Microbiology Monographs

Series Editor: Alexander Steinbüchel

Bernd H. A. Rehm
David Wibowo *Editors*

Microbial Production of High-Value Products

Springer

Microbiology Monographs

Volume 37

Series Editor

Alexander Steinbüchel, Münster, Germany

The Springer book series presents carefully refereed volumes on selected microbiological topics. Microbiology is a still rapidly expanding field with significant impact on many areas of basic and applied science. The growth in knowledge of microbial physiology, cell structure, biotechnological capabilities and other aspects of microorganisms is increasing dramatically even in the face of the breakthroughs that already have been made. ***Microbiology Monographs***

Reflecting these most recent achievements, the series' wide scope encompasses such topics as inclusions in prokaryotes, predatory prokaryotes, magnetoreception and magnetosomes in bacteria, uncultivable microorganisms, microbial endosymbionts, bacterial resistance, extremophilic microorganisms, analyses of genome sequences and structures, microorganisms as cell factories for chemicals and fuels, metabolic engineering, gene transfer and expression systems and distinct physiological groups of bacteria.

The volume editors are well-known experts in their particular fields, and each volume offers 10 to 20 comprehensive review articles covering all relevant aspects of the topic in focus. All chapters are systematically reviewed by the series editor and respective volume editor(s).

Bernd H. A. Rehm • David Wibowo
Editors

Microbial Production of High-Value Products

Springer

Editors
Bernd H. A. Rehm
Centre for Cell Factories and Biopolymers,
Griffith Institute for Drug Discovery
Menzies Health Institute Queensland
Griffith University
Brisbane, QLD, Australia

David Wibowo
Centre for Cell Factories and Biopolymers,
Griffith Institute for Drug Discovery
Griffith University
Brisbane, QLD, Australia

ISSN 1862-5576 ISSN 1862-5584 (electronic)
Microbiology Monographs
ISBN 978-3-031-06599-6 ISBN 978-3-031-06600-9 (eBook)
https://doi.org/10.1007/978-3-031-06600-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Advancements in Inducer Systems for Recombinant Protein Production in *E. coli*

Jaya A. Gupta, Kathiresan Pandi, and Anurag S. Rathore

Contents

Abstract *Escherichia coli* (*E. coli*) is the predominant industrial microorganism and cost-effective microbial cell "factory" for producing commercial products, such as metabolites, enzymes, biochemicals, and high-value biotherapeutics. It is the primary workhorse for non-glycosylated recombinant proteins production as it is the production organism used to produce around 30% of biopharmaceutical products. Heterologous protein expression in *E. coli* is often a preferred choice for many reasons: it is economical, provides a fast growth rate, and can reach high cell density and high product titers. Being widely studied, an abundance of biochemical and physiological knowledge is also available. Also, various genetic tools are available to manipulate *E. coli*, including a significantly extensive catalog of expression plasmids and engineered strains and several cultivation strategies. Selecting appropriate expression plasmids and promoter strength is crucial for efficient protein production. Promoter strength depends on the promoter components that can be either constitutive or inducible. Constitutive promoters allow continuous gene expression, which can cause a metabolic burden affecting cell growth, whereas inducible promoters control gene expression through specific inducers. The existing expression systems are diverse and complex, and its selection depends on the nature

J. A. Gupta · K. Pandi · A. S. Rathore (✉)
Department of Chemical Engineering, Indian Institute of Technology, Hauz Khas, India
e-mail: asrathore@biotechcmz.com

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*,
Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_1

of target protein that requires in particular cases to be produced at high concentrations, or at different cell growth phases, or at particular location whether in the cytoplasmic or periplasmic space or in some cases, secretion into medium.

1 Tuning Recombinant Protein Production in *E. coli*

E. coli served as the workhorse of modern biotechnology. It is the predominant industrial microorganism and most frequently used prokaryotes to produce commercial products, such as metabolites, enzymes, biochemicals, and high-value biotherapeutics (Gupta and Shukla 2017; Walsh 2018). A substantial amount of knowledge has been generated over the past 30 years about developing *E. coli* strains for recombinant proteins expression in cytoplasm, periplasm, or secretion into medium (Rosano et al. 2019). Moreover, many expression vectors containing regulated promoters, signal sequences, antibiotic selection and tags for efficient protein purification have been developed (Rosano et al. 2019). Despite all these advancements, process optimization is required to achieve a high titer of recombinant proteins. Due to variations in vectors, gene expression and its products, promoter strength, plasmid copy number, and host–vector interactions, process optimization is often a tedious and time-consuming task (Sahdev et al. 2008). Moreover, there are recombinant proteins that are toxic to the cell. In *E. coli*, it has been confirmed that the expression of heterologous proteins in uncontrolled fashion results in the initiation of metabolic stress responses, undesired product loss, and eventually cell death (Bentley et al. 1990; Bonomo and Gill 2005). Therefore, the complex interplay of media processing, the nature of the product, and the expression system determines productivity, expression level, and product quality. A high level of foreign gene expression that results in 25–35% of total cellular protein may sometimes lead to significant genetic instability (Bentley et al. 1990; Hoffmann and Rinas 2004). When biomass is high, promoters in the plasmid should be tightly regulated (Saïda et al. 2006). Multiple different transcriptional regulations of the promoter are found in *E. coli*. Perhaps, the most common control mechanism is the binding of regulatory proteins to the specific region of the promoter. However, there exist other mechanisms of regulation, such as transcriptional attenuation, anti-termination, anti-sense RNA, variation in sigma factors, and anti-sigma-factors (Tropp 2008). In practice, based on carbon source availability, promoter activity can be controlled by regulatory proteins (Tropp 2008). Transcriptional proteins control the activity of the promoter in two ways. Negatively controlled systems prevent RNA polymerase binding to the promoter using regulatory elements, thereby repressing transcription. On the other hand, regulatory elements in positively controlled systems allow the binding of RNA polymerase to the promoter through activator proteins, thus permitting transcription (Tropp 2008). Either inducer or repressor can control the negatively and positively regulated promoters (Rosano and Ceccarelli 2014). In addition, a negatively regulated inducer system is activated by inhibiting the binding of the repressor to the operator. In contrast, a positively

regulated inducer system is activated in the presence of an effector by inducer binding to the activator (Tropp 2008). Both inducer systems have been successfully used to produce various heterologous proteins from *E. coli* (Sørensen and Mortensen 2005; Rosano and Ceccarelli 2014).

This chapter reviews the most popular and industrially useful promoter systems, such as lactose, arabinose, and rhamnose.

2 Promoter Systems of *E. coli*

A commercially useful promoter must be strong, tightly regulated, have a low basal expression level, must be strain independent, require simple and cost-effective induction method for gene expression, and should be independent on the commonly used ingredients of culturing media. Among the various promoter systems available for protein expression in *E. coli*, only a few of them are used. One example is constitutive promoters that have been opted for heterologous gene expression in *E. coli*. As gene expression in *E. coli* is based on multicopy plasmids, maintaining high-level gene expression constitutively along with cellular growth would cause a metabolic burden to the cells leading to the premature arrest of the system. Therefore, the constitutive promoters would not be ideal for heterologous protein production.

Another example is auto-inducible promoters, which do not require any external inducer for gene expression. Auto-inducible promoters aim toward large-scale protein production as they get induced either at the late log phase or at the stationary phase (i.e., a growth stage when cells are not dividing but remain metabolically active, and the genes required for the cell survival are expressed during this phase). One category of auto-inducible promoters is stationary-phase promoters. It has been reported that nearly 20% of the *E. coli* genes are expressed at an increased level in this phase (Rava et al. 1999). Stationary phase promoters show enhanced activity in the stationary phase, but little or no activity in the exponential phase. Nonetheless, many of these promoters exhibit low strength and, therefore are not popular for constructing expression vectors (Jaishankar and Srivastava 2017).

Starvation promoters constitute another category that has been widely used in *E. coli* (Shin and Seo 1990). These promoters are induced by starvation of a particular molecule/metabolite promoting the induction of heterologous genes. Alkaline phosphate promoter (phoA) is an example of starvation promoters, where phosphate starvation can be used to control the expression of heterologous genes during the non-growth phase (Keasling 1999). It is the starvation of external inorganic phosphate (P_i) in growth medium which induces the phosphate starvation response in the cell. This causes a release of alkaline phosphatase in periplasmic space to generate P_i from organophosphates because the physiological role of this enzyme is to perform hydrolysis of organic phosphate compounds present in growth medium to fulfill the cell's requirement of inorganic phosphate (Wanner 1996). Also, there is no known internal sensor for phosphate, so internal phosphate stores

cannot be sensed by the cells. It is the external phosphate starvation that drives the heterologous protein production from these promoters. Starvation promoters are particularly anticipated in bioremediation-related applications, where removal of environmental contaminants can control induction from these promoters (Keasling 1999).

A controlled expression system that could be induced at a desired time and condition is advantageous for efficient recombinant production (Nieto et al. 2000).

3 Inducer Systems in *E. coli*

Repressors, activators, and inducers are three important regulatory proteins that control gene expression in operons. Repressor protein inhibits gene transcription in response to an external stimulus. Activator protein increases gene transcription in response to an external signal. Inducers activate or deactivate transcription depending on the sugar availability and requirement in the cells (Tropp 2008). In *E. coli,* many genes are mostly expressed in a constitutive manner, meaning that the gene expression is always switched "on." In other cases, when their transcribed proteins are required by the cells, their expression will be controlled and hence referred as inducible expression. The most studied and popular promoters in bacteria are those regulating operons for sugar metabolism, e.g., lacZYA operon, araBAD operon, and rhaBAD operon (Rosano and Ceccarelli 2014). These promoters are inducible promoter system which is beneficial for recombinant protein production.

3.1 Lactose Induction System

Inducible sugars can control both the negatively and positively regulated promoters. Lac operon is one of the well-studied systems in *E. coli*. Regulation of lac operon is directed by the lac repressor gene, lacI (Kercher et al. 1997). In the absence of lactose, lacI binds to the promoter and prevents the transcription of the lactose operon genes (Fig. 1). In the presence of inducer, allolactose produced by galactosidase (lacZ) using lactose, the repressor weakens its binding affinity to the operator (Wheatley et al. 2013) (Fig. 2). Galactosidase also hydrolyzes lactose to galactose and glucose, and metabolizes them into the central carbon metabolism. Lactose permease (lacY) of lac operon allows lactose transport into the cell through membranes. The galactoside acetyltransferase (lacA) enzyme metabolizes acetyl-CoA to galactoside by transferring the acetyl group (Juers et al. 2012). Catabolite activator protein (CAP, also known as cyclic adenosine monophosphate (cAMP) receptor protein or CRP) positively controls its transcription (Wheatley et al. 2013). Efficient transcription is possible only when the CAP-cAMP complex binds to the -35 region upstream of the promoter (Simpsonl 1980; Malan and McClure 1984). A low concentration of glucose facilitates the increase of cAMP levels in the cells and

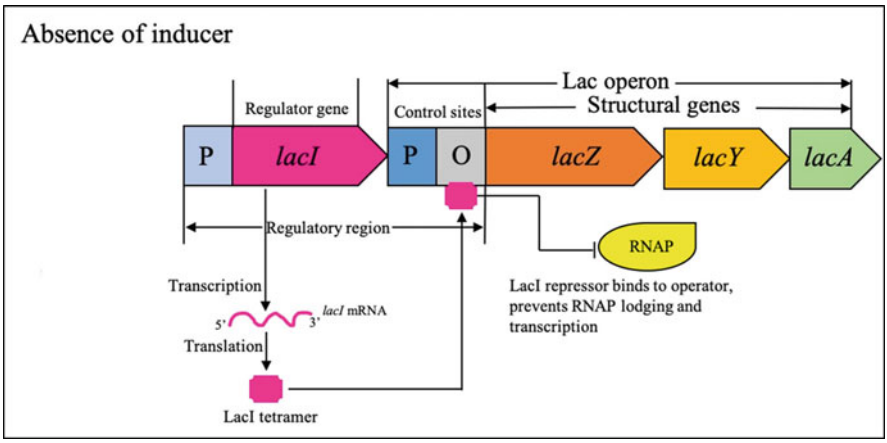


Fig. 1 Repressed lactose operon. This shows the repressed state of lactose operon, where absence of inducer allows binding of LacI repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

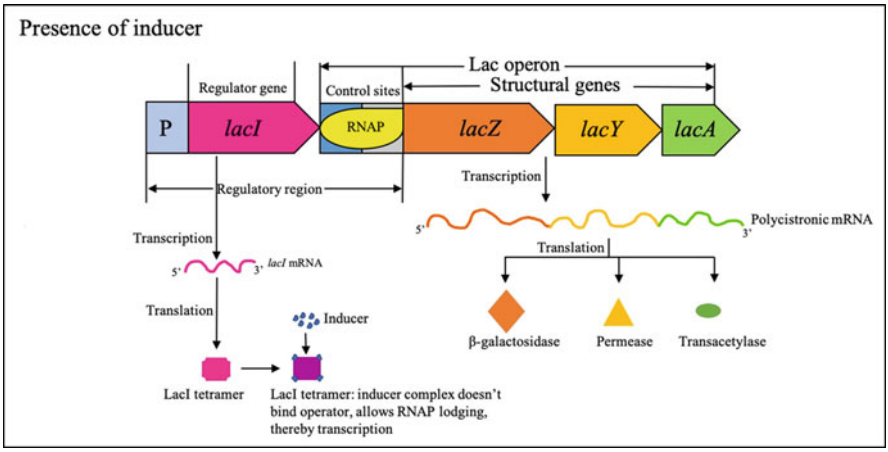


Fig. 2 Induced lactose operon. This shows the induced state of lactose operon, where presence of inducer forms a complex with LacI repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

allows transcription of lac operon genes (Inada et al. 1996). Modified lac promoters are routinely used in the *E. coli* vectors for heterologous gene expression. The strength of the lactose promoter was modified by fusing the -35 region of the tryptophan promoter from trp operon with the -10 region of the lactose promoter (De Boer et al. 1983; Neubauer et al. 1991). This modified promoter is referred to as tac promoter, which is tenfold more efficient than lac promoter, particularly on

multicopy plasmids. Therefore, tac promoter is used for recombinant proteins production at a commercial scale (De Boer et al. 1983).

