary relationships of proteins within a superfamily, to study its reaction mechanism, but also catalytic promiscuity.

Successful protein engineering for industrial applications should target these various properties simultaneously keeping in mind that changing one property (i.e., substrate scope) can also affect other properties (e.g., expression, stability). For instance, in one of our projects, we identified a highly enantioselective esterase variant obtained by directed evolution bearing three amino acid substitutions [1]. Unfortunately, although *E. coli* produced high amounts of the soluble wildtype esterase, only small amounts of the triple mutant were produced, most of it in an unfolded insoluble form. Back mutations identified a double mutant that was both highly enantioselective and efficiently produced in soluble form.

Modern tools developed for protein engineering helped to solve various issues related to the application of enzymes, which has been summarized in our recent review [2] stating that "in the past, an enzyme-based process was designed around the limitations of the enzyme; today, the enzyme is engineered to fit the process specifications."

In this contribution first some basic concepts for protein engineering are briefly given together with some selected examples. Finally, its potential for novel bio-production systems is elucidated. Readers are referred to a broad number of books, book chapters, and review articles for a broader coverage of protein engineering [2–12].

2 Basic Strategies in Protein Engineering

Protein engineering usually follows two major concepts: rational design or directed (molecular) evolution [5, 6, 12] (Fig. 1).

Rational protein design requires the 3D-structure of a protein or at least a good homology model. A high-resolution structure of an enzyme is usually determined by x-ray crystallography and is deposited at the Brookhaven protein database (www.pdb.org). Preferentially, the structure also contains a substrate (or substrate analogue) bound to the active site as this provides detailed information about the residues involved in substrate binding and pockets binding substrate substituents. A model of the protein structure can be generated automatically based on a known
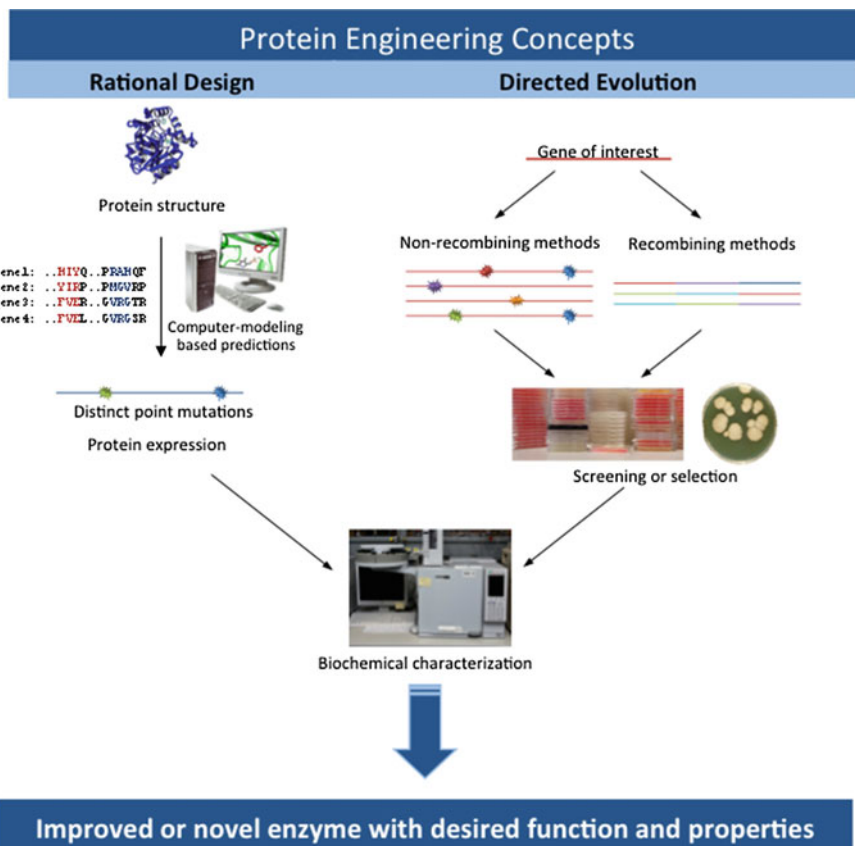
Fig. 1 Overview of protein engineering strategies

homologous x-ray structure (e.g., by Phyre [13] or Robetta [14]). More recently, a method was proposed to predict a protein structure from evolutionary and sequence variations [15]. These structures then serve as a basis to perform molecular modeling for the identification of key residues to be changed in order to improve the protein in the desired direction. Once amino acid residues are identified from a structural analysis these are introduced into the protein on the gene level by site-directed mutagenesis and after production of the mutant it can be analyzed for altered function. The success of this approach largely depends on the information available and the modeling approach used. For instance, if the substrate scope or stereoselectivity of the enzyme should be altered, a simple homology model without detailed knowledge about the binding pockets and active site region will barely enable correct predictions. If, however, a 3D-structure with a bound transition state analogue is available, then modern modeling software should allow a rather reliable prediction, which amino acids should be changed. Software for molecular modeling now runs on standard personal computers and is

rather easy to use (e.g., Yasara; [16]), tools for sequence comparison (e.g., BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi [17]) and alignment (e.g., ClustalW: http://www.ebi.ac.uk/Tools/msa/clustalw2/ [18, 19]) enable quick identification of novel sequences and conserved motifs. The most sophisticated rational protein design uses QM/MM (quantum mechanical/molecular mechanics) simulations, but this requires extended skills and large computational resources.

In contrast, directed evolution only requires a functional expression of the gene to produce the protein of interest. The protein-encoding gene is then subjected to random mutagenesis (e.g., by error-prone PCR) or gene shuffling resulting in very large mutant libraries. These are then either screened for the desired property—usually 10^3–10^4 variants in the microtiter plate format—or subjected to selection methods (e.g., by complementation of pathways in the host organism, but also FACS or microfluidic systems can be used; see Sects. 3.1 and 3.2). Best hits are usually subjected to further rounds of directed evolution until the protein shows the desired properties. Finally, the position and type of amino acid substitutions are determined and the enzyme is biochemically characterized in detail. Often back mutations or combinations are performed to further improve the enzyme or elucidate the specific influence of a given amino acid substitution. Many detailed protocols for directed evolution are available, including several books [20–23] and reviews [3, 9, 24–26].

Both strategies have their advantages and limitations: rational protein design depends on the availability of the protein 3D structure and sufficient knowledge of the contribution of residues to enzyme properties whereas directed evolution strongly depends on suitable mutagenesis methods and high-throughput screening or selection methods to identify the desired variants in a fast and reliable manner. More recently, researchers combined both methods and dubbed this focused directed evolution or semi-rational design. Thus, information available from related protein structures, families and mutants already identified, is combined and then used for targeted randomization of certain areas of the protein [8]; see below.

One challenge in directed evolution experiments is the coverage of a sufficiently large sequence space: the creation and analysis of as many variants as possible. When considering a protein (enzyme) consisting of 200 amino acids, the number of possible variants of a protein by introduction of M substitutions in N amino acids can be calculated from the formula $19^M[N!/(N-M)!M!]$ (from the 20 proteinogenic amino acids, one is already present in the wildtype enzyme; therefore the value 19). Thus, for two random mutations already more than seven million variants are possible; with three or more substitutions, the creation and screening of a library becomes very challenging.

2.1 Methods for the Creation of Mutant Libraries

The still most often used method for the creation of libraries is the error-prone PCR in which conditions are used that lead to the introduction of approximately

1–10 nucleotide mutations per 1,000 base pairs [27]. This is achieved by changing the reaction conditions, that is, to use Mn^{2+} salts instead of Mg^{2+} salts (the polymerase is magnesium-dependent), the use of *Taq* polymerase from *Thermomyces aquaticus*, and variations in the concentrations of the four desoxynucleotides (dNTPs). It should be noted, that due to the biased exchange of nucleotides and the degenerate genetic code, only an average of 6 of all possible 19 amino acid substitutions are accessible by this method. This can be overcome by the use of an optimized polymerase such as Mutazyme from Stratagene. Another approach utilizes mutator strains, for example, the *E. coli* derivative *Epicurian coli* XL1-Red, which lacks DNA repair mechanisms [28]. The introduction of a plasmid bearing the gene encoding the protein of interest leads to mutations during replication. Both methods introduce point mutations, and several iterative mutation rounds followed by identification of best variants are usually required to obtain a biocatalyst with the desired properties.

Alternatively, methods of recombination (also referred to as sexual mutagenesis) can be used. The first example was the DNA (or gene-) shuffling developed by Stemmer, in which DNAse degrades the gene followed by recombination of the fragments using PCR with and without primers [29, 30]. This process mimics natural recombination and has been proven in various examples as a very effective tool to create desired enzymes. Later this method was further refined and termed "DNA family shuffling" or "molecular breeding," enabling the creation of chimeric libraries from a family of genes.

The Arnold laboratory developed two variations of DNA shuffling: The staggered extension process (StEP) is based on a modified PCR protocol using a set of primers and short reaction times for annealing and polymerization. Truncated oligomers dissociate from the template and anneal randomly to different templates leading to recombination. Several repetitions allow the formation of full-length genes [31]. Many other methods are covered in the reviews and books cited above.

In the past few years researchers started to combine rational protein design and random mutagenesis. The major advantage is that if one knows that certain enzyme regions are linked to a desired property without a hint as to which specific amino acid substitution is the best, randomizing only a few amino acids can drastically reduce the screening effort. For instance, randomization of three positions corresponds to $20^3 = 8,000$ combinations. Keeping in mind the degenerate genetic code, about 100,000 variants need to be screened to cover all of the combinations. Although this number is still very high, it is substantially smaller than the sequence space to be covered in standard error-prone PCR libraries. Furthermore, the number of mutants can be further reduced by using the NDT codon (encoding only 12 of the 20 proteinogenic amino acids, but all types of amino acids). Another concept is the combinatorial active-site saturation test (CAST). For this, a small set of amino acids in the vicinity of the active site is chosen and mutated randomly followed by screening for best hits. One single CAST approach can result in an impressive enhancement of enzyme features as demonstrated for an epoxide hydrolase [32, 33]. ISM, iterative saturation mutagenesis, introduces an evolutionary factor into the CASTing strategy. In one

example, the best variants from a first CASTing round served as templates for the next round. Bearing in mind that mutations can have additive or cooperative effects, the order of positions to be mutated and screened might be crucial, leading to different pathways. Following all pathways is work intensive, as each step in a pathway includes a whole library that needs to be screened [34]. Reetz and coworkers have shown that only 8 paths increased the enantioselectivity of an epoxide hydrolase. ISM was also successfully used to improve the thermoactivity (from 45 to 93 °C) [35] and organic solvent stability of a lipase drastically from *Bacillus subtilis* [36]. These and further examples are nicely summarized in recent reviews by Reetz [9, 37, 38].

A bioinformatics-based approach is the ProSAR-strategy (protein sequence activity relationships) developed by Fox et al. [39]. This was used to guide the design of a halohydrin dehalogenase by data obtained from different protein engineering approaches (rational design, random mutagenesis, gene shuffling, saturation mutagenesis), sorting the single mutations into beneficial, neutral, and deleterious ones and keeping only the beneficial mutations for further rounds of mutagenesis. The best halohydrin dehalogenase contained 35 amino acid substitutions leading to a 4,000-fold increased volumetric productivity in the synthesis of a Lipitor (Atorvastatin) side chain [39]. In my group we have used the software 3DM to guide protein engineering of enzymes from the α/β-hydrolase enzyme superfamily [40, 41] and this served as basis to create "small, but smart" libraries. The 3DM allows the design of mutant libraries in a manner so that only amino acids frequently occurring at a given position are covered. This concept enabled us to improve the enantioselectivity ($E_{Mutant} = 80$, $E_{WT} = 3.2$) of an esterase from *Pseudomonas fluorescens* (PFE) by mutating only four residues near the active site. Also activity was improved 240-fold [42]. The same approach was used to improve the thermostability of PFE by 8 °C [43].

2.2 Assay Systems

The major challenge in directed evolution is the identification of desired variants within the mutant libraries. Suitable assay methods should enable a fast, very accurate and targeted identification of desired biocatalysts out of libraries comprising usually 10^3–10^6 mutants. In principle, two different approaches can be applied, namely screening or selection.

Selection-based systems have been used traditionally to enrich certain microorganisms. For in vitro evolution, selection methods are less frequently used as they usually can only be applied to enzymatic reactions that occur in the metabolism of the host strain. On the other hand, selection-based systems allow a considerably higher throughput compared to screening systems (see below). Often, selection is performed as a complementation: an essential metabolite is produced only by a mutated enzyme variant. For instance, a growth assay was used to identify monomeric chorismate mutases. Libraries were screened using media

lacking L-tyrosine and L-phenylalanine [44]. In a similar manner, complementation of biochemical pathways has also been used to identify mutants of an enzyme involved in tryptophan biosynthesis. HisA and TrpF (isomerases involved in the biosynthesis of histidine and tryptophan, respectively) have a similar tertiary structure, and the aminoaldose substrates used are very similar except for a different residue at the amino functionality. Using random mutagenesis and selection by complementation on media lacking tryptophan, several HisA variants that catalyze the TrpF reaction both in vivo and in vitro were identified [45]. One of these variants also retained significant HisA activity.

Stemmer's group subjected four genes of cephalosporinases (enzyme causing antibiotic resistance by hydrolysis of cephalosporins) from *Enterobacter*, *Yersinia*, *Citrobacter*, and *Klebsiella* species to error-prone PCR or DNA-shuffling. Libraries from four generations (a total of 50,000 colonies) were assayed by selection on agar plates with increasing concentrations of Moxalactam (a β-lactam antibiotic). Only those clones could survive that were able to hydrolyze the β-lactam antibiotic. The best variants from error-prone PCR gave only an eightfold increased activity, but the best chimeras from multiple gene-shuffling showed 270–540-fold resistance to Moxalactame [46]. Sequencing of a mutant revealed low homology compared to the parental genes and a total of 33 amino acid substitutions and seven crossovers were found. These changes would have been rather impossible to achieve using error-prone PCR and single-gene shuffling only; thus, these investigations demonstrated the power of DNA-shuffling.

Mutants of an esterase from *Pseudomonas fluorescens* (PFE) produced by directed evolution using the mutator strain *Epicurian coli* XL1-Red were assayed for altered substrate specificity using a selection procedure [28]. The key to the identification of improved variants acting on a sterically hindered 3-hydroxy ester—which was not hydrolyzed by the wild-type esterase—was an agar plate assay system based on pH-indicators, thus leading to a change in color upon hydrolysis of the ethyl ester. Parallel assaying of replica-plated colonies on agar plates supplemented with the glycerol derivative of the 3-hydroxy ester was used to refine the identification, because only *E. coli* colonies producing active esterases had access to the carbon source glycerol, thus leading to enhanced growth and, in turn, larger colonies. By this strategy, a double mutant was identified that efficiently catalyzed hydrolysis.

Screening-based systems (not to be confused with the use of the term "screening" for the identification of microorganisms) are much more frequently used. Due to the very high number of variants generated by directed evolution, common analytical tools such as gas chromatography and HPLC are less useful, as they are usually too time consuming. High-throughput GC–MS or NMR techniques have also been described, but these require the availability of rather expensive equipment and, in the case of screening for enantioselective biocatalysts, the use of deuterated substrates. Phage display, ribosome display, and fluorescense-activated cell sorting (FACS) are the methods of choice to screen very large mutant libraries containing on the order of $>10^6$ variants [47–49], but they are not generally applicable. Examples for the use of FACS are given in Sect. 3.2.

The most frequently used methods are based on photometric and fluorimetric assays performed in microtiter plate (MTP)-based formats in combination with high-throughput robot assistance. These allow a rather accurate screening of several tens of thousands of variants within a reasonable time, and provide information about the enzymes investigated, notably their activity, by determining initial rates or endpoints and their stereoselectivity by using both enantiomers of the compound of interest. One versatile example is the use of umbelliferone derivatives. Esters or amides of umbelliferone are rather unstable, especially at extreme pH and at elevated temperatures, but these ether derivatives are very stable as the fluorophore is linked to the substrate via an ether bond. Only after the enzymatic reaction and treatment with sodium periodate and bovine serum albumin (BSA), is the fluorophore released [50]. A few other techniques are outlined in Sects. 3.1 and 3.2. Three recent extensive reviews cover the use of protein engineering methods to obtain better enzymes for organic synthesis [7–9].

3 Cell-Free Protein Expression

Protein engineering is commonly based on the use of a microorganism (*Escherichia coli* in most cases) to create the mutant library. This, however, can be encountered with disadvantages. Even if the enzyme of interest is functionally expressed without formation of inclusion bodies, the presence of many other enzymes in the host cell can lead to substantial background activities affecting the assay to identify the desired mutants, undesired side reactions can occur, or the amount of the recombinant enzyme is too low to measure the activity over the background. Also it can be laborious to prepare samples for the measurement as steps such as cell disruption, centrifugation, and transfer of cell lysate to an assay plate are required. Furthermore, if the protein of interest is toxic to the host, then this approach will not be applicable.

Consequently, researchers developed alternatives such as display of the mutant library on the cell surface [51–53], which could be successfully established for *E. coli* and *Saccharomyces cerevisiae*. With this surface display, for instance, enantioselective variants of an esterase could be identified [54, 55]. A more widely used method is cell free protein expression. This is based on in vitro transcription and translation (IVTT) of a DNA (or RNA) template using extracts from *E. coli* or wheat germ containing all components necessary for protein biosynthesis [56]. IVTT is nowadays a well-established method: several systems are commercially available, it permits the direct use of PCR templates of mutant libraries, and the protein of interest is produced in reasonable amounts and at high purity. Thus IVTT can help to overcome the issues mentioned above for whole cell expression. Furthermore, a direct linkage between genotype and phenotype can be established facilitating the identification of the encoding gene once the desired enzyme variant has been found in an assay. More details about IVTT can be found elsewhere in this issue (see Cell-Free Systems: Functional Modules for Synthetic and

Chemical Biology). IVTT also turned out to be a very useful concept for protein engineering where in vitro compartmentalization (IVC) or microfluidic systems are used for library production and screening of desired variants.

3.1 In Vitro *Compartmentalization*

In vitro compartmentalization (IVC) allows the physical separation of each enzyme variant in a protein-engineering experiment. In contrast to standard screening methods (see Sect. 2.2) where mutant libraries are transferred to microtiter plates to enable separate measurement of each variant, in IVC either whole cells or single proteins from IVTT productions are separated. Initially IVC was developed using *E. coli* by creating water-in-oil emulsions of tiny droplets containing statistically a single *E. coli* cell, each expressing a different enzyme variant. These droplets can have volumes in the femtoliter range and diameters around 1 μm. A 1 ml *E. coli* culture can contain $> 10^8$ cells and it was shown that IVC allows $>10^{10}$ droplets in 1 ml of emulsion, the equivalent of $>100,000$ microtiter plates. Later, the same principle was transferred to IVTT systems in which the components needed for protein expression, the template gene, and a substrate are encapsulated in the emulsified droplets (Fig. 2) [47, 57]. Combined with the relative ease of preparation, high stability over a broad range of temperature, pH, and salt concentrations, IVC represents a useful tool especially for miniaturization and parallelization of enzymatic reactions. However, IVC also has disadvantages. First, the amount of active protein produced per droplet is very small and hence the detection limit for the enzymatic activity can be too low to be measured properly. Second, the substrate must be entrapped in the droplets (independent whether an *E. coli* cell or IVTT are used for protein expression) when the emulsion is prepared. If the substrate is unstable over a longer period of incubation time, then the activity measurement may give misleading results. Third, a fluorescent product is usually formed to overcome the issues with the detection limit and to enable identification of desired proteins with a fluorescence-activated cell sorter (FACS). Substrate and product differ substantially, however, from the structure of the real target molecule and thus the risk is high that mutants are identified, which show optimal activity or selectivity for the surrogate substrate, but do not show identical properties towards the real substrate of interest in accordance to the first law of directed evolution "You get what you screen for." Finally, the IVC approach only allows endpoint measurements, but barely the determination of enzyme kinetics. This becomes very important, if the wildtype enzyme already shows activity for the substrate used and variants are searched for with improved activity increasing the risk for false-positive results.
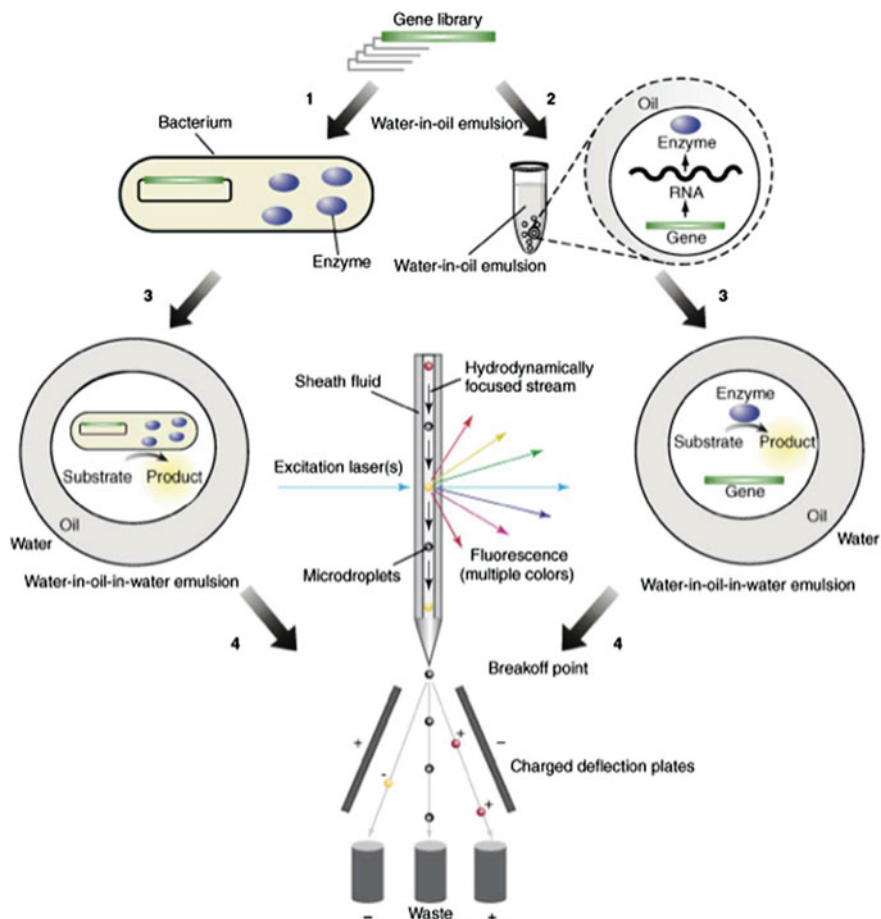
Fig. 2 Principle of compartmentalization for the selection of mutant enzymes. A gene library is either transformed into a bacterium (*1*) or subjected to IVTT in water-in-oil emulsions in the presence of the substrate (*2*). In the next step, water-in-oil-in-water emulsions are prepared (*3*). Enzyme variants produced in each of these compartmentalized systems convert the substrate into a fluorescent product (*4*). Finally, these droplets are then analyzed by FACS. Droplets containing the desired product are sorted out maintaining the physical linkage between phenotype and genotype. The nucleotide and protein sequence of the mutant can then be determined and the protein produced on a larger scale to perform proper biochemical characterization (reproduced from Ref. [57] with permission from Elsevier)

3.2 Microfluidic Systems

Microfluidic systems have recently also become important for protein engineering. This tool allows the creation of microdroplets in a controlled and sophisticated manner. For instance, aqueous microdroplets can be formed at a rate of up to 10,000 per second to form a monodisperse single emulsion or double emulsion.
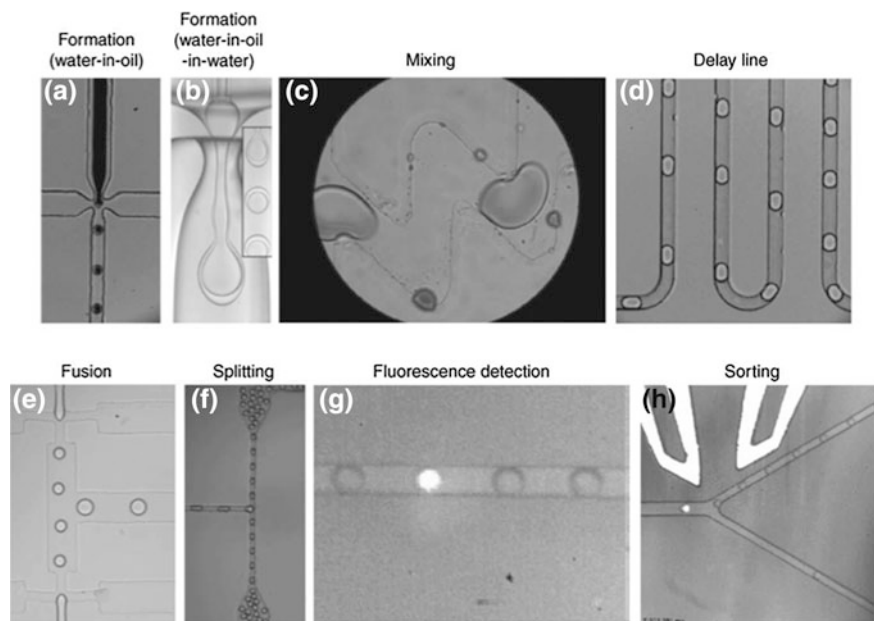
Fig. 3 Formation and manipulation of microdroplets in microfluidic systems (reproduced from Ref. [57] with permission from Elsevier)

Major advantages are that systems can be designed to fuse droplets, allowing the addition of substrate from one droplet to another droplet containing the protein of interest, then subdividing them, incubating them in delay lines to prolong enzymatic reaction times, and they can be sorted, for example, by FADS, dielectrophoresis, or electrocoalesence (Fig. 3). Thus, microfluidic systems can overcome some limitations of IVC while keeping its advantages.

The application of microfluidic systems for protein engineering has been successfully demonstrated for the directed evolution of horseradish peroxidase (HRP) [58]. Mutant libraries created by error-prone PCR were expressed on the surface of the yeast *S. cerevisiae* EBY100, encapsulated, and subjected to the microfluidic device. Analysis of an initial population of $\sim 10^7$ cells by fluorescence measurement allowed sorting the brightest droplets. Biochemical characterization revealed that after a second round of sorting, variants of HRP could be identified as being 10 times faster than the wildtype with activities close to the diffusion limit. The entire screening took less than 10 h using <150 μl total reagent volume. This without doubt represents a dramatic reduction of costs and time compared to standard robotic systems. In another example, a microfluidic workflow was developed in which cytoplasmatic protein expression in *E. coli*, cell lysis, and reaction progress by fluorescent measurement could be established at the picoliter droplet scale using water-in-oil-droplets. This system could be used to identify arylsulfatase variants with significantly improved activity and expression levels [58]. More recently, the

Griffiths group could demonstrate that the microfluidic system can also be combined with IVTT [59]. Here single genes were compartmentalized in aqueous droplets, dispersed in an inert carrier oil, and amplified by PCR. The droplets containing 30,000 copies of each gene were then fused with droplets containing the IVTT ingredients and the substrate for the fluorogenic assay. With this concept, the authors were able to separate mixtures of *lac*Z genes encoding an active β-galactosidase from *lac*Zmut genes encoding an inactive variant at an impressive speed of 2,000 droplets per second.

4 Conclusions

The methods and examples covered in this short contribution should demonstrate that protein engineering has reached a mature level to improve properties of the enzyme of interest in a rather straightforward manner. Researchers now can choose from a variety of concepts for the creation and design of mutant libraries. Furthermore, the recently developed tools for in vitro transcription and translation in combination with in vitro compartmentalization now enable searching within very large libraries ($>10^6$ members) using either FACS or microfluidic systems to find the improved variants. This can lead to a substantial reduction in costs and labor to solve a given protein engineering problem and might become competitive to rather classical systems where high-throughput screening is performed in microtiter plates.

References

1. Schmidt M, Hasenpusch D, Kähler M, Kirchner U, Wiggenhorn K, Langel W, Bornscheuer UT (2006) Directed evolution of an esterase from *Pseudomonas fluorescens* yields a mutant with excellent enantioselectivity and activity for the kinetic resolution of a chiral building block. Chem Bio Chem 7:805
2. Bornscheuer UT, Huisman G, Kazlauskas RJ, Lutz S, Moore J, Robins K (2012) Engineering the third wave in biocatalysis. Nature 485:185
3. Davids T, Schmidt M, Böttcher D, Bornscheuer UT (2013) Strategies for the discovery and engineering of enzymes for biocatalysis. Curr Opin Chem Biol 17:125. http://dx.doi.org/10.1016/j.cbpa.2013.02.022
4. Buchholz K, Kasche V, Bornscheuer UT (2012) Biocatalysts and enzyme technology, 2nd edn. Wiley-VCH, Weinheim
5. Lutz S, Bornscheuer UT (eds) (2009) Protein engineering handbook, vol 1 and 2. Wiley-VCH, Weinheim
6. Lutz S, Bornscheuer UT (eds) (2012) Protein engineering handbook, vol 3. Wiley-VCH, Weinheim
7. Strohmeier GA, Pichler H, May O, Gruber-Khadjawi M (2011) Application of designed enzymes in organic synthesis. Chem Rev 111:4141
8. Behrens GA, Hummel A, Padhi SK, Schätzle S, Bornscheuer UT (2011) Discovery and protein engineering of biocatalysts for organic synthesis. Adv Synth Catal 353:2191

9. Reetz MT (2011) Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. Angew Chem Int Ed 50:138

10. Steiner K, Schwab H (2012) Recent advances in rational approaches for enzyme engineering. Comput Struct Biotechnol J 2:e201209010

11. Wenda S, Illner S, Mell A, Kragl U (2011) Industrial biotechnology—the future of green chemistry? Green Chem 13:3007

12. Kazlauskas RJ, Bornscheuer UT (2009) Finding better protein engineering strategies. Nat Chem Biol 5:526

13. Kelley LA, Sternberg MJ (2009) Protein structure prediction on the web: a case study using the Phyre server. Nat Protoc 4:363

14. Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res 32:W526

15. Marks DS, Hopf TA, Sander C (2012) Protein structure prediction from sequence variation. Nat Biotechnol 30:1072

16. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K (2009) Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins 77(Suppl 9):114

17. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403

18. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38:W695

19. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947

20. Arnold FH, Georgiou G (eds) (2003) Directed enzyme evolution: screening and selection methods. Methods Mol Biol Humana Press, Totawa

21. Arnold FH, Georgiou G (eds) (2003) Directed evolution library creation: methods and protocols., Methods in Molecular BiologyHumana Press, Totawa

22. Brakmann S, Johnsson K (eds) (2002) Directed molecular evolution of proteins. Wiley-VCH, Weinheim

23. Brakmann S, Schwienhorst A (eds) (2004) Evolutionary methods in biotechnology: clever tricks for directed evolution. Wiley-VCH, Weinheim

24. Turner NJ (2009) Directed evolution drives the next generation of biocatalysts. Nat Chem Biol 5:567

25. Neylon C (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. Nucleic Acid Res 32:1448

26. Böttcher D, Bornscheuer UT (2010) Protein engineering of microbial enzymes. Curr Opin Microbiol 13:274

27. Cadwell RC, Joyce GF (1992) Randomization of genes by PCR mutagenesis. PCR Methods Appl 2:28

28. Bornscheuer UT, Altenbuchner J, Meyer HH (1998) Directed evolution of an esterase for the stereoselective resolution of a key intermediate in the synthesis of Epothilones. Biotechnol Bioeng 58:554

29. Stemmer WPC (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Nat Acad Sci U S A 91:10747

30. Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389

31. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. Nat Biotechnol 16:258

32. Reetz MT, Bocola M, Carballeira JD, Zha DX, Vogel A (2005) Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. Angew Chem Int Ed 44:4192

33. Reetz MT, Wang LW, Bocola M (2006) Directed evolution of enantioselective enzymes: iterative cycles of CASTing for probing protein-sequence space. Angew Chem Int Ed 45:1236

34. Reetz MT, Sanchis J (2008) Constructing and analyzing the fitness landscape of an experimental evolutionary process. Chem Bio Chem 9:2260

35. Reetz MT, Carballeira JD, Vogel A (2006) Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angew Chem Int Ed 45:7745

36. Reetz MT, Soni P, Fernandez L, Gumulya Y, Carballeira JD (2010) Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. Chem Commun 46:8657

37. Reetz MT (2013) The Importance of additive and non-additive mutational effects in protein engineering. Angew Chem Int Ed. 52:2658. doi:10.1002/anie.201207842

38. Acevedo-Rocha CG, Kille S, Reetz MT (2013) Iterative saturation mutagenesis: a powerful approach to engineer proteins by simulating Darwinian evolution. Methods Mol Biol (submitted)

39. Fox RJ, Davis SC, Mundorff EC, Newman LM, Gavrilovic V, Ma SK, Chung LM, Ching C, Tam S, Muley S, Grate J, Gruber J, Whitman JC, Sheldon RA, Huisman GW (2007) Improving catalytic function by ProSAR-driven enzyme evolution. Nat Biotechnol 25:338

40. Kuipers RKP, Joosten HJ, van Berkel WJ, Leferink NG, Rooijen E, Ittmann E, van Zimmeren F, Jochens H, Bornscheuer U, Vriend G, dos Santos VA, Schaap PJ (2010) 3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities. Proteins 78:2101

41. Kourist R, Jochens H, Bartsch S, Kuipers R, Padhi SK, Gall M, Bottcher D, Joosten HJ, Bornscheuer UT (2010) The alpha/beta-hydrolase fold 3DM database (ABHDB) as a tool for protein engineering. Chem Bio Chem 11:1635

42. Jochens H, Bornscheuer UT (2010) Natural diversity to guide focused directed evolution. Chem Bio Chem 11:1861

43. Jochens H, Aerts D, Bornscheuer UT (2010) Thermostabilization of an esterase by alignment-guided focussed directed evolution. Prot Eng Des Sel 23:903

44. MacBeath G, Kast P, Hilvert D (1998) Redesigning enzyme topology by directed evolution. Science 279:1958

45. Juergens C, Strom A, Wegener D, Hettwer S, Wilmanns M, Sterner R (2000) Directed evolution of a (beta-alpha) 8-barrel enzyme to catalyze related reactions in two different metabolic pathways. Proc Nat Acad Sci U S A 97:9925

46. Crameri A, Raillard SA, Bermudez E, Stemmer WP (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288

47. Aharoni A, Griffiths AD, Tawfik DS (2005) High-throughput screens and selections of enzyme-encoding genes. Curr Opin Chem Biol 9:210

48. Becker S, Hobenreich H, Vogel A, Knorr J, Wilhelm S, Rosenau F, Jaeger KE, Reetz MT, Kolmar H (2008) Single-cell high-throughput screening to identify enantioselective hydrolytic enzymes. Angew Chem Int Ed 47:5085

49. Fernández-Àlvaro E, Snajdrova R, Jochens H, Davids T, Böttcher D, Bornscheuer UT (2011) A combination of in vivo selection and cell sorting for the identification of enantioselective biocatalysts. Angew Chem Int Ed 50:8584

50. Reymond JL, Wahler D (2002) Substrate arrays as enzyme fingerprinting tools. Chem Bio Chem 3:701

51. Jose J (2006) Autodisplay: efficient bacterial surface display of recombinant proteins. Appl Microbiol Biotechnol 69:607

52. Kondo A, Ueda M (2004) Yeast cell-surface display—applications of molecular display. Appl Microbiol Biotechnol 64:28

53. Becker S, Schmoldt HU, Adams TM, Wilhelm S, Kolmar H (2004) Ultra-high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of noval biocatalysts. Curr Opin Biotechnol 15:323

54. Becker S, Hobenreich H, Vogel A, Knorr J, Wilhelm S, Rosenau F, Jaeger KE, Reetz MT, Kolmar H (2008) Single-cell high-throughput screening to identify enantioselective hydrolytic enzymes. Angew Chem Int Ed 47:5085
55. Becker S, Michalczyk A, Wilhelm S, Jaeger KE, Kolmar H (2007) Ultrahigh-throughput screening to identify *E. coli* cells expressing functionally active enzymes on their surface. Chem Bio Chem 8:943
56. Shimizu Y, Kanamori T, Ueda T (2005) Protein synthesis by pure translation systems. Methods 36:299
57. Griffiths AD, Tawfik DS (2006) Miniaturising the laboratory in emulsion droplets. Trends Biotechnol 24:395
58. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret JC, Marquez M, Klibanov AM, Griffiths AD, Weitz DA (2010) Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. Proc Nat Acad Sci U S A 107:4004
59. Fallah-Araghi A, Baret JC, Ryckelynck M, Griffiths AD (2012) A completely in vitro ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. Lab Chip 12:882

Compartmentalization and Metabolic Channeling for Multienzymatic Biosynthesis: Practical Strategies and Modeling Approaches

U. Jandt, C. You, Y. H.-P. Zhang and A.-P. Zeng

Abstract The construction of efficient enzyme complexes for multienzymatic biosynthesis is of increasing interest in order to achieve maximum yield and to minimize the interference due to shortcomings that are typical for straightforward one-pot multienzyme catalysis. These include product or intermediate feedback inhibition, degeneration, and diffusive losses of reaction intermediates, consumption of co-factors, and others. The main mechanisms in nature to tackle these effects in transient or stable protein associations are the formation of metabolic channeling and microcompartments, processes that are desirable also for multienzymatic biosynthesis in vitro. This chapter provides an overview over two main aspects. First, numerous recent strategies for establishing compartmentalized multienzyme associations and constructed synthetic enzyme complexes are reviewed. Second, the computational methods at hand to investigate and optimize such associations systematically, especially with focus on large multienzyme complexes and metabolic channeling, are discussed. Perspectives on future studies of multienzymatic biosynthesis concerning compartmentalization and metabolic channeling are presented.

Keywords Compartmentalization · Metabolic channeling · Modeling · Multienzymatic synthesis · Synthetic biology

U. Jandt (✉) · A.-P. Zeng
Institute of Bioprocess and Biosystems Engineering, Hamburg University of Technology, Denickestr. 15, 21071 Hamburg, Germany
e-mail: uwe.jandt@tu-harburg.de

A.-P. Zeng
e-mail: aze@tu-harburg.de

C. You · Y. H.-P. Zhang
Virginia Tech, Biological Systems Engineering Department, 210-A Seitz Hall, Blacksburg, VA 24061, USA
e-mail: yoshion@vt.edu

Y. H.-P. Zhang
e-mail: ypzhang@vt.edu

Contents

1 Introduction

Multienzymatic biosynthesis or biotransformation refers to the use of several enzymes and co-enzymes to carry out chemical reactions in a cascade for the production of a desired compound. For in vitro reaction cascades, this mostly leads to typical shortcomings such as product or intermediate feedback inhibition, diffusive losses, consumption of co-factors, and others. On the other hand, such cascade enzyme biosynthesis featuring substrate channeling is ubiquitous inside living cells. Cells face many challenges including rate-limited enzymes, competing metabolic reactions, and labile or toxic intermediates in enzymatic biotransformation (Fig. 1) [174, 175]. For example, some enzymes, such as aldolase (ALD, EC 4.1.2.13) in the glycolysis and gluconeogenesis pathways, suffer from slow turnover, which may result in metabolic flux disequilibrium or rate-limited steps in pathways [35, 36]. Some metabolites can participate in many pathways, leading to their reduced availability for the desired pathway. An example of this is dihydroxy-acetone phosphate (DHAP), an intermediate that is involved in the glycolysis metabolic pathway and the Calvin cycle, and DHAP is also extremely unstable [151].

To deal with these challenges, nature has evolved two main strategies. One is compartmentalization which is the physical separation of biological reactions to bring the cascade enzymes close to each other. Examples of compartmentalization include membrane-bound organelles and bacterial microcompartments [35, 36]. Membrane-bound organelles are common in eukaryotes, such as peroxisomes, which encapsulate reactions that generate or consume the toxic hydrogen peroxide in oxidative reactions [64]. One of the examples of bacterial microcompartments is carboxysome (Fig. 1a), which encapsulates RuBisCO and carbonic anhydrase (CA) to accelerate the reaction rate by providing a high local concentration of carbon dioxide to the enzyme ([21]; Yeates et al. 2008). The membranes of
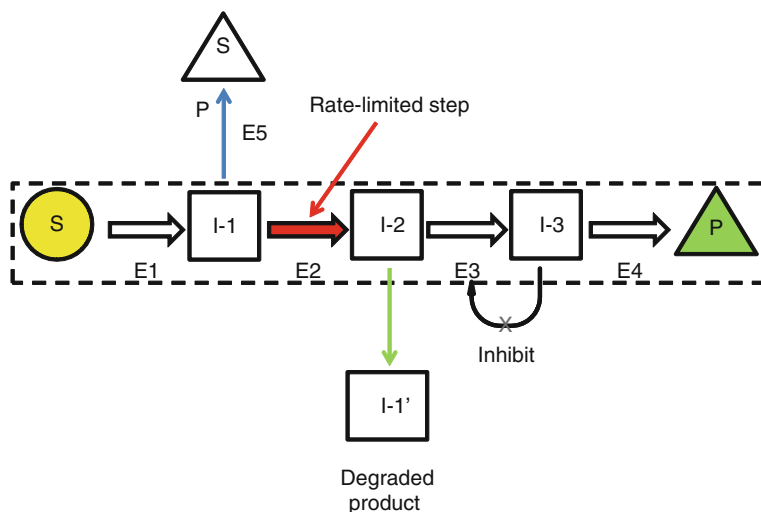
Fig. 1 The scheme of multienzymatic biotransformation and challenges. *S* substrate, I-1, I-2, I-3, I-1′, intermediates, *P* product, *SP* side product, *E* enzyme

organelles and protein shells of bacterial microcompartments can also protect the cell from toxic intermediates [124, 136]. Another way is the formation of metabolic channeling in enzyme complexes, which link cascade enzymes together. Such enzyme complexes found in nature include tryptophan synthase [52], polyketide synthase [159], the family of 2-oxoacid dehydrogenase complexes such as the pyruvate dehydrogenase complex (PDC; [123, 162]) or ketoglutarate dehydrogenase complex (OGDC; [107]), and cellulosomes [13, 14, 48]. A typical example is tryptophan synthase (Fig. 2b), a two-enzyme complex that catalyzes the last two reactions in the biosynthesis of ʟ-tryptophan [52]. The reactive intermediate of insole is channeled from the active site of the α enzyme to the active site of the β enzyme without being released into the surrounding environment.

Enzyme cascades in living cells, cell compartments, and also organisms as seen in nature rely upon a well-orchestrated regulation machinery that features information, energy, and mass transfer at very different time and size scales. These include information dissipation within enzymes, local diffusion of small messenger molecules, and metabolic channeling within temporary or permanently formed enzyme complexes and nanocompartments. The exploitation of many of these evolutionarily developed features for technical application drives the need for a deeper mechanistic understanding of these interactions and thus the development of multiscaled modeling methods [119, 129]. In practice, biologists have devised many nature-inspired synthetic systems to mimic cellular compartmentalization and enzyme complex systems featuring substrate channeling. The key point of these systems is to bring the cascade enzymes spatially close, resulting in an increased reaction rate and reduced diffusion of intermediates to the surrounding

(a) Carboxysome

RuBisCO

Protein shell

CA

(b) Tryptophan synthase

α

β
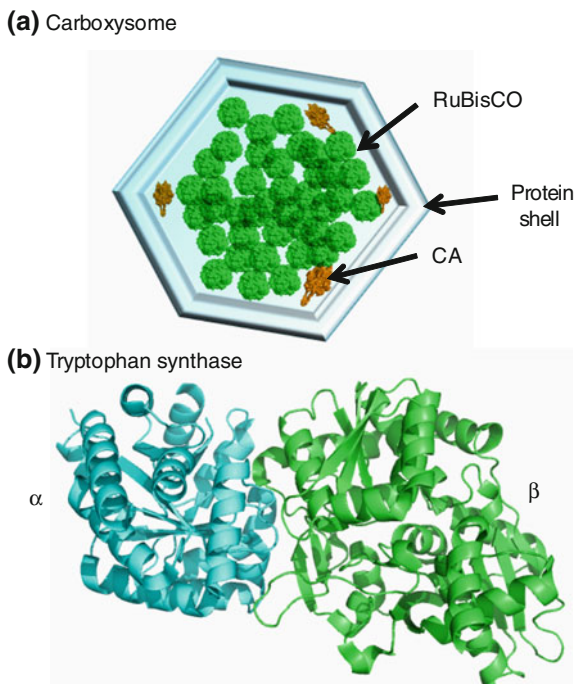
environments. Many of these systems are basis technologies for multienzymatic biotransformation to increase the production of industrially and commercially important chemicals [95, 138, 173].

In this chapter, we describe molecular principles and recently attempted strategies to build multienzymatic compartments and multienzyme complexes, including enzyme encapsulation, fusion proteins, co-immobilization, and protein/nuclear acid scaffolds [4]. For a more technologically oriented review of some of these strategies the reader is referred to the chapter of Ardao and colleagues [4] in this volume. Furthermore, the latest advances of modeling large multienzyme ensembles and complexes, such as directed metabolic channeling, self-assembly, and information transfer, are addressed and discussed. Also we provide perspectives on multienzymatic biosynthesis concerning compartmentalization and synthetic enzyme complexes (metabolons).

2 Practical Strategies

Inspired by natural bacterial microcompartments [84], microcompartments have been heterologously expressed and could be used to encapsulate foreign pathways, including enzymes and substrates (Fig. 2a). Carboxysomes from cyanobacteria were heterologously expressed in *Escherichia coli,* and the encapsulated RuBisCO
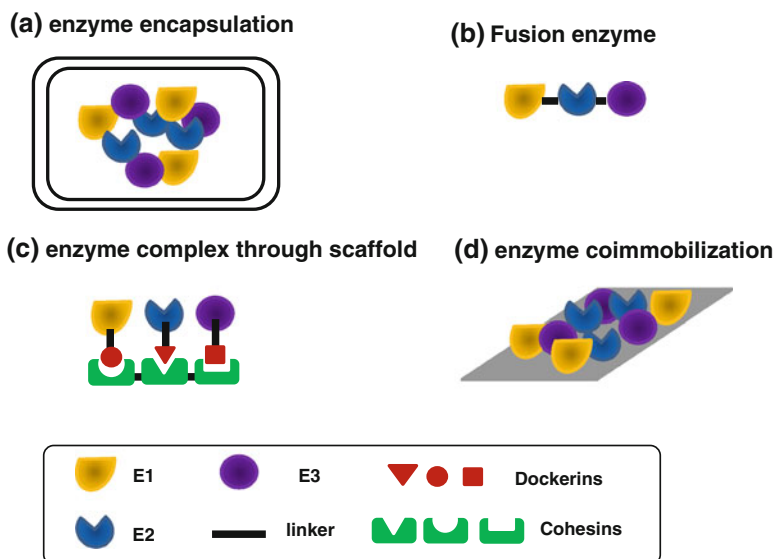
Fig. 3 Strategy for compartmentalization of cascade enzymes: enzyme encapsulation by membrane or protein shell (**a**), a fusion protein of multiple functions (**b**), an enzyme complex through the interaction with the scaffold (**c**), co-immobilization of several enzymes (**d**)

in purified synthetic carboxysomes was still capable of fixing carbon dioxide in vitro [21]. The ethanolamine utilization (Eut) and 1,2-propanediol utilization (Pdu) microcompartment from *Salmonella* have also been expressed in *E. coli* [41, 122]. These heterologously expressed microcompartments have the potential to increase the metabolic flux in enzymatic biosynthesis to the desired product, making this method a useful tool for biotransformation. Another compartmentalization strategy utilizes naturally existing membrane-bound organelles. The advantages of using membrane-bound organelles is that the localization mechanisms are well studied [147] and they allow us to engineer some model microorganisms such as *E. coli* and yeast. For example, a methyl halide transferase was introduced into yeast vacuoles, resulting in the increase of methyl iodide production [16]. The introduction of terpenoid production pathway components to the mitochondria significantly increased product yield [57] (Fig. 3).

Another encapsulation device was recently introduced using a lumazine synthase capsid [168]. Native lumazine synthase, an enzyme catalyzing riboflavin synthesis, does not encapsulate any other enzymes. However, directed evolution was used to engineer electrostatic interactions between this capsid and encapsulated proteins, rendering this capsid the ability to capsulate a toxic enzyme, HIV protease. The ability to insulate toxic compounds from the cell by using a designable compartmentalization system could be very important for living-cell biotransformation.

There are many challenges, such as targeting of more foreign enzymes to microcompartments, and controlling the architecture of artificial microcompartments and stoichiometries of encapsulated enzymes before heterologously expressed microcompartments can be used for multienzymatic biosynthesis. Another problem is the limited knowledge of the mechanism of substrate transport and secretion of final products across the microcompartment shell [90]. A better understanding of these factors and studies in natural compartmentalization systems may help develop better synthetic microcompartments [35, 36].

2.1 Fusion Proteins

The fusion of multiple proteins together is the creation of chimeric proteins, in which two or more cascade enzymes are combined by a linker to form a multifunctional single polypeptide (Fig. 2b) [43, 174]. For example, Bulow et al. constructed a bifunctional enzyme containing cascade reactions mediated by *E. coli* β-galactosidase (LacZ) and galactokinase (GalK) [27]. The resulting fusion protein displayed the enzymatic activity of both gene products. Later, the same group produced another fusion protein containing LacZ and galactose dehydrogenase from *Pseudomonas fluorescens* for the sequential hydrolysis of lactose followed by the oxidation of the galactose, forming the corresponding lactone. This synthetic fusion enzyme displayed kinetic advantages (1.5 to 2.4-fold) over free enzyme mixtures. Such enhancement increased when lactose concentration decreased [97]. This suggests that the proximity conferred by the artificial fusion can enable some amount of synergistic action of the cascade reaction. Such effect has also been revealed in other fusion enzymatic systems [104, 120, 132, 142, 145].

The construction of multifunctional fusion proteins has been proposed to have a potential application in industrial production because it may be beneficial in directing metabolic flux to a preferred pathway [28]. Two genes encoding glycerol 3-phosphate dehydrogenase and glycerol 3-phosphatase were reported to be spontaneously fused into an open reading frame encoding a bifunctional enzyme with a 4-AA linker by removing a small C-terminal fragment of glycerol-3-phosphate dehydrogenase [109]. The high efficiency of this fusion protein enables partial glycerol 3-phosphate channeling between two active sites, resulting in higher glycerol yields. By utilizing a chemostat, the evolved *E. coli* strains that have a spontaneous mutation of this fusion enzyme can produce glycerol from glucose in high yield, concentration, and productivity. The effects of different linker lengths (e.g., 2, 14, 24, 46, and 104 amino acids) of a fusion of hydrogenase and ferredoxin for hydrogen production were tested by Silver and her coworkers [2]. They found that the optimal linker length was 14 amino acids in vivo with an enhanced factor of more than four and no substrate channeling was observed by in vitro tests.

Although building fusion protein is a very simple way to build an enzyme complex [77], the misfolding of large multidomain proteins often happens in hosts [34, 117], resulting in large uncertainty regarding this strategy.

2.2 Building Enzyme Complexes Using Protein Scaffolds

Inspired by extracellular hydrolase enzyme complexes, cellulosomes [15, 49, 146, 170, 171], a synthetic scaffold containing different cohesins from different microorganisms was first used for the assembly of cascade metabolic enzymes (Fig. 2c) [169]. Triosephosphate isomerase (TIM), ALD, and fructose 1,6-bis-phosphatase (FBP) were engineered to have a dockerin at their C-terminals. These three dockerin-containing enzymes can be self-assembled into a static trifunctional enzyme complex through interaction with a miniscaffold protein consisting of three different matching cohesins. The formation of this three-enzyme complex was validated by the result of native PAGE. The particle size of this synthetic enzyme complex was increased approximately fivefold relative to the scaffold only (the particle sizes were estimated by their Zeta potentials with the Malvern Zetasizer Nano ZS system). Such an enzyme complex features substrate channeling because of the spatial proximity of the cascade enzymes [174]. The synthetic metabolon showed more than one order of magnitude enhancements on reaction rates compared to the noncomplexed TIM, ALD, and FBP mixture. This result suggests that using cohesins and dockerins from cellulosomes would be a powerful tool for building an enzyme complex for multienzymatic biotransformation.

Other protein scaffolds from metazoan signaling proteins have been tested for synthetic metabolons for enhancing product titers [51]. The mevalonate pathway suffers from a toxic intermediate and flux imbalance due to unmatched enzyme activities in the pathway. The best-designed synthetic scaffoldin increased product yields 77-fold [51]. The enzymes responsible for the synthesis of the glucaric acid pathway were scaffolded in the same way, resulting in fivefold enhancement in the product titer [114]. Another pathway is the biological hydrogen pathway containing [Fe–Fe]-hydrogenase, ferredoxin, and pyruvate–ferredoxin oxidoreductase (PFOR), which suffers from competing reactions. The synthetic enzymatic pathway containing a protein scaffold, hydrogenase, and PFOR has an approximately fivefold reaction rate enhancement [2].

The degree of enhancement in product yield or reaction rate is believed to be related to enzyme/enzyme ratios, their orientation, and scaffold stoichiometry. However, there is no clear rule to estimate it before testing. Increasing the number of the interaction domain in the scaffold for rate-limited enzymes can dramatically increase the product titers [51]. Via it, optimized stoichiometries and geometries of synthetic protein scaffolds may make multienzyme biotransformation more effective.

Fig. 4 Enzyme complexes are hypothesized to oligomerize into large complexes due to the oligomerization of enzymes
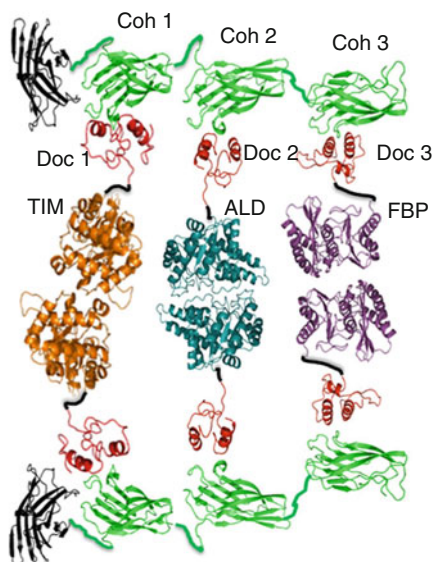
Understanding the structure of enzyme complexes may help elucidate the mechanism pertaining to substrate channeling among enzyme cascades. It has been hypothesized that they may form higher-order complexes due to enzyme oligomerization [94]. For example, in the synthetic TIM, ALD, and FBP enzyme complex, it is known that only dimeric TIM is active [101]; all monomeric, dimeric, and tetrameric ALDs are active [144]; and both dimeric and tetrameric FBPs are active [65]. So this three-enzyme complex may form a dimer with TIM, ALD, and FBP facing one another, as hypothesized previously (Fig. 4) [28].

2.3 Building Enzyme Complexes Using Nucleic Acid Scaffolds

Nucleic acids, both DNA and RNA, can be used as scaffolds to recruit enzymes to form multienzyme complexes. DNA or RNA can be designed to fold into various structures in vitro, forming simple structures such as sheets to more complicated structures such as tubes and capsules [3, 165]. Therefore it is more convenient to synthesize DNA or RNA scaffolds in vitro than protein scaffolds. However, the cost of DNA and RNA is more expensive than that of protein scaffolds, which may impair their application on a large scale.

Several multienzyme systems have been successfully assembled using a DNA scaffold. NADH-flavin mononucleotide (FMN) oxidoreductase and luciferase were assembled onto DNA scaffolds by using streptavidin–biotin linkages, leading to enhanced enzymatic activity relative to the free enzyme mixture [118]. Also,

plasmids can be used as a DNA scaffold, which recruits enzymes via zinc finger domains that bound to specific motifs on the scaffold [44]. Several enzymes were tested using this system, including pathways for the production of resveratrol, 1,2-propanediol, and mevalonate [44]. Yields were found to increase as a function of scaffold architecture, similar to protein scaffolds.

Recently, RNA has drawn huge attention as a scaffold to recruit enzymes for enzyme complexes [1, 46, 71]. The RNA molecules can comprise multiple different aptamer motifs, which can recruit proteins, as well as have complementation regions, which form scaffolds. Such scaffolds are relatively easy to build and characterize, and their expression level can be controlled. So RNA scaffolds are suitable for the expression of specific stoichiometries of enzymes in one RAN chain. For the proof-of-concept experiment, [Fe–Fe]-hydrogenase and ferredoxin, enzymes involved in the biological production of hydrogen, were recruited to RNA scaffolds for hydrogen production [50]. This cascade suffers from side reactions and requires the enzymes to be in close proximity to increase the metabolite flux [2]. Hydrogen production was also found to be related to scaffold architecture.

Regardless of protein or nucleic acid scaffolds, the generality of such scaffolds needs more investigation than only a few proof-of-concept experiments. The mechanism of scaffold action must be further elucidated [35, 36, 94].

2.4 Co-Immobilization of Enzymes

Another approach for facilitating multienzymatic biotransformation is immobilizing cascade enzymes by tethering enzymes on one solid support (Fig. 3d) [20, 55, 160, 172] or direct cross-linking (Mateo et al. 2006; Moehlenbrock et al. 2010). The solid support can be synthetic organic polymers (e.g., Amberlite XAD-7 and Eupergit®C) (Katchalski-Katzir and Kraemer 2000; Kirk and Christensen 2002), inorganic polymers (e.g., silica, zeolites, and mesoporous silicas; Díaz and Balkus 1996; Petri et al. 2005; Zhou and Hartmann 2012), and natural polymers (e.g., cellulose, starch, agarose, and chitosan; Krajewska 2004). Regenerated amorphous cellulose (RAC) made from Avicel (Zhang et al. 2006) is an excellent solid support with a large external surface for enzyme immobilization (Hong et al. 2008; Liao et al. 2012). The entire binding surface of RAC is externally accessible to the large-size enzymes, so the enzymes immobilized on RAC rarely lose their apparent activity (Hong et al. 2008; [169]). An enzyme complex containing TIM, ALD, and FBP immobilized on RAC through a CBM-containing scaffold showed higher activity towards the simple enzyme complex linked by the scaffolding [172]. Multiple enzymes can be precipitated by adding salts, organic solvents, or nonionic polymers followed by cross-linking, resulting in cross-linked enzyme aggregates (CLEAs; Sheldon and van Pelt 2013). A combined CLEA containing an S-selective hydroxynitrile lyase from *Manihot esculenta* and a nonselective nitrilase from *Pseudomonas fluorescens*, catalyzed one-pot conversion of

benzaldehyde to S-mandelic acid (Mateo et al. 2006). This combined CLEA showed higher activity than the mixture of the two separate CLEAs.

Enzyme components during co-immobilization can be randomly distributed [20, 55, 160], positionally assembled [87, 161, 163], and even the active site of an enzyme face that of another enzyme [103]. The enhanced reaction rates among co-immobilization of cascade enzymes have been observed for several systems ([103]; Mateo et al. 2006; [115, 152]), but direct cross-linking could lead to loss of enzyme activity [174].

3 Modeling Approaches

The vastly differing time and size scales of regulation and interaction processes that are necessary for specific and efficient biocatalysis within (and between) living cells prohibited extensive computational simulations in the past (Fig. 5). Generally, strongly reductive approaches were (and are still) widely used to describe only partial processes. With ever-increasing computational power, simulations on a systems level of larger time- and size-scaled regions even up to cells become more realistic in the foreseeable future [158]. Many efforts have been made in systems biology in recent years in order to understand living cells better and to build whole cell models [63, 83].

Despite the fact that computational power will still increase in the next years, pure computation power alone will not be sufficient. Instead, versatile theoretical and mathematical frameworks to describe multilevel interactions are needed and are often yet to be developed [119]. In-depth discussions of various modeling approaches (phenomenological vs. mechanistic, continuous vs. stochastic) describing enzyme interactions [167], followed by detailed considerations of the combination of different scales and methods [129], for example, to capture the hierarchical organization of cells [81] contribute to that development. It has become clear in the last decades that the dynamical self-organization of cells and cell metabolism is fundamentally driven by functional enzymatic associations, catalytic reactions produced during metabolite channeling, microcompartmentalization, and the development of dissipative networks [105, 106]. In this section, the focus is therefore laid on modeling aspects to describe dynamics in multienzyme complexes on a molecular level (Sect. 3.1), in microcompartments (Sect. 3.2), and the determination of fluctuating enzyme–enzyme collaborations from a systems level perspective (Sect. 3.3).

Various spatiotemporally resolved modeling approaches on a larger scale, covering different aspects in mammalian cell compartments and whole cells, have been described elsewhere [47, 79, 150].
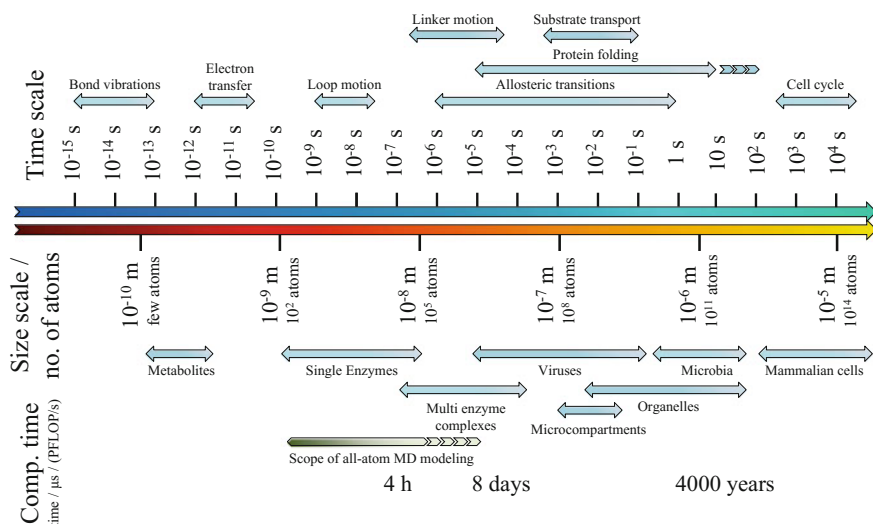
Fig. 5 Typical time and size scales of biologically relevant events and corresponding estimated computation time for coverage of 1 μs using extensive simulation on a molecular level per available petaFLOPS/s (1 petaFLOPS/s $= 10^{15}$ floating point operations per second) computation power (Sect. 3.1.4). Computation performance with respect to system size and parallel computation. Units has been estimated to scale by O($n \log n$)

3.1 Molecular Level

Molecular modeling methods represent a powerful tool to elucidate the regulation and interplay of different molecular regions within enzymes and partly also enzyme complexes. The classical approaches perform all-atom simulations of biomolecules in explicit solvents, usually applying several restraints, simplifications, and cut-offs to accelerate the simulation [108]. The numerically simulated timestep represents the fastest relevant motion (bond vibrations) and usually ranges between 1–4 fs [93]. The most popular methods and program packages are CHARMM [23], AMBER [33], GROMACS [73], and NAMD [128]. Structure data at atom resolution (5 Å) employed in these simulations can be determined using, for example, X-ray crystallography, although it is usually not directly available for (large) enzyme complexes. The availability of atomic resolution structure data has been drastically increasing in recent years. For example, the stable pyruvate dehydrogenase multienzyme complex (PDC) consists of multiple copies of three main enzymes (not considering regulatory kinases and phosphatases attached to the complex) with a total weight of 9 MDa. For the main constituting enzymes, more than 80 different datasets are currently available.

3.1.1 Structure Determination

An important point to be considered in the context of molecular-level modeling is the structure determination of enzyme complexes under physiologic conditions, especially after application of variations such as point mutations. Classical atom-resolution structure determination methods, such as X-ray spectrography [17, 40], rely on the ability to establish a regular crystal structure of a single enzyme. Since many available structure data originate from variable species and also variable tissues [123], the data are usually homogenized using homology modeling methods [86].

The structure extraction of crystallized enzymes is not possible in every case because it can hamper the ternary or quarternary conformation of the enzyme and it is not possible for larger multienzyme complexes. Methods such as small-angle X-ray scattering (SAXS; [18, 69]) and small-angle neutron scattering (SANS; [162]) provide coarse structure data also for large complexes, but with no information about atomic distribution. The gap can be closed with fitting algorithms that systematically optimize orientations and positions of subcomponent atomic structure data to match the low-resolution measured data of the whole complex, obtained, for example, from electron microscopy [56] or SAXS [126]. The dynamics of proteins and protein folding can be observed with NMR spectroscopy, especially in-cell NMR of labeled biomolecules [19].

3.1.2 Molecular Dynamics Simulations

It can be estimated from Fig. 4 that all-atom simulations of higher-order structure changes of enzymes, such as involved in allosteric regulation, protein folding, and also substrate transport via channeling are yet beyond reach. Instead, simplified models, such as steered molecular dynamics simulations, umbrella sampling methods, and implicit solvent models among others, are often employed to reach longer timescales close to the millisecond range [70, 91]. For example, the model-based examination of fast-folding proteins (100 μs–3 ms) showed a set of common principles ruling the folding process of 12 structurally different proteins. Such long simulations, however, increasingly reveal shortcomings of the simplifications applied to the underlying force fields [91], as has been suspected for the apparent over-stabilization of helices during long timescale rearrangements compared to experimental data [62].

For comparably small enzyme systems, molecular modeling methods are already widely used for rational protein design [37, 40], which is also addressed in the chapter by Bornscheuer et al. in this volume. Efforts and successes in design and characterization of synthetic compartmentalization have been reviewed in Chen and Silver [35, 36]. Molecular modeling methods as well as other bioinformatics methods were, for example, applied for medical applications such as antiviral research [88] and the identification of novel anticancer targets [102]. Recent successful attempts include, for example, the introduction of novel

catalytic reactions including the Diels–Alder reaction [148], design of retro-ALDs [80], or the reduction of allosteric product inhibition of aspartokinase for lysine production [38, 39]. The complete dynamic protein-inhibitor binding process for trypsinbenzamidine has been elucidated based on numerous molecular dynamics simulations, revealing three different and partially metastable intermediate binding states between the free and bound states [26]. Generally, the transfer of MD simulation predictions to experimental procedures is challenging, as has been discussed extensively for the example of protein folding [139].

Large enzyme complexes have been less subjected to systematic reconnaissance of structure–function relationships, partly due to their size and corresponding computation power demands and partly due to their complex function involving metabolic channeling, information transfer between subcomponents, self-assembly, and others. Novel inhibitory compounds against the E1 subcomponent of the PDC have been identified with molecular docking methods [130]. It was shown that the dynamic of loops in the active center of the E1 component of PDC is likely to control the decarboxylation step rates of this component, which was revealed by ^{19}F NMR analysis [9, 82]. The influence of structure on the self-assembly of protein complexes has been addressed for pore-forming complexes including the TatA complex, identifying an electrostatic charge zipper mechanism that is also proposed for other membrane-bound proteins [60, 164]. Mostly, symmetry plays a pivotal role in the assembly of large multienzyme structures that are constituted of multiple copies of enzymes (review in [92]). For example, it has been shown earlier that Euclidean geometric considerations predicted the formation of the dodecahedric shape of the PDC E2p component almost exactly [78]. Specific efforts to evaluate the mechanisms of substrate channeling in multienzyme complexes are discussed separately in the following section.

3.1.3 Metabolic Channeling

Not many, although increasing, efforts have been made to elucidate systematically mechanisms of metabolic channeling within enzymes and enzyme complexes by means of structural modeling methods. A straightforward approach to facilitate metabolic channeling is to bring the active sites of corresponding enzymes close to each other, an effect referred to as "channeling by proximity" [12]. In this context, the optimal orientation leading to partial channeling between subsequent enzymes of the glycolytic pathway has been quantitatively simulated [72]. One of the most extensively studied natural channeling mechanisms is in tryptophan synthase [74]. Dunn et al. resolved the chemical background controlling the allosteric interactions responsible for activity switching of the site depending on the stage of the subunit catalytic cycle and also the conformation changes preventing the escape of substrate [52, 53]. These efficiently coordinated cycles obviously rely on directed information transfer for synchronization between the subunits, which may be more systematically elucidated with recently developed directed energy dissipation models [100], especially because long-range interactions seem to play an

important role for allosteric networks in tryptophan synthase [5]. Shorter-range synchronization processes between different active sites, such as proton channels (or "wires") in thiamine enzymes, have also been described [59]. The mechanisms of aldehyde channeling involving two gating residues (His-20 and Tyr-290) in BphI-BphJ, an aldolase–dehydrogenase complex, have been examined [31], as well as the translocation of glutamate or aspartate in a glutamate transporter along an energetically preferred permeation pathway of ≈ 23 Å and two phases [70]. Further molecular dynamics simulation efforts in transporters and membrane channels have been reviewed in Khalili-Araghi et al. [85] and Bahar et al. [8].

3.1.4 Aspects of Computation Power

With growing computational power and availability of multiscale structure data, the structure–function relationships of single parts or aspects of multienzyme complexes can be increasingly elucidated, such as structure stability and self-assembly properties of PDC [162]. However, comprehensive structure-based modeling of enzyme complexes as large as PDC are still beyond the scope of available computation power, instead extensive simplifications (e.g., normal mode analysis, NMA, or a quantized elastic deformational model, QEDM; [8, 24, 68, 110] can be applied to evaluate global conformation variations.

With presently available average-scale computation clusters (2–4 teraFLOPS/s, 1 teraFLOPS/s $= 10^{12}$ floating point operations per second), the computation time for a molecular dynamics calculation of a 30,000-atom system covering 100 ns is roughly 1 day [62, 91]. In theory, an all-atom simulation of a large multienzyme complex (PDC: 9 MDa \approx 10^6 atoms) covering 1 µs assuming an approximate O ($n \log n$) scaling of computation efficiency would take more than 450 days, which is impractical. On the other hand, the available computation power is still increasing more or less exponentially. The classic notion of "Moore's law," the postulated doubling of computation capacity every 18–24 months, seems to have held true, at least until the end of this decade [25]. For comparison: the fastest computer today (Spring 2013) reaches a computation speed of $\approx 17,500$ teraF-LOPS/s ("Titan," Oak Ridge National Laboratory, USA), whereas 1 exaFLOPS/s (1 million teraFLOPS/s) performance computers are already in the pipeline for 2018–2020 approximately [25, 131].

However, the faster nominal computation speeds usually do not translate linearly to decreased simulation times: up-scaling of molecular dynamics calculations to larger, that is, more parallelized, computer systems are far from trivial because of the increased communication overhead between the parallel computation threads. This is mainly caused by long-range and slowly decaying electrostatic interactions that need to be considered in the simulations and prohibit simple subdivision of the calculated systems. More efficient approaches involve fast Ewald mesh/multigrid methods [143] or the reaction field (RF) method that allows improved scaling behavior [O (n) instead of O ($n \log n$)] up to more than 30,000 cores, yielding 30 ns/day for 3–5 million atom systems [140].

Another problem that emerges during massively parallel simulations of very large molecular systems is the limited available memory (few gigabytes) per single computation core, principally increasing the communication overhead drastically. Sophisticated communication management and load balancing methods, however, make the simulation of very large systems (up to 100 million atoms) tractable and efficiently scalable [108].

Extensive comparisons of widely used simulation packages (among others: AMBER, CHARMM, GROMACS) have been performed by Loeffler and Winn [98], with total system sizes (including solvent and lipids) ranging from 2×10^5 to 3×10^6 atoms.

In a few singular cases, biological simulations at scales comparable to the largest multienzyme complexes have already been performed, for example, for the comparably small and simply structured Satellite tobacco mosaic virus (STMV; 50 ns with \approx 10^6 atoms at a computation speed of 0.7–1.1 ns/day) [61].

3.2 Modeling of Microcompartments

In natural systems, another widespread mechanism to minimize diffusive losses, metabolic competition, and the like in densely crowded environments such as the cytosol is the formation of subcompartments of usually only few dozen nanometers in size. With respect to the organization of membranes, this refers to functional rafts of 50 nm diameter with specialized functions [149] whose formations are highly dynamic [96]. Also within compartments, microcompartmentalization of substrates and metabolites occurs and is increasingly considered relevant [134]. The motion of large solutes is hampered by molecular crowding and sieving, whereas the mobility of small solutes is strongly restricted by binding [121]. In some cases, subcompartments and a microheterogeneity can be visualized [11] very well, for example, for local Ca^{2+} domains [29, 133]. Other examples are the regulated ADP [112] and ATP [113] diffusion within isolated brain mitochondria and cardiomyocytes.

A correct mechanistic model understanding of enzymatic reaction and metabolite transfer within or between such microcompartments consequently enforces descriptions beyond bulk ordinary differential equation systems that imply perfectly mixed compartments without diffusive barriers at "mesoscopic" scale (i.e., of some μm size), and a very high (principally infinite) number of target molecules. This has led to fundamentally different views on how to approach subcompartmentalization effects from the modeling side [10, 134]. Stochastic methods are becoming more prevalent. For example, the diffusion restrictions of macromolecules and small signaling molecules in cells have been studied with stochastic simulation [89]. Also, enzyme reaction kinetics can be described more accurately with Monte Carlo methods [66, 67] if few substrate molecules are present. For this purpose, stochastic simulation tools for biological systems have been developed, such as Stochsim [58] or MCell [32].

3.3 Interactome Network

A vast variety of multienzyme combinations has developed in nature. However, many of them form temporary metabolic channels (metabolons) between mutually communicating enzymes or enzyme cascades [99, 153, 154], which are hard to detect and isolate. It seems therefore worthwhile to scan available omics datasets systematically—especially covering the interactome [76, 137]—on a systems level for candidates forming such temporary channels. Recently, large-scale protein–protein interaction (PPI) databases have become available, indicating potential interactions between protein pairs, in some cases weighted by reliability of that interaction [42, 157]. The amount of potential interactions often exceeds 10^5; for the human interactome, it has been estimated to be around 650,000 [155]. Usually, curation based on the literature is performed to reduce the number of false positives; however, the benefit of such curations has been questioned [45].

Generally, the identification of protein complex candidates is assumed to correlate with connectivity in the interactome and can be performed with high-throughput clustering and network analysis methods by means of neighborhood analysis (e.g., ClusterONE for weighted networks; [116], MCL or molecular complex detection, MCODE, based on unweighted networks; [6], and others). Such methods may further be used for global identification of regulatory enzymes at metabolic branching points, involving metabolic channeling [75]. Also, metabolic flux rates and the degree of complexification can be estimated [54]. Furthermore, based on interactome analysis, it has been argued recently that numerous protein–protein channeling interactions do not necessarily rely on physically close location of the enzymes, instead nonenzymatic mediator proteins might be responsible for associating subsequent enzymes in a pathway [125].

The principal problem of such methods is the yield of many false positives (among a smaller amount of false negatives). A comprehensive comparison of frequently used methods of interactome network analysis with respect to robustness and sensitivities is given in Brohee and Van Helden [22].

4 Conclusions and Perspectives

In this contribution, practical and modeling aspects of understanding and exploiting the principles of compartmentalization and metabolic channeling have been discussed. It shows that many methods and approaches exist to facilitate artificial multienzymatic biosynthesis, however, the details are often not systematically understood and metabolic channeling is often simply triggered by proximity of two active sites. Because naturally evolved systems, such as tryptophane synthase or the family of stable and big 2-oxoacid dehydrogenase complexes, exhibit much more sophisticated channeling and regulation behavior, it is worthwhile to dig deeper to reveal their secrets.

It is expected that in the coming years, breakthroughs will presumably be made in the following areas: (1) development of novel and powerful algorithms for systematic identification of multienzyme complexes by integrating different networks such as metabolic and interactome networks; (2) multiscale modeling approaches from the atomic to compartment level, increasingly covering biologically relevant timescales; (3) computational design of new functional biomolecules with desired metabolic channeling; (4) development of novel nanomaterials that are capable of encapsulating multienzymes through co-immobilization; (5) discovery of new scaffolds or development of scaffold-free technologies for the construction of multienzyme complexes.

References

1. Afonin KA, Bindewald E, Yaghoubian AJ, Voss N, Jacovetty E, Shapiro BA, Jaeger L (2010) In vitro assembly of cubic RNA-based scaffolds designed in silico. Nat Nano 5(9):676–682
2. Agapakis CM, Ducat DC, Boyle PM, Wintermute EH, Way JC, Silver PA (2010) Insulation of a synthetic hydrogen metabolism circuit in bacteria. J Biol Eng 4:3
3. Aldaye FA, Palmer AL, Sleiman HF (2008) Assembling materials with DNA as the guide. Science 321(5897):1795–1799
4. Ardao I, Hwang E-T, Zeng A-P (2013) Multi-enzymatic bioreaction systems at micro/nano scale. Adv Biochem Eng/Biotechnol (Accepted)
5. Axe JM, Boehr DD (2013) Long-range interactions in the alpha subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling. J Mol Biol 425(9):1527–1545
6. Bader G, Hogue C (2003) An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4(1):2
7. Badger MR, Price GD (2003) CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot 54(383):609–622
8. Bahar I, Lezon TR, Bakan A, Shrivastava IH (2010) Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. Chem Rev 110(3):1463–1497
9. Balakrishnan A, Nemeria NS, Chakraborty S, Kakalis L, Jordan F (2012) Determination of pre-steady-state rate constants on the escherichia coli pyruvate dehydrogenase complex reveals that loop movement controls the rate-limiting step. J Am Chem Soc 134(45): 18644–18655
10. Barros L, Martínez C (2007) An enquiry into metabolite domains. Biophys J 92(11):3878
11. Barros LF (2010) Towards single-cell real-time imaging of energy metabolism in the brain. Front Neuroenergetics 2:4
12. Bauler P, Huber G, Leyh T, McCammon J (2010) Channeling by proximity: the catalytic advantages of active site colocalization using Brownian dynamics. J Phys Chem 1(9):1332–1335
13. Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521–554
14. Bayer EA, Morag E, Lamed R (1994) The cellulosome—a treasure-trove for biotechnology. Trends Biotechnol 12(9):379–386
15. Bayer EA, Shimon LJ, Shoham Y, Lamed R (1998) Cellulosomes-structure and ultrastructure. J Struct Biol 124(2–3):221–234

16. Bayer TS, Widmaier DM, Temme K, Mirsky EA, Santi DV, Voigt CA (2009) Synthesis of methyl halides from biomass using engineered microbes. J Am Chem Soc 131(18): 6508–6515

17. Berger I, Blanco A, Boelens R, Cavarelli J, Coll M, Folkers G, Nie Y, Pogenberg V, Schultz P, Wilmanns M et al (2011) Structural insights into transcription complexes. J Struct Biol 175(2):135–146

18. Bernadó P, Svergun D (2011) Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Mol BioSyst 8(1):151–167

19. Bertini I, Felli IC, Gonnelli L, Kumar MVV, Pierattelli R (2011) 13 C direct-detection biomolecular NMR spectroscopy in living cells. Angew Chem Int Ed Engl 50(10): 2339–2341

20. Betancor L, Berne C, Luckarift HR, Spain JC (2006) Coimmobilization of a redox enzyme and a cofactor regeneration system. Chem Comm 0(34):3640–3642

21. Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF (2012) Modularity of a carbon-fixing protein organelle. Proc Natl Acad Sci U S A 109(2):478–483

22. Brohee S, Van Helden J (2006) Evaluation of clustering algorithms for protein–protein interaction networks. BMC Bioinformatics 7:488

23. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4(2):187–217

24. Brooks G (2009) Cell-cell and intracellular lactate shuttles. J Physiol 587(23):5591–5600

25. Brown M (2010) DARPA plans exaflop supercomputer for 2018. Wired

26. Buch I, Giorgino T, De Fabritiis G (2011) Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. Proc Natl Acad Sci U S A 108(25):10184–10189

27. Bulow L, Ljungcrantz P, Mosbach K (1985) Preparation of a soluble bifunctional enzyme by gene fusion. Nat Biotech 3(9):821–823

28. Bülow L, Mosbach K (1991) Multienzyme systems obtained by gene fusion. Trends Biotechnol 9(1):226–231

29. Cancela JM, Van Coppenolle F, Galione A, Tepikin AV, Petersen OH (2002) Transformation of local Ca2+ spikes to global Ca2+ transients: the combinatorial roles of multiple Ca2+ releasing messengers. EMBO J 21(5):909–919

30. Cannon GC, Shively JM (1983) Characterization of a homogenous preparation of carboxysomes from Thiobacillus neapolitanus. Arch Microbiol 134(1):52–59

31. Carere J, Baker P, Seah SY (2011) Investigating the molecular determinants for substrate channeling in BphI-BphJ, an aldolase-dehydrogenase complex from the polychlorinated biphenyls degradation pathway. Biochemistry 50(39):8407–8416

32. Casanova H, Berman F, Bartol T, Gokcay E, Sejnowski T, Birnbaum A, Dongarra J, Miller M, Ellisman M, Faerman M, Obertelli G, Wolski R, Pomerantz S, Stiles J (2004) The virtual instrument: support for grid-enabled MCell simulations. Int J High Perf Comput Appl 18:3–17

33. Case DA, Cheatham TE III, Darden TA, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling CL, Wang B, Woods RJ (2005) The amber biomolecular simulation programs. J Comput Chem 26(16):1668–1688

34. Chang H-C, Kaiser CM, Hartl FU, Barral JM (2005) De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. J Mol Biol 353(2):397–409

35. Chen AH, Silver PA (2012) Designing biological compartmentalization. Trends Cell Biol 22(12):662–670

36. Chen AH, Silver PA (2012) Designing biological compartmentalization. Trends Cell Biol 22(12):662–670

37. Chen Z, Jandt U, Rappert S, Zeng A-P (2013) Proteindesign fuer die Entwicklung von industriellen Mikroorganismen. BIOspektrum 19(1):99–101

38. Chen Z, Meyer W, Rappert S, Sun J, Zeng A (2011) Coevolutionary analysis enabled rational deregulation of allosteric enzyme inhibition in corynebacterium glutamicum for lysine production. Appl Environ Microbiol 77(13):4352

39. Chen Z, Rappert S, Sun J, Zeng A (2011) Integrating molecular dynamics and co-evolutionary analysis for reliable target prediction and deregulation of the allosteric inhibition of aspartokinase for amino acid production. J Biotechnol 154(4):248–254

40. Chen Z, Wilmanns M, Zeng A (2010) Structural synthetic biotechnology: from molecular structure to predictable design for industrial strain development. Trends Biotechnol 28(10):534–542

41. Choudhary S, Quin MB, Sanders MA, Johnson ET, Schmidt-Dannert C (2012) Engineered protein nano-compartments for targeted enzyme localization. PLoS ONE 7(3):e33342

42. Collins SR, Kemmeren P, Zhao X-C, Greenblatt JF, Spencer F, Holstege FC, Weissman JS, Krogan NJ (2007) Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics 6(3):439–450

43. Conrado RJ, Varner JD, DeLisa MP (2008) Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. Curr Opin Biotechnol 19(5):492–499

44. Conrado RJ, Wu GC, Boock JT, Xu H, Chen SY, Lebar T, Turnšek J, Tomšič N, Avbelj M, Gaber R, Koprivnjak T, Mori J, Glavnik V, Vovk I, Benčina M, Hodnik V, Anderluh G, Dueber JE, Jerala R, DeLisa MP (2012) DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. Nucleic Acids Res 40(4):1879–1889

45. Cusick ME, Yu H, Smolyar A, Venkatesan K, Carvunis A-R, Simonis N, Rual J-F, Borick H, Braun P, Dreze M et al (2008) Literature-curated protein interaction datasets. Nature 6(1):39–46

46. Delebecque CJ, Lindner AB, Silver PA, Aldaye FA (2011) Organization of intracellular reactions with rationally designed RNA assemblies. Science 333(6041):470–474

47. Dinh A-T, Pangarkar C, Theofanous T, Mitragotri S (2007) Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. Biophys J 92(3):831–846

48. Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat Rev Microbiol 2(7):541–551

49. Doi RH, Kosugi A, Murashima K, Tamaru Y, Han SO (2003) Cellulosomes from mesophilic bacteria. J Bacteriol 185(20):5907–5914

50. Ducat DC, Sachdeva G, Silver PA (2011) Rewiring hydrogenase-dependent redox circuits in cyanobacteria. Proc Natl Acad Sci U S A 108(10):3941–3946

51. Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KLJ, Keasling JD (2009) Synthetic protein scaffolds provide modular control over metabolic flux. Nat Biotech 27(8):753–759

52. Dunn MF (2012) Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. Arch Biochem Biophys 519(2):154–166

53. Dunn M, Niks D, Ngo H, Barends T, Schlichting I (2008) Tryptophan synthase: the workings of a channeling nanomachine. Trends Biochem Sci 33(6):254–264

54. Durek P, Walther D (2008) The integrated analysis of metabolic and protein interaction networks reveals novel molecular organizing principles. BMC Syst Biol 2:100

55. El-Zahab B, Jia H, Wang P (2004) Enabling multienzyme biocatalysis using nanoporous materials. Biotechnol Bioeng 87(2):178–183

56. Fabiola F, Chapman MS et al (2005) Fitting of high-resolution structures into electron microscopy reconstruction images. Structure 13(3):389–400

57. Farhi M, Marhevka E, Masci T, Marcos E, Eyal Y, Ovadis M, Abeliovich H, Vainstein A (2011) Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab Eng 13(5):474–481

58. Firth C, Bray D (2004) Stochastic simulation of cell signaling pathways. In Bower JM, Bolouri H (eds) Computational modeling of genetic and biochemical networks, pp 263–286. MIT Press, Cambridge

59. Frank RA, Titman CM, Pratap JV, Luisi BF, Perham RN (2004) A molecular switch and proton wire synchronize the active sites in thiamine enzymes. Science 306(5697):872–876

60. Frank R, Pratap J, Pei X, Perham R, Luisi B (2005) The molecular origins of specificity in the assembly of a multienzyme complex. Structure 13(8):1119–1130

61. Freddolino P, Arkhipov A, Larson S, McPherson A, Schulten K (2006) Molecular dynamics simulations of the complete satellite tobacco mosaic virus. Structure 14(3):437–449

62. Freddolino P, Liu F, Gruebele M, Schulten K (2008) Tenmicrosecond molecular dynamics simulation of a fast-folding WW domain. Biophys J 94(10):L75

63. Freddolino PL, Tavazoie S (2012) The Dawn of virtual cell biology. Cell 150(2):248–250

64. Gehrmann W, Elsner M (2011) A specific fluorescence probe for hydrogen peroxide detection in peroxisomes. Free Radic Res 45(5):501–506

65. Gerbling K-P, Steup M, Latzko E (1985) Fructose-1,6-bisphosphatase from Synechococcus leopoliensis. Eur J Biochem 147(1):207–215

66. Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem 104(9):1876–1889

67. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361

68. Gorba C, Miyashita O, Tama F (2008) Normal-mode flexible fitting of high-resolution structure of biological molecules toward one-dimensional low-resolution data. Biophys J 94(5):1589–1599

69. Gorynia S, Bandeiras T, Pinho F, McVey C, Vonrhein C, Round A, Svergun D, Donner P, Matias P, Carrondo M et al (2011) Structural and functional insights into a dodecameric molecular machine - the RuvBL1/RuvBL2 complex. J Struct Biol 176(3):279

70. Gu Y, Shrivastava IH, Amara SG, Bahar I (2009) Molecular simulations elucidate the substrate translocation pathway in a glutamate transporter. Proc Natl Acad Sci U S A 106(8):2589–2594

71. Guo P (2010) The emerging field of RNA nanotechnology. Nat Nano 5(12):833–842

72. Hakobyan, D. and Nazaryan, K. 2010. Molecular dynamics study of interaction and substrate channeling between neuron-specific enolase and b-type phosphoglycerate mutase. Proteins Struct Funct Bioinformatics 78(7):1691–1704

73. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447

74. Huang X, Holden H, Raushel F (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions. Annu Rev Biochem 70(1):149–180

75. Huthmacher C, Gille C, Holzhütter H (2008) A computational u analysis of protein interactions in metabolic networks reveals novel enzyme pairs potentially involved in metabolic channeling. J Theor Biol 252(3):456

76. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. Natl Acad Sci 98(8):4569–4574

77. Iturrate L, Sanchez-Moreno I, Doyaguez EG, Garcia-Junceda E (2009) Substrate channelling in an engineered bifunctional aldolase/kinase enzyme confers catalytic advantage for C–C bond formation. Chem Commun 2009(13):1721–1723

78. Izard T, Aevarsson A, Allen MD, Westphal AH, Perham RN, de Kok A, Hol WGJ (1999) Principles of quasi-equivalence and Euclidean geometry govern the assembly of cubic and dodecahedral cores of pyruvate dehydrogenase complexes. Proc Natl Acad Sci U S A 96(4):1240–1245

79. Jandt U, Zeng A-P (2012) Modeling of intracellular transport and compartmentation. In: Hu W-S, Zeng A-P (eds) Advances in biochemical engineering and biotechnology: genomics and systems biology of mammalian cell culture, pp 221–249

80. Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF et al (2008) De novo computational design of retro-aldol enzymes. Science 319(5868):1387–1391

81. Joshi H, Singharoy A, Sereda Y, Cheluvaraja S, Ortoleva P (2011) Multiscale simulation of microbe structure and dynamics. Prog Biophys Mol Biol 107(1):200–217
82. Kale S, Ulas G, Song J, Brudvig GW, Furey W, Jordan F (2008) Efficient coupling of catalysis and dynamics in the e1 component of escherichia coli pyruvate dehydrogenase multienzyme complex. Proc Natl Acad Sci U S A 105(4):1158–1163
83. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, Assad-Garcia N, Glass JI, Covert MW (2012) A whole-cell computational model predicts phenotype from genotype. Cell 150(2):389–401
84. Kerfeld CA, Heinhorst S, Cannon GC (2010) Bacterial microcompartments. Annu Rev Microbiol 64:391–408
85. Khalili-Araghi F, Gumbart J, Wen P-C, Sotomayor M, Tajkhorshid E, Schulten K (2009) Molecular dynamics simulations of membrane channels and transporters. Curr Opin Struct Biol 19(2):128–137
86. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The u SWISS-MODEL repository and associated resources. Nucleic Acids Res 37(suppl 1):D387
87. Kim MI, Kim J, Lee J, Jia H, Na HB, Youn JK, Kwak JH, Dohnalkova A, Grate JW, Wang P, Hyeon T, Park HG, Chang HN (2007) Crosslinked enzyme aggregates in hierarchically-ordered mesoporous silica: a simple and effective method for enzyme stabilization. Biotechnol Bioeng 96(2):210–218
88. Kirchmair J, Distinto S, Markt P, Schuster D, Wolber G (2011) Development of anti-viral agents using molecular modeling and virtual screening techniques. Infectious Disorders-Drug Targets 11(1):64
89. Klann MT, Lapin A, Reuss M (2009) Stochastic simulation of signal transduction: impact of the cellular architecture on diffusion. Biophys J 96(12):5122–5129
90. Klein MG, Zwart P, Bagby SC, Cai F, Chisholm SW, Heinhorst S, Cannon GC, Kerfeld CA (2009) Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. J Mol Biol 392(2):319–333
91. Klepeis J, Lindorff-Larsen K, Dror R, Shaw D (2009) Longtimescale molecular dynamics simulations of protein structure and function. Curr Opin Struct Biol 19(2):120–127
92. Lai Y-T, King NP, Yeates TO (2012) Principles for designing ordered protein assemblies. Trends Cell Biol 22(12):653–661
93. Larsson P, Hess B, Lindahl E (2011) Algorithm improvements for molecular dynamics simulations. Wiley Interdiscip Rev Comput Mol Sci 1(1):93–108
94. Lee H, DeLoache WC, Dueber JE (2012) Spatial organization of enzymes for metabolic engineering. Metab Eng 14(3):242–251
95. Lee SK, Chou H, Ham TS, Lee TS, Keasling JD (2008) Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. Curr Opin Biotechnol 19(6):556–563
96. Lingwood D, Ilya Levental H-JK, Simons K (2009) Lipid rafts as functional heterogeneity in cell membranes. Biochem Soc Trans 37(Pt 5):955–960
97. Ljungcrantz P, Carlsson H, Mansson MO, Buckel P, Mosbach K, Bulow L (1989) Construction of an artificial bifunctional enzyme, beta-galactosidase/galactose dehydrogenase, exhibiting efficient galactose channeling. Biochemistry 28(22):8786–8792
98. Loeffler HH, Winn M (2012) Large biomolecular simulation on HPC platforms III. AMBER, CHARMM, GROMACS, LAMMPS and NAMD. Technical report, STFC Daresbury Laboratory, Warrington WA4 4AD, UK
99. Lyubarev A, Kurganov B (1989) Supramolecular organization of tricarboxylic acid cycle enzymes. Biosystems 22(2):91–102
100. Ma C, Xiu Z, Zeng A (2011) A new concept to reveal protein dynamics based on energy dissipation. PLoS ONE 6(10):e26453
101. Maes D, Zeelen JP, Thanki N, Beaucamp N, Alvarez M, Thi MH, Backmann J, Martial JA, Wyns L, Jaenicke R, Wierenga RK (1999) The crystal structure of triosephosphate isomerase (TIM) from Thermotoga maritima: a comparative thermostability structural analysis of ten different TIM structures. Proteins 37(3):441–453

102. Mak L, Liggi S, Tan L, Kusonmano K, Rollinger JM, Koutsoukas A, Glen RC, Kirchmair J (2013) Anti-cancer drug development: computational strategies to identify and target proteins involved in cancer metabolism. Curr Pharm Design 19(4):532–577

103. Mansson MO, Siegbahn N, Mosbach K (1983) Site-to-site directed immobilization of enzymes with bis-NAD analogues. Proc Natl Acad Sci U S A 80(6):1487–1491

104. Mao Q, Schunk T, Gerber B, Erni B (1995) A String of Enzymes, Purification and Characterization of a Fusion Protein Comprising the Four Subunits of the Glucose Phosphotransferase System of Escherichia coli. J Biol Chem 270(31):18295–18300

105. Martínez de la Fuente I (2010) Quantitative analysis of cellular metabolic dissipative, self-organized structures. Int J Mol Sci 11(9):3540–3599

106. Martínez de la Fuente I, Cortes JM, Perez-Pinilla MB, Ruiz-Rodriguez V, Veguillas J (2011) The metabolic core and catalytic switches are fundamental elements in the self-regulation of the systemic metabolic structure of cells. PLoS ONE 6(11):e27224

107. McLain AL, Szweda PA, Szweda LI (2011) Ketoglutarate dehydrogenase: a mitochondrial redox sensor. Free Radic Res 45(1):29–36

108. Mei C, Sun Y, Zheng G, Bohm E, Kale L, Phillips J, Harrison C (2011) Enabling and scaling biomolecular simulations of 100 million atoms on petascale machines with a multicore-optimized message-driven runtime. In: High performance computing, networking, storage and analysis (SC), 2011 International conference for, pp 1–11. IEEE

109. Meynial Salles I, Forchhammer N, Croux C, Girbal L, Soucaille P (2007) Evolution of a Saccharomyces cerevisiae metabolic pathway in Escherichia coli. Metab Eng 9(2):152–159

110. Ming D, Kong Y, Lambert MA, Huang Z, Ma J (2002) How to describe protein motion without amino acid sequence and atomic coordinates. Proc Natl Acad Sci U S A 99(13):8620–8625

111. Miyazaki H, Kusano Y, Shinjou N, Shoji F, Yokokawa M, Watanabe T (2012) Overview of the K computer system. FUJITSU Sci Tech J 48(3):255–265

112. Monge C, Beraud N, Kuznetsov AV, Rostovtseva T, Sackett D, Schlattner U, Vendelin M, Saks VA (2008) Regulation of respiration in brain mitochondria and synaptosomes: restrictions of ADP diffusion in situ, roles of tubulin, and mitochondrial creatine kinase. Mol Cell Biochem 318(1):147–165

113. Monge C, Grichine A, Rostovtseva T, Sackett D, Kuznetsov A, Saks V (2009) Compartmentation of ATP in cardiomyocytes and mitochondria kinetic studies and direct measurements. Biophys J 96:241

114. Moon TS, Dueber JE, Shiue E, Prather KLJ (2010) Use of modular, synthetic scaffolds for improved production of glucaric acid in engineered E. coli. Metab Eng 12(3):298–305

115. Mosbach K, Mattiasson B (1970) Matrix-bound enzymes. II. Studies on a matrix-bound two-enzyme-system. Acta Chem Scand 24(6):2093–2100

116. Nepusz T, Yu H, Paccanaro A (2012) Detecting overlapping protein complexes in protein–protein interaction networks. Nat Methods 9(5):471–472

117. Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. Nature 388(6640):343–349

118. Niemeyer CM, Koehler J, Wuerdemann C (2002) DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. ChemBioChem 3(2–3):242–245

119. Noble D (2010) Biophysics and systems biology. Philos Transact A Math Phys Eng Sci 368(1914):1125–1139

120. Orita I, Sakamoto N, Kato N, Yurimoto H, Sakai Y (2007) Bifunctional enzyme fusion of 3-hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase. Appl Microbiol Biotechnol 76(2):439–445

121. Ovádi J, Saks V (2004) On the origin of intracellular compartmentation and organized metabolic systems. Mol Cell Biochem 256–257(1–2):5–12

122. Parsons JB, Frank S, Bhella D, Liang M, Prentice MB, Mulvihill DP, Warren MJ (2010) Synthesis of empty bacterial microcompartments, directed organelle protein incorporation, and evidence of filament-associated organelle movement. Mol Cell 38(2):305–315

123. Patel M, Korotchkina L (2006) Regulation of the pyruvate dehydrogenase complex. Biochem Soc Trans 34(2):217–222
124. Penrod JT, Roth JR (2006) Conserving a volatile metabolite: a role for carboxysome-like organelles in Salmonella enterica. J Bacteriol 188(8):2865–2874
125. Pérez-Bercoff A, McLysaght A, Conant GC (2011) Pate e terns of indirect protein interactions suggest a spatial organization to metabolism. Mol BioSyst 7(11):3056–3064
126. Petoukhov MV, Svergun DI (2005) Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys J 89(2):1237–1250
127. Pettersson H, Pettersson G (2001) Kinetics of the coupled reaction catalysed by a fusion protein of (beta)-galactosidase and galactose dehydrogenase. Biochim Biophys Acta: Protein Struct Mol Enzymol 1549(2):155–160
128. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comp Chem 26(16):1781–1802
129. Qu Z, Garfinkel A, Weiss J, Nivala M (2011) Multi-scale modeling in biology: how to bridge the gaps between scales? Prog Biophys Mol Biol 107(1):21–31
130. Ren Y, He J, Feng L, Liao X, Jin J, Li Y, Cao Y, Wan J, He H (2011) Structure-based rational design of novel hit compounds for pyruvate dehydrogenase multienzyme complex E1 components from Escherichia coli. Bioorg Med Chem 19(24):7501–7506
131. Resch MM (2013) Beyond exaflop computing: reaching the frontiers of high performance computing. In: Sust Sim Perf 2012, pp 3–12. Springer
132. Riedel K, Bronnenmeier K (1998) Intramolecular synergism in an engineered exo-endo-1,4-β-glucanase fusion protein. Mol Microbiol 28(4):767–775
133. Rizzuto R, Pozzan T (2006) Microdomains of intracellular Ca2+: molecular determinants and functional consequences. Physiol Rev 86(1):369–408
134. Saks V, Beraud N, Wallimann T (2008) Metabolic compartmentation- a system level property of muscle cells: real problems of diffusion in living cells. Int J Mol Sci 9(5):751
135. Saks V, Monge C, Guzun R (2009) Philosophical basis and some historical aspects of systems biology: from Hegel to noble-applications for bioenergetic research. Int J Mol Sci 10(3):1161–1192
136. Sampson EM, Bobik TA (2008) Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. J Bacteriol 190(8):2966–2971
137. Sanchez C, Lachaize C, Janody F, Bellon B, Röder L, Euzenat J, Rechenmann F, Jacq B (1999) Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an internet database. Nucleic Acids Res 27(1):89–94
138. Savage DF, Way J, Silver PA (2008) Defossiling fuel: how synthetic biology can transform biofuel production. ACS Chem Biol 3(1):13–16
139. Schaeffer RD, Fersht A, Daggett V (2008) Combining experiment and simulation in protein folding: closing the gap for small model systems. Curr Opin Struct Biol 18(1):4–9
140. Schulz R, Lindner B, Petridis L, Smith J (2009) Scaling of multimillion-atom biological molecular dynamics simulation on a petascale supercomputer. J Chem Theor Comp 5(10):2798–2808
141. Selivanov V, Krause S, Roca J, Cascante M (2007) Modeling of spatial metabolite distributions in the cardiac sarcomere. Biophys J 92(10):3492–3500
142. Seo HS, Koo YJ, Lim JY, Song JT, Kim CH, Kim JK, Lee JS, Choi YD (2000) Characterization of a bifunctional enzyme fusion of trehalose-6-phosphate synthetase and trehalose-6-phosphate phosphatase of Escherichia coli. Appl Environ Microbiol 66(6):2484–2490
143. Shan Y, Klepeis J, Eastwood M, Dror R, Shaw D (2005) Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. J Chem Phys 122:054101
144. Sherawat M, Tolan DR, Allen KN (2008) Structure of a rabbit muscle fructose-1,6-bisphosphate aldolase A dimer variant. Acta Cryst D 64(5):543–550

145. Shibuya I, Tamura G, Shima H, Ishikawa T, Hara S (1992) Construction of an α-amylase/glucoamylase fusion gene and its expression in Saccharomyces cerevisiae. Biosci Biotechnol Biochem 56(6):884–889

146. Shoham Y, Lamed R, Bayer EA (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends Microbiol 7(7):275–281

147. Siddiqui MS, Thodey K, Trenchard I, Smolke CD (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. FEMS Yeast Res 12(2):144–170

148. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, Clair JLS, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL et al (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction. Science 329(5989):309–313

149. Simons K, Ikonen E (1997) Functional rafts in cell membranes. Nature 387:569–572

150. Slepchenko BM, Schaff JC, Macara I, Loew LM (2003) Quantitative cell biology with the virtual cell. Trends Cell Biol 13(11):570–576

151. Ślepokura K, Lis T (2010) Dihydroxyacetone phosphate, DHAP, in the crystalline state: monomeric and dimeric forms. Carbohyd Res 345(4):512–529

152. Srere PA, Mattiasson B, Mosbach K (1973) An immobilized three-enzyme system: a model for microenvironmental compartmentation in mitochondria. Proc Natl Acad Sci U S A 70(9):2534–2538

153. Srere P (1987) Complexes of sequential metabolic enzymes. Annu Rev Biochem 56(1):89–124

154. Srere P, Ovadi J (1990) Enzyme-enzyme interactions and their metabolic role. FEBS Lett 268(2):360–364

155. Stumpf MP, Thorne T, De Silva E, Stewart R, An HJ, Lappe M, Wiuf C (2008) Estimating the size of the human interactome. PNAS 105(19):6959–6964

156. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO (2008) Atomic-level models of the bacterial carboxysome shell. Science 319(5866):1083–1086

157. Tarassov K, Messier V, Landry CR, Radinovic S, Molina MMS, Shames I, Malitskaya Y, Vogel J, Bussey H, Michnick SW (2008) An in vivo map of the yeast protein interactome. Science 320(5882):1465–1470

158. Tomita M (2001) Whole-cell simulation: a grand challenge of the 21st century. Trends Biotechnol 19(6):205–210

159. Tran L, Broadhurst RW, Tosin M, Cavalli A, Weissman KJ (2010) Insights into protein–protein and enzyme-substrate interactions in modular polyketide synthases. Chem Biol 17(7):705–716

160. van de Velde F, Lourenço ND, Bakker M, van Rantwijk F, Sheldon RA (2000) Improved operational stability of peroxidases by coimmobilization with glucose oxidase. Biotechnol Bioeng 69(3):286–291

161. van Dongen SFM, Nallani M, Cornelissen JJLM, Nolte RJM, van Hest JCM (2009) A three-enzyme cascade reaction through positional assembly of enzymes in a polymersome nanoreactor. Chem Eur J 15(5):1107–1114

162. Vijayakrishnan S, Kelly S, Gilbert R, Callow P, Bhella D, Forsyth T, Lindsay J, Byron O (2010) Solution structure and characterisation of the human pyruvate dehydrogenase complex core assembly. J Mol Biol 399(1):71–93

163. Vriezema DM, Garcia PML, Sancho Oltra N, Hatzakis NS, Kuiper SM, Nolte RJM, Rowan AE, van Hest JCM (2007) Positional assembly of enzymes in polymersome nanoreactors for cascade reactions. Angew Chem Int Ed 46(39):7378–7382

164. Walther TH, Gottselig C, Grage SL, Wolf M, Vargiu AV, Klein MJ, Vollmer S, Prock S, Hartmann M, Afonin S et al (2013) Folding and self-assembly of the tata translocation pore based on a charge zipper mechanism. Cell 152(1):316–326

165. Wei B, Dai M, Yin P (2012) Complex shapes self-assembled from single-stranded DNA tiles. Nature 485(7400):623–626

166. Welch G, Easterby J (1994) Metabolic channeling versus free diffusion: transition-time analysis. Trends Biochem Sci 19(5):193

167. Westerhoff HV, Kolodkin A, Conradie R, Wilkinson SJ, Bruggeman FJ, Krab K, van Schuppen JH, Hardin H, Bakker BM, Moné MJ et al (2009) e Systems biology towards life in silico: mathematics of the control of living cells. J Math Biol 58(1):7–34
168. Wörsdörfer B, Woycechowsky KJ, Hilvert D (2011) Directed evolution of a protein container. Science 331(6017):589–592
169. You C, Myung S, Zhang YHP (2012) Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angew Chem Int Edit 51(35):8787–8790
170. You C, Zhang XZ, Sathitsuksanoh N, Lynd LR, Zhang Y-HP (2012) Enhanced microbial utilization of recalcitrant cellulose by an ex vivo cellulosome-microbe complex. Appl Environ Microbiol 78(5):1437–1444
171. You C, Zhang X-Z, Zhang Y-HP (2012) Mini-scaffoldin enhanced mini-cellulosome hydrolysis performance on low-accessibility cellulose (Avicel) more than on high-accessibility amorphous cellulose. Biochem Eng J 63:57–65
172. You C, Zhang YHP (2012) Self-assembly of synthetic metabolons through synthetic protein scaffolds: one-step purification, co-immobilization, and substrate channeling. ACS Synth Biol 2:102–110
173. Zhang Y-HP (2010) Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol Bioeng 105(4):663–677
174. Zhang YHP (2011) Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol Adv 29(6):715–725
175. Zhang YHP, Myung S, You C, Zhu Z, Rollin JA (2011) Toward low-cost biomanufacturing through in vitro synthetic biology: bottom-up design. J Mater Chem 21(47):18877–18886

Cell-Free Systems: Functional Modules for Synthetic and Chemical Biology

Marlitt Stech, Andreas K. Brödel, Robert B. Quast, Rita Sachse and Stefan Kubick

Abstract The main goal of cell-free protein synthesis is to produce correctly folded and functional proteins in reasonable amounts for further downstream applications. Especially for eukaryotic proteins, functionality is often directly linked to the presence of posttranslational modifications. Thus, it is of highest interest to develop novel cell-free expression systems that enable the synthesis of posttranslationally modified proteins. Here we present recent advances for the synthesis of glycoproteins, proteins containing disulfide bridges, membrane proteins, and fluorescently labeled proteins. The basis for the expression of these difficult-to-express target proteins is a translationally active cell extract which can be prepared from eukaryotic cell lines such as *Spodoptera frugiperda* 21 (*Sf*21) and Chinese hamster ovary (CHO) cells. Due to a very mild lysate preparation procedure, microsomal vesicles derived from the endoplasmic reticulum (ER) can be maintained in the eukaryotic lysate. These vesicles are translocationally active and serve as functional modules facilitating protein translocation and enrichment as well as posttranslational modification of de novo synthesized proteins. In particular, for the synthesis of membrane proteins microsomal vesicles are the essential prerequisite for the insertion of the desired protein into a biologically active membrane scaffold providing a natural environment. We anticipate that the use of such translationally active eukaryotic cell lysates containing translocationally active vesicles may solve a large number of problems still persistent when expressing eukaryotic proteins in vitro.

Keywords Cell-free protein synthesis · Fluorescent label · GUV · In vitro translation · Insect cell lysate · Membrane protein · Microsome · Noncanonical amino acid · Orthogonal tRNA/Synthetase pair · Posttranslational modification · Single-chain antibody (scFv) · Vesicle

M. Stech · A. K. Brödel · R. B. Quast · R. Sachse · S. Kubick (✉)
Fraunhofer IBMT, Branch Potsdam, Group of Cell-free Protein Synthesis,
Am Mühlenberg 13, 14476 Potsdam, Germany
e-mail: Stefan.Kubick@ibmt.fraunhofer.de

Abbreviations

ATP	Adenosine-5'-triphosphate
AzPhe	*p*-azido-*L*-phenylalanine
AzPheRS	Orthogonal mutant synthetase specific for AzPhe
BP	Bandpass filter
CECF	Continuous-exchange cell-free
CFCF	Continuous-flow cell-free
CHO	Chinese hamster ovary
CLSM	Confocal laser scanning microscopy
CrPV	Cricket paralysis virus
DDM	*n*-dodecyl-*ß*-maltoside
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EndoH	Endoglycosidase *H*
E-PCR	Expression-PCR
EPO	Erythropoietin
ER	Endoplasmic reticulum
eYFP	Enhanced yellow fluorescent protein
Fab	Fragment, antigen-binding
FITC	Fluorescein isothiocyanate
Fv	Antibody variable fragment
GTP	Guanosine-5'-triphosphate
GUV	Giant unilamellar vesicle
IGR	Intergenic region
IRES	Internal ribosome entry site
KD	Association equilibrium constant
Mel	Melittin
PAGE	Polyacrylamide gel electrophoresis
PDI	Protein disulfide isomerase
PI	Protease inhibitors
PNGaseF	*N*-glycosidase *F*
PTM	Posttranslational modification
RU	Resonance units
scFv	Single-chain antibody fragment
SDS	Sodium dodecyl sulfate
*Sf*21	*Spodoptera frugiperda 21*
SN1	Supernatant 1
SN2	Supernatant 2
SPR	Surface plasmon resonance
TCA	Trichloroacetic acid
TM	Translation mixture
$tRNA_{CUA}$	Amber suppressor tRNA
VF1	Vesicular fraction 1
VF2	Vesicular fraction 2

Contents

1 Introduction

In the last decades cell-free protein synthesis, also termed in vitro translation, has become a valuable tool for structural and functional proteomics. Starting as a research tool to investigate the fundamentals of translation processes in vitro [1] it has now evolved to serve as a reliable and versatile protein production platform [2, 3]. In particular, cell-free protein synthesis has facilitated the successful production of pharmaceutically relevant target proteins [4–7], such as membrane proteins [8–14] or the high-throughput production of protein libraries [15–17]. The strategy of cell-free protein expression shows crucial advantages over conventional in vivo production methods [2, 3]. Protein production in living cells is always a compromise between the conditions that are required for cell growth and viability and the conditions necessary for the synthesis of a functional target protein. In contrast, cell-free systems represent open systems, wherein reaction conditions can be adjusted according to the needs of the individual protein without consideration of the conditions necessary for cell cultivation. Moreover, cell-free protein synthesis can be performed using linear DNA templates omitting the need for time- and labor-consuming cloning steps. For protein production in vivo an average working time of 14 days is required, starting from plasmid construction, cloning, transfection, and cell cultivation to product purification. In contrast, cell-free protein

synthesis is much faster and can be performed in one to two days, including experimental steps as template generation, protein expression, and purification, thus decreasing protein production costs dramatically [5, 7, 17].

The basis of a cell-free expression system is a crude cell extract that can be obtained from all different types of living cells. Lysates gained from *Escherichia coli* (*E. coli*), wheat germ, rabbit reticulocytes, and insect cells [17] are very common and widely used. Inasmuch as every cell lysate has defined characteristics, one has to consider carefully which type of lysate shall be used to express a desired protein. One of the first decisions to make is whether a prokaryotic or a eukaryotic in vitro translation system is the better choice. Usually this decision will be influenced by the following considerations: protein characteristics such as origin, complexity, and solubility as well as the yield of protein needed; downstream processing; the availability of resources and time and costs [17]. Prokaryotic cell lysates have the advantage of a long-lasting expertise, low costs, and high production yields but only very few posttranslational modifications (PTMs) can be carried out on target proteins [2, 18]. This represents a major drawback in comparison to the use of eukaryotic cell lysates as these lysates are well-suited for the synthesis of complex proteins and proteins with PTMs not found in bacteria [12, 19, 20].

One of the main demands of cell-free systems is the potential to produce functional proteins. Covalent PTMs such as glycosylation, signal peptide cleavage, phosphorylation, farnesylation, acetylation, palmitoylation, and disulfide bond formation have a great impact on protein folding, localization, and activity [21]. Thus, it is important to ensure PTM formation also in cell-free systems [12]. The majority of PTMs are processed inside the endoplasmic reticulum (ER), therefore it is desirable to implement parts of the ER in a cell-free system. This was achieved either by developing heterogeneous or homogeneous translation systems that both implement microsomal vesicles in a translationally active cell lysate. In heterogeneous systems the cell lysate and the microsomal fraction are derived from different types of cells [22, 23], whereas in homogeneous systems both fractions are gained from a single cell type. An example of such a homogeneous in vitro translation system is represented by the insect cell-free expression system [11, 14, 24]. Due to a sophisticated extract preparation procedure that retains the integrity of subcellular components, vital parts of the ER can be maintained in the lysate as translocationally active vesicles. Although the ER in its native structure is ripped during cell disruption, membrane structures rearrange and form smaller compartments designated as microsomal vesicles. The enzymatic activity of the biological membranes survives this process as protein translocation and membrane protein insertion have been reported. Using this type of lysate PTMs such as glycosylation, lipidation, phosphorylation, signal peptide cleavage, and disulfide bond formation can be performed on de novo synthesized proteins [11, 12, 25–28].

Cell-free protein synthesis has made remarkable progress in the expression of complex proteins including proteins with disulfide bridges. To ensure the synthesis of correctly folded proteins with disulfide bridges the applied cell-free system must fulfill certain requirements. All needs considered, the expression system should

provide an oxidizing environment, implement enzymes that facilitate disulfide bond formation and may apply folding helpers such as chaperones [17]. Of course, the extent of modification depends on the type of lysate to be used. In particular, for the synthesis of antibody fragments cell-free expression systems have gained growing attention. Recently, the production of Fab and single-chain antibody fragments (scFv) has been reported in insect cell lysates which only need moderate modification of the reaction conditions while yielding highly functional molecules [25, 27]. In contrast to *E. coli*-based systems, these lysates do not need to be supplemented with endogenous enzymes and cofactors. Moreover, they allow a production of functional Fab and scFv fragments without endotoxin contamination and thereby may enable a direct application of cell-free synthesized molecules, for example, in cell-based assays.

Currently, there is a large gap between the number of sequenced and annotated DNA sections and the systematic study of the encoded proteins. In particular, the expression and characterization of membrane proteins lags far behind the analysis of cytosolic proteins. This imbalance gets even more relevance when one considers that three quarters of all potential drug targets account for membrane proteins [29]. Membrane proteins constitute a difficult-to-express class of proteins as their overexpression in vivo very often leads to cytotoxic effects or results in protein aggregation, misfolding, and low yields [2]. However, in vitro translation can circumvent these obstacles. As cell-free systems provide open reaction conditions, they can be supplemented with any kind of reagent that allows a solubilization of membrane proteins and supports their correct folding and assembly [30, 31]. The supplementation of lysates with natural or synthetic lipids and/or detergents [32], purified *E. coli* phospholipid bilayer vesicles [33], nanolipoprotein particles [34], or unilamellar liposomes [35] have shown to be promising strategies. An alternative strategy is represented by the utilization of eukaryotic cell compartments in order to provide a native environment that allows a cotranslational translocation and insertion of membrane proteins. This strategy has been successfully performed by using a homogeneous cell-free translation system based on insect cell lysates that contain microsomal vesicles derived from the ER [11]. As reported in the literature, these membrane vesicles can be manipulated regarding their size [26] and physicochemical properties. Such a manipulation allows the implementation of membrane vesicles in technical systems such as biochips or microarrays, opening up new strategies to screen for pharmacologically relevant target proteins.

In the near future, we anticipate a growing need for powerful and flexible eukaryotic protein expression systems. Insect cell lysates proved to be a valuable production platform, but other eukaryotic cell lysates are also of great interest. We have addressed this need by developing cell-free systems based on CHO cells and also human cell lines. These expression systems are very beneficial, especially for the production of posttranslationally modified proteins such as glycosylated proteins.

The incorporation of unnatural amino acids into cell-free synthesized proteins, including eukaryotic proteins harboring PTMs, is another important issue. The natural repertoire of primary building blocks for protein synthesis is limited to the

22 standard or proteinogenic amino acids. In spite of this moderate number underlying the broad diversity of protein functions found in nature, great efforts have been made in order to expand this natural repertoire. To date, a large number of nonnatural or noncanonical amino acids with diverse characteristics has been cotranslationally incorporated into proteins. The physicochemical properties of these unnatural building blocks range from being closely related to the standard amino acids to being uniquely artificial ones [36–38]. Incorporation of noncanonical amino acids bearing chemically or physically reactive side chains represents a promising strategy to efficiently introduce novel characteristics to mimic or further enhance protein properties in vivo and in vitro [37, 39]. This toolbox of new properties has facilitated many different areas of research as well as industrial applications and will most likely have a great impact on future perspectives in chemical and synthetic biology.

2 Production of Functional Antibody Fragments in Cell-Free Systems

2.1 Antibody Engineering

Antibody molecules consist of different domains, each exhibiting specific functions. Based on genetic engineering techniques, the subdomains of an antibody molecule can be rearranged in order to develop new and advantageous recombinant antibody formats. In particular, genetically truncated variants of antibodies such as Fab or Fv fragments are of great importance for pharmaceutical research. The antigen-binding surface of Fab fragments and Fv fragments is identical to the one of their corresponding full-length antibody. Therefore, they provide the full binding specificity with the advantage of being less complex and easier to synthesize. Moreover, they exhibit improved behavior regarding their pharmacokinetics for tissue penetration. However, as these formats of engineered antibodies are monovalent, fast dissociation rates and poor retention times on the target have been documented [40–42]. Nevertheless, recombinant DNA technology has found new ways to further improve their behavior by changing the format to dimeric, trimeric, or tetrameric conjugates [40, 43–45].

2.2 Cell-Free Synthesis of Proteins with Disulfide Bridges in Prokaryotic Cell Lysates

Cell-free systems used for the synthesis of functional Fab and Fv fragments need to provide an environment that facilitates the formation of disulfide bonds. In most cases immunoglobulin domains are stabilized by an intramolecular disulfide bond.

This disulfide bond connects two ß-sheets that form one immunoglobulin domain [46, 47]. As reported, the existence of the intramolecular disulfide bond mainly influences the stability, folding, and functionality of the antibody fragment [48, 49]. Until now, cell-free synthesis of functional proteins containing disulfide bonds was successfully performed using modified *E. coli* cell extracts [50–56], wheat germ extracts [57], and insect cell lysates [25, 27, 58].

Several publications have shown that the formation of disulfide bonds in *E. coli* cell lysates requires additional precautions. Modifications of the lysate include the supplementation of molecular chaperones as well as the adjustment of the redox potential by adding mixtures of oxidized [54] and reduced glutathione in combination with protein disulfide isomerase (PDI) [51, 53, 55, 56]. In addition, the chemical pretreatment of lysates with iodoacetamide can also be combined with the above-mentioned modifications [50, 52, 59]. In the literature, the expression of several disulfide bridge-containing proteins in *E. coli* cell lysates has been reported, for example, Fab fragments [51, 55], scFv fragments [53, 56], recombinant plasminogen activator [52, 54], and urokinase protease [52] as well as virus like particles [50].

However, the use of prokaryotic cell-free systems also implies significant disadvantages. *E. coli* cell lysates have to be supplemented with exogenous proteins such as PDI and chaperones in order to allow disulfide bond formation, thus implicating additional working steps. The high productivity of *E. coli* cell lysates may be accompanied by aggregation of insoluble target proteins. Furthermore, one has to consider that proteins synthesized in *E. coli* cells and their extracts are contaminated with endotoxins that have to be removed before the expressed target protein can be applied in cell-based assays. Hence, eukaryotic cell-free translation systems are of great interest and may overcome these obstacles.

2.3 Production of Recombinant Antibody Fragments in Eukaryotic Cell Lysates

In this section we focus on the cell-free expression of antibody fragments using insect cell lysates as a source for protein production. As reported recently, insect cell lysates only need moderate modification including mainly the adjustment of redox conditions to enable the synthesis of functional antibody fragments [27, 58]. Due to their open nature, cell-free translation systems are readily accessible. Hence, redox conditions can be easily adjusted using defined mixtures of reduced and oxidized glutathione. These substances can be added to the cell lysate itself or to the buffer used for the translation reaction.

Eukaryotic cells comprise defined subcellular compartments resulting in many different organelles, including the ER, the cell nucleus, the Golgi apparatus, and mitochondria. These organelles can be considered as individual cell compartments usually surrounded by a single- or double-lipid bilayer harboring a specific protein

translocation machinery [60]. The evolution of cell compartments is a key feature of eukaryotic cells and an attribute of higher organisms [61]. When creating novel cell-free systems one could utilize these natural subcellular structures in order to create a nature like environment, thus ensuring optimum conditions for the expressed protein. Taking these structures as a basis, they can be further developed and modified. In the end a system evolves that has nature as a model while overcoming its limitations. The open nature of cell-free systems is the prerequisite for rapid modifications and adjustments of the reaction conditions depending on the individual requirements of the protein to be synthesized. In addition, Fab fragment and scFv production in eukaryotic lysates containing microsomal vesicles has the particular advantage that the cellular natural compartment used for in vivo disulfide bond formation can be utilized for the same purpose in vitro. This may then have a beneficial effect on the efficiency and quality of disulfide bond formation.

2.4 Using microsomal vesicles as protein microcontainers

The fusion of a protein sequence to a specific signal sequence enables the translocation of proteins into defined cell compartments [22]. The signal peptide of honeybee melittin (Mel) has been shown to be well suited to promote the translocation of de novo synthesized proteins into the lumen of the ER in insect cells [62]. Hence, this phenomenon can be used in cell-free reactions to target proteins into microsomal structures derived from the ER. According to the characteristics of the protein of interest, these so-called microsomes can then be used as microcontainers enabling the enrichment of the translocated target protein. They also serve as the prerequisite for PTMs such as glycosylation, lipidation, and signal peptide cleavage [11, 12, 26]. Furthermore, insect vesicles may have a positive effect on the formation of disulfide bonds in proteins inasmuch as they mimic the appropriate environment present in living cells. Insect microsomes can be separated by a single centrifugation step from the translation mixture (TM), thus enabling a rapid separation and enrichment of the translocated target protein [27].

As reported, scFv fragments fused to a melittin signal sequence (Mel-Anti-FITC scFv) can be efficiently translocated into the lumen of insect vesicles [27]. This observation was shown by confocal laser scanning microscopy (CLSM) and trichloroacetic acid (TCA) precipitation results combined with autoradiography. As demonstrated by fluorescence analysis, insect vesicles containing the fluorescent target protein can be quantitatively separated from the cytosolic fraction by centrifugation (Fig. 1). In order to release the translocated and vesicle-trapped antibody molecules, a detergent-containing buffer was used. N-dodecyl-β-maltoside (DDM) was shown to be the best-suited detergent, allowing for an efficient release of the target proteins while preserving their functionality.

Insect vesicles can maintain their mechanical and functional stability over a long period of time as demonstrated by the cumulative enrichment of Mel-Anti-FITC scFv in the lumen of these vesicles after several consecutive runs of protein
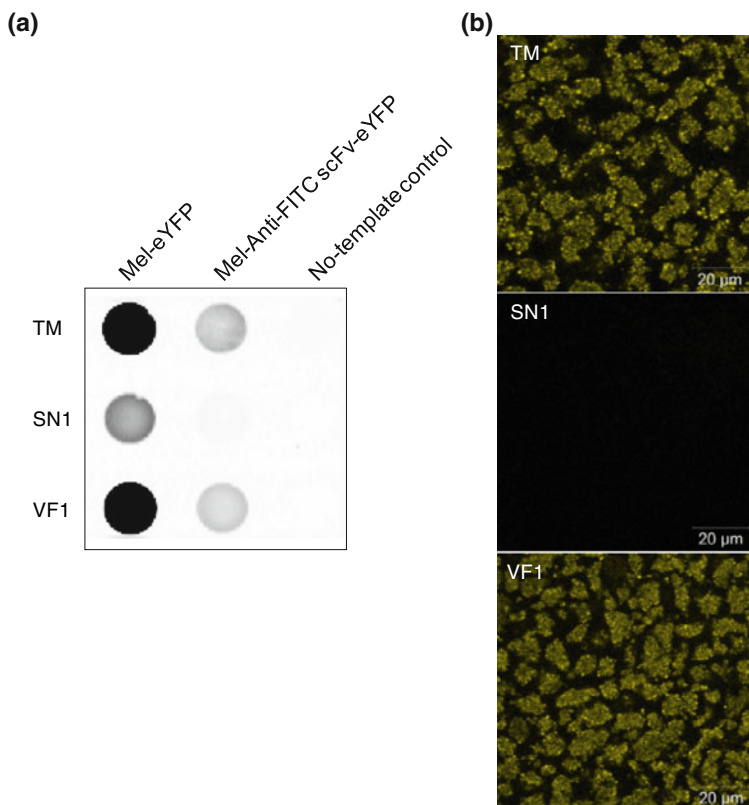
Fig. 1 Fluorescence analysis of Mel-eYFP and Mel-Anti-FITC scFv-eYFP synthesized in insect cell lysate using the phosphorimager system (Typhoon TRIO+ Imager, GE Healthcare) (**a**) and CLSM (LSM 510 Meta microscope, Zeiss) (**b**). **a** Mel-eYFP, Mel-Anti-FITC scFv-eYFP, and a no-template control were analyzed in the translation mixture (TM), the supernatant (SN1), and the vesicular fraction (VF1). Strong emission intensity arising from VF1 indicates translocation of Mel-eYFP and Mel-Anti-FITC scFv-eYFP into the lumen of the insect vesicles. **b** Samples were excited at 488 nm using an argon laser and fluorescence emission was recorded with a longpass filter in the wavelength range above 505 nm. Fluorescent insect vesicles in TM and VF1 indicate efficient translocation of Mel-Anti-FITC scFv into their lumen

synthesis and translocation (Fig. 2). Up to five translation reactions were performed using the same set of vesicles originating from the first batch of translation. As presented in Fig. 2, protein yields increase from synthesis step to step in the translation mixture (TM) and the vesicular fraction (VF1). These results demonstrate the successful enrichment of the target proteins in the lumen of the insect vesicles as well as the remarkable mechanical and functional stability of the microsomal vesicles.
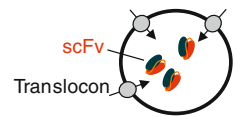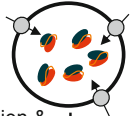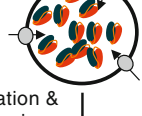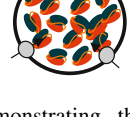
Synthesis step	ER-Vesicle	Protein yield TM [μg/ml]	Protein yield VF1 [μg/ml]
1	scFv / Translocon	12.6 ± 0.3	5.0 ± 0.2
	Centrifugation & Resuspension		
2		18.6 ± 2.0	8.8 ± a 0.6
	Centrifugation & Resuspension		
3		20.9 ± 1.7	12.5 ± 0.6
	Centrifugation & Resuspension		
4		26.8 ± 2.7	14.3 ± 0.3
	Centrifugation & Resuspension		
5		28.6 ± 3.6	23.5 ± 1.5

Fig. 2 Schematic overview demonstrating the cumulative enrichment of scFv antibody fragments fused to a melittin signal sequence in insect vesicles. Cell-free synthesis and translocation of Mel-Anti-FITC scFv into the lumen of insect vesicles was performed in five sequential protein synthesis steps using the same set of vesicles that originated from the initial translation reaction. The first translation batch was performed using insect cell lysate including endogenous insect vesicles. After each synthesis step, total translation reactions (TM) were centrifuged and separated into a supernatant and a vesicular fraction (VF1). VF1 was resuspended in a fresh translation mixture including insect cell lysate without vesicles. The concentration of ^{14}C-leucine-labeled Mel-Anti-FITC scFv increased from synthesis one to five in TM and VF1 demonstrating the successful accumulation of the target proteins in the lumen of the insect vesicles

A typical time course analysis of cell-free synthesized Mel-Anti-FITC scFv is given in Fig. 3a indicating the highest protein yields after 240 min of incubation. Time-dependent protein quantification by liquid scintillation counting supports the assumption of an efficient translocation of Mel-Anti-FITC scFv into the lumen of insect vesicles, as higher protein yields were reported in VF1 compared to the supernatant fraction (SN1; Fig. 3b). Translocated and trapped scFv molecules present in VF1 were then subjected to resolubilization using DDM-containing
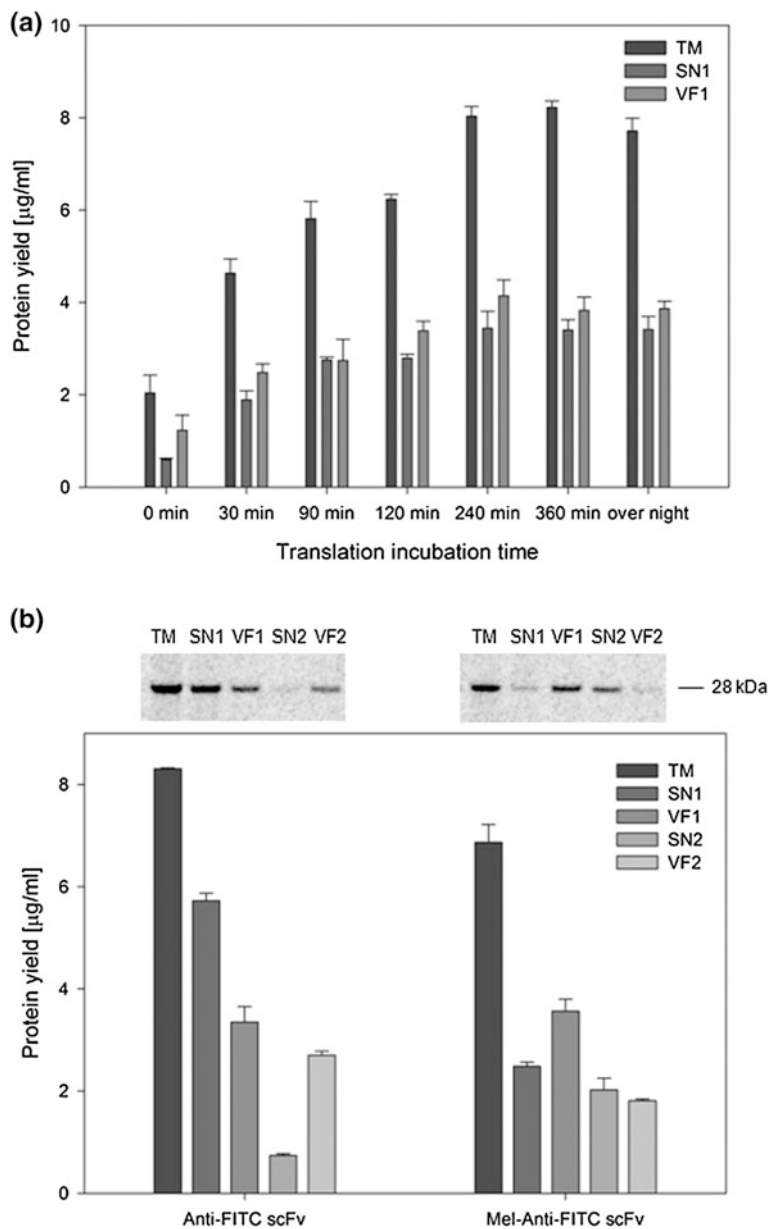
Fig. 3 Cell-free expression of Anti-FITC scFv and Mel-Anti-FITC scFv in insect cell lysate in the presence of ^{14}C-leucine, assigned by hot TCA precipitation and autoradiography. **a** Time course analysis of cell-free expressed Mel-Anti-FITC scFv analyzed in the translation mixture (TM), the supernatant (SN1) after centrifugation and the vesicular fraction (VF1). **b** Comparative analysis of Anti-FITC scFv and Mel-Anti-FITC scFv. TMs were fractionated into four different lysate fractions by centrifugation: supernatant 1 (SN1), vesicular fraction 1 (VF1), supernatant 2 (SN2), and vesicular fraction 2 (VF2). SN2 was gained after resuspension of VF1 in DDM-containing buffer. The molecular mass of Anti-FITC scFv and Mel-Anti-FITC scFv detected in the autoradiograph corresponds to the calculated molecular mass of 26.8 kDa
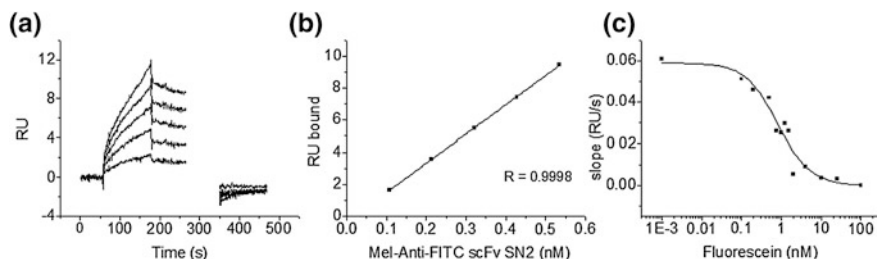
Fig. 4 Functional analysis of Mel-Anti-FITC scFv (SN2) synthesized in insect cell lysate by surface plasmon resonance (Biacore T100 system). **a** Specific binding of Mel-Anti-FITC scFv to immobilized FITC depending on the applied scFv concentration. Injected concentrations of Mel-Anti-FITC scFv increase from bottom to top: 0.2, 0.4, 0.6, 0.8, and 1 nM Mel-Anti-FITC scFv in SN2. **b** Calibration curve derived from the Biacore sensorgrams presented in (**a**). **c** Inhibition curves of Mel-Anti-FITC scFv premixed with increasing concentrations of fluorescein (0.0, 0.1, 0.2, 0.5, 0.75, 1.0, 1.25, 1.5, 2.0, 4.0, and 10 nM fluorescein). For Mel-Anti-FITC scFv a K_D of 1.1 nM, for Anti-FITC scFv a K_D of 1.2 nM was determined. RU: resonance units

buffer. In order to separate released antibody fragments from vesicular remnants, an additional centrifugation step was performed. The resulting supernatant, referred to as supernatant 2 (SN2), was applied to functional analysis. Two independent label-free interaction analysis methods, fluorescence quenching analysis (data not shown) and surface plasmon resonance (Biacore T100 system), demonstrated the functionality of the model scFvs Anti-FITC and Mel-Anti-FITC [27]. By surface plasmon resonance a K_D value of 1.1 and 1.2 nM was determined for Anti-FITC scFv and Mel-Anti-FITC scFv, respectively (Fig. 4). These results show the successful adaption of the insect cell-free expression system to produce functional active scFv molecules, thus demonstrating the high flexibility and capability of this eukaryotic translation system.

3 Development of a Continuous-Exchange Cell-Free Translation System Based on Insect Cell Lysates

Batch-formatted cell-free translation reactions are well suited for the fast, reliable, and easy-to-handle production of a given target protein. Nevertheless, the rapid depletion of energy resources and the accumulation of inhibitory reaction by-products such as free phosphates lead to a short lifetime of the system and comparatively low protein yields [2]. An experimental strategy to overcome these problems was first introduced by Spirin and coworkers, termed a continuous-flow cell-free (CFCF) translation system [63]. Although increased protein yields could be reached with this system its operational complexity was rather challenging.

This disadvantage was first overcome by the development of the continuous-exchange cell-free (CECF) system [64, 65]. The principle of a cell-free translation

reaction carried out in a CECF system is straightforward: protein synthesis takes place in a reaction chamber that is separated from a second compartment, the so-called feeding compartment, by a semipermeable dialysis membrane. Energy and substrates are continuously supplied via the dialysis membrane and reaction by-products diffuse into the feeding compartment. Reaction times exceeded 24 h and significantly higher protein yields have been reported [63, 66, 67]. The combination of CECF translation systems and *E. coli* cell lysates is already widely used and kits are commercially available. Yields up to milligrams of protein per 1 ml reaction volume have been documented thus far [18]. Nowadays, a combination of high protein yields and the possibility of producing posttranslationally modified proteins is favored. As eukaryotic cell lysates enable the synthesis of various posttranslationally modified proteins and CECF-systems are a valuable tool to increase the protein yield, a combination of both advantages seems to overcome longstanding limitations. A schematic overview of a batch-formatted and a CECF-formatted cell-free transcription/translation reaction is illustrated in Fig. 5a and b, respectively.

Here we present results on the cell-free expression of two model proteins combining a coupled cell-free translation system based on insect lysate with a commercially available CECF device (two-chamber microplate, 5PRIME). Two proteins, the enhanced yellow fluorescent protein (eYFP), N-terminally fused to a Strep-tag (SII-eYFP), and the human glycoprotein erythropoietin (EPO), with its native signal sequence exchanged by the melittin signal sequence, were chosen to demonstrate the proof of principle. As expected, both proteins were expressed with higher protein yields in the CECF-formatted reaction compared to the batch-formatted reaction as shown by fluorescence analysis (SII-eYFP; Fig. 5c) and autoradiography (EPO; Fig. 5d). Most important, these experiments have shown that the microsomal vesicles present in the insect cell lysate are completely compatible with the applied dialysis membrane (Fig. 5c). To achieve a further increase in the protein yield and to prolong the lifetime of the reaction we supplemented the cell-free reaction with protease inhibitors (PI). Using these inhibitors the reaction lifetime was extended from 2 h (batch) to 48 h (CECF) and protein yields of SII-eYFP increased significantly as demonstrated by fluorescence analysis (Fig. 5c). In addition, TCA precipitation of ^{14}C-leucine-labeled SII-eYFP revealed an increase in the amount of synthesized protein from 13 µg/ml, synthesized in a batch-reaction supplemented with PI (batch, + PI), to 260 µg/ml, synthesized in a CECF-system supplemented with PI (CECF, + PI) (Fig. 6). The translocation of EPO into the lumen of the insect vesicles present in the eukaryotic CECF system and its subsequent glycosylation was verified by the enzymatic digestion of the sugar moieties and visualized by autoradiography (Fig. 5d).