The lacUV5 promoter is another modified promoter derived from the lac operon to guide the expression of heterologous genes on a plasmid. It is similar to the lac promoter, comprising just two base-pair mutations in the -10 hexamer region. It independently works regardless of activators or other cis-regulatory elements (Brosius et al. 1983; Shibui et al. 1988). LacI repressor alone can control lacUV5 promoter expression. The lacI gene found in the chromosome leads to very few lac repressor molecules (Brosius et al. 1983). Repressors will not occupy the lac operators due to low concentration, resulting in basal level expression from lac-based promoters. To control basal level expression, three operators were inserted into a plasmid with the correct spacing leads to complete inhibition of the promoter (Oehler et al. 1990, 1994). Several studies have been performed to enhance the strength of lac repression in lac promoter system. The first study was the isolation of the lacIq mutants, which resulted in a tenfold increase in LacI expression (Calos and Miller 1981). When plasmids carry the lacIq gene, they are less dependent on *E. coli* (Brosius et al. 1983). That *E. coli* strains with a high concentration of LacI are still not sufficient for full lactose induction but are still fully inducible with non-metabolizable IPTG. Nevertheless, IPTG is not recommended for large-scale recombinant proteins production owing to its high cost, cellular toxicity and regulatory issues.

Another modified lac promoter system with widespread use is pET vector-based lac promoter. The pET vectors contain a strong T7 promoter controlled by T7 RNA polymerase (Studier 2014). The T7 RNA polymerase is chromosomally integrated using the prophage DE3 and controlled by lacUV5 promoter in the genome (Tabor and Richardson 1985). However, leakiness is a significant issue in the pET vectors that can be improved by adding either a copy of lacI gene in the expression vector, a lac operator downstream of the T7 promoter, or T7 lysozyme that controls the expression of T7 RNA polymerase independently (Studier 1991, 2014). These modifications lead to the development of commercially successful promoters to produce many heterologous proteins in *Escherichia coli* (Samuelson 2011). The lactose-based promoters are predominantly used to produce high expression levels of recombinant proteins in *E. coli*, and many lactose- or IPTG-based vectors have been designed to optimize recombinant proteins expression (Samuelson 2011). Hundreds of studies on lactose-based medium and process optimization have been performed to improve recombinant protein expression.

Recently, auto-induction has been gaining attention to produce recombinant proteins in *E. coli*. Auto-induction works when cells shift from an un-induced to an induced gene expression under metabolic control of the cell. The auto-inducing method comprises an optimal combination of glucose, lactose, glycerol and other essential nutrients in the medium (Chen et al. 2014). After glucose is exhausted, lactose is converted to allolactose, thus inducing the expression of recombinant protein through lactose-based promoter systems (Nie et al. 2013). Many proteins have been successfully produced using auto-induction in *E. coli*. For instance, auto-induction method has been demonstrated with enhanced expression of recombinant

tissue plasminogen activator (tPA) (Fathi-Roudsari et al. 2018). By optimizing parameters like temperature and medium composition, recombinant protein expression was significantly improved. This study found that a decrease in temperature and auto-induction medium highly influenced the yield of active soluble tissue plasminogen activator in *E. coli*. Using a highly enriched auto-induction medium, biologically active tPA was improved by 30%. Another example is the recombinant production of pullulanase which functionally participates in specific hydrolysis of starch processing. Recombinant pullulanase was successfully expressed in *E. coli* with 14 U/mL pullulanase activity when induced with IPTG (Nie et al. 2013).

Basal expression of recombinant proteins is prevalent in the lactose-based expression system, which might be disadvantageous to the host and ultimately lead to system instability issues and negatively affect the accumulation of recombinant protein. Therefore, repression of lactose operator in plasmid and genome of *E. coli* was constructed, and lactose operator was bound with lactose repressor to prevent T7 RNA polymerase activity and expression of target protein before induction. Auto-induction strategy has been used to considerably improve the recombinant pullulanase activity and yields (i.e., up to 580 U/mL) (Nie et al. 2013). Nitrile hydratase has been used to produce valuable chemicals such as acrylamide, nicotinamide, and 5-cyanovaleramide (Gupta et al. 2010). Recombinant nitrile hydratase was expressed and produced in *E. coli* using auto-induction. A glycerol limited fed-batch process produced 2170 U/mL of nitrile hydratase and biotransformation with recombinant nitrile hydratase exhibited a productivity of 187 g of nicotinamide/g dry cell weight/h. Auto-induction has been also used to produce industrial enzyme nitrile hydratase. *E. coli* has been considered the industrial workhorse for producing N-glycosylated proteins since the breakthrough work of engineering the glycosylation pathway from *Campylobacter jejuni* (Wacker et al. 2002). It is well known that N-glycosylation decides the biological activity of glycoproteins. Therefore, efficient glycosylation is a prerequisite, and the developing process is critically important. Different glycosylation sequons in *E. coli* carrying the heterologous pgl locus from *C. jejuni* have been used. Compared to IPTG induction, auto-induction method produced glycoproteins with 100% glycosylation efficiency (Ding et al. 2017).

3.2 Arabinose Induction System

Positively regulated systems are defined by transcriptional regulatory elements and activators, coupled with RNA polymerase binding to the specific region in the promoter, thus guiding transcription. Positively regulated systems are functionally characterized by a slower induction response.

pBAD is a classic model of a positively regulated promoter system induced by L-arabinose (Schleif 2000). araE and araFGH are two arabinose transport systems present in the arabinose operon and function to convert arabinose into xylulose-5-phosphate by the action of enzymes encoded by ribulokinase (araB), isomerase (araA), and epimerase (araD) (Schleif 2000). The araBAD genes are located close

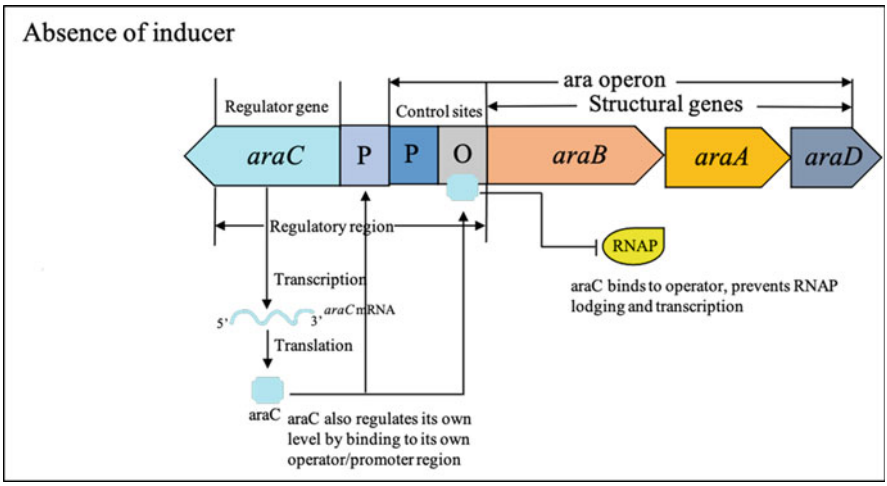


Fig. 3 Repressed arabinose operon. This shows the repressed state of arabinose operon, where the absence of inducer allows binding of araC repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

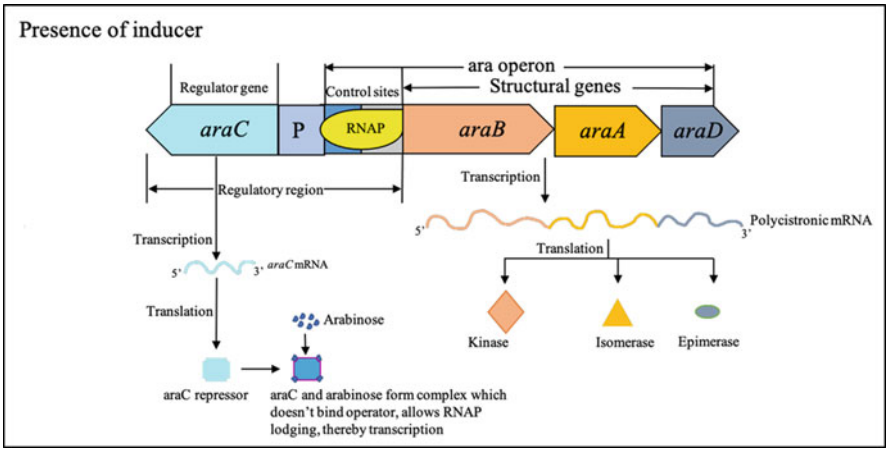


Fig. 4 Induced arabinose operon. This shows the induced state of arabinose operon, where the presence of inducer forms a complex with araC repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

to araFGH and araE in the operon. AraC protein regulates its own expression and ara operon genes (Figs. 3 and 4). The araC gene is located close to the araBAD operon but in opposite orientation (Johnson and Schleif 1995). It belongs to the araC/xylS family, which belongs to a positively regulated system (Mari et al. 1997). The araBAD gene products are induced approximately 300 times by AraC compared to the un-induced level. Arabinose promoter has the lowest basal expression level, but

the efficiency of repression is gene-dependent, and the repression is not tightly regulated. Induction of the genes also reflects catabolic repression, and induction is delayed due to the presence of glucose and some other sugars. Repression is controlled by the cAMP, which in turn affects the CRP activity. Studies of the araBAD promoter confirm that to activate transcription araC protein binding site must overlap the -35 region of the promoter by 4 bp. Moreover, the two half transcription sites recognized by AraC protein must be in the same direct repeat orientation to activate transcription. This is due to specific contacts made between RNA polymerase and AraC protein at pBAD promoter. In vivo studies in arabinose promoters found that araFGH expression, another arabinose promoter, is more sensitive to catabolite repression but not to arabinose concentration in comparison with araE and araBAD promoters. The relative levels of inducibility in wild-type cells of araBAD, araFGH, and araE have been reported to be 6.5, 5, and 1, respectively (Johnson and Schleif 1995). Researchers have concluded that all promoter systems are rapidly responsive and inducible by 0.53 mM of arabinose, and the arrangement of araC binding site is different among araBAD, araFGH, and areE promoters (Johnson and Schleif 1995). It is mechanistically known that CRP-cAMP and araC bind to specific promoter sites. AraC protein is known to bind to three sites within promoters such as araI, araO1, and araO2 (Ogden et al. 1980; Lee et al. 1981). However, initial studies found that only the transcriptionally active AraC protein could bind to the araI site (Ogden et al. 1980; Lee et al. 1981; Miyada et al. 1984). Later studies have proven that when arabinose is present, AraC protein can bind to all three sites (Hahn et al. 1984). It is known that CRP can bind to pBAD promoter region in *E. coli* and cause repression. Few mutational studies show that even in the absence of CRP, arabinose promoter is still threefold repressed when AraC protein is present (Hahn et al. 1984). Results suggest that some other mechanism may be involved in the repression in the absence of CRP-cAMP. When cells lack adenyl cyclase (cya), pBAD promoter expression is poor upon arabinose addition. However, when pBAD repression is abolished by deleting the operator (areO2) site in the promoter, elevated expression of the promoter in cya deleted strain is observed. Also, when araC expression was increased, it proportionally increased the expression of the pBAD promoter (Hahn et al. 1984). This observation is likely due to stimulation by CRP. Therefore, Hahn et al. (1984) concluded that pBAD promoter expression is reduced in the absence of CRP-cAMP upon the addition of arabinose. In the absence of CRP-cAMP repression, araC-arabinose is prevented from functioning at araI site to induce the expression.

The pBAD promoter from *E. coli* is used to express many recombinant proteins. Most recombinant plasmids contain araC gene due to low araC expression from a single chromosomal copy in *E. coli*, considering that many operators present on multicopy plasmids would require more araC repressor. Arabinose promoter varies between 250- and 1300-fold induction ratio in expression vectors and the basal level is very low due to carbon catabolite repression when glucose is present (Guzman et al. 1995). Although pBAD promoter is around 2.5- to 4.5-fold weaker than the tac promoter. Despite their low strength, recombinant proteins can accumulate up to 30% of total cell protein, depending on translation and kinetics. The induction can be

modulated by adding suboptimal concentrations of arabinose. Arabinose-based induction system has been successfully used to produce many recombinant products. Researchers have developed a high cell density fed-batch process to produce 5-hydroxymethylfurfural oxidase (HMFO) and eugenol oxidase (EUGO) (Román et al. 2020). They achieved high cell density culture (HCDC) using two stages of fed-batch process cultivated with glucose for improving biomass. Once glucose was exhausted, arabinose was added for induction. After inducing with arabinose, glycerol was used as an additional carbon source, and this resulted in an eightfold improvement in protein yield compared to IPTG-based inducer system (Román et al. 2020). Arabinose-based promoter expression has been tested for difficult to express proteins such as antibody fragments, membrane proteins, and vaccines. In one study, the expression of human recombinant tetanus toxoid and antibody fragment has been expressed under the control of pBAD and compared with the Lac promoter (Clark et al. 1997). Recombinant proteins were accumulated in a soluble fraction of periplasmic space, which is desirable for downstream processing. Compared with lac promoter induction using IPTG, production of Fab could be more strictly repressed under the control of arabinose promoter and low concentrations of arabinose were required, which is a significant advantage where production of a highly expressed Fab is toxic to the *E. coli* host.