Taken together, our experiments show the successful development of a coupled eukaryotic transcription/translation system that can be used in combination with a commercially available CECF device. Combination of both the vesicle-containing eukaryotic translation system and the CECF system enables the production of cytosolic proteins as well as posttranslationally modified proteins with significantly increased protein yields.
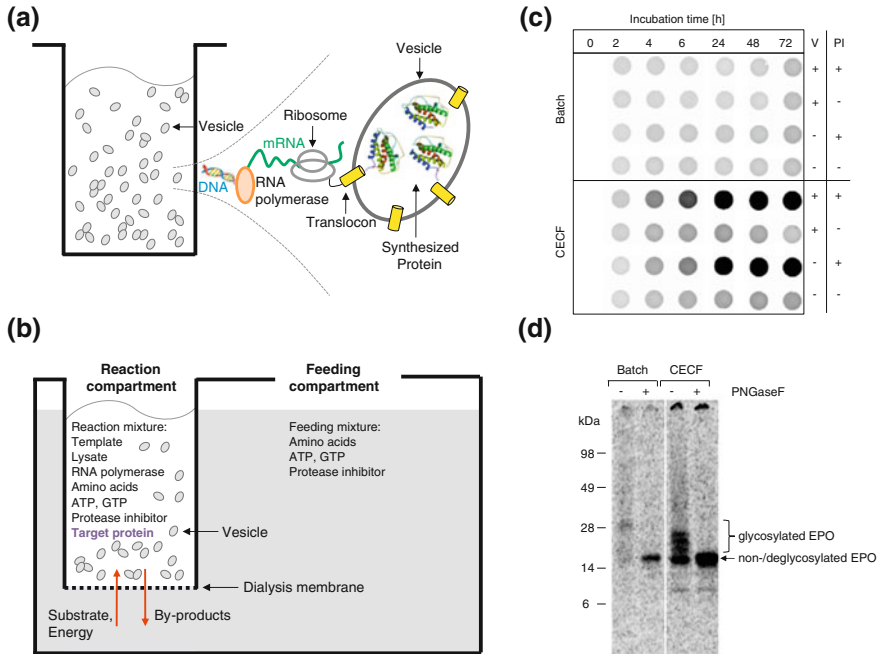
Fig. 5 Schematic presentation of a batch-formatted (**a**) and a CECF-formatted (**b**) cell-free transcription/translation reaction. **a** Synthesized protein: human erythropoietin (EPO) (NMR structure EPO, source: Protein Data Bank). **b** The CECF device (5PRIME) is composed of two chambers, the reaction compartment and the feeding compartment, which are separated by a semipermeable dialysis membrane (molecular weight cut-off: 10 kDa). In the reaction compartment protein synthesis takes place as the cell lysate containing the translational machinery and the microsomal vesicles are trapped in this compartment. The feeding compartment includes amino acids, ATP, GTP, and additional low molecular mass substances that can diffuse through the membrane. **c** Fluorescence analysis of SII-eYFP synthesized in a batch-formatted and a CECF-formatted coupled cell-free translation system using either insect lysate with (+) or without (–) vesicles (V). Cell-free reactions were supplemented with (+) protease inhibitors (PI) in order to prolong the reaction lifetime. Control reactions were carried out without (–) the addition of PI. Samples were analyzed using the phosphorimager system (Typhoon TRIO+ Imager, GE Healthcare). **d** Autoradiograph showing ^{14}C-leucine-labeled EPO synthesized in a batch-formatted and a CECF-formatted coupled cell-free translation system using insect lysate. EPO is presented before (–) and after (+) deglycosylation with PNGaseF. The migration pattern corresponds to the expected molecular mass of glycosylated and nonglyco-sylated EPO (21 kDa, nongylcosylated)

4 Tailored Microsomes in Cell-Free Systems

The complexity and variability of membrane proteins significantly hamper their expression and functional analysis. Inasmuch as membrane proteins are naturally present in low abundance in living cells, common in vivo overexpression systems often reach their limits caused by cytotoxic effects or aggregation of membrane
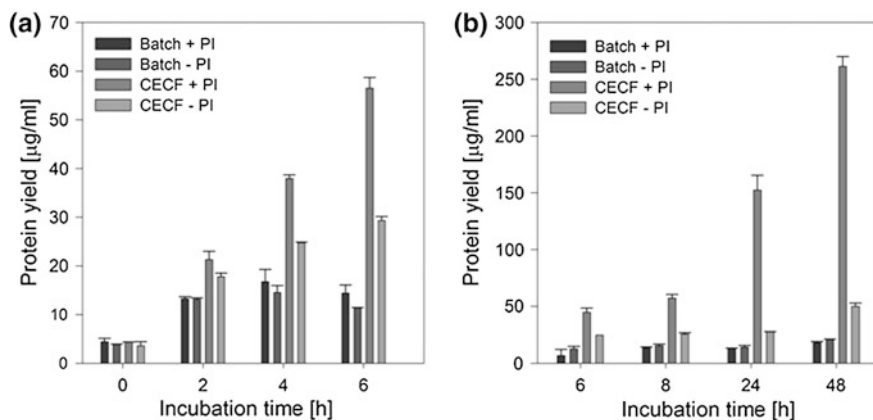
Fig. 6 Diagrams showing the increase of the ^{14}C-leucine-labeled protein SII-eYFP synthesized in a CECF-formatted cell-free translation reaction in comparison to a batch-formatted reaction. Coupled cell-free reactions using insect lysate were carried out in the presence (+) and in the absence (−) of protease inhibitors (PI), respectively. Aliquots of the total translation mixture were precipitated at the indicated time points and radiolabeled protein was quantified by liquid scintillation counting. Aliquots of the time points 0–6 h were taken from one reaction (**a**); aliquots from 6 to 48 h were taken from a second reaction (**b**)

proteins. Thus, the reduction of the biological protein production machinery to a viable, cell-independent, cell-free expression system seems to be advantageous. Nevertheless, membrane proteins require a complex and dynamic environment and in many cases versatile PTMs are essential to exhibit their full functionality. Typical approaches to mimic the amphiphilic environment of a membrane for protein expression in cell-free systems are either detergent- or lipid-based or apply liposomes as well as nanodiscs in different reaction modes [68].

A major advantage of the insect cell-free system is the expression of membrane proteins in the presence of microsomes. These vesicles enable the biological process of cotranslational protein integration into a lipid bilayer and PTMs such as glycosylation or disulfide bond formation. Thus, the system is capable of synthesizing a correctly folded, functional target membrane protein. In this section we focus on the potential of microsomes for technical implementation and their utilization for functional studies of membrane proteins.

To date, the expression of many functional and structural divergent membrane proteins in the insect cell-free system has been reported [11, 12, 26]. As examples, we present a repertory of integral membrane proteins ranging from type I membrane proteins (Hb-EGF and EGFR) across a six-transmembrane ion-channel (Aqp1) to a seven-transmembrane G-protein coupled receptor (ETB; Table 1). All proteins were fused to eYFP for CLSM analysis. Furthermore, the cleavable signal sequence of the type I membrane proteins was replaced with the melittin signal sequence to address the insect translocon efficiently and favor cotranslational integration into the membrane. Membrane proteins were synthesized in the presence of ^{14}C-leucine and subsequently analyzed by autoradiography.

Table 1 Overview of cell-free synthesized membrane proteins. Summarized are their molecular mass in kDa, number of transmembrane domains (TMD), function, and the UniProt Knowledgebase (*UniProtKB*) number. Data source: http://www.uniprot.org

Protein	Size [kDa]	TMD	Function	UniProtKB
Mel-Hb-EGF-eYFP				
Proheparin-binding EGF-like growth factor fused to a melittin signal sequence and eYFP	51	1	Growth factor	Q06175
Aqpl-eYFP				
Aquaporin-1 fused to eYFP	55	6	Water channel	P29972
ETB-eYFP				
EndothelinB receptor fused to eYFP	78	7	Endothelin B receptor	P24530
Mel-EGFR-eYFP				
Epidermal growth factor receptor fused to a melittin signal sequence and eYFP	162	1	Receptor tyrosine kinase	Q504U8

Fig. 7 Qualitative analysis of the cell-free expressed membrane proteins Mel-Hb-EGF-eYFP (51 kDa), Aqp1-eYFP (55 kDa), ETB-eYFP (78 kDa), and Mel-EGFR-eYFP (162 kDa) by SDS-PAGE and autoradiography
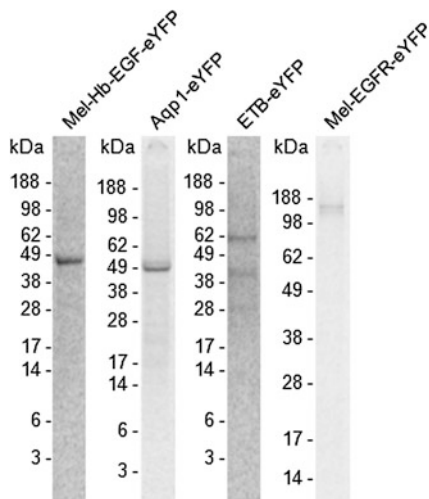
Compared to globular marker proteins, membrane proteins migrate with lower apparent molecular mass in sodium dodecyl sulfate (SDS) polyacrylamide gels during electrophoresis. This is caused by their hydrophobic transmembrane domains and their interaction with SDS [69]. However, in the autoradiograph shown in Fig. 7 distinct predominant protein bands are separated, indicating efficient protein expression and separation. The localization of the fluorescent membrane proteins was probed via CLSM analysis. As expected, the microsomes appear with a fluorescent membrane, surrounding a dark lumen (pinpointed by red arrows in Fig. 8) due to the insertion of the target membrane proteins.

Typically, the vesicles of the insect cell-free expression system show a diameter of 1–10 μm and tend to form irregular aggregates of much bigger size including undefined amounts of single microsomal vesicles (Fig. 8). This obstacle makes it difficult to integrate them into technical systems and apply them in functional studies. The giant unilamellar vesicle (GUV) formation from microsomes containing cell-free synthesized membrane anchored proteins is a suitable method to enlarge the size of vesicles [26].

The formation of well-defined synthetic membrane model systems to study elementary biological processes in a less complex system compared to a viable cell is widely used in the field of biophysics [70–73]. The formation of GUVs is one option to build such a reduced biomimetic system. Among other techniques, the rehydration of dried lipid films and voltage application are commonly used to grow vesicles up to 100 μm in diameter [74]. In addition to the general characterization of membranous structures, GUVs are commonly used for the integration of target membrane proteins and the analysis of the protein's functionality and interaction with other molecules. A major advantage of GUVs as a model system for functional studies is their size, allowing for microscopic investigations and physical manipulation, such as patch clamp measurements, or studies of optical signal generation, for instance, by fluorescence intensity changes during substrate metabolism.
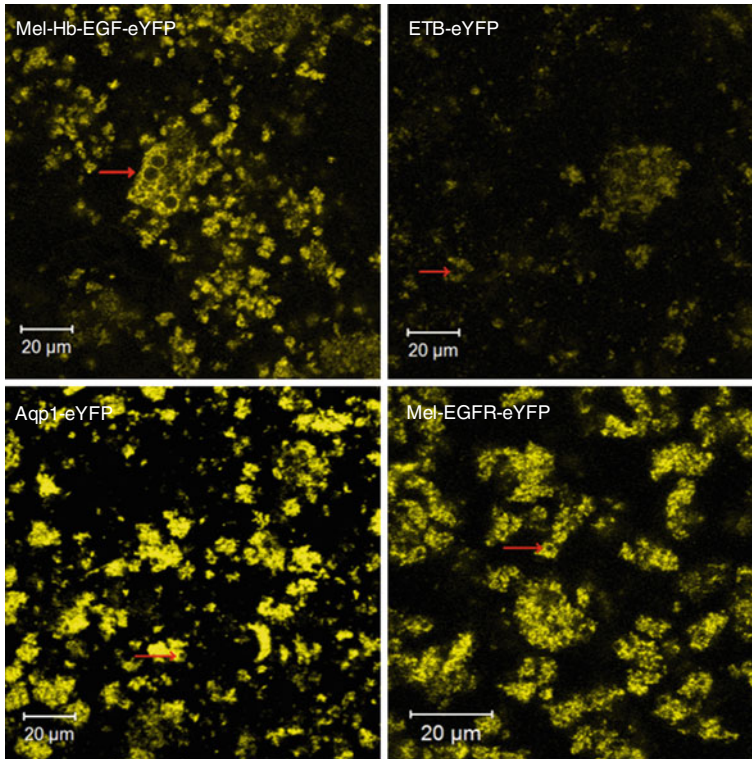
Fig. 8 CLSM images of cell-free synthesized transmembrane proteins Mel-Hb-EGF-eYFP, ETB-eYFP, Aqp1-eYFP, and Mel-EGFR-eYFP integrated into microsomal membranes. Samples were excited at 488 nm using an argon laser and fluorescence emission was recorded with a longpass filter in the wavelength range above 505 nm (LSM 510 Meta microscope, Zeiss). Red arrows indicate microsomal membranes formed in vesicular shape harboring fluorescent target proteins

Currently, several methods for the reconstitution of membrane proteins into GUVs have been established. However, the conventional dehydration and subsequent rehydration are derogatory for a membrane protein's functionality. In contrast, milder processes such as partial drying [26, 75] or sucrose-supported GUV formation [76] are preferred in order to preserve the protein's function. For controlled incorporation of membrane proteins, a peptide-induced fusion mechanism is described, requiring several preparation steps to integrate the target membrane protein into the GUV formation process [77]. To reduce the complexity of the target membrane protein integration, we focus on the combination of different mild approaches to gain GUVs from insect-based microsomes (schematically depicted in Fig. 9). While maintaining their functionality, the generated GUVs will incorporate membrane proteins via the biological process of cotranslational translocation in an active and directed manner. In this context, the
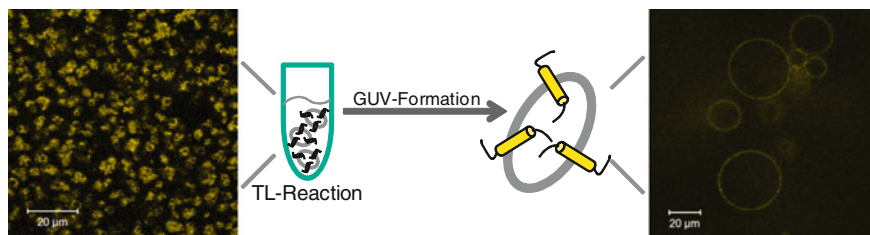
Fig. 9 From small microsomes to GUVs. Scheme of the GUV formation process for size enlargement of microsomes containing the cell-free synthesized membrane protein Mel-Hb-EGF-eYFP (CLSM image on the *left*) to gain target protein harboring GUVs (CLSM image on the *right*). Microsomes derived from the insect cell-free system were used in the formation process under physiological conditions. Samples were excited at 488 nm using an argon laser and fluorescence emission was recorded with a longpass filter in the wavelength range above 505 nm (LSM 510 Meta microscope, Zeiss)

combination of cell-free protein expression and the GUV formation process enables us not only to enrich the insect-based microsomes with target membrane proteins but also to tailor their size and composition according to the requirements for further downstream analyses.

5 A Cell-Free System Based on Extracts from Cultured CHO Cells

CHO cells are the most widely used mammalian host for recombinant protein production to date [78]. They are the preferred choice for the high-yield synthesis of biologically active proteins requiring proper folding and PTMs such as glycosylation [79]. This potential turns CHO cells into a valuable candidate for a productive in vitro protein expression system. We have addressed this fact and have developed a cell-free system based on extracts from CHO cells. The system combines transcription and translation in a single step (coupled system) to save the user's time and effort (Fig. 10). The preparation of a translationally active lysate, the construction of a suitable expression vector, and the optimization of the reaction conditions were in the scope of the investigations demonstrated below.

5.1 Lysate Preparation

The key element of cell-free protein production is the translationally active cell extract, including ribosomes, tRNAs, translation factors, and enzymes. Crude extracts were obtained from CHO cells according to a procedure previously described for insect cells [24]. CHO cells were grown in suspension at 37 °C to a final density of four million cells per milliliter. Cells were disrupted mechanically
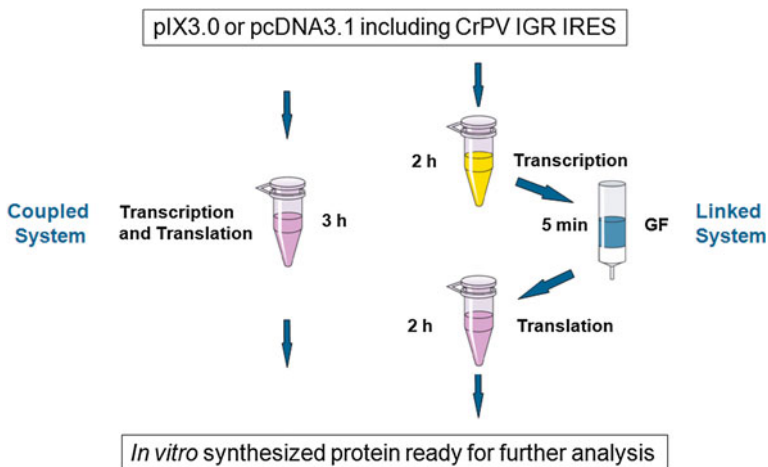
Fig. 10 Workflow of the cell-free protein synthesis reactions using the CHO cell-free system. The coupled system combines transcription and translation in a single batch reaction. Alternatively, protein expression can be performed in a linked system where transcription and translation reactions are separated by an intermediate gel filtration step

and cell extracts were further processed by removing nuclei and cell debris. Endogenous amino acids were removed by size-exclusion chromatography and residual mRNA was digested by S7-nuclease treatment. Translationally active CHO extracts were stored at −80 °C.

5.2 Construction of Efficient Expression Vectors

Translation initiation is one of the key challenges when it comes to the activation of a cell-free translation system. In eukaryotic cells, mRNA capping and poly-adenylation are essential modifications in order to activate translation initiation efficiently [80, 81]. In accordance, cell-free systems accept capped and poly-adenylated mRNA for protein synthesis. However, in prokaryotic and eukaryotic translation systems secondary and tertiary structures of mRNA are often the cause for inefficient translation initiation [82]. Recent advances have been made regarding the preparation of "universal" translation initiation sequences for pro-karyotic and eukaryotic cell-free translation systems [82, 83]. Another promising approach to yield highly flexible DNA templates is the utilization of internal ribosome entry sites (IRES). IRES sequences are structured mRNA elements reaching several tens to hundreds of nucleotides and allowing for direct binding to ribosomes [84]. Therefore IRESs function as direct initiators of protein translation thus neglecting the need for mRNA capping and polyadenylation.

Noncoding regions flanking the coding sequence of the gene of interest significantly affect the translation initiation rate and hence the efficiency of protein
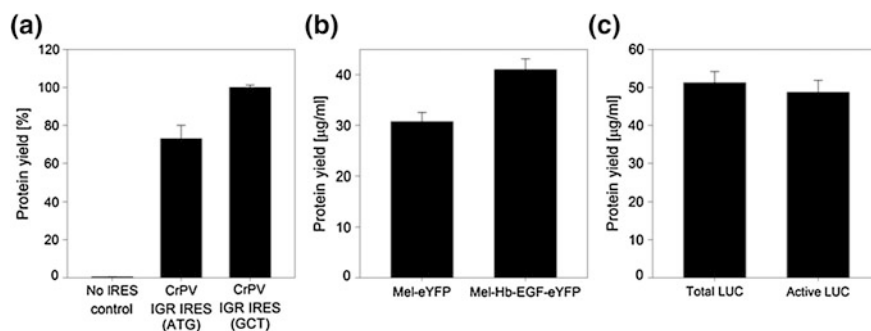
Fig. 11 Protein synthesis in a CHO cell-derived in vitro translation system. **a** Comparative analysis of different luciferase (LUC) encoding expression constructs. The amount of active LUC was determined using a LUC reporter assay. Protein yields were normalized to the reaction with the highest protein yield (=100 %), which was achieved using the optimized vector including the CrPV IGR IRES with an ATG-to-GCT mutation of the initiation codon in the EasyXpress pIX3.0 backbone. **b** Quantitative analysis of the ^{14}C-leucine-labeled proteins Mel-eYFP and Mel-Hb-EGF-eYFP [in pIX3.0 CrPV IGR IRES (GCT)] by liquid scintillation counting. **c** Comparison of total and active LUC. The amount of total LUC [in pIX3.0 CrPV IGR IRES (GCT)] was calculated after liquid scintillation counting, the amount of active LUC was determined by using a LUC reporter assay. Cell-free reactions were incubated at 33 °C for 3 h (**a**, **b**) and 4.5 h (**c**), respectively

production in cell-free systems. Most eukaryotic in vitro translation systems utilize IRES sequences for efficient translation of the protein of interest. In this way, costly capping of mRNA is circumvented and the inhibitory effect of free caps on the initiation factor eIF-4E is avoided. We analyzed specific IRESs from the intergenic region (IGR) of the *Dicistroviridae* family and we demonstrated their function in CHO extracts. IGR IRES-driven translation bypasses all steps of conventional translation initiation [85] that are known as major bottlenecks in cell-free protein expression [82, 86]. Cloning of the IGR IRES from cricket paralysis virus (CrPV) upstream of the firefly luciferase (LUC) encoding sequence in the EasyXpress pIX3.0 (Qiagen) vector backbone dramatically increased the expression level of active LUC (Fig. 11a). The majority of IGR IRESs display maximum efficiency in the context of a GCU (alanine) as the first translated codon [87]. Replacement of the initial ATG-codon to a GCT-codon resulted in an additional increase in the protein expression level (Fig. 11a). As T7 RNA polymerase is used for transcription, DNA constructs must contain a T7 promoter. The T7 promoter-driven vector EasyXpress pIX3.0 but also the pcDNA3.1 (Life Technologies) vector backbone, well known for protein production in vivo, facilitated cell-free protein synthesis when combined with the CrPV IGR IRES. Therefore the same pcDNA3.1 construct can be used for protein expression in vivo and in vitro without recloning. This enables convenient in vitro screening of suitable templates prior to their use in cultured and transfected CHO cells.

5.3 Optimization of Reaction Conditions

The optimization of reaction conditions is essential for obtaining an effective in vitro translation system. Basic parameters that have to be optimized are: energy supply (ATP and GTP), incubation temperature, reaction time, amino acid concentration, and ion concentrations (Mg^{2+} and K^+). The optimized cell-free CHO-based system led to protein yields of approximately 50 μg/ml active LUC within 3 h.

5.4 Performance of the CHO-Based Cell-Free System

Multiple proteins were synthesized in vitro to evaluate the performance of the cell-free system. Reactions were carried out for 3–4.5 h at 33 °C. Synthesized proteins were labeled with the radioactive isotope ^{14}C-leucine and their yield and quality of expression were monitored. The cytosolic protein LUC, the secreted protein Mel-eYFP as well as the integral membrane protein Mel-Hb-EGF-eYFP were investigated quantitatively and the productivity of the CHO-based cell-free system reached approximately 30–50 μg/ml of the target proteins (Fig. 11b, c). Data were complemented by qualitative analysis (Fig. 12). In this respect, the above-mentioned proteins (LUC, Mel-eYFP, and Mel-Hb-EGF-eYFP) as well as the cytosolic protein eYFP and the glycoprotein EPO were investigated using autoradiography. Full-length proteins were observed without any indication of proteolysis or fragmentation in all cases. Therefore, the described CHO cell-free system can be used to produce a broad range of target proteins including membrane proteins and glycoproteins. Cell-free synthesized EPO resulted in two bands on the

Fig. 12 Qualitative analysis of the cytosolic proteins luciferase (LUC, 61 kDa) and eYFP (27 kDa), the secreted protein Mel-eYFP (29 kDa), the membrane protein Mel-Hb-EGF-eYFP (51 kDa), and the glycoprotein EPO (21 kDa, nonglycosylated), expressed in the CHO cell-free system. ^{14}C-leucine-labeled, de novo synthesized proteins were visualized by autoradiography after gel electrophoresis (Typhoon TRIO+ Imager, GE Healthcare)
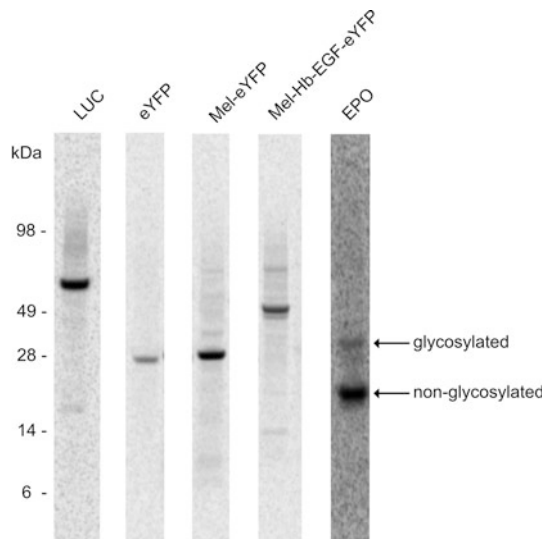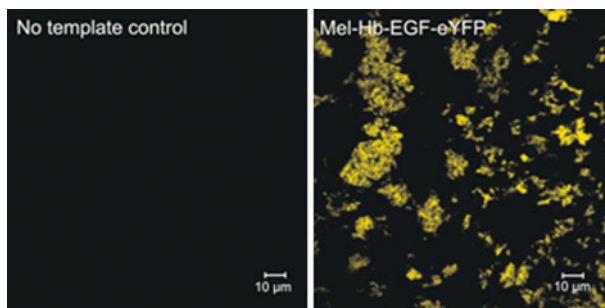
Fig. 13 Incorporation of a membrane protein into microsomal membranes in CHO-lysates. CLSM images depict de novo synthesized Mel-Hb-EGF-eYFP. Fluorescent vesicles indicate the translocation of the target protein into the microsomal vesicles present in the CHO cell-free system. eYFP was excited at 488 nm and fluorescence emission was recorded with a longpass filter in the wavelength range above 505 nm (LSM 510 Meta microscope, Zeiss)

autoradiograph (Fig. 12). The upper band (approximately 32 kDa) was identified by glycosidase digestion as the glycosylated protein, depicting the potential of the system to perform N-linked glycosylation. In eukaryotic cell-free systems, N-glycosylation is usually realized in the lumen of microsomes, originating from the rough ER [88]. In addition, these microsomes can be used to incorporate membrane proteins into a native like lipid environment suitable for further downstream functional studies. Membrane incorporation of cell-free expressed proteins was tested by studying the production of the type I transmembrane protein heparin-binding EGF-like growth factor [89] using the Mel-Hb-EGF-eYFP construct described previously in Sect. 4 . Fluorescent vesicles indicate the translocation and incorporation of the target protein into the microsomal vesicles present in the CHO cell-free system (Fig. 13). Translocation efficiency was not affected by the ATG-to-GCT mutation of the initiation codon. Cell-free production of transmembrane proteins represents the basis for further downstream characterization of their functionality.

We have demonstrated the development of a cell-free system based on extracts derived from cultured CHO cells. The system enables the production of a wide range of target proteins, including membrane proteins and glycoproteins. Synthesis cannot only be accomplished with specific in vitro protein expression vectors but also with in vivo constructs, such as pcDNA3.1. This provides the potential to conveniently optimize a template DNA in vitro prior to expression in vivo without recloning. Some conceivable applications of this method are the evaluation of gene-specific mutations, the evaluation of purification tags including their terminal position, codon optimization, and the analysis of signal peptides. In all cases vector backbones should be equipped with the CrPV IGR IRES which is essential for high-yield protein synthesis in vitro.

The broad range of applications turns the CHO cell-free system into a powerful alternative to already established systems. It offers a tool for the fast and easy

synthesis of recombinant proteins, not only for the time-saving evaluation of efficient expression constructs, but also for functional and structural studies.

6 Protein Modification in Cell-Free Systems Through Incorporation of Noncanonical Amino Acids

The manifold functions of proteins expressed in all known organisms can be traced back to a subset of only 20 + 2 primary building blocks, the standard or so-called proteinogenic amino acids. Nevertheless, fruitful efforts with a great impact on research as well as industrial applications have been made in order to expand this natural repertoire. For this purpose, amino acid analogues with a variety of different characteristics have been generated and employed as noncanonical building blocks in different in vivo and in vitro protein expression systems [37, 39]. Their applications range from mimicking natural physicochemical properties and PTMs to biophysical probing, cross-linking, and further site-directed modification through unique reactivity [38].

The utilization of novel primary building blocks in the context of cell-free protein synthesis is accompanied by several advantages compared to their use in vivo [2]. Cell-free protein synthesis systems are based on an open mode of reaction that enables a fast and easy manipulation of the synthesis conditions. To account for the requirements of an efficient and directed amino acid incorporation and corresponding production of modified proteins, the reaction can be directly supplemented with additional components in optimal concentrations including novel protein building blocks, tRNAs, enzymes, and ions.

6.1 Fidelity of Protein Synthesis

The high fidelity that leads to a correctly synthesized protein can be mainly accredited to a certain class of enzymes responsible for charging the right amino acid onto its cognate tRNAs [90]. These so-called aminoacyl-tRNA synthetases exhibit a high degree of specificity towards their amino acid and tRNA substrates. The resulting aminoacyl-tRNA then functions as a transfer molecule that delivers the appropriate amino acid to the ribosome. There, the amino acid is incorporated into the nascent amino acid chain in correspondence to the tRNA's anticodon recognizing its complementary triplet codon in the mRNA. To circumvent the quality control of tRNA aminoacylation by the corresponding endogenous synthetases, two different methodologies are widely employed that enable the "misacylation" of tRNAs with noncanonical amino acids. In the following their application in the context of eukaryotic cell-free protein synthesis is presented in more detail.

6.2 Random and Site-Directed Cotranslational Incorporation of Fluorescent Amino Acids into Proteins

One possibility for charging tRNAs with a desired amino acid is based on a semisynthetical methodology. In short, a dinucleotide corresponding to the 3'-end of the mature tRNA is connected to the desired amino acid in several steps and then subsequently ligated enzymatically to the tRNA's shortened 3'-end [36]. In the case of loading a tRNA with a noncognate amino acid, the resulting aminoacyl-tRNA is referred to as "misacylated." However, the tRNA to be misacylated must not be a substrate for endogenous synthetases. This could lead to deacylation of the misacylated tRNA and aminoacylation with an undesired standard amino acid. Consequently, an inefficient incorporation of the desired amino acid and, with regard to a site-directed protein modification, a heterogeneous product would be synthesized due to the incorporation of undesired standard amino acids at unwanted positions. These obstacles can be overcome by application of tRNAs derived from bacteria exhibiting significantly differing properties and therefore not cross-reacting with eukaryotic synthetases [91]. Depending on the misacylated tRNA's anticodon the incorporation of a noncanonical amino acid can be directed towards a desired codon.

By semisynthetically charging a tRNA that comprises an anticodon that decodes one of the 61 sense codons, a desired noncanonical amino acid can be randomly incorporated into the protein of interest inasmuch as the respective tRNA competes with the corresponding endogenous tRNA bearing the same anticodon for decoding the sense codon [92]. Here we exemplify this methodology by incorporation of a lysine derivative containing the fluorophore Bodipy TMR into proteins connected to the EGF-receptor signaling pathway. For this purpose we applied an engineered tRNA decoding UUC normally employed by the genetic code to incorporate phenylalanine (Fig. 14a–d). This amino acid derivative is readily compatible with the translational machinery and moreover does not influence the translocation of proteins into the insect microsomes which additionally enables its incorporation into secretory and membrane proteins (Fig. 14b, c). Depending on the ratio of engineered tRNA to the respective endogenous tRNA, the random incorporation leads to a protein with the noncanonical amino acid statistically incorporated at multiple sites.

Furthermore, a desired noncanonical amino acid can also be incorporated into the protein of interest in a site-directed manner (Fig. 14e, f). For this purpose, a suppressor tRNA is employed, containing an anticodon complementary to one of the stop codons normally functioning as a signal to terminate the protein's translation. The so-called amber stop codon UAG is the most widely used codon because its suppression is tolerated to a higher degree than the suppression of the other two stop codons (ochre, UAA; opal, UGA) in living organisms. Moreover, the amber stop codon exhibits higher suppression efficiencies [93, 94]. Application of such a suppressor tRNA is used to achieve a stoichiometric incorporation of one noncanonical amino acid per protein. Furthermore, the incorporation at a desired
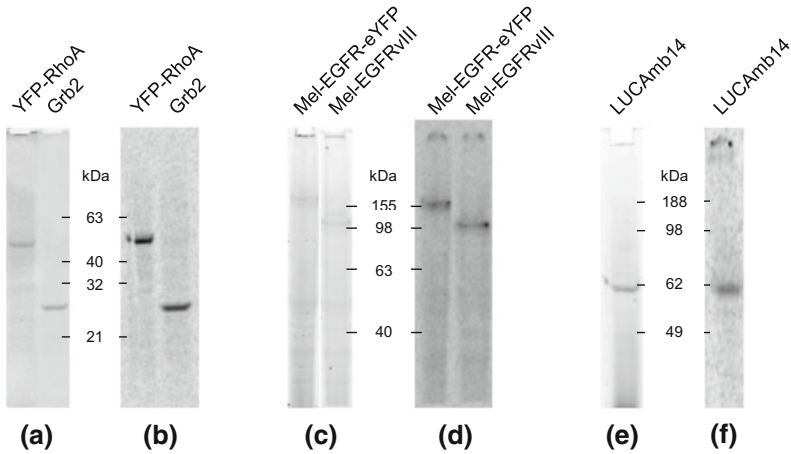
Fig. 14 Random and site-directed incorporation of BODIPY TMR-lysine into cell-free synthe-sized proteins. Semisynthetically aminoacylated tRNAs were probed in the insect cell-free system. Random labeling of soluble proteins (**a** and **b**; YFP-RhoA: Ras homologue gene family member A, N-terminally fused to YFP, 56 kDa; Grb2: Growth factor receptor bound protein 2, 24 kDa) and membrane proteins (**c** and **d**; Mel-EGFR-eYFP: epidermal growth factor receptor, N-terminally fused to a melittin signal sequence and C-terminally to eYFP, 163 kDa; Mel-EGFRvIII: Epidermal growth factor receptor deletion mutant vIII N-terminally fused to a melittin signal sequence, 105 kDa). Site-directed labeling of luciferase (LUC) containing an internal amber stop codon at the position corresponding to amino acid 14 (**e** and **f**; LUCAmb14, 61 kDa). De novo synthesized proteins were visualized by in-gel fluorescence (**a, c, e**, excitation 532 nm, emission filter 580 nm BP, Typhoon TRIO+ Imager, GE Healthcare) and autoradiography after cotranslational incorporation of ^{14}C-leucine (**b, d, f**) and separation in a SDS-PAGE

position is predefined by the location of the amber stop codon in the mRNA transcript. Hence, any given gene can be easily equipped with an amber codon by mutagenesis to allow for a site-directed incorporation of noncanonical amino acids in the resultant protein.

6.3 Orthogonal tRNA/Synthetase Pairs in the Cell-Free Insect System

Orthogonal tRNA/synthetase pairs are utilized to enzymatically charge a cognate tRNA with the desired amino acid. These orthogonal pairs can be considered an expansion of the translational machinery as they supplement the endogenous components and provide the utilization of an additional protein building block. Nevertheless, certain fundamental prerequisites have to be met in order to account for the efficient and site-directed incorporation of a noncanonical amino acid. The orthogonal synthetase has to exhibit a distinct specificity towards its cognate
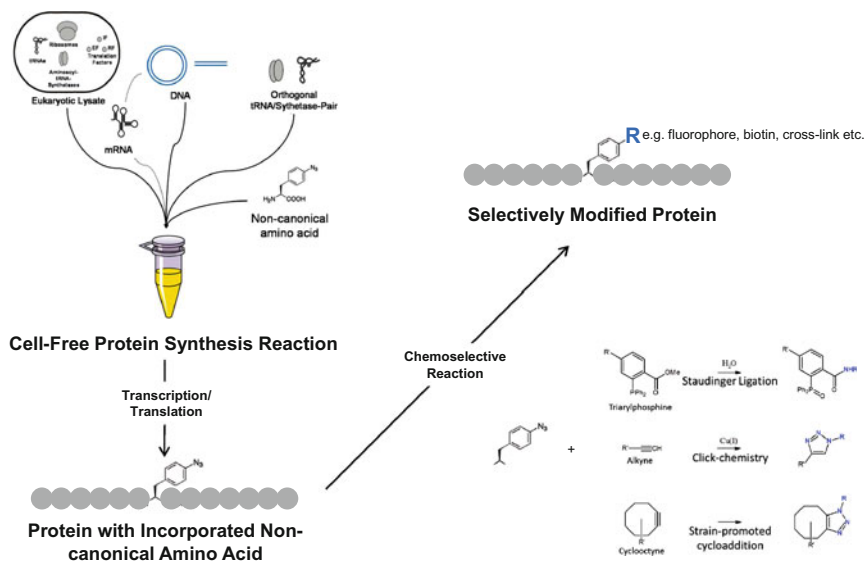
Fig. 15 Orthogonal cell-free protein synthesis system based on insect cell lysates. The cell-free protein synthesis system is complemented with in vitro synthesized orthogonal synthetase (AzPheRS), tRNA$_{CUA}$, and *p*-azido-*L*-phenylalanine (AzPhe). Protein synthesis can be carried out in a linked and a coupled reaction mode, respectively. De novo synthesized proteins harbor the noncanonical amino acid at positions predefined by the location of the amber codon in the respective gene template. Proteins with incorporated AzPhe can then be further modified based on the chemoselectivity of the azido group towards, for example, triarylphosphines and alkynes allowing for a site-directed introduction of desired characteristics into the cell-free synthesized proteins

noncanonical amino acid and the corresponding tRNA to account for a homogeneous product. Therefore, orthogonal tRNA/synthetase pairs are usually evolved by screening mutant libraries based on a tRNA/synthetase pair which occurs in one of the three domains of life and is therefore potentially orthogonal in one of the other two domains. For example, the *E. coli* tyrosyl-tRNA synthetase in combination with one of its cognate tRNAs (tRNA$_{CUA}$) has been found to be orthogonal in eukaryotes [95]. Moreover, the synthetase has been evolved in yeast to specifically aminoacylate a suppressor tRNA bearing an anticodon complementary to the amber stop codon with a variety of different noncanonical amino acids.

We employed a mutant tyrosyl-tRNA synthetase (AzPheRS) to incorporate *p*-azido-*L*-phenylalanine (AzPhe) into proteins in our cell-free insect system (Fig. 15). This noncanonical amino acid can be used as a chemical handle to further modify de novo synthesized proteins in a directed and bioorthogonal manner through the unique reactivity of the azido function. To assure the compatibility of the incorporated noncanonical amino acid with the translocation and the glycosylation machinery present in a functional state inside the insect microsomes, we expressed a mutant of the human glycoprotein EPO harboring an
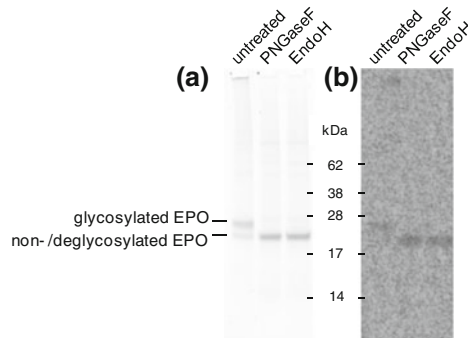
Fig. 16 Cotranslational incorporation of *p*-azido-*L*-phenylalanine (AzPhe) into the human glycoprotein erythropoietin (EPO, 21 kDa, nonglycosylated) and subsequent site-directed modification with fluorescent phosphine dye. A mutant EPO gene fused to a melittin signal sequence with an amber codon at the position corresponding to amino acid 23 was expressed in the cell-free insect lysate in the presence of AzPhe, orthogonal mutant synthetase (AzPheRS) and tRNA$_{CUA}$. The reaction was subsequently treated with DyLight650-phosphine for 2 h at 37 °C. Aliquots of the complete reaction mixture were applied to PNGaseF and EndoH digestion to confirm glycosylation. The fluorescently labeled glycosylated and deglycosylated EPO was visualized by in-gel fluorescence (**a**, excitation 633 nm, emission 670 nm BP, Typhoon TRIO+ Imager, GE Healthcare) and autoradiography after cotranslational incorporation of ^{14}C-leucine (**b**) and separation in a SDS-PAGE

internal amber codon in the coding sequence at the position corresponding to amino acid 23 (Fig. 16, untreated). Glycosylation of the suppression product and the corresponding cotranslational translocation of the protein into the insect microsomes were shown by digestion of the glycans with *N*-glycosidase *F* (Fig. 16, PNGaseF) and endoglycosidase H (Fig. 16, EndoH). To further demonstrate the successful incorporation of the noncanonical amino acid into the glycoprotein, the suppression reaction was treated with the commercially available fluorescent DyLight650-phosphine which has been described as chemoselectively reacting with azides via the so-called Staudinger ligation [96]. The modified glycoprotein was visualized "in-gel" after separation in a SDS-PAGE (Fig. 16a).

The magnesium concentration is a critical parameter in cell-free protein synthesis and strongly correlates with the nature of the template encoding the gene of interest [97]. Moreover, the magnesium concentration significantly influences the integrity of the synthesized protein, in particular, in the case of mutant genes with an internal amber stop codon (Fig. 17). To achieve optimal results, the influence of different magnesium concentrations on the protein's integrity as well as the yield of the suppression product was routinely analyzed.

The presented results demonstrate the potential of combining cotranslational noncanonical amino acid incorporation and cell-free protein synthesis covering a broad range of structural and functional divergent proteins. Our insect-based cell-free system enables the synthesis of secretory proteins, membrane proteins, and proteins that require PTMs. In combination with the site-directed incorporation of
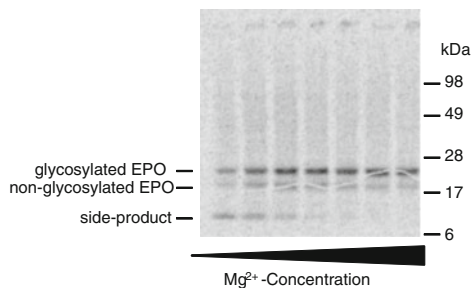
Fig. 17 Influence of magnesium concentration on the integrity of glycosylated erythropoietin (EPO) suppression product. A mutant EPO gene fused to a melittin signal sequence with an amber codon at the position corresponding to amino acid 23 was expressed using insect cell lysate in the presence of *p*-azido-*L*-phenylalanine (AzPhe), orthogonal mutant synthetase (AzPheRS), and tRNA$_{CUA}$. Reactions were supplemented with increasing amounts of Mg(OAc)$_2$. De novo synthesized proteins were analyzed by autoradiography (Typhoon TRIO+ Imager, GE Healthcare) after cotranslational ^{14}C-leucine incorporation and separation in a SDS-PAGE

noncanonical amino acids, the orthogonal cell-free system represents a versatile platform for functional in vitro studies as well as protein engineering.

7 Summary

Cell-free protein synthesis has become a valuable tool for the synthesis of many different classes of proteins such as membrane proteins [8–14, 98–100], post-translationally modified proteins [11, 26], and toxic proteins [101]. In particular, in the field of antibody engineering, cell-free protein synthesis offers several advantages as recombinant antibody formats can be produced in a time-saving and therefore cost-effective manner when using appropriate in vitro translation systems [25, 27, 51, 53, 56, 57]. In this context, it has been shown that novel cell-free systems based on translationally active eukaryotic cell lysates are well suited to enable the synthesis of functional and endotoxin-free antibody formats [25, 27]. The possibility of facilitating the synthesis of toxic proteins is one of the advantages of cell-free protein synthesis [101]. Therefore we anticipate that this method could become a valuable source for the future production of antibody–drug and antibody–toxin conjugates.

Cell-free systems are well suited for an automated and high-throughput production of recombinant proteins. The reasons behind this are manifold. First, linear DNA templates can be used directly as expression templates. A method to produce these templates in a fast and easy manner is the so-called expression-PCR (E-PCR) where the gene of interest is amplified and regulatory elements needed for an efficient production of the target protein are added [102]. The use of PCR amplified DNA templates avoids the need for time- and labor-intensive cloning

steps. In addition, these templates can be used in prokaryotic as well as in eukaryotic cell-free translation systems if the templates contain a Shine–Dalgarno sequence [11]. Second, in contrast to living cells that are naturally surrounded by a protective cell membrane, cell extracts are completely accessible and thus can be modified. Given this prerequisite, it seems to be very likely that in the near future the toolbox of in vitro translation systems can be further expanded and evolved in order to facilitate the expression of any desired protein. Third, cell-free reactions are scalable and range from the microliter to liter scale [7]. In particular, miniaturization is a promising strategy in order to develop microchips that can be analyzed in an automated environment, whereas scaling-up of cell-free production systems to the liter scale is the basis for large-scale production of pharmacologically relevant proteins. Fourth, cell-free synthesized proteins can be modified co- and posttranslationally. Modifications of proteins may include the introduction of fluorescent labels, isotope-labeled amino acids, and/or chemical moieties. The production of chemically modified proteins may then simplify and accelerate screening procedures. In particular, in the field of antibody engineering, fluorescently labeled antibodies are a valuable prerequisite to identifying novel binders against desired targets.

Although the research field of cell-free protein synthesis has made rapid progress in the last decade, some challenges remain to be overcome. Currently, there is no universal in vitro translation system that enables the synthesis of any desired target protein in a correctly folded and active manner [17]. In particular, the synthesis of posttranslationally modified and functional eukaryotic proteins is still a major obstacle. In this context it is highly desirable to develop in vitro translation systems that facilitate the synthesis of human proteins exhibiting human glycosylation patterns. When it comes to the comparison of total protein yields, prokaryotic systems still outperform their eukaryotic relatives [2]. However, the research field of cell-free protein synthesis is developing rapidly and many of these challenges have already been addressed. We anticipate that eukaryotic lysates based on insect, CHO, and human cells have the potential to solve these problems, in particular, if these eukaryotic systems contain microsomal fractions derived from the ER or Golgi apparatus to ensure appropriate PTM formation. Having these translationally active lysates on hand, they can be easily adjusted to different reaction formats (e.g. CECF systems) in order to increase protein yields significantly, thus combining high production rates with the potential to produce posttranslationally modified proteins. In eukaryotic lysates translocationally active membranous vesicles serve as functional modules representing the prerequisite for PTMs such as glycosylation as well as membrane protein insertion, protein entrapment, and enrichment. Moreover, the possibility of modifying and manipulating these functional modules allows their implementation in biochip-based technical systems for further downstream procedures.

Cell-free systems are derived from living cells and represent minimal biological systems. Several approaches have shown that biological systems can be further developed and modified in order to understand and utilize but also expand nature′s facilities [103]. The method of cell-free protein synthesis is a convenient tool in

order to study the dynamics of complex biological protein translation systems. Due to their potential to synthesize difficult-to-express proteins, in vitro translation systems are well suited to enlarge the repertoire of proteins that can be addressed in pharmacological characterizations. In particular, the analysis of membrane protein complexes that were embedded in microsomes by cotranslational translocation and the assembly of individual membrane-spanning proteins in cell-free systems are of increasing interest. All things considered, cell-free translation systems provide a promising basis in order to analyze the interaction landscape of membrane protein complexes as well as their participation in cellular signaling pathways. Furthermore, in vitro translation systems enable the implementation of an enzyme-mediated protein modification machinery introducing nonnatural structural elements into protein sequences such as fluorescent and chemically modified moieties as well as biophysical labels. In particular, the introduction of site-specific modifications into proteins may have a great impact on the investigation of biological questions regarding cellular signal transduction. Taken together, our approach provides a versatile and flexible platform to synthesize proteins with modified and enhanced functionality, thus expanding the natural repertoire of protein species.

References

1. Nirenberg M, Matthaei J (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. Proc Natl Acad Sci U S A 47:1588–1602
2. Katzen F, Chang G, Kudlicki W (2005) The past, present and future of cell-free protein synthesis. Trends Biotechnol 23:150–156
3. Swartz J (2006) Developing cell-free biology for industrial applications. J Ind Microbiol Biotechnol 33:476–485
4. Goerke AR, Swartz JR (2008) Development of cell-free protein synthesis platforms for disulfide bonded proteins. Biotechnol Bioeng 99:351–367

5. Kanter G, Yang J, Voloshin A, Levy S, Swartz J, Levy R (2007) Cell-free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. Blood 109:3393–3399

6. Yang J, Kanter G, Voloshin A, Michel-Reydellet N, Velkeen H, Levy R, Swartz JR (2005) Rapid expression of vaccine proteins for B-cell lymphoma in a cell-free system. Biotechnol Bioeng 89:503–511

7. Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ (2011) Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. Biotechnol Bioeng 108:1570–1578

8. Gourdon P, Alfredsson A, Pedersen A, Malmerberg E, Nyblom M, Widell M, Berntsson R, Pinhassi J, Braiman M, Hansson Ö, Bonander N, Karlsson G, Neutze R (2008) Optimized in vitro and in vivo expression of proteorhodopsin: a seven-transmembrane proton pump. Protein Expr Purif 58:103–113

9. Kalmbach R, Chizhov I, Schumacher MC, Friedrich T, Bamberg E, Engelhard M (2007) Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. J Mol Biol 371:639–648

10. Klammt C, Löhr F, Schäfer B, Haase W, Dötsch V, Rüterjans H, Glaubitz C, Bernhard F (2004) High level cell-free expression and specific labeling of integral membrane proteins. Eur J Biochem 271:568–580

11. Kubick S, Gerrits M, Merk H, Stiege W, Erdmann VA (2009) In vitro synthesis of posttranslationally modified membrane proteins. In: "Membrane Protein Crystallization" current topics in membranes, Chapter 2, vol 63. Elsevier, Burlington

12. Sachse R, Wüstenhagen D, Šamalíková M, Gerrits M, Bier FF, Kubick S (2012) Synthesis of membrane proteins in eukaryotic cell-free systems. Eng Life Sci 12:1–10

13. Savage DF, Anderson CL, Robles-Colmenares Y, Newby ZE, Stroud RM (2007) Cell-free complements in vivo expression of the *E. coli* membrane proteome. Protein Sci 16:966–976

14. von Groll U, Kubick S, Merk H, Stiege W, Schäfer F (2007) Advances in insect-based cell-free protein expression. In: Kudlicki W, Katzen F, Bennett P (eds) Cell-free protein expression. Landes Bioscience, Austin

15. Goshima N, Kawamura Y, Fukumoto A, Miura A, Honma R, Satoh R, Wakamatsu A, Yamamoto J, Kimura K, Nishikawa T, Andoh T, Iida Y, Ishikawa K, Ito E, Kagawa N, Kaminaga C, Kanehori K, Kawakami B, Kenmochi K, Kimura R, Kobayashi M, Kuroita T, Kuwayama H, Maruyama Y, Matsuo K, Minami K, Mitsubori M, Mori M, Morishita R, Murase A, Nishikawa A, Nishikawa S, Okamoto T, Sakagami N, Sakamoto Y, Sasaki Y, Seki T, Sono S, Sugiyama A, Sumiya T, Takayama T, Takayama Y, Takeda H, Togashi T, Yahata K, Yamada H, Yanagisawa Y, Endo Y, Imamoto F, Kisu Y, Tanaka S, Isogai T, Imai J, Watanabe S, Nomura N (2008) Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat Methods 5:1011–1017

16. Griffiths AD, Tawfik DS (2003) Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. Embo J 22:24–35

17. Carlson ED, Gan R, Hodgman CE, Jewett MC (2012) Cell-free protein synthesis: applications come of age. Biotechnol Adv 30:1185–1194

18. Schwarz D, Klammt C, Koglin A, Löhr F, Schneider B, Dötsch V, Bernhard F (2007) Preparative scale cell-free expression systems: new tools for the large scale preparation of integral membrane proteins for functional and structural studies. Methods 41:355–369

19. Chang H-C, Kaiser CM, Hartl FU, Barral JM (2005) De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. J Mol Biol 353:397–409

20. Hillebrecht JR, Chong S (2008) A comparative study of protein synthesis in in vitro systems: from the prokaryotic reconstituted to the eukaryotic extract-based. BMC Biotechnol 8:1790–1793

21. Mann M, Jensen ON (2003) Proteomic analysis of post-translational modifications. Nat Biotech 21:255–261

22. Blobel G, Dobberstein B (1975) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol 67:835–851
23. Walter P, Blobel G (1983) Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol 96:84–93
24. Kubick S, Schacherl J, Fleischer-Notter H, Royall E, Roberts LO, Stiege W (2003) In vitro translation in an insect-based cell-free system. In: Swartz JR (ed) Cell-free protein expression. Springer, Berlin, pp 209–217
25. Merk H, Gless C, Maertens B, Gerrits M, Stiege W (2012) Cell-free synthesis of functional and endotoxin-free antibody Fab fragments by translocation into microsomes. Biotechniques 53:153–160
26. Shaklee PM, Semrau S, Malkus M, Kubick S, Dogterom M, Schmidt T (2010) Protein incorporation in giant lipid vesicles under physiological conditions. Chembiochem 11:175–179
27. Stech M, Merk H, Schenk J, Stöcklein W, Wüstenhagen D, Micheel B, Duschl C, Bier F, Kubick S (2012) Production of functional antibody fragments in a vesicle-based eukaryotic cell-free translation system. J Biotechnol 164:220–231
28. Brödel AK, Raymond JA, Duman JG, Bier FF, Kubick S (2013) Functional evaluation of candidate ice structuring proteins using cell-free expression systems. J Biotechnol 163:301–310
29. Khnouf R, Olivero D, Jin S, Coleman MA, Fan ZH (2010) Cell-Free expression of soluble and membrane proteins in an array device for drug screening. Anal Chem 82:7021–7026
30. Casteleijn MG, Urtti A, Sarkhel S (2012) Expression without boundaries: cell-free protein synthesis in pharmaceutical research. Int J Pharm 440(1):39–47
31. Zanders ED, Kai L, Roos C, Haberstock S, Proverbio D, Ma Y, Junge F, Karbyshev M, Dötsch V, Bernhard F (2012) Systems for the cell-free synthesis of proteins, chemical genomics and proteomics. Humana Press, New York, pp 201–225
32. Klammt C, Schwarz D, Fendler K, Haase W, Dötsch V, Bernhard F (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. FEBS J 272:6024–6038
33. Wuu JJ, Swartz JR (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. Biochim Biophys Acta 1778:1237–1250
34. Doyle S, Cappuccio J, Hinz A, Kuhn E, Fletcher J, Arroyo E, Henderson P, Blanchette C, Walsworth V, Corzett M, Law R, Pesavento J, Segelke B, Sulchek T, Chromy B, Katzen F, Peterson T, Bench G, Kudlicki W, Hoeprich P Jr, Coleman M (2009) Cell-free expression for nanolipoprotein particles: building a high-throughput membrane protein solubility platform, high throughput protein expression and purification, Humana Press, Totowa, pp 273–295
35. Goren MA, Nozawa A, Makino S, Wrobel RL, Fox BG (2009) Cell-free translation of integral membrane proteins into unilamelar liposomes. Methods Enzymol 463:647–673
36. Budisa N (2004) Prolegomena to future experimental efforts on genetic code engineering by expanding its amino acid repertoire. ChemInform 43:6426–6463
37. Wang L, Xie J, Schultz PG (2006) Expanding the genetic code. Annu Rev Biophys Biomol Struct 35:225–249
38. Xie J, Schultz PG (2006) A chemical toolkit for proteins—an expanded genetic code. Nat Rev Mol Cell Biol 7:775–782
39. Wang Q, Parrish AR, Wang L (2009) Expanding the genetic code for biological studies. Chem Biol 16:323–336
40. Hudson P, Souriau C (2003) Engineered antibodies. Nat Med 9:129–134
41. Goel A, Baranowska-Kortylewicz J, Hinrichs SH, Wisecarver J, Pavlinkova G, Augustine S, Colcher D, Booth BJM, Batra SK (2001) 99mTc-labeled divalent and tetravalent CC49 single-chain Fv's: novel imaging agents for rapid in vivo localization of human colon carcinoma. J Nucl Medicine 42:1519–1527

42. Adams GP, Schier R, McCall AM, Simmons HH, Horak EM, Alpaugh RK, Marks JD, Weiner LM (2001) High affinity restricts the localization and tumor penetration of single-chain Fv antibody molecules. Cancer Res 61(12):4750–4755

43. Souriau C, Hudson PJ (2001) Recombinant antibodies for cancer diagnosis and therapy. Expert Opin Biol Ther 1:845–855

44. Carter P (2001) Improving the efficacy of antibody-based cancer therapies. Nat Rev Cancer 1:118–129

45. Todorovska A, Roovers RC, Dolezal O, Kortt AA, Hoogenboom HR, Hudson PJ (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting. J Immunol Methods 248:47–66

46. Alzari P, Lascombe M, Poljak R (1988) Three-dimensional structure of antibodies. Annu Rev Immunol 6:555–580

47. Davies D, Padlan E, Sheriff S (1990) Antibody-antigen complexes. Annu Rev Biochem 59:439–473

48. Glockshuber R, Schmidt T, Plückthun A (1992) The disulfide bonds in antibody variable domains: effects on stability, folding in vitro, and functional expression in *Escherichia coli*. Biochemistry 31:1270–1279

49. Goto Y, Hamaguchi K (1979) The role of the intrachain disulfide bond in the conformation and stability of the constant fragment of the immunoglobulin light chain. J Biochem 86:1433–1441

50. Bundy BC, Swartz JR (2011) Efficient disulfide bond formation in virus-like particles. J Biotechnol 154:230–239

51. Jiang X, Ookubo Y, Fujii I, Nakano H, Yamane T (2002) Expression of Fab fragment of catalytic antibody 6D9 in an *Escherichia coli* in vitro coupled transcription/translation system. FEBS Lett 514:290–294

52. Kim D-M, Swartz JR (2004) Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of *Escherichia coli*. Biotechnol Bioeng 85:122–129

53. Merk H, Stiege W, Tsumoto K, Kumagai I, Erdmann VA (1999) Cell-free expression of two single-chain monoclonal antibodies against lysozyme: effect of domain arrangement on the expression. J Biochem 125:328–333

54. Oh I-S, Kim D-M, Kim T-W, Park C-G, Choi C-Y (2006) Providing an oxidizing environment for the cell-free expression of disulfide-containing proteins by exhausting the reducing activity of *Escherichia coli* S30 extract. Biotechnol Prog 22:1225–1228

55. Oh I-S, Lee J-C, Lee M-s, Chung J-h, Kim D-M (2010) Cell-free production of functional antibody fragments. Bioprocess Biosyst Eng 33:127–132

56. Ryabova LA, Desplancq D, Spirin AS, Plückthun A (1997) Functional antibody production using cell-free translation: effects of protein disulfide isomerase and chaperones. Nat Biotechnol 15:79–84

57. Kawasaki T, Gouda MD, Sawasaki T, Takai K, Endo Y (2003) Efficient synthesis of a disulfide-containing protein through a batch cell-free system from wheat germ. Eur J Biochem 270:4780–4786

58. Ezure T, Suzuki T, Shikata M, Ito M, Ando E, Nishimura O, Tsunasawa S (2007) Expression of proteins containing disulfide bonds in an insect cell-free system and confirmation of their arrangements by MALDI-TOF MS. Proteomics 24:4424–4434

59. Yin G, Swartz JR (2004) Enhancing multiple disulfide bonded protein folding in a cell-free system. Biotechnol Bioeng 86:188–195

60. Rapoport TA (2007) Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450:663–669

61. Bromley EHC, Channon K, Moutevelis E, Woolfson DN (2008) Peptide and protein building blocks for synthetic biology: from programming biomolecules to self-organized biomolecular systems. ACS Chem Biol 3:38–50

62. Tessier DC, Thomas DY, Khouri HE, Laliberié F, Vernet T (1991) Enhanced secretion from insect cells of a foreign protein fused to the honeybee melittin signal peptide. Gene 98:177–183

63. Spirin A, Baranov V, Ryabova L, Ovodov S, Alakhov Y (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242:1162–1164
64. Alakhov J, Baranov V, Ovodov S, Ryabova L, Spirin A, Morozov I (1995) Method of preparing polypeptides in cell-free translation system. Institute of Protein Research, Moscow
65. Kim D, Choi C (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. Biotechnol Prog 12:645–649
66. Kigawa T, Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. J Biochem 110:166–168
67. Grandi G, Shirokov VA, Kommer A, Kolb VA, Spirin AS (2007) Continuous-exchange protein-synthesizing systems, In vitro transcription and translation protocols. Humana Press, pp 19–55
68. Junge F, Haberstock S, Roos C, Stefer S, Proverbio D, Dötsch V, Bernhard F (2011) Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. N Biotechnol 28:262–271
69. Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc Natl Acad Sci U S A 106:1760–1765
70. Fenz SF, Sengupta K (2012) Giant vesicles as cell models. Integr Biol 4:982–995
71. Liu A, Fletcher D (2009) Biology under construction: in vitro reconstitution of cellular function. Nat Rev Mol Cell Biol 10:644–650
72. Schwille P, Diez S (2009) Synthetic biology of minimal systems. Crit Rev Biochem Mol Biol 44:223–242
73. Smith A-S (2010) Physics challenged by cells. Nat Phys 6:726–729
74. Walde P, Cosentino K, Engel H, Stano P (2010) Giant vesicles: preparations and applications. Chembiochem 11:848–865
75. Girard P, Pecreaux J, Lenoir G, Falson P, Rigaud JL, Bassereau P (2004) A new method for the reconstitution of membrane proteins into giant unilamellar vesicles. Biophys J 87:419–429
76. Doeven MK, Folgering JH, Krasnikov V, Geertsma ER, van den Bogaart G, Poolman B (2005) Distribution, lateral mobility and function of membrane proteins incorporated into giant unilamellar vesicles. Biophys J 88:1134–1142
77. Kahya N, Pécheur E-I, de Boeij WP, Wiersma DA, Hoekstra D (2001) Reconstitution of membrane proteins into giant unilamellar vesicles via peptide-induced fusion. Biophys J 81:1464–1474
78. Kim J, Kim Y-G, Lee G (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 93:917–930
79. Jayapal K, Wlaschin K, Hu W, Yap M (2007) Recombinant protein therapeutics from CHO cells—20 years and counting. Chem Eng Prog 103:40–47
80. Gebauer F, Hentze MW (2004) Molecular mechanisms of translational control. Nat Rev Mol Cell Biol 5:827–835
81. Hershey J, Merrick WC (2000) The initiation pathway and mechanism of initiation of protein synthesis. In: Sonenberg N, Hershey JWB, Mathews MB (eds) Translational control of gene expression. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp 33–88
82. Swartz JR (2009) Universal cell-free protein synthesis. Nat Biotech 27:731–732
83. Mureev S, Kovtun O, Nguyen UTT, Alexandrov K (2009) Species-independent translational leaders facilitate cell-free expression. Nat Biotech 27:747–752
84. Fitzgerald KD, Semler BL (2009) Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. Biochim Biophys Acta 1789:518–528
85. Wilson JE, Pestova TV, Hellen CUT, Sarnow P (2000) Initiation of protein synthesis from the a site of the ribosome. Cell 102:511–520
86. Zeenko VV, Wang C, Majumder M, Komar AA, Snider MD, Merrick WC, Kaufman RJ, Hatzoglou M (2008) An efficient in vitro translation system from mammalian cells lacking the translational inhibition caused by eIF2 phosphorylation. RNA 14:593–602

87. Hertz MI, Thompson SR (2011) In vivo functional analysis of the *Dicistroviridae* intergenic region internal ribosome entry sites. Nucleic Acids Res 39(16):7276–7288

88. Blobel G, Dobberstein B (1975) Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. J Cell Biol 67:852–862

89. Higashiyama S, Abraham J, Miller J, Fiddes J, Klagsbrun M (1991) A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF. Science 251:936–939

90. Ibba M, Soll D (2000) Aminoacyl-tRNA synthesis. Annu Rev Biochem 69:617–650

91. Kwok Y, Wong JT-F (1980) Evolutionary relationship between *Halobacterium cutirubrum* and eukaryotes determined by use of aminoacyl-tRNA synthetases as phylogenetic probes. Can J Biochem 58:213–218

92. Kudlicki WA, Katzen F, Bennett RP (2007) Cell-free protein expression. Landes Bioscience, Austin, Tex

93. Gerrits M (2010) Funktion und Effizienz von *amber*-Suppressor-tRNAs in der zellfreien Proteinbiosynthese, Fachbereich Biologie, Chemie, Pharmazie der Freien Universität Berlin, Freie Universität Berlin, Berlin

94. Noren CJ, Anthony-Cahill SJ, Noren KA, Griffith MC, Schultz PG (1990) In vitro suppression of an amber mutation by a chemically aminoacylated transfer RNA prepared by runoff transcription. Nucleic Acids Res 18:83–88

95. Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG (2003) An expanded eukaryotic genetic code. Science 301:964–967

96. Saxon E, Bertozzi CR (2000) Cell surface engineering by a modified Staudinger reaction. Science 287:2007–2010

97. Swartz JR (2003) Cell-free protein expression. Springer, New York

98. Miot M, Betton J-M (2010) Reconstitution of the Cpx signaling system from cell-free synthesized proteins. New Biotechnol 28:277–281

99. Katzen F, Fletcher JE, Yang J-P, Kang D, Peterson TC, Cappuccio JA, Blanchette CD, Sulchek T, Chromy BA, Hoeprich PD, Coleman MA, Kudlicki W (2008) Insertion of membrane proteins into discoidal membranes using a cell-Free protein expression approach. J Proteome Res 7:3535–3542

100. Katzen F, Peterson TC, Kudlicki W (2009) Membrane protein expression: no cells required. Trends Biotechnol 27:455–460

101. Orth JHC, Schorch B, Boundy S, Ffrench-Constant R, Kubick S, Aktories K (2011) Cell-free synthesis and characterization of a novel cytotoxic pierisin-like protein from the cabbage butterfly *Pieris rapae*. Toxicon 57:199–207

102. Merk H, Meschkat D, Stiege W (2003) Expression-PCR: from gene pools to purified proteins within 1 day. In: Swartz JR (ed) Cell-free protein expression. Springer, Berlin, pp 15–23

103. Lee K, Lee K, Park J, Kim D (2012) Flexible programming of cell-free protein synthesis using magnetic bead-immobilized plasmids. PLoS One 7(3):e34429

New Bioproduction Systems: From Molecular Circuits to Novel Reactor Concepts in Cell-Free Biotechnology

Steffen Rupp

Abstract The last decades witnessed a strong growth in several areas of biotechnology, especially in fields related to health, as well as in industrial biotechnology. Advances in molecular engineering now enable biotechnologists to design more efficient pathways in order to convert a larger spectrum of renewable resources into industrially used biofuels and chemicals as well as into new pharmaceuticals and therapeutic proteins. In addition material sciences advanced significantly making it more and more possible to integrate biology and engineering. One of the key questions currently is how to develop new ways of engineering biological systems to cope with the complexity and limitations given by the cell. The options to integrate biology with classical engineering advanced cell free technologies in the recent years significantly. Cell free protein production using cellular extracts is now a well-established universal technology for production of proteins derived from many organisms even at the milligram scale. Among other applications it has the potential to supply the demand for a multitude of enzymes and enzyme variants facilitating in vitro metabolic engineering. This review will briefly address the recent achievements and limitations of cell free conversions. Especially, the requirements for reactor systems in cell free biotechnology, a currently underdeveloped field, are reviewed and some perspectives are given on how material sciences and biotechnology might be able to advance these new developments in the future.

Keywords Molecular circuits · Cell-free conversion · Cell-free reactor · Energy regeneration · ATP-synthase

S. Rupp (✉)
Department of Molecular Biotechnology, Fraunhofer IGB, Nobelstr. 12,
70569 Stuttgart Germany
e-mail: rupp@igb.fraunhofer.de

Contents

1 Introduction

The last decades witnessed strong growth in several areas of biotechnology. In the last decades of the twentieth century red biotechnology started its rise with remarkable success especially in the field of therapeutic proteins. Within this century the growing awareness of the need to change from fossil to sustainable resources induced huge efforts in industrial biotechnology with the goal of securing on a long-term perspective the supply of fuels as well as the supply of basic compounds for the chemical industry. Well-established bioprocesses have been expanded, including bioethanol production with efforts to shift from first-generation resources such as sugars to second-generation resources such as lignocellulosics. Furthermore advances in molecular engineering now enable biotechnologists to design more efficient metabolic pathways in order to convert a larger spectrum of renewable resources into already industrially used biofuels or chemicals as well as new pharmaceuticals and therapeutic proteins [1–3]. Currently operating processes are in general large-scale fermentation processes both in pharmaceutical and industrial biotechnology. Especially for industrial biotechnology, only some processes leading to valuable molecules are able to compete with the petro-based product. Usually these are well-known biotechnological processes such as ethanol, citric acid, or vitamin production to name a few.

The development of novel bioprocesses, for example, employing reprogrammed biological systems for production of artemisinin in yeast or 1,3-propanediol in *Escherichia coli*, took more than a decade and involved hundreds of person-years with the respective associated cost [1, 4]. In addition, for many fermentation processes already working in the lab a transfer to industrial scale currently is not economically feasible. One of the drawbacks of fermentation processes is that in general most of the energy supplied in the feed as sugars or fatty acids is used for cell proliferation and keeping the cell alive and not for

production of the desired compound. Therefore, a whole metabolism has to be supported which in general includes the production of a whole array of side-products. Some of these compounds may cause significant problems in downstream processing (DSP), often impairing the establishment of a fermentation process on an industrial scale or increasing the cost of the product substantially. In addition, depending on the underlying pathway, fermentation can be time consuming, resulting in a lower space–time yield. Also microorganisms usually have a limited capacity to take up or release compounds both with regard to quantity and specificity. The membranes of the microorganisms act as barriers both for substrates and products, thereby preventing efficient conversion unless efficient transport mechanisms are present in the cell. One of the key questions therefore is how to develop new ways of engineering biological systems to cope with the complexity and limitations given by the cell. Several key developments in recent decades allow us to address these questions in a direct and comprehensive way. The ability to analyze transcriptional networks on a pan-genomic level using next-generation sequencing technologies as well as the ability to synthesize DNA at a genome level plays a major role in accelerating the development of novel biotechnological production systems or making them possible at all. These developments resulted in completely new fields in life science and accelerated others tremendously, including systems biology and synthetic biology. In parallel, several concepts for cell-free biosynthesis have been developed. Cell-free protein production using cellular extracts is now a well-established universal technology for the production of proteins derived from many organisms, even on a milligram scale. It has the potential to supply the demand for a multitude of enzymes and enzyme variants also for in vitro metabolic engineering.

This review briefly addresses the recent achievements and limitations of cell-free conversions. Recent developments in design and modeling of molecular circuits are especially addressed. The options cell-free biotechnology offers for the realization of synthetic molecular circuits using purified compounds as well as their implementation in a complex cell-free context, making use of existing protein synthesis and regulation machinery are addressed. Furthermore, the requirements of reactor systems for cell-free biotechnology to be used, a currently underdeveloped field, are reviewed and some perspectives are given on how material sciences and biotechnology might be able to develop new reactor concepts in the future.

2 Cell-Free Conversion: Achievements and Limitations

Cell-free biosynthesis using enzymes for conversion of various substrates in pharmaceuticals, food, and feed as well as industrial biotechnology actually has a long history [5]. Due to the availability of only one catalyst present in the reaction in general pure products without side-products are obtained. Also barriers given by membranes in fermentation processes do not exist. Thus single enzymatic conversion steps are also very common in large-scale industrial applications. The

origins of cell-free conversions can be dated back to the nineteenth century. Eduard Buchner showed in 1897 that yeast extracts are able to convert sugar to ethanol and CO_2. For this discovery he was rewarded with the Nobel Prize for Chemistry in 1907. Nowadays a plethora of different applications of enzymatic conversions exist. One of the most currently widely used cell-free processes is PCR using polymerases for DNA-amplification. Enzymes are also used in the food and feed industry for processing of the respective products at larger scales [6–8]. The recently developed strategies for sustainable production of chemicals and fuels in the last decade focusing on renewable resources within the current biorefinery concepts have further increased the importance of enzymatic conversions in processing of biomass to fermentable substrates [9].

New concepts for bioproduction involve the use of enzymatic chains converting substrates in a stepwise manner to the final product. The first reports describing the reconstruction of entire metabolic pathways in vitro to generate the respective products via a multienzyme reaction have already been published [10–14]. However, chains of enzymatic reactions in a complex cell-free environment are still at an early developmental stage. Nevertheless, the combination of defined enzymes or enzymatic chains both in vivo and in vitro for the implementation of complex or completely novel, synthetic biochemical reactions is a promising avenue for next-generation biotechnologies.

Setting up entire biosynthetic pathways de novo will benefit greatly from the ability to have access to a large portfolio of enzymes on demand, including the option to test variants with optimized performance. Cell-free protein synthesis could be an option for rapid protein production in order to generate and test the enzymes required and to optimize them in cell-free screening procedures. The syntheses of large nucleic acid molecules nowadays also enable the creation of new combinations of modules of enzymes with different regulatory and catalytic properties including enzymatic chains composed of several individual or even fused molecules resulting in directed reaction chains with optimized kinetics and without the accumulation of potentially toxic agents (see, e.g., [15, 16]).

Cell-free protein synthesis nowadays has been established at the mg/ml scale for many different proteins [17–19] (and Kubick, this book). Both systems derived from cellular extracts from several different organisms as well as largely synthetic systems using heterologously expressed purified compounds have been developed [20–22]. The most commonly used cellular extracts are derived from *E. coli*, wheat germ, and insect cells. Whereas *E. coli* lysates still have an unmatched expression level for a multitude of proteins, eukaryotic lysates provide additional posttranslational modifications and are often used for membrane protein expression.

Interestingly, recent reports showed that cell-free protein synthesis using *E. coli* lysates is almost linearly scalable from 100 μl to 100 L, over six orders of magnitude. This remarkable result was achieved in the process of up-scaling of rhGM-CSF (granulocyte–macrophage colony-stimulating factor) expression. This cytokine could be expressed at a 700-mg/L level after a 10-h reaction in a stirred fermenter without any additional process control [23].

One of the main limitations in cell-free synthesis is the lifetime of cellular extracts and the lack of reactor concepts adapted to cell-free protein synthesis. Currently no reactor concept including active control or bioprocess development common for all [24] fermentation processes is visible. A main short-term goal for cell-free protein synthesis therefore should be the development of cell-free reactor systems including the development of appropriate control systems and reaction devices enabling a controlled feed of the required substrates as well as supplying the required energy and removing the products generated.

Another limitation is the current focus on the production of individual proteins in cell-free protein synthesis. To be able to produce complex biomolecules requiring several synthesis steps, the setup of synthetic pathways, including several controllable gene expression systems in vitro will be required. In addition new reactor concepts enabling longer synthesis periods will be necessary. This development needs a thorough understanding of molecular circuits. A combination of the efforts in cell-free protein synthesis, the development of molecular circuits and the required reactor concepts may allow such a development. Several aspects focusing on these three topics are reviewed and discussed below.

3 The Modular Setup of Molecular Circuits

The advent of genome-scale investigations in the late twentieth century laid the foundations for a thorough understanding of molecular pathways regulating cellular metabolism. The ability to sequence the genomes was the first breakthrough enabling the researcher to look at the entire genetic information of any desired organism. The first genome sequences from *Haemophilus influenza* [25] and especially from *Saccharomyces cerevisiae* [26] advanced the molecular understanding of the sequenced organisms tremendously and resulted finally in the sequence of the human genome [27, 28]. The knowledge of the genome of an organism enabled the comprehensive setup of transcriptomic profiling and their analysis at least for all potential protein coding genes and facilitated the analysis of proteome data significantly. Through these technical advances detailed pathway analyses of entire organisms are feasible [29].

Comparisons between different but related organisms were crucial for understanding the modular structure of molecular circuits and their evolution. In particular, genome-wide comparative analyses of transcriptional circuits across different yeast species have been performed. Large-scale rewiring in which individual genes move in and out of regulons through changes in *cis*-regulatory sequences has been identified to be a general phenomenon. Transcription factor substitution and the formation of new combinatorial interactions are also important contributors to the rewiring. This leads to the situation that orthologous transcription factors of high sequence identity might be able to substitute for each other in different organisms but regulate a different set of genes and regulons [30, 31]. These findings clearly showed the modular setup of regulatory circuits and their evolutionary plasticity.

The availability of genomes and transcriptomes is not a limiting factor anymore. Since the development of the next-generation sequencing technologies all kinds of nucleic acid analysis are possible at a Tera-base level within a few days. Thus in-depth studies of molecular circuits on a cellular level based on nucleic acids (including mRNA, ncRNA, and miRNA among others) are possible without any bottlenecks at the level of data generation. This is manifested, for example, by the public release of sequencing data from more than 1,000 human genomes collected between 2008 and 2010 and their analysis to set up an integrated map of genetic variation by the end of 2012 [32].

4 Synthetic Molecular Circuits

In synthetic biology not only molecular circuits from nature were adapted to artificial systems, but also completely artificial regulatory elements based on nucleic acids were developed. In 2004 one of the first conferences on synthetic biology, the Synthetic Biology 1.0 Conference at the Massachusetts Institute of Technology, brought forward the idea to reverse-engineer life [33]. The general concept here is to identify a class of standard operational components in a multidisciplinary attempt that can be assembled into functioning molecular machines, for example, molecular circuits representing a molecular pathway or key regulatory elements to switch on and off cellular functions.

The design of new molecular circuits ranges from purely theoretical designs mostly based on nucleic acid systems of reduced complexity to the implementation of molecular circuits in complex in vitro systems or even within cells. Nucleic acid systems have been used to understand chemical circuits based on diffusion and stochastic binding in order to translate their interactions into logical circuits [34]. The use of so-called "seesaw" gates for displacement reactions of DNA strands [35] has been very successful to engineer circuits generating in and outputs reflecting logical functions such as OR and AND (Fig. 1). More complex systems are even able to generate NOT, NAND, NOR, and XOR functions capable of computing even calculations such as the square roots of a four-digit binary number [36]. Although the development of completely artificial networks based mostly on enzyme-free DNA reaction systems may be still not close to implementation into
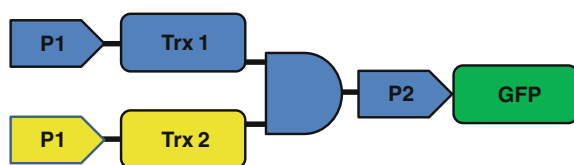

Fig. 1 AND gate circuit representing the co-activation of a reporter gene GFP by the concerted action of two transcription factors expressed from promoter P1 and acting in combination on promoter P2 to express GFP; adapted from Shin and Noireaux [39]

biotechnology, they provide valuable know-how in understanding the dynamics of genetic circuits, which may advance the reliable construction of regulatory networks in biotechnological production systems (for review see [37]).

5 Synthetic Molecular Circuits in Complex Cell-Free Systems

Regulatory circuits in cellular systems in addition to DNA include a variety of enzymatic and transcriptionally active proteins that enable much more complex regulatory circuits. This includes time-controlled assembly and disassembly processes or pattern formation in the development of multicellular organisms. Several approaches to set up oscillating regulatory circuits using a defined set of enzymes and nucleic acids were reported that move in this direction [38]. Recently the implementation of synthetic gene circuits regulated by several transcriptional units—a transcription factor and its specific target site acting as promoter for transcription—in cell-free systems derived from complete cellular extracts has been presented [39]. The general idea is to use a framework that in principle is able to self-reproduce and by that enables the continuous production and degradation of regulatory elements, such as transcription factors and their modulators. This could be possible by "hijacking" the transcription-translation machinery of the cell. Cell-free extracts as described above provide such a system. The combination of cell-free protein synthesis with synthetic molecular circuits might in the future enable the synthesis of artificial self-reproducing entities. Setting up such a system, however, requires the ability to control not only one transcription factor, or transcriptional unit such as T7- or T3-based systems, which currently are generally used in cell-free biotechnology to produce one protein for biotechnological purposes, but a multitude of transcriptional units. The availability of multiple regulatory systems in living cells is obvious and many of them have been studied extensively. Recently the regulons of *E. coli* sigma factors have been analyzed in detail with the help of in vitro run-off transcription experiments [40]. Using these data the consensus binding sites for the individual sigma factors could be pinpointed in more detail as well as the regulons addressed. In addition it was shown recently that the endogenous *E. coli* Sigma factor 70 together with *E. coli* RNA polymerase can be used to express proteins in cell-free systems efficiently [41]. With this knowledge it is now possible to set up molecular cascades in cell-free systems by combining the different transcription systems.

A first step in setting up synthetic molecular circuits based on the *E. coli* sigma factor system has been reported recently by Shin and Noireaux [39]. In this work the construction and the phenomenological characterization of synthetic gene circuits engineered with a cell-free expression toolbox is reported. The toolbox works with the seven *E. coli* sigma factors and their target sites as well as the bacteriophage polymerases T3 and T7. They use the *E. coli* endogenous holoenzyme E70 as the primary transcription machinery expressing all sigma factors and the bacteriophage-polymerases. Also transcriptional repression units derived from

E. coli (*lacI* repressor and arabinose inducible system) as well as the tetracycline repressor system (*tetR* repressor) have been established in this cell-free expression system. As reporter genes for readout of the circuits, fluorescent proteins were used. In this study the individual activity and the cross-talk between the individual transcription activation units designed were studied in detail as a prerequisite to setting up molecular circuits. In addition to the regulation of protein expression, using the respective transcription factors, proteolytic degrons were introduced into sigma factors in order to generate degradable transcription factors. Furthermore, the general half-life of mRNA in cellular extracts was manipulated in order to analyze the effect of mRNA stability on the output of transcriptional activation cascades set up by a series of transcription factors. Using this controllable system Shin and Noireaux [41] could set up elementary circuit motifs, such as multiple stage cascades and AND gate and negative feedback loops by using the evaluated sigma factors, a set of repressors, and two heterologous RNA polymerases. These circuits could also be run in synthetic liposomes in which the prepared extracts were encapsulated. Interestingly, these vesicles can be addressed from the outside. Addition of inducer molecules such as arabinose added to vesicles containing an arabinose-inducible circuit resulted in expression of the reporter gene to a similar extent as if directly introduced within the vesicle. Using this toolbox the authors could show the importance of the global mRNA turnover rate, protein turnover, and of passive competition-induced transcriptional regulation, thereby providing a model system to study complex regulatory circuits on a transcription/translation level in vitro [39].

A key feature for further development of complex synthetic circuits that eventually could result in the production of complex biomolecules requiring several synthesis steps or even self-replicating units is the period in which cell-free reaction systems remain active. The current lifetime of a few hours to approximately a day reported for some systems is certainly a major limitation for the development of more elaborate systems. One way to alleviate this could be the development of novel reactor systems for cell-free biotechnology.

6 Reactor Types for Cell-Free Synthesis: State of the Art and Perspectives

In contrast to the well-established controllable reactor systems for microbial or mammalian fermentations, currently no corresponding reactor types for cell-free protein synthesis are available on the market. The predominant way to perform cell-free protein production is a simple batch mode in which all components are mixed at the beginning of the reaction. The reaction then takes place until one of the components essential for protein synthesis is exhausted or inactivated. Although many limitations in cell-free protein synthesis have been removed or alleviated by now we are still far from knowing all mechanisms crucial for this

process (see, e.g., [42]). Currently, this type of batch mode production, running for several hours, provides protein levels at the mg/ml level and is very well suited for providing proteins for test or screening purposes. However, large-scale production of individual proteins in an economic way or setting up complex biochemical pathways requiring a timely orchestration of expression events or even the vision to create self-replicating units is not possible with this concept.

Several key issues for reactor design were addressed over 10 years ago [43, 44]. One of the key issues is the supply of energy for the reaction and feeding substrates such as amino acids. Several different reactor concepts beyond a simple batch process have been evaluated. Currently only passive diffusion systems are on the market. In order to supply the reaction with additional substrates, including amino acids and compounds for energy regeneration, membrane-based exchange systems were developed. These systems are not controlled actively but use dialysis or ultrafiltration membranes to exchange small molecules and retain the proteins of the cellular extract within the reaction compartment. The membrane in general is connected to an additional buffer reservoir in order to replace used compounds and extract deleterious compounds by diffusion that may accumulate during cell-free protein production (CECF, continuous exchange cell-free reactor) [45, 46]. These passive reactor systems are available on the market up to a volume of 10 ml and are commercialized as kits with the respective extracts and buffer systems. Interestingly these passive diffusion systems were developed as an improvement to continuous flow (continuous mode) cell-free reactors (CFCF), in which a direct continuous input of substrates (buffer) and removal of an equivalent amount of reaction solution via a membrane was realized [47, 48]. The main reason for moving to a passive diffusion system was the inefficient use of expensive substrates provided in the feeding buffer to an extent prohibitive for scaling up (Fig. 2).

A modification of these systems is a hollow fiber membrane reactor that can be operated in a cross-flow mode and has been used previously in many other biotechnological applications. Nakano et al. [43] describe a hollow fiber reactor in which the feeding buffer is conducted through a UF membrane fiber that is placed in the cell-free protein synthesis reaction compartment. Although this type of reactor was reported to increase protein production twofold compared to a batch reaction without exchange, further reports using this type of reactor in the field of cell-free protein synthesis are scarce, in contrast to reports from other areas of biotechnology.

The need for new reactor concepts is obvious considering the large improvements in cell-free protein synthesis witnessed in recent years and nicely illustrated in the chapters of this book. New reactor concepts certainly have to focus on energy regeneration systems and feeding and harvesting strategies, as well as avoiding accumulation of compounds inhibiting any of the crucial steps in a transcription–translation system.
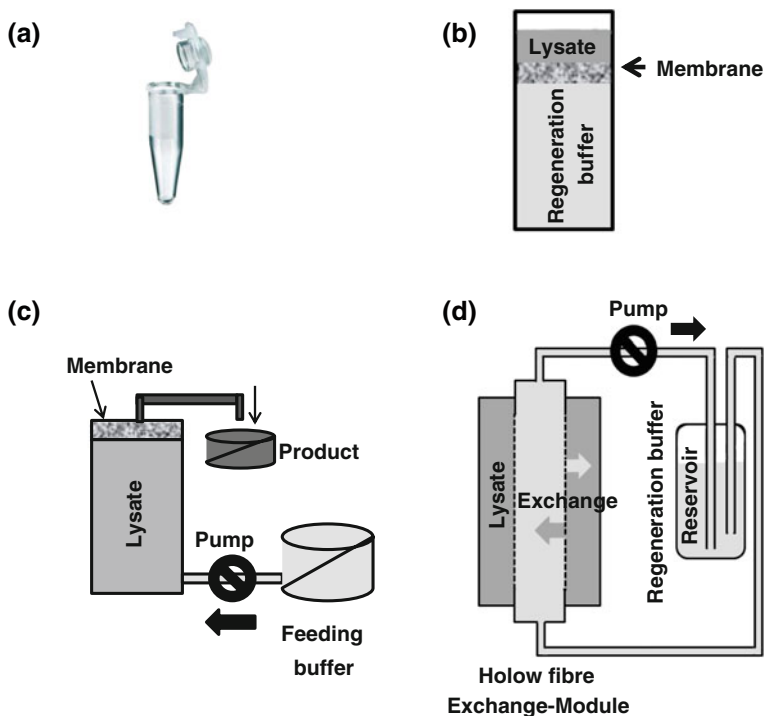
Fig. 2 Overview of current reactor types for cell-free protein synthesis. **a** The reactor type currently most commonly used in cell-free protein synthesis (CFPS). Simple containers, such as an Eppendorf tube, for batch mode reactions. **b** CECF reactor: batch reactor connected via a membrane to a buffer reservoir for regeneration of used-up components, energy, and removal of waste products generated during CFPS. **c** CFCF reactor: reactor with a constant feed of reaction buffer, including substrates and energy compounds among others; removal of low molecular weight compounds or products via an ultrafiltration membrane. **d** Variation of CFCF and CECF reactor using a hollow fiber module in cross-flow as exchange unit. The regeneration buffer flows through the reaction chamber resulting in exchange of substrate and waste

7 Integration of Energy Regeneration Concept into a Reactor Design

One of the key aspects for cell-free protein synthesis was the addition of energy regenerating systems to the cellular lysates. After recognizing that addition of ATP is not helpful but rather detrimental, several different energy regeneration systems focusing on balanced ATP-regeneration have been developed [49]. One of the key advances in improving energy regeneration systems was the direct use of metabolic pathways including oxidative phosphorylation from the already present endogenous energy regeneration system in *E. coli* lysates [50, 51]. Optimizing lysate production conditions to keep these pathways active resulted in the ability to use inexpensive substrates including pyruvate and glutamate or even
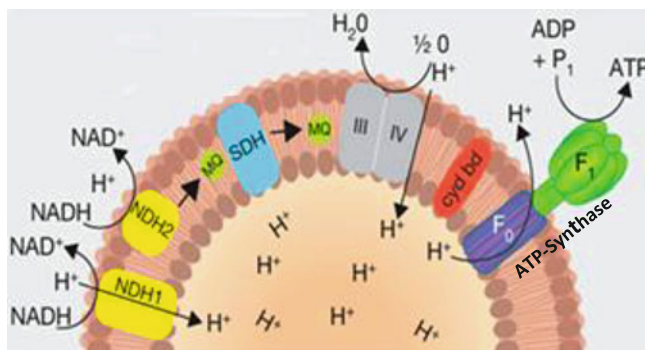
Fig. 3 Inverted inner membrane vesicle: NADH generates the proton gradient via the respiratory chain. ATP-synthase regenerates ATP from ADP

maltodextrose [52]. These compounds are metabolized by the enzymes of glycolysis and the TCA-cycle to provide reduction equivalents (NADH/FADH2). These reduction equivalents are, as in a living cell, converted to ATP within the lysate using oxidative phosphorylation on inverted inner membrane vesicles (Fig. 3). The presence of inverted inner membrane vesicles is a result of the cell homogenization procedure of *E. coli*, thereby providing access of the catalytic domain of the ATP-synthase to the outer environment of the vesicles, the lysate. The reduction equivalents drive the respiratory chain which by using oxygen as the final electron acceptor creates the proton gradient enabling ATP-regeneration using ATP-synthase. Thus the only requirement for a reactor in keeping this energy regeneration system active for a cell-free protein synthesis is the sufficient supply of oxygen. This is possible with standard fermentation equipment. Indeed, regular stirred and aerated fermentation reactors have been used to scale up oxygen-dependent cell-free protein synthesis to the liter scale. Recent reports show the production of a pharmaceutical protein rhGM-CSF [23] using a modified *E. coli* lysate (OCSF) at the 100-L scale. This reaction also was performed as a batch reaction in a standard stirred fermentation tank with the pH maintained at 7.0 and dissolved oxygen concentration of approximately 20 %, however, without any additional regulation of the system, such as feeding of substrates or other components. The stirred reactor was able to produce protein for approximately 10 h to a final concentration of 700 mg/L.

8 Energizing Membranes Without Central Carbon Metabolism

Using standard fermentation equipment for cell-free protein synthesis might optimize the productivity but gives only a limited new quality to this process. One of the key problems using cellular metabolic pathways for energy regeneration as described above is the inactivation of the system by thus far only partially

understood mechanisms. Secondary products during metabolism of the substrates could be one cause of the limited lifetime of the system. This is a well-known problem in classical fermentations, where elaborated bioprocess schemes are conducted to avoid metabolic overflow, for example. Utilizing cellular metabolism in a cell-free system, however, is not mandatory. This is also true if ATP-synthase as the most efficient energy regenerating system of the cell is to be used for energy regeneration. The only requirement for using ATP-synthase as the energy regeneration system is the creation of a pH-gradient across a membrane in which the ATP-synthase is residing. Cell-free protein synthesis also has been shown to work without the majority of proteins in a cytosolic lysate. The PURE (protein synthesis using recombinant elements) system consists of a preparation of ribosomes to which purified recombinant compounds of the translation machinery, including the required 20 aminoacyl-t-RNA synthetases, initiation-, elongation- and release-factors and polymerases for transcription are added [20, 21]. This limited amount of compounds both enables less complex purification of the product and simpler data generation for process control. Although the PURE system shows a significantly higher mRNA stability than cell-free extracts, because of the absence of RNAses the overall protein yield is still significantly lower than in the cell-free protein production systems using complete cellular lysates. One reason for this might be the use of less-efficient energy regeneration systems. A combination of new energy regeneration systems with these more defined cell-free production systems could be employed to realize novel concepts that will be able to supersede the use of cellular metabolism as regeneration systems and might be more flexible and easier to control.

9 Reactor Concepts for Use of ATP-Synthase as an Energy Regeneration System

As described above, ATP-synthase is a highly efficient system for energy regeneration. It requires a proton gradient for activation that can be generated by several different means. The natural way is to use reduction equivalents (NADH, FADH2) generated by the central carbon metabolism and the respiratory chain embedded in the membrane. In order to avoid the use of barely controllable metabolic pathways in a lysate these reduction equivalents could be provided by other means. The use of alternative co-factor regeneration systems would render the enzymes of the central carbon metabolism obsolete. The inner membrane vesicles carrying the respiratory chain and the ATP-synthase would be sufficient for energy regeneration.

10 Inner Membrane Vesicles

Inner membrane vesicles can be isolated from *E. coli* using a similar effort to that of isolating the ribosomal fraction required for cell-free protein synthesis. In order to use the inner membrane vesicles for ATP production the only requirement is the supply of NADH. Purified inner membrane vesicles have been shown to support ATP regeneration after the addition of NADH (Fig. 3) (see, e.g., [53]). This option, however, would require a regeneration system for NADH for long-term use of the production system. Co-factor regeneration is crucial also for many enzymatic reactions in pharmaceutical or industrial biotechnology. Therefore several options exist to set up simple enzymatic or electrochemical NADH regeneration systems for driving the proton gradient in inner membrane vesicles required for ATP generation. These options span from enzymatic reactions to electrochemical regeneration of NADH, providing options to control the NADH level directly (for review see [54]). One of the most elegant enzymatic regeneration systems is for-mate dehydrogenase, which converts formate into CO_2. CO_2 can be removed easily from the reaction, resulting in an irreversible reaction without accumulation of the product in the system. Regulated addition of formate to the reaction might enable NADH-level adjustments appropriate for cell-free protein synthesis and avoid the detrimental effects of formate as has been reported for other applications [55].

A second option would be the electrochemical regeneration of NADH. Here significant advances have been made recently. Using Rh(BiPy) complexes several limitations in electrochemical reduction of NAD^+ could be overcome [56]. Also new synthetic analogues of NAD^+ with higher stability and improved properties in electrochemical reduction have been shown to open new ways for substitution of NAD^+/NADH [57]. Combining electrochemical reaction systems with inner membrane vesicles might be an interesting option for ATP regeneration.

Reactor concepts using the respiration chain to generate the driving force for ATP-synthesis still need to take the oxygen demand of the system into account, including the required aeration. However, introduction of electrochemical or enzymatic approaches for NADH generation might give much easier options for controlling the production rates of ATP and thereby protein production, without the problem of side reactions in complete lysates hampering the process.

11 NADH-Independent ATP-Synthase Activation

In order to control the ATP-synthase function more directly and to avoid the dependence on oxygen, several other ideas of generating H^+-gradients have been developed. The most prominent ones are the use of different types of proton pumps, in general the rhodopsins which are light-driven proton pumps. The motivation of this research originally was to develop tools needed for the mech-anistic understanding of the ATP-synthase machinery. However, the knowledge of

how to energize membranes and get membrane proteins such as ATP-synthase to work could open new ways in generating energy regeneration systems also for cell-free protein synthesis.

Several examples for NADH independent ATP production by ATP-synthases have been presented already. The first reports for light-driven proton pumps that, reconstituted together with ATP-synthase in liposomes, are able to generate ATP were published almost 40 years ago using bacteriorhodopsin and ATP-synthase from *Rhodospirillum rubrum* (see Fig. 4). The stability and productivity of these systems have been improved constantly. This includes the variation of several different types of rhodopsins and ATP-synthases as well as the lipid composition to generate the reconstituted liposomes [58–63]. To reconstitute both bacteriorhodopsin and ATP-synthases in liposomes, the individual proteins in general have to be isolated independently and reconstituted in artificial liposomes. Critical points here are the ratio of ATP-synthase and rhodopsin as well as the composition of the lipids. Using these systems, ATP-synthesis rates have been reported to reach
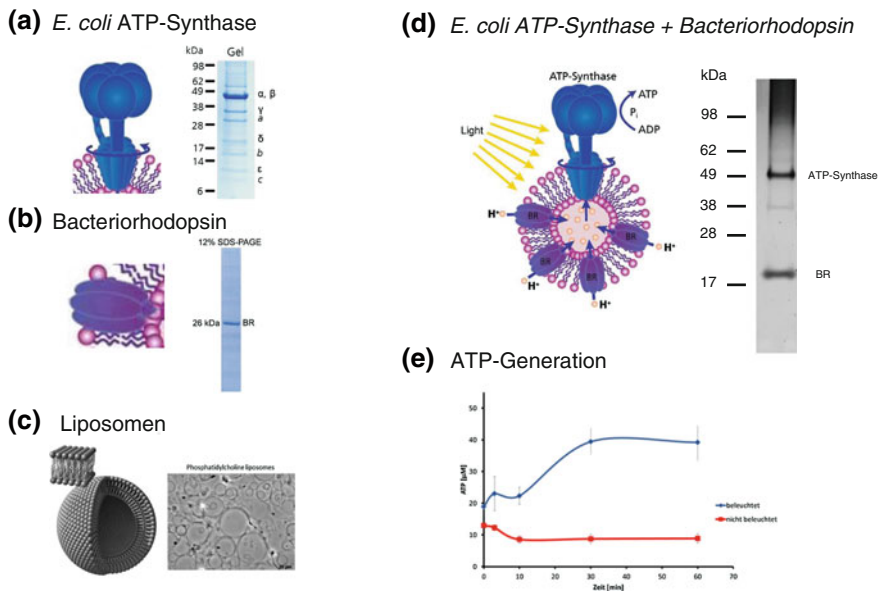
Fig. 4 Light-driven ATP-synthesis using bacteriorhodopsin as proton pump: **a** Isolation of ATP-synthase from *E. coli*. Schematic representation of ATP-synthase complex and Coomassii gel of the purified complex on SDS-PAGE. **b** Isolation of bacteriorhodopsin from *Halobacter salinarum*. Schematic representation of bacteriorhodopsin and Coomassii gel of the purified protein on SDS-PAGE. **c** Liposomes, large unilamellar vesicles (LUV); schematic illustration and microscopic picture. **d** Reconstituted ATP-synthase and bacteriorhodopsin in LUVs; schematic representation and SDS-PAGE silver stained showing both ATP-synthase and bacteriorhodopsin. **e** Light-induced ATP synthesis of reconstituted vesicles containing both bacteriorhodopsin and ATP-synthase. *Red line*: no light, the added ADP containing ∼10 uM ATP. The ATP is hydrolyzed by ATP-synthase. *Blue line*: light: bacterorhodopsin generates proton gradient able to drive ATP-synthesis from ADP

several hundred nmole/min mg [61]. Figure 4 shows the general setup of such a system composed of bacteriorhodopsin, *E. coli*, F0F1 ATP-synthase, and artificial liposomes. These systems are completely independent of oxygen, however, they require a light source for activating the rhodopsins which would have to be present in a respective energy regeneration module of the reactor.

Most recently an interesting report described the expression of a novel light-driven H+ transporter, delta-rhodopsin from *Haloterrigena turkmenica* in *E. coli* [64]. In contrast to bacteriorhodopsin, delta-rhodopsin can be functionally expressed in *E. coli* under normal growth conditions. Using a recombinant *E. coli* strain expressing delta-rhodopsin, inverted inner membrane vesicles could be isolated that act as light-driven ATP-regeneration systems. To increase stability and lifetime, the ATP-synthase expressed in this strain was derived from thermostable F0F1-ATP synthase from thermophilic Bacillus PS3 TF0F1; [65]. The ATP generated (40 μM) was used to drive an ATP-dependent enzymatic reaction, the conversion of glucose to glucose-6-phosphate using hexokinase. This system was designed to give a proof of concept that light-driven ATP-generation can be used for providing energy to enzymatic conversions. In contrast to the complex isolation of ATP-synthase, rhodopsin, and reconstitution in liposomes, this system requires only the isolation of inverted inner membrane vesicles, thereby reducing the required purification steps drastically without changes for a reactor concept as described above. A similar system has been reported using a rhodopsin from *Gloeobacter violaceus* [59].

12 Integration of ATP-Synthase into Synthetic Membranes

The option of directly integrating ATP-synthases in polymer-supported membranes with the associated stability and flexibility in order to create artificial ATP-regeneration membranes would be an interesting alternative for novel reactor concepts. Successful and functional integration would allow the direct use of chemically maintained proton gradients across a technical membrane, such as a hollow fiber or planar membrane, to regenerate the ATP demand of cell-free protein synthesis or other ATP-demanding processes.

Several approaches have been reported to integrate ATP-synthase into polymer membranes. In one approach liposomes containing rhodopsins and ATP-synthase were directly embedded in sol–gel matrices employing silica sol and PEG [66]. This resulted in a transparent inorganic matrix in which the liposome provides the membrane structure and protein orientation to two transmembrane proteins, bacteriorhodopsin (bR) and F(0)F(1)-ATP synthase. Sol–gel encapsulation converts the liposomal solution into a robust material without compromising the intrinsic activity of the incorporated proteins. Thus light could drive a proton gradient that was converted into regenerating ATP from ADP to a similar extent as the systems described above.

In another approach nanoporous silica particles have been used as support for lipid bilayers into which membrane proteins can be integrated [67]. Using an ATPase from *I. tratricus* expressed in *E. coli* and Na+ gradients set up by the passive ion-channel gramicidin A across the lipid bilayer of the silica particle it was possible either to generate ATP by transportation of Na^+ ions through the ATPase or vice versa to use the ATP-hydrolysis reaction to pump Na^+ ions. Nanoporous silica particles also have been used to support lipid membranes into which cytochrome *c* oxidase was incorporated [68]. This enzyme was functional with respect to catalysis of O_2 reduction to water, and charge separation across the membrane. Thereby it could maintain a proton electrochemical gradient across the supported proteomembrane. This showed that the membrane system was proton tight, and thereby in combination with ATP-synthase might be another option to regenerate ATP. In this case again supply of O_2 is essential for functionality which would have to be considered if implemented into a reactor concept.

Direct integration onto polymer membranes also has been reported for ATP-synthase. In this report ATP-synthase–containing vesicles were spread on cellulose membranes. The authors could show an oriented incorporation onto the supporting membrane, however, activity of the ATP-synthase has not been reported in this publication [69].

Block-copolymers, both planar and vesicular, have been proposed to be able to integrate proteins and by that may act as a support structure for membrane proteins. Recently this approach yielded several interesting results with regard to functional integration of membrane proteins into planar block copolymer membranes. Kumar et al. reported the functional incorporation of a complex water channel into such a membrane [70]. Thus incorporation of complex membrane channels in polymer membranes has been shown already. Further research will be needed to see if this is possible for other more complex proteins such as ATP-synthase as well. This would also open up new perspectives for simplified reactor concepts utilizing planar-block-copolymer membranes regenerating ATP employing a purely chemical proton gradient across the membrane to drive ATP-synthesis (Fig. 5).

The different concepts of energy regeneration systems discussed above could represent a possible evolution of reactor systems for cell-free protein synthesis as well as other ATP-dependent processes. The charm of this development would be the continuous reduction of complexity of the system—and thereby our ability to control it—which in the end would boil down to a single protein complex converting an electrochemical gradient into chemical energy that could fuel the translation machinery (Fig. 6). Bringing this concept to reality will still require strong efforts in joining membrane technology and biochemistry to realize these ambitious goals.
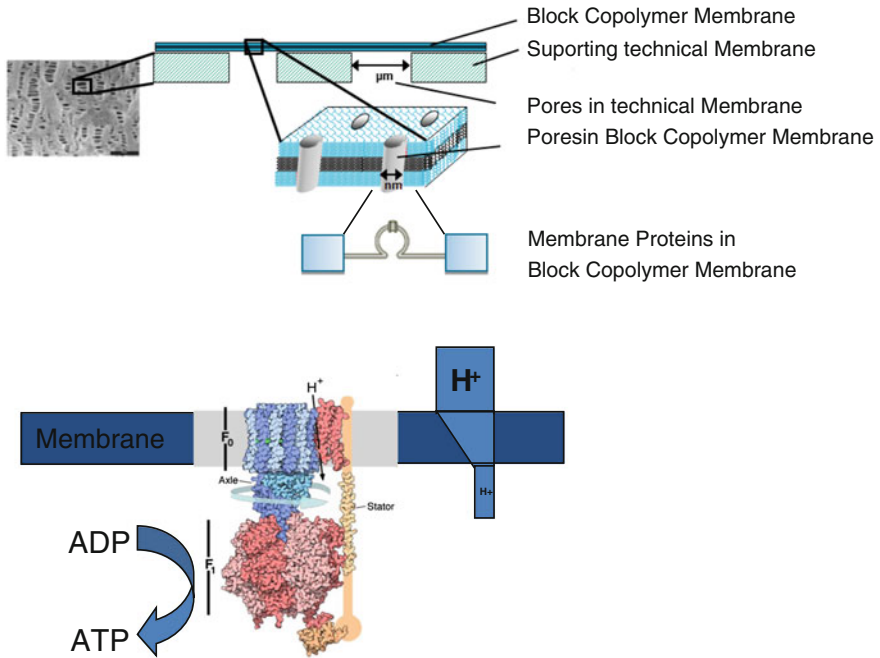
Fig. 5 Model of membrane protein in block-copolymer membranes: ATP-synthase directly converting a proton gradient into chemical energy, ATP. *Upper scheme*: Structure of supported block copolymer membranes with pores to integrate proteins. *Lower scheme*: Vision of direct activation of ATP-synthase in block copolymer membranes using a chemical proton gradient

13 Conclusions

The development of biotechnology has gained considerable pace in recent decades, both fueled by the development in medicine and pharmaceutical biotechnology as well as by the recent developments in industrial biotechnology. Advances in bioprocess technologies and metabolic engineering made it possible to improve existing products and develop novel ones using fermentation processes. However, the inherent limitations of cells fostered new concepts in biotechnology focusing on cell-free production systems. Especially the lack of cellular barriers, such as membranes and direct access to the product are a significant advantage over classical fermentation processes. The increasing wealth of databases furthermore enables the identification and realization of novel synthetic pathways for production of new compounds using partially or completely in vitro generated synthetic pathways. Thereby these technologies will be a powerful platform for contributing to a next-generation biotechnology. Nevertheless, the complex interactions within a cell are far from understood at a level to transfer them fully in vitro. This results in an only limited ability to control cell-free processes in an efficient way. Optimizing cell-free production systems by control of the critical
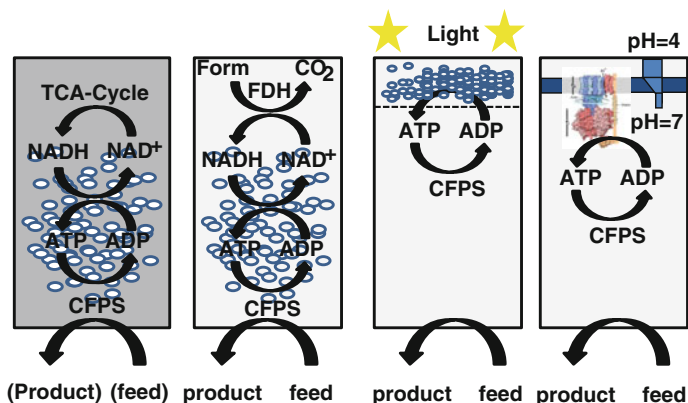
Fig. 6 Energy regeneration concepts with decreasing complexity for ATP regeneration: The depicted schemes represent potential energy modules for cell-free protein synthesis (CFPS) that can be embedded in reactors with regulated feed and product recovery processes. **a** Represents the state of the art. Utilizing metabolic pathways for energy regeneration, including inverted inner membrane vesicles (cytomim). Currently operated in batch mode. **b** Using purified compounds for CFPS and NADH generation, for example, PURE system, inner membrane vesicle, and formate dehydrogenase (FDH). Reduced amount of compounds should simplify reactor control. **c** Light-driven ATP regeneration using reconstituted vesicles of rhodopsins and ATP-synthase in an illuminated compartment. **d** ATP regeneration using chemical proton gradients across block copolymer membranes harboring ATP-synthase

parameters and efficient regeneration systems both on a structural and energetic level will be a necessary step to further improve these technologies. This requires new reactor concepts that enable the control of critical factors, such as generation of side-products hampering cell-free protein synthesis, and include energy regeneration concepts, either by utilizing existing pathways or generating novel energy sources for energy regeneration. These developments will further pave the way for cell-free processes to industrial applications.

References

1. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J et al (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–943
2. Tang WL, Zhao H (2009) Industrial biotechnology: tools and applications. Biotechnol J 4:1725–1739
3. Wohlgemuth R (2009) The locks and keys to industrial biotechnology. N Biotechnol 25:204–213

4. Sauer M, Marx H, Mattanovich D (2008) Microbial production of 1,3-propanediol. Recent Pat Biotechnol 2:191–197
5. Zaks A (2001) Indstrial biocatalysis. Curr Opin Chem Biol 5:130–136
6. Gassen HG (2006) High-tech bioactives and novel enzymes from BRAIN, a European pioneer in white biotechnology. Biotechnol J 1:752–755
7. Panke S, Wubbolts M (2005) Advances in biocatalytic synthesis of pharmaceutical intermediates. Curr Opin Chem Biol 9:188–194
8. Taylor MJ, Richardson T (1979) Applications of microbial enzymes in food systems and in biotechnology. Adv Appl Microbiol 25:7–35
9. Chandel AK, Chandrasekhar G, Silva MB, da Silva SS (2012) The realm of cellulases in biorefinery development. Crit Rev Biotechnol 32:187–202
10. Bujara M, Schumperli M, Billerbeck S, Heinemann M, Panke S (2010) Exploiting cell-free systems: implementation and debugging of a system of biotransformations. Biotechnol Bioeng 106:376–389
11. Bujara M, Schumperli M, Pellaux R, Heinemann M, Panke S (2011) Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. Nat Chem Biol 7:271–277
12. Zhang YH, Evans BR, Mielenz JR, Hopkins RC, Adams MW (2007) High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. PLoS One 2:e456
13. Zhang YH, Huang WD (2012) Constructing the electricity-carbohydrate-hydrogen cycle for a sustainability revolution. Trends Biotechnol 30(6):301–306
14. Zhang YH, Sun J, Zhong JJ (2010) Biofuel production by in vitro synthetic enzymatic pathway biotransformation. Curr Opin Biotechnol 21:663–669
15. Mullaney JA, Rehm BH (2010) Design of a single-chain multi-enzyme fusion protein establishing the polyhydroxybutyrate biosynthesis pathway. J Biotechnol 147:31–36
16. Wang HH, Huang PY, Xu G, Haas W, Marblestone A, Li J, Gygi S, Forster AC, Jewett MC, Church GM (2012) Multiplexed in vivo His-tagging of enzyme pathways for in vitro single-pot multi-enzyme catalysis. ACS Synth Biol 1:43–52
17. Baranov VI, Morozov I, Ortlepp SA, Spirin AS (1989) Gene expression in a cell-free system on the preparative scale. Gene 84:463–466
18. Stech M, Brodel AK, Quast RB, Sachse R, Kubick S (2013) Cell-free systems: functional modules for synthetic and chemical biology. Adv Biochem Eng Biotechnol 8(1):e53134 [E-Pub ahead of print]
19. Swartz JR (2009) Universal cell-free protein synthesis. Nat Biotechnol 27:731–732
20. Ohashi H, Kanamori T, Shimizu Y, Ueda T (2010) A highly controllable reconstituted cell-free system—a breakthrough in protein synthesis research. Curr Pharm Biotechnol 11:267–271
21. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755
22. Yang WC, Patel KG, Wong HE, Swartz JR (2012) Simplifying and streamlining *Escherichia coli*-based cell-free protein synthesis. Biotechnol Prog 28:413–420
23. Zawada JF, Yin G, Steiner AR, Yang J, Naresh A, Roy SM, Gold DS, Heinsohn HG, Murray CJ (2011) Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. Biotechnol Bioeng 108:1570–1578
24. Brodel AK, Raymond JA, Duman JG, Bier FF, Kubick S (2013) Functional evaluation of candidate ice structuring proteins using cell-free expression systems. J Biotechnol 163:301–310
25. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496–512
26. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M et al (1996) Life with 6000 genes. Science 274(546):563–567
27. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

28. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al (2001) The sequence of the human genome. Science 291:1304–1351

29. Khatri P, Sirota M, Butte AJ (2012) Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol 8:e1002375

30. Martchenko M, Levitin A, Hogues H, Nantel A, Whiteway M (2007) Transcriptional rewiring of fungal galactose-metabolism circuitry. Curr Biol 17:1007–1013

31. Scannell DR, Wolfe K (2004) Rewiring the transcriptional regulatory circuits of cells. Genome Biol 5:206

32. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65

33. Voigt CA, Keasling JD (2005) Programming cellular function. Nat Chem Biol 1:304–307

34. Adleman LM (1994) Molecular computation of solutions to combinatorial problems. Science 266:1021–1024

35. Qian L, Winfree E (2011) A simple DNA gate motif for synthesizing large-scale circuits. J R Soc Interface 8:1281–1297

36. Qian L, Winfree E (2011) Scaling up digital circuit computation with DNA strand displacement cascades. Science 332:1196–1201

37. Zhang DY, Seelig G (2011) Dynamic DNA nanotechnology using strand-displacement reactions. Nat Chem 3:103–113

38. Kim J, White KS, Winfree E (2006) Construction of an in vitro bistable circuit from synthetic transcriptional switches. Mol Syst Biol 2:68

39. Shin J, Noireaux V (2012) An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. ACS Synth Biol 1:29–41

40. Maciag A, Peano C, Pietrelli A, Egli T, De Bellis G, Landini P (2011) In vitro transcription profiling of the σS subunit of bacterial RNA polymerase: re-definition of the σS regulon and identification of σS-specific promoter sequence elements. Nucleic Acids Res 39:5338–5355

41. Shin J, Noireaux V (2010) Efficient cell-free expression with the endogenous E. coli RNA polymerase and sigma factor 70. J Biol Eng 4:8

42. Kim DM, Swartz JR (2000) Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol Prog 16:385–390

43. Nakano H, Shinbata T, Okumura R, Sekiguchi S, Fujishiro M, Yamane T (1999) Efficient coupled transcription/translation from PCR template by a hollow-fiber membrane bioreactor. Biotechnol Bioeng 64:194–199

44. Spirin AS (2004) High-throughput cell-free systems for synthesis of functionally active proteins. Trends Biotechnol 22:538–545

45. Kim DM, Choi CY (1996) A semi continuous prokaryotic coupled transcription/translation system using a dialysis membrane. Biotechnol Prog 12:645–649

46. Martemyanov KA, Shirokov VA, Kurnasov OV, Gudkov AT, Spirin AS (2001) Cell-free production of biologically active polypeptides: application to the synthesis of antibacterial peptide cecropin. Protein Expr Purif 21:456–461

47. Kigawa T, Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. J Biochem 110:166–168

48. Spirin AS, Baranov VI, Ryabova LA, Ovodov SY, Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242:1162–1164

49. Calhoun KA, Swartz JR (2007) Energy systems for ATP regeneration in cell-free protein synthesis reactions. Methods Mol Biol 375:3–17

50. Jewett MC, Calhoun KA, Voloshin A, Wuu JJ, Swartz JR (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. Mol Syst Biol 4:220

51. Jewett MC, Swartz JR (2004) Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol Bioeng 86:19–26

52. Wang Y, Zhang YH (2009) Cell-free protein synthesis energized by slowly-metabolized maltodextrin. BMC Biotechnol 9:58

53. Etzold C, Deckers-Hebestreit G, Altendorf K (1997) Turnover number of Escherichia coli F0F1 ATP synthase for ATP synthesis in membrane vesicles. Eur J Biochem 243:336–343

54. Weckbecker A, Groger H, Hummel W (2010) Regeneration of nicotinamide coenzymes: principles and applications for the synthesis of chiral compounds. Adv Biochem Eng Biotechnol 120:195–242

55. Neuhauser W, Steininger M, Haltrich D, Kulbe KD, Nidetzky B (1998) A pH-controlled fed-batch process can overcome inhibition by formate in NADH-dependent enzymatic reductions using formate dehydrogenase-catalyzed coenzyme regeneration. Biotechnol Bioeng 60:277–282

56. Walcarius A, Nasraoui R, Wang Z, Qu F, Urbanova V, Etienne M, Gollu M, Demir AS, Gajdzik J, Hempelmann R (2011) Factors affecting the electrochemical regeneration of NADH by (2,2'-bipyridyl) (pentamethylcyclopentadienyl)-rhodium complexes: impact on their immobilization onto electrode surfaces. Bioelectrochemistry 82:46–54

57. Lee HJ, Lee SH, Park CB, Won K (2011) Coenzyme analogs: excellent substitutes (not poor imitations) for electrochemical regeneration. Chem Commun (Camb) 47:12538–12540

58. Freisleben HJ, Zwicker K, Jezek P, John G, Bettin-Bogutzki A, Ring K, Nawroth T (1995) Reconstitution of bacteriorhodopsin and ATP synthase from micrococcus luteus into liposomes of the purified main tetraether lipid from thermoplasma acidophilum: proton conductance and light-driven ATP synthesis. Chem Phys Lipids 78:137–147

59. Lee KA, Jung KH (2011) ATP regeneration system using E. coli ATP synthase and Gloeobacter rhodopsin and its stability. J Nanosci Nanotechnol 11:4261–4264

60. Pitard B, Richard P, Dunach M, Girault G, Rigaud JL (1996) ATP synthesis by the F0F1 ATP synthase from thermophilic Bacillus PS3 reconstituted into liposomes with bacteriorhodopsin. 1. Factors defining the optimal reconstitution of ATP synthases with bacteriorhodopsin. Eur J Biochem 235:769–778

61. Pitard B, Richard P, Dunach M, Rigaud JL (1996) ATP synthesis by the F0F1 ATP synthase from thermophilic Bacillus PS3 reconstituted into liposomes with bacteriorhodopsin. 2. Relationships between proton motive force and ATP synthesis. Eur J Biochem 235:779–788

62. Racker E, Stoeckenius W (1974) Reconstitution of purple membrane vesicles catalyzing light-driven proton uptake and adenosine triphosphate formation. J Biol Chem 249:662–663

63. Wagner N, Gutweiler M, Pabst R, Dose K (1987) Coreconstitution of bacterial ATP synthase with monomeric bacteriorhodopsin into liposomes. A comparison between the efficiency of monomeric bacteriorhodopsin and purple membrane patches in coreconstitution experiments. Eur J Biochem 165:177–183

64. Hara KY, Suzuki R, Suzuki T, Yoshida M, Kino K (2011) ATP photosynthetic vesicles for light-driven bioprocesses. Biotechnol Lett 33:1133–1138

65. Suzuki T, Ueno H, Mitome N, Suzuki J, Yoshida M (2002) F(0) of ATP synthase is a rotary proton channel. Obligatory coupling of proton translocation with rotation of c-subunit ring. J Biol Chem 277:13281–13285

66. Luo TJ, Soong R, Lan E, Dunn B, Montemagno C (2005) Photo-induced proton gradients and ATP biosynthesis produced by vesicles encapsulated in a silica matrix. Nat Mater 4:220–224

67. Oliynyk V, Mille C, Ng JB, von Ballmoos C, Corkery RW, Bergstrom L (2013) Selective and ATP-driven transport of ions across supported membranes into nanoporous carriers using gramicidin A and ATP synthase. Phys Chem Chem Phys 15:2733–2740

68. Nordlund G, Sing Ng JB, Bergstrom L, Brzezinski P (2009) A membrane-reconstituted multisubunit functional proton pump on mesoporous silica particles. ACS Nano 3:2639–2646

69. Tutus M, Rossetti FF, Schneck E, Fragneto G, Forster F, Richter R, Nawroth T, Tanaka M (2008) Orientation-selective incorporation of transmembrane F0F1 ATP synthase complex from micrococcus luteus in polymer-supported membranes. Macromol Biosci 8:1034–1043

70. Kumar M, Habel JE, Shen YX, Meier WP, Walz T (2012) High-density reconstitution of functional water channels into vesicular and planar block copolymer membranes. J Am Chem Soc 134:18631–18637

Cell-free Biosystems in the Production of Electricity and Bioenergy

Zhiguang Zhu, Tsz Kin Tam and Y.-H. Percival Zhang

Abstract Increasing needs of green energy and concerns of climate change are motivating intensive R&D efforts toward the low-cost production of electricity and bioenergy, such as hydrogen, alcohols, and jet fuel, from renewable sugars. Cell-free biosystems for biomanufacturing (CFB2) have been suggested as an emerging platform to replace mainstream microbial fermentation for the cost-effective production of some biocommodities. As compared to whole-cell factories, cell-free biosystems comprised of synthetic enzymatic pathways have numerous advantages, such as high product yield, fast reaction rate, broad reaction condition, easy process control and regulation, tolerance of toxic compound/product, and an unmatched capability of performing unnatural reactions. However, issues pertaining to high costs and low stabilities of enzymes and cofactors as well as compromised optimal conditions for different source enzymes need to be solved before cell-free biosystems are scaled up for biomanufacturing. Here, we review the current status of cell-free technology, update recent advances, and focus on its applications in the production of electricity and bioenergy.

Keywords Bioenergy · Biofuels · Biomanufacturing · Cell-free biosystems · Enzymatic fuel cell · In vitro synthetic biology · Synthetic enzymatic pathway

---

Z. Zhu · Y.-H. P. Zhang (✉)
Biological Systems Engineering Department, Virginia Tech, 304 Seitz Hall,
Blacksburg, VA 24061, USA
e-mail: ypzhang@vt.edu

T. K. Tam · Y.-H. P. Zhang
Cell Free Bioinnovations Inc., Blacksburg, VA 24060, USA

Y.-H. P. Zhang
Gate Fuels Inc., Blacksburg, VA 24060, USA

Y.-H. P. Zhang
Institute of Critical Technology and Applied Science, Virginia Tech,
Blacksburg, VA 24061, USA

Contents

1 Introduction

1.1 Cell-free Biosystems for Biomanufacturing

Enzymes as biocatalysts are highly selective and greatly accelerate reaction rates under mild conditions. So far, many biomanufacturing processes have been implemented by using a single enzyme for producing sugars, semisynthetic antibiotics, and other high-value compounds [1–4]. In order to meet bioprocessing needs, tremendous efforts have been made pertaining to the discovery of desired enzymes, enhancement in enzyme stability by protein engineering and/or enzyme immobilization, reduction in enzyme production and separation costs, and enzyme recycling [5–7]. With increased system complexity, multienzyme one pot has been used in biomanufacturing, for instance, the synthesis of chiral alcohols involving in situ cofactor regeneration [8], synthesis of saccharides [9], and cellulosic biomass hydrolysis via a few synergistic cellulase components [10].

Complicated biotransformation through more than three enzymes has been mainly used to produce high-value products in the pharmaceutical and biomedical industries, such as the synthesis of polypeptides, antibodies, chiral compounds, polysaccharides, and so on [11–15]. In these cases, costs of enzymes and cofactors and/or their recycling may not be essential to the success of their production

Fig. 1 The conceptual scheme of cell-free biosystems for the production of electricity, hydrogen, and liquid biofuels
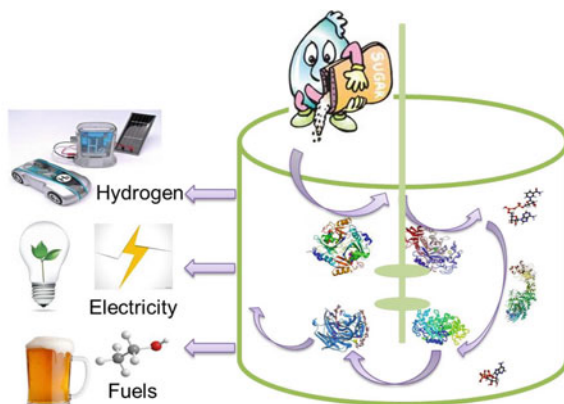
because the prices of these high-value products could vary from millions to billions of dollars per kilogram [9].

Recently cell-free biosystems for biomanufacturing (CFB2) have been suggested to produce low-value biocommodities as an alternative to whole-cell fermentation [9, 16–19]. In these cell-free biosystems, a number of enzymes and/or coenzymes are put together in one reactor or even separated by a selective cellular or synthetic membrane for implementing desired biotransformation (Fig. 1). It is essential to decrease enzyme costs and increase enzyme stability, when low-cost commodities, such as electricity, liquid biofuels, and hydrogen, are produced by using enzymes on a large scale. The design principles of cell-free biosystems have been summarized elsewhere [9, 16, 17]. Cell-free biosystems have produced electricity, hydrogen, and liquid biofuels better than microbial cells in terms of product yield, rate, and/or titer, suggesting their great industrial potential [18, 20].

1.2 Comparison of Cell-free Biosystems and Microbial Fermentation

Microorganisms have been our friends for thousands of years since our ancestors used them to make beer, wine, bread, cheese, and so on. With the development of genetic engineering, metabolic engineering, and synthetic biology, recombinant microorganisms are playing more important roles in producing fuels, energy, food, chemicals, and medicines [21, 22]. Although microbial fermentation dominates the biomanufacturing industry, many inherent drawbacks of whole-cell microorganisms cannot be addressed due to limits of the microorganisms' nature and thermodynamics. For instance, anaerobic microorganisms are able to produce a maximum of 4 units of dihydrogen per glucose unit, called the Thauer limit [23]. In contrast, cell-free biosystems without cellular membrane, gene regulation, or undesired pathways can produce nearly 12 mol of dihydrogen per glucose unit
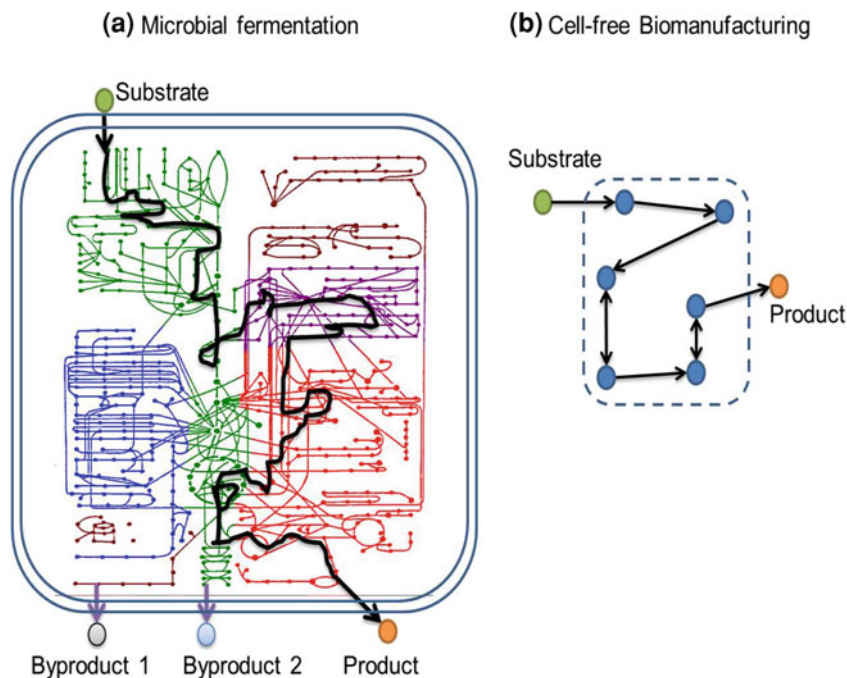
Fig. 2 Comparison of microbial cell factories (**a**) and cell-free biosystems (**b**)

(Fig. 2b) [24–26]. Another example is transformation of cellulose to starch. It is difficult for cellulolytic microorganisms to convert cellulose to starch at high yields, while achieved by a novel cell-free biosystem through restricted hydrolysis of cellulose plus cellobiose phosphorylase and alpha-glucan phosphorylase [27].

Cell-free biosystems feature several advantages over microbial cell factories (Fig. 2). (1) High product yield could be achieved with cell-free biosystems because neither by-products nor cell mass are produced. Microbes have to spend a lot of energy and resources on duplicating themselves, maintaining basic metabolisms, and often producing undesired products. As a result, in microbial cell factories, high-yield products could be achieved only by resting cells or anaerobic fermentations when the formation of cell mass is minimal [28]. (2) Reaction rates per unit area or volume of enzymatic biosystems are usually higher than those based on microbes because of higher loadings of desired enzymes per unit area or volume and no cellular membrane that causes reduced mass diffusion rates. (3) Enzymatic biosystems usually work under broader reaction conditions, such as high temperature and low pH [29]. (4) Cell-free biosystems involve the in vitro assembly of enzymes and modules only. The utilized enzymes can be either purified or crude cell lysate containing the desired proteins (Fig. 2b). By contrast, complicated microbial cell metabolism involves multigene assembling and regulation, multiprotein expression and folding, and numerous intermediates and

metabolite fluxes (Fig. 2a). As a result, the optimization of complex cellular networks could be time-consuming and labor-intensive. Engineered microorganisms frequently exhibit a performance far from that designed or expected [30, 31]. (5) Enzymes usually tolerate organic solvents and ionic liquids more than microorganisms because microorganisms should keep the cellular membrane intact for their basic functions [20, 32]. (6) Cell-free biosystems enable the construction of novel unnatural pathways implementing some biological reactions that living microbes or chemical catalysts could not previously implement. For example, β-1, 4-glucosidic bond linked cellulose has been converted to α-1,4-glucosidic bond linked starch by a mixture of intracellular and extracellular enzymes in one vessel [27, 33]. This achievement implies an out-of-the-box solution to feeding the world by the utilization of the most abundant nonfood biomass resources.

Although cell-free biosystems have numerous appealing features as mentioned above, their commercialization has a long way to go to really compete with fermentation processes, especially for some well-developed fermentations for the production of ethanol, amino acids, antibiotics, bulk enzymes such as amylase and protease, and so on. It is worth noting that some microbial fermentations may not be replaced with cell-free biosystems in the far future. For example, it is economically infeasible to produce low-cost bulk enzymes from amino acids, ammonia, and sugars by using cell-free biosystems because feedstocks are more costly than products, although it is operative to produce high-value special protein drugs and enzymes. Because it is too costly to generate a large amount of ATP in vitro at low cost [16, 34], microbial fermentations that require plenty of ATP inputs may not be replaced by cell-free biosystems, such as antibiotics or ATP-intensive tryptophan (i.e., up to 70 ATPs per tryptophan synthesized) [35]. Although in vitro ATP could be produced from low-cost polyphosphate [36] or in vitro oxidative phosphorylation of NADH [37], the former could result in an accumulation of phosphate ions, which may be an inhibitor to cell-free biosystems [4] and the latter's reconstructed cellular membrane cannot last long. Some biosynthetic pathways require a coenzyme CoA or CoA-containing metabolites, which are easily degraded in vitro; the replacement of CoA by some chemical compounds has not been explored. Therefore, some microbial fermentation involving CoA cannot be replaced with cell-free biosystems, such as polyhydroxyalkanoates (PHA) [38].

1.3 Patterns of Cell-free Synthetic Enzymatic Pathways

Patterns of multienzyme biocatalysis can be classified into four different "designs": linear, orthogonal, parallel, and cyclic according to interactions among substrates, intermediates, and products [39]. According to product generation, we would like to propose three patterns of synthetic enzymatic pathways (Fig. 3). The first pattern is *linear*, where the product from the first reaction becomes the substrate for the next reaction and a series of cascade enzymes together have a linear
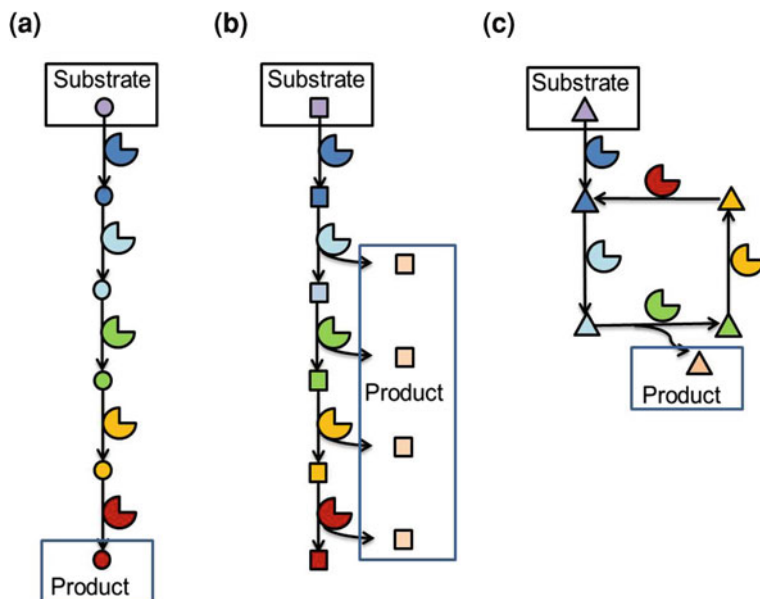
Fig. 3 Patterns of cell-free synthetic enzymatic pathways: linear (**a**), linear branch (**b**), and cyclic branch (**c**)

reaction system. Most synthetic processes regardless of chemical or biological conversions belong to this pattern [40, 41]. The second pattern is *linear branch*, where desired products are generated from "branches" in a linear enzymatic pathway. For example, three NADH molecules can be generated in a linear reaction containing three cascade enzymes [8]. In enzymatic fuel cells, a few pathways have been designed in a linear-branched pattern for in-depth oxidation of organic fuels [42–44]. The third pattern is *cyclic branch* for the generation of "branch products" through a cyclic pathway. One example is high-yield hydrogen generation from starch and water, in which one glucose-6-phosphate can generate two NADPH molecules by two dehydrogenases through the cyclic pathway [24]. Twelve units of dihydrogen can be produced per unit of glucose equivalent if the cycle repeats six times. The above classification is not strict. Sometimes, complicated pathway designs can be hybridized based on several patterns.

1.4 Cell-free Biosystems for Biomanufacturing: Problems and Their Respective Solutions

Possible doubts of CFB2 include the high production cost of many enzymes, poor enzyme stability, costly and labile coenzymes, and different optimal conditions for different source enzymes.

Cell-free biosystems can utilize relatively low-quality purified enzymes or even crude cell extracts, different from the high-purity enzymes used in academic labs and the pharmaceutical industry [2]. Therefore, there is no need to use a series of expensive chromatographic methods to obtain high-purity enzymes. Many less costly scalable protein purification techniques, such as salt precipitation, heat precipitation [45, 46], and adsorption on low-cost absorbent [47], have been developed so that enzyme purification costs can be significantly decreased. Actually, several industrial enzyme production costs have already been decreased greatly. For example, industrial cellulases have selling prices of approximately $10 US/kg [9, 48]. Transgenic plants can produce bulk enzymes at very low cost, for example, $7.8 US/kg protein as claimed by Infinite Enzymes LLC (http://www.infiniteenzymes.com/).

Enzyme stability is another important cost-determining factor. Many efforts have been focused on discovering thermostable enzymes, enzyme immobilization, and enzyme engineering via rational design or directed evolution for increasing enzyme stability [49–52]. During the past half century, a number of enzyme immobilization technologies have been developed, such as adsorption, cross-linking, encapsulation, and their combinations [53–55]. For instance, a one-step protein purification and immobilization method has been developed by using low-cost, ultrahigh adsorption capacity cellulose to adsorb cellulose-binding–module tagged thermophilic enzymes, resulting in a highly active and ultrastable enzyme with a total turnover number of more than 10^9 mol of product per mole of enzyme [56]. Akermin reported that three immobilized enzymes used for enzymatic fuel cells can actively oxidize the fuel for over three years and still retain their activities (http://www.fuelcelltoday.com/news-events/news-archive/2008/june/akermin-achieves-record-enzyme-stability). New nanomaterials provide a new opportunity of stabilizing enzymes on their surfaces for a long time [57, 58].

Costs and stability of cofactors may be another issue for cell-free biosystems because biomolecules such as ATP and NAD(P) are expensive and unstable. One of the solutions is careful design of synthetic pathways that have balanced cofactors or do not require cofactors. For example, both enzymatic pathways for hydrogen generation from polysaccharides and alcohol production from glucose do not need ATP for sugar phosphorylation, completely different from pathways occurring in microbes [20, 24]. Another solution is the use of the low-cost bio-mimetic cofactors that have similar functions with natural cofactors [59–61]. Redox enzyme engineering has achieved the change of cofactor preference from one to another.

An individual enzyme in cell-free biosystems has to work under compromised conditions, where all enzymes can work [24–26, 32, 62]. The trade-off among enzymes might be solved by using enzymes from the same organism or by engineering enzymes to have similar optimal working conditions.

We believe that cell-free biosystems have the potential to replace some of microbial processes, especially for the production of low-value biocommodities.

Several examples of bioelectricity, biohydrogen, and biofuel production based on cell-free biosystems are delineated below.

2 Bioelectricity Generation

2.1 Overview of Enzymatic Fuel Cells

Biological fuel cells (BFCs) are bioelectrochemical systems that use biocatalysts to harness chemical energy in fuels and convert it to electricity [63]. Compared to traditional fuel cells that utilize precious metal catalysts, BFCs feature modest reaction conditions, utilization of complex organic compounds (e.g., sugars, alcohols, and mixtures such as wastewater) or hydrogen, high catalytic specificity, renewable biocatalysts, nearly zero carbon emissions, biodegradability, and low safety concerns [64]. According to the biocatalysts, two types of BFCs can be classified: microbial fuel cells (MFCs) and enzymatic fuel cells (EFCs).

EFCs utilize oxidoreductases as biocatalysts to generate electricity through the oxidation of chemical fuels at the anode, whereas oxidants are reduced at the cathode with protons transferred across the membrane from the anode (Fig. 4) [63]. EFCs are regarded as next-generation, environmentally friendly, micropower sources [64] because they could have higher power output in terms of normalized electrode areas than MFCs mainly due to the lack of cellular membranes that slow down mass transfer and higher enzyme loading per volume [65–67]. With the development of more stable enzymes and better pathways, sugar-powered EFCs may have unmatched energy storage densities of approximately 3,000 Ah/kg of glucose, which are suitable for powering small or micro portable electronics in the near future [68].

2.2 Problems of EFCs

Four key challenges must be solved before a large-scale implementation of EFCs. These are incomplete oxidation of fuels (e.g., hexoses), short enzyme lifetimes, low power densities, and high costs associated with enzymes, mediators, and systems. Each challenge can be addressed with multiple approaches (Table 1), but some solutions may be in conflict. Therefore, it is very important to systematically analyze the pros and cons of each solution to determine the best choices.

Energy density is a crucial criterion for batteries and power sources. EFCs can utilize a large range of chemical compounds as fuels, but most EFCs suffer from the incomplete oxidation of fuels, leading to a waste of energy utilization. For example, most glucose-powered EFCs are based on one oxidoreductase (i.e., glucose dehydrogenase or glucose oxidase), resulting in only two electrons

Fig. 4 General configuration
of an enzymatic fuel cell
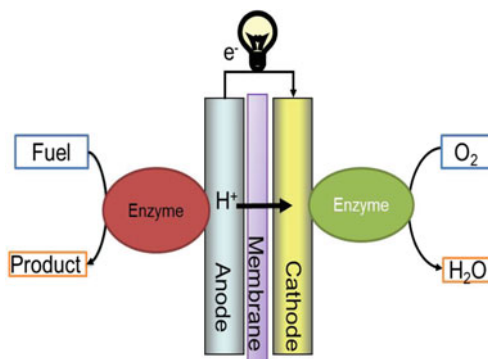
generated per glucose unit [69, 70]. Incomplete oxidation dramatically increases the fuel cost of EFCs and results in product inhibition to enzymes. To increase fuel utilization efficiency and boost energy density of EFCs, a complete oxidation of fuels mediated by cascade enzymes is highly desired [71–73]. For example, hexoses may be completely oxidized to produce electrons at theoretical yield (i.e., 24 electrons per hexose), resulting in a very high energy density of 3574 Ah/kg glucose, which could be much higher than lithium rechargeable batteries.