E. coli NEB10β strain (Miret et al. 2020) has been used for the expression of the complex fusion protein phosphite dehydrogenase-cyclohexanone monooxygenase (PTDH–CHMO) using arabinose promoter. The fed-batch process was developed using a chemically defined medium supplemented with amino acids and glycerol. It resulted in a 9.2-fold improvement of the recombinant protein yields than in a complex medium and an accumulation of up to 2 g/L of PTDH-CHMO fusion protein after 6 h of induction. Arabinose-based expression was also demonstrated for the production of the alcohol dehydrogenase (ADH) enzyme using the same NEB10β strain (Miret et al. 2020). Arabinose expression system (Eberhardt et al. 2017) was tested along with T7, T5, tac, lactose expression system to produce thermostable steryl glucosidase in *Escherichia coli*. Based on comparative analysis, arabinose expression system offered a 40% improvement in shake flasks. Further process optimization in bioreactor enhanced the production of steryl glucosidase to 200-fold with a maximum activity of 260 U/mL after 6 h of arabinose induction in high cell density culture. Arabinose expression system has also been successfully demonstrated to produce high titer of inclusion bodies in *E. coli*. Human antimicrobial peptide LL-37 has been produced as a fusion protein using arabinose. The results show that active LL-37 can be produced at 1 g/L (Krahulec et al. 2010).

Many commercial vectors are constructed to induce recombinant proteins using arabinose. Apart from gene expression vectors, *E. coli* BL21-AI strain has been specifically developed to utilize the arabinose expression system. This strain carries a genomic insertion of a cassette containing the T7 RNA polymerase gene in the araB locus, allowing the expression of T7 RNA polymerase to be regulated by the araBAD promoter. BL21-AI strain is suitable for high-level recombinant protein expression from any T7-based expression vector. Levels of T7 RNA polymerase can be tightly regulated under arabinose control (Narayanan et al. 2011). High levels of

heterologous gene expression from T7 promoter led to low basal expression and makes it possible to express recombinant genes whose products are toxic for the host cells (Muntari et al. 2012). Another merit of arabinose promoter is its ability to fine-tune the expression level of recombinant proteins depending on its concentration in the medium (Guzman et al. 1995). The efficiency of the arabinose expression system is further increased by the higher stability of the plasmid DNA. These characteristic features of arabinose induction could be potentially used to produce recombinant proteins, which tend to form inclusion bodies when produced at high levels in *E. coli*.

3.3 Rhamnose Induction System

Rhamnose induction is another model for a positive regulation system. Rhamnose induction is a two-step tight control induction process. Its transporter (rhaT) controls L-rhamnose uptake and rhamnose is first converted to l-rhamnulose by an isomerase (rhaB), and then metabolized to rhamnulose-5-phosphate by a kinase (rhaA), and hydrolyzed by an aldolase (rhaD). The products dihydroxyacetone phosphate and L-lactaldehyde are consumed by other metabolic pathways (Tobin and Schleif 1987; Moralejo et al. 1993). The gene rhaBAD forms an operon located close to the rhaT gene (Fig. 5). Two activator proteins, rhaR and rhaS, are located upstream to rhaBAD which are arranged in opposite orientations. When rhamnose is consumed by the cell, rhaR protein binds with the rhamnose and becomes activated (Fig. 6). rhaR induces its own operon rhaSR. Once rhaS and rhaR accumulate high levels in the cell, then rhaS begins to activate the rhaBAD expression (Giacalone et al. 2006).

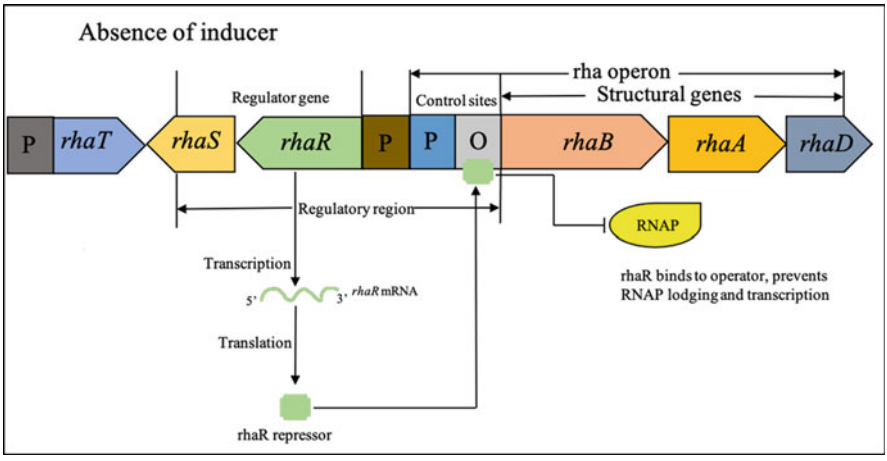


Fig. 5 Repressed rhamnose operon. This shows the repressed state of rhamnose operon, where the absence of inducer allows binding of rhaR repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

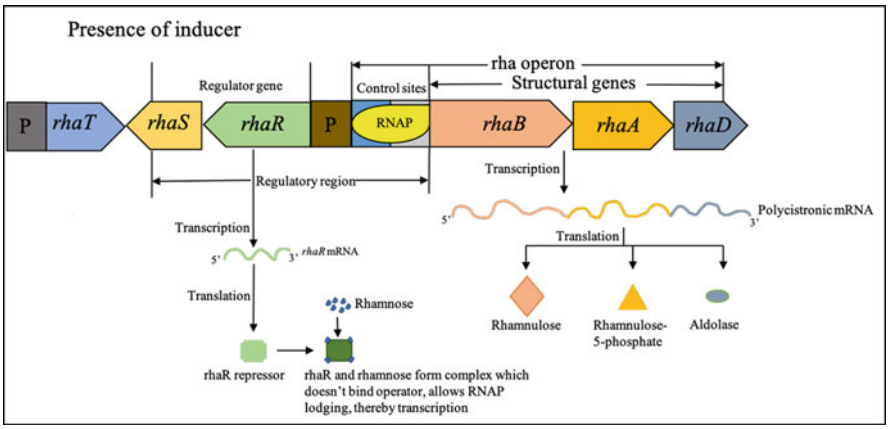


Fig. 6 Induced rhamnose operon. This shows the induced state of rhamnose operon, where the presence of inducer forms a complex with rhaR repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

Preliminary studies found that rhamnose induction can induce up to 30,000-fold. rhaS was produced at high levels even in the absence of rhamnose. Complementation analysis showed that rhamnose (rhaBAD) induction requires RhaS, not RhaR. Gene deletion and invitro transcription-translation assay studies found that RhaS protein is an essential regulator in rhaBAD induction. Promoter deleted strains found that two cis-acting elements are involved in rhaBAD induction. Deletion of upstream element resulted in 60-fold decrease in rhaBAD induction. DNA mobility shift assays concluded that CRP protein binds to the DNA region in rhaBAD operon. Thus, complete induction of rhaBAD expression requires CRP-cAMP complex (Badía et al. 1989; Via et al. 1996). When compared to araBAD promoter, rhaBAD promoter is tightly regulated. The basal level of rhaBAD is about tenfold lower than the lactose promoter but rhaBAD and araBAD promoter induction strengths are similar (Haldimann et al. 1998). In contrast to the case with the araBAD promoter, the regulatory genes rhaRS are critical for the function of rhaBAD and the chromosomal copy seems to produce enough regulatory proteins to saturate the binding sites on the expression vector. The induction times for attaining a high level of recombinant protein expression varied between 4 and 12 h. Therefore, prolonged response rates of rhaBAD promoter were found to be highly beneficial for proteins that have to be transferred into the periplasm or secretion into the medium (Giacalone et al. 2006). The rhaBAD regulatory system has been successfully used in HCD fermentation due to its low basal level of expression. A high cell-density fed-batch process for producing heterologous proteins in *E. coli* has been developed using rhamnose. The optimized process with rhamnose resulted in the production of 100 g/L cell dry weight and 3.8 g/L of recombinant L-N-carbamoylase from *E. coli* (Wilms et al. 2001). Production rates of membrane and secretory recombinant proteins need to be controlled in *E. coli* for obtaining high yields (Schlegel et al. 2013). rhaBAD based

promoter has been used to produce difficult to express proteins in *E. coli*, since the rhamnose promoter system allows precise tuning of the expression levels in a concentration-dependent fashion (Schlegel et al. 2013). In this study, sfGFP as a model protein was used for monitoring the in vivo kinetics of rhaBAD induction. Moreover, the researchers monitored the production of both the membrane and the secretory proteins in wild-type *E. coli* with the RhaT-mediated L-rhamnose uptake deficient single and double mutants (Hjelm et al. 2017). rhaB rhaT deleted strain improved production yields of all membrane and secretory proteins tested (Hjelm et al. 2017). Penicillin G amidase (PGA) from *Alcaligenes faecalis* was produced using rhamnose in *E. coli*. Production of PGA was induced by repeated addition of the inducer rhamnose. A biomass yield of 13.5 g/L of dry weight and PGA yield of 4500 units per liter were obtained (Deak et al. 2003).

Rhamnose induction system has been successfully used to manufacture recombinant therapeutic proteins. High cell density cultivation was developed in *E. coli* using the rhaBAD expression system. Single-chain antibody fragment was produced in fed-batch cultivation, and a specific product concentration of up to 20 mg/g was obtained. Slow and tight induction of rhamnose produces periplasmic 700 mg/L of antibody fragment in *E. coli* within 4 h (Lindner et al. 2014). Therefore, rhaBAD promoter has some characteristic features that make it exceptionally well suited for expressing the recombinant proteins. These include tight control of expression, low basal activity, and production rates controlled by rhamnose concentration, and it functions independent of any *E. coli* strain (Kelly et al. 2016). However, rhamnose suffers from one major limitation and that is its consumption as sugar by a specific pathway in *E. coli*. Due to its consumption in the central carbon metabolism, a decrease in inducer concentration is observed over time, resulting in lower expression and eventually leading to a decrease in product yield (Kelly et al. 2016). The problem of transient expression caused by inducer degradation has been addressed for some other promoter systems by using nonmetabolized inducer analogs such as IPTG and anhydrotetracycline (ATC). These inducer analogs will not be metabolized and have proven widely valuable for the production of recombinant products. Sugars resembling rhamnose in structure have been studied as potential inducers of the rhaBAD promoter system. Fluorescence studies revealed that L-mannose is the potential inducer among 35 sugars tested, as it provides a broader window of expression levels, good graded response to inducer concentration, and prolong induction (Kelly et al. 2016). L-Lyxose sugar may also be useful at lower expression levels, where it may perform better than L-mannose with respect to induction regulation. Due to commercial availability, L-Mannose and L-lyxose might be useful for bioprocess applications in the future (Kelly et al. 2016).

Most recombinant protein-producing genetic modules and expression systems were constructed in the last decade. Large amounts of soluble recombinant proteins are obtained by a high level of expression using different induction systems. However, the expression of difficult protein targets such as antibody fragments, membrane proteins, large proteins with many disulfide bonds using the expression systems continues to be a challenge with significant losses incurred during protein refolding (Sandomenico et al. 2020). Many strategies can be employed to slow down

the rate of recombinant protein expression and increase the expression level in the soluble fraction, e.g. lowering the growth temperature, optimizing inducer concentration, using an auto-induction medium, or employing a weaker promoter (Sandomenico et al. 2020). Although such tinkering can be very successful, determining the optimal strategy can be laborious as it is dependent on the particular target protein in question (Sandomenico et al. 2020).