Enzymes suffer from low stability, low volumetric catalytic ability compared to metal catalysts, and may require tedious enzyme isolation and purification. To prolong enzyme lifetime in EFCs, most efforts have been focused on enzyme immobilization on electrodes, where enzymes usually originate from mesophilic organisms. Such immobilization not only prolongs the lifetime of enzymes but also enhances the electron transfer among enzymes and electrodes [74]. Typical immobilization methods include simple adsorption, entrapment, wiring, covalent linking, sandwiching, and so on [75]. Therefore, nearly all current EFCs are based on immobilized enzymes. However, this immobilization often severely impairs the apparent enzyme activities, which results in very low power densities [76–79]. Because enzyme stability can be enhanced by enzyme engineering, optimization of the enzyme complex, and the use of thermostable enzymes from thermophiles, other synergetic solutions could be (i) to extend the enzyme lifetime by using thermostable enzymes (thermoenzymes) and/or protein engineering [80, 81], (ii) to retain the enzyme activity and increase the mass transfer without immobilization, and (iii) to facilitate the electron transfer by using electron mediators between the enzymes and the electrodes [82, 83].

Power densities of EFCs are much lower compared to traditional fuel cell technology, resulting in their niche applications. Most EFCs can generate power at the μW–mW/cm^2 levels, which are enough to power biosensors and very small electronics [84]. Further increase in the power output of EFCs requires a combinatorial strategy of better cell configuration by using parallel stacking electrode plates [70], better enzymatic electrode design, enhanced performances of enzymes, and so on.

Table 1 Challenges and problems of EFCs and their respective solutions

Challenges	Problems	Solutions
Incomplete oxidation of fuels	Low energy density, Product inhibition	Synthetic cascade enzymes [1, 43, 62, 73, 86]
Enzyme stability	Short working lifetime, Short shelf time	Enzyme immobilization [150] Enzyme engineering [82] Thermostable enzymes [62]
Low power density Current density Voltage	Limited applications	Free enzyme for intact enzyme activity and better substrate mass transfer [84] Synthetic enzymatic complexes [151] Better redox immobilization [59] Stirring or flow through Decrease inherent ohmic resistance Better buffer [70] Better mediators [152] Directed electron transfer (DET) redox enzymes [63]
Costly EFC component Nafion membrane Mediator	High costs Complicated configuration Degradation Reduction of open circuit potential	Membrane-free by the use of both bioanode and biocathode [77] More stable mediators [59] Mediator-free DET redox enzymes [63]

2.3 Complete Oxidation of Fuels

EFCs can utilize synthetic enzymatic pathways for the complete oxidation of fuels. Synthetic enzymatic pathways have already been constructed or suggested to perform the complete or deep oxidation of methanol, ethanol, glycerol, and pyruvate [42, 44, 85, 86]. Methanol is the first fuel that has been completely oxidized to CO_2 by three NAD^+ based dehydrogenases working along a linear pattern [44] (Fig. 5a). Ethanol, glycerol, and pyruvate are demonstrated later for deep or complete oxidization through biomimetic pathways (i.e., the citric acid cycle).

Glucose is the most common fuel in nature because it is the least costly, abundant, renewable, and safe. Several studies have shown four electrons generated from one glucose unit by using two dehydrogenases [62, 87]. To completely oxidize glucose, two synthetic pathways have been suggested. The first is based on a nonnatural, minimal six-enzyme cascade pathway containing PQQ-dependent dehydrogenases, aldolase, and oxalate oxidase (Fig. 5b) [43]. This pathway utilizes low chemical specificities of some enzymes that can work on multiple substrates. The final product CO_2 was detected. However, the problems such as low power output and lack of quantitative evidence of coulombic efficiency remain
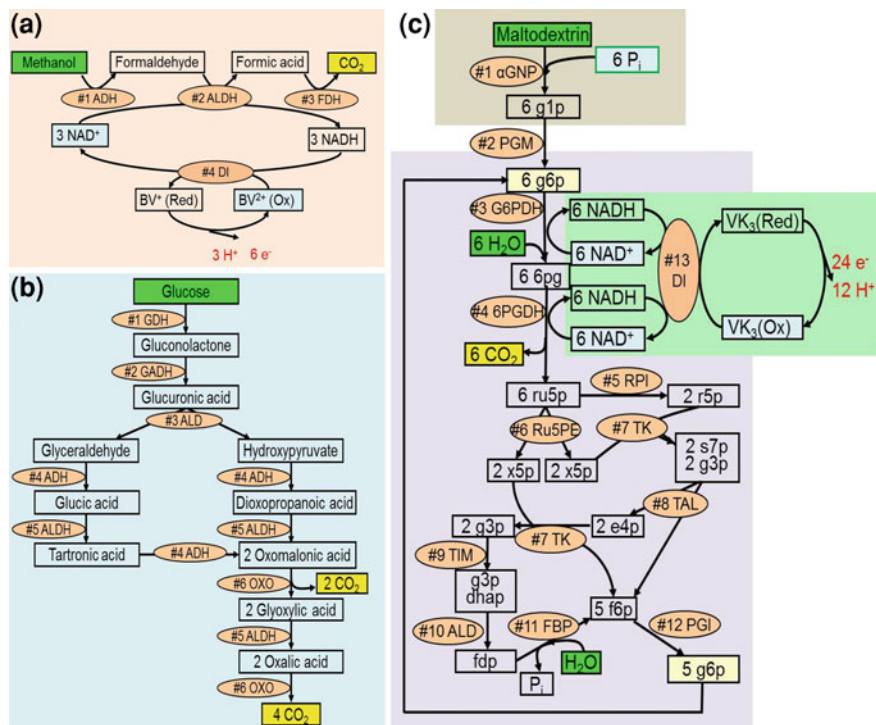
Fig. 5 Synthetic enzymatic pathways for the complete oxidation of methanol (**a**), glucose (**b**), and maltodextrin (**c**). **a** *ADH* Alcohol dehydrogenase; *ALDH* Aldehyde dehydrogenase; *FDH* Formate dehydrogenase; *DI* Diaphorase; and *BV* Benzyl viologen. **b** *GDH* Glucose dehydrogenase; *GADH* Gluconate 2-dehydrogenase; *ALD* Aldolase; *ADH* Alcohol dehydrogenase; *ALDH* Aldehyde dehydrogenase; and *OXO* Oxalate oxidase. **c** *αGNP* α-glucan phosphorylase; *PGM* Phosphoglucomutase; *G6PDH* Glucose-6-phosphate dehydrogenase; *6PGDH* 6-phosphogluconate dehydrogenase; *RPI* ribose-5-phosphate isomerase; *Ru5PE* ribulose-5-phosphate 3-epimerase; *TK*: Transketolase, *TAL* Transaldolase; *TIM* Triose phosphate isomerase; *ALD* Aldolase; *FBP* Fructose-1,6-bisphosphatase; *PGI*: Phosphoglucose isomerase; and *DI* Diaphorase. Metabolites are *g1p* Glucose-1-phosphate; *g6p* Glucose-6-phosphate; *6pg* 6-phosphogluconate; *ru5p* Ribulose-5-phosphate; *x5p* Xylulose-5-phosphate; *r5p* Ribose-5-phosphate; *s7p* Sedoheptulose-7-phosphate; *g3p* Glyceraldehyde-3-phosphate; *e4p* Erythrose-4-phosphate; *dhap* Dihydroxyacetone phosphate; *fdp* Fuctose-1,6-diphosphate; *f6p* Fructose-6-phosphate; *Pᵢ* Inorganic phosphate, and *VK₃* Vitamin K₃

unsolved. The other 13-enzyme pathway contains two dehydrogenases that are responsible for NADH generation, one diaphorase for electron transfer from NADH to electrode, and the ten other enzymes for sugar recycling in the pentose phosphate pathway, and glycolysis and gluconeogenesis pathways (Fig. 5c). This synthetic pathway utilizes phosphate instead of expensive and unstable ATP as the phosphate source for the activating glucose unit of maltodextrin. At a low maltodextrin concentration (i.e., 2 mM glucose equivalent), a constant current density

of more than 0.25 mA/cm^2 was generated (unpublished). The overall coulombic efficiency was more than 80 % of the theoretical yields, 24 electrons per glucose (in preparation for publication).

To increase the power density, free enzymes were used in EFCs [44] instead of immobilized enzymes which have significantly reduced activities due to conformation structural change and mobility restriction from the immobilization process [84]. When using a high maltodextrin concentration (i.e., 100 mM), the current density of the EFC was approximately 2 mA/cm^2 and remained constant for a few days, where all enzymes are from thermophiles. A further increase in the power density of EFCs will be focused on optimizing enzyme ratios, developing enzyme complexes featuring substrate channeling, and designing a better EFC configuration.

2.4 Switchable EFCs

Katz and his coworkers designed EFCs with a switchable power release controlled by biochemical signals, which can be logically processed by enzyme-based logic gates [88] or networks [89]. The enzyme-based logic networks are multisignal processing enzymatic systems with a biocatalytic cascade enzymatic system generating pH as the final output dependent on the whole set of the biochemical input signals [89]. The switchable properties of the EFCs were based on the polymer-brush-modified electrodes with the bioelectrocatalytic activity dependent on the solution pH value. The pH changes generated in situ by biocatalytic cascade reactions allowed the reversible activation/inactivation of the bioelectrocatalytic interfaces, resulting in a tunable activity of the entire biofuel cells.

Figure 6 shows the schematic representation of a sequence of biochemical reactions designed to produce in situ pH changes as the final output of the biochemical cascade. The reactions are activated by four chemical input signals. They are NADH, acetaldehyde, maltose, and sucrose (input signals A, B, C, and D) processed by four enzymes, which are alcohol dehydrogenase (ADH), amyloglucosidase (AGS), invertase (Inv), and glucose dehydrogenase (GDH), respectively. The biochemical input signals are defined as digital 0 in the absence of the respective chemicals and digital 1 in the optimized experimental concentration. All possible 16 combinations of the biochemical input signals were examined for changing pH in the system, and only three of them (1, 1, 0, 1; 1, 1, 1, 0; and 1, 1, 1, 1) can acidify the electrolyte solution from the initial neutral value. All "wrong" combinations of the input signals keep the initial pH value of the electrolyte solution, preserving the inactive biofuel cell that generates the low power output, whereas the acidified solution caused by "correct" combinations of the input signals can result in the transition of the pH-switchable oxygen cathode from an inactive state to an electrochemically active one, causing the activation of the entire biofuel cell. The multisignal processing enzyme system can make the "decision" to switch on and off the biofuel cell autonomously by logical processing of the biochemical input signals.
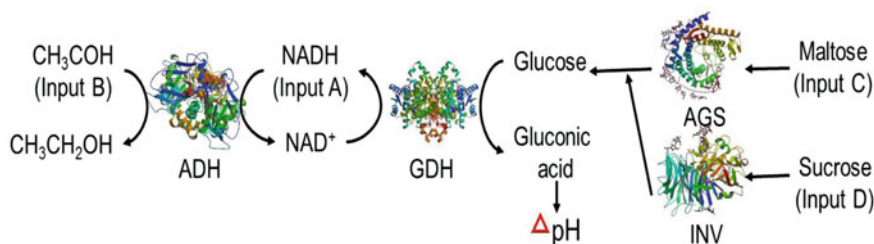
Fig. 6 A switchable EFC. *ADH* Alcohol dehydrogenase, *GDH* Glucose dehydrogenase, *AGS* Amyloglucosidase, *INV* Invertase

3 Bioenergy Production

Cell-free biosystems have been investigated to produce gaseous hydrogen, liquid alcohols, and key precursors for jet fuels from biomass sugars.

3.1 Hydrogen

Hydrogen has been proposed as a future transportation fuel due to the higher energy conversion efficiencies through polymer electrolyte membrane (PEM) fuel cell systems than those for traditional internal combustion engines and fewer pollutants generated in end users. Numerous approaches of hydrogen production have been developed from different natural sources (i.e., water, fossil fuels, solar energy). Considering the resource availability, price, renewability, and scalability, it is appealing to produce hydrogen from low-cost biomass sugars [90]. Three main approaches to generate hydrogen from biomass are chemical, biological, and the combination of both [91]. Chemical approaches include gasification, pyrolysis, and aqueous phase reforming, which suffer relatively poor selectivity, low hydrogen yields (e.g., six to eight hydrogen per glucose unit), and harsh reaction conditions. In contrast, biological approaches mediated by microorganisms or enzymes feature ambient reaction temperature and atmospheric pressure, high selectivity, and low energy input. Dark fermentation by some anaerobic microorganisms is a dominant biohydrogen production in nature. The Thauer limit (i.e., low H_2 yield) prevents its large-scale production [92]. Electrohydrogenesis using exoelectrogenic bacteria in microbial fuel cells is able to produce nearly nine hydrogen per glucose unit by using electricity as an input [93]. Photosynthetic bacteria can utilize organic acids and solar energy for hydrogen production, but suffer from very slow generation rates and potentially high hydrogen collection cost. Other issues, such as reactor engineering, system design, process control, and operation and maintenance have to be solved as well [94].

A cell-free synthetic enzymatic pathway has been proposed to produce biohydrogen from starch-based substrates with high production yields (Fig. 7) [24].
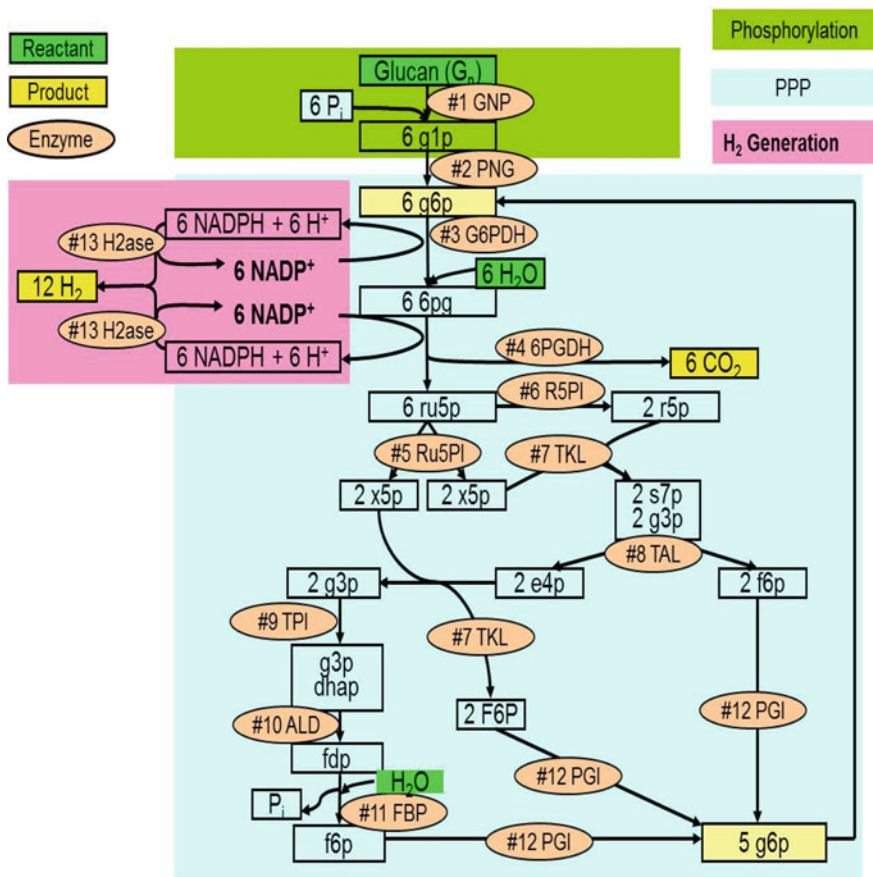
Fig. 7 Biohydrogen production from α-glucan through a cell-free synthetic enzymatic pathway, *αGNP* α-glucan phosphorylase, *PGM* Phosphoglucomutase, *G6PDH* Glucose-6-phosphate dehydrogenase, *6PGDH* 6-phosphogluconate dehydrogenase, *RPI* Ribose-5-phosphate isomerase, *Ru5PE* Ribulose-5-phosphate 3-epimerase, *TK* Transketolase, *TAL* Transaldolase, *TIM* Triose phosphate isomerase, *ALD* Aldolase, *FBP* Fructose-1,6-bisphosphatase, *PGI* Phosphoglucose isomerase, *H2ase* hydrogenase, *g1p* Glucose-1-phosphate, *g6p* Glucose-6-phosphate, *6pg* 6-phosphogluconate, *ru5p* Ribulose-5-phosphate, *x5p* Xylulose-5-phosphate, *r5p* Ribose-5-phosphate, *s7p* Sedoheptulose-7-phosphate, *g3p* Glyceraldehyde-3-phosphate, *e4p* Erythrose-4-phosphate, *dhap* d phosphate, *fdp* Fructose-1,6-diphosphate, *f6p* Fructose-6-phosphate, *P_i* Inorganic phosphate

A nearly theoretical yield of 12 mol of hydrogen per mole of glucose unit has been demonstrated [24]. This synthetic pathway contains three main modules: (1) a chain-shortening phosphorylation reaction for producing glucose-1-phosphate (G1P) catalyzed by glucan phosphorylase and conversion from G1P to glucose-6-phosphate (G6P) catalyzed by phosphoglucomutase, (2) a pentose phosphate pathway containing ten enzymes together that produce 12 NADPH per G6P unit, and (3) 12 mol of hydrogen generation from 12 mol of NADPH catalyzed by

hydrogenase. One of the most important features of this endothermic reaction is that the chemical energy output/input (hydrogen/carbohydrate) is more than 100 % (i.e., 122 %) [24]. This reaction is the first chemical reaction that can generate hydrogen energy by absorbing low-temperature waste heat, which no one utilized before. In addition, hydrogen produced through this pathway is easy to separate and collect due to the gaseous product generated from the aqueous reactants at temperatures below the water boiling point. Furthermore, enzymatic hydrogen is extremely pure so that it may be used by a polymer electrolyte membrane fuel cell (PEMFC) directly without further purification. This pathway does not involve ATP-driven reactions. Later, high-yield hydrogen production from cellulosic substrates and water in batch reactions was demonstrated [25]. It is expected that 12 hydrogen units per glucose equivalent can be achieved in a continuous reactor. In addition, a nearly 20-fold reaction rate increase has been shown (i.e., from 0.21 [95] to 3.92 mmol/L/h [25]) by optimizing rate-limiting enzyme loadings and enzyme ratios, increasing substrate concentrations, and elevating reaction temperature slightly. This reaction rate is comparable to those reported for dark fermentation and electrohydrogenesis with the help of microorganisms [93, 96]. Further rate enhancements have led to rates of up to 160 mmol of H_2/L/h [33]. The in silico kinetic model has predicted that the enzymatic hydrogen generation rate could be enhanced to 355 mmol of H_2/L/h without a discount of the hydrogen yield [97].

Furthermore, another nonnatural cell-free enzymatic pathway containing 13 enzymes has been used to produce nearly 100 % theoretical yield from a monomer pentose: xylose and water, that is, 9.6 mol of hydrogen per mole of xylose. By using a novel polyphosphate xylulokinase, we are able to convert xylose to hydrogen and carbon dioxide without ATP [26].

3.2 Alcohols

Liquid alcohols, such as ethanol and butanol, are liquid fuel additives and precursors for the production of other green biochemicals. Currently bioconversion of biomass sugars to ethanol is based on well-established microbial fermentation [98]. Unintended nonproductive reaction pathways could lead to low product yields [99]. Complicated metabolic networks could have great engineering challenges in modifying pathways as expected for the maximum formation of desired products. Different from ethanol, butanol fermentation often suffers from low product yield, slow productivity, and low product titers due to product inhibition.

Cell-free ethanol production has a long history. More than 100 years ago, Buchner discovered ethanol production from glucose with a crude yeast cell extract. The reconstruction of an in vitro yeast glycolytic enzyme system by Welch and Scopes has been demonstrated for high-yield ethanol production (i.e., 99 %) [100]. More recently, the production of cell-free ethanol or isobutanol through nonnatural pathways has been developed [20] (Fig. 8). Glucose is converted to pyruvate using four cascade enzymes, and then pyruvate is converted to ethanol by
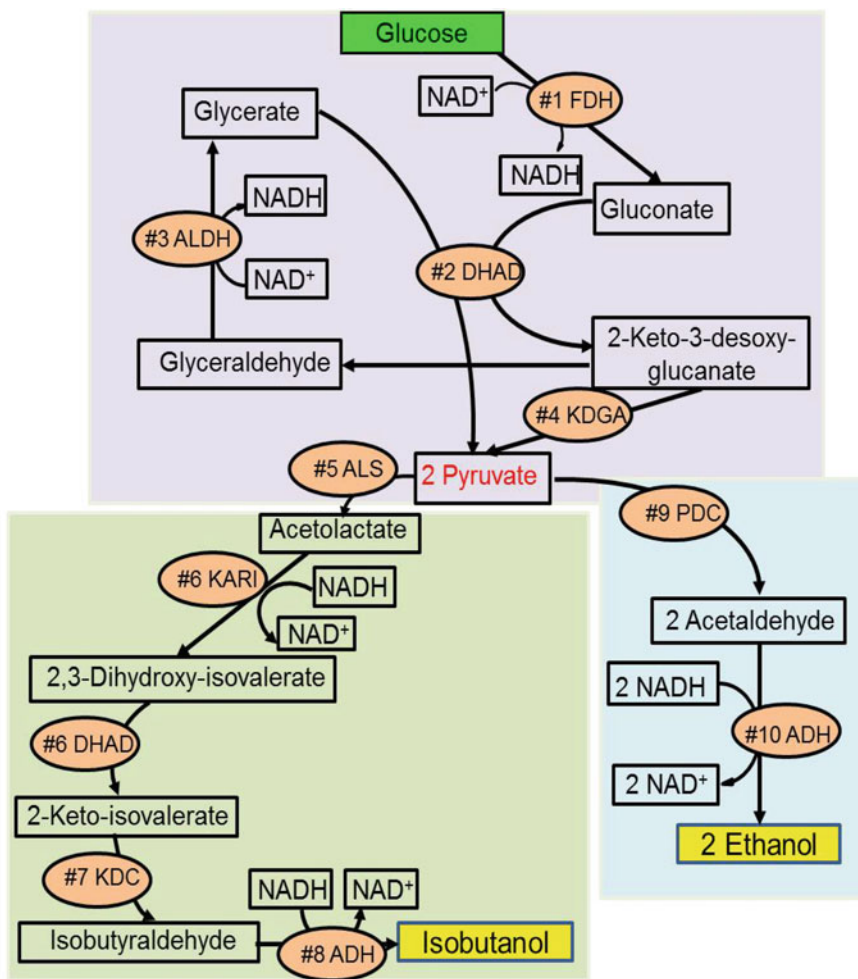
Fig. 8 Isobutanol and ethanol productions through a minimal-enzyme pathway, *GDH* Glucose dehydrogenase, *DHAD* (gluconate/glycerate) Dihydroxy acid dehydratase, *ALDH* Glyceraldehyde dehydrogenase, *KDGA* 2-keto-3-desoxygluconate aldolase, *ALS* Acetolactate synthase, *KARI* Ketolacid reductoisomerase, *KDC* 2-ketoacid decarboxylase, *PDC* Pyruvate decarboxylase complex, and *ADH* Alcohol dehydrogenase

a two-step reaction, or isobutanol by a cascade enzymatic reaction comprised of four enzymes (Fig. 8). It is noted that the above pathway does not require the involvement of ATP. Pronounced tolerance to the presence of high-level alcohols is a key characteristic of this cell-free system. In contrast to microbial cells that have reduced performances at 1 % isobutanol concentration [101], the cell-free biosystem was able to tolerate isobutanol up to 4 % [20].

3.3 Jet Fuel

Long chain fatty acids and their derivatives can be used as or upgraded to jet fuels. Because of imbalance of NAD(P)H in the anabolism and catabolism, long chain fatty acid esters (microdiesel) must be produced from sugars through semiaerobic fermentation [28, 102, 103]. Because semiaerobic fermentation consumes a more significant amount of sugar (e.g., >20 %) for the synthesis of cell mass than anaerobic fermentation does, less carbohydrate would be allocated to the production of microdiesel [28]. Such low energy-retaining efficiencies result in great challenges in economically viable production. It is one of the reasons that Amyris and LS9 gave up their biofuel production efforts.

Another high energy-retaining approach was proposed to produce jet fuel by a hybrid of cell-free biosystems and aqueous phase reforming (Fig. 9). High-yield generation of NAD(P)H from a renewable biomass sugar, cellobiose, through in vitro synthetic enzymatic pathways consisting of 12 purified enzymes and coenzymes has been designed [32] (Fig. 9). When the NAD(P)H generation system was coupled with its consumption reaction mediated by xylose reductase, the NADPH yield was as high as 11.4 mol NADPH per cellobiose (i.e., 95 % of theoretical yield; 12 NADPH per glucose unit) in a batch reaction. Consolidation of endothermic reactions and exothermic reactions in one pot results in a very high energy-retaining efficiency of 99.6 % from xylose and cellobiose to xylitol. The combination of this high-yield and projected low-cost biohydrogenation and aqueous phase reforming may be important for the production of sulfur-free liquid jet fuel in the future [32].

4 Strategies in Cell-free Biosystems for Biomanufacturing

Enzymes are a basic building block for constructing cell-free complicated synthetic pathways. Two ways for constructing cell-free biosystems are whole crude cell extracts and purified enzymes. The former is easily prepared but most enzyme components may not last long and the existing undesired pathways may cause side reactions [4, 104–106]. The latter requires more labor and cost for enzyme preparation but it has better access control and easy optimization of numerous building blocks with different characteristics. Sometimes, these two different ways may be combined for meeting different needs and retaining some advantages.

4.1 Individual Enzymes

Direct use of individual free enzymes is the most popular and straightforward method. It is desired to improve activities and stabilities of those enzyme building
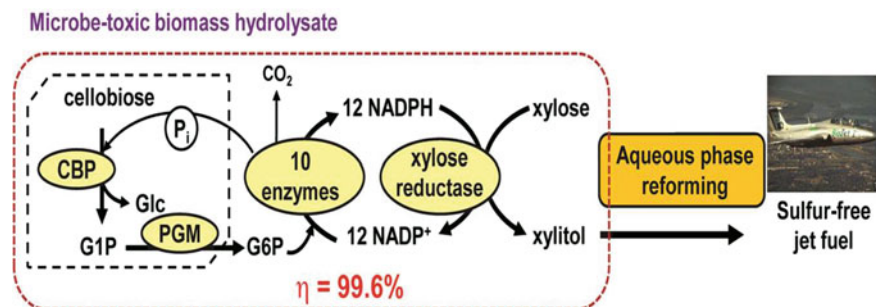
Fig. 9 Scheme of NADH generation from cellobiose through a 12-enzyme pathway. A combination of biohydrogenation and aqueous phase reforming could be used to produce jet fuel with the highest energy-retaining efficiency including production and separation costs (Modified from Ref. [32])

blocks. Typical approaches include the discovery of novel enzymes with better performances; engineering of existing enzymes with protein engineering, such as rational design, directed evolution, or a combinatorial way of using both [52, 107]; and enzyme immobilization.

The discovery and utilization of thermophilic enzymes from thermophilic microorganisms may be a shortcut. For example, a number of thermophilic enzymes have been cloned, expressed, and characterized in *E. coli*, such as *Clostridium thermocellum* cellodextrin phosphorylase [25], *C. thermocellum* cellobiose phosphorylase [25], *C. thermocellum* alpha-glucan phosphorylase [108], *C. thermocellum* phosphoglucomutase [109], *Thermotoga maritima* 6-phophogluconate dehydrogenase (Tm6PGDH) [110], *T. maritima* fructose bisphosphatase [111], and *T. maritima* pentose phosphate isomerase [45]. We have succeeded in producing more than 250 mg of Tm6PGDH per L of the culture, accounting for >30 % of the total cellular protein, using low-cost lactose as the inducer and heat-precipitation method for purification [110]. This enzyme is very active ($k_{cat} = 325$ s^{-1} at 80 °C) and ultrathermostable in terms of total turnover number (e.g., a half lifetime of 140 h at 80 °C). Adams and his coworkers have succeeded in producing the recombinant *Pyroccocus furiosus* hydrogenase, a very complicated enzyme [112]. To purify recombinant proteins at low cost, several scalable protein purification methods have been developed, such as heat precipitation [45, 110], ammonia sulfate precipitation, and cellulose-binding module-based adsorption and immobilization [36, 56].

It is difficult to estimate accurately the recombinant protein production costs, which are closely associated with the production host, expression level, purification method, and so on. As a result, current recombinant protein costs could vary by several orders of magnitude. Low-cost production of bulk enzymes has been achieved in industry, such as protease and amylase produced by *Bacillus* sp., cellulase produced by *Trichoderma* and *Aspergillus* sp., at selling prices of

approximately \$5–10 US per kg of dry protein [9, 48]. It is estimated that current costs of recombinant proteins produced by *E. coli* BL21(DE3) are approximately \$100 US per kg of enzyme [113]. Dr. Tao at EnzymeWorks (China) pointed out that current enzyme costs in his company were approximately \$70 US per kg of enzyme because they can grow the *E. coli* cell densities of more than 100 g dry cell weight per liter without the use of costly pure oxygen for aeration (personal communication). It is anticipated that the cost of bulk recombinant enzymes will decrease greatly when their markets are ready.

4.2 Enzyme Complexes as Biocatalytic Modules

Cell-free biosystems often require intensive efforts in the preparation of numerous purified enzymes as building blocks, especially when most building blocks are not commercially available. Enzyme complexes such as metabolon and other structured or synthetic enzyme systems can be used for cell-free biosynthesis. More details about this topic are available in another chapter of this book [114].

4.3 Redox Enzyme Engineering

Most redox enzymes can bind with the nicotinamide nucleotides and/or the flavins, both of which act as ancillary electron carriers in metabolisms. NAD and NADP have distinct metabolic functions, yet they differ only by an additional phosphate group esterified at the 2′-hydroxyl group of the AMP moiety of NADP. For example, NAD^+ is used exclusively in the oxidative degradations, whereas NADPH is confined with few exceptions to the reactions of reductive biosynthesis.

The first paper pertaining to redox enzyme engineering was published in1990 by Perham and his coworkers [115]. By using molecular modeling and comparing amino acid sequences responsible for cofactor binding sites, they changed NADP-preferred glutathione dihydrogen to NAD-preferred by site-directed mutagenesis [114]. Later, a number of studies were conducted by changing cofactor preferences from NADP to NAD [116–122], from NAD to NADP [123–127], and relaxing or broadening cofactor specificity [128–132] based on rational design.

Nicotinamide mononucleotide (NMN) is a precursor of NAD(P) with a smaller size as compared to NAD(P). Several wildtype redox enzymes could work with it, including liver alcohol dehydrogenase [133] and glutamic dehydrogenase [134]. A mutant with changes in three amino acids can work on NMN [135]. Recently, Scott et al. demonstrated that engineered *Pyrococcus furiosus* alcohol dehydrogenase has an ability to work with NMN [59]. By following the same strategy, our lab has modified a thermophilic G6PDH that can work with NMN although its activity is very low (data not published).

Because natural cofactors are not stable in vitro and are costly, it is important to replace them with biomimetic cofactors. In 1990s, Lowe and his coworkers developed a series of biomimetic analogs of NAD(P) based on triazine dyes [136]. Some natural dehydrogenases, such as horse liver alcohol dehydrogenase, are able to utilize such biomimetic cofactors for implementing redox reactions [136]. Later, Fish et al. found out that the pyrophosphate and adenosine groups associated with NAD are not essential in the hydride transfer and proposed the use of a low-cost and stable biomimetic, 1-benzyl-3-carbamoyl-pyridium chloride, to replace natural cofactors [137]. Later, they also showed that wildtype horse liver alcohol dehydrogenase [138] and monooxygenase [139] could weakly work with this biomimetic. However, most wildtype redox enzymes may not work on such biomimetics although several wildtype enoate reductases can work on it and its derivatives [61]. Furthermore, Clark and Fish worked together and demonstrated that a P450 mutant with two amino acid changes can work on this biomimetic [140]. Also, another group demonstrated that engineered P450 can utilize Zn dust as an electron source rather than natural cofactors [141, 142]. In 2011, Ji et al. [143] presented a bio-orthogonal system that catalyzed the oxidative decarboxylation of l-malate with a dedicated biomimetic cofactor, nicotinamide flucytosine dinucleotide (NFCD), where the redox enzymes were engineered by using saturation mutagenesis of the key amino acid sites [144]. In our lab, we have engineered a NAD(P)-preferred glucose-6-phosphate dehydrogenase working on 1-benzyl-3-carbamoyl-pyridium chloride by changing three amino acids, but its specific activity was very different (unpublished). Further tuning of redox enzymes is needed due to its important application. Although the importance of redox enzyme engineering is recognized more and more to have a place in the future of biomanufacturing [9, 104, 145, 146], redox enzyme engineering remains in its early stages because there is no framework or a general rule for engineering redox enzymes with nonnatural cofactors [59].

4.4 Treated Cell Extracts

Crude cell extracts have been used for cell-free protein synthesis for a long time due to its easy preparation and low costs [4, 104, 143, 147]. However, the other remaining cellular enzymes may cause some side reactions. Honda and his coworkers overexpressed a series of thermophilic enzymes in a mesophilic host *E. coli*. At the first step, *E. coli* cells were recombinant thermophilic enzymes at low temperatures. At the second step, the *E. coli* cell extract was treated at 70 °C for 30 min, resulting in denaturation of indigenous proteins and elimination of undesired side reactions. Consequently, highly selective and stable cell-free synthetic pathways were readily prepared [46]. By utilizing this simple enzyme cocktail preparation, they were able to produce malate [148], lactate [46], and 2-deoxyribose 5-phosphate [149] from sugars.

5 Summary

The concept of cell-free biosystems for producing high-value products is widely accepted. This concept has now been expanded to produce low-value products including electricity and bioenergy, due to its advantages of high product yields, fast reaction rate, broad reaction condition, easy process control and regulation, tolerance of toxic compound/product, and capability of performing unnatural reactions. The most challenging obstacle to CFB2 is the development of integrated, standardized, predictable, highly efficient, and stable enzymatic building blocks at low cost and on large scales. Other efforts should be made on efficient recycling of the coenzymes and use of low-cost and stable biomimetic coenzyme analogs. With more and more effective and stable enzymes and coenzymes, cell-free manufacturing will eventually replace microbial fermentation in the production of most low-value biocommodities.

References

1. Chen X, Fang JW, Zhang JB, Liu ZY, Shao J, Kowal P, Andreana P, Wang PG (2001) Sugar nucleotide regeneration beads (superbeads): a versatile tool for the practical synthesis of oligosaccharides. J Am Chem Soc 123:2081–2082
2. Swartz J (2006) Developing cell-free biology for industrial applications. J Ind Microbiol Biotechnol 33:476–485
3. Schultheisz HL, Szymczyna BR, Scott LG, Williamson JR (2008) Pathway engineered enzymatic de novo purine nucleotide synthesis. ACS Chem Biol 3:499–511
4. Wang Y, Zhang Y-HP (2009) Cell-free protein synthesis energized by slowly metabolized maltodextrin. BMC Biotechnol 9:58
5. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K (2012) Engineering the third wave of biocatalysis. Nature 485:185–194
6. Bauer MW (2005) Distinguishing red and green biotechnology: cultivation effects of the elite press. Int J Public Opin Res 17:63–89
7. Zaks A (2001) Industrial biocatalysis. Curr Opin Chem Biol 5:130–136
8. Wichmann R, Vasic-Racki D (2005) Cofactor regeneration at the lab scale. Adv Biochem Eng Biotechnol 92:225–260
9. Zhang Y-HP (2010) Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol Bioeng 105:663–677
10. Liao HH, Zhang XZ, Rollin JA, Zhang Y-HP (2011) A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. Biotechnol J 6:1409–1418
11. Fenical W, Jensen PR (2006) Developing a new resource for drug discovery: marine actinomycete bacteria. Nat Chem Biol 2:666–673
12. Hodgman CE, Jewett MC (2012) Cell-free synthetic biology: thinking outside the cell. Metab Eng 14:261–269

13. Jewett MC, Calhoun KA, Voloshin A, Wuu JJ, Swartz JR (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. Mol Syst Biol 4:220

14. Halpin DR, Harbury PB (2004) DNA display II: genetic manipulation of combinatorial chemistry libraries for small-molecule evolution. PloS Biol 2:1022–1030

15. Xu Y, Masuko S, Takieddin M, Xu H, Liu R, Jing J, Mousa SA, Linhardt RJ, Liu J (2011) Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. Science 334:498–501

16. Zhang YHP, Myung S, You C, Zhu ZG, Rollin JA (2011) Toward low-cost biomanufacturing through in vitro synthetic biology: bottom-up design. J Mater Chem 21:18877–18886

17. Zhang YHP, Sun JB, Zhong JJ (2010) Biofuel production by in vitro synthetic enzymatic pathway biotransformation. Curr Opin Biotechnol 21:663–669

18. You C, Zhang Y-HP (2013) Cell-free biosystems for biomanufacturing. Adv Biochem Eng Biotechnol 131:89–119

19. Guterl J-K, Sieber V (2013) Biosynthesis "debugged": novel bioproduction strategies. Eng Life Sci 13:4–18

20. Guterl J-K, Garbe D, Carsten J, Steffler F, Sommer B, Reiße S, Philipp A, Haack M, Rühmann B, Koltermann A et al (2012) Cell-free metabolic engineering: production of chemicals by minimized reaction cascades. ChemSusChem 5:2165–2172

21. Lee SY, Park JH, Jang SH, Nielsen LK, Kim J, Jung KS (2008) Fermentative butanol production by Clostridia. Biotechnol Bioeng 101:209–228

22. Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of E-coli. Science 291:1790–1792

23. Dutta D, De D, Chaudhuri S, Bhattacharya SK (2005) Hydrogen production by Cyanobacteria. Microb Cell Fact 4:36

24. Zhang Y-HP, Evans BR, Mielenz JR, Hopkins RC, Adams MWW (2007) High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. PLoS One 2:e456

25. Ye X, Wang Y, Hopkins RC, Adams MWW, Evans BR, Mielenz JR, Zhang Y-HP (2009) Spontaneous high-yield production of hydrogen from cellulosic materials and water catalyzed by enzyme cocktails. ChemSusChem 2:149–152

26. del Campo JSM, Rollin J, Myung S, Chun Y, Chandrayan S, Patiño R, Adams MW, Zhang Y-HP (2013) High-yield production of dihydrogen from xylose by using a synthetic enzyme cascade in a cell-free system. Angew Chem Int Ed. doi:10.1002/ange.201300766

27. You C, Chen H, Myung S, Sathitsuksanoh N, Ma H, Zhang X-Z, Li J, Zhang Y-HP (2013) Enzymatic Transformation of Non-Food Biomass to Starch. Proc Natl Acad Sci USA 110:7182–7187

28. Huang WD, Zhang Y-HP (2011) Analysis of biofuels production from sugar based on three criteria: thermodynamics, bioenergetics, and product separation. Energy Environ Sci 4:784–792

29. Panke S, Held M, Wubbolts M (2004) Trends and innovations in industrial biocatalysis for the production of fine chemicals. Curr Opin Biotechnol 15:272–279

30. Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G (2006) Engineering yeast transcription machinery for improved ethanol tolerance and production. Science 314:1565–1568

31. Gitai Z (2005) The new bacterial cell biology: moving parts and subcellular architecture. Cell 120:577–586

32. Wang Y, Huang W, Sathitsuksanoh N, Zhu Z, Zhang Y-HP (2011) Biohydrogenation from biomass sugar mediated by in vitro synthetic enzymatic pathways. Chem Biol 18:372–380

33. Zhang Y-HP (2013) Next generation biorefineries will solve the food, biofuels, and environmental trilemma in the energy-food-water nexus. Energy Sci Eng. doi: 10.1002/ese3.2

34. Resnick SM, Zehnder AJB (2000) In vitro ATP regeneration from polyphosphate and AMP by polyphosphate: AMP phosphotransferase and adenylate kinase from *Acinetobacter johnsonii* 210A. Appl Environ Microbiol 66:2045–2051
35. Akashi H, Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. Proc Natl Acad Sci USA 99:3695–3700
36. Liao HH, Myung S, Zhang Y-HP (2012) One-step purification and immobilization of thermophilic polyphosphate glucokinase from *Thermobifida fusca* YX: glucose-6-phosphate generation without ATP. Appl Microbiol Biotechnol 93:1109–1117
37. Jewett MC, Calhoun KA, Voloshin A, Wuu JJ, Swartz JR (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. Mol Syst Biol 4:57
38. Chen G-Q, Wu Q (2005) The application of polyhydroxyalkanoates as tissue engineering materials. Biomaterials 26:6565–6578
39. Ricca E, Brucher B, Schrittwieser JH (2011) Multi-enzymatic cascade reactions: overview and perspectives. Adv Synth Catal 353:2239–2262
40. Seisser B, Lavandera I, Faber K, Spelberg JHL, Kroutil W (2007) Stereo-complementary two-step cascades using a two-enzyme system leading to enantiopure epoxides. Adv Synth Catal 349:1399–1404
41. Ingram CU, Bommer M, Smith MEB, Dalby PA, Ward JM, Hailes HC, Lye GJ (2007) One-pot synthesis of amino-alcohols using a de-novo transketolase and beta-alanine: pyruvate transaminase pathway in *Escherichia coli*. Biotechnol Bioeng 96:559–569
42. Arechederra RL, Treu BL, Minteer SD (2007) Development of glycerol/O-2 biofuel cell. J Power Sources 173:156–161
43. Xu S, Minteer SD (2012) Enzymatic biofuel cell for oxidation of glucose to CO2. ACS Catal 2:91–94
44. Palmore GTR, Bertschy H, Bergens SH, Whitesides GM (1998) A methanol/dioxygen biofuel cell that uses NAD$^+$-dependent dehydrogenases as catalysts: application of an electro-enzymatic method to regenerate nicotinamide adenine dinucleotide at low overpotentials. J. Electroanal. Chem. 443:155–161
45. Sun FF, Zhang XZ, Myung S, Zhang Y-HP (2012) Thermophilic *Thermotoga maritima* ribose-5-phosphate isomerase RpiB: optimized heat treatment purification and basic characterization. Protein Expr Purif 82:302–307
46. Ye X, Honda K, Sakai T, Okano K, Omasa T, Hirota R, Kuroda A, Ohtake H (2012) Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb Cell Fact 11:120
47. Hong J, Wang YR, Ye XH, Zhang YHP (2008) Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. J Chromatogr A 1194:150–154
48. Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW (2012) The challenge of enzyme cost in the production of lignocellulosic biofuels. Biotechnol Bioeng 109:1083–1087
49. Eijsink VGH, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G (2004) Rational engineering of enzyme stability. J Biotechnol 113:105–120
50. Eijsink VGH, Gaseidnes S, Borchert TV, van den Burg B (2005) Directed evolution of enzyme stability. Biomol Eng 22:21–30
51. Bloom JD, Arnold FH (2009) In the light of directed evolution: pathways of adaptive protein evolution. Proc Natl Acad Sci USA 106:9995–10000
52. Ye X, Zhang C, Zhang Y-HP (2012) Engineering a large protein by combined rational and random approaches: stabilizing the *Clostridium thermocellum* cellobiose phosphorylase. Mol BioSyst 8:1815–1823
53. Kim J, Jia HF, Wang P (2006) Challenges in biocatalysis for enzyme-based biofuel cells. Biotechnol Adv 24:296–308
54. Hanefeld U, Gardossi L, Magner E (2009) Understanding enzyme immobilisation. Chem Soc Rev 38:453–468

55. Hwang ET, Gu MB (2013) Enzyme stabilization by nano/microsized hybrid materials. Eng Life Sci 13:49–61
56. Myung S, Zhang X-Z, Zhang Y-HP (2011) Ultra-stable phosphoglucose isomerase through immobilization of cellulose-binding module-tagged thermophilic enzyme on low-cost high-capacity cellulosic adsorbent. Biotechnol Prog 27:969–975
57. Gao Y, Zhao F, Wang Q, Zhang Y, Xu B (2010) Small peptide nanofibers as the matrices of molecular hydrogels for mimicking enzymes and enhancing the activity of enzymes. Chem Soc Rev 39:3425–3433
58. Lin Y, Zhou S, Sheehan SW, Wang D (2011) Nanonet-based hematite heteronanostructures for efficient solar water splitting. J Am Chem Soc 133:2398–2401
59. Campbell E, Meredith M, Minteer SD, Banta S (2012) Enzymatic biofuel cells utilizing a biomimetic cofactor. Chem Commun 48:1898–1900
60. Maid H, Bohm P, Huber SM, Bauer W, Hummel W, Jux N, Groger H (2011) Iron catalysis for in situ regeneration of oxidized cofactors by activation and reduction of molecular oxygen: a synthetic metalloporphyrin as a biomimetic NAD(P)H oxidase. Angew Chem Int Ed 50:2397–2400
61. Paul CE, Gargiulo S, Opperman DJ, Lavandera I, Gotor-Fernández V, Gotor V, Taglieber A, Arends IWCE, Hollmann F (2012) Mimicking nature: synthetic nicotinamide cofactors for C=C bioreduction using enoate reductases. Org Lett 15:180–183
62. Zhu ZG, Sun F, Zhang X, Zhang Y-HP (2012) Deep oxidation of glucose in enzymatic fuel cells through a synthetic enzymatic pathway containing a cascade of two thermostable dehydrogenases. Biosens Bioelectron 36:110–115
63. Minteer SD, Liaw BY, Cooney MJ (2007) Enzyme-based biofuel cells. Curr Opin Biotechnol 18:228–234
64. Davis F, Higson SPJ (2007) Biofuel cells: recent advances and applications. Biosens Bioelectron 22:1224–1235
65. Cooney MJ, Svoboda V, Lau C, Martin G, Minteer SD (2008) Enzyme catalysed biofuel cells. Energy Environ Sci 1:320–337
66. Rabaey K, Rozendal RA (2010) Microbial electrosynthesis: revisiting the electrical route for microbial production. Nat Rev Microbiol 8:706–716
67. Harnisch F, Schroder U (2010) From MFC to MXC: chemical and biological cathodes and their potential for microbial bioelectrochemical systems. Chem Soc Rev 39:4433–4448
68. Zhu ZG, Wang YR, Minteer SD, Zhang YHP (2011) Maltodextrin-powered enzymatic fuel cell through a non-natural enzymatic pathway. J Power Sources 196:7505–7509
69. Sato A, Kano K, Ikeda T (2003) Diaphorase/naphthoquinone derivative-modified electrode as an anode for diffusion-controlled oxidation of NADH in electrochemical cells. Chem Lett 32:880–881
70. Sakai H, Nakagawa T, Tokita Y, Hatazawa T, Ikeda T, Tsujimura S, Kano K (2009) A high-power glucose/oxygen biofuel cell operating under quiescent conditions. Energy Environ Sci 2:133–138
71. Arechederra RL, Minteer SD (2009) Complete oxidation of glycerol in an enzymatic biofuel cell. Fuel Cells 9:63–69
72. Sokic-Lazic D, Arechederra RL, Treu BL, Minteer SD (2010) Oxidation of biofuels: fuel diversity and effectiveness of fuel oxidation through multiple enzyme cascades. Electroanalysis 22:757–764
73. Sokic-Lazic D, de Andrade AR, Minteer SD (2011) Utilization of enzyme cascades for complete oxidation of lactate in an enzymatic biofuel cell. Electrochim Acta 56:10772–10775
74. Moehlenbrock MJ, Minteer SD (2008) Extended lifetime biofuel cells. Chem Soc Rev 37:1188–1196
75. Park BW, Yoon DY, Kim DS (2010) Recent progress in bio-sensing techniques with encapsulated enzymes. Biosens Bioelectron 26:1–10
76. Gao F, Viry L, Maugey M, Poulin P, Mano N (2010) Engineering hybrid nanotube wires for high-power biofuel cells. Nat Commun 1:2

77. Zebda A, Gondran C, Le Goff A, Holzinger M, Cinquin P, Cosnier S (2011) Mediatorless high-power glucose biofuel cells based on compressed carbon nanotube-enzyme electrodes. Nat Commun 2:370

78. Johnston W, Cooney MJ, Liaw BY, Sapra R, Adams MWW (2005) Design and characterization of redox enzyme electrodes: new perspectives on established techniques with application to an extremeophilic hydrogenase. Enzyme Microb Technol 36:540–549

79. Johnston W, Maynard N, Liaw BY, Cooney MJ (2006) In situ measurement of activity and mass transfer effects in enzyme immobilized electrodes. Enzyme Microb Technol 39:131–140

80. Güven G, Prodanovic R, Schwaneberg U (2010) Protein engineering: an option for enzymatic biofuel cell design. Electroanalysis 22:765–775

81. Wong TS, Schwaneberg U (2003) Protein engineering in bioelectrocatalysis. Curr Opin Biotechnol 14:590–596

82. Sugiyama T, Goto Y, Matsumoto R, Sakai H, Tokita Y, Hatazawa T (2010) A mediator-adapted diaphorase variant for a glucose dehydrogenase-diaphorase biocatalytic system. Biosens Bioelectron 26:452–457

83. Yu EH, Scott K (2010) Enzymatic biofuel cells-fabrication of enzyme electrodes. Energies 3:23–42

84. Zhao XY, Jia HF, Kim J, Wang P (2009) Kinetic limitations of a bioelectrochemical electrode using carbon nanotube-attached glucose oxidase for biofuel cells. Biotechnol Bioengin 104:1068–1074

85. Akers NL, Moore CM, Minteer SD (2005) Development of alcohol/O-2 biofuel cells using salt-extracted tetrabutylammonium bromide/Nafion membranes to immobilize dehydrogenase enzymes. Electrochim Acta 50:2521–2525

86. Sokic-Lazic D, Minteer SD (2009) Pyruvate/air enzymatic biofuel cell capable of complete oxidation. Electrochem Solid-State Lett 12:F26–F28

87. Tasca F, Gorton L, Kujawa M, Patel I, Harreither W, Peterbauer CK, Ludwig R, Noll G (2010) Increasing the coulombic efficiency of glucose biofuel cell anodes by combination of redox enzymes. Biosens Bioelectron 25:1710–1716

88. Amir L, Tam TK, Pita M, Meijler MM, Alfonta L, Katz E (2009) Biofuel cell controlled by enzyme logic systems. J Am Chem Soc 131:826–832

89. Tam TK, Pita M, Ornatska M, Katz E (2009) Biofuel cell controlled by enzyme logic network: approaching physiologically regulated devices. Bioelectrochemistry 76:4–9

90. Zhang YHP (2010) Renewable carbohydrates are a potential high-density hydrogen carrier. Int J Hydrogen Energy 35:10334–10342

91. Zhang YHP (2009) A sweet out-of-the-box solution to the hydrogen economy: is the sugar-powered car science fiction? Energy Environ Sci 2:272–282

92. Thauer RK, Jungermann K, Decker K (1977) Energy-conservation in chemotropic anaerobic bacteria. Bacteriol Rev 41:100–180

93. Cheng S, Logan BE (2007) Sustainable and efficient biohydrogen production via electrohydrogenesis. Proc Natl Acad Sci USA 104:18871–18873

94. Holladay JD, Hu J, King DL, Wang Y (2009) An overview of hydrogen production technologies. Catal Today 139:244–260

95. Woodward J, Orr M, Cordray K, Greenbaum E (2000) Enzymatic production of biohydrogen. Nature 405:1014–1015

96. Das D, Veziroglu TN (2001) Hydrogen production by biological processes: a survey of literature. Int J Hydrogen Energy 26:13–28

97. Ardao I, Zeng A-P (2013) In silico evaluation of a complex multi-enzymatic system using one-pot and modular approaches: application to the high-yield production of hydrogen from a synthetic metabolic pathway. Chem Eng Sci 87:183–193

98. Lau MW, Bals BD, Chundawat SPS, Jin MJ, Gunawan C, Balan V, Jones AD, Dale BE (2012) An integrated paradigm for cellulosic biorefineries: utilization of lignocellulosic biomass as self-sufficient feedstocks for fuel, food precursors and saccharolytic enzyme production. Energy Environ Sci 5:7100–7110

99. Nicolaou SA, Gaida SM, Papoutsakis ET (2010) A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. Metab Eng 12:307–331

100. Welch P, Scopes RK (1985) Studies on cell-free metabolism: ethanol-production by a yeast glycolytic system reconstituted from purified enzymes. J Biotechnol 2:257–273

101. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature 451:86–89

102. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, del Cardayre SB, Keasling JD (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. Nature 463:559–562

103. Kalscheuer R, Stolting T, Steinbuchel A (2006) Microdiesel: *Escherichia coli* engineered for fuel production. Microbiology 152:2529–2536

104. Swartz JR (2011) Transforming biochemical engineering with cell-free biology. AIChE J 58:5–13

105. Hold C, Panke S (2009) Towards the engineering of in vitro systems. J R Soc Interface 6:S507–S521

106. Bujara M, Schümperli M, Pellaux R, Heinemann M, Panke S (2011) Optimization of a blueprint for in vitro glycolysis by metabolic real-time analysis. Nat Chem Biol 7:271–277

107. Zhou H, Xie XK, Tang Y (2008) Engineering natural products using combinatorial biosynthesis and biocatalysis. Curr Opin Biotechnol 19:590–596

108. Ye X, Rollin J, Zhang Y-HP (2010) Thermophilic α-glucan phosphorylase from *Clostridium thermocellum*: cloning, characterization and enhanced thermostability. J Mol Cat B Enzymatic 65:110–116

109. Wang Y, Zhang Y-HP (2010) A highly active phosphoglucomutase from *Clostridium thermocellum*: cloning, purification, characterization, and enhanced thermostability. J Appl Microbiol 108:39–46

110. Wang Y, Zhang Y-HP (2009) Overexpression and simple purification of the *Thermotoga maritima* 6-phosphogluconate dehydrogenase in *Escherichia coli* and its application for NADPH regeneration. Microb Cell Fact 8:30

111. Myung S, Wang YR, Zhang Y-HP (2010) Fructose-1,6-bisphosphatase from a hyper-thermophilic bacterium *Thermotoga maritima*: characterization, metabolite stability and its implications. Proc Biochem 45:1882–1887

112. Sun J, Hopkins RC, Jenney FE, McTernan PM, Adams MWW (2010) Heterologous expression and maturation of an NADP-dependent [NiFe]-hydrogenase: a key enzyme in biofuel production. PLoS One 5:e10526

113. Santacoloma PA, Gr Sin, Gernaey KV, Woodley JM (2010) Multienzyme-catalyzed processes: next-generation biocatalysis. Org Proc Res Dev 15:203–212

114. Jandt U, You C, Zhang Y-HP, Zeng A-P (2013) Compartmentation and metabolic channeling: practical and modeling aspects for multienzymatic biosynthesis. Adv Biochem Eng Biotechnol

115. Scrutton NS, Berry A, Perham RN (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. Nature 343:38–43

116. Zhang L, Ahvazi B, Szittner R, Vrielink A, Meighen E (1999) Change of nucleotide specificity and enhancement of catalytic efficiency in single point mutants of *Vibrio harveyi* aldehyde dehydrogenase. Biochemistry 38:11440–11447

117. Yaoi T, Miyazaki K, Oshima T, Komukai Y, Go M (1996) Conversion of the coenzyme specificity of isocitrate dehydrogenase by module replacement. J Biochem 119:1014–1018

118. Bastian S, Liu X, Meyerowitz JT, Snow CD, Chen MMY, Arnold FH (2011) Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. Metab. Eng. 13:345–352

119. Rosell A, Valencia E, Ochoa WF, Fita I, Pares X, Farres J (2003) Complete reversal of coenzyme specificity by concerted mutation of three consecutive residues in alcohol dehydrogenase. J Biol Chem 278:40573–40580

120. Döhr O, Paine MJI, Friedberg T, Roberts GCK, Wolf CR (2001) Engineering of a functional human NADH-dependent cytochrome P450 system. Proc Natl Acad Sci USA 98:81–86
121. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S (2002) Alteration of the specificity of the cofactor-binding pocket of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A. Protein Eng Des Sel 15:131–140
122. Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S (2002) Optimizing an artificial metabolic pathway: engineering the cofactor specificity of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase for use in vitamin C biosynthesis. Biochemistry 41:6226–6236
123. Bocanegra JA, Scrutton NS, Perham RN (1993) Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. Biochemistry 32:2737–2740
124. Mittl PRE, Berry A, Scrutton NS, Perham RN, Schulz GE (1993) Structural differences between wild-type NADP-dependent glutathione reductase from *Escherichia coli* and a redesigned NAD-dependent mutant. J Mol Biol 231:191–195
125. Steen IH, Lien T, Madsen MS, Birkeland N-K (2002) Identification of cofactor discrimination sites in NAD-isocitrate dehydrogenase from *Pyrococcus furiosus*. Arch Microbiol 178:297–300
126. Watanabe S, Kodaki T, Makino K (2005) Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural Zinc. J Biol Chem 280:10340–10349
127. Glykys DJ, Banta S (2009) Metabolic control analysis of an enzymatic biofuel cell. Biotechnol Bioeng 102:1624–1635
128. Woodyer RD, van der Donk WA, Zhao H (2003) Relaxing the nicotinamide cofactor specificity of phosphite dehydrogenase by rational design. Biochemistry 42:11604–11614
129. Wiegert T, Sahm H, Sprenger GA (1997) The substitution of a single amino acid residue (Ser-116 → Asp) alters NADP-containing glucose-fructose oxidoreductase of *Zymomonas mobilis* into a glucose dehydrogenase with dual coenzyme specificity. J Biol Chem 272:13126–13133
130. Katzberg M, Skorupa-Parachin N, Gorwa-Grauslund M-F, Bertau M (2010) Engineering cofactor preference of ketone reducing biocatalysts: a mutagenesis study on a γ-diketone reductase from the yeast *Saccharomyces cerevisiae* serving as an example. Int J Mol Sci 11:1735–1758
131. Sanli G, Banta S, Anderson S, Blaber M (2004) Structural alteration of cofactor specificity in *Corynebacterium* 2,5-diketo-D-gluconic acid reductase. Protein Eng 13:504–512
132. Campbell E, Wheeldon IR, Banta S (2010) Broadening the cofactor specificity of a thermostable alcohol dehydrogenase using rational protein design introduces novel kinetic transient behavior. Biotechnol Bioeng 107:763–774
133. Plapp BV, Sogin DC, Dworschack RT, Bohlken DP, Woenckhaus C, Jeck R (1986) Kinetics and native and modified liver alcohol dehydrogenase with coenzyme analogs: isomerization of enzyme-nicotinamide adenine dinucleotide complex. Biochemistry 25:5396–5402
134. Fisher HF, McGregor LL (1969) The ability of reduced nicotinamide mononucleotide to function as a hydrogen donor in the glutamic dehydrogenase reaction. Biochem Biophys Res Commun 34:627–632
135. Flores H, Ellington AD (2005) A modified consensus approach to mutagenesis inverts the cofactor specificity of *Bacillus stearothermophilus* lactate dehydrogenase. Protein Eng Des Sel 18:369–377
136. Ansell RJ, Lowe CR (1999) Artificial redox coenzymes: biomimetic analogues of NAD$^+$. Appl Microbiol Biotechnol 51:703–710
137. Lo HC, Leiva C, Buriez O, Kerr JB, Olmstead MM, Fish RH (2001) Bioorganometallic chemistry. 13. regioselective reduction of NAD$^+$ models, 1-benzylnicotinamde triflate and beta-nicotinamide ribose-5′-methyl phosphate, with in situ generated [Cp*Rh(Bpy)H]$^+$: structure–activity relationships, kinetics, and mechanistic aspects in the formation of the 1,4-NADH derivatives. Inorg Chem 40:6705–6716

138. Lo HC, Fish RH (2002) Biomimetic NAD$^+$ models for tandem cofactor regeneration, horse liver alcohol dehydrogenase recognition of 1,4-NADH derivatives, and chiral synthesis. Angew Chem Int Ed 41:478–481

139. Lutz J, Hollmann F, Ho TV, Schnyder A, Fish RH, Schmid A (2004) Bioorganometallic chemistry: biocatalytic oxidation reactions with biomimetic NAD$^+$/NADH co-factors and [Cp*Rh(bpy)H]$^+$ for selective organic synthesis. J Organomet Chem 689:4783–4790

140. Ryan JD, Fish RH, Clark DS (2008) Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1,4-NADH cofactors. ChemBioChem 9:2579–2582

141. Nazor J, Schwaneberg U (2006) Laboratory evolution of P450 BM-3 for mediated electron transfer. ChemBioChem 7:638–644

142. Nazor J, Dannenmann S, Adjei RO, Fordjour YB, Ghampson IT, Blanusa M, Roccatano D, Schwaneberg U (2008) Laboratory evolution of P450 BM3 for mediated electron transfer yielding an activity-improved and reductase-independent variant. Protein Eng Des Sel 21:29–35

143. Sachse R, Wüstenhagen D, Šamalíková M, Gerrits M, Bier FF, Kubick S (2013) Synthesis of membrane proteins in eukaryotic cell-free systems. Eng Life Sci 13:39–48

144. Ji D, Wang L, Hou S, Liu W, Wang J, Wang Q, Zhao ZK (2011) Creation of bioorthogonal redox systems depending on nicotinamide flucytosine dinucleotide. J Am Chem Soc 133:20857–20862

145. Zhang Y-HP (2011) Simpler is better: high-yield and potential low-cost biofuels production through cell-free synthetic pathway biotransformation (SyPaB). ACS Catal 1:998–1009

146. Zhang Y-HP, Huang W-D (2012) Constructing the electricity-carbohydrate-hydrogen cycle for a sustainability revolution. Trends Biotechnol 30:301–306

147. Carlson ED, Gan R, Hodgman CE, Jewett MC (2012) Cell-free protein synthesis: applications come of age. Biotechnol Adv 30:1185–1194

148. Ye X, Honda K, Morimoto Y, Okano K, Ohtake H (2013) Direct conversion of glucose to malate by synthetic metabolic engineering. J Biotechnol 164:34–40

149. Honda K, Maya S, Omasa T, Hirota R, Kuroda A, Ohtake H (2010) Production of 2-deoxyribose 5-phosphate from fructose to demonstrate a potential of artificial biosynthetic pathway using thermophilic enzymes. J Biotechnol 148:204–207

150. Kim J, Grate JW, Wang P (2006) Nanostructures for enzyme stabilization. Chem Eng Sci 61:1017–1026

151. You C, Myung S, Zhang YHP (2012) Facilitated substrate channeling in a self-assembled trifunctional enzyme complex. Angew Chem Int Ed 51:8787–8790

152. Chakraborty D, Barton SC (2011) Influence of mediator redox potential on fuel sensitivity of mediated laccase oxygen reduction electrodes. J Am Chem Soc 158:B440–B447

In Vitro Multienzymatic Reaction Systems for Biosynthesis

Inés Ardao, Ee Taek Hwang and An-Ping Zeng

Abstract In vitro multienzymatic bioreaction systems are attracting increasing attention for the development of bioproduction systems. The de-coupling of the biocatalytic pathway from the cellular machinery aimed at growth and survival allows achievement of high product yields and thus reduces byproduct or waste generation. Additionally, the use of several enzymes allows the realization of much more complex synthetic schemes, thus expanding the chemical diversity of synthetic compounds with new chemical properties or bioactivities. This chapter provides a survey of in vitro multienzymatic bioreaction systems used for biosynthesis, discusses process design aspects for the technical realization of multienzymatic bioreactions, highlights advantages and recent developments in using multienzyme microreactors, and finally reviews examples and strategies of the co-immobilization of multienzymes in nano/microsized materials.

Abbreviations

APO	2-aminophenoxazin-3-one
NADH	Reduced form of nicotinamide adenine dinucleotide
NADPH	Reduced form of nicotinamide adenine dinucleotide phosphate
Ni–NTA	Nickel-nitrilotriacetic acid

I. Ardao · E. T. Hwang · A.-P. Zeng (✉)
Institute of Bioprocess and Biosystems Engineering,
Technische Universität Hamburg-Harburg, Denickestr. 15,
21073 Hamburg, Germany
e-mail: AZE@TUHH.de

I. Ardao
e-mail: ines.ardao@uclouvain.be

I. Ardao
Earth and Life Institute–Bioengineering Laboratory, Université catholique de Louvain,
Croix du Sud 2-L7.05.19, 1348 Louvain-la-Neuve, Belgium

LbL	Layer-by-layer
RAMA	Fructose-1,6-bisphosphate aldolase from rabbit muscle
PDMS	Polydimethylsiloxane
LTCC	Low temperature co-fired ceramics
CIM®	Convective Interaction Media
GA	Glutaraldehyde
PTFE	Poly(tetrafluoroethene)
PS	Polystirene
PB	Prussian Blue
THN	1,3,6,8-tetrahydroxynaphthalene
Au-MSNP	Gold-magnetic silica nanoparticles
GOD	Glucose oxidase
POD	Horseradish peroxidase
BSA	Bovine serum albumin
FateDH	Formate dehydrogenase
FaldDH	Formaldehyde dehydrogenase
ADH	Alcohol dehydrogenase

Contents

1 Introduction

In vitro multienzymatic processes attempt to mimic the effective enzymatic reaction pathways or networks that take place in vivo inside the cells by coupling two or more enzymes via a cascade, a parallel, or a network configuration [1]. These processes are becoming increasingly attractive for the development of new bioproduction systems owing to their great potential to achieve very high product yield and thus to reduce byproduct or waste generation, by de-coupling of the biocatalytic pathway from the cellular machinery aimed at growth and survival [2]. Additionally, the use of several enzymes allows the realization of much more complex synthetic schemes, thus expanding the chemical diversity of synthetic

compounds with new chemical properties or bioactivities [3]. They could be used to replace traditional multistep chemical processes for the production of special chemicals and pharmaceuticals that are mainly carried out by isolating the intermediary product of each step and using it as a substrate for the next step [4, 5]. Many advantages have been identified for the use of multienzyme systems in biosynthesis, such as defined enantioselectivity and stereoselectivity, high yields, easier downstream processing, and equilibrium shifting by removing the reaction products with a subsequent reaction [4, 6].

There is also an increased interest in using multienzymatic reaction systems to examine and optimize new and synthetic metabolic pathways in vitro in the context of synthetic biology, which is an emerging discipline aimed at the design and engineering of new biological components and systems to expand the biosynthetic and therapeutic applications of natural biological systems [2, 3]. Development of in vitro and in vivo multienzymatic pathways requires different steps of engineering the new metabolic routes, optimization of the pathway flux by selecting the right proportion of the enzymes involved and integration of the engineered pathway into the host's metabolism (in the case of in vivo metabolic networks) or the selection of the biocatalyst form (soluble or (co)immobilized enzymes, in the case of in vitro pathways) [3].

Connected to that, microfluidics as microbioreactors for multienzymatic reaction systems are gaining attention to aid in the synthetic pathway development. Their small size allows a better control on the reaction time than in conventional reactors, high efficiency, and outstanding repeatability and reproducibility. Additionally, the small volumes and the high flexibility and modularity have been exploited with success for designing synthetic pathways and screening reaction conditions and process development [7, 8]. Another important development in recent years for improving the efficiency of multienzymatic systems is the use of micro- and nanostructured materials for the co-immobilization of enzymes on solid supports. The benefits of the use of immobilized enzymes (e.g. enhanced thermal and operational stability, easy separation and reuse) are joined together to the improvement of reaction rates caused by the physical proximity between the enzymes, mimicking the multienzyme complexes and metabolic channeling in cells [3].

This chapter provides a survey of in vitro multienzymatic bioreaction systems used for biosynthesis, discusses the process design aspects for the technical realization of multienzymatic bioreactions, highlights advantages and recent developments in using multienzyme microreactors, and finally reviews examples and strategies of co-immobilization of multienzymes in nano/microsized materials. Readers are referred to several recent excellent reviews on multienzymatic bioreaction systems [2, 3, 6] and related issues on the applications of multienzymatic systems in synthetic biology [9–11]. Many of the other aspects of synthetic biology, such as genetic circuits programming, minimal cells, in vitro DNA or RNA synthesis and replication, synthetic genomes, and enzyme co-localization by protein scaffolding, fusion proteins or CLEAs, are therefore either not discussed or only briefly covered here.

2 Multienzyme Bioreaction Systems for Biosynthesis

In vitro multienzymatic systems present several advantages over engineered cellular systems for synthetic applications due to the reduction of the metabolic network complexity, the elimination of viability, and toxicity constraints. They allow greater reaction flexibility and reduced competing side reactions and accumulation of metabolites, which lead to higher yields. An easier reaction optimization can be done by selecting suitable enzyme and substrate concentrations, pH, temperature, and addition of co-substrates or solvents without the need for complex molecular biology techniques. However, the lack of long-term stable purified enzymes in large quantities and at a reasonable cost hinders the commercial applications of these systems presently [2, 3].

In vitro multienzymatic systems for product biosynthesis, either by reconstruction of natural metabolic pathways or de novo design of synthetic enzyme networks, have been developed in recent years (Fig. 1).

Many of the two-enzyme systems reported in the literature aim at the regeneration of expensive cofactors, namely NADH (reduced form of nicotinamide adenine dinucleotide) or NADPH (reduced form of nicotinamide adenine dinucleotide phosphate) via coupled reactions using dehydrogenases (e.g., formate, glucose, and alcohol dehydrogenases), oxidoreductases, oxygenases, and hydrogenases [12–15] or regeneration of nucleoside triphosphate, nucleoside diphosphate, and nucleoside monophosphate (N = A: Adenosine, C: Cytidine, U: Uridine) via coupling with kinases (e.g., pyruvate, polyphosphate, creatine and nucleoside monophosphate kinases) [16, 17]. Another important group of two-enzyme systems aims at coupling an additional enzyme to an enzymatic reaction to remove a byproduct in order to shift the reaction equilibrium or to eliminate an inhibitor or an undesired product [16]. A reported example is the use of catalase to remove the hydrogen peroxide formed by oxidases and reduce its inhibitory and harmful effect. This approach was used in the production of optically pure amino acids by L-amino acid oxidases [18] and in the removal of dissolved oxygen from liquid solutions by glucose oxidase to reduce oxygen interference in biosensors [19]. Examples of three-enzyme systems include the combination of three dehydrogenases for the production of oxaloacetate [20], production of (R)-1-phenylethanol and (R)-α-methylbenzylamine by two dehydrogenases coupled with a ω-transaminase [21], determination of sucrose with a three-enzyme membrane immobilized biosensor [22], or the conversion of glucose into fructose-1,6-diphosphate [23]. Four-enzyme systems have been studied for the production of aromatic D-aminoacids as D-pheylanaline and D-tyrosine [24], conversion of D-methionine into L-methionine [25], or synthesis of carbohydrates from glycerol [26].