3.4 pNEW Induction System

To broaden the applicability of the expression system across a wide range of *E. coli* strains a cumate (p-isopropylbenzoate)-regulated expression system has been developed by researchers (Choi et al. 2010). It was first developed for mammalian cells and later for methylotrophic bacteria (Mullick et al. 2006; Choi et al. 2006). A cumate-inducible expression system adapted for *E. coli* is designated as pNEW. It carries a synthetic operator and the regulator (*cymR*) of the *Pseudomonas putida* F1 *cmt* operon (Choi et al. 2006; Eaton 1996; Mullick et al. 2006). The target gene expression is controlled transcriptionally with these regulatory elements through chemical inducer cumate. Generally, the cumate switch requires four major components: a strong promoter, a repressor-binding DNA sequence or operator, expression of a repressor-encoding gene (*cymR*), and the chemical inducer cumate (Choi et al. 2006; Mullick et al. 2006). The absence of cumate keeps the system in an off state (Fig. 7) while the addition of cumate quickly changes the binding of the repressor CymR to the operator which further releases the repressor from the operator, resulting in the formation of the CymR-cumate complex and the expression of the

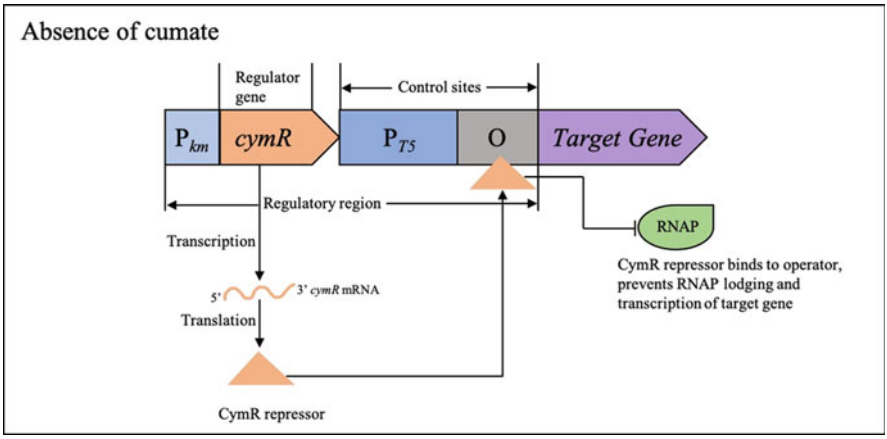


Fig. 7 Repressed cumate gene switch. This shows switched-off state of cumate system, *cymR* gene is under the control of weak constitutive promoter P_{km}. The levels of CymR produced is sufficient to inhibit transcription of target gene in absence of cumate

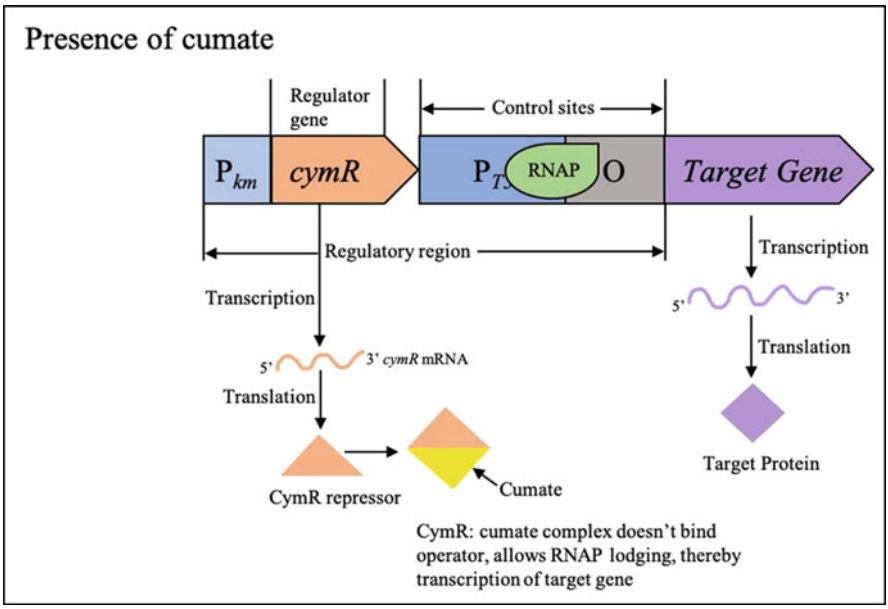


Fig. 8 Induced cumate gene switch. This shows switched on state of cumate system, where the presence of cumate, forms a complex with CymR repressor, induces conformational changes and repressor can no longer be able to bind with operator, thus allowing transcription of target gene

downstream target gene (Fig. 8). In pNEW plasmid, a 26 bp truncated operator fragment of the cmt operon was inserted downstream of the strong T5 promoter (PT5), which is mostly recognized by RNA polymerases in *E. coli* strains (Bujard et al. 1987). To minimize transcriptional read-through from PT5, the T7 terminator was provided just downstream of MCS, and for the selection of recombinant clones, kanamycin resistance gene (Kmr) was used. It was demonstrated that pNEW system is tightly regulated and operational across all *E. coli* strains, recommending its broad applicability (Choi et al. 2010). The rheostat (dose-dependent) control through cumate switch has also been determined in *E. coli* where authors measured the specific GFP yields of cells harboring the pNEW-gfp construct in which recombinant culture was induced with a range of cumate concentrations (0–122 mM). A linear, exponential specific GFP yield response (rheostat mechanism) was observed in the range of 0–10 mM cumate (subsaturating concentrations), which culminated in a wide range of specific GFP yields (0–90 mg/g dry weight of cells). The GFP yield of 90 mg/g was achieved at the higher limit of the linear response range, which signifies almost 50% of maximal achievable target protein during high-cell density fermentations induced with 100 mM (Choi et al. 2010). Expression of target proteins directed by pLac and pBAD promoters can also be controlled by changing the inducer concentrations, but the rheostat control functions only at much higher concentrations ($\geq$100 μM) of arabinose and IPTG (Giacalone et al. 2006; Hashemzadeh-Bonehi et al. 1998). Using the cumate switch, the rheostat zone is

feasible with more economical and lower concentrations of chemical inducer (cumate). If required, it also holds the potential to achieve high heterologous protein yields (50% of theoretical maximum) at higher inducer concentrations of the linear rheostat range (Choi et al. 2010). This tunability of production yield can be beneficial for process development (Kaur et al. 2009; Otto et al. 1995) and developing desired production strategies. Another important characteristic of the cumate induction system is a homogenous expression of microbial population post-induction with the inducer cumate, mainly when the purpose is to work through metabolically altered pathways. However, under similar culture conditions, pET-gfp resulted in lesser GFP yield than pNEW-gfp, which may be due to homogenous and consistent induction trait of cumate inducible system. Lower leakiness compared to pET system is the next important trait of cumate switch, which is mainly important when the expressed target protein is toxic. Therefore, cumate switch offers the margin to induce the recombinant cells in the desired growth phase and time, which results in optimum recombinant protein yield and optimum growth of the production host. The relatively quick response of cumate-based induction as compared to IPTG-based induction using pET-system is the next important characteristic that results in optimum protein yield. The plausible reason for relatively quick induction by cumate switch is attributed to its direct and simpler induction which only requires displacing the repressor CymR from the operator upon cumate addition thereby promoting the downstream transcription of target genes by existing *E. coli* RNA polymerases. However, IPTG-based induction is indirect, as it first induces the transcription of T7 RNA polymerase, which then further induces the transcription of heterologous target genes from the T7 promoter of pET-system (Choi et al. 2006). The specific target protein yield and biomass yield in high cell density fermentation of the recombinant *E. coli* resulted in higher recombinant protein (GFP) yield and biomass induced with similar concentrations of cumate and IPTG at the end of fermentation. The demonstration of cumate switch by Choi and co-workers showed that cumate offers several advantages over IPTG-based induction such as lower cost, non-toxic, non-leaky, ability to modulate induction using different inducer concentrations and it can be used across all *E. coli* hosts targeted for recombinant protein production. Hence, cumate based induction system can offer cost-effective benefits, particularly for large-scale bioprocesses compared to an IPTG-based induction system (Choi et al. 2006).

4 Conclusions and Future Trends

This chapter describes the different induction systems and genetically modified promoters used to facilitate the expression of recombinant proteins *in E. coli*. Problems concomitant with high-level overexpression have also been addressed. Several categories of promoter-inducer systems for heterologous production have been presented, each of them offering a variety of advantages and limitations. While

homogenous and long-term induction of recombinant proteins continues to be a challenge, significant advancements have been made in the past few decades.

IPTG-based induction from T7 expression system remains the gold standard as it ensures strong induction and also because IPTG is not metabolized by the cells. Nonetheless, it is known to cause metabolic burden and toxicity (Dvorak et al. 2015). Lactose-based induction offers several advantages where comparable product titers have been achieved without affecting cellular fitness (Bashir et al. 2016). The selection of an appropriate inducer system for target therapeutic production is a prerequisite. Development in inducer systems (arabinose, rhamnose, and cumate) allows precise regulation of recombinant protein production rates in a concentration-dependent manner. Recently, production using auto-induction is in trend where a combination of appropriate promoter and media/feed is designed in such a way that programmed induction can be achieved.

The prevalence of protein-based approved biopharmaceuticals seems to remain an industry reality in the near future. Amongst biopharmaceuticals approved, antibodies continue to dominate, but new gene-based and nucleic acid-based products and cellular therapies are also being launched. Also, there is a trend toward mammalian-based production, given that many recently approved products belong to the class containing post-translational modifications and therefore require a mammalian system (Walsh 2018). With increasing interest in the expression of more complex proteins, pressure continues for the development of microbial systems that can balance the quantity and quality of the product and its production cost (Castiñeiras et al. 2018). Additionally, along with the improvement in engineered N-glycosylation pathway in *E. coli*, recent developments in systems and synthetic biology arms (Gardner 2013) allow complete characterization of pathways that should accurately predict design models. *E. coli* remains a significant contributor to the biopharmaceutical industry.

References

Badía J, Baldomà L, Aguilar J, Boronat A (1989) Identification of the rhaA, rhaB and rhaD gene products from *Escherichia coli* K-12. FEMS Microbiol Lett 65:253–257

Bashir H, Ahmed N, Khan MA, Zafar AU, Tahir S, Khan MI et al (2016) Simple procedure applying lactose induction and one-step purification for high-yield production of rhCIFN. Biotechnol Appl Biochem 63(5):708–714

Bentley WE, Mirjalili N, Andersen DC et al (1990) Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnol Bioeng 35:668–681

Bonomo J, Gill RT (2005) Amino acid content of recombinant proteins influences the metabolic burden response. Biotechnol Bioeng 90:116–126

Brosius J, Amann E, Brosius J, Ptaslme M (1983) Vectors bearing a hybrid Trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene 25(2–3):167–178

Bujard H, Gentz R, Lanzer M et al (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. Methods Enzymol 155:416–433

Calos MP, Miller JH (1981) The DNA sequence change resulting from the IQ1mutation, which greatly increases promoter strength. MGG Mol Gen Genet 183:559–560

Castiñeiras TS, Williams SG, Hitchcock AG, Smith DC (2018) *E. coli* strain engineering for the production of advanced biopharmaceutical products. FEMS Microbiol Lett 365:1–10

Chen WB, Nie Y, Xu Y, Xiao R (2014) Enhancement of extracellular pullulanase production from recombinant *Escherichia coli* by combined strategy involving auto-induction and temperature control. Bioprocess Biosyst Eng 37:601–608

Choi YJ, Morel L, Bourque D et al (2006) Bestowing inducibility on the cloned methanol dehydrogenase promoter (PmxaF) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1. Appl Environ Microbiol 72(12):7723–7729

Choi YJ, Morel L, Le FT et al (2010) Novel, versatile, and tightly regulated expression system for *Escherichia coli* strains. Appl Environ Microbiol 76(15):5058–5066

Clark MA, Hammond FR, Papaioannou A et al (1997) Regulation and expression of human fabs under the control of the *Escherichia coli* arabinose promoter, PBAD. Immunotechnology 3: 217–226

De Boer HA, Comstock LJ, Vassert M (1983) The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc Natl Acad Sci U S A 80(1):21–25

Deak PM, Lutz-Wahl S, Bothe H, Fischer L (2003) Bioreactor cultivation of *Escherichia coli* for production of recombinant penicillin G amidase from *Alcaligenes faecalis*. Biotechnol Lett 25: 397–400

Ding N, Yang C, Sun S et al (2017) Increased glycosylation efficiency of recombinant proteins in Escherichia coli by auto-induction. Biochem Biophys Res Commun 485:138–143

Dvorak P, Chrast L, Nikel PI, Fedr R, Soucek K, Sedlackova M et al (2015) Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21 (DE3) carrying a synthetic metabolic pathway. Microb Cell Factories 14(1):1–15

Eaton RW (1996) p-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the cmt operon. J Bacteriol 178(5):1351–1362

Eberhardt F, Aguirre A, Menzella HG, Peiru S (2017) Strain engineering and process optimization for enhancing the production of a thermostable steryl glucosidase in Escherichia coli. J Ind Microbiol Biotechnol 44:141–147

Fathi-Roudsari M, Maghsoudi N, Maghsoudi A et al (2018) Auto-induction for high level production of biologically active reteplase in *Escherichia coli*. Protein Expr Purif 151:18–22

Gardner TS (2013) Synthetic biology: from hype to impact. Trends Biotechnol 31(3):123–125

Giacalone MJ, Gentile AM, Lovitt BT et al (2006) Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system. BioTechniques 40:355–364

Gupta SK, Shukla P (2017) Microbial platform technology for recombinant antibody fragment production: a review. Crit Rev Microbiol 43:31–42

Gupta N, Balomajumder C, Agarwal VK (2010) Enzymatic mechanism and biochemistry for cyanide degradation: a review. J Hazard Mater 176(1–3):1–13

Guzman L-M, Belin D, Carson MJ, Beckwith JON (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121–4130

Hahn S, Dunn T, Schleif R (1984) Upstream repression and CRP stimulation of the *Escherichia coli* l-arabinose operon. J Mol Biol 180:61–72

Haldimann A, Daniels LL, Wanner BL (1998) Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. J Bacteriol 180:1277–1286

Hashemzadeh-Bonehi L, Mehraein-Ghomi F, Mitsopoulos C et al (1998) Importance of using lac rather than ara promoter vectors for modulating the levels of toxic gene products in *Escherichia coli*. Mol Microbiol 30(3):676–678

Hjelm A, Karyolaimos A, Zhang Z et al (2017) Tailoring *Escherichia coli* for the l -rhamnose PBAD promoter-based production of membrane and secretory proteins. ACS Synth Biol 6:985–994

Hoffmann F, Rinas U (2004) Stress induced by recombinant protein production in *Escherichia coli*. Adv Biochem Eng Biotechnol 89:73–92

Inada T, Kimata K, Aiba H (1996) Mechanism responsible for glucose-lactose diauxie in *Escherichia coli*: challenge to the cAMP model. Genes Cells 1(3):293–301

Jaishankar J, Srivastava P (2017) Molecular basis of stationary phase survival and applications. Front Microbiol 8:2000

Johnson CM, Schleif RF (1995) In vivo induction kinetics of the arabinose promoters in *Escherichia coli*. J Bacteriol 177(12):3438–3442

Juers DH, Matthews BW, Huber RE (2012) LacZ β-galactosidase: structure and function of an enzyme of historical and molecular biological importance. Protein Sci 21(12):1792–1807