Natural metabolic pathways in vivo have been the inspiration for the design of more complex in vitro multienzymatic systems [27, 28]. The general strategy was to first identify the desired route to the target compound and then use the corresponding isolated enzymes to perform the reaction in vitro, avoiding other enzymatic reactions to compete for the substrates or intermediates and thus helping to
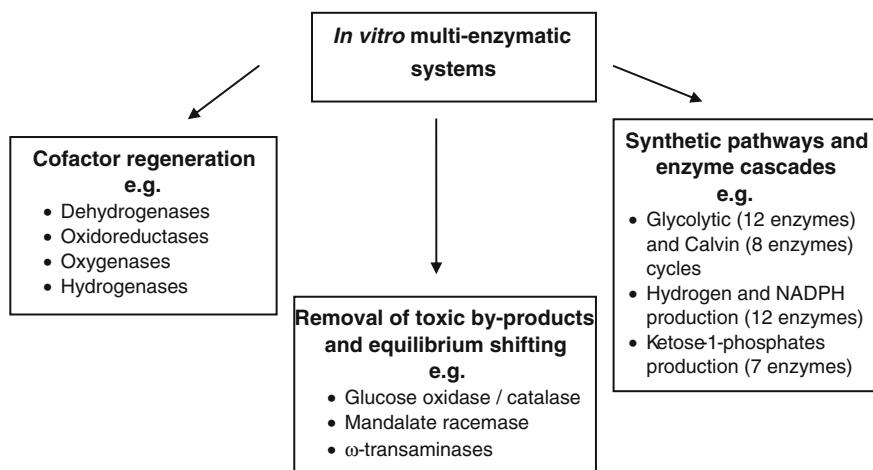
Fig. 1 Examples of different types of in vitro multienzymatic bioreaction systems for biosynthesis

obtain higher yields than in vivo. This strategy has been used to design multienzymatic pathways using enzymes for the production of ketose-1-phosphates by catalysis using 7 enzymes [29], the synthesis of corrin hydrogenobyrinic acid catalyzed by 12 enzymes [27], the reconstruction of glycolytic pathway from 12 purified enzymes [30], and the Calvin cycle from 8 enzymes to enable artificial photosynthesis [31]. A related approach is the so-called cell-free synthetic pathway biotransformation, in which enzymes from different microorganisms are combined into a de novo designed route in order to improve the system catalytic performance [28, 32]. By means of this approach, a multienzymatic system composed of 14 enzymatic reactions catalyzed by 13 enzymes has been developed for the high-yield production of hydrogen from renewable biomass with three times higher theoretical yields than natural fermentations [33, 34]. Another multienzymatic pathway designed following this approach is the in vitro synthetic enzymatic pathway for CO_2 utilization powered by electricity [35] and the production of NADPH from cellobiose [36].

The so-called one-pot batch process (all the enzymes in the same batch reactor) is the traditional reactor configuration for performing multienzymatic reactions. This configuration implies that all the enzymes encounter the same operation conditions (reaction medium, concentration of substrates and products, ionic strength, pH, and temperature). Therefore, one of the main challenges of the development of this kind of process is to select enzymes with compatible optimal reaction conditions to fully exploit the biocatalytic potential of all the enzymes. Moreover, the increase in protein concentration due to the addition of a high number of enzymes may compromise the enzyme stability. Another challenge is to avoid the presence of crossover inhibitions. This is especially important in multienzymatic systems inspired in natural metabolic pathways where many enzymes

may be inhibited by intermediates of the pathway for regulatory purposes. The use of mathematical models and simulations has become an essential tool for process analysis and optimization [5, 16, 37].

3 Process Design

One of the easiest reactor configurations to perform a multienzyme process is the use of soluble enzymes all together in a batch reactor, the so-called one-pot process. These processes present several advantages such as simple reactor operation, smaller reactor volume, less waste generation, lower investment and operation costs, and the shifting of unfavorable equilibriums toward the formation of desired products [3, 5, 6, 16]. For that reason, this is the most used approach in proof-of-concept experiments for testing the feasibility of enzymes of a reaction cascade [29, 34].

Despite the usefulness of the one-pot process as a first approach to a multienzymatic reaction, it presents a number of limitations regarding technical development of such processes. The most obvious limitation is the need to select enzymes with a compatible operational window of process conditions such as pH, temperature, and ionic strength. However, an increasing number of enzymes with enhanced biocatalytic characteristics or even exhibition of completely new activities have been discovered from thermophilic or hyperthermophilic microorganisms. One of the most useful properties for industrial applications is their usually outstanding thermostability, compared with enzymes from mesophilic organisms, which allows them to work at higher temperatures—with the consequent increase in reaction rate and yields and improved mass transfer of substrates and products [38]. For that reason, it may be advantageous in some applications to couple enzymes from thermophilic and hyperthermophilic microorganisms with other enzymes from mesophilic sources. This is the case, for example, with the synthetic pathway for high-yield hydrogen production that makes use of a hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus* coupled with 12 enzymes from the pentose phosphate and gluconeogenesis pathways of different mesophilic sources [33, 34]. In these systems, a compromise should be taken when selecting the reaction conditions (especially regarding the temperature), which leads to work under suboptimal conditions for many of the involved enzymes.

Apart from the need for finding compatible operation conditions for obtaining suitable enzyme activity in one-pot processes, the selected conditions have to be also compatible with a high stability of both the enzymes and the substrates, intermediates, or cofactors involved on the process. This is particularly important when the system involves the use of expensive and labile cofactors, such as NADH and NADPH.

Another important constraint to the use of one-pot processes is that all the enzymes are in contact with the mixture of all substrates and intermediates of the
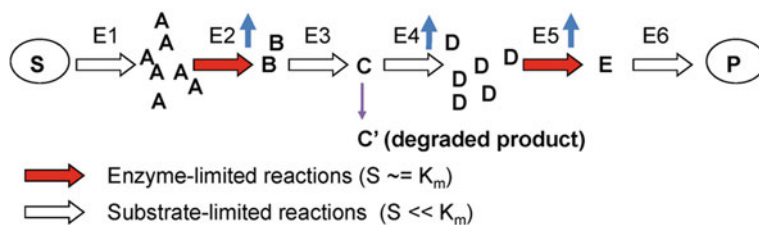
Fig. 2 Scheme of enzymatic pathway containing enzyme-limited reactions and substrate-limited reactions. *S* an initial substrate; *P* a final product; *E1–E6* enzymes; *A–E* each metabolites (*Source* [39])

reaction and, therefore, the possibility of cross-over inhibitions significantly increases. This limitation is specifically important when the multienzymatic system intends to mimic a metabolic pathway in the cell. In many metabolic pathways, the key enzymes are normally inhibited by the final product (feedback inhibition) in order to regulate the metabolic flow through this specific pathway. Feedback inhibition aims at regulating the metabolism of the cells and avoids overproduction of metabolites, but it severely hinders the production of the target compound when the multienzymatic system is used for synthetic purposes.

Feedback inhibition in a one-pot process can be somehow relieved by a careful selection of the enzymatic loadings of the cascade. It is possible to maintain the concentration of an inhibiting intermediate in values close to zero if the rate of the production of this compound is far lower than the rate of the inhibitor-consuming reactions (Fig. 2). However, it implies the use of an excess of the enzymes downward in the pathway to keep the inhibitor concentration at very low levels, which would increase the costs of the process. The right enzyme ratio between the inhibitor-producing and the inhibitor-consuming enzymes should also be maintained all over the process operation to avoid accumulation of the inhibitor. Maintaining this ratio is difficult if the enzymes present very different stabilities, which implies the need for use of an even greater excess of the less stable enzyme.

The use of soluble enzymes in one-pot systems presents some other limitations. High enzyme loadings may be necessary for achieving high reaction rates and productivities to make the process competitive and cost-effective. However, a high concentration of enzymes in solution, especially when a high number of enzymes are involved, increases the chances of molecule collisions and protein aggregation that usually leads to enzyme inactivation and reduction of enzyme stability. Because the cost of the enzymes usually represents a significant fraction of the total production costs of a biocatalytic process, recycling and reusing the enzymes is highly advantageous in order to decrease the enzyme costs. Soluble enzymes can be separated from the reaction mixture by the use of ultrafiltration membranes. However, continuous membrane reactors present several drawbacks, such as membrane fouling and high energy consumption to provide a sufficient pressure difference across the membrane [16].

One-pot

batch
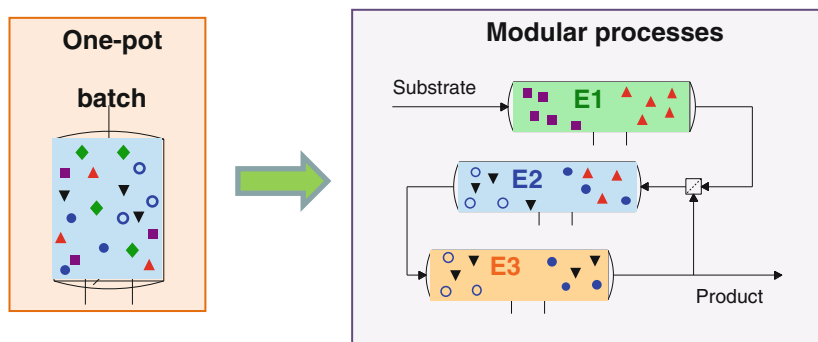
- Same conditions for the biocatalysts
- Crossover and feedback inhibitions
- Limited enzyme solubility
- No reuse of enzymes

Modular processes

Substrate → E1

E2

E3 → Product

- Biocatalysts at optimal conditions for each
- Feed-back inhibitions mitigated
- Higher enzyme loading and stability
- Well suitable for continuous process

Fig. 3 Scheme of advantages of modular processes compared to one-pot processes

An alternative to the use of soluble enzymes is to immobilize the enzymes in solid carriers. Enzyme immobilization allows an easy separation of the enzymes from the reaction mixture to be used in subsequent cycles and it also contributes to enhance the enzyme stability by maintaining the enzyme structure in the active conformation for a longer time. In the case of multienzymatic systems, enzymes can be immobilized in different carriers or co-immobilized in the same carrier for being used in a batch reactor or in a packed bed reactor. However, these systems still suffer from other limitations of the one-pot process, such as the need for compatible process conditions and the possibility of feedback or cross-over inhibitions [16].

An interesting alternative to overcome the limitations of the one-pot process is to have the enzymes physically separated in different compartments (Fig. 3). This approach is inspired by the compartmentalization of natural cells [40, 41], in which different metabolic processes are performed in different compartments such as mitochondria, lysosomes, or the endoplasmic reticulum. Different operation conditions (temperature, pH, ionic strength) can be maintained in the process compartments according to the optimal conditions of the involved enzymes. Another advantage is in the case of different enzyme stabilities of the enzymes in the multienzymatic system because the inactive reaction module can be easily replaced without affecting the other reaction modules of the process. A few examples of multienzymatic systems performed in separate reactors have been reported in literature. One example is the use of two membrane reactors, which were placed in series to physically separate the oxidative and the reductive enzymes in an enzymatic cascade for the production of 12-ketoursodeoxycholic acid from cholic acid [42]. Another example is the use of immobilized acid

phosphatase and RAMA or a thermophilic rhamnulose-1-phosphate aldolase in three packed-bed reactors in series for the synthesis of complex carbohydrates analogues from dihydroxyacetone [43].

The selection of the most suitable process configuration in multienzymatic systems increases in complexity with the number of enzymes involved: for a multienzymatic system involving n reaction steps, the theoretical number of possible process configurations is 2^{n-1} [16]. For that reason, having detailed and reliable data for the reactions (kinetic information, stability data) is very advisable in order to in silico simulate and screen the different process configurations, thus saving time, effort, and costs [16]. Microreactors (reactors that handle microliter and nanoliter volumes) are other very advantageous alternatives for process screening with reduced costs in reactants, enzymes, and resources. They also allow for an increased production by simply numbering up instead of the conventional process of scale-up [44].

4 Multienzyme Microreactors

Microreactors are usually defined as miniaturized reaction devices with dimensions in the millimeter, micrometer, and nanometer range [45]. This definition mainly encompasses microfluidic devices but, in some cases, self-organized systems such as reverse micelles, liposomes, and microemulsions are also considered as microreactors as well [45]. Due to their small size, they present a larger surface-to-volume ratio and better defined reaction times than conventional reactors [46]. The possibility of handling small amounts of reactants (microliter or nanoliter volumes) allows the achievement of high efficiency and outstanding repeatability and reproducibility [7, 47]. For that reason—and for the considerable reduction of costs that the use of small volumes entails—microreaction technologies have been exploited with success for the screening of reaction conditions and process development [44]. Additional advantages of microreactors in chemical reactions, compared with conventional ones, are an enhanced heat and mass transfer and a defined laminar flow, which allows a better control of the reaction conditions [7]. Extraction and chemical catalytic processes are favored in microreactors due to the much larger surface-to-volume ratio. From an industrial point of view, the use of microreactors presents many benefits: faster technology transfer from process development into production, earlier start of production at lower costs, easier scale-up of production capacity, smaller equipment size, lower costs for transportation, materials, and energy, and more flexible response to market demands [48].

Microfluidic devices are generally classified into two types: chip-type microreactors and microcapillary devices [7]. Chip-type reactors offer several advantages, such as easy control of the microfluidics and the possibility of integrating different process-units at the microscale in the same device. This characteristic of chip-type microreactors has been exploited for the development of bioanalytical devices [45] or for the cell disruption for separation of undamaged

Fig. 4 Scheme of different
types of enzyme retention
inside microreactors
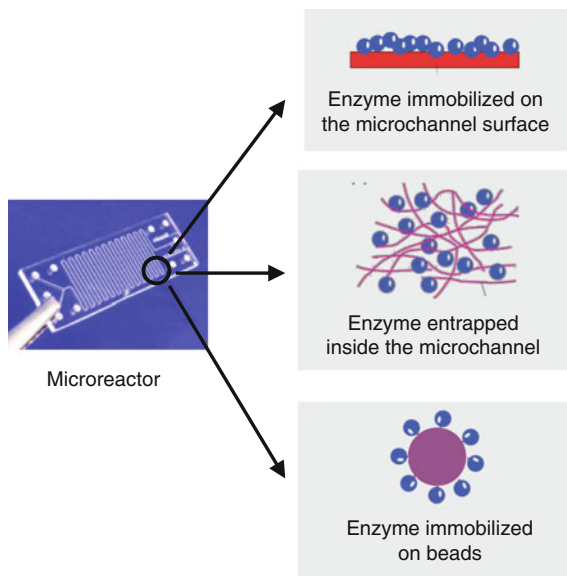
cell organelles [49]. Microreactors of the microcapillary type consist of a single microchannel where the reaction takes place. The main advantage is an easy scale-up of the process by increasing the number of microcapillaries assembled [7].

Microreactors allow better control of the flow pattern [7]. Some methods have been developed to stabilize multiple laminar flows in chip-type reactors, such as guide structures at the bottom of the microchannels [50] or partial surface modification [51–53], as well as to provide a good mixing on the system by breaking the laminar flow in different types of micromixers such as electrokinetic mixing [54], microbeads [55], chaotic mixers [56], separated serpentine flow path [57], and zig-zag microchannels [58].

The use of enzymes in microfluidic applications is challenging due to the need to maintain the enzyme structure to preserve its activity when attached to different surface geometries and chemistries at the micro/nanoscale [59]. Different techniques have been developed for the retention of enzymes in microfluidic devices, such as immobilization on beads of small dimensions, enzyme entrapment on different matrices, and immobilization on the microchannel (Fig. 4). Table 1 shows a series of examples of enzyme microreactors. Glucose oxidase, horseradish peroxidase, and alkaline phosphatase have been extensively used as model enzymes for the development of microreactors either in soluble form, immobilized on beads, or attached to the microreactor walls [45, 47].

The number of multienzymatic processes performed in microfluidic devices in the literature is growing. A major opportunity of multienzyme reactions in microreactor systems is the ability to perform in vitro metabolic pathway manipulation for high-throughput bioconversions [47]. Another feature addressed is that spatially controlled temperature regions and/or gradients according to the different

Table 1 Enzyme-immobilization techniques used in enzyme microreactors (*Source* [6])

I. Immobilization by particle entrapment

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Glass	Cross-linking (3-aminopropylsilane/ glutaraldehyde)	Xantin oxidase Horseradish peroxidase	Ease in preparation Enable multistep reaction Limited number of enzymes are applicable due to denaturation Pressure gain	[60]
Polystyrene	Biotin-Avidin (Avidin-coated beads were used)	Horseradish peroxidase	Ease in preparation Enable multistep reaction Biotin-label is required Pressure gain	[55]
Agarose	Complex formation (Ni–NTA and His-tag)	Horseradish peroxidase	Ease in preparation Applicable for engineered enzymes Higher pressure by increasing flow rate and particles may be crushed	[61]
Polystyrene	Complex formation (Ni–NTA and His-tag)	Glucose oxidase	Ease in preparation Applicable for engineered enzymes Higher pressure by increasing flow rate and particles may be crushed	[62]
Magnetic bead	Cross-linking (3-aminopropylsilane/ glutaraldehyde)	Bacterial P450	Preparation is easy Enzyme can be immobilized on any place by placing a magnet Amount of enzyme particle is limited because of plugging	[59, 60]

(continued)

Table 1 (continued)

I. Immobilization by particle entrapment

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Polymer monolith	Entrapment(2-vinyl-4,4-dimethylazlactone, ethylenedimethacrylate, 2-hydroxyethyl methacrylate, acrylamide)	Benzaldehyde liase	Stabilization of enzyme structure and activity Requirement of skill in preparation Denaturation during entrapment process	[63]
Silica monolith	Entrapment within porous silica	p-Nitrobenzyl esterase	Stabilization of enzyme structure and activity Compatibility in organic solvent Requirement of skill in preparation Denaturation possible during entrapment process	[52–54]
Aluminium oxide	Cross-linking (3-aminopropylsilane/glutaraldehyde)	Glucose oxidase	Large surface area due to porous nature Applicable for heterogeneous reactions Complicated preparation Not applicable for large-scale processing	[64]
Porous polymer monolith	Multistep photografting	Trypsin LysC	Eliminate nonspecific adsorption of proteins and peptides	[65]
Convective Interaction Media (CIM)-disk epoxy monolith	Entrapment within monolith	Glycosyltransferases	CIM® Epoxy Disk Monolithic Column is available for purchase	[66]
Caged mesoporous silica in Ca-alginate fiber	Entrapment within amine- modified mesoporous silica	Glucose oxidase	Reduced leakage and improved activity and stability of the immobilized enzyme	[67]

(continued)

Table 1 (continued)

I. Immobilization by particle entrapment

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Low-temperature co-fired ceramics multilayer substrates	Cross-linking (Glyoxal-agarose gels)	β-galactosidase	Stable operation for 6 months	[68]

II. Immobilization on microchannel surfaces

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
SiO2 surface	Physical adsorption of biotinylated poly-lysine/biotin-avidin	Alkaline phosphatase	Ease in preparation Requirement for avidin-conjugation Possible occurrence of detachment	[69]
Polydimethylsiloxane (PDMS) (O₂ plasma treated)	Physical adsorption of lipid bilayer/biotin-avidin	Alkaline phosphatase	Enable immobilization of enzyme on plastic surface Possible occurrence of detachment Expensive reagents Requirement for avidin-conjugation	[70]
PDMS	Physical adsorption of fibrinogen/ Photochemical reaction of fluorescein- biotin	Alkaline phosphatase	Enable partial modification of microchannel Special equipment is required	[71]
Silicon	Cross-linking (3-aminopropylsilane/ glutaraldehyde)	Trypsin	Simple operation Difficulty in channel preparation Poor reproducibility	[72]

(continued)

Table 1 (continued)

II. Immobilization on microchannel surfaces

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Fused silica (sol–gel modified)	Cross-linking (3-aminopropylsilane/ glutaraldehyde)	Cucumisin Lipase L-Lactic dehydrogenase	Simple operation Immobilizes ~10 times more enzymes than single-layer immobilization and therefore performs with higher reaction efficiency Several chemistry is available (amide, disulfide, His-tag) Needs several steps for immobilization Reproducibility strongly affected by characteristics of silica surface	[64–67]
Poly(methyl methacrylate)	Cross-linking (Si–O bond between modified surface and silica monolith)	Trypsin	Stabilize enzyme under denaturation condition Complicated preparation method	[73]
PDMS (O₂ plasma treated)	Cross-linking (Si–O-Ti or Si–O-Al bond between titania or alumina monolith)	Trypsin	Stabilizes enzyme under denaturation condition Complicated preparation method	[74]
Poly(ethylene terephthalate) microchip	Entrapment within nanozeolite-assembled network	Trypsin	Large surface/volume network by layer-by-layer technique	[75]
Silicon rubber	Cross-linking (3-aminopropyltrieth-oxysilane and glutaraldehyde)	Thermophilic β-glycosidase	Reaction can be performed at 80 °C Complicated preparation method Reaction is slow because not much enzyme can be immobilized	[76]

(continued)

Table 1 (continued)

II. Immobilization on microchannel surfaces

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Fused silica	Cross-linking between physically-immobilized Silica particle (3-aminopropylsilane/succinate)	Lipase	Much larger surface area (1.5 times greater than sol–gel modified surface) and higher efficiency Complicated preparation method Unstable withed physical force (bending etc.)	[77]
SiO₂ nanospring	Disulfide bond	β-galactosidase	High solvent-accessible surface area permeability and mechanical stability Repeatability of re-immobilization was poor	[78]
Photopatternin g onto poly(ethylene glycol)-grafted surface	Cross-linking by photo-patterned vinylazlactone	Horseradish peroxidase Glucose oxidase	Reduced nonspecific absorption Sequentially multistep reaction could be achieved Requires special equipment	[8]
PDMS	Entrapment within hydrogel formed on surface	Alkaline phosphatase Urease	Quite fast reaction (90 % conversion at 10 min reaction) Immobilization of multiple enzyme Complicated preparation method Not applicable for higher flow rate	[79]

III. Immobilization on membranes

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
PDMS/Glass	Placement of polyvinyl difluoride membrane that adsorbs enzymes	Trypsin	Easy preparation Less efficiency Possibility of leakage at higher flow rate	[80]

(continued)

Table 1 (continued)

III. Immobilization on membranes

Media	Immobilization method	Enzyme	Advantage and disadvantage	References
Glass	Covalent cross-linking with Nylon membrane formed at liquid–liquid interface (glutaraldehyde)	Horseradish peroxidase	Integration of membrane permeation\ and enzyme reaction Preparation of multiple membrane Complicated preparation method Unstable membrane at higher flow rat	[79]
PTFE	Enzyme-embedded membrane formation using glutaraldehyde/paraformaldehyde	α-Chimotrypsin Trypsin α-Aminoacylase Other various enzymes	Easy preparation Durable (>40 days) Applicable in organic solvents Almost all enzymes can be immobilized by adding poly-Lys	[78–80]

enzyme optima may be better controlled in microfluidic devices. The ability to rapidly evaluate the effects of reaction conditions and synthetic multienzyme pathways on a microfludic platform provides a new paradigm for performing metabolic pathway engineering—namely, the reconstruction of pathways for use in new compound discovery [82].

Sequential reactions of two- and three-coupled enzymes have been reported. A synthetic metabolic pathway consisting of a type III polyketide synthase known as 1,3,6,8-tetrahydroxynaphthalene (THN) synthase from *Streptomyces coelicolor* and soybean peroxidase was carried out in a microfluidic platform [82]. THN synthase was immobilized to Ni–NTA agarose beads prepacked into a microfluidic channel, whereas the second enzyme was covalently attached to the walls of a second microfluidic channel precoated with a reactive poly(maleic anhydride) derivative. Novel polyketide derivatives such as flaviolin and biflavolin were synthesized with a yield of 40 %.

In another application, enzymes were immobilized into modified glass microfluidic channels to form bi- and trienzymic systems [59]. Glass microfluidic channels were modified using a highly reactive poly(maleic anhydre-alt-α-olefin) (PMA)-based coating that reacts with the free amino groups of the enzymes to generate a mixed covalent-noncovalent immobilization support. Soybean peroxidase was coupled with *Candida antarctica* lipase in a bienzymic system or with invertase from *Saccharomyces cerevisiae* and glucose oxidase from *Aspergillus niger* to form a trienzymic system. In the sequential bienzymic system, the lipase was used to catalyse the hydrolysis of *p*-tolyl acetate to yield *p*-cresol, with the substrate of the reaction catalyzed by the peroxidase, whereas the three-enzyme system was aimed at generating the cosubstrate of the peroxidase, H_2O_2, by the production of glucose molecules from sucrose catalyzed by invertase and subsequent reaction with oxygen to yield H_2O_2 catalyzed by glucose oxidase. Another example for a two-enzyme system with surface-immobilized enzymes consisted on immobilized avidinD-conjugated glucose oxidase and streptavidin-conjugated horseradish peroxidase linked in series to biotinylated phospholipid bilayers coated inside poly(dimethylsiloxane) microchannels [70].

Another interesting example of compartmentalization in microfluidic devices is the local separation of enzymes on the surface of the microchannel by photopatterning [8]. Invertase, horseradish peroxidase, and glucose oxidase were patterned on porous polymer monoliths in separate regions of a single channel. All possible arrangements of the three enzymes were tested, only observing product formation when the enzymes were placed in the correct sequential order. According to these results, photopatterning enzymes on polymer monoliths were claimed as a simple technique for preparing spatially located multienzyme microreactors for directional synthesis.

An immobilized enzyme microreactor for the screening of multistep bioconversions was developed and tested for the characterization of a de novo transketolase-ω-transaminase pathway to synthesize chiral amino alcohols [83]. The prototype dual microreactor was composed by two surface-derivatized capillaries connected in series that contain the His6-tagged enzymes immobilized by

reversible attachment via Ni-NTA linkage. Transketolase was employed in the first reaction to catalyze the production of a chiral ketone from an achiral substrate. In a second reaction, a transaminase was used for the catalysis of the biocatalytic addition of an amine to the previous product to generate chiral amino alcohols, such as 2-amino-1,3,4-butanetriol. The authors claimed that this microreactor configuration presents a high utility for in vitro multistep pathway evaluation.

A chemoenzymatic pathway for the synthesis of 2-aminophenoxazin-3-one (APO) from nitrobenzene was also performed in a microfluidic device. In this cascade, zinc catalyzes the initial reduction of nitrobenzene to hydroxylamino-benzene, which undergoes a biocatalytic conversion to 2-aminophenol followed by enzymatic polymerization to APO. Individual microfluidic chips containing metallic zinc, silica-immobilized hydroxylaminobenzene mutase, and silica-immobilized soybean peroxidase connected in series were developed and proven to be suitable for carrying out the chemoenzymatic cascade.

The detection of compounds and operation parameters in microfluidic systems remains a challenge, although great progress has been made in the last few years to adapt many of the common analytical detection techniques to the microreactor scale [47]. An integrated microfluidic chip combined with mass spectrometry (matrix assisted laser desorption/ionization-mass spectrometry) was developed for parallel and multiple analyses with attractive features such as an inherent simplicity, low sample consumption, and high sensitivity [84]. This device was tested with a two-enzyme system formed by aspartate aminotransferase and ω-amino-transferase to produce an aliphatic (S)-amino acid and an (R)-amine using aspartate, α-ketoglutarate, and α-methylbenzylamine as substrates.

The aforementioned examples of multienzyme systems performed in micro-fluidic devices show the potential of this technology for pathway development and process configuration screening and set the basis for further development of high-throughput processes with an easy and fast scale-up to industrial scale.

5 Co-immobilization of Multienzyme Systems in Nano/Microsized Materials

One of the most promising novel immobilization techniques is the enzyme immobilization on nano/microsized particles and structures. The use of specific carriers allows the achievement of very high specific loadings while contributing to reduce the mass-transfer limitations. Additionally, an enhancement of the enzyme stability such as during operation, storage, and recycling has been reported for many systems [85]. In the case of coimmobilization of different enzymes of a multienzymatic system, the immobilized system could elevate the local concentration of the intermediate products and substantially shorten the distance the intermediates need to travel from one active site to another active site of the enzymes.
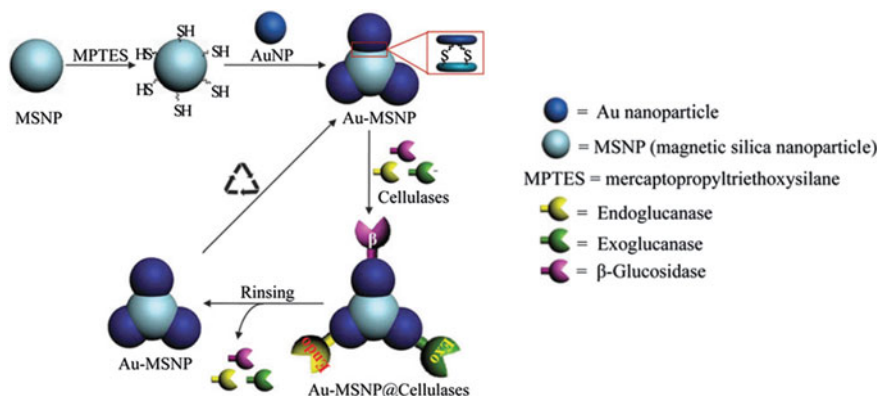
Fig. 5 Example of random co-immobilization of cystein-tagged cellulases (endoglucanase, exoglucanase and β-glucosidase) on gold nanoparticles and gold-magnetic silica nanoparticles (*Source* [88])

Multienzyme immobilization can be designed via random coimmobilization [86–92] and compartmentalization [93–99]. Random coimmobilization is the simplest way. Generally, mixed enzyme solutions are added with a specific carrier and then classical immobilization methods, such as physical adsorption, covalent attachment, or cross-linking, are applied. Of course, according to the reaction rate or pathway, the loading of enzymes and their ratios need to be controlled and optimized. A representative example is shown in Fig. 5. In this study, three cystein-tagged cellulases are coimmobilized on gold nanoparticles and gold-magnetic silica nanoparticles (Au-MSNP) particles for the hydrolytic degradation of cellulose [88]. After treatment with cellulases coimmobilized on Au-MSNP, the yields of cellobiose and glucose increased by 158 and 179 %, respectively. This demonstrates an enhancement in enzyme activity as a result of coimmobilization on Au-MSNP.

The cellular environment can be regarded as highly organized and complex compartments in which multienzyme cascade reactions can take place simultaneously with unsurpassed efficiency and specificity. To mimic natural enzyme organization to create delicate microenvironments, immobilization by compartmentalization has been applied. For example, several innovative materials, such as polymer capsules, phospholipid liposome vesicles, and polymersomes were developed and used with specific sol–gel chemistry and layer-by-layer deposition methods [100]. Until now, polymersome-based nanoreactors or layer-by-layer (LBL) based bioinspired biomineral microcapsules were developed and reported consistently with several cascade model enzyme reactions, as shown in Fig. 6.

Glucose oxidase and horseradish peroxidase have been widely used as model enzymes for the development of different coimmobilization methods via layer-by-layer, entrapment on polymersomes, encapsulation, and attachment to nanoparticle surfaces (Table 2). A layer-by-layer approach was used to immobilize these enzymes on magnetic particles, allowing an easy and rapid separation from the
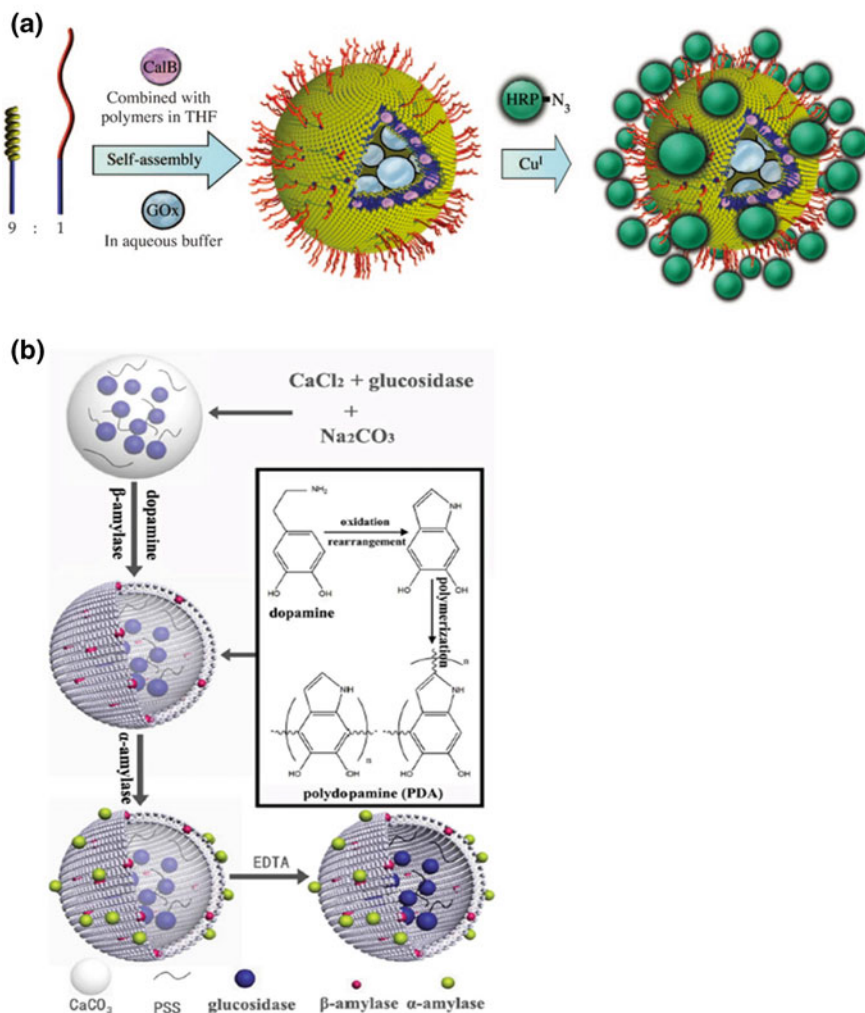
Fig. 6 Examples of enzyme immobilization by compartmentalization. **a** Polymersome nanore-actors based on block copolymers of isocyanopeptides and styrene with immobilized glucose oxidase in their lumen, *Candida antarctica* lipase B in their bilayer membrane, and horseradish peroxidase on their surface (*Source* [98]). **b** Layer-by-layer based bioinspired biomineral polydopamine microcapsule with α-amylase, β-amylase and glucosidase immobilized through physical encapsulation in the lumen, in situ entrapment within the wall, and chemical attachment on the out surface under extremely mild conditions. Abbreviations: GOx, glucose oxidase; HRP, horseradish peroxidase. (*Source* [99])

media by a magnet [87]. Micrometer-sized shell-in-shell polyelectrolyte capsules containing the enzymes of the cascade were also produced [101]. Improved overall reaction rates (a 2.5-times higher rate than free enzymes) were obtained by coimmobilization of enzymes in separated polymer layers on the surface of silica

Table 2 Examples of immobilization of multienzyme systems in nanoparticles and nanostructures

I. Glucose oxidase/Horseradish peroxidase

Type of cascade	Enzymes involved	Immobilization approach and materials	Reference
Glucose oxidase/horseradish peroxidase	Glucose oxidase (GOD) and horseradish peroxidase (POD)	Coimmobilization of enzyme on polystyrene (PS) particles (470 or 640 nm) by adsorption and layer-by-layer (LbL) coating layer	[87]
		POD was associated with the membrane and GOD was entrapped within the water pool of the PS–polyisocyanopeptide polymersomes (517 nm)	[101]
		Coencapsulation: GOD is encapsulated in the outer surrounding compartment and POD in the inner capsule of shell-in-shell microcapsules. The ball-in-ball particles had a total size ranging from 8–10 μm with an inner core diameter of 3–4 μm.	[95]
		Coimmobilization of enzyme in seperated polymer layer on the surface of silica microparticles (3 μm) by LbL film assembly	[91]
		Coimmobilization of enzyme on Au–PB–Fe_3O_4 three-layer composite magnetic particles (80 nm) by covalent attachment	[92]
		Coencapsulation: A nanocontainer composed of a bilayer of POD giant amphiphiles containing glucose oxidase in the inner aqueous compartment	[94]
		Coimmobilization of enzymes in glutaraldehyde (GA)-crosslinked bovine serum albumin (BSA) in the spacing compartments $CaCO_3$ microparticles (2–5 μm)	[86]
		Coimmobilization of enzyme on block copolymer-quantum dot micelles (145 nm) by adsorption	[89]
		Coimmobilization of enzyme in macroporous silica foam by adsorption	[110]
		Coimmobilization of enzyme on polystyrene nanoparticles with carboxyl groups on the surface by covalent attachment	[111]

II. Other bienzymatic systems

Type of cascade	Enzymes involved	Immobilization approach and materials	Reference
Synthesis of phydroxyphenylglycine (optically pure amino acids)	D-hydantoinase and N-carbamyl-D-amino acid amidohydrolase	Physical adsorption and covalent linking of cell free extract on chitin	[103]
Direct conversion of cephalosporin C to 7-aminocephalosporanic acid	D-amino acid oxidase was co-immobilized with catalase	Covalent immobilization on glyoxyl-agarose and poly(ethylene imine)-agarose	[104]

(continued)

Table 2 (continued)

II. Other bienzymatic systems

Type of cascade	Enzymes involved	Immobilization approach and materials	Reference
Sequential enzymatic reactions	Acetylcholinesterase and choline oxidase	Covalent immobilization on polystyrene core nanoparticles with a phospholipid polymer surface (300–500 nm)	[105]
Sequential enzymatic reactions	Glucose oxidase and lignin peroxidase	Coimmobilization on nanoporous gold (40–50 nm) via physical adsorption	[111]
Metabolic pathway	Malate dehydrogenase and citrate synthase	Coimmobilization on gold nanoparticles (30 nm) via physical adsorption	[106]
Sequential enzymatic reactions	Superoxide dismutase and lactoperoxidase	Coencapsulation of enzymes in polymeric nanocontainers (95 nm)	[97]
An immunosensor based on the electrochemiluminescence of peroxydisulfate	Horseradish peroxidase and glucoseoxidase	Coimmobilization of enzyme via physical adsorption on palladium nanoparticles attached to functional carbon nanotubes	[90]

III. Cofactor regeneration multienzymatic cascades

Type of cascade	Enzymes involved	Immobilization approach and materials	References
An NAD +/NADH recycling system for the synthesis of L-glutamate and lactate	Lactate dehydrogenase and glutamate dehydrogenase	Co-immobilization of two enzyme on Amberlite XAD-7 via physical adsorption (polymer beads)	[107]
Cofactor regeneration	Glutamate dehydrogenase, lactate dehydrogenase	Coimmobilization of two enzymes on silica nanoparticles (30 nm) by covalent attachment	[108]
Cofactor regeneration	Lactate dehydrogenase, glucose dehydrogenase	Coimmobilization of two enzymes in porous spherical glass particles (60 μm) via covalent attachment	[112]
Design of combined nanoscale and microscale for cofactor regeneration	Lactate dehydrogenase, glucose dehydrogenase	Co-immobilization on PS nanoparticles (109–114 nm) and then mixing in microcapsule (3–10 μm)	[113]
Continuous conversion of nitrobenzene to hydroxylaminobenzene with NADPH recycling	Nitrobenzene nitroreductase and glucose-6-phosphate dehydrogenase	Coencapsulatation in poly(ethylene imine)-induced silica particles	[93]
Simultaneous production of 1,3-dihydroxyacetone and xylitol with in situ cofactor regeneration	Glycerol dehydrogenase, xylose reductase	Coimmobilization of two-enzyme P(MMA–EDMA–MAA) nanoparticles (160 nm) by covalent attachment	[114]

(continued)

Table 2 (continued)

III. Cofactor regeneration multienzymatic cascades

Type of cascade	Enzymes involved	Immobilization approach and materials	References
Enzymatically catalysed Baeyer–Villiger reactions with cofactor regeneration	Baeyer–Villiger monooxygenases and glucose-6-phosphate dehydrogenase	Encapsulation of enzyme polymersome nanoreactors with immobilized enzyme onto surface of polymersomes	[96]

IV. CO$_2$ reduction systems

Type of cascade	Enzymes involved	Immobilization approach and materials	Reference
CO$_2$ to methanol	Formate dehydrogenase (FateDH), formaldehyde dehydrogenase (FaldDH), and alcohol dehydrogenase (ADH)	Coencapsulation *in silica* sol–gel matrices	[115]
		Coencapsulated in an alginate-silica (ALG-SiO$_2$) hybrid gel	[116]
		Coimmobilization of three enzymes on polystyrene particles (500 nm) via covalent attachment	[117]
		Coencapsulation of three dehydrogenases within titania particles through biomimetic mineralization process (400–600 nm)	[118]
		Coencapsulation of three enzymes in protamine–silica hybrid microcapsules (3 μm)	[119]
		Coencapsulation of three enzymes into a hybrid matrix made of alginate and silicate	[120]
		LbL self-assembly on the protamine-soaked enzyme (the first enzyme)-containing CaCO$_3$ microspheres (4 μm)	[109]
CO$_2$ to formaldehyde	Formate dehydrogenase (FateDH), formaldehyde dehydrogenase (FaldDH)	FateDH is entrapped accompanying the formation of titania nanoparticles (NPs) FaldDH is immobilized on the functionalized surface of NPs	[121]

(continued)

Table 2 (continued)

V. Other multienzyme systems

Type of cascade	Enzymes involved	Immobilization approach and materials	References
Sequential enzymatic reactions	Lipase, glucose oxidase, and horseradish peroxidase	Immobilized glucose oxidase in their lumen, *Candida antarctica* lipase B in their bilayer membranes, and horseradish peroxidase on their surface	[98]
Sequential enzymatic reactions	β-glucosidase, glucose oxidase, and horseradish peroxidase	Coimmobilization of enzymes in GA-cross-linked BSA in the spacing compartments CaCO$_3$ microparticles (2–5 μm)	[86]
In vitro 1, 3-popanediol (1, 3-PDO) production	Glyceroldehydratase, 1,3-propanediol-oxidoreductase	Coimmobilization on TiO$_2$/enzyme/ PAMhybrid materials (500 ∼ 800 nm) by adsorption	[122]
Degradation of cellulose	Endo-glucanase, exo-glucanase, and β-glucosidase	Coimmobilization of three cellulases on Au-doped magnetic silica nanoparticles (25–27 nm) by covalent attachment	[88]
Converting starch to isomaltooligosaccharide (IMOs)	α-amylase, β-amylase, and glucosidase	Three enzymes are respectively immobilized through physical encapsulation in the lumen, in situ entrapment within the wall, and chemical attachment on the out surface of the hybrid microcapsules (3–4 μm)	[123]
		The three enzymes are respectively immobilized through physical encapsulation in the lumen, in situ entrapment within the wall, and chemical attachment on the out surface of polydopamine (PDA) microcapsule (3–5 μm)	[99]

microparticles of 3 μm [91]. A more drastic increase of reaction rates (100 % increase) was obtained by coimmobilization of these enzymes on polystyrene nanoparticles to form a biomimetic dual-functionalized nanoparticle-based platform [102] and, in the case of coimmobilization on Block copolymer-quantum dot micelles of 145 nm, via physical adsorption [89]. Glucose oxidase and horseradish peroxidase were also coimmobilized for biosensor applications following immobilization by adsorption in macroporous silica foams [89] and by covalent attachment to Au-PB-Fe_3O_4 three-layer composite magnetic nanoparticles [92].

Other bienzymatic systems have been also developed using hydrolases, oxidases, catalase, hydratases, and synthases (Table 2). Enhanced enzyme stabilities were obtained by covalent attachment on chitin, agarose, and polystyrene particles due to three different causes: (i) higher enzyme activities at lower temperatures, which allows reducing the operation temperature with a subsequent increase of the operational stability [103]; (ii) removal of the H_2O_2 produced by a coupled enzyme, which negatively affects the activity of an oxidase [104]; and (iii) a stabilizing effect of the covalent attachment to the particle [105]. Different bienzymatic systems have been immobilized by nanoporous gold, palladium nanoparticles, polymeric nanocontainers, and gold nanoparticles. In the case of the immobilization on the surface of nanoparticles, the organization and ratio of activity of the enzymes is important for the sequential rate of the reaction [106].

Different approaches have been developed that use multienzymatic systems for cofactor regeneration (Table 2). Lactate dehydrogenase and glutamate dehydrogenase for the regeneration of NADH were coimmobilized on Amberlite via physical adsorption obtaining high reuse stability [107] and on silica nanoparticles via covalent attachment with high cofactor turnover numbers [108]. These high turnover numbers were obtained by particle–particle collision driven by the Brownian motion of the nanoparticles containing the enzyme and the nanoparticles containing the immobilized cofactor. Lactate dehydrogenase was coupled with glucose dehydrogenase in porous spherical glass particles and on polystirene nanoparticles inside a microcapsule. In the first case, the use of long spacers between the enzyme and the surface of the glass particle enhance the reaction rate. In the second case, a multilevel design was performed to match reaction and mass transfer rates in order to maximize productivity. NADPH regeneration was achieved by nitrobenzene nitroreductase and glucose-6-phosphate dehydrogenase co-encapsulation in poly(ethylene imine)-induced silica particles, with up to 125-fold higher product formation compared with a noncoupled system. Improved productivity and stability was obtained as well in the immobilization of glycerol dehydrogenase and xylose reductase on affinity magnetic nanoparticles.

An active research field on the coimmobilization of multienzymatic systems focuses on the reduction of CO_2. A bienzymatic system composed of formate dehydrogenase and formaldehyde dehydrogenase catalyzes the reduction of CO_2 to generate formaldehyde. Further coupling with alcohol dehydrogenase results in the formation of methanol as the main product of the CO_2 reduction. Different approaches have been developed for the coimmobilization of these enzymes on nanoparticles and nanostructures (Table 2). An overall equilibrium shifting to the

product formation has been observed after immobilization of the three-enzyme system via coencapsulation in silica sol–gel matrices and after immobilization of the two-enzyme systems via layer-by-layer self-assembly, probably as a result of the substantially reduced diffusion distance for the reaction intermediate between the active sites of the enzymes [109].

Other three-enzyme multienzymatic systems have been coimmobilized for various applications (Table 2). Good enzyme reusability and stability was observed in many of the examples as well, highlighting the feasibility of the coimmobilization technology for carrying out multienzymatic cascades of up to three enzymes. However, space and mass-transfer limitations hinder the applicability of this technology for multienzymatic pathways involving a high number of enzymes.

6 Conclusion

Multienzymatic systems present a great potential for the biocatalytic synthesis of many compounds and can overcome the major problems of limited product yields and operational complexity of many fermentation processes. Metabolic pathways can be mimicked and studied in vitro by proper selection and combination of purified or semipurified enzymes. Furthermore, completely new or synthetic bioreaction pathways designed de novo can be studied in vitro to examine their feasibility and determine optimization. However, as the number of enzymes increases, so does the complexity of the process design. Performing multienzymatic cascades in a one-pot system is the easiest process configuration, but it presents several limitations. Enzymes with compatible optima process conditions should be selected. The presence of crossover and feedback inhibitions can also hamper the product formation and cause low production yields. Distribution of the enzymes in physically separated compartments is a process alternative that has lead to higher productivities in a number of multienzymatic systems. Microfluidic devices allow the creation of enzyme patterns inside miniaturized reactors and, as such, they are useful tools for the design of multienzymatic pathways and screening of process configurations. Finally, coimmobilization of multienzymatic systems with a low number of enzymes is a promising alternative as well. This technique allows the reduction of the diffusion path of the intermediates between the active site of the enzymes, contributing to the increase of the reaction rates.

References

1. Cornish-Bowden A (2004) Fundamentals of enzyme kinetics. Portland Press, London
2. Hodgman CE, Jewett MC (2012) Cell-free synthetic biology: thinking outside the cell. Metab Eng 14:261–269

3. Lopez-Gallego F, Schmidt-Dannert C (2010) Multi-enzymatic synthesis. Curr Opin Chem Biol 14:174–183
4. Findrik Z, Vasić-Rački Đ (2009) Overview on reactions with multi-enzyme systems. Chem Biochem Eng Q 23:545–553
5. Santacoloma PA, Sin G, Gernaey KV, Woodley JM (2010) Multienzyme-catalyzed processes: next-generation biocatalysis. Org Process Res Dev 15:203–212
6. Ricca E, Brucher B, Schrittwieser JH (2011) Multi-Enzymatic cascade reactions: overview and perspectives. Adv Synth Catal 353:2239–2262
7. Asanomi Y, Yamaguchi H, Miyazaki M, Maeda H (2011) Enzyme-immobilized microfluidic process reactors. Molecules 16:6041–6059
8. Logan TC, Clark DS, Stachowiak TB, Svec F, Fréchet JMJ (2007) Photopatterning enzymes on polymer monoliths in microfluidic devices for steady-state kinetic analysis and spatially separated multi-enzyme reactions. Anal Chem 79:6592–6598
9. Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Mol Syst Biol 2(2006):0028
10. Forster AC, Church GM (2006) Synthetic biology projects in vitro. Genome Res 17:000
11. Keasling JD (2008) Synthetic biology for synthetic chemistry. ACS Chem Biol 3:64–76
12. Chenault HK, Simon ES, Whitesides GM (1988) Cofactor regeneration for enzyme-catalysed synthesis. Biotechnol Genet Eng Rev 6:221–270
13. Lauterbach L, Lenz O, Vincent KA (2013) H_2-driven cofactor regeneration with NAD(P)$^+$-reducing hydrogenases. FEBS J, n/a-n/a.
14. Liu W, Wang P (2007) Cofactor regeneration for sustainable enzymatic biosynthesis. Biotechnol Adv 25:369–384
15. Wichmann R, Vasic-Racki D (2005) Cofactor regeneration at the lab scale. Springer, Berlin
16. Xue R, Woodley JM (2012) Process technology for multi-enzymatic reaction systems. Bioresour Technol 115:183–195
17. Zhao H, van der Donk WA (2003) Regeneration of cofactors for use in biocatalysis. Curr Opin Biotechnol 14:583–589
18. Findrik Z, Vasić-Rački Đ (2008) Mathematical modelling of amino acid resolution catalyzed by l-amino acid oxidases from *Crotalus adamanteus* and *Crotalus atrox*. Process Biochem 43:1186–1194
19. Plumeré N, Henig J, Campbell WH (2012) Enzyme-catalyzed O_2 removal system for electrochemical analysis under ambient air: Application in an amperometric nitrate biosensor. Anal Chem 84:2141–2146
20. Srere PA, Mattiasson B, Mosbach K (1973) An immobilized three-enzyme system: a model for microenvironmental compartmentation in mitochondria. P Natl Acad Sci 70:2534–2538
21. Yun H, Yang Y-H, Cho B-K, Hwang B-Y, Kim B-G (2003) Simultaneous synthesis of enantiomerically pure (R)-1-phenylethanol and (R)-α-methylbenzylamine from racemic α-methylbenzylamine using ω-transaminase/alcohol dehydrogenase/glucose dehydrogenase coupling reaction. Biotechnol Lett 25:809–814
22. Soldatkin OO, Peshkova VM, Dzyadevych SV, Soldatkin AP, Jaffrezic-Renault N, El'skaya AV (2008) Novel sucrose three-enzyme conductometric biosensor. Mat Sci Eng C 28:959–964
23. Ishii N, Suga Y, Hagiya A, Watanabe H, Mori H, Yoshino M, Tomita M (2007) Dynamic simulation of an in vitro multi-enzyme system. FEBS Lett 581:413–420
24. Bae H-S, Lee S-G, Hong S-P, Kwak M-S, Esaki N, Soda K, Sung M-H (1999) Production of aromatic D-amino acids from α-keto acids and ammonia by coupling of four enzyme reactions. J Mol Catal B Enzym 6:241–247
25. Findrik Z, Vasić-Rački Đ (2007) Biotransformation of D-methionine into L-methionine in the cascade of four enzymes. Biotechnol Bioeng 98:956–967
26. Gijsen HJM, Qiao L, Fitz W, Wong C-H (1996) Recent advances in the chemoenzymatic synthesis of carbohydrates and carbohydrate mimetics. Chem Rev 96:443–474
27. Roessner CA, Scott AI (1996) Achieving natural product synthesis and diversity via catalytic networking ex vivo. Chem Biol 3:325–330

28. Zhang Y-HP (2010) Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: challenges and opportunities. Biotechnol Bioeng 105:663–677

29. Fessner W-D, Walter C (1992) "Artificial metabolisms" for the asymmetric one-pot synthesis of branched-chain saccharides. Angew Chem Int Edit 31:614–616

30. Itoh A, Ohashi Y, Soga T, Mori H, Nishioka T, Tomita M (2004) Application of capillary electrophoresis-mass spectrometry to synthetic in vitro glycolysis studies. Electrophoresis 25:1996–2002

31. Wendell D, Todd J, Montemagno C (2010) Artificial photosynthesis in Ranaspumin-2 Based Foam. Nano Lett 10:3231–3236

32. Zhang YHP (2009) Using extremophile enzymes to generate hydrogen for electricity. Microbe 4:560–565

33. Woodward J, Orr M, Cordray K, Greenbaum E (2000) Biotechnology: enzymatic production of biohydrogen. Nature 405:1014–1015

34. Zhang YHP, Evans BR, Mielenz JR, Hopkins RC, Adams MWW (2007) High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. PLoS ONE 2:e456

35. Zhang YHP, Huang W-D (2012) Constructing the electricity–carbohydrate–hydrogen cycle for a sustainability revolution. Trends Biotechnol 30:301–306

36. Wang Y, Huang W, Sathitsuksanoh N, Zhu Z, Zhang YHP (2011) Biohydrogenation from biomass sugar mediated by in vitro synthetic enzymatic pathways. Chem Biol 18:372–380

37. Ardao I, Zeng A-P (2013) In silico evaluation of a complex multi-enzymatic system using one-pot and modular approaches: application to the high-yield production of hydrogen from a synthetic metabolic pathway. Chem Eng Sci 87:183–193

38. Bouzas TdM, Barros-Velázquez J, González Villa T (2006) Industrial applications of hyperthermophilic enzymes: a review. Protein Pept Lett 13:645–651

39. Myung S, Wang Y, Zhang YHP (2010) Fructose-1,6-bisphosphatase from a hyper-thermophilic bacterium Thermotoga maritima: characterization, metabolite stability, and its implications. Process Biochem 45:1882–1887

40. Chen AH, Silver PA (2012) Designing biological compartmentalization. Trends Cell Biol 22:662–670

41. Jandt U, You C, Zhang YHP, Zeng AP (2013) Compartmentalization and metabolic channeling for multienzymatic biosynthesis: practical strategies and modeling approaches. Adv Biochem Eng/Biotechnol (accepted)

42. Monti D, Ferrandi EE, Zanellato I, Hua L, Polentini F, Carrea G, Riva S (2009) One-pot mutienzymatic synthesis of 12-ketoursodeoxycholic acid: Subtle cofactor specificities rule the reaction equilibria of five biocatalysts working in a row. Adv Synth Catal 351:1303–1311

43. Babich L (2013) Enzymatic cascade reactions involving phosphorylated intermediates: immobilization and process optimization. Faculty of Science, University of Amsterdam, Amsterdam (The Netherlands)

44. Zhang X, Stefanick S, Villani FJ (2004) Application of microreactor technology in process development. Org Process Res Dev 8:455–460

45. Urban PL, Goodall DM, Bruce NC (2007) Enzymatic microreactors in chemical analysis and kinetic studies. Biotechnol Adv 24:42–57

46. Miyazaki M, Maeda H (2006) Microchannel enzyme reactors and their applications for processing. Trends Biotechnol 24:463–470

47. Matosevic S, Szita N, Baganz F (2011) Fundamentals and applications of immobilized microfluidic enzymatic reactors. J Chem Technol Biotechnol 86:325–334

48. Ehrfeld W, Hessel V, Löwe H (2000) Microreactors. Wiley-VCH, Weinheim

49. Rajabi N, Hoffmann M, Bahnemann J, Zeng A-P, Schl, uuml, ter M, ller J, ouml rg (2012) A Chaotic Advection Enhanced Microfluidic Split-and-Recombine Mixer for the Preparation of Chemical and Biological Probes. J Chem Eng Jpn 45:703–707

50. Tokeshi M, Minagawa T, Uchiyama K, Hibara A, Sato K, Hisamoto H, Kitamori T (2002) Continuous-flow chemical processing on a microchip by combining microunit operations and a multiphase flow network. Anal Chem 74:1565–1571

51. Hibara A, Nonaka M, Hisamoto H, Uchiyama K, Kikutani Y, Tokeshi M, Kitamori T (2002) Stabilization of liquid interface and control of two-phase confluence and separation in glass microchips by utilizing octadecylsilane modification of microchannels. Anal Chem 74:1724–1728

52. Yamaguchi Y, Ogino K, Takagi F, Honda T, Yamashita K, Miyazaki M, Nakamura H, Maeda H (2005) Partial chemical modification of a microchannel and stabilization of water-oil phase separation, Proceedings of the 8th International Conference on Microreaction Technology (IMRET 8). American Institute of Chemical Engineers, New York

53. Zhao B, Moore JS, Beebe DJ (2002) Principles of surface-directed liquid flow in microfluidic channels. Anal Chem 74:4259–4268

54. Erickson D, Li D (2002) Influence of surface heterogeneity on electrokinetically driven microfluidic mixing. Langmuir 18:1883–1892

55. Seong GH, Crooks RM (2002) Efficient mixing and reactions within microfluidic channels using microbead-supported catalysts. J Am Chem Soc 124:13360–13361

56. Stroock AD, Dertinger SKW, Ajdari A, Mezić I, Stone HA, Whitesides GM (2002) Chaotic mixer for microchannels. Science 295:647–651

57. Kim J-H, Kin B-G, La M, Yoon J-B, Yoon E (2002) A disposssable passive microfluidic system integrated with micromixer and DNA purification chip for DNA sample preparation. Micro total analysis sytem. Dordrecht, The Netherlands, pp 224–226

58. Mengeaud V, Josserand J, Girault HH (2002) Mixing processes in a zigzag microchannel: finite element simulations and optical study. Anal Chem 74:4279–4286

59. Lee M-Y, Srinivasan A, Ku B, Dordick JS (2003) Multienzyme catalysis in microfluidic biochips. Biotechnol Bioeng 83:20–28

60. Richter T, Shultz-Lockyear LL, Oleschuk RD, Bilitewski U, Harrison DJ (2002) Bi-enzymatic and capillary electrophoretic analysis of non-fluorescent compounds in microfluidic devices: determination of xanthine. Sens Actuators B: Chem 81:369–376

61. Srinivasan A, Bach H, Sherman DH, Dordick JS (2004) Bacterial P450-catalyzed polyketide hydroxylation on a microfluidic platform. Biotechnol Bioeng 88:528–535

62. Drager G, Kiss C, Kunz U, Kirschning A (2007) Enzyme-purification and catalytic transformations in a microstructured PASSflow reactor using a new tyrosine-based Ni-NTA linker system attached to a polyvinylpyrrolidinone-based matrix. Org Biomol Chem 5:3657–3664

63. Sakai-Kato K, Kato M, Ishihara K, Toyo'oka T (2004) An enzyme-immobilization method for integration of biofunctions on a microchip using a water-soluble amphiphilic phospholipid polymer having a reacting group. Lab Chip 4:4–6

64. Heule M, Rezwan K, Cavalli L, Gauckler LJ (2003) A miniaturized enzyme reactor based on hierarchically shaped porous ceramic microstruts. Adv Mater 15:1191–1194

65. Krenkova J, Lacher NA, Svec F (2009) Highly efficient enzyme reactors containing trypsin and endoproteinase Lysc immobilized on porous polymer monolith coupled to ms suitable for analysis of antibodies. Anal Chem 81:2004–2012

66. Delattre C, Vijayalakshmi MA (2009) Monolith enzymatic microreactor at the frontier of glycomic toward a new route for the production of bioactive oligosaccharides. J Mol Catal B Enzym 60:97–105

67. Yang C, Zhang Z, Shi Z, Xue P, Chang P, Yan R (2010) Development of a novel enzyme reactor and application as a chemiluminescence flow-through biosensor. Anal Bioanal Chem 397:2997–3003

68. Baeza M, López C, Alonso Jn, López-Santín J, Álvaro G (2009) Ceramic microsystem incorporating a microreactor with immobilized biocatalyst for enzymatic spectrophotometric assays. Anal Chem 82:1006–1011

69. Gleason NJ, Carbeck JD (2004) Measurement of enzyme kinetics using microscale steady-state kinetic analysis. Langmuir 20:6374–6381

70. Mao H, Yang T, Cremer PS (2001) Design and characterization of immobilized enzymes in microfluidic systems. Anal Chem 74:379–385

71. Holden MA, Jung S-Y, Cremer PS (2004) Patterning enzymes inside microfluidic channels via photoattachment chemistry. Anal Chem 76:1838–1843

72. Ekström S, Önnerfjord P, Nilsson J, Bengtsson M, Laurell T, Marko-Varga G (1999) Integrated microanalytical technology enabling rapid and automated protein identification. Anal Chem 72:286–293

73. Qu H, Wang H, Huang Y, Zhong W, Lu H, Kong J, Yang P, Liu B (2004) Stable microstructured network for protein patterning on a plastic microfluidic channel: strategy and characterization of on-chip enzyme microreactors. Anal Chem 76:6426–6433

74. Wu H, Tian Y, Liu B, Lu H, Wang X, Zhai J, Jin H, Yang P, Xu Y, Wang H (2004) Titania and alumina sol − gel-derived microfluidics enzymatic-reactors for peptide mapping: design, characterization, and performance. J Proteome Res 3:1201–1209

75. Ji, Zhang Y, Zhou X, Kong J, Tang Y, Liu B (2008) Enhanced protein digestion through the confinement of nanozeolite-assembled microchip reactors. Anal Chem 80:2457–2463

76. Thomsen MS, Polt P, Nidetzky B (2007) Development of a microfluidic immobilised enzyme reactor. Chemical Communi 0:2527–2529

77. Nakamura H, Li X, Wang H, Uehara M, Miyazaki M, Shimizu H, Maeda H (2004) A simple method of self assembled nano-particles deposition on the micro-capillary inner walls and the reactor application for photo-catalytic and enzyme reactions. Chem Eng J 101:261–268

78. Schilke KF, Wilson KL, Cantrell T, Corti G, McIlroy DN, Kelly C (2010) A novel enzymatic microreactor with *Aspergillus oryzae* *β*-galactosidase immobilized on silicon dioxide nanosprings. Biotechnol Prog 26:1597–1605

79. Koh W-G, Pishko M (2005) Immobilization of multi-enzyme microreactors inside microfluidic devices. Sens Actuators B: Chem 106:335–342

80. Gao J, Xu J, Locascio LE, Lee CS (2001) Integrated microfluidic system enabling protein digestion, peptide separation, and protein identification. Anal Chem 73:2648–2655

81. Hisamoto H, Shimizu Y, Uchiyama K, Tokeshi M, Kikutani Y, Hibara A, Kitamori T (2002) Chemicofunctional membrane for integrated chemical processes on a microchip. Anal Chem 75:350–354

82. Ku B, Cha J, Srinivasan A, Kwon SJ (2006) Chip-based polyketide biosynthesis and functionalization. Biotechnol Prog 22:1102–1107

83. Matosevic S, Lye GJ, Baganz F (2011) Immobilised enzyme microreactor for screening of multi-step bioconversions: Characterisation of a de novo transketolase-*ω*-transaminase pathway to synthesise chiral amino alcohols. J Biotechnol 155:320–329

84. Lee S-H, Lee C-S, Kim B-G, Kim Y-K (2008) An integrated microfluidic chip for the analysis of biochemical reactions by MALDI mass spectrometry. Biomed Microdevices 10:1–9

85. Hwang ET, Gu MB (2013) Enzyme stabilization by nano/microsized hybrid materials. Eng Life Sci 13:49–61

86. Bäumler H, Georgieva R (2010) Coupled enzyme reactions in multicompartment microparticles. Biomacromolecules 11:1480–1487

87. Caruso F, Schüler C (2000) Enzyme multilayers on colloid particles: assembly, stability, and enzymatic activity. Langmuir 16:9595–9603

88. Cho EJ, Jung S, Kim HJ, Lee YG, Nam KC, Lee H-J, Bae H-J (2012) Co-immobilization of three cellulases on Au-doped magnetic silica nanoparticles for the degradation of cellulose. Chem Commun 48:886–888

89. Jia F, Zhang Y, Narasimhan B, Mallapragada SK (2012) Block copolymer-quantum dot micelles for multienzyme colocalization. Langmuir 28:17389–17395

90. Niu H, Yuan R, Chai Y, Mao L, Liu H, Cao Y (2013) Highly amplified electrochemiluminescence of peroxydisulfate using bienzyme functionalized palladium nanoparticles as labels for ultrasensitive immunoassay. Biosens Bioelectron 39:296–299

91. Pescador P, Katakis I, Toca-Herrera JL, Donath E (2008) Efficiency of a bienzyme sequential reaction system immobilized on polyelectrolyte multilayer-coated colloids. Langmuir 24:14108–14114

92. Zhuo Y, Yuan P-X, Yuan R, Chai Y-Q, Hong C-L (2009) Bienzyme functionalized three-layer composite magnetic nanoparticles for electrochemical immunosensors. Biomaterials 30:2284–2290

93. Betancor L, Berne C, Luckarift HR, Spain JC (2006) Coimmobilization of a redox enzyme and a cofactor regeneration system. Chem Commun 0:3640–3642

94. Delaittre G, Reynhout IC, Cornelissen JJLM, Nolte RJM (2009) Cascade reactions in an all-enzyme nanoreactor. Chem A Eur J 15:12600–12603

95. Kreft O, Prevot M, Möhwald H, Sukhorukov GB (2007) Shell-in-shell microcapsules: a novel tool for integrated, spatially confined enzymatic reactions. Angew Chem Int Ed 46:5605–5608

96. Meeuwissen SA, Rioz-Martinez A, de Gonzalo G, Fraaije MW, Gotor V, van Hest JCM (2011) Cofactor regeneration in polymersome nanoreactors: enzymatically catalysed baeyer-villiger reactions. J Mater Chem 21:18923–18926

97. Tanner P, Onaca O, Balasubramanian V, Meier W, Palivan CG (2011) Enzymatic cascade reactions inside polymeric nanocontainers: a means to combat oxidative stress. Chem A Eur J 17:4552–4560

98. van Dongen SFM, Nallani M, Cornelissen JJLM, Nolte RJM, Van Hest JCM (2009) A three-enzyme cascade reaction through positional assembly of enzymes in a polymersome nanoreactor. Chem A Eur J 15:1107–1114

99. Zhang L, Shi J, Jiang Z, Jiang Y, Qiao S, Li J, Wang R, Meng R, Zhu Y, Zheng Y (2011) Bioinspired preparation of polydopamine microcapsule for multienzyme system construction. Green Chem 13:300–306

100. Schoffelen S, van Hest JCM (2012) Multi-enzyme systems: bringing enzymes together in vitro. Soft Matter 8:1736–1746

101. Vriezema DM, Garcia PML, Sancho|Oltra N, Hatzakis NS, Kuiper SM, Nolte RJM, Rowan AE, Van|Hest JCM (2007) Positional assembly of enzymes in polymersome nanoreactors for cascade reactions. Angew Chem Int Ed 46:7378–7382

102. Jia F, Narasimhan B, Mallapragada SK (2013) Biomimetic multienzyme complexes based on nanoscale platforms. AIChE J 59:355–360

103. Aranaz I, Ramos V, De La Escalera S, Heras A (2003) Co-immobilization of d-hydantoinase and d-carboamylase on Chitin: Application to the Synthesis of p-hydroxyphenylglycine. Biocatal Biotransform 21:349–356

104. Lopez-Gallego F, Batencor L, Hidalgo A, Mateo C, Fernandez-Lafuente R, Guisan JM (2005) One-pot conversion of cephalosporin c to 7-aminocephalosporanic acid in the absence of hydrogen peroxide. Adv Synth Catal 347:1804–1810

105. Watanabe J, Ishihara K (2005) Sequential enzymatic reactions and stability of biomolecules immobilized onto phospholipid polymer nanoparticles. Biomacromolecules 7:171–175

106. Keighron JD, Keating CD (2010) Enzyme: nanoparticle bioconjugates with two sequential enzymes: stoichiometry and activity of malate dehydrogenase and citrate synthase on au nanoparticles. Langmuir 26:18992–19000

107. Le M, Means GE (1998) NAD +/NADH recycling by coimmobilized lactate dehydrogenase and glutamate dehydrogenase. Enzyme Microbial Technol 23:49–57

108. Liu W, Zhang S, Wang P (2009) Nanoparticle-supported multi-enzyme biocatalysis with in situ cofactor regeneration. J Biotechnol 139:102–107

109. Shi J, Zhang L, Jiang Z (2011) Facile construction of multicompartment multienzyme system through layer-by-layer self-assembly and biomimetic mineralization. ACS Appl Mater Interfaces 3:881–889

110. Cao X, Li Y, Zhang Z, Yu J, Qian J, Liu S (2012) Catalytic activity and stability of glucose oxidase/horseradish peroxidase co-confined in macroporous silica foam. Analyst 137:5785–5791

111. Qiu H, Li Y, Ji G, Zhou G, Huang X, Qu Y, Gao P (2009) Immobilization of lignin peroxidase on nanoporous gold: Enzymatic properties and in situ release of H2O2 by co-immobilized glucose oxidase. Bioresour Technol 100:3837–3842
112. El-Zahab B, Jia H, Wang P (2004) Enabling multienzyme biocatalysis using nanoporous materials. Biotechnol Bioeng 87:178–183
113. Wang P, Ma G, Gao F, Liao L (2005) Enabling multienzyme bioactive systems using a multiscale approach. China Particuology 3:304–309
114. Zhang Y, Gao F, Zhang S-P, Su Z-G, Ma G-H, Wang P (2011) Simultaneous production of 1,3-dihydroxyacetone and xylitol from glycerol and xylose using a nanoparticle-supported multi-enzyme system with in situ cofactor regeneration. Bioresour Technol 102:1837–1843
115. Obert R, Dave BC (1999) Enzymatic conversion of carbon dioxide to methanol: enhanced methanol production in silica sol−gel matrices. J Am Chem Soc 121:12192–12193
116. Xu S-w, Lu Y, Li J, Jiang Z-Y, Wu H (2006) Efficient conversion of CO_2 to methanol catalyzed by three dehydrogenases co-encapsulated in an alginate − Silica (ALG − SiO_2) hybrid gel. Ind Eng Chem Res 45:4567–4573
117. El-Zahab B, Donnelly D, Wang P (2008) Particle-tethered NADH for production of methanol from CO_2 catalyzed by coimmobilized enzymes. Biotechnol Bioeng 99:508–514
118. Sun Q, Jiang Y, Jiang Z, Zhang L, Sun X, Li J (2009) Green and efficient conversion of CO_2 to methanol by biomimetic coimmobilization of three dehydrogenases in protamine-templated Titania. Ind Eng Chem Res 48:4210–4215
119. Jiang Y, Sun Q, Zhang L, Jiang Z (2009) Capsules-in-bead scaffold: a rational architecture for spatially separated multienzyme cascade system. J Mater Chem 19:9068–9074
120. Dibenedetto A, Stufano P, Macyk W, Baran T, Fragale C, Costa M, Aresta M (2012) Hybrid technologies for an enhanced carbon recycling based on the enzymatic reduction of CO_2 to methanol in water: chemical and photochemical NADH regeneration. ChemSusChem 5:373–378
121. Shi J, Wang X, Jiang Z, Liang Y, Zhu Y, Zhang C (2012) Constructing spatially separated multienzyme system through bioadhesion-assisted bio-inspired mineralization for efficient carbon dioxide conversion. Bioresour Technol 118:359–366
122. Wu M, He Q, Shao Q, Zuo Y, Wang F, Ni H (2011) Preparation and characterization of monodispersed microfloccules of TiO_2 nanoparticles with immobilized multienzymes. ACS Appl Mater Interfaces 3:3300–3307
123. Zhang L, Shi J, Jiang Z, Jiang Y, Meng R, Zhu Y, Liang Y, Zheng Y (2011) Facile preparation of robust microcapsules by manipulating metal-coordination interaction between biomineral layer and bioadhesive layer. ACS Appl Mater Interfaces 3:597–605

Directed Multistep Biocatalysis Using Tailored Permeabilized Cells

**Steffen Krauser, Christian Weyler, Lisa Katharina Blaß
and Elmar Heinzle**

Abstract Recent developments in the field of biocatalysis using permeabilized cells are reviewed here, with a special emphasis on the newly emerging area of multistep biocatalysis using permeabilized cells. New methods of metabolic engineering using in silico network design and new methods of genetic engineering provide the opportunity to design more complex biocatalysts for the synthesis of complex biomolecules. Methods for the permeabilization of cells are thoroughly reviewed. We provide an extended review of useful available databases and bioinformatics tools, particularly for setting up genome-scale reconstructed networks. Examples described include phosphorylated carbohydrates, sugar nucleotides, and polyketides.