Kaur P, Agarwal S, Datta S (2009) Delineating bacteriostatic and bactericidal targets in mycobacteria using IPTG inducible antisense expression. PLoS One 4(6):e5923

Keasling JD (1999) Gene-expression tools for the metabolic engineering of bacteria. Trends Biotechnol 17(11):452–460

Kelly CL, Liu Z, Yoshihara A et al (2016) Synthetic chemical inducers and genetic decoupling enable orthogonal control of the rhaBAD promoter. ACS Synth Biol 5:1136–1145

Kercher MA, Lu P, Lewis M (1997) Lac repressor-operator complex. Curr Opin Struct Biol 7:76–85

Krahulec J, Hyršová M, Pepeliaev S et al (2010) High level expression and purification of antimicrobial human cathelicidin LL-37 in *Escherichia coli*. Appl Microbiol Biotechnol 88: 167–175

Lee NL, Gielow WO, Wallace RG (1981) Mechanism of araC autoregulation and the domains of two overlapping promoters, Pc and PBAD, in the L-arabinose regulatory region of *Escherichia coli*. Proc Natl Acad Sci 78:752–756

Lindner R, Moosmann A, Dietrich A et al (2014) Process development of periplasmatically produced single chain fragment variable against epidermal growth factor receptor in *Escherichia coli*. J Biotechnol 192:136–145

Malan TP, McClure WR (1984) Dual promoter control of the *Escherichia coli* lactose operon. Cell 39:173–180

Mari M, Gallegos M-T, Schleif R et al (1997) AraC/XylS family of transcriptional regulators

Miret J, Román R, Benito M et al (2020) Development of a highly efficient production process for recombinant protein expression in *Escherichia coli* NEB10β. Biochem Eng J 159:107612

Miyada CG, Stoltzfus L, Wilcox G (1984) Regulation of the araC gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on araBAD expression. Proc Natl Acad Sci 81:4120–4124

Moralejo P, Egan SM, Hidalgo E, Aguilar J (1993) Sequencing and characterization of a gene cluster encoding the enzymes for L-rhamnose metabolism in *Escherichia coli*. J Bacteriol 175: 5585–5594

Mullick A, Xu Y, Warren R et al (2006) The cumate gene-switch: a system for regulated expression in mammalian cells. BMC Biotechnol 6:43

Muntari B, Amid A, Mel M et al (2012) Recombinant bromelain production in *Escherichia coli*: process optimization in shake flask culture by response surface methodology. AMB Express 2: 12

Narayanan A, Ridilla M, Yernool DA (2011) Restrained expression, a method to overproduce toxic membrane proteins by exploiting operator–repressor interactions. Protein Sci 20:51–61

Neubauer P, Wolff C, Hecker M et al (1991) Introduction of the tac-promoter by lactose under fermentation conditions. Acta Biotechnol 11:23–29

Nie Y, Yan W, Xu Y et al (2013) High-level expression of *Bacillus naganoensis* pullulanase from recombinant *Escherichia coli* with auto-induction: effect of lac operator. PLoS One 8(10): e78416

Nieto C, Fernández De Palencia P, López P, Espinosa M (2000) Construction of a tightly regulated plasmid vector for *Streptococcus pneumoniae*: controlled expression of the green fluorescent protein. Plasmid 43(3):205–213

Oehler S, Amouyal M, Kolkhof P et al (1994) Quality and position of the three lac operators of *E. coli* define efficiency of repression. EMBO J 13:3348–3355
Oehler S, Eismann ER, Kramer H, Muller-Hill B (1990) The three operators of the lac operon cooperate in repression. EMBO J 9:973–979
Ogden S, Haggerty D, Stoner CM et al (1980) The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. Proc Natl Acad Sci U S A 77(6):3346–3350
Otto CM, Niagro F, Su X, Rawlings CA (1995) Expression of recombinant feline tumour necrosis factor is toxic to *Escherichia coli*. Clin Diagn Lab Immunol 2(6):740–746
Rava PS, Somma L, Steinman HM (1999) Identification of a regulator that controls stationary-phase expression of catalase-peroxidase in *Caulobacter crescentus*. J Bacteriol 181(19):6152–6159
Román R, Lončar N, Casablancas A et al (2020) High-level production of industrially relevant oxidases by a two-stage fed-batch approach: overcoming catabolite repression in arabinose-inducible Escherichia coli systems. Appl Microbiol Biotechnol 104:1–9
Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:172
Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: a 5-year update. Protein Sci 28:1412–1422
Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307: 249–264
Saïda F, Uzan M, Odaert B, Bontems F (2006) Expression of highly toxic genes in *E. coli*: special strategies and genetic tools. Curr Protein Pept Sci 7(1):47–56
Samuelson JC (2011) Recent developments in difficult protein expression: a guide to *E. coli* strains, promoters, and relevant host mutations. Methods Mol Biol 705:195–209
Sandomenico A, Sivaccumar JP, Ruvo M (2020) Evolution of *Escherichia coli* expression system in producing antibody recombinant fragments. Int J Mol Sci 21:1–39
Schlegel S, Rujas E, Ytterberg AJ et al (2013) Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. Microb Cell Factories 12:1–12
Schleif R (2000) Regulation of the L-arabinose operon of *Escherichia coli*. Trends Genet 16(12): 559–565
Shibui T, Uchida M, Teranishi Y (1988) A new hybrid promoter and its expression vector in *Escherichia coli*. Agric Biol Chem 52:983–988
Shin PK, Seo J-H (1990) Analysis of *E. coli* phoA-lacZ fusion gene expression inserted into a multicopy plasmid and host cell's chromosome. Biotechnol Bioeng 36(11):1097–1104
Simpsonl RB (1980) Interaction of the cAMP receptor protein with the lac promoter. Nucleic Acids Res 8(4):759–766
Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol 115:113–128
Studier FW (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J Mol Biol 219(1):37–44
Studier FW (2014) Stable cultures and auto-induction for inducible protein production in *E. coli*. Struct Genomics Methods Protoc 1091:17–32
Tabor S, Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc Natl Acad Sci 82:1074–1078
Tobin JF, Schleif RF (1987) Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. J Mol Biol 196:789–799
Tropp BE (2008) Molecular biology: genes to proteins. Jones and Bartlett Publishers
Via P, Badia J, Baldomá L et al (1996) Transcriptional regulation of the *Escherichia coli* rhaT gene. Microbiology 142:1833–1840
Wacker M, Linton D, Hitchen PG et al (2002) N-linked glycosylation in *campylobacter jejuni* and its functional transfer into *E. coli*. Science 298(5599):1790–1793

Walsh G (2018) Biopharmaceutical benchmarks 2018. Nat Biotechnol 36:1136–1145

Wanner BL (1996) Phosphorus assimilation and control of the phosphate regulon. In: Neidhardt FC, Curtiss RI, Gross CA et al (eds) *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. 2nd ed, chapter 84. Am Soc Microbiol, Washington, DC

Wheatley RW, Lo S, Jancewicz LJ et al (2013) Structural explanation for allolactose (lac operon inducer) synthesis by lacZ β-galactosidase and the evolutionary relationship between allolactose synthesis and the lac repressor. J Biol Chem 288:12993–13005

Wilms B, Hauck A, Reuss M et al (2001) High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. Biotechnol Bioeng 73:95–103

Microbial Biosynthesis of Straight-Chain Aliphatic Carboxylic Acids

Lei Zhuang, Yuxin Liu, and Haoran Zhang

Contents

Abstract Straight-chain aliphatic carboxylic acids are valuable biomolecules with wide applications in various industries. Microbial biosynthesis offers an effective and environmentally friendly approach to the production of bioproducts from renewable feedstock materials. This book chapter reviews the current research progress of microbial biosynthesis of straight-chain aliphatic carboxylic acids. The biosynthesis pathways for exemplary products and associated engineering efforts for production improvement are highlighted to demonstrate the research status in this promising and developing field.

L. Zhuang · Y. Liu · H. Zhang (✉)
Department of Chemical and Biochemical Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
e-mail: Haoran.Zhang@Rutgers.edu

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_2

1 Overview

Straight-chain aliphatic carboxylic acids are a group of important molecules with various chemical properties and a wide range of industrial values. Traditionally, many of these molecules have large global markets and are produced mainly from the petroleum industry. However, large-scale production of these products often requires high reaction temperature and harsh conditions, consumption of nonrenewable crude oil, and use of reactant/catalytic materials that are not environmentally friendly. In recent years, microbial biosynthesis has been developed as a robust alternative technology for the production of a variety of bioproducts. The biosynthesis processes recruit microbes to convert renewable feedstocks to desired products through enzymatic reactions, which can be conducted under mild conditions without using any harsh chemicals. As such, microbial biosynthesis offers great advantages in process costs, sustainability, and environmental friendliness.

To date, a variety of microbes, either natural producers or metabolically engineered heterologous hosts, have been widely used for biosynthesis of straight-chain aliphatic carboxylic acids from renewable feedstocks; these include popularly used bacteria (e.g., *Escherichia coli*, *Corynebacterium glutamicum*), yeast (e.g., *Saccharomyces cerevisiae*, *Yarrowia lipolytica*), and microalgae (e.g., *Phaeodactylum tricornutum*). Assisted by the advances of related research disciplines, such as synthetic biology, metabolic engineering, and directed evolution, there is a boom of studies on microbial biosynthesis of straight-chain aliphatic carboxylic acids in recent years. Although the bioproduction performance for many of these molecules is still not yet ready for meeting the requirements of industrial production, the accumulated knowledge, techniques, and experience pave the way for its practical applications in the future.

This chapter provides a summary of the research status of microbial biosynthesis of straight-chain aliphatic carboxylic acids. In particular, we review representative groups of acids with high research and industrial values. The biosynthesis pathway development, metabolic engineering strategies adopted for biosynthesis improvement, and recent achievements of large-scale bioproduction are discussed. The challenges and opportunities for future development are also discussed.

2 Biosynthesis of Straight-Chain Aliphatic Carboxylic Acids

2.1 C1 Formic Acid

Formic acid is a single carbon organic acid and has a long history of being used in different industries. For example, it is a food preservative and antibacterial agent, and it has been used as an important intermediate for the synthesis of other valuable chemicals. What makes formic acid biosynthesis special is that this molecule can be

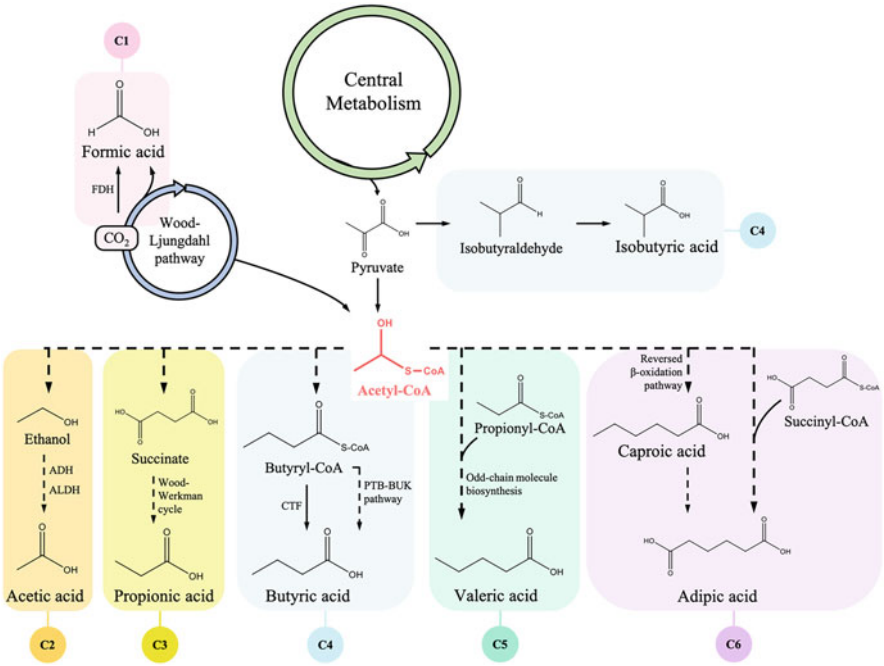


Fig. 1 Schematic of representative pathways for C1-C6 straight-chain aliphatic carboxylic acid biosynthesis. Solid and dash arrows represent single-step and multi-step enzymatic conversion, respectively. *FDH* formate dehydrogenase, *ADH* alcohol dehydrogenase, *ALDH* aldehyde dehydrogenase, *CTF* CoA-transferase

produced from carbon dioxide. As such, it carries strong research significance for green-house gas fixation.

CO_2 can be reduced to formic acid by formate dehydrogenase using NADH or electrode-derived electrons as the reducing power (Choe et al. 2014; Le et al. 2018). Alternatively, the reaction can be carried out by other enzymes such as carbon dioxide reductase using hydrogen gas as the reducing power (Schuchmann and Müller 2013; Eguchi et al. 1985; Alissandratos et al. 2014). The CO_2 reduction is also part of the Wood–Ljungdahl pathway that is used for other acid biosynthesis as discussed in the sections below (Fig. 1). Compared with other organic acids, the research and application of formic acid bioproduction is very limited. This is mainly because of the low efficiency of the bioconversion process due to technical difficulties such as CO_2 mass transfer limitations. To date, only a small number of microbes, such as *A. woodii*, *E. coli*, and *S. oneidensis*, have been investigated for formic acid bioproduction. For example, one recent study reported the 20 mM formic acid biosynthesis from CO_2, and the production was improved when HCO_3^- was used as the carbon substrate.