Keywords Biocatalysis · In situ biocatalysis · Metabolic engineering · Network design · One-pot synthesis · Permeabilization

Contents

S. Krauser · C. Weyler · L. K. Blaß · E. Heinzle (✉)
Biochemical Engineering Institute, Saarland University, 66123 Saarbrücken, Germany
e-mail: e.heinzle@mx.uni-saarland.de

1 Introduction

In billions of evolutionary steps, nature developed an impressive set of strategies to create molecules with a wide range of structures. Nearly every carbon-, nitrogen-, oxygen-, or sulfur-containing skeleton and functional group can be assembled in principle by bioconversions. A large number of different enzymatically catalyzed reactions support cellular growth and survival [1]. Not only the substrate and reaction specificity but also the efficiency of enzymatic reactions are usually far beyond manmade chemical processes. Recent developments in biochemical research not only support a detailed mechanistic understanding of relationship of structure and reactivity, but they also allow extended targeted redesign and modification of enzymes. Even completely new functionalities can be designed and created with modern molecular and modeling tools (e.g. Diels-Alder synthesis with a de novo designed enzyme [2]). The development and present status concerning biocatalysis—mostly based on engineered single enzymes—have been reviewed thoroughly elsewhere [3]. Recent developments in the field of metabolic network research, both experimentally as well as computationally, open up new potentials for multistep biocatalysis both in vivo as well as in vitro.

Presently, slightly more than 100 commercial applications use enzymes in industrial-scale processes [4]. Due to the usually high price, the time required for improving enzymes genetically, the often shorter development times required for organic chemistry alternatives, and the still widespread ignorance of biocatalysis in the field of organic chemistry, bioconversion processes are often not considered [5]. There is, however, a trend towards biotechnological processes as the ecological impact (E-factor) of industrial productions is gaining weight and public pressure demands a sustainable industry [6–9]. Biocatalytic processes often have a very low ecological impact, such as with selective oxidation of carbohydrates [10], but in some cases chemical alternatives are similar or even better [11].

Although in vivo synthesis using whole, viable microorganisms provides complex products from simple and cheap raw materials by fermentation, it is limited by the fitness and tolerance of the organism and by cellular transport processes. Modern metabolic engineering methods provide a whole toolbox comprising computational and molecular tools for directed design and optimization of production pathways. In most cases, these allow the conversion of a poorly producing native organism into a highly efficient producer strain. However, transport barriers, bottlenecks in the metabolism, toxic side effects, and the usually required complex downstream processing of resulting mixtures of product and growth medium limit the industrial applications (Fig. 1, Case C). In vitro
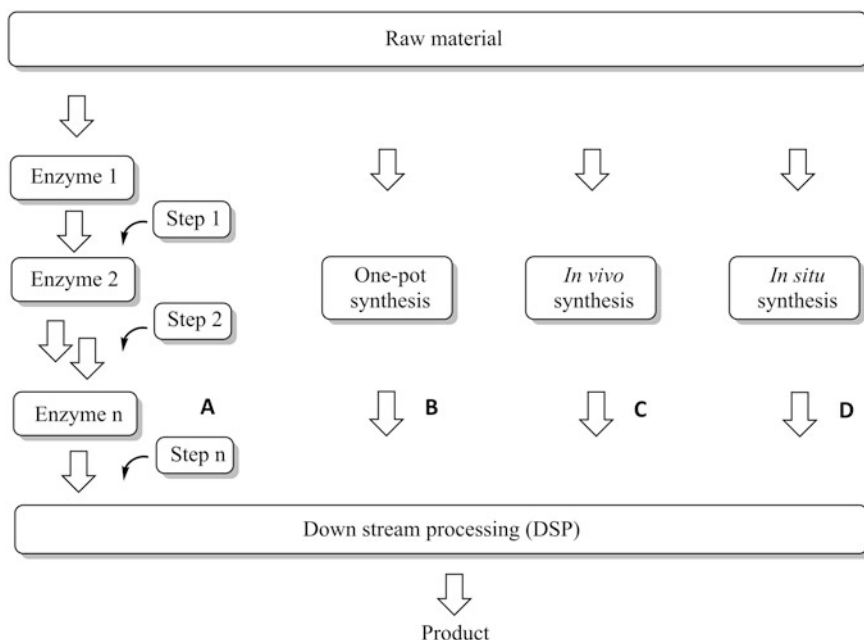
Fig. 1 Types of multistep biosynthetic processes. *A* Synthesis using multiple enzymes in separate processes, *B* Synthesis with all enzymes reacting in one-pot, *C* In vivo synthesis using living cells in fermentation processes, *D* In situ synthesis using permeabilized cells

synthesis, on the other hand, serves as a biotechnological alternative to the classic chemical catalysts. Engineered for the highest activity, stability, and substrate spectrum, enzymes provide the highest turnover rates, simultaneously working with outstanding selectivity [3]. However, in case of complex syntheses, the use of enzymes is restricted by required optimal conditions for each enzyme and potential intermediate clean-up or buffer change between individual steps (Fig. 1, Case A). Additionally, the regeneration of cofactors, such as nicotinamide adenine dinucleotide phosphate [NAD(P)H], limits this type of application. A further alternative is the one-pot synthesis using multiple enzymes; however, they require extended optimization of enzymes to operate at the same pH and buffer concentrations (Fig. 1, Case B). This approach can also be taken using cell hydrolysates of suitable strains, as has been reviewed elsewhere [12]. On the other hand, various approaches use synthetic assemblies of enzymes, such as in emulsions, using scaffolds, tethering to surfaces, or covalent binding to achieve one-pot synthesis biocatalysis [13].

Yet another strategy uses permeabilized cells—often called in situ synthesis (Fig. 1, Case D). Permeabilized cell membranes allow diffusion of low-weight compounds between the intracellular space and the surrounding reaction buffer while large biopolymers (i.e. proteins and DNA) remain trapped inside the microenvironment of the cell. Contrary to using cell hydrolysates, which has a

very long tradition [12], optimal permeabilization will keep the enzymes in their native macromolecular environment and not cause any denaturation of enzymes by the permeabilizing agent. In this way, the macromolecular crowding effects that are expected to modify protein activities [14], such as channeling [15], are preserved in their original status. Removing all small metabolites and cofactors represents a kind of reset of the metabolic network, permitting directed conversions by the selection of appropriate substrate combinations. In general, this can be combined with careful tailoring of the enzymatic outfit of a cell, thus increasing selectivity of bioconversion using permeabilized cells. Network changes may involve gene deletions, gene amplification, or heterologous gene expression. Additionally, selective inhibitors might be used to block undesired sidereactions [16]. An interesting alternative concept [17] uses enzymes from thermophilic organisms that are expressed in a mesophilic organism. Cells are cultivated and then heated to rupture the cells and inactivate enzymes that are not desired for the in vitro biocatalytic conversion.

This review highlights the opportunities that synthesis with permeabilized wholecells provides for the production of pharmaceuticals and fine chemicals. The term "in situ synthesis" was introduced in the early 1960s, indicating that macromolecules remain in their original intracellular environment. Prokaryotic and eukaryotic cells can be permeabilized, but the permeabilization procedure depends on the composition of the cell wall and has to be optimized for each cell type. Early studies on permeabilized cells by Felix showed promising results. Felix concluded that permeabilized cells can be produced quickly and simply and can be used several times, thus requiring less energy for the synthesis of biomass [18]. These studies on synthesis with permeabilized cells never achieved appropriate acknowledgment, however—likely because of the missing genetic and metabolic engineering tools at that time. The available tools have dramatically changed since then, and it seems obvious that synthesis with permeabilized cells will provide an alternative method, thus closing the gap between in vivo and in vitro biosynthesis.

The enormous increase in DNA sequencing power has recently created an overwhelming wealth of genome and metabolic network information of a large number of single (micro)organisms but also of microbial habitats using metagenome analysis. In parallel, computational tools for handling and exploring this vast amount of data have been developed at a high rate. However, detailed biochemical knowledge of enzyme characteristics is lagging far behind. Nevertheless, genome and enzyme databases provide an enormous amount of data that may be explored for permeabilized cell synthesis. Whole genome metabolic networks become increasingly available—a few of them already carefully curated. Metabolic regulation is also increasingly explored, but it requires considerably higher effort compared to sequencing. For some microorganisms, such as *Escherichia coli* and *Saccharomyces cerevisiae,* metabolic and regulatory networks are already fairly well understood, but we are still quite far away from the comprehensive understanding required for creating fully predictive models. This is even more the case for the majority of microorganisms. The metabolism of microorganisms may differ considerably. Nevertheless, they all share large parts of their central metabolism,
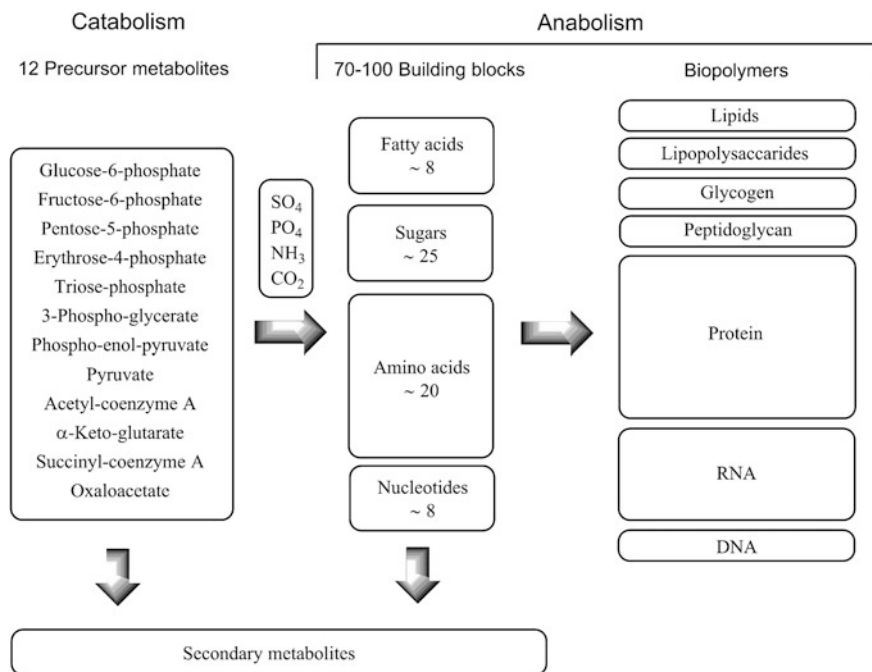
Fig. 2 Overall structure of metabolism

particularly the 12 small precursor molecules representing the bottleneck of the bow-tie-shaped structure of metabolic networks [19]. These precursor molecules serve as starting materials for all building blocks and polymers that can be synthesized in the metabolic network (Fig. 2). Although biopolymers constitute the major fraction of cellular biomass, secondary metabolites are of great interest as pharmaceutically active compounds or precursors thereof.

With the present knowledge, molecular and computational methods, and the advent of new possibilities of designing and engineering enzymes and whole metabolic pathways, a large field of applications opens up. Together with the long-known technique of permeabilization of cellular barriers (i.e. cell membranes and cell walls), new and intriguing opportunities for designing tailored biocatalyst and bioprocesses become accessible. Once such a biocatalyst is established, it can be produced easily by simple cultivation followed by permeabilization. Downstream processing would simply start with the removal of the biocatalyst, such as by centrifugation. There are, however, major hurdles to be overcome as far as more complex biosynthesis is concerned. The most important are the supply of precursors and cosubstrates, such as adenosine triphosphate (ATP) or NAD(P)H; the achievement of selective biocatalysts, meaning the elimination of the manifold possible undesirable side reactions; and the intensification of processes to obtain high final product concentrations. To reach these goals, it is important to understand (i) the permeabilization process on a molecular basis, (ii) biosynthetic

Fig. 3 Cell envelopes of
gram-positive and gram-
negative bacteria. *CAP*
covalently attached protein,
IMP integral membrane
protein, *LP* lipoprotein, *LPS*
lipopolysaccharide, *LTA*
lipoteichoic acid, *OMP* outer
membrane protein, *WTA* wall
teichoic acid. Adapted
from [21]
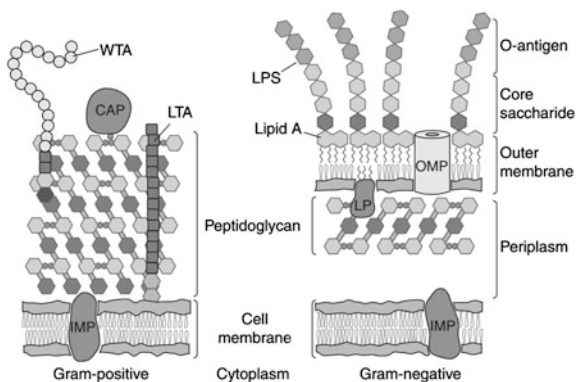
pathways and their regulation, (iii) supply of precursor molecules, (iv) regenera-
tion of cosubstrates, (v) design and selection of enzymes as part of the biosynthetic
processes, and (vi) design of biocatalytic pathways on the basis of the vast
genomic and biochemical knowledge. In this review, we will discuss some prin-
ciples of these points, particularly the permeabilization process and the design of
biocatalytic pathways. We also will provide a few examples illustrating this
emerging biocatalytic technology.

2 Permeabilization

2.1 Lipids and Membrane Structures

All cells are confined by envelope structures that create cellular compartments,
which allow a controlled interior separated from the surrounding environment.
Although the diversity of these structures is huge, they share common elements,
such as lipids and integral membrane proteins, which constitute the core structures
of membranes that are often linked to additional carbohydrate and protein struc-
tural elements.

   As shown in Fig. 3, there are two major classes of bacterial envelopes. Gram-
positive organisms have a thick layer of peptidoglycan as the main compound of
their cell wall, providing them with great mechanical stability. Gram-negative
organisms have an additional outer membrane, which protects the cell by pre-
venting nonpolar toxic substances from entering. Compared to the cell membrane,
it is less selective and more permeable to small molecules [20].

   Yeast cells also have a robust cell wall (Fig. 4), characterized by a strongly
branched glucan layer. Due to the robustness of their cell wall structures, yeast and
bacteria have great mechanical stability, which they often preserve after perme-
abilization [22–24]. This facilitates the handling and especially the washing of
these cells. Most permeabilization methods disturb or partly disrupt the cellular

Fig. 4 Structure of the cell surface of yeast. *CW* Cell wall, *CM* cell membrane, *PMP* plasma membrane protein (From [30] with permission from Elsevier)
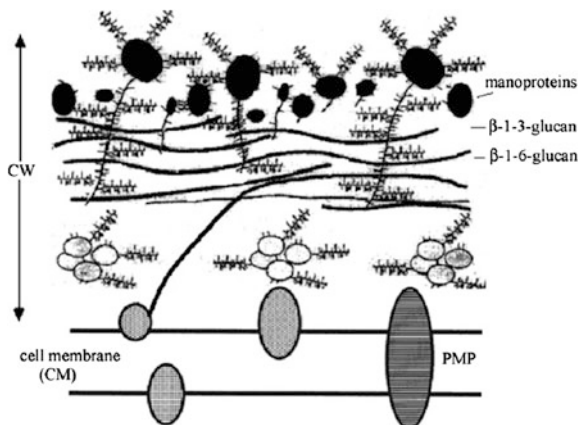
Fig. 5 Structure of the plasma membrane
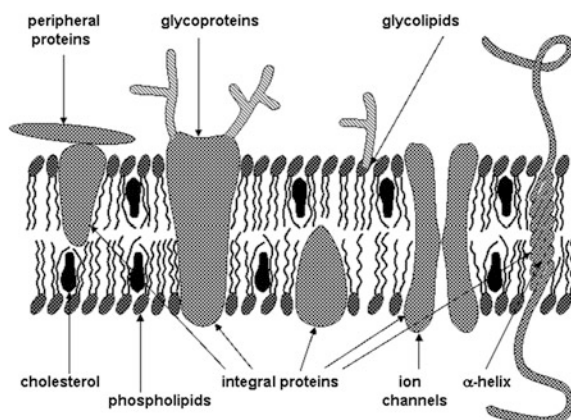
membranes while leaving the remaining cell wall structures largely intact in most cases [24–26]. For this reason, bacterial and yeast cells preserve their macroscopic shape and dimensions after permeabilization. Although it does not block the passage of small molecules, the remaining cell envelope largely prevents the diffusion of cytosolic enzymes into the surrounding medium [27–29]. At the same time, the typical mechanical robustness is usually maintained, simplifying downstream processing as centrifugation.

Because mammalian cells do not have a cell wall, they are more susceptible to mechanical stress. The permeabilization further weakens the cell membrane (Fig. 5). For this reason, permeabilized mammalian cells can only be centrifuged at relatively low speeds [31]. In contrast to bacterial and yeast cells, permeabilization of mammalian cell membranes normally leads to the release of a large fraction or even all cytosolic proteins into the medium because there is no cell wall retaining these proteins [31, 32]. When treating mammalian cells with cross-linking reagents prior to permeabilization, this loss can be avoided [33]. It has

been shown that the cell membrane of Chinese hamster ovary (CHO) cells can be selectively permeabilized while leaving the mitochondrial membrane intact, such as by using digitonin [31]. Full permeabilization is obtained using other surfactants, such as triton X-100, and other permeabilization conditions [31].

2.2 Permeabilization Methods/Agents

It needs to be mentioned in advance that there is no generally optimal permeabilization method. It depends on the organism to be permeabilized and has to be tailored to meet the requirements of the following processes. Normally, it is not possible to predict which method is suitable for a certain organism. Even when working with different strains of the same organism, very different approaches may be necessary for successful permeabilization [16, 22, 25, 34, 35]. So far, optimal conditions have to be identified experimentally by a trial-and-error procedure. However, as our knowledge of membrane processes on molecular scale keeps growing [36], it might be possible to predict suitable permeabilization conditions in the future.

Organic Solvents

Toluene is the most commonly used solvent for cell permeabilization. Studies on its effect on *E. coli* were carried out as early as 1965 [37]. During permeabilization, cells released 25 % of total cell protein into the surrounding medium. In the presence of Mg^{2+} only very little cell material was released. Even nucleotides remained within the cells. The lipopolysaccharide layer of the outer membrane is bound to the membrane by divalent cations. Their presence therefore protects the cell from nonpolar compounds (such as toluene) [21]. It was shown by electron microscopy that toluene mainly damages the cell membrane, whereas the outer membrane was left intact. More recent publications point out that permeabilization with solvents is more harmful to some cells and less reproducible than the use of surfactants [37]. On the other hand, *E. coli* and two *Bacillus* strains were successfully permeabilized using a combination of ethylenediaminetetraacetic acid (EDTA) and low toluene concentrations in 2009 [25, 34]. Furthermore, yeast cells could be permeabilized by treatment with mixtures of chloroform, ethanol, and toluene [38].

Surfactants

Currently, surfactants are the most commonly used permeabilizing agents. There are three general mechanisms by which surfactants can interfere with the cellular membrane. If applied in concentrations above their critical micellar concentration, they can form micelles consisting of a mix of membrane compounds and surfactant molecules. The cationic part of surfactants such as cetyltrimethylammonium bromide (CTAB) can interact with the negatively charged membrane, which disturbs its structure and causes permeability [29]. The nonpolar part, which is of course found in all surfactants, can integrate into the membrane due to its high

affinity to the lipids nonpolar tail, disturbing its transversal organization. In practice, it is to be expected that permeabilization occurs due to a combination of these effects [24, 39, 40].
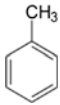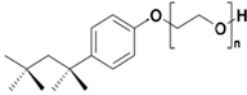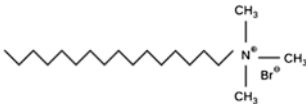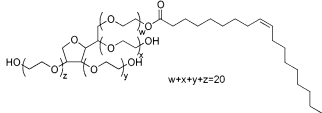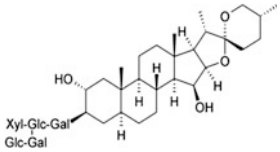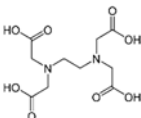
Most proteins contain polar and nonpolar parts due to the respective amino-acid side-chains. This makes them susceptible to the binding of surfactants, which can lead to conformational changes and ultimately denaturation of the protein if the surfactant concentration is sufficiently high [41–43]. The most notable example of the affinity of surfactants for proteins is sodium dodecyl sulfate (SDS) poly-acrylamide gel electrophoresis, in which proteins are "covered" with the surfactant SDS [43].

When permeabilized with nonionic surfactant Triton X-100, *Yarrowia lipolytica* released only very small amounts of protein. Electron microscopy confirmed that permeabilization with this surfactant is a mild process [44]. In 2011, Niklas et al. permeabilized the cell membrane and the mitochondrial membrane of CHO cells with Triton X-100, whereas treatment with digitonin (a saponin) selectively permeabilized the cell membrane while leaving the mitochondria intact [31, 45]. Digitonin does not primarily act as a surfactant but binds to cholesterol [46], which is more abundant in the cell membrane than in the mitochondrial membrane [47]. The previously mentioned surfactant CTAB is the most commonly used permeabilizing agent harboring a permanent positive charge. By its strong binding to lipids and phospholipids, it induces pore formation in cellular membranes [48]. When permeabilizing *S. cerevisiae* with CTAB, Gowda proved that permeabilization was dependent on the ratio of amount of surfactant to cells rather than surfactant concentration in the medium [49]. The Upadhyas group treated yeast cells with 1 % glutaraldehyde prior to permeabilization with CTAB. In this way, protein leakage was greatly reduced and the cells could be reused 20 times as catalyst in the oxidation of D-phenylalanine to phenylpyruvate [50]. An overview of surfactants used for permeabilization is given in Table 1.

Chelating Agents
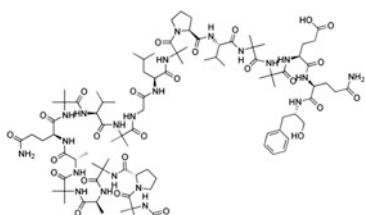
The structural stability of membranes and most cell wall structures is based on electrostatic interactions and van der Waals forces. Divalent cations counterbalance the negative charges of phosphate groups, allowing these groups to stay densely packed. Furthermore, the lipopolysaccharide layer of Gram-negative bacteria is attached to the outer membrane by calcium and magnesium ions. Chelating agents, such as EDTA [51], form stable complexes with these cations, separating them from the cells. This process disturbs the cell membrane and causes the release of lipopolysaccharides [20]. It has been shown that cells of *B. pumilus*, *B. subtilis*, and *E. coli* permeabilized with EDTA and low (7 % v/v) toluene concentration are effective catalysts for NADPH regeneration and asymmetric reductions [25]. These cells showed high stability and could be reused three times as catalyst in 68-h batch reactions [25, 34]. *G. oxidans* and *R. eutropha* were permeabilized with EDTA and toluene (2 % v/v) and then used as catalyst in a reduction reaction using hydrogen as reductant [52]. Combining EDTA treatment with low toluene concentrations seems to permeabilize cells effectively without the

Table 1 Permeabilizing agents

Notable examples (methods/agents)	Structures	Mechanism of permeabilization/notes
Toluene (OS) [22, 34, 52, 73–76]		Disrupt the cell membrane Often combined with EDTA because the amount of damage done to membranes is strongly dependent on divalent cation concentrations
Ethanol (OS) [22, 40, 50, 77, 78]		
Chloroform (OS) [74]	$CHCl_3$	
Triton X-100 (D) [22, 23, 31, 40, 71, 72, 76]		$n = 8–9$ Nonionic surfactant Partly dissolves membranes unspecifically
CTAB (D) [21, 29, 50, 79, 80]		Cationic surfactant Partly dissolves membranes unspecifically
Tween 80 (D) [22, 58, 60, 71, 77, 81–83]		Nonionic branched polyether Partly dissolves membranes unspecifically
Digitonin (D) [31–33, 45]		Steroid glycoside Selectively permeabilizes membranes with high cholesterol content
EDTA (C) [20, 25, 26, 34, 50–52, 84, 85]		Complexation and thus removal of divalent cations required for membrane stability

(continued)

Table 1 (continued)

Notable examples (methods/agents)	Structures	Mechanism of permeabilization/notes
Alamethicin (P) [60, 86, 87]		Channel forming

EDTA ethylenediaminetetraacetic acid, *OS* organic solvent, *D* surfactant, *P* peptide, *C* chelating agent

disruptive effect caused by permeabilization with toluene alone (at high concentrations). The permeabilization of *P. putida* with EDTA was optimized with regard to EDTA concentration, temperature, treatment time, and pH with the support of a statistical experimental design with the goal of conversion of limonene [26].

Peptides and Proteins

Eukaryotes synthesize numerous peptides and small proteins with antimicrobial activity. Many of these compounds act by permeabilizing microbial, fungal, and viral cellular envelopes [53, 54]. Defensins are small, rigid (internal disulfide bonds), and positively charged (high arginine content) peptides. Due to their positive charge, they can integrate into the negatively charged membranes and associate to form multimeric pores [55–57]. It should be mentioned that polyethylenimine, a polymer that is positively charged under physiological conditions, exhibits a similar effect [22]. Besides the pore-forming activity, cationic peptides can displace calcium and magnesium from the outer membrane, thus destabilizing this region. Different mechanisms of action have been discussed and attempts of establishing structure–activity relationships have been described [58]. The protein psoriasin has been proven to kill *E. coli* at neutral pH without damaging its membrane, whereas at low pH it acts only against *B. megaterium* by membrane permeabilization [59]. In 2009, the channel-forming peptide alamethicin was used to permeabilize *A. thaliana* and tobacco suspension cells to study callose synthase activity. In these experiments, it was superior to surfactant digitonin, which showed an inhibitory effect on the involved enzymes [60].

Other Permeabilizing Agents

Ionophores are compounds that can carry ions through biological membranes. Nigericin was used to permeabilize HeLa cells to release metabolites and small proteins without interfering with protein synthesis [61]. When desiring to remove

soluble proteins from the cytoplasm, the cell membrane can be partly removed by enzymatic digestion. This approach has been applied to different yeast strains [62, 63].

Physical Methods

Electric Pulse Methods

Electroporation is a well-established method for introducing foreign DNA into organisms [64–66]. However, it can also be used to permeabilize cells for the enhancement of their catalytic properties. A general advantage of the electropermeabilization compared to the previously described chemical methods is that no foreign substances, such as surfactants or solvents that can interfere with enzymes, need to be added [67]. Because the conductivity of lipid bilayers is very low, the application of an external electric field creates a potential difference across the membrane [68]. This causes alterations within the membrane, disturbing its structure and leading to pore formation. An additional nanoscale fragmentation of the membrane has been proven [69]. In 2008, Tryfona et al. reported the enhancement of amino acid release by *C. glutamicum* after electropermeabilization, demonstrating the potential of this approach for production purposes [27].

Other Physical Methods

An osmotic shock causes cells to change their size rapidly, thus causing stress to the cellular membranes and potentially leading to permeabilization. This approach has been used to permeabilize yeast [70] and fungi [28] for enzymatic activity studies in situ. Because mammalian cell envelopes do not contain a peptidoglycan layer protecting them from osmotic lysis, the osmotic shock method is not suitable to permeabilize these cells [22, 71]. Freezing and thawing of cells has been studied as a method for cell permeabilization as well. It is generally combined with the application of surfactants [22, 71, 72].

3 Enzymatic Synthesis of Complex Molecules

Although in vitro biotransformation of all kinds of simple and complex molecules has reached a very high standard with numerous applications [3, 4], the biosynthesis of complex molecules from simple and inexpensive precursors is still dominated by in vivo systems (i.e. whole living cells in fermentation processes). Although the main focus in in vitro synthesis is mostly the characteristics of the enzyme applied [3], in vivo production focuses on a whole network or organism characteristics. Up-to-date design of synthesis using whole metabolic networks starts with stoichiometry-based methods, usually flux balance analysis [88–91]. Similarly, the design of more complex networks of cell extracts [92] or permeabilized cells will start out from information that is readily available from public databases (see Table 4). The operability of a path or network is additionally

dependent on favorable thermodynamics, which is increasingly accessible via databases, and even more with the development of methods for molecular energy calculations (e.g. using group contribution methods) [92, 93]. A full quantitative description for a fully model-based design would also include kinetics, which are much more difficult to describe.

3.1 Enzymatic Reactions

Living organisms rely on metabolic networks with amazingly efficient synthetic routes to provide essential intermediates and building blocks for cellular polymers but also for other low-molecular-weight compounds—so-called secondary metabolites [94]. Enzymes catalyzing these reactions are highly conserved in their three-dimensional structure and chemical function despite the genomic diversity among organisms [95]. Thousands of biologically synthesized molecules that are already known and those that are not yet discovered give an image of the powerful synthetic machinery that cells can rely on. For conducting chemical reactions, enzymes often include cofactors (coenzymes or cosubstrates as well as prosthetic groups).

When using permeabilized cells, it is important to distinguish carefully between cosubstrates that are dissolved in cellular compartments, such as cytosol, and prosthetic groups that are integral parts of enzymes. Permeabilization will allow cosubstrates to diffuse freely out of the cells, whereas prosthetic groups would remain within the cells attached to their protein hosts. As a consequence, cosubstrates can be added in a targeted way to a biosynthetic network captured within the permeabilized cell membrane and are thus a design variable. Cofactors are necessary to provide energy, redox equivalents, and carbon-units and are evolutionary preserved (Table 2). In some cases, enzymes can even operate in a thermodynamically unfavorable direction, provided resulting products are removed by the subsequent exergonic processes [96].

Water plays an essential role in all naturally occurring enzymatic reactions. By entropy effects and acid/base catalysis as well as regulating the tertiary structure of enzyme folding, it influences the binding of substrates and actual catalytic activity [97]. Nonaqueous solvents do not only change the solubility of substrates, products, and intermediates, but they also have a strong impact on the structure of enzymes and therefore their catalytic activity [98]. The application of nonaqueous solvents in multienzyme biocatalysis seems more difficult because of the expected influences on individual enzymes of such a biocatalytic network. Enzymes do usually not only convert the major substrates listed in databases but may act on a whole spectrum of compounds; for example, many proteases can also act as esterases (a well-known textbook example). This complicates the design of more complex networks, but it also opens up lots of opportunities to create variations of metabolites, potentially leading to new pharmaceuticals or to precursors for further

Table 2 Cofactors in enzyme reactions

Cofactor	Group transfer	Species
3′-Phosphoadenosine-5′-phosphosulfate	Sulfate group	Bacteria, archaea and eukaryotes
Adenosine triphosphate	Phosphate group	Bacteria, archaea and eukaryotes
Ascorbic acid	Electrons	Bacteria, archaea and eukaryotes
Biotin	CO_2	Bacteria, archaea and eukaryotes
Cobalamine	Hydrogen, alkyl groups	Bacteria, archaea and eukaryotes
Coenzyme A	Acetyl group and other acyl groups	Bacteria, archaea and eukaryotes
Coenzyme B	Electrons	Methanogens
Coenzyme F420	Electrons	Methanogens and some bacteria
Coenzyme M	Methyl group	Methanogens
Coenzyme Q	Electrons	Bacteria, archaea and eukaryotes
Cytidine triphosphate	Diacylglycerols and lipid head groups	Bacteria, archaea and eukaryotes
Flavin adenine dinucleotide	Electrons	Bacteria, archaea and eukaryotes
Flavin mononucleotide	Electrons	Bacteria, archaea and eukaryotes
Glutathione	Electrons	Some bacteria and most eukaryotes
Heme	Electrons	Bacteria, archaea and eukaryotes
Lipoamide	Electrons, acyl groups	Bacteria, archaea and eukaryotes
Menaquinone	Carbonyl group and electrons	Bacteria, archaea and eukaryotes
Methanofuran	Formyl group	Methanogens
Methylcobalamin	Acyl groups	Bacteria, archaea and eukaryotes
Molybdopterin	Oxygen atoms	Bacteria, archaea and eukaryotes
NAD^+ and $NADP^+$	Electrons	Bacteria, archaea and eukaryotes
Nucleotide sugars	Monosaccharides	Bacteria, archaea and eukaryotes
Pyridoxal phosphate	Amino and carboxyl groups	Bacteria, archaea and eukaryotes
Pyrroloquinoline quinone	Electrons	Bacteria
S-Adenosyl methionine	Methyl group	Bacteria, archaea and eukaryotes

Table 2 (continued)

Cofactor	Group transfer	Species
Tetrahydrobiopterin	Oxygen atom and electrons	Bacteria, archaea and eukaryotes
Tetrahydrofolic acid	Methyl, formyl, methylene and formimino groups	Bacteria, archaea and eukaryotes
Tetrahydromethanopterin	Methyl group	Methanogens
Thiamine pyrophosphate[l]	2-carbon groups, α cleavage	Bacteria, archaea and eukaryotes

synthesis of pharmaceuticals. In some cases, the native enzyme easily accepts substrate analogues, whereas others have to be tailored [99, 100].

3.2 Enzyme Kinetics

Although the kinetics of biosynthetic networks generally follow the same principle as that of isolated enzymes, some peculiarities have to be considered because of macromolecular crowding [14]. A known but quantitatively little-characterized effect is channeling—that is, the direct transfer of substrates to the following enzyme in a cascade reaction [101].

Biocatalytic network kinetics can be modeled in various ways. In the simplest approach, one attempts to identify the limiting catalytic step and use the kinetics of this step to describe the whole process. This approach is adequate as long as the step remains the limiting one; this may change during optimization. In a second alternative approach, full kinetics of all relevant reaction steps are applied. This requires, however, the elaboration of the kinetics of all individual enzymes, and it will only be valid if no secondary effects (e.g. metabolite channeling) occur. Wilson et al. have characterized the kinetics of a dimodular nonribosomal peptide synthetase [102]. A new powerful method for larger scale kinetic analysis has been presented that seems applicable to permeabilized cells as well [103]. Network-oriented approaches use power-law kinetics, such as secondary metabolism in plants [104] or lin-log kinetics [105], which do not require a full kinetic analysis of each single enzyme but an evaluation of the whole network of interest.

In permeabilized cells, it is possible to add desired substrates and omit undesired ones in a directed manner, thus preventing activity of undesired reactions [23, 106]. However, all enzymes present remain ready but have to be considered sleeping unless their corresponding substrates are present. In more complex cases, it will usually not be avoidable to create intermediates that can react with present substrates, intermediates, or products. In this case, it is generally possible to delete the genes coding for undesired enzymes or to inhibit them by selective inhibitors. Because most enzyme-catalyzed conversions are equilibrium reactions, excess supply of reactants can push them in the desired direction. This extends the

operability of bioconversions while the working range of native enzymes has to be respected [107].

Enzymes will usually be inactivated, particularly in a nonphysiological environment. An advantage of permeabilized cells with respect to catalytic activity is the stabilizing effect of the protective environment inside the cell. Enzyme activity can be retained for several days in a suitable buffer media [23]. Even multiple uses of permeabilized cells in some cases did not affect the activity of the enzymes significantly [18]. Enzyme stability very much depends on the nature of the enzyme and the environment to which it is exposed. Often, storage stability is significantly different from operational stability, such as for oxidases [108]. This is in contrast to isolated enzymes, which are exposed to damaging influences and often have to be modified genetically to increase their stability [109].

3.3 Synthesis of Complex Metabolites

Natural products represent powerful lead structures for new pharmaceutically active agents [110]. Methods for discovering and producing new secondary metabolites are increasingly powerful [111, 112]. There is an increasing interest to also produce modifications of natural metabolites [113, 114], which would certainly be stimulated if cell extracts or permeabilized cells could carry out such complex syntheses.

Some biosynthetic enzymes form megasynthetases by aggregation and post-translational modification and are mostly coded by contiguous operons on the genome. Those clusters build up biosynthetic machineries for larger, complex molecules, such as secondary metabolites [115–119]. So far, it is hardly possible to isolate such complexes in an active form, and rebuilding them in vitro only works for smaller, simple aggregates [120, 121]. However, substrate channeling in enzyme clusters enables effective bioconversion, limiting side reactions [15]. Intermediates towards the product are passed among the subunits of a megasynthetase and shielded from the surrounding environment, such as the cytosol. This allows high reaction rates and improves selectivity. Figure 6 shows the biosynthesis of the secondary metabolite oxytetracycline in *Streptomyces rimosus*, elucidated by Pickens et al. in 2009 [115, 122].

Because most of the pharmaceutical bioactives that are produced biotechnologically target microbial growth, production of such compounds in microorganisms is limited. Resistance concerning the synthesized product and secretion from the cellular plasma is highly demanded to preserve the fitness of the producing strain. This causes a significant loss in efficiency [123]. Although living organisms have to be prepared for the stress resulting from toxic products, using them in a permeabilized state can reduce this problem if toxic effects do not directly involve the actual biosynthesis but only other essential functions, such as respiration. Cells can be easily grown in their preferred conditions. Triggered expression of corresponding synthesis genes in the late, stationary growth phase does not affect
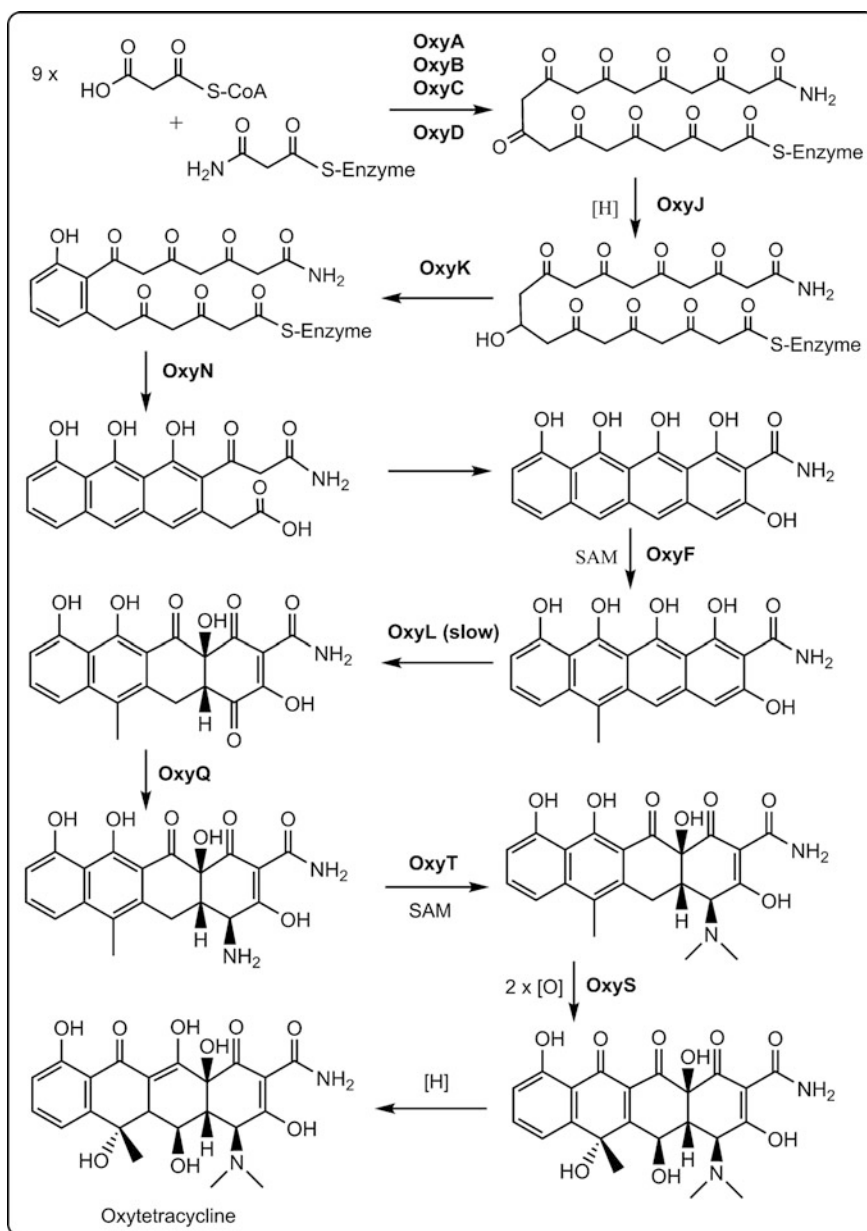
Fig. 6 Biosynthesis of oxytetracycline performed by the oxytetracycline mega synthetase cluster (from [122] with permission from the American Society for Biochemistry and Molecular Biology)
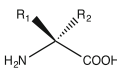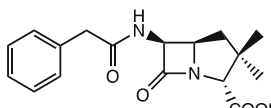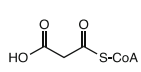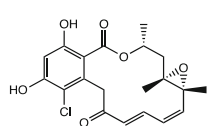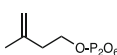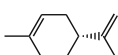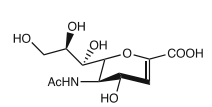
replication very much. In the case of megasynthetase clusters, controlled induction is even more advantageous because they are naturally expressed only at very low copy numbers. High expression rates may produce dysfunctional aggregates, such as inclusion bodies. On the other hand, when using permeabilized cells, it is easy to set cell concentrations to a desired value [106]. In that way, it is possible to reach catalytic activity that would not be possible by regular growth, thus enhancing the volume-related synthesis rate compared to in vivo production during fermentation.

The use of such large enzyme clusters for synthetic applications in permeabilized cells has diverse advantages. Accessing synthesis clusters via permeabilized cell membranes is not known to harm the native assembly of subunits [31]. The surfactant itself might affect the complex's structure. The synthetic machinery can thus be addressed in vitro so that biosynthesis can occur in an environment that can be easily controlled and influenced—provided that the corresponding route and co-factors are known and available. Because many secondary metabolite synthetase clusters are spread all over microbial genomes, their identification is considerably more difficult. Only very few have been investigated in more detail concerning their mechanism and kinetics [102]. Resolving reaction mechanisms and intermediates from known starting monomers to the final compound is still a lot of work. Research will not only provide some insight into how the biosynthetic toolbox evolved but may also inspire scientists to copy and modify enzyme complexes by known or new engineering tools. Table 3 lists major secondary metabolite classes and their link to primary metabolism.

In nature, megasynthetases can combine several subunits synthesizing hybrids of those classes, such as for alkaloid synthesis in plants [124]. Research in this area underwent astonishing and promising progress in the last decade. The growing number of known biosynthetic paths enlarges the range of potential synthetic applications as well as the understanding of genomic coding across species [48, 125]. Extensive knowledge of genomes increasingly allows the identification and later activation of silent gene clusters [126]. Engineering of secondary metabolite production in streptomycetes [127] and myxobacteria [128] is increasingly providing new opportunities for the application of permeabilized cells. The previously mentioned synthesis gene cluster of Oxytetracycline of *Streptomyces rimosus* was transferred from a *Streptomyces* strain into the myxobacterium *Myxococcus xanthus* and also *E. coli* using the Red/ET cloning technique (Gene Bridges, Heidelberg, Germany) [123, 129]. In this case, the necessary σ^{54} transcription factor is co-expressed with needed chaperones present in engineered *E. coli* BAPI strain for correct folding. Such efficient cloning tools enable biotechnologists to use specific synthetic skills of hard-to-handle organisms in robust production organisms, such as *E. coli*. There is a great interest in creating natural and modified versions of complex low molecular weight compounds, such as polyketides or nonribosomal peptides and mixtures thereof [113, 114].

Another very fast growing field is the in vitro reconstruction of such synthetic pathways. Enzymes are brought together in one pot to perform multistep

Table 3 Classes of secondary metabolites, building blocks, and common examples

Secondary metabolites	Enzymes	Monomer	Example
Peptides	Nonribosomal peptide synthetases	Amino acid	Penicillin G (antibiotic)
Polyketides	Polyketide synthetases types I–III	Malonyl-CoA	Radicicol (antibiotic)
Isoprenoids	Prenylating enzymes	Isoprenyl-pyrophosphate	S (−) Limonen (flavor)
Carbohydrates	Carbohydrate synthetases	$R\text{-}(C_nH_{2n\text{-}1}O_{n\text{-}1})$ Activated sugars R = activating residue	Neu5Ac2en (antiviral)

syntheses, either in microreactors (i.e. nanocontainers) or arranged on preset scaffolds (i.e. attached to a DNA linker) [13, 130–134].

3.4 Enzymes Beyond Nature

A very interesting field in engineering bioconversions is tailoring enzymes for specific tasks. The fast-growing community for enzyme engineering is aiming to manipulate enzyme properties for applications in chemical processes. Mainly, the resistance against organic solvents and high temperatures as well as stability and the ability to immobilize them attached to supports are driving the development of

enzyme engineering [100]. In silico designed enzymes carrying out so-far-unknown reactions also are targets of research [2, 135]. Increasing knowledge of enzyme-substrate interaction at the transition state enables scientists to simulate reactions that we so far would not think of being carried out by enzymes, thus extending the space reachable via biosynthesis. A very interesting biosynthetic concept was recently described by Ye et al., expressing pathways from thermophilic organisms in a mesophilic one with subsequent inactivation of mesophilic activities at higher temperatures [17, 136].

Modern genetic tools enable tailoring sequences of native enzymes, thus modifying their activity profiles; for example, one can extend their substrate spectrum, switch cofactor requirements, and enhance thermostability and solvent stability. This is done either by using evolutionary methods supported by large screening platforms or applying rational design methods [137]. Screening methods are becoming more automated and allow efficient high-throughput analysis of large libraries of mutants. Computer models have become increasingly powerful and can predict the relationship between the structure of an enzyme and its function. Enzyme engineering is increasingly introduced into the engineering of whole secondary metabolite pathways [138].

Not every enzyme needs to be engineered to carry out unnatural reactions. In some cases, enzymes convert derivatives of their main substrate like the original [96, 113]. Living cells, with their distinct metabolism, limit the supply of unnatural substrates; however, with permeabilized cells, it is possible to expose enzymes with unnatural reagents, thus exploring pathways that do not occur in nature. This might permit the use of existing biosynthetic pathways for synthesizing artificial analogs of secondary metabolites via a supply of unnatural substrates.

4 Network Design

Directed, selective biosynthesis using either cell hydrolysates or permeabilized cells can be a straightforward process for shorter paths or small networks. It is, however, becoming increasingly challenging for longer biosynthetic paths with more compounds involved. The number of potential undesired side reactions increases dramatically. Therefore, there is a great need to guide the design of such complex biocatalysts by using adequate computational tools and the rapidly increasing information available on mostly public databases. Panke and Bujara proposed an in silico tool for network topology analysis based on genome-scale metabolic network models to be applied for in vitro biocatalysis in cell-free systems [92]. Starting out from the whole-genome scale metabolic reconstruction of *E. coli* [139], they introduced several changes, particularly concerning transport and other membrane processes. Considering basic thermodynamic data and expression data from *E. coli,* they arrived at a model that could eventually predict interfering pathways for the production of dihydroxyacetone phosphate starting from glucose. The presently available examples of pathway prediction for

Table 4 Biological database resources

	URL	Content
Biochemical databases		
PubChem [140, 141]	http://pubchem.ncbi.nlm.nih.gov/	Chemical molecules and their activities against biological assays
ChEBI [142]	http://www.ebi.ac.uk/chebi/	Chemical Entities Of Biological Interest
TCDB [143]	http://www.tcdb.org/	Transporter Classification Database
Transport DB [144]	http://www.membranetransport.org/	Transporter protein analysis database
SABIO-RK [145]	http://sabio.villa-bosch.de/	Biochemical reaction kinetics
Rhea [146]	http://www.ebi.ac.uk/rhea/	Manually annotated database of chemical reactions
Genome databases		
GEM	https://www.gem.re.kr/	Genome Encyclopedia of Microbes
GenBank® [147]	http://www.ncbi.nlm.nih.gov/Genbank/	Annotated collection of all publicly available DNA sequences
NCBI Entrez Genome [148]	http://www.ncbi.nlm.nih.gov/sites/genome	Sequence and map data from whole genomes of over 1,000 species and strains
NCBI Entrez Gene [149]	http://www.ncbi.nlm.nih.gov/gene	Database of genes
GO [150]	http://www.geneontology.org/	The Gene Ontology
ENA [151]	http://www.ebi.ac.uk/ena/	European Nucleotide Archive
Protein and enzyme databases		
BRENDA [152]	http://www.brenda-enzymes.info/	Comprehensive enzyme information System
Expasy—ENZYME [153]	http://www.expasy.org/enzyme	Enzyme nomenclature database
UniProt [154]	http://www.uniprot.org/	The Universal Protein Resource
PSORTdb [155]	http://db.psort.org/	Protein subcellular localizations for bacteria and archaea
ProLinks [156]	http://prl.mbi.ucla.edu/prlbeta/	Inferring functional linkages between proteins
STRING [157]	http://string-db.org/	Search Tool for the Retrieval of Interacting Genes/Proteins
IntAct [158]	http://www.ebi.ac.uk/intact/	Molecular interaction database

(continued)

Table 4 (continued)

	URL	Content
Pathway databases		
KEGG [159, 160]	http://www.genome.jp/kegg/	Kyoto Encyclopedia of Genes and Genomes
BioCyc [161]	http://biocyc.org/	Collection of 1,962 Pathway/Genome Databases
BioPath	http://www.molecular-networks.com/biopath3	Biochemical molecules, reactions and pathways
ExPASy—Biochemical Pathways [162]	http://web.expasy.org/pathways/	Digitized version of the Roche Applied Science "Biochemical Pathways" wall chart
UniPathway [163]	http://www.grenoble.prabi.fr/obiwarehouse/unipathway	A resource for the exploration of metabolic pathways
UM-BBD [164]	http://umbbd.ethz.ch/	University of Minnesota Biocatalysis/Biodegradation Database
MetaNetEx.org [165]	http://metanetx.org/	Automated Model Construction and Genome Annotation for Large-Scale Metabolic Networks
Model databases		
BiGG [166]	http://bigg.ucsd.edu/	Knowledgebase of Biochemically, Genetically, and Genomically Structured Genome-Scale Metabolic Network Reconstructions
BioModels [167]	http://www.ebi.ac.uk/biomodels-main/	Annotated Published Models
Organism specific databases		
EcoCyc [168]	http://ecocyc.org/	*E. coli* K-12 MG1655
SGD [161]	http://www.yeastgenome.org/	*S. cerevisiae* Genome Database

biosynthesis using either cell extracts or permeabilized cells are still very limited, but we expect that there will be a rapid increase of such studies in the very near future. There is, however, a whole series of studies available for living cells that are separated from the environment by their envelopes, providing selective transport of molecules in and out of the cells.

Modern planning and development of biochemical syntheses or novel synthesis routes in living organisms is effectively supported by the use of appropriate in silico tools. Such tools are increasingly available for pathway design in microorganisms and allow quick and directed engineering of living cells [89, 91]. These tools rely heavily on the existence and quality of the numerous biological databases containing information on different aspects such as genome sequences, enzyme data, or even whole pathways (Table 4). Together with data from primary literature and further sources, this information can be used for the composition of network reconstructions of the organism of interest. Such networks can then be conveniently analyzed and developed further with different bioinformatic tools (Table 5). In particular, they can be used to design pathways or biosynthetic subnetworks useful for biocatalytic purposes, such as the in situ synthesis of a desired primary or secondary metabolite. With an increasing number of steps and increasing numbers of metabolites and coenzymes, the involved design becomes an increasingly complex task.

4.1 Databases

Biological databases can be classified into several different categories, such as biochemical databases, genome databases, protein or enzyme databases, pathway databases, or model databases (Table 4). This classification is based on the biological content of the respective databases. However, an overlap of information can occur. For example, genome databases (Table 4) also contain protein sequence information.

Biochemical Databases

Rhea [146] is a manually annotated, expert-curated reaction database with a main focus on enzyme-catalyzed reactions. It also contains other types of reactions. All reaction participants are linked to ChEBI [142], which provides data such as structure, formula, and charge. All reactions in the database are stoichiometrically and charge-balanced and reaction directionality is added if it is available. *SABIO-RK* [145], the biochemical reaction kinetics database, is a curated database containing biochemical reactions and their corresponding kinetics. It describes the participants and modifiers of the reactions as well as measured kinetic data, such as kinetic rate equations, embedded in an experimental and environmental context. The *Transporter Classification Database* [143] (TCDB) provides a functional and

phylogenetic classification of membrane transport proteins. The classification system used is the transporter classification (TC) system that is analogous to the Enzyme Commission (EC) number for enzymes. The database is curated with data from over 10,000 published references. It contains over 5,600 unique protein sequences that are classified in more than 600 transporter families. *TransportDB* [144] contains the predicted cell membrane transport protein complement for over 365 organisms (bacteria, archaea, and eukaryota). The protein classification is done according to the TC classification system.

Genome Databases

Genome databases contain nucleotide sequences and functional annotations. GenBank [147], run by the National Center for Biotechnology Information, is a genetic sequence database of all publicly available DNA sequences. It contains the bibliographic and biological annotated sequences from more than 260,000 organisms.

Protein and Enzyme Databases

Protein and enzyme databases collect functional information from proteins and enzymes. BRENDA, the Braunschweig Enzyme Database [152], is a collection of functional and property data of enzymes. The majority of the contained data is manually extracted from primary literature and covers information in over 50 data fields, such as classification and nomenclature; reaction and specificity; information on function, structure, occurrence, preparation, and application of enzymes; and properties of mutants and engineered variants. Enzymes in BRENDA are linked to their respective pathways, source organism, and protein sequence, if deposited. *UniProt*, the universal protein resource, contains information on protein sequences and annotation data. It comprises four databases, namely the UniProt Knowledgebase (UniProtKB), the UniProt Reference Clusters (UniRef), the UniProt Archive (UniParc), and the UniProtMetagenomic and Environmental Sequences (UniMES) database on metagenomic and environmental data. UniProtKB is a collection of functional information on proteins together with annotation. The core data available for each protein are its amino acid sequence, protein name or description, taxonomic data, and citation information. Additionally, it contains as much annotation information as possible, such as ontologies, classifications, and cross-references, together with an indication of annotation quality. The database consists of two sections: UniProtKB/SwissProt contains reviewed and manually annotated records, whereas UniProtKB/TrEMBL has data records that are unreviewed and automatically annotated and still await full manual annotation. In the April 2013 release, UniProtKB/Swiss-Prot contained 539,829 sequence entries and UniProtKB/TrEMBL contained 33,106,277 sequence entries. UniRef is a database providing clustered sets of sequences from UniProtKB, including splice variants and isoforms and selected UniParc records. Its purpose is to obtain the complete coverage of sequence space at several resolutions. UniParc contains most of the publicly available protein sequences.

The planning of a biochemical synthesis involves, besides other aspects, the determination of possible side reactions. Those unwanted reactions may lead to a decrease in the yield of the desired product and complicate the downstream processing. It would thus be favorable to find information on all reactions catalyzed by the enzyme of interest including thermodynamic and kinetic parameters. Some of the enzyme resources presented in Table 4 contain information on single enzymes. However, as of 2013, there is no database that presents such information in a systematic manner.

Pathway Databases

Pathway databases contain data on chemical pathways, their reactions, and components that are involved in them and their corresponding interactions, thus describing the biochemistry of metabolic processes. These databases offer the possibility of providing several types of information in the context of graphical representation of the pathways in pathway maps. BioCyc [161] is a collection of 2038 pathway/genome databases in version 16.5. Each database contains the genome and metabolic pathways of a single organism. Based on the quality of the data, the databases are divided into three tiers. Tier 1 contains databases that are curated based on literature data. Tier 2 and tier 3 databases contain computationally predicted metabolic pathways, predictions as to which genes code for missing enzymes in metabolic pathways, and predicted operons. Tier 2 undergoes moderate curation and tier 3 is not curated at all.

The *Kyoto Encyclopedia of Genes and Genomes* (KEGG) is a curated database resource that integrates genomic, chemical, and systemic function information of various organisms. Its knowledge base consists of 15 main databases in the three categories systems information (pathway, brite, module, disease, drug, environ), genomic information (orthology, genome, genes) and chemical information (compound, glycan, reaction, rpair, rclass, enzyme). *BioCyc* is a collection of 2,038 pathway/genome databases in version 16.5. Each database contains the genome and metabolic pathways of a single organism.

The *BioPath* database contains molecules, reactions, and biological pathways. Its first version is based on the Roche Applied Science "Biochemical Pathways" wall chart and is extend with additional information from literature. In its release SVN-1285 from December 24, 2012, *BioPath* contained 14,368 molecules, 3,912 reactions, and 638 pathways. A new version of the book form was also published [169].

The University of Minnesota Biocatalysis/Biodegradation database *UM-BBD* contains information about microbial biocatalytic reactions and biodegradation pathways for xenobiotic compounds [164]. Information on microbial enzyme-catalyzed reactions that are important for biotechnology can also be found. A Swiss bioinformatics group has opened their database for automated model construction and genome annotation for large-scale metabolic networks, providing links to several hundred genome-scale metabolic networks [165].

Table 5 Bioinformatic tools for automated reconstruction of metabolic network models and their analysis

	URL	Content
Automated reconstruction		
Model SEED [170, 171]	http://www.theseed.org/	A comparative genomics environment for curation of genomic data
Pathway Tools [172]	http://bioinformatics.ai.sri.com/ptools/	A symbolic systems biology software system
KOBAS [173]	http://kobas.cbi.pku.edu.cn/	KEGG Orthology-Based Annotation System
ASGARD	http://sourceforge.net/projects/asgard-bio/	Software for metabolic pathway reconstruction and sequence annotation
GLAMM [174]	http://glamm.lbl.gov/	Genome-Linked Application for Metabolic Maps
GEMSiRV [175]	http://sb.nhri.org.tw/GEMSiRV/en/GEMSiRV	Software platform for genome-scale metabolic models simulation, reconstruction, and visualization
ERGO	http://ergo.integratedgenomics.com	Genome Analysis and Discovery System for the in silico analysis of organisms
Analysis		
CellNetAnalyzer [176]	http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html	MATLAB package for structural and functional analysis of biochemical networks
Metatool [177]	http://pinguin.biologie.uni-jena.de/bioinformatik/networks/	Computation of structural properties of biochemical reaction networks
Efmtool [178]	http://www.csb.ethz.ch/tools/efmtool	Computation of elementary flux modes of metabolic networks
COBRA Toolbox [179, 180]	http://opencobra.sourceforge.net/	COnstraints-Based Reconstruction and Analysis

Model Databases

Model databases are repositories of mathematical models of biological systems. They contain models ranging from reconstructions of individual pathways up to genome-scale metabolic networks of organisms. The *BiGG* database is a knowledgebase of biochemically, genetically, and genomically structured genome-scale metabolic network reconstructions [166]. It currently contains 10 different genome-scale reconstructions from eight organisms that are developed on the Genomatica Simpheny platform and share a standard nomenclature. BiGG allows browsing the model contents, the visualization of metabolic pathway maps, as well as the export of Systems Biology Markup Language (SBML) files of all models. *BioModels* [167] is a repository for computational models of biological systems. In its 24th release in December 2012, it contained 923 curated and noncurated models published in peer-reviewed literature and 142,050 models built from the *Path2Models* project that generates models from pathways. The database features browsing of models through lists, based on Gene Ontology (GO) terms or using the taxonomy annotations as well as model search. The model presentation gives access to all information stored about a model. All models can be exported in various file formats or represented graphically. Basic model simulation is also possible. The functionality of *BioModels* can also be accessed by other software tools through its web services.

Organism-Specific Databases

Information about specific organisms are often collected in organism-specific databases. Examples for such databases are *EcoCyc* [168], which contains *E. coli* data, or the *S. cerevisiae* Genome Database *SGD* [161].

4.2 Bioinformatic Tools

For the automated reconstruction of networks and analysis of network reconstructions, various bioinformatic tools are available. A selection is listed in Table 5.

Model SEED [170, 171] is a web-based resource for the high-throughput generation, optimization, and analysis of genome-scale metabolic network models. It integrates and augments technologies for the genome annotation, the construction of gene–protein reaction associations, the generation of biomass reactions, reaction network assembly, thermodynamic analysis of reaction reversibility, and model optimization to generate draft genome-scale metabolic network models. The generation of a metabolic network reconstruction from the assembled genome sequence takes about 48 h and automates nearly every step.

PathwayTools [172] is a software environment for the creation of pathway/ genome databases (PGDB) such as *EcoCyc* [168]. It allows the prediction of metabolic pathways and operons and network gap filling. Curators can interactively

edit PGDBs. A large number of query and visualization tools as well as tools for comparative and systems biology analyses are available. *Pathway Tools* consists of three components. *PathoLogic* is used to create new PGDBs from annotated genomes. The *Pathway/Genome Editors* allow for the refinement of PGDBs. With the *Pathway/Genome Navigator* querying, visualization and analyses of PGDBs can be carried out.

Metatool [177] is an user-friendly tool for the calculation of elementary flux modes, conservation relations, and enzyme subsets in metabolic networks. Version 5.1 can be embedded into GNU Octave and Matlab through script files and shared libraries. For calculations, the metabolic network data can be supplied to the program through the Metatool input format, as an SBML file or as stoichiometric matrix directly. *CellNetAnalyzer* [176] is a Matlab package with tools for the structural and functional analysis of different types of biochemical networks. For all computations, only the network topology is needed. *CellNetAnalyzer* allows the construction, input, and output of network projects via the *Network Composer*, text files, or SBML. Furthermore, it is possible to visualize network maps, either through import from KEGG or TRANSPATH or with external drawing tools. The functional network analysis covers the characterization of functional states of a network, the detection of functional dependencies, or qualitative predictions on effects of perturbations. For mass flow networks, there are two kinds of methods—namely constraint-based approaches and graph-theoretical analysis. Features are topological properties of the network such as dead-end metabolites, blocked or parallel reactions, and enzyme subsets. Metabolic flux analysis is also covered with the computation of steady-state flux distributions, feasibility check of flux scenarios, or optimal flux distributions for arbitrary linear objective functions. The computation of elementary modes for the metabolic path analysis is also possible. Minimal cutset analysis can help to detect strategies for the repression of certain network functionality. From the graph-theoretical side, network properties such as shortest path lengths, connectivity of the network, or network diameter can be computed.

The *Genome-linked Application for Metabolic Maps* [174] (*GLAMM*) is a web interface unifying different tools for the reconstruction of metabolic networks from annotated genome data, visualization of metabolic networks together with experimental data, and investigation of the construction of novel transgenic pathways. GLAMM supports biological retrosynthesis and integration with tools of *MicrobesOnline*. The genome analysis and discovery system ERGO developed by Integrated Genomics is a systems biology informatics toolkit for comparative genomics. With ERGO, one can capture, query, and visualize sequenced genomes and assign functions to genes, integrate genes into pathways, and identify unknown genes, gene products, and pathways. Its genomic database integrates with a collection of microbial metabolic and nonmetabolic pathways and proprietary algorithms. ERGO allows automated or manual annotation of genomes and genes, pathway analysis, multiple genome comparison, functional analysis of microarray data, data mining for the discovery of target genes, and in silico metabolic engineering and strain improvement.

4.3 Network Reconstruction

A metabolic network reconstruction is a structured database combining the available genetic, genomic, and biochemical data of an organism [181]. In general, a genome-scale metabolic network reconstruction consists of a list of reactions including their stoichiometry, the specific genes whose gene products are associated with these reactions, supporting annotation, and literature references. The fundamental goal of a network reconstruction is the accurate definition of the chemical transformations that take place among the chemical components of the network [181]. The construction and curation of a computational network links the organism's genome and expression to metabolic reaction fluxes, biomass, and energy production and consumption and enables the mathematical representation of the reactions and metabolic processes occurring in the organism. Metabolic networks can thus be used for in silico experiments [182]. The process of compiling a (genome-scale) metabolic network can be broken up into five major stages [183], as depicted in Fig. 7. Briefly, in the first stage, a draft reconstruction of the network is built, which is refined in the second step. Then the network is converted to a mathematical model. In the fourth step, the reconstructed network is validated and can then be used for further experiments.

The first stage consists of creating a draft reconstruction of the network, at minimum containing a list of genes with their associated reactions and the corresponding EC numbers. The draft is based on the genome annotation for the most recent version of the target organism's genome and data from biological databases (Table 4) and results in a collection of the genome-encoded functions of the metabolism. The important information for each gene is its function, its position in the genome and coding region, the strand and locus names, and the protein it codes for. For eukaryotes, information regarding alternative transcripts is also of interest because they may have distinct functions or a different cellular localization [183]. Candidate metabolic functions for the draft reconstructions can be retrieved by using GO categories, EC numbers, and biochemical databases [183] (Table 4). In general, the creation of a network draft is carried out automatically with the help of different software tools (Table 5). The automated genome annotation process with ERGO provides a draft annotation that requires manual curation to add organism-specific information. BLAST is used to annotate gene function based on orthology with other annotated genomes provided in online databases as well as phylogenetic approaches. *Model SEED* starts with an unannotated genome sequence and builds a draft metabolic network with gap filling and verification features [184]. To refine this draft, a manual reconstruction refinement step is necessary.