2.2 C2 Acetic Acid

Acetic acid is an industrially important compound that has been widely used for the production of synthetic polymers (e.g., polyvinyl acetate), photographic film materials (e.g., cellulose acetate), organic solvent (e.g., ethyl acetates), vinegar, and other products.

Acetic acid biosynthesis has been well studied. There have been extensive efforts for the utilization of various microbes for acetic acid production. The main microbes recruited for the bioproduction are native producers, namely acetic acid bacteria. So far, a few different pathways have been discovered and utilized for acetic acid biosynthesis. The first pathway is ethanol oxidation in which acetyl-CoA, a central metabolite, is converted to ethanol and subsequently undergoes two consecutive oxidation reactions catalyzed by alcohol dehydrogenase and aldehyde dehydrogenase, respectively (Fig. 1). Since the pathway precursor ethanol is derived from glycolysis, strong metabolic flux can be attracted to acetic acid formation and the overall biosynthesis performance is high. For example, using this pathway, acetic acid bacteria have been adopted for high-level bioproduction of acetic acid at industrially relevant scales (Gullo et al. 2014). On the other hand, acetyl-CoA can be converted to acetate phosphate and then acetic acid. This route adopted the Wood–Ljungdahl pathway for fixation of CO_2 to acetic acid (Ragsdale and Pierce 2008). Specifically, this pathway comprises two branches that convert CO_2 to formic acid and CO, respectively, which are then combined to generate acetyl-CoA and acetic acid. Notably, the reducing equivalent needed for Wood–Ljungdahl pathway can be provided by glycolysis and subsequent conversion of acetyl-CoA to acetic acid. As such, the coupling of Wood–Ljungdahl pathway with glycolysis/acetyl-CoA conversion will lead to the formation of three acetic acid molecules from one glucose molecule, eliminating the CO_2 emission.

Besides the Wood–Ljungdahl pathway, it has been found that CO_2 can be fixed to make acetic acid through the glycine synthase pathway. Notably, it has been found that the Wood–Ljungdahl pathway and the glycine synthase pathway are not totally independent; instead, they are functionally interconnected for CO_2 fixation (Song et al. 2020).

Large-scale cultivation has increased the acetic acid biosynthesis level to dozens of grams per liter in many reports (Vidra and Németh 2018). One challenge for acetic acid bioproduction is product toxicity. To this end, it is important to increase the resistance of the recruited microbial host (Nakano et al. 2004) or remove acetic acid in situ using bioprocess engineering techniques (Karekar et al. 2020).

2.3 C3 Propionic Acid

Propionic acid is one of the top value-added compounds from biomass (Werpy and Petersen 2004). It is a Generally Recognized as Safe (GRAS) compound with

various important chemical properties. Its derivatives have been widely used as food preservatives, personal care products, and pharmaceutical substances. Moreover, propionic acid can be used as a building block molecule for the synthesis of many other industrially important molecules. The biosynthesis of this C3 organic acid has been well studied and widely achieved using various microbes (Ranaei et al. 2020; Eş et al. 2017).

Several pathways have been discovered for propionic acid biosynthesis. For example, succinate in the TCA cycle can be decarboxylated to make propionic acid via propionyl-CoA. This pathway also called Wood–Werkman cycle, is a popular pathway used by many native producers. Another well-studied pathway uses lactic acid to produce propionic acid via lactoyl-CoA and acryloyl-CoA. In addition, malonyl-CoA of the fatty acid pathway can be converted to 3-hydroxyproionic acid which is then used to generate propionic acid. Acetyl-CoA can be combined with pyruvate to make propionic acid via citramalate. Some simple amino acids, such as valine, isoleucine, threonine, and methionine, can also be converted to propionic acid through a series of enzymatic reactions.

Propionic acid can be produced by a variety of microbes, such as several species of *Propionibacterium* and *Veillonella*. Most of the previous studies focused on using bioprocess engineering techniques, e.g., manipulation of the cultivation conditions, to improve the bioproduction performance (Ranaei et al. 2020; Eş et al. 2017; Gonzalez-Garcia et al. 2017). To this end, extensive efforts have been made for cultivation optimization, which led to high-level production. For example, 97 g/L propionic acid production with a yield of 0.54 g/g glucose was achieved using a *Propionibacterium acidipropionici* mutant strain (Zhang and Yang 2009). Also, a variety of different carbon substrates, such as glucose, glycerol, and whey lactose, have been used for propionic acid biosynthesis (Ranaei et al. 2020). It is noteworthy that most of the native producers are difficult to modify using genetic or metabolic engineering tools. So far, only a few microbes, such as *P. acidipropionici* and *P. jensenii* have been genetically engineered for propionic acid biosynthesis. As such, strain evolution by random mutagenesis and genome shuffling is an alternative approach for production improvement. On the other hand, there are an increasing number of studies using non-native producers, such as model microbe *E. coli*, for propionic acid biosynthesis (Akawi et al. 2015; Kandasamy et al. 2013). These microbes are more genetically tractable and thus can be rationally engineered for bioproduction. Yet, the bioproduction performance of these microbes so far is still not comparable with the native producers.

2.4 C4 (1) Butyric Acid

Butyric acid has been widely used in food and pharmaceutical industries (Wang et al. 2016); it is also an important precursor for many valuable chemicals such as butanol (biofuel) and cellulose acetate butyrate (thermoplastic) (Richter et al. 2012; Zhang

et al. 2009). Although butyric acid has an unpleasant odor, its derivatives are popular fragrance molecules (Jiang et al. 2018).

Butyric acid biosynthesis has been established in a variety of different microbes. The development of bioproduction using sustainable materials, such as lignocellulose, industrial byproducts, and municipal waste, has also been the focus of recent studies. Generally, the first step of using sugar-based feedstocks is to generate simple carbon substrates such as glucose or xylose. The carbon substrates are converted to pyruvate and further to acetyl-CoA through the central metabolism. Acetyl-CoA is then converted to butyryl-CoA by several enzymes, including thiolase and butyryl-CoA dehydrogenase. From butyryl-CoA, butyric acid is produced either through the phosphotransbutyrylase(PTB)-butyrate kinase(BUK) pathway or by the CoA-transferase(CTF) (Jiang et al. 2018) (Fig. 1). The competing pathways include the conversion from pyruvate to lactate, from acetyl-CoA to acetate and ethanol, and the generation of butanol from butyryl-CoA.

The main butyric acid-producing microbes are *Clostridium*, *Butyrivibrio*, *Butyribacterium*, *Eubacterium*, and *Fusobacterium (*Duncan et al. 2002*;* Pituch et al. 2013*). Clostridium*, which has been genome-sequenced and widely studied, is promising for industrial applications due to its ability of using various feedstocks and high production performance (Dwidar et al. 2012). For example, *Clostridium tyrobutyricum* ATCC 25755 was able to use xylose to produce 57.9 g/L butyric acid with a yield of up to 0.59 g/g (Zhu and Yang 2004). When using glucose as the substrate, the production of butyric acid reached 73.38 g/L with a yield of 0.46 g/g (Song et al. 2010). However, the difficulties associated with genetic engineering of *Clostridium* remain a big challenge. *E. coli*, has also been metabolically engineered to produce 10 g/L butyric acid from glucose with a high product selectivity (butyric acid/acetic acid ratio: 143 g/g) (Saini et al. 2014). The highest production by *E. coli* is 14.3 g/L, although the yield is still low (0.11 g/g glucose). This is probably due to the poor butyric acid tolerance (Jawed et al. 2016). *Thermobifida fusca*, a microbe with robust cellulases and hemicelluloses for lignocellulose degradation, has been metabolically engineered to express an exogenous acetate CoA-transferase to produce butyric acid from lignocellulose. Grown on 90 g/L corn stover as the substrate, such engineered *T. fusca* achieved a production of 17.1 g/L butyric acid (Deng et al. 2015).

2.5 C4 (2) Isobutyric Acid

Isobutyric acid, also known as 2-methylpropanoic acid, is one critical chemical for the production of plastics, resins, pharmaceuticals, and food industry products. Traditional manufacturing of isobutyric acid relies on the oxidation of petroleum-derived isobutyraldehyde.

For the isobutyric acid biosynthesis, glucose can enter the central metabolism to produce pyruvate and then isobutyraldehyde, which is converted to isobutyric acid by alcohol dehydrogenase. Using this pathway, engineered *E. coli* was reported to

produce 11.7 g/L isobutyric acid (Zhang et al. 2011). Moreover, by deleting the competing isobutanol pathway and overexpressing additional enzymes for isobutyraldehyde conversion to isobutyric acid, a high production yield of 0.39 g/g glucose was achieved (Xiong et al. 2015). After the process scale-up and cultivation condition optimization, the production was boosted to 90 g/L (Xiong et al. 2015). Other microbes, such as *Pseudomonas*, were also metabolically engineered to produce isobutyric acid, although the production performance was relatively low (Lang et al. 2014).

2.6 C5 Valeric Acid

Valeric acid, also called pentanoic acid, has long been used as a food additive or flavoring agent. Valeric acid has an unpleasant odor, but many of its esters are pleasant-smelling chemicals and therefore popularly used in perfumes and cosmetics industries. In addition, valeric acid and its derivatives can be used for making biofuel additives and bioplastics.

Valeric acid biosynthesis is under development, as there are only limited efforts for biosynthesis studies compared with other straight-chain carboxylic acids. This is largely due to the difficulties associated with engineering the biosynthesis pathway (s) and relatively limited industrial values of this C5 molecule. It has been reported that valeric acid can be biosynthesized from pyruvate through an engineered pathway (Dhande et al. 2012). Such an artificial pathway uses threonine, 2-ketobutyrate, and 2-ketovalerate as the intermediates and has been successfully reconstituted in *E. coli* for 2.58 g/L production from glucose. Alternatively, valeric acid can be biosynthesized through the combination of propionyl-CoA and acetyl-CoA, which generates a series of intermediates including ketovaleryl-CoA, trans-2-pentenoyl-CoA, and valeryl-CoA and finally results in the formation of valeric acid. Using this designed pathway, 398 mg/L valeric acid was produced from 10 g/L glycerol (Tseng and Prather 2012). Notably, there are also efforts using simple organic acids as the substrates for microbial biosynthesis. For example, propionate and methanol were used for valeric acid production in an anaerobic chain elongation open-culture bioreactor (De Smit et al. 2019). Large-scale production of valeric acid using engineering microbes has not been extensively studied, yet.

2.7 C6 (1) Caproic Acid

Caproic acid, also called hexanoic acid, is an oily acrid liquid with wide application in different industries. For example, it can be used as an antimicrobial agent, food additive, lubricating agent, etc. Moreover, it is a building block molecule for the chemical synthesis of various downstream products. Its esters are particularly useful for food and pharmaceutical industries.

A variety of natural anaerobic bacteria, such as *Clostridium kluyveri*, *Megasphaera elsdenii*, and *Eubacterium limosum*, are capable of producing caproic acid (Cavalcante et al. 2017). The caproic acid biosynthesis uses the mechanism of carboxylic acid chain elongation, which is essentially a reverse β-oxidation pathway. Specifically, acetyl-CoA is used as the starting molecule for the carboxylic acid chain formation, which proceeds via intermediate metabolites such as acetoacetyl-CoA, butyryl-CoA, and hexanoyl-CoA. Since many intermediates are involved in the pathway, biosynthesis requires a large number of enzymes, which are mostly active in the native producers. Lactic acid and ethanol are often used as electron donors for this biosynthesis process under anaerobic conditions. For example, ethanol can be converted to acetic acid to generate the energy needed for biosynthesis; acetic acid is subsequently used as the carbon source for the caproic acid chain formation. Notably, a similar pathway can be used for butyric acid as discussed in the section above.

Many caproic acid biosynthesis studies utilize the methods that are less commonly used for other carboxylic acid bioproduction (e.g., open cultures, mixed cultures, unsterile conditions). For example, it has been reported that, using a microbiome obtained from a fermentation pit, the caproic acid bioproduction reached 23 g/L under optimized conditions (Zhu et al. 2015). Most of the current studies rely on the use of a mixed substrate, such as acetic acid and ethanol, for high bioproduction.

2.8 C6 (2) Adipic Acid

Adipic acid, also named hexanedioic acid, is the most important dicarboxylic acid with a huge global market. This C6 molecule has a carboxyl group at each end of the linear chain, providing special chemical properties (e.g., pKa of 4.43 and 5.41) and outstanding industrial values. For example, adipic acid is a precursor for the manufacturing of commodity chemical nylon. It is also used as a food ingredient and a material for the production of polyurethanes.

There are two major pathways for adipic acid biosynthesis. The first pathway is oxidation of caproic acid (discussed in the section above) via the fatty acid ω-oxidation pathway. Via this pathway, the production of more than 50 g/L adipic acid was achieved using an industrial yeast strain (Beardslee and Picataggio 2012). Another pathway is the combination of acetyl-CoA and succinyl-CoA to generate 3-oxoadipyl-CoA which is subsequently converted to 3-hydroxyadipyl-CoA, 5-carboxy-2-pentenyl-CoA, adipyl-CoA, and finally adipic acid. This is also called the reverse adipate degradation pathway. In a recent report, 2.23 g/L adipic acid was produced from 50 g/L glucose in a *T. fusca* strain harboring this pathway (Deng and Mao 2015). Similarly, Cheong et al. biosynthesized 2.5 g/L adipic acid from succinyl-CoA and acetyl-CoA using engineered *E. coli (*Cheong et al. 2016*)*.

In addition, there are other adipic acid production routes that use both biosynthesis and chemical reactions. For example, glucaric acid and *cis,cis*-muconic acid,

both of which can be obtained from simple carbon substrates by microbial biosynthesis, can be converted to adipic acid using catalyzed hydrogenation reactions. Notably, *cis,cis*-muconic acid can also be produced from lignin through depolymerization reactions and enzymatic aromatic degradation, which greatly expands the source for adipic acid production. These pathways have been recruited for the production using different microbes (Kruyer and Peralta-Yahya 2017; Deng et al. 2016). However, depending on the starting materials and the adopted biosynthesis conditions, the production efficiency varies significantly between them.

2.9 C7 Heptanoic Acid

Heptanoic acid, or enanthic acid, is an odd-chain acid with an unpleasant odor. It can be used for the production of biodiesel, antibiotic agents, fragrances, and bioplastics.

Compared with other acids, there are only a few potential routes for heptanoic acid biosynthesis. The main route is through the fatty acid biosynthesis mechanism (Fig. 2). However, instead of acetyl-CoA, this pathway uses propionyl-CoA as the starting molecule; and by incorporation of C2 extension unit, it results in the formation of C5 and then C7 acids. Some microbes have the enzymes needed for the utilization of this pathway and thus have been used for microbial bioproduction.

It should be noted that, due to the nature of the heptanoic acid biosynthesis pathway, many microbial biosynthesis processes tend to produce a mixture of different medium-chain carboxylic acids. For example, a mixed culture cultivated in an upflow anaerobic filter reactor was engineered to produce valerate, caproate, and heptanoate all at the gram per liter scale (Grootscholten et al. 2013). Co-production of multiple acid products is also very common for the mono-culture of one single microbe. Jeon and coworkers utilized the culture of a *Megasphaera* strain and produced 5.7 g/L pentanoic acid, 9.7 g/L hexanoic acid, 3.2 g/L heptanoic acid, and 1.2 g/L octanoic acid (Jeon et al. 2016). However, the efforts for high production of heptanoic acid as the specific main bioproduct are still limited. Also, separating heptanoic acid from other byproducts presents an additional challenge of high-level bioproduction.

2.10 Medium-Chain Straight Aliphatic Carboxylic Acids

2.10.1 Medium-Chain Fatty Acids

The definition of medium-chain fatty acids, i.e., the length of the carbon chain, varies in different reports. For this work, medium-chain fatty acids refer to the C8-C12 fatty acids. Medium-chain saturated fatty acids possess stronger hydrophobicity than the short-chain acids discussed above. As such, they have distinct chemical properties and industrial values.

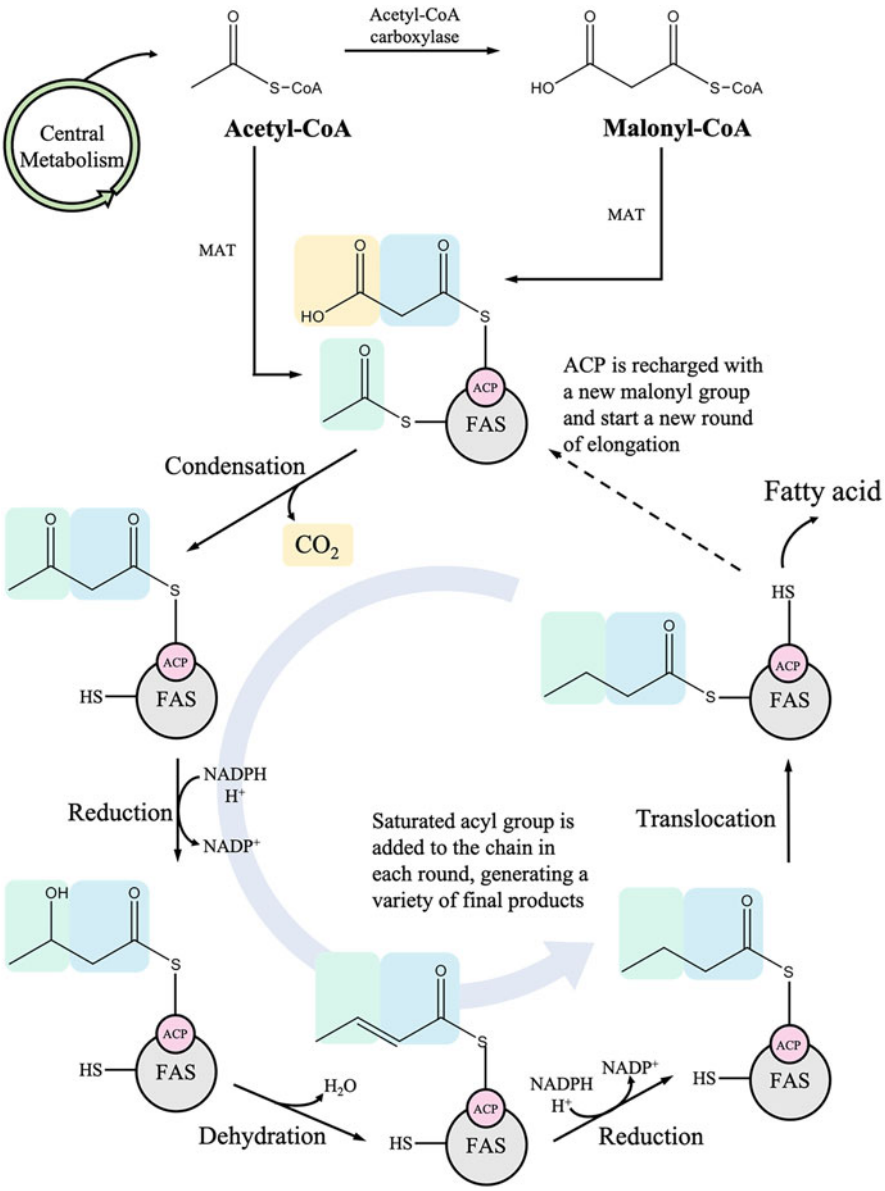


Fig. 2 Schematic of the fatty acid pathway for biosynthesis of straight-chain aliphatic carboxylic acids. *MAT* malonyl/acetyl transferase, *FAS* fatty acid synthase, *ACP* acyl carrier protein

C8 Octanoic acid (caprylic acid) is an oily liquid with very low water solubility and light odor. The main use of octanoic acid is to produce esters that can be used in perfume and dye industries. Octanoic acid itself has antimicrobial activities and thus

has been used as a disinfectant. In addition, as a medium-chain fatty acid, it can be taken as a dietary supplement.

C9 Nonanoic acid (pelargonic acid) is a naturally occurring fatty acid found in many plants. It is mainly used as an herbicide and blossom thinner in agriculture and an additive for personal-care products. Its synthetic esters are commonly used for food flavoring.

C10 Caprici acid (decanoic acid or decylic acid) is a C10 saturated fatty acid. Its esters are often used for the production of perfumes, lubricants, food additives, and plastics. In addition, it can be linked to pharmaceutical molecules to generate prodrugs with better lipophilicity that is particularly useful for drug injection.

C11 Undecylic acid (undecanoic acid) is a highly hydrophobic compound with poor water solubility. It is also a good antifungal agent.

C12 Lauric acid (dodecanoic acid) is a twelve-carbon fatty acid with special chemical properties that makes it an effective surfactant. It is historically used for the manufacturing of soaps, detergents, pesticides, and cosmetics.

For medium-chain acids, biosynthesis often utilizes the fatty acid biosynthesis pathway, which is a ubiquitous process in many organisms and has been extensively studied in the open literature. Specifically, acetyl-CoA and malonyl-CoA (also derived from acetyl-CoA) are the substrates for fatty acid biosynthesis. A set of repetitive decarboxylative Claisen condensation reactions are used for carbon chain elongation via ACP-intermediates. Meanwhile, enzymatic reactions, such as keto-reduction, dehydration, and enoyl reduction, take place on the extending chain and result in the formation of saturated or unsaturated fatty acids.

For medium-chain acid biosynthesis using the fatty acid biosynthesis pathway, the key challenge is how to control the length of the chain, i.e., how to stop the chain elongation process at the proper timing to make the desired products (Heil et al. 2019). To this end, it is important to engineer the reaction specificity of the enzymes involved in the chain elongation steps. On the other hand, despite the efforts in engineering the pathway enzymes, it is almost entirely unavoidable that some fatty acid byproducts, either straight or branched, are co-produced by the recruited microbial host. This, therefore, presents a challenge for the products' downstream separation and purification. In fact, compared with short-chain acid, medium-chain fatty acid biosynthesis using microbial hosts has not been extensively studied. Nevertheless, there have been attempts to explore this area, which are nicely summarized in some review articles (Sarria et al. 2017). For example, octanoic acid is biosynthesized by adding a C2 unit (from acetyl-CoA) to C6 hexanoic acid. Tan et al. reported that, using metabolically engineered *E. coli*, the fatty acid biosynthesis pathway can be manipulated to produce octanoic acid from glucose. Several engineering strategies, such as computational strain design, pathway gene regulation, medium optimization, were adopted. The highest production reached 1 g/L with greater than 70% selectivity using a fed-batch culture. In another report, it was found that the rapid, irreversible elongation of fatty acid acyl-ACP precursors is

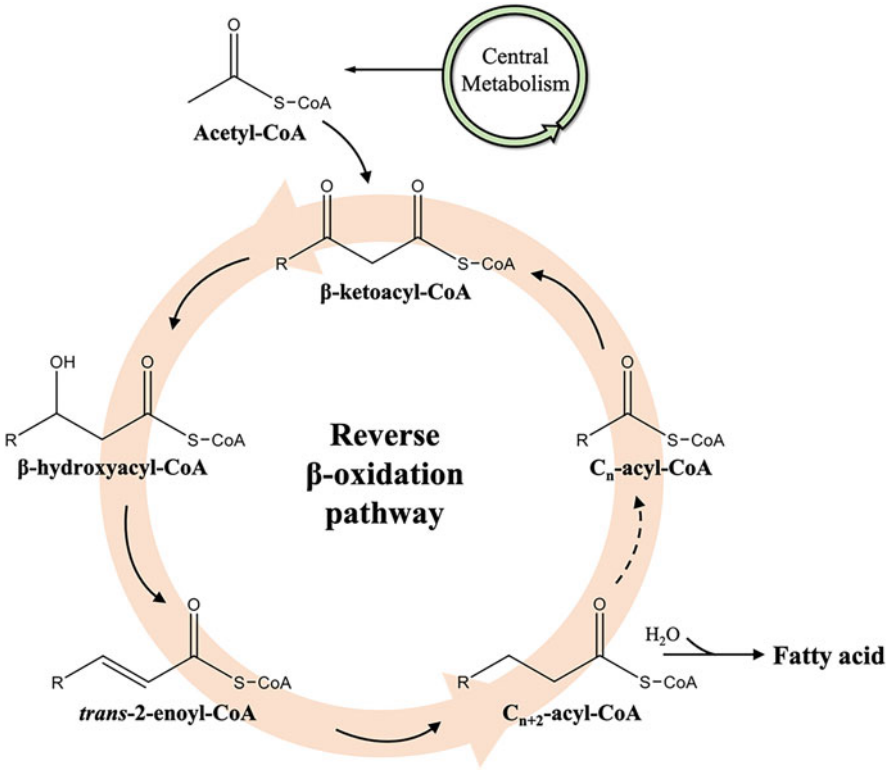


Fig. 3 Schematic of the reverse β-oxidation pathway for biosynthesis of straight-chain aliphatic carboxylic acids

a critical issue for medium-chain fatty acid biosynthesis. As such, a key pathway enzyme, ketoacyl synthase, was engineered for degradation using an exogenous inducer. This strategy ingeniously slowed down the chain elongation and improved the production of medium-chain fatty acid in engineered *E. coli*. Based on these efforts, the authors were able to achieve comprehensive production of all even- and odd-chain-length fatty acids with high efficiency (Torella et al. 2013). On the other hand, yeast, such as *S. cerevisiae* and *Y. lipolytica*, has also been recruited for octanoic acid biosynthesis by using the eukaryotic fatty acid biosynthesis enzymes (Leber and Da Silva 2014; Rutter et al. 2015; Rigouin et al. 2017).

Alternatively, biosynthesis can be achieved using the reverse β-oxidation pathway, as illustrated in Fig. 3. The pathway allows the use of acetyl-CoA as a direct substrate for chain elongation, bypassing the ATP-requiring conversion to malonyl-CoA in the regular fatty acid biosynthesis pathway. Similar to the fatty acid biosynthesis pathway, the chain length control for this pathway is challenging, and the biosynthesis efforts often lead to the production of acid mixtures. For example, Wu et al. utilized the reverse β-oxidation pathway to produce medium-chain fatty acids in *E. coli* (Wu et al. 2017). In particular, the rate-limiting steps of the pathway

were identified and addressed using metabolic engineering strategies. After the stepwise optimization, the production reached 3.8 g/L from glucose. Notably, the products contained a mixture of C6, C8, and C10 fatty acids, due to the limitation of the specificity of the pathway enzymes. The reverse β-oxidation pathway has also been established in *S. cerevisiae* for the production of medium-chain fatty acids (Lian and Zhao 2015).

It should be noted that there is a lack of successful studies for converting simple carbon materials to a specific medium-chain saturated fatty acid without byproduct acids. The separation of the desired acid products thus needs to overcome considerable technical challenges to isolate them from byproducts with similar chemical properties. In this sense, more research efforts should be dedicated with ingenious design and effective engineering approaches.

2.10.2 Medium-Chain Dicarboxylic Acids

Medium-chain dicarboxylic acids are widely used for the manufacturing of valuable industrial products, including detergents, surfactants, lubricants, perfumes, and pharmaceuticals. Moreover, the presence of the two carboxyl groups makes them outstanding materials for the production of various synthetic polymers, such as nylon, polyamides, and polyesters.