The manual refinement step starts with an initial evaluation of the completeness of the draft reconstruction for identifying missing functions in the network. The draft can be reviewed pathway by pathway, starting from canonical pathways. The reactions of the model are evaluated in their metabolic context such that missing gene annotations and missing reactions can be identified easier. The use of network maps that show the environment of reactions is also convenient. Such maps can be
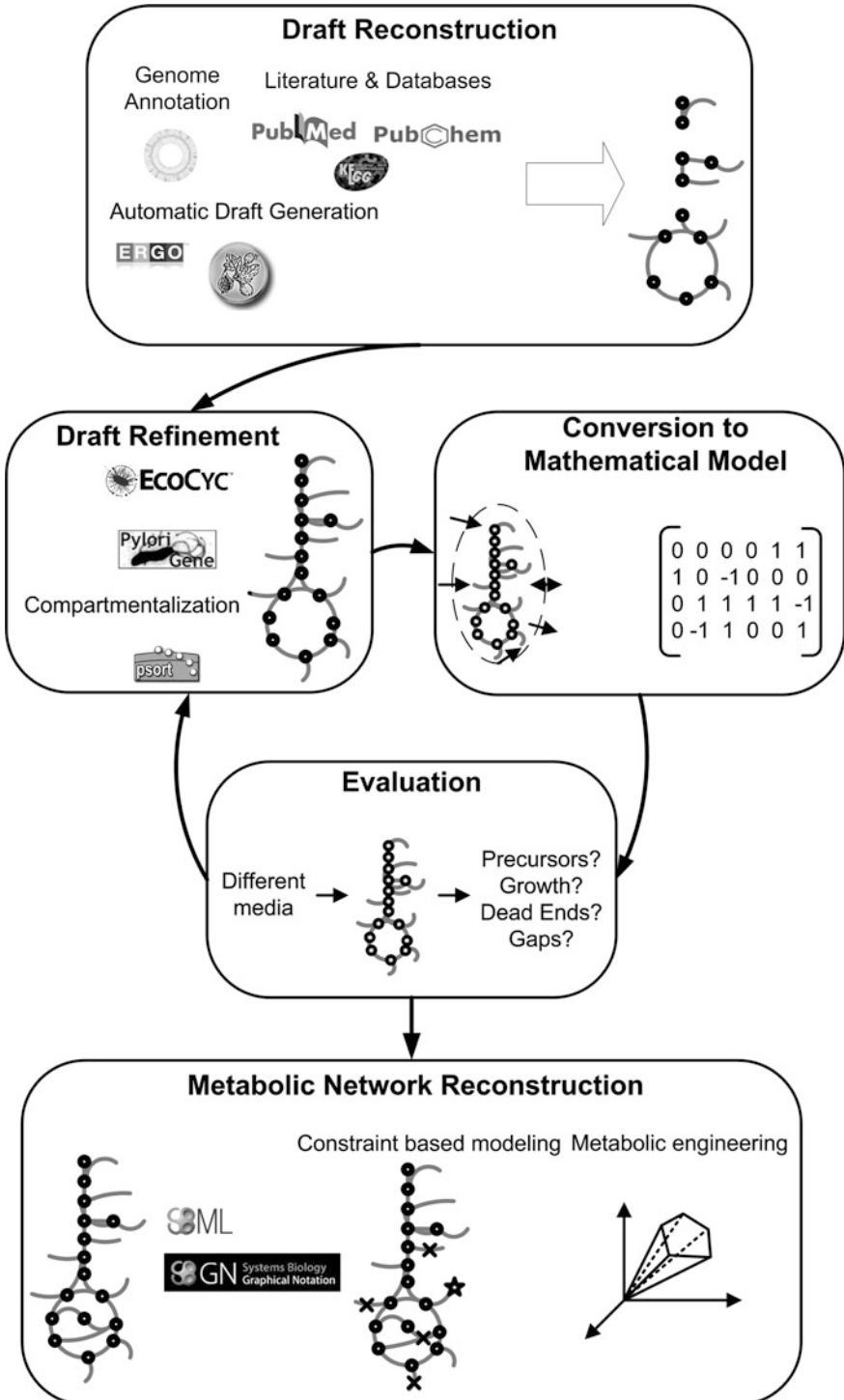
Fig. 7 Workflow of an iterative network reconstruction process

found in databases such as KEGG or in organism-specific literature [183]. Correct stoichiometry requires complete balancing of elements and charges. Some databases may lack information on protons and water, for example. The reaction directions and reversibilities should also be evaluated. If no information is available, the reaction should be left reversible. Organism-specific functions should also be taken into account, such as the use of substrate and cofactors, which can differ between different organisms. The review of primary literature dealing with the metabolism and function of the organism is necessary to identify these organism-specific characteristics. For a growing number of organisms, specific books exist, which are a resource for additional information. When organism-specific information is not available to the desired extent, data from phylogenetic neighbors can be taken into account. Gene–protein reaction relationships, which connect the genes with their associated enzymes via Boolean logic, allow the simulation of phenotypic effects of gene knockouts. The incorporation of thermodynamic information is also of great value for the network model. The directionality and feasible rates of reactions are based on the thermodynamic favorability, which can be determined from Gibbs free-energy changes. This information is obtained from literature. However, the available data are rarely sufficient for genome-scale reconstructions but rather for smaller models. Also, the compartmentalization information for metabolite and reaction localization as well as intracellular transport reactions have to be checked. If no sufficient data are available, the respective proteins should be assumed to reside in the cytosol. However, incorrect assignment can lead to additional network gaps [183]. Furthermore, biomass composition, maintenance parameters, and growth conditions of the organism are to be determined by different experimental and computational methods.

The conversion from the network reconstruction to a mathematical model for validation and in silico applications consists of three steps, which can mostly be automated with suitable tools. The first step is the mathematical representation of the network as a stoichiometric matrix. In the second step, the boundaries of the system are defined. For each metabolite that can be consumed or secreted, an exchange reaction is added to allow the definition of environmental conditions for in silico simulations. Finally, constraints are added to the model to turn it into a condition-specific model. Thermodynamic data for enzyme capacities or regulation help to determine a set of feasible steady-state flux solutions.

The evaluation stage includes network verification, evaluation, and validation steps to help detect gaps in the network. To find candidates for filling gaps, an intensive literature search is needed that helps to identify the environment of the dead-end metabolites. One must also take care of stoichiometrically balanced cycles formed by internal network reactions that can carry fluxes despite closed exchange reactions [183]. The model must be tested for its ability to synthesize all biomass precursors, such as amino acids, nucleotide triphosphates, or lipids with different medium compositions. This can be done by growing the organism on specific carbon sources [184]. It should be checked if the model could reproduce known incapabilities of the organism. It is also advised to compare the predicted

physiological properties with known properties such as carbon splits in the central metabolic pathways of the organism.

Network reconstructions can be used for several major applications that address different aims of these models [185], such as using metabolic network reconstructions for putting high-throughput experimental data into context. They can also be used for discovery of network properties, hypothesis-driven discovery, and exploration of multispecies relationships. Network reconstructions also have applications in metabolic engineering, where they can be used for constraint-based modeling and the in silico prediction of possible cellular phenotypes without the need for kinetic data. The main concept behind network metabolic modeling is the identification and mathematical definition of constraints for the separation of feasible and infeasible metabolic behavior. These constraints are usually much easier to identify than kinetic parameters. There are three types of constraints: Physicochemical constraints deal with mass and energy conservation, the dependency of reaction rates on metabolite concentrations, and the negative free-energy change for spontaneous reactions. Environmental constraints are imposed as a result of specific conditions such as nutrients, whereas regulatory constraints express the effects of gene expression and enzyme activity regulation properties.

4.4 Network Representation

Metabolic network models can formally be described as graphs $G(V, E)$, where V is the set of vertices and E is the set of edges connecting node pairs. In a directed graph, the edges are ordered, whereas in an undirected graph an edge is represented by an unordered node pair. There are different possibilities to represent a graph. In a *compound graph,* the nodes represent the chemical compounds. An (un)directed edge connects two compounds if they are involved in the same reaction. The dual form of the compound graph is the *reaction graph*, where the nodes represent reactions. Edges in this graph connect two reactions that share compounds as products and substrates. Both types of representations have similar limitations, as they are both ambiguous and do not represent all information of the network. Another type of graph representation is the *bipartite graph*, in which there are two classes of nodes representing compounds and reactions in the graph. Directed or undirected edges in a bipartite graph are only possible between two nodes of different classes. A substrate is defined by a directed edge from a compound to a reaction node, a product by an edge from a reaction to a compound node. An equivalent representation of a bipartite graph and also the generalization of a compound graph is a *hypergraph* with directed or undirected edges. In such a graph, a hyperedge relates a set of substrates to a set of products. This graph type allows an unambiguous representation of reactions and compounds, but it has limited coverage because reaction control factors cannot be represented. Diverse graph types and data models for biochemical pathways are reviewed in more detail in [186].

A basic mathematical representation of metabolic networks is the stoichiometric matrix. It represents its charge and elementary balanced metabolic reactions and thus quantifies the stoichiometric relationship between the metabolites in a reaction. The rows and columns of the matrix correspond to the metabolites and reactions of the network. Its nonzero elements are the stoichiometric coefficients, which are positive for products and negative for substrates. For genome-scale metabolic networks, the stoichiometric matrix is sparse because relatively few metabolites participate in a given reaction. Another quite simple representation method for metabolic networks is compiling all information in spreadsheets. The spreadsheet should contain all gene names and their abbreviations. For each reaction, the reactant, substrate, product symbols, balanced stoichiometry, reversibility, compartment, associated protein, and its EC number should be included. Also of importance are literature references and a confidence rating for each annotation entry as well as comments.

SBML is a machine-readable format for representing biological models. Its basic idea is to cast a network reconstruction into a formal, computable form, thus allowing network analysis using simulations and other mathematical methods. SBGN [187], the *Systems Biology Graphical Notation*, is a project that aims to standardize the graphical notation used in maps of biological processes. Currently, there are three different languages for different types of network maps. The Process Description language can be used to depict temporal courses of biochemical interactions of a network. Relationships between entities of a network can be modeled with the Entity Relationship language. The Activity Flow language visualizes the information flow between biochemical entities in the network, such as for representation of the effects of perturbations on the network.

4.5 Metabolic Network Design and Manipulation

The planning of a biochemical synthesis in permeabilized cells involves primarily finding a synthetic route starting from available, inexpensive, and stable substrates. The overall goal is to obtain high concentration with high yield and selectivity in the shortest possible time. High yield and selectivity can only be obtained if undesired side reactions do not take place. The products of undesired side reactions will also complicate downstream processing. Side reactions do not occur if one of the substrates required is missing. The substrate composition is a design variable when using permeabilized cells as biocatalysts. Side reactions can be eliminated by the deletion of the corresponding gene or by the addition of a selective inhibitor. A major engineering task of biosynthesis in permeabilized cells is the regeneration of cofactors, such as NAD(P)H [188] or ATP [189, 190].

Metabolic engineering is the manipulation of enzymatic, transport, and regulatory functions of a cell through recombinant DNA technologies. One of its important objectives is the improvement of the cellular phenotype or the yield of a

desired product. Traditionally, this is done by rationally selected gene deletions or overexpression of native and heterologous genes in an organism. To remove undesirable metabolic pathways in an organism, site-directed mutagenesis or homologous recombination can be used. To increase biochemical yields and add new functions, heterologous genes or even complete pathways can be introduced into the organism. In silico metabolic models allow rational predictions of the phenotypical response of changes in culture media, gene knockouts, and the incorporation of heterologous enzymes and pathways into an organism [191].

Flux balance analysis is a widely used constraint-based method in metabolic engineering for studying biochemical networks. It allows for the in silico prediction of flux profiles that optimize a cellular objective, depending on the problem. Often, the biomass production or the production rate of a certain metabolite of interest is used as an objective. The fundamental assumption for flux balance analysis is that the metabolism in the cell is at steady state as well as all reaction fluxes and metabolite concentrations [184]. The input for flux balance analysis is the mathematical representation of the metabolic network as a stoichiometric matrix. The stoichiometric coefficients in the matrix constrain the flow of metabolites through the network. These steady-state mass balance equations for each metabolite and the environmental and growth conditions can be described mathematically in the form of constraints for an optimization problem. The metabolite balance equation is a homogeneous system of linear equations:

$$S \cdot v(t) = 0$$

where S is the stoichiometric matrix and $v(t)$ is the vector of reaction rates. It requires that each metabolite is consumed at the same rate as it is produced [192]. For the quantitative determination of the metabolic fluxes, linear programming can be used to solve the stoichiometric matrix for a given objective function under various constraints [193]. The constraints of the problem describe the space of all eligible possibilities from which an optimal solution can be selected. They are generally given in the form of equalities and/or inequalities.

Constraint-based methods focus only on reaction fluxes, neglecting enzyme kinetics and regulations that can influence the actual fluxes; therefore, they have some limitations in their predictive capabilities [194]. However, they can be computed very efficiently, even for large networks. *Elementary mode analysis* is an important method for metabolic network studies [177], allowing the enumeration of all independent minimal pathways in the network that are stoichiometrically and thermodynamically feasible. Elementary flux modes are independent flux distributions of a metabolic network at steady state. The inputs for elementary mode analysis are the reaction stoichiometries and reversibilities. All metabolites in the network are classified as either internal or external. The internal metabolites are balanced and the external metabolites are assumed to be buffered. The computation of elementary modes in large networks is difficult due to its combinatorial complexity [176]. Once elementary modes are computed, the deletions necessary for the elimination of undesired side product formation can be directly identified. It
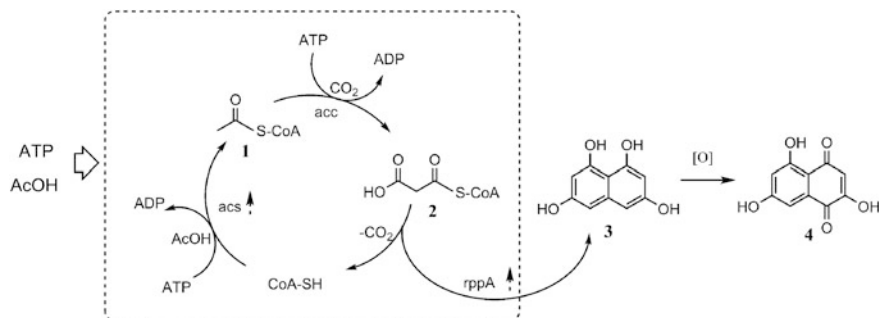
Fig. 8 Multienzyme synthesis of tetra-hydroxy-naphthalene (THN, *3*) and subsequent oxidation to flaviolin (*4*). ATP and acetate are supplied with catalytic amounts of coenzyme A. Acetyl-CoA (*1*) is formed rapidly by overexpressed acetyl-CoA synthase (acs) and carboxylated by the acetyl-CoA carboxylase complex (acc) to yield malonyl-CoA (*2*), which serves as monomer for the polyketide backbone of THN
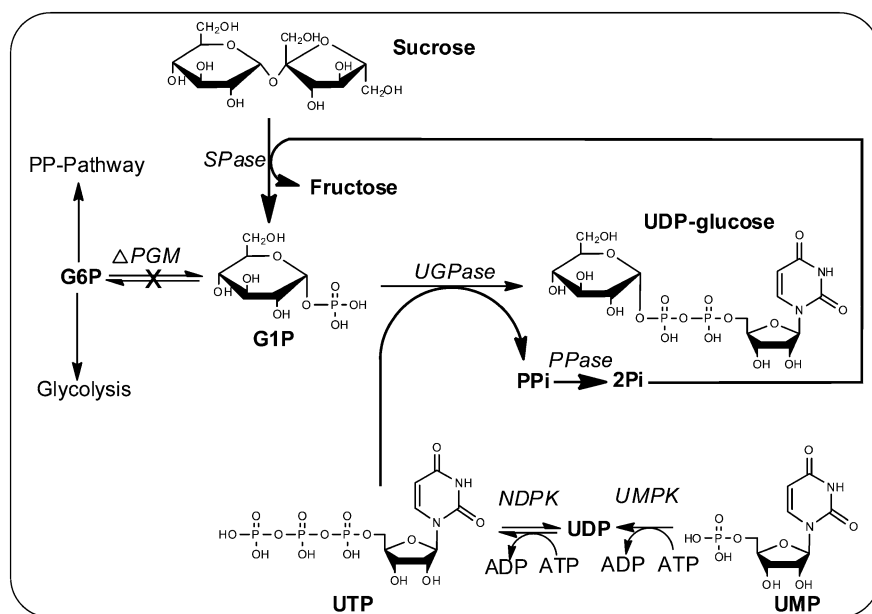
Fig. 9 Reaction network of UDP-glucose biosynthesis in *E.coli* JW 0675-1 SP [198]. Enzymes: *UMK* uridine monophosphate kinase (EC 2.7.4.22), *NDK* nucleoside-diphosphate kinase (EC 2.7.4.6), *UGPase* UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9), *PPase* inorganic diphosphatase (EC 3.6.1.1), *SPase* sucrose phosphorylase (EC 2.4.1.7), *Pgm* phosphoglucomutase (EC 5.4.2.2). Compounds: *UMP* uridine monophosphate, *g1p* glucose-1-phosphate, *UDP-glucose* uridine diphosphate glucose

has been shown that it is even possible to directly identify successfully targets for the overexpression of enzymes just based on the known stoichiometry of a network [89, 90, 195].
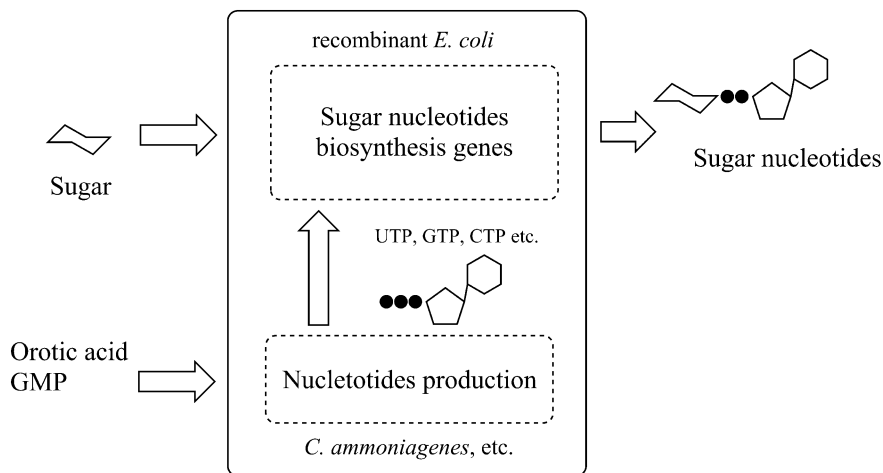
Fig. 10 Artificial biosynthesis of sugar-nucleotides by Kyowa Hakko Bio. Nucleotide-rich strains, such as *C. ammoniagenes*, are used for the biosynthesis of nucleotide precursors. The charged metabolites are enabled to diffuse between the producer strain and tailored *E. coli*, carrying synthesis genes for the corresponding sugar nucleotides (Figure adapted from [199–201])

5 Applications

Directed multistep biocatalysis with permeabilized whole cells is a new strategy for the synthesis of fine chemicals. However, only a few publications are currently available. In this section, a few successful cases are described briefly.

Multienzyme whole-cell in situ biocatalysis for the production of the polyketide secondary metabolite flaviolin [196] was realized in tailored *E. coli* (Fig. 8) [16]. The permeabilized *E. coli* strain synthesized significant amounts of the target molecule while consuming acetate and ATP in the presence of coenzyme A. The productivity of 1 g of permeabilized cells could be determined as 0.025 µmol of flaviolin per minute. Cells were able to produce half the amount of flaviolin synthesized by a corresponding fermentation with the original producer strain but within one-third of time. The isolation of flaviolin was very easy because the reaction buffer only contained a few defined components with known properties. Only a small amount of protein—approximately 0.5 % of the total cellular protein—was found in the reaction buffer after separating the biocatalyst. It might derive from released membrane protein.

Using a similar approach, a different *E. coli* strain was tailored for the synthesis of uridine diphosphate (UDP)-glucose. Sucrose phosphorylase (SPase) from *Leuconostoc mesenteroides* [197] was overexpressed in a strain lacking phosphoglucomutase. Sucrose phosphorylase catalyzes the formation of glucose-1-phosphate from sucrose and phosphate. Because phosphoglucomutase was knocked out, glucose-1-phosphate could not be isomerized to glucose-6-phosphate, thus
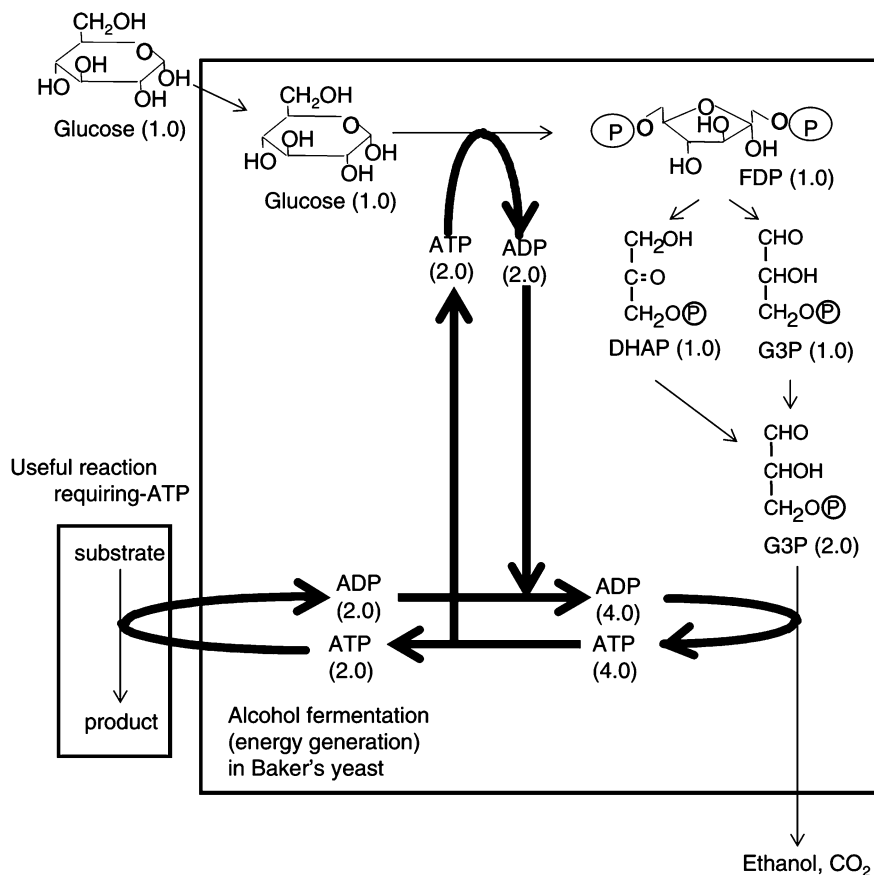
Fig. 11 Glucose-based ATP regeneration system using permeabilized yeast cells, as described by Horinouchi [190]

preventing any direct carbon flux into the glycolysis [198]. Using this strain, we could directly produce UDP-glucose using sucrose, ATP, and uridine monophosphate (Fig. 9).

Reaction conditions such as temperature, pH, and cell density were optimized for this synthesis. Surprisingly, it was found that manganese was better suited than magnesium as metal cofactor for the synthesis. According to literature data on the involved enzymes, this was unexpected and indicated that enzymes in permeabilized cells can behave differently than the isolated enzymes. The rate of UDP-glucose formation was 0.052 mmol min^{-1} L^{-1}. The catalyst could be separated from the reaction mixture by short centrifugation. A comparison of the protein contents of cell lysate and of samples taken after 2 h of reaction revealed that less than 2 % of cellular protein was released.

Kyowa Hakko Bio uses a similar approach for the synthesis of nucleotide sugars and peptides. A cofermentation of *Corynebacterium ammoniagenes* and

tailored *E. coli* is supplemented with a minimum amount of surfactant [199–201]. This strategy is quite similar to the multistep approach and benefits substrate interchange between the organisms (Fig. 10). However, use of the permeabilization procedure enhances the efficiency of nucleotide-triphosphate uptake by *E. coli* cells, which perform one biosynthetic step. To our knowledge, this is the only industrial application where permeabilized cells are synthetically used. Although the principle of permeabilization is advantageous in this case, the process lacks selectivity. The product is diluted in the reaction media containing the surfactant and nutrients.

Horinouchi et al. developed an efficient system for ATP regeneration and coupled it to the synthesis of 2-deoxyribose 5-phosphate and further to 2′-deoxyribonucleoside [189, 190]. ATP regeneration is driven by glycolysis of permeabilized baker's yeast as depicted in Fig. 11.

The combination of NADPH regeneration with permeabilized cells has been examined by Zhang et al. [25, 34]. While reducing β keto esters to β hydroxy esters, NADPH is consumed. This expansive cofactor is then regenerated by coupling $NADP^+$ reduction to the oxidation of glucose to gluconolactone. Permeabilized cells of *Bacillus pumilus*, *Bacillus subtilis,* and *E. coli* were used in these studies. Using yeast *Candida ontarioensis*, Ni et al. reduced several acetophenones to the corresponding alcohols while regenerating NADPH with glucose oxidation. However, significant product yields were also observed when no cosubstrate was added [202].

6 Concluding Remarks

Directed multistep biocatalysis with permeabilized cells promises to close the gap between in vivo and in vitro biotechnology. The combination of two major developments significantly extends the application of permeabilized cells for more complex biosynthesis beyond single-step or dual-step coupled bioconversions. These developments are in the fields of genome sequencing and related bioinformatic developments, including databases and computational tools. Only with the advent and operability of whole-genome scale metabolic models, it becomes feasible to predict the function of a metabolic network in silico with high success rates [92]. However, methods of molecular biology for manipulating metabolic networks have become increasingly powerful and efficient [203]. Methods to study metabolic pathway activities have developed significantly, particularly metabolic flux analysis, thus providing reliable information for validating in silico models.

Multistep biosynthesis provides opportunities to design and develop one-pot biosynthetic processes for more molecules with more complex structures [16]. Particularly for the synthesis of complex secondary metabolites and derivatizations of such, directed multistep approaches could simplify processes and avoid extensive waste production. This technology is still on the edge of a breakthrough and needs considerable research efforts to become more applicable on a larger

scale. Although the main effort for multistep biocatalyst design is on the development of a functioning reaction setup, most permeabilization procedures are very simple and applicable to any cellular organism. However, there is still limited knowledge on the molecular-scale interaction of surfactants with biological membranes. Molecular processes have only been studied in detail for a few combinations of organisms and permeabilization procedures. A more profound and general understanding of these mechanisms might help to optimize production processes and aid in finding optimal conditions for the permeabilization of specific organisms in a more directed way. New methods of nanostructure imaging might resolve the unknowns of interaction mechanisms. New effective and highly selective permeabilization methods might further improve the handling of permeabilized cells and allow on-shelf storage for better availability. Increasing knowledge on the enzymatic potential further gives opportunities to tailor megasynthetase biocatalysts for the synthesis of specific compounds. A remaining bottleneck is in the limited availability of detailed biochemical knowledge, such as the selectivity of certain enzymes and their regulation. Resolving new biosynthetic routes in enzyme clusters could lead to designed megasynthetases for hybrid structures, as the modular building principle can allow the combination of several clusters for completely new biosynthetic routes without a major engineering effort.

References

1. Lopez-Gallego F, Schmidt-Dannert C (2010) Multi-enzymatic synthesis. Curr Opin Chem Biol 14(2):174–183
2. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, Clair JLS, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL, Houk KN, Michael FE, Baker D (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. Science 329(5989):309–313. doi:10.1126/science.1190239
3. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K (2012) Engineering the third wave of biocatalysis. Nature 485(7397):185–194. doi:10.1038/Nature11117
4. Liese A, Seelbach K, Wandrey C (2000) Industrial biotransformations. Wiley-VCH, New York
5. Wohlgemuth R (2011) Molecular and engineering perspectives of the biocatalysis interface to chemical synthesis. Chem Biochem Eng Q 25(1):125–134
6. Wohlgemuth R (2010) Biocatalysis—key to sustainable industrial chemistry. Curr Opin Biotechnol 21(6):713–724. doi:10.1016/j.copbio.2010.09.016
7. Heinzle E, Biwer AP, Cooney CL (2006) Development of sustainable bioprocesses. Wiley-VCH, New York
8. Drauz K, Waldmann H (2012) Enzyme catalysis in organic synthesis. Wiley-VCH, New York
9. Aehle W (2004) Enzymes in industry. Wiley-VCH, New York
10. Schneider K, Dorscheid S, Witte K, Giffhorn F, Heinzle E (2012) Controlled feeding of hydrogen peroxide as oxygen source improves production of 5-ketofructose From L-sorbose using engineered pyranose 2-oxidase from Peniophora gigantea. Biotechnol Bioeng 109(11):2941–2945. doi:10.1002/bit.24572

11. Kuhn D, Kholiq MA, Heinzle E, Buehler B, Schmid A (2010) Intensification and economic and ecological assessment of a biocatalytic oxyfunctionalization process. Green Chem 12(5):815–827. doi:10.1039/b921896c

12. You C, Zhang YH (2013) Cell-free biosystems for biomanufacturing. Adv Biochem Eng Biotechnol 131:89–119. doi:10.1007/10_2012_159

13. Moses T, Pollier J, Thevelein JM, Goossens A (2013) Bioengineering of plant (tri)terpenoids: from metabolic engineering of plants to synthetic biology in vivo and in vitro. New Phytol. doi:10.1111/nph.12325

14. Minton AP (2006) How can biochemical reactions within cells differ from those in test tubes? J Cell Sci 119(14):2863–2869. doi:10.1242/Jcs.03063

15. Monti D, Ferrandi EE, Zanellato I, Hua L, Polentini F, Carrea G, Riva S (2009) One-pot multienzymatic synthesis of 12-ketoursodeoxycholic acid: subtle cofactor specificities rule the reaction equilibria of five biocatalysts working in a row. Adv Synth Catal 351(9):1303–1311. doi:10.1002/adsc.200800727

16. Krauser S, Kiefer P, Heinzle E (2012) Multienzyme whole-cell in situ biocatalysis for the production of flaviolin in permeabilized cells of Escherichia coli. ChemCatChem 4(6):786–788. doi:10.1002/cctc.201100351

17. Ye X, Honda K, Sakai T, Okano K, Omasa T, Hirota R, Kuroda A, Ohtake H (2012) Synthetic metabolic engineering-a novel, simple technology for designing a chimeric metabolic pathway. Microb Cell Fact 11:120. doi:10.1186/1475-2859-11-120

18. Felix H (1982) Permeabilized cells. Anal Biochem 120(2):211–234. doi:10.1016/0003-2697(82)90340-2

19. Ma HW, Zeng AP (2003) The connectivity structure, giant strong component and centrality of metabolic networks. Bioinformatics 19(11):1423–1430

20. Vaara M (1992) Agents that increase the permeability of the outer-membrane. Microbiol Rev 56(3):395–411

21. Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. Cold Spring Harb Perspect Biol 2(5):a000414. doi:10.1101/cshperspect.a000414

22. Cánovas M, Torroglosa T, Iborra JL (2005) Permeabilization of Escherichia coli cells in the biotransformation of trimethylammonium compounds into l-carnitine. Enzyme and Microbial Technology 37(3):300–308

23. Yuan YB, Heinzle E (2009) Permeabilization of Corynebacterium glutamicum for NAD(P)H-dependent intracellular enzyme activity measurement. C R Chim 12(10–11):1154–1162. doi:10.1016/j.crci.2009.09.006

24. Abraham J, Bhat S (2008) Permeabilization of baker's yeast with N-lauroyl sarcosine. J Ind Microbiol Biotechnol 35(8):799–804. doi:10.1007/s10295-008-0350-9

25. Zhang J, Witholt B, Li Z (2006) Coupling of permeabilized microorganisms for efficient enantioselective reduction of ketone with cofactor recycling. Chem Commun 4:398–400. doi:10.1039/b515721h

26. Malik M, Ganguli A, Ghosh M (2012) Modeling of permeabilization process in Pseudomonas putida G7 for enhanced limonin bioconversion. Appl Microbiol Biotechnol 95(1):223–231. doi:10.1007/s00253-012-3880-z

27. Tryfona T, Bustard MT (2008) Impact of pulsed electric fields on Corynebacterium glutamicum cell membrane permeabilization. J Biosci Bioeng 105(4):375–382. doi:10.1263/jbb.105.375

28. Sestak S, Farkas V (2001) In situ assays of fungal enzymes in cells permeabilized by osmotic shock. Anal Biochem 292(1):34–39

29. Kaur G, Panesar PS, Bera MB, Kumar H (2009) Hydrolysis of whey lactose using CTAB-permeabilized yeast cells. Bioprocess Biosyst Eng 32(1):63–67. doi:10.1007/s00449-008-0221-9

30. Ding Y, Lu H, Shi G, Liu J, Shen G, Yu R (2008) Cell-based immobilization strategy for sensitive piezoelectric immunoassay of total prostate specific antigen. Biosens Bioelectron 24(2):228–232. doi:10.1016/j.bios.2008.03.020

31. Niklas J, Melnyk A, Yuan YB, Heinzle E (2011) Selective permeabilization for the high-throughput measurement of compartmented enzyme activities in mammalian cells. Anal Biochem 416(2):218–227. doi:10.1016/j.ab.2011.05.039

32. Cook GA, Gattone VH, Evan AP, Harris RA (1983) Structural-changes of isolated hepatocytes during treatment with digitonin. Biochim Biophys Acta 763(4):356–367. doi:10.1016/0167-4889(83)90097-6

33. Aragon JJ, Feliu JE, Frenkel RA, Sols A (1980) Permeabilization of animal-cells for kinetic studies of intracellular enzymes—insitu behavior of the glycolytic-enzymes of erythrocytes. Proc Natl Acad Sci U S A (Biol Sci) 77(11):6324–6328. doi:10.1073/pnas.77.11.6324

34. Zhang W, O'Connor K, Wang DIC, Li Z (2009) Bioreduction with efficient recycling of NADPH by coupled permeabilized microorganisms. Appl Environ Microbiol 75(3):687–694. doi:10.1128/aem.01506-08

35. Zhang H, Witholt B, Li Z (2006) Efficient NADPH recycling in enantioselective bioreduction of a ketone with permeabilized cells of a microorganism containing a ketoreductase and a glucose 6-phosphate dehydrogenase. Adv Synth Catal 348(4–5):429–433. doi:10.1002/adsc.200505439

36. Owen DM, Williamson DJ, Magenau A, Gaus K (2012) Sub-resolution lipid domains exist in the plasma membrane and regulate protein diffusion and distribution. Nat Commun 3:1256. doi:10.1038/ncomms2273

37. Daines AM, Maltman BA, Flitsch SL (2004) Synthesis and modifications of carbohydrates, using biotransformations. Curr Opin Chem Biol 8(2):106–113. doi:10.1016/j.cbpa.2001.02.003

38. Datsenko KA (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97(12):6640–6645

39. Zaragoza A, Aranda FJ, Espuny MJ, Teruel JA, Marques A, Manresa A, Ortiz A (2009) Mechanism of membrane permeabilization by a bacterial trehalose lipid biosurfactant produced by Rhodococcus sp. Langmuir 25(14):7892–7898. doi:10.1021/la900480q

40. Jamur MC, Oliver C (2010) Permeabilization of cell membranes. Methods Mol Biol (Clifton, NJ) 588:63–66. doi:10.1007/978-1-59745-324-0_9

41. Womack MD, Kendall DA, Macdonald RC (1983) Detergent effects on enzyme-activity and solubilization of lipid bilayer-membranes. Biochim Biophys Acta 733(2):210–215. doi:10.1016/0005-2736(83)90524-2

42. Tandon S, Horowitz PM (1987) Detergent-assisted refolding of guanidinium chloride-denatured rhodanese—the effects of the concentration and type of detergent. J Biol Chem 262(10):4486–4491

43. Weber K, Kuter DJ (1971) Reversible denaturation of enzymes by sodium dodecyl sulfate. J Biol Chem 246(14):4504–4509

44. Galabova D, Tuleva B, Spasova D (1996) Permeabilization of Yarrowia lipolytica cells by Triton X-100. Enzyme Microb Technol 18(1):18–22. doi:10.1016/0141-0229(96)00063-4

45. Fiskum G, Craig SW, Decker GL, Lehninger AL (1980) The cytoskeleton of digitonin-treated rat hepatocytes. Proc Natl Acad Sci U S A (Biol Sci) 77(6):3430–3434. doi:10.1073/pnas.77.6.3430

46. Sperry WM, Webb M (1950) A revision of the Schoenheimer-Sperry method for cholesterol determination. J Biol Chem 187(1):97–106

47. Colbeau A, Nachbaur J, Vignais PM (1971) Enzymic characterization and lipid composition of rat liver subcellular membranes. Biochim Biophy Acta 249(2):462–492. doi:10.1016/0005-2736(71)90123-4

48. Cheng Q, Xiang L, Izumikawa M, Meluzzi D, Moore BS (2007) Enzymatic total synthesis of enterocin polyketides. Nat Chem Biol 3(9):557–558. doi:10.1038/nchembio.2007.22

49. Gowda LR, Bachhawat N, Bhat SG (1991) Permeabilization of Bakers' yeast by cetyltrimethylammonium bromide for intracellular enzyme catalysis. Enzyme Microb Technol 13(2):154–157

50. Upadhya R, Nagajyothi HBhat SG, Bhat SG (2000) Stabilization of D-amino acid oxidase and catalase in permeabilized Rhodotorula gracilis cells and its application for the

preparation of alpha-ketoacids. Biotechnol Bioeng 68(4):430–436. doi:10.1002/(sici)1097-0290(20000520)68:4<430:aid-bit8>3.0.co;2-3

51. McCord JM, Day ED (1978) Superoxide-dependent production of hydroxyl radical catalyzed by iron-EDTA complex. FEBS Lett 86(1):139–142. doi:10.1016/0014-5793(78)80116-1

52. Rundback F, Fidanoska M, Adlercreutz P (2012) Coupling of permeabilized cells of Gluconobacter oxydans and Ralstonia eutropha for asymmetric ketone reduction using H-2 as reductant. J Biotechnol 157(1):154–158. doi:10.1016/j.jbiotec.2011.09.029

53. Lehrer RI, Barton A, Daher KA, Harwig SSL, Ganz T, Selsted ME (1989) Interaction of human defensins with Escherichia-coli—mechanism of bactericidal activity. J Clin Invest 84(2):553–561. doi:10.1172/jci114198

54. Hill CP, Yee J, Selsted ME, Eisenberg D (1991) Crystal-structure of defensin HNP-3, an amphiphilic dimer—mechanisms of membrane permeabilization. Science 251(5000):1481–1485. doi:10.1126/science.2006422

55. Lichtenstein A (1991) Mechanism of mammalian-cell lysis mediated by peptide defensins—evidence for an initial alteration of the plasma-membrane. J Clin Invest 88(1):93–100. doi:10.1172/jci115310

56. White SH, Wimley WC, Selsted ME (1995) Structure, function, and membrane integration of defensins. Curr Opin Struct Biol 5(4):521–527. doi:10.1016/0959-440X(95)80038-7

57. van Kan EJM, Demel RA, Breukink E, van der Bent A, de Kruijff B (2002) Clavanin permeabilizes target membranes via two distinctly different ph-dependent mechanisms. Biochemistry 41(24):7529–7539. doi:10.1021/bi012162t

58. Powers J-PS, Hancock REW (2003) The relationship between peptide structure and antibacterial activity. Peptides 24(11):1681–1691. doi:10.1016/j.peptides.2003.08.023

59. Michalek M, Gelhaus C, Hecht O, Podschun R, Schroder JM, Leippe M, Grotzinger J (2009) The human antimicrobial protein psoriasin acts by permeabilization of bacterial membranes. Dev Comp Immunol 33(6):740–746. doi:10.1016/j.dci.2008.12.005

60. Aidemark M, Andersson CJ, Rasmusson AG, Widell S (2009) Regulation of callose synthase activity in situ in alamethicin-permeabilized Arabidopsis and tobacco suspension cells. BMC Plant Biol 9:27. doi:2710.1186/1471-2229-9-27

61. Alonso MA, Carrasco L (1982) Molecular basis of the permeabilization of mammalian cells by ionophores. Eur J Biochem 127(3):567–569. doi:10.1111/j.1432-1033.1982.tb06909.x

62. Kearsey SE, Brimage L, Namdar M, Ralph E, Yang X (2005) In situ assay for analyzing the chromatin binding of proteins in fission yeast. Methods Mol Biol (Clifton, NJ) 296:181–188

63. Ferrer P, Hedegaard L, Halkier T, Diers I, Savva D, Asenjo JA (1996) Molecular cloning of a lytic beta-1,3-glucanase gene from Oerskovia xanthineolytica LLG109—a beta-1,3-glucanase able to selectively permeabilize the yeast cell wall. In: Asenjo JA, Andrews BA (eds) Recombinant DNA biotechnology III: The integration of biological and engineering sciences, vol 782. Annals of the New York Academy of Sciences. New York Academy of Sciences, New York, pp 555–565. doi:10.1111/j.1749-6632.1996.tb40593.x

64. Dower WJ, Miller JF, Ragsdale CW (1988) High-efficiency transformation of Escherichia-coli by high-voltage electroporation. Nucleic Acids Res 16(13):6127–6145. doi:10.1093/nar/16.13.6127

65. Fromm M, Callis J, Taylor LP, Walbot V (1987) Electroporation of DNA and RNA into plant-protoplasts. Methods Enzymol 153:351–366

66. Becker DM, Guarente L (1991) High-efficiency transformation of yeast by electroporation. Methods Enzymol 194:182–187

67. Tryfona T, Bustard MT (2006) Enhancement of biomolecule transport by electroporation: a review of theory and practical application to transformation of Corynebacterium glutamicum. Biotechnol Bioeng 93(3):413–423. doi:10.1002/bit.20725

68. Escoffre JM, Dean DS, Hubert M, Rols MP, Favard C (2007) Membrane perturbation by an external electric field: a mechanism to permit molecular uptake. Eur Biophys J Biophys Lett 36(8):973–983. doi:10.1007/s00249-007-0194-7

69. Pliquett U, Joshi RP, Sridhara V, Schoenbach KH (2007) High electrical field effects on cell membranes. Bioelectrochemistry 70(2):275–282. doi:10.1016/j.bioelechem.2006.10.004

70. Crotti LB, Drgon T, Cabib E (2001) Yeast cell permeabilization by osmotic shock allows determination of enzymatic activities in situ. Anal Biochem 292(1):8–16. doi: 10.1006/abio.2001.5051

71. Bernal V, Sevilla A, Canovas M, Iborra JL (2007) Production of L-carnitine by secondary metabolism of bacteria. Microb Cell Fact 6:1–17. doi:3110.1186/1475-2859-6-31

72. Miozzari GF, Niederberger P, Hütter R (1978) Permeabilization of microorganisms by Triton X-100. Anal Biochem 90(1):220–233

73. Jackson RW, Demoss JA (1965) Effects of toluene on Escherichia coli. J Bacteriol 90(5):1420–1425

74. Flores MV, Voget CE, Ertola RJJ (1994) Permeabilization of yeast-cells (Kluyveromyces) with organic-solvents. Enzyme Microb Technol 16(4):340–346. doi:10.1016/0141-0229(94)90177-5

75. Thedei G, Leitao DPS, Bolean M, Paulino TP, Spadaro ACC, Ciancaglini P (2008) Toluene permeabilization differentially affects F- and P-type ATPase activities present in the plasma membrane of Streptococcus mutans. Braz J Med Biol Res 41(12):1047–1053

76. Martin LM, Ruiz CA, Andres M, Catalan J (2011) Permeabilization of Trigonopsis variabilis for enhanced D-amino acid oxidase activity. Chem Eng Commun 198(4):516–529. doi:10.1080/00986445.2010.512530

77. Tokuda H, Matsuyama S (2004) Sorting of lipoproteins to the outer membrane in E-coli. Biochim Biophys Acta (Mol Cell Res) 1693(1):5–13. doi:10.1016/j.bbamer.2004.02.005

78. De Groeve M, Depreitere V, Desmet T, Soetaert W (2009) Enzymatic production of α-D-galactose 1-phosphate by lactose phosphorolysis. Biotechnol Lett 31(12):1873–1877

79. Foulstone M, Reading C (1982) Assay of amoxicillin and clavulanic acid, the components of Augmentin, in biological fluids with high-performance liquid chromatography. Antimicrob Agents Chemother 22(5):753–762. doi:10.1128/aac

80. Cheng S, Wei D, Song Q, Zhao X (2006) Immobilization of permeabilized whole cell penicillin G acylase from Alcaligenes faecalis using pore matrix crosslinked with glutaraldehyde. Biotechnol Lett 28(14):1129–1133

81. Elling L (1996) Kinetic characterization of UDP-glucose pyrophosphorylase from germinated barley (malt). Phytochemistry 42(4):955–960

82. Shen B, Du LC, Sanchez C, Edwards DJ, Chen M, Murrell JM (2002) Cloning and characterization of the bleomycin biosynthetic gene cluster from Streptomyces verticillus ATCC15003. J Nat Prod 65(3):422–431. doi:10.1021/np010550q

83. Bogman K, Erne-Brand F, Alsenz J, Drewe J (2003) The role of surfactants in the reversal of active transport mediated by multidrug resistance proteins. J Pharm Sci 92(6):1250–1261. doi:10.1002/jps.10395

84. Ryan W, Parulekar SJ (1991) Immobilization of Escherichia-Coli JM103[pUC8] in kappa-carrageenan coupled with recombinant protein release by insitu cell-membrane permeabilization. Biotechnol Prog 7(2):99–110. doi:10.1021/bp00008a004

85. Rutter GA, Denton RM (1992) Effects of insulin and guanosine 5′-gamma-thio triphosphate on fatty-acid synthesis and lipolysis within electropermeabilized fat-cells. Biochem J 281:431–435

86. Bonnafous J-C, Dornand J, Mani J-C (1982) Alamethicin or detergent permeabilization of the cell membrane as a tool for adenylate cyclase determination: application to the study of hormone responsiveness in lymphocytes. Biochim Biophys Acta (Mol Cell Res) 720(3):235–241. doi:10.1016/0167-4889(82)90046-5

87. Johansson FI, Michalecka AM, Moller IM, Rasmusson AG (2004) Oxidation and reduction of pyridine nucleotides in alamethicin-permeabilized plant mitochondria. Biochem J 380:193–202. doi:10.1042/bj20031969

88. Curran KA, Crook NC, Alper HS (2012) Using flux balance analysis to guide microbial metabolic engineering. Methods Mol Biol 834:197–216. doi:10.1007/978-1-61779-483-4_13

89. Neuner A, Heinzle E (2011) Mixed glucose and lactate uptake by Corynebacterium glutamicum through metabolic engineering. Biotechnol J 6(3):318–329. doi:10.1002/biot.201000307

90. Becker J, Zelder O, Hafner S, Schroder H, Wittmann C (2011) From zero to hero-Design-based systems metabolic engineering of Corynebacterium glutamicum for L-lysine production. Metab Eng 13(2):159–168. doi:10.1016/j.ymben.2011.01.003

91. Kroemer JO, Wittmann C, Schroeder H, Heinzle E (2006) Metabolic pathway analysis for rational design of L-methionine production by Escherichia coli and Corynebacterium glutamicum. Metab Eng 8(4):353–369. doi:10.1016/j.ymben.2006.02.001

92. Bujara M, Panke S (2012) In silico assessment of cell-free systems. Biotechnol Bioeng 109(10):2620–2629. doi:10.1002/Bit.24534

93. Soh KC, Miskovic L, Hatzimanikatis V (2012) From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome-scale metabolic networks. FEMS Yeast Res 12(2):129–143. doi:10.1111/j.1567-1364.2011.00771.x

94. Yuryev R, Liese A (2010) Biocatalysis: the outcast. ChemCatChem 2(1):103–107. doi:10.1002/cctc.200900126

95. Weinig S, Mahmud T, Muller R (2003) Markerless mutations in the myxothiazol biosynthetic gene cluster: a delicate megasynthetase with a superfluous nonribosomal peptide synthetase domain. Chem Biol 10(10):953–960. doi:10.1016/j.chembiol.2003.09.013

96. Wessjohann L, Vogt T, Julia K, Robert K (2012) Prenyl- und Methyltransferasen in Natur- und Synthese. Biospektrum 18:22–25. doi:10.1007/s12268-012-0137-4

97. Branco RJF, Graber M, Denis V, Pleiss J (2009) Molecular mechanism of the hydration of Candida antarctica lipase B in the gas phase: water adsorption isotherms and molecular dynamics simulations. ChemBioChem 10(18):2913–2919. doi:10.1002/cbic.200900544

98. Trodler P, Pleiss J (2008) Modeling structure and flexibility of Candida antarctica lipase B in organic solvents. BMC Struct Biol 8:9. doi:10.1186/1472-6807-8-9

99. Struck A-W, Thompson ML, Wong LS, Micklefield J (2012) S-adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. Chembiochem 13(18):2642–2655. doi:10.1002/cbic.201200556

100. Thirlway J, Lewis R, Nunns L, Al Nakeeb M, Styles M, Struck A-W, Smith CP, Micklefield J (2012) Introduction of a non-natural amino acid into a nonribosomal peptide antibiotic by modification of adenylation domain specificity. Angew Chem-Int Ed 51(29):7181–7184. doi:10.1002/anie.201202043

101. Zhang YH (2011) Substrate channeling and enzyme complexes for biotechnological applications. Biotechnol Adv 29(6):715–725. doi:10.1016/j.biotechadv.2011.05.020

102. Wilson DJ, Shi C, Teitelbaum AM, Gulick AM, Aldrich CC (2013) Characterization of AusA: a dimodular nonribosomal peptide synthetase responsible for the production of aureusimine pyrazinones. Biochemistry 52(5):926–937. doi:10.1021/bi301330q

103. Link H, Kochanowski K, Sauer U (2013) Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. Nat Biotechnol 31(4):357–361. doi:10.1038/nbt.2489

104. Heinzle E, Matsuda F, Miyagawa H, Wakasa K, Nishioka T (2007) Estimation of metabolic fluxes, expression levels and metabolite dynamics of a secondary metabolic pathway in potato using label pulse-feeding experiments combined with kinetic network modelling and simulation. Plant J 50(1):176–187. doi:10.1111/j.1365-313X.2007.03037.x

105. Wu L, Wang WM, van Winden WA, van Gulik WM, Heijnen JJ (2004) A new framework for the estimation of control parameters in metabolic pathways using lin-log kinetics. Eur J Biochem 271(16):3348–3359. doi:10.1111/j.1432-1033.2004.04269.x

106. Krauser S, Kiefer P, Heinzle E (2012) Multienzyme whole-cell in situ biocatalysis for the production of flaviolin in permeabilized cells of Escherichia coli. ChemCatChem 4(6):786–788. doi:10.1002/cctc.201100351

107. Broadwater SJ, Roth SL, Price KE, Kobaslija M, McQuade DT (2005) One-pot multi-step synthesis: a challenge spawning innovation. Org Biomol Chem 3(16):2899–2906
108. Treitz G, Maria G, Giffhorn F, Heinzle E (2001) Kinetic model discrimination via step-by-step experimental and computational procedure in the enzymatic oxidation of D-glucose. J Biotechnol 85(3):271–287. doi:10.1016/s0168-1656(00)00371-0
109. Wohlgemuth R (2007) Interfacing biocatalysis and organic synthesis. J Chem Technol Biotechnol 82(12):1055–1062. doi:10.1002/jctb.1761
110. Clardy J, Walsh C (2004) Lessons from natural molecules. Nature 432(7019):829–837. doi:10.1038/nature03194
111. Bode HB, Muller R (2005) The impact of bacterial genomics on natural product research. Angew Chem-Int Ed 44(42):6828–6846. doi:10.1002/anie.200501080
112. Wu M-C, Law B, Wilkinson B, Micklefield J (2012) Bioengineering natural product biosynthetic pathways for therapeutic applications. Curr Opin Biotechnol 23(6):931–940. doi:10.1016/j.copbio.2012.03.008
113. Walsh CT, O'Brien RV, Khosla C (2013) Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. Angew Chem. doi:10.1002/anie.201208344
114. Wong FT, Khosla C (2012) Combinatorial biosynthesis of polyketides—a perspective. Curr Opin Chem Biol 16(1–2):117–123. doi:10.1016/j.cbpa.2012.01.018
115. Pickens LB, Tang Y (2009) Decoding and engineering tetracycline biosynthesis. Metab Eng 11(2):69–75. doi:10.1016/j.ymben.2008.10.001
116. Pulsawat N, Kitani S, Nihira T (2007) Characterization of biosynthetic gene cluster for the production of virginiamycin M, a streptogramin type A antibiotic, in Streptomyces virginiae. Gene 393(1–2):31–42. doi:10.1016/j.gene.2006.12.035
117. Schultz AW, Oh DC, Carney JR, Williamson RT, Udwary DW, Jensen PR, Gould SJ, Fenical W, Moore BS (2008) Biosynthesis and structures of cyclomarins and cyclomarazines, prenylated cyclic peptides of marine actinobacterial origin. J Am Chem Soc 130(13):4507–4516. doi:10.1021/ja711188x
118. Tang GL, Cheng YQ, Shen B (2004) Leinamycin biosynthesis revealing unprecedented architectural complexity for a hybrid polyketide synthase and nonribosomal peptide synthetase. Chem Biol 11(1):33–45. doi:10.1016/j.chembiol.2003.12.014
119. Kim ES, Bibb MJ, Butler MJ, Hopwood DA, Sherman DH (1994) Sequences of the oxytetracycline polyketide synthase-encoding otc genes from Streptomyces-rimosus. Gene 141(1):141–142. doi:10.1016/0378-1119(94)90144-9
120. Schlichting I, Cryle MJ (2009) Macromolecular complexes involving cytochrome P450 enzymes. FEBS J 276:21–22
121. Cryle MJ, Schlichting I (2008) Structural insights from a P450 Carrier Protein complex reveal how specificity is achieved in the P450(Biol) ACP complex. Proc Natl Acad Sci U S A 105(41):15696–15701. doi:10.1073/pnas.0805983105
122. Pickens LB, Tang Y (2010) Oxytetracycline biosynthesis. J Biol Chem 285(36):27509–27515. doi:10.1074/jbc.R110.130419
123. Stevens DC, Henry MR, Murphy KA, Boddy CN (2010) Heterologous expression of the oxytetracycline biosynthetic pathway in Myxococcus xanthus. Appl Environ Microbiol 76(8):2681–2683. doi:10.1128/aem.02841-09
124. Weissman KJ, Muller R (2008) Protein-protein interactions in multienzyme megasynthetases. ChemBioChem 9(6):826–848. doi:10.1002/cbic.200700751
125. Findrik Z, Vasic-Racki D (2009) Overview on reactions with multi-enzyme systems. Chem Biochem Eng Q 23(4):545–553
126. Ochi K, Hosaka T (2013) New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. Appl Microbiol Biotechnol 97(1):87–98. doi:10.1007/s00253-012-4551-9
127. Medema MH, Breitling R, Takano E (2011) Synthetic biology in Streptomyces bacteria. Methods Enzymol 497:485–502. doi:10.1016/B978-0-12-385075-1.00021-4

128. Wenzel SC, Muller R (2009) The biosynthetic potential of myxobacteria and their impact in drug discovery. Curr Opin Drug Discov Devel 12(2):220–230

129. Boddy CN, Garza A (2010) System and method for the heterologous expression of polyketide synthase gene clusters. US Patent (20100184038 A1)

130. Akai S, Hanada R, Fujiwara N, Kita Y, Egi M (2010) One-pot synthesis of optically active allyl esters via lipase-vanadium combo catalysis. Org Lett 12(21):4900–4903. doi:10.1021/ol102053a

131. Caiazzo A, Garcia PML, Wever R, van Hest JCM, Rowan AE, Reek JNH (2009) Synergy between chemo- and bio-catalysts in multi-step transformations. Org Biomol Chem 7(14):2926–2932. doi:10.1039/b901592b

132. Fowler ZL, Koffas MAG (2009) Biosynthesis and biotechnological production of flavanones: current state and perspectives. Appl Microbiol Biotechnol 83(5):799–808. doi:10.1007/s00253-009-2039-z

133. Santacoloma PA, Sin G, Gernaey KV, Woodley JM (2011) Multienzyme-catalyzed processes: next-generation biocatalysis. Org Process Res Dev 15(1):203–212. doi:10.1021/op1002159

134. Tanner P, Onaca O, Balasubramanian V, Meier W, Palivan CG (2011) Enzymatic cascade reactions inside polymeric nanocontainers: a means to combat oxidative stress. Chem-Eur J 17(16):4552–4560. doi:10.1002/chem.201002782

135. Richter F, Leaver-Fay A, Khare SD, Bjelic S, Baker D (2011) De novo enzyme design using Rosetta3. PLoS ONE 6(5):e19230. doi:10.1371/journal.pone.0019230

136. Ye X, Honda K, Morimoto Y, Okano K, Ohtake H (2013) Direct conversion of glucose to malate by synthetic metabolic engineering. J Biotechnol 164(1):34–40. doi:10.1016/j.jbiotec.2012.11.011

137. Nestl BM, Nebel BA, Hauer B (2011) Recent progress in industrial biocatalysis. Curr Opin Chem Biol 15(2):187–193. doi:10.1016/j.cbpa.2010.11.019

138. Bar-Even A, Salah Tawfik D (2013) Engineering specialized metabolic pathways—is there a room for enzyme improvements? Curr Opin Biotechnol 24(2):310–319. doi:10.1016/j.copbio.2012.10.006

139. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO (2007) A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol 3:18. doi:10.1038/msb4100155

140. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH (2009) PubChem: a public information system for analyzing bioactivities of small molecules. Nucleic Acids Res 37:W623–W633. doi:10.1093/nar/gkp456

141. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Zhou Z, Han L, Karapetyan K, Dracheva S, Shoemaker BA, Bolton E, Gindulyte A, Bryant SH (2012) PubChem's bioassay database. Nucleic Acids Res 40(D1):D400–D412. doi:10.1093/nar/gkr1132

142. de Matos P, Adams N, Hastings J, Moreno P, Steinbeck C (2012) A database for chemical proteomics: ChEBI. Methods Mol Biol (Clifton, NJ) 803:273–296

143. Saier MH Jr, Yen MR, Noto K, Tamang DG, Elkan C (2009) The Transporter Classification Database: recent advances. Nucleic Acids Res 37:D274–D278. doi:10.1093/nar/gkn862

144. Ren Q, Chen K, Paulsen IT (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res 35:D274–D279. doi:10.1093/nar/gkl925

145. Wittig U, Kania R, Golebiewski M, Rey M, Shi L, Jong L, Algaa E, Weidemann A, Sauer-Danzwith H, Mir S, Krebs O, Bittkowski M, Wetsch E, Rojas I, Mueller W (2012) SABIO-RK-database for biochemical reaction kinetics. Nucleic Acids Res 40(D1):D790–D796. doi:10.1093/nar/gkr1046

146. Alcantara R, Axelsen KB, Morgat A, Belda E, Coudert E, Bridge A, Cao H, de Matos P, Ennis M, Turner S, Owen G, Bougueleret L, Xenarios I, Steinbeck C (2012) Rhea-a manually curated resource of biochemical reactions. Nucleic Acids Res 40(D1):D754–D760. doi:10.1093/nar/gkr1126

147. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. Nucleic Acids Res 41(D1):D36–D42. doi:10.1093/nar/gks1195

148. Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Federhen S, Feolo M, Fingerman IM, Geer LY, Helmberg W, Kapustin Y, Krasnov S, Landsman D, Lipman DJ, Lu ZY, Madden TL, Madej T, Maglott DR, Marchler-Bauer A, Miller V, Karsch-Mizrachi I, Ostell J, Panchenko A, Phan L, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Slotta D, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Wang YL, Wilbur WJ, Yaschenko E, Ye J (2012) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 40(D1):D13–D25. doi:10.1093/nar/gkr1184

149. Maglott D, Ostell J, Pruitt KD, Tatusova T (2011) Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res 39:D52–D57. doi:10.1093/nar/gkq1237

150. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene Ontology: tool for the unification of biology. Nat Genet 25(1):25–29

151. Cochrane G, Alako B, Amid C, Bower L, Cerdeno-Tarraga A, Cleland I, Gibson R, Goodgame N, Jang M, Kay S, Leinonen R, Lin X, Lopez R, McWilliam H, Oisel A, Pakseresht N, Pallreddy S, Park Y, Plaister S, Radhakrishnan R, Riviere S, Rossello M, Senf A, Silvester N, Smirnov D, ten Hoopen P, Toribio A, Vaughan D, Zalunin V (2013) Facing growth in the European Nucleotide Archive. Nucleic Acids Res 41(D1):D30–D35. doi:10.1093/nar/gks1175

152. Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, Soehngen C, Stelzer M, Thiele J, Schomburg D (2011) BRENDA, the enzyme information system in 2011. Nucleic Acids Res 39:D670–D676. doi:10.1093/nar/gkq1089

153. Bairoch A (2000) The ENZYME database in 2000. Nucleic Acids Res 28(1):304–305. doi:10.1093/nar/28.1.304

154. Magrane M, Consortium U (2011) UniProt Knowledgebase: a hub of integrated protein data. Database (J Biol Databases Curation) 14:bar009-bar009

155. Yu NY, Laird MR, Spencer C, Brinkman FSL (2011) PSORTdb-an expanded, auto-updated, user-friendly protein subcellular localization database for Bacteria and Archaea. Nucleic Acids Res 39:D241. doi:10.1093/nar/gkq1093

156. Bowers PM, Pellegrini M, Thompson MJ, Fierro J, Yeates TO, Eisenberg D (2004) Prolinks: a database of protein functional linkages derived from coevolution. Genome Biol 5(5):R35. doi:10.1186/gb-2004-5-5-r35

157. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41(D1):D808–D815. doi:10.1093/nar/gks1094

158. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, Jandrasits C, Jimenez RC, Khadake J, Mahadevan U, Masson P, Pedruzzi I, Pfeiffenberger E, Porras P, Raghunath A, Roechert B, Orchard S, Hermjakob H (2012) The IntAct molecular interaction database in 2012. Nucleic Acids Res 40(D1):D841–D846. doi:10.1093/nar/gkr1088

159. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res 40(D1). doi:10.1093/nar/gkr988

160. Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 28(1):27–30. doi:10.1093/nar/28.1.27

161. Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang PF, Karp PD (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 40(D1):D742–D753. doi:10.1093/nar/gkr1014

162. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, Stockinger H (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40(W1):W597–W603. doi: 10.1093/Nar/Gks400

163. Morgat A, Coissac E, Coudert E, Axelsen KB, Keller G, Bairoch A, Bridge A, Bougueleret L, Xenarios I, Viari A (2012) UniPathway: a resource for the exploration and annotation of metabolic pathways. Nucleic Acids Res 40(D1). doi:10.1093/nar/gkr1023

164. Gao J, Ellis LBM, Wackett LP (2010) The University of Minnesota Biocatalysis/ Biodegradation Database: improving public access. Nucleic Acids Res 38:D488–D491. doi: 10.1093/nar/gkp771

165. Ganter M, Bernard T, Moretti S, Stelling J, Pagni M (2013) MetaNetX.org: a website and repository for accessing, analysing and manipulating metabolic networks. Bioinformatics 29(6):815-816. doi:10.1093/bioinformatics/btt036

166. Schellenberger J, Park JO, Conrad TM, Palsson BO (2010) BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinformatics 11:213. doi:10.1186/1471-2105-11-213

167. Li C, Donizelli M, Rodriguez N, Dharuri H, Endler L, Chelliah V, Li L, He E, Henry A, Stefan MI, Snoep JL, Hucka M, Le Novere N, Laibe C (2010) BioModels Database: an enhanced, curated and annotated resource for published quantitative kinetic models. BMC Syst Biol 4:92. doi:10.1186/1752-0509-4-92

168. Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muniz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M, Altman T, Kaipa P, Spaulding A, Pacheco J, Latendresse M, Fulcher C, Sarker M, Shearer AG, Mackie A, Paulsen I, Gunsalus RP, Karp PD (2011) EcoCyc: a comprehensive database of Escherichia coli biology. Nucleic Acids Res 39:D583–D590. doi:10.1093/nar/gkq1143

169. Michal G, Schomburg D (2012) Biochemical pathways. An atlas of biochemistry and molecular biology. Wiley, New York

170. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crecy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Ruckert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33(17):5691–5702. doi: 10.1093/nar/gki866

171. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28(9):977–982. doi:10.1038/nbt.1672

172. Karp PD, Paley SM, Krummenacker M, Latendresse M, Dale JM, Lee TJ, Kaipa P, Gilham F, Spaulding A, Popescu L, Altman T, Paulsen I, Keseler IM, Caspi R (2010) Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. Brief Bioinformatics 11(1):40–79. doi:10.1093/bib/bbp043

173. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li C-Y, Wei L (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic Acids Res 39:W316. doi:10.1093/nar/gkr483

174. Bates JT, Chivian D, Arkin AP (2011) GLAMM: Genome-Linked Application for Metabolic Maps. Nucleic Acids Res 39:W400–W405. doi:10.1093/nar/gkr433

175. Liao Y-C, Tsai M-H, Chen F-C, Hsiung CA (2012) GEMSiRV: a software platform for GEnome-scale metabolic model simulation, reconstruction and visualization. Bioinformatics 28(13):1752–1758. doi:10.1093/bioinformatics/bts267

176. Klamt S, Saez-Rodriguez J, Gilles ED (2007) Structural and functional analysis of cellular networks with Cell NetAnalyzer. BMC Syst Biol 1:2. doi:10.1186/1752-0509-1-2

177. von Kamp A, Schuster S (2006) Metatool 5.0: fast and flexible elementary modes analysis. Bioinformatics 22(15):1930-1931. doi:10.1093/bioinformatics/btl267
178. Terzer M, Stelling J (2008) Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 24(19):2229–2235. doi:10.1093/bioinformatics/btn401
179. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat Protoc 2(3):727–738. doi:10.1038/nprot.2007.99
180. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BO (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat Protoc 6(9):1290–1307. doi:10.1038/nprot.2011.308
181. Reed JL, Famili I, Thiele I, Palsson BO (2006) Towards multidimensional genome annotation. Nat Rev Genet 7(2):130–141. doi:10.1038/nrg1769
182. Zomorrodi AR, Suthers PF, Ranganathan S, Maranas CD (2012) Mathematical optimization applications in metabolic networks. Metab Eng 14(6):672–686. doi:10.1016/j.ymben.2012.09.005
183. Thiele I, Palsson BO (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat Protoc 5(1):93–121. doi:10.1038/nprot.2009.203
184. Haggart CR, Bartell JA, Saucerman JJ, Papin JA (2011) Whole-genome metabolic network reconstruction and constraint-based modeling. In: Jameson D, Verma M, Westerhoff HV (eds) Methods in systems biology. Methods in enzymology, vol 500. Elsevier Academic Press Inc, San Diego, pp 411–433. doi:10.1016/b978-0-12-385118-5.00021-9
185. Oberhardt MA, Palsson BO, Papin JA (2009) Applications of genome-scale metabolic reconstructions. Mol Syst Biol 5:320. doi:10.1038/msb.2009.77
186. Deville Y, Gilbert D, van Helden J, Wodak S (2003) An overview of data models for the analysis of biochemical pathways. In: Gordon P (ed) Proceedings of computational methods in systems Biology, vol 2602. Springer, Berlin
187. Le Novere N, Hucka M, Mi HY, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem MI, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villeger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman TC, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu GM, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H (2009) The systems biology graphical notation (vol 27, pg 735, 2009). Nat Biotechnol 27(9):864. doi:10.1038/nbt0909-864d
188. Lee W-H, Kim M-D, Jin Y-S, Seo J-H (2013) Engineering of NADPH regenerators in Escherichia coli for enhanced biotransformation. Appl Microbiol Biotechnol 97(7):2761–2772. doi:10.1007/s00253-013-4750-z
189. Horinouchi N, Ogawa J, Kawano T, Sakai T, Saito K, Matsumoto S, Sasaki M, Mikami Y, Shimizu S (2006) Efficient production of 2-deoxyribose 5-phosphate from glucose and acetaldehyde by coupling of the alcoholic fermentation system of baker's yeast and deoxyriboaldolase-expressing Escherichia coli. Biosci Biotechnol Biochem 70(6):1371–1378. doi:10.1271/bbb.50648
190. Horinouchi N, Sakai T, Kawano T, Matsumoto S, Sasaki M, Hibi M, Shima J, Shimizu S, Ogawa J (2012) Construction of microbial platform for an energy-requiring bioprocess: practical 2'-deoxyribonucleoside production involving a C-C coupling reaction with high energy substrates. Microb Cell Fact 11:82. doi:10.1186/1475-2859-11-82
191. Blazeck J, Alper H (2010) Systems metabolic engineering: genome-scale models and beyond. Biotechnol J 5(7):647–659. doi:10.1002/biot.200900247
192. Terzer M, Maynard ND, Covert MW, Stelling J (2009) Genome-scale metabolic networks. Wiley Interdiscip Rev Syst Biol Med 1(3):285–297 doi:10.1002/wsbm.37

193. Kim TY, Sohn SB, Bin Kim Y, Kim WJ, Lee SY (2012) Recent advances in reconstruction and applications of genome-scale metabolic models. Curr Opin Biotechnol 23(4):617–623. doi:10.1016/j.copbio.2011.10.007

194. Durot M, Bourguignon P-Y, Schachter V (2009) Genome-scale models of bacterial metabolism: reconstruction and applications. FEMS Microbiol Rev 33(1):164–190. doi: 10.1111/j.1574-6976.2008.00146.x

195. Neuner A, Wagner I, Sieker T, Ulber R, Schneider K, Peifer S, Heinzle E (2013) Production of L-lysine on different silage juices using genetically engineered Corynebacterium glutamicum. J Biotechnol 163(2):217–224. doi:10.1016/j.jbiotec.2012.07.190

196. Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebizuka Y, Horinouchi S (1999) A new pathway for polyketide synthesis in microorganisms. Nature 400(6747):897–899. doi:10.1038/23748

197. Goedl C, Schwarz A, Minani A, Nidetzky B (2007) Recombinant sucrose phosphorylase from Leuconostoc mesenteroides: characterization, kinetic studies of transglycosylation, and application of immobilised enzyme for production of [alpha]-D-glucose 1-phosphate. J Biotechnol 129(1):77–86

198. Weyler C, Heinzle E (2013) submitted

199. Koizumi S (2003) Large-scale production of oligosaccharides using bacterial functions. Trends Glycosci Glyc 15(82):65–74

200. Koizumi S, Endo T, Tabata K, Nagano H, Ohnishi J, Ozaki A (2000) Large-scale production of GDP-fucose and Lewis X by bacterial coupling. J Ind Microbiol Biotechnol 25(4):213–217. doi:10.1038/sj.jim.7000055

201. Koizumi S, Endo T, Tabata K, Ozaki A (1998) Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria. Nat Biotech 16(9):847–850

202. Ni Y, Zhang B, Sun Z (2012) Efficient synthesis of (R)-2-Chloro-1-(3-chlorophenyl)ethanol by permeabilized whole-cells of Candida ontarioensis. Chin J Catal 33(4):681–687. doi: 10.1016/s1872-2067(11)60363-x

203. Fu J, Bian XY, Hu SB, Wang HL, Huang F, Seibert PM, Plaza A, Xia LQ, Muller R, Stewart AF, Zhang YM (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat Biotechnol 30(5):440–446. doi:10.1038/Nbt.2183

Index