There are a few different pathways that can be used for microbial biosynthesis of medium-chain dicarboxylic acids. For the first pathway, medium-chain fatty acids can be hydroxylated at the ω position (the terminal carbon of the aliphatic chain) by ω-specific enzymes, generating ω-hydroxy fatty acids (Song et al. 2014). Subsequently, the terminal carbon is further oxidized by appropriate alcohol and aldehyde dehydrogenases to produce medium-chain dicarboxylic acids. This route can be coupled with the regular fatty acid biosynthesis pathway or the reverse β-oxidation pathway to enable the de novo bioproduction from simple carbon substrates such as glucose and glycerol. For example, Bowen et al. used a bioinformatics approach to identify key enzymes such as a ω-specific cytochrome P450 enzyme and selected alcohol and aldehyde dehydrogenases, which were then utilized to effectively oxidize medium-chain fatty acids at the terminal carbon to generate corresponding alcohol, aldehyde, and eventually dicarboxylic acids. This pathway was introduced into *E. coli* and integrated with the engineered fatty acid biosynthesis, which enabled the conversion of glucose to medium-chain α, ω-dicarboxylic acids (Bowen et al. 2016; Haushalter et al. 2017). The de novo production reached 600 mg/L after optimization of the cultivation conditions. Similarly, *S. cerevisiae* has been engineered to produce medium-chain α, ω-dicarboxylic acids through functional expression of a heterologous cytochrome P450 and cytochrome reductase. Coupled with the engineered fatty acid pathway, this approach resulted in the de novo dicarboxylic acid production from simple renewable sugars (Han et al. 2017).

Alternatively, some medium-chain dicarboxylic acids can be biosynthesized using non-decarboxylative Claisen condensation and β-reduction reactions, which have been established in *E. coli* (Cheong et al. 2016). To this end, selected

functionalized primers and functionalized extender units (e.g., analogs of fatty acid biosynthesis precursors) were combined to generate the desired α, ω-dicarboxylic acid products. By selection of the pathway enzymes with desired specificity and activities, suberic acid (C8) and sebacic acid (C10) has been successfully produced from glycerol. Notably, a similar strategy was also used for the adipic acid biosynthesis, as discussed in the previous section.

Straight-chain aliphatic dicarboxylic acids can also be produced by direct oxidation of alkanes using the engineered microbial alkane degradation pathway. For example, it has been found that a wide range of microbes is able to oxidize different alkanes to carboxylic acids (Wentzel et al. 2007). However, this pathway is promiscuous and can generate a variety of different products beside carboxylic acids. Also, the bioconversion efficiency is relatively low and hard to control. As such, this pathway is not popularly adopted for making straight-chain aliphatic dicarboxylic acids.

2.11 Long Straight-Chain Aliphatic Carboxylic Acids

2.11.1 Long-Chain Fatty Acids (>C12)

Long-chain fatty acids are essential constituent components of lipids and play an important role in cell membrane integrity. They are also critical substances for cellular energy storage. In addition, they carry other essential biological functions, such as gene expression regulation, protein modification, metabolic signaling, etc. Long-chain fatty acids, either saturated or unsaturated, can be used in various industries. For example, many of them have high nutraceutical values and thus have been widely used as food supplements.

Most long-chain fatty acids studied by the research communities are straight acids. As such, their biosynthesis can use the regular straight-chain fatty acid biosynthesis pathway (Fig. 2). Similar to the discussion above, due to the issue of chain elongation and termination control, the products are usually a mixture of products with different lengths. Recently, there have been a lot of new achievements in using various microbes as the host of biosynthesis. For example, *Yarrowia lipolytica*, a non-pathogenic native oleaginous yeast, has been widely utilized and engineered for fatty acid biosynthesis due to its outstanding native capability for lipid formation (Liu et al. 2021; Ledesma-Amaro and Nicaud 2016). In fact, a variety of engineering tools and strategies have been developed for this microbe, making it great a platform host for the biosynthesis of fatty acids and other related product biosynthesis. It should be also noted that bioproduction scale-up based on the use of *Y. lipolytica* is also straightforward, offering important benefits for industrially relevant production. It has been reported that, by blocking fatty acids' peroxisomal uptake and conversion to lipid, *Y. lipolytica* can produce 2.3 g/L free fatty acids in shake flask, and the long-chain fatty acid content was improved significantly (Ghogare et al. 2020). In another study, *Y. lipolytica's* glycerol metabolism was

eliminated to direct more flux toward the lipid pathway. When combined with other engineering strategies, 2 g/L free fatty acids, mostly C16-C24 fatty acids, were produced from glucose (Yuzbasheva et al. 2018). *Y. lipolytica* has also been used for the production of other unusual carboxylic acid products, such as polyunsaturated fatty acids (Gemperlein et al. 2019).

Long-chain fatty acid biosynthesis in *E. coli* has also been achieved (Lu et al. 2008). For example, by integrating the selected plant genes into *E. coli*'s fatty acid biosynthesis pathway, Kassab et al. established the production of long-chain fatty acids with yields greater than 200 mg/g dried cell weight (Kassab et al. 2019). Lu et al. reported the production of 2.5 g/L fatty acids (C12 and above) by engineering the native fatty acid biosynthesis and competing pathways and optimizing the resulting strain's cultivation conditions. Moreover, microalgae are another efficient workhorse for fatty acid biosynthesis. For example, through engineering glucose-6-phosphate, Xue et al. reprogrammed and improved the fatty acid bioproduction profile in oleaginous microalga *Phaeodactylum tricornutum* (Xue et al. 2017). Notably, microalgae can directly use photosynthesis to fix CO_2 and produce desired bioproducts, offering an intriguing route for renewable bioproduction. However, challenges, such as the development of genetic engineering tools and environmental concerns of using genetically modified organisms, need to be addressed (Blatti et al. 2013; Tang et al. 2020).

The reverse β-oxidation pathway is an alternative route for long-chain fatty acid bioproduction. In fact, it has been reported that, by controlling the expression of different thioesterases of this pathway, several long-chain fatty acids were produced in engineered *E. coli* (Dellomonaco et al. 2011). Further, individual functional units of the pathway were characterized in vitro and in vivo, which provided new insights for long-chain fatty acid biosynthesis.

2.11.2 Long-Chain Dicarboxylic Acids

Long-chain dicarboxylic acids are versatile raw materials used in cosmetic and plastic industries for the manufacturing of polymers, lubricants, perfumes, UV blocker in sun cream, and other products. Current production of long-chain dicarboxylic acids mostly relies on chemical synthesis. However, the formation of unwanted byproducts and multistep purification processes make the cost of production rise dramatically with the length of the carbon chain. Efficient biosynthesis provides a cost-effective manner for long-chain dicarboxylic acid production from renewable substrates.

Similar to their medium-chain counterparts, long-chain dicarboxylic acids can be produced by oxidation of long-chain alkanes, fatty acids, or fatty acid esters. To this end, alkane can be first oxidized to fatty alcohol by selected hydroxylase complexes. Fatty alcohol is further oxidized to fatty aldehyde by fatty alcohol oxidase, followed by the oxidation of the aldehyde group to carboxyl group at both ends of the carbon chain. The first step of the series of the oxidation reaction is also known as the rate-limiting step, as the corresponding enzyme is composed of hard-to-express Cytochrome P450 (CYP) monooxygenase and a NADPH-CYP reductase.

The conversion of fatty acid follows a similar pathway, although the oxidation only occurs at one end of the carbon chain. In an industrial yeast strain *Candida tropicalis*, four genes encoding isozymes of the acyl-CoA oxidase involved in the β–oxidation pathway were disrupted to direct more alkane and fatty acid substrates to enter the ω–oxidation pathway. Further overexpression of P450 and NADPH-cytochrome reductase genes of the pathway increased the productivity, and dicarboxylic acids ranging from C12 to C22 were biosynthesized with high conversion (80–100%) from dodecane or methyl myristate (Picataggio et al. 1992). Another study showed that knocking out the carnitine acetyltransferase (CAT) gene, which is responsible for transporting the acetyl-CoA into the mitochondrion for the TCA cycle, resulted in a 21% increase of dicarboxylic acids production in *Candida tropicalis* (Cao et al. 2006).

E. coli has also been engineered to produce long-chain α, ω-dicarboxylic acids from fatty acids by utilization of a CYP450-dependent ω-oxidation pathway (Sathesh-Prabu and Lee 2015). To reduce the damage to the cells by H_2O_2, an undesired byproduct of the terminal oxidation, thiourea was added to the cell culture. In addition, 5-aminolevulinic acid was supplemented as a heme precursor to enhance the activity of a hemi-dependent monooxygenase of the pathway. Based on these strategies, the production of 159 mg/L of C12 dicarboxylic acids and 410 mg/L of C14 dicarboxylic acids was achieved. For biosynthesis using fatty acid ester, Lee et al. genetically engineered yeast *Wickerhamiella sorbophila* to block the β-oxidation pathway and enabled the bioconversion of methyl laurate to dodecanedioic acid (Lee et al. 2018).

3 The Challenges and Future Directions

3.1 Product Toxicity

Despite the efforts of establishing effective biosynthesis pathways in selected microbial hosts, straight-chain aliphatic acid biosynthesis is often limited by the product toxicity (Warnecke and Gill 2005). In fact, for most of the compounds discussed above, high-level accumulation in the cell culture almost unavoidably leads to the inhibition of microbial growth and other metabolic activities. To address this challenge, many innovative strategies have been developed and implemented in recent years. For example, product secretion can largely help reduce the toxic products' intracellular accumulation. Zhu et al. metabolically engineered *S. cerevisiae* for medium-chain fatty acid biosynthesis and conducted directed evolution for a membrane transporter to enhance the secretion of medium-chain fatty acids (Zhu et al. 2020). In addition, strain adaptive laboratory evolution was also utilized to improve the robustness of *S. cerevisiae*, leading to the production of greater than 1 g/L extracellular medium-chain fatty acids.

Another popularly used strategy is in situ removal of the toxic compounds from the culture during the cultivation process (De Brabander et al. 2021). For example,

organic solvents can be added to the cell culture during the cultivation process to extract the toxic product from the aqueous phase, which also enriches the product in the organic phase and facilitates the downstream purification efforts. It has been reported that, using an alamine 336 and oleyl alcohol mixture as the extractant, hexanoic acid produced by *Megasphaera elsdenii* was effectively removed from the cell culture, which improved the cell growth and the bioproduction performance (Choi et al. 2013). Membrane-supported reactive extraction and resin adsorption have also been used to achieve the in situ removal of the toxic metabolic products and byproducts (Gössi et al. 2020; Ataei and Vasheghani-Farahani 2008).

Strain evolution is also a viable option for increasing the cell resistance to the product toxicity. Fletcher et al. found that *S. cerevisiae* could be adapted to low pH and high concentration of lactic acid by utilizing evolutionary paths to develop the tolerance (Fletcher et al. 2017). Notably, the strain adaption is largely dependent upon the cultivation conditions, such as carbon source, the product type, the recruited microbial species, etc.

3.2 Genetic Engineering Tools for Microbial Biosynthesis

Model microbes, such as *E. coli* and *S. cerevisiae*, have been widely used for the biosynthesis of straight-chain aliphatic carboxylic acids (Abbott et al. 2009; Liu and Jarboe 2012; Yu et al. 2011). Their rich engineering toolkits allow for advanced bioproduction manipulation. However, these microbes are only recruited for the biosynthesis of a portion of the carboxylic acid products. For many other microbes, reliable genetic engineering tools are still largely under-developed. As such, they are highly difficult to engineer for improving biosynthesis capabilities. This includes not only the alteration of the native pathways but also the introduction of heterologous enzymes from different organisms. One typical example is propionic acid producer *Propionibacterium*. Despite its wide application for propionic acid bioproduction, genetic engineering protocols for this microbe are still not well established.

This situation presents an outstanding challenge and calls for the development of sophisticated biotechniques. For example, clustered regularly interspaced short palindromic repeats (CRISPR) are expected to provide new and powerful tools to enable the engineering of the target microbes that are hard to engineer before. On the other hand, utilizing microbial consortia or co-cultures adds a new dimension to address the current challenge, as it extends the biosynthesis capability of one microbe by supplementing one or more biosynthesis partners (Brenner et al. 2008). Specifically, one hard-to-engineer microbe can utilize its native biosynthesis capability to generate a particular pathway intermediate/precursor, whereas another microbe(s) with established genetic tools can be used to convert the intermediate/precursor to the desired final product. This strategy enables biosynthesis that cannot be achieved using either strain alone.

3.3 Efficiency of the Biosynthesis Pathways

Microbial biosynthesis performance is largely dependent on the efficiency of the selected pathways. Straight-chain aliphatic carboxylic acid biosynthesis often involves both endogenous and heterologous enzymes of the recruited microbial hosts. As such, it is challenging to ensure coordinated expression and functions of all pathway enzymes. In fact, it is quite common that the biosynthesis efficiency is restricted by the low activities of one or few pathway enzymes or by distraction of metabolic flux into the competing pathways.

Resolution of these issues requires the adoption of new techniques and methods. For example, transcriptomics and metabolomics analyses will play an important role in identifying the limiting steps of the biosynthesis pathway. Consequently, protein engineering can be adopted to improve the bioconversion efficiency of a particular enzyme responsible for the limiting step. On the other hand, the metabolic flux distribution at the branching points between the desired and competing pathways can be controlled by manipulating the activities of the corresponding enzymes through directed enzyme evolution or the utilization of alternative enzymes from other organisms.

It should be noted that dynamic control of the pathway precursor supply and consumption is another robust approach for improving the overall biosynthesis efficiency. For instance, Xu et al. established a metabolic regulation network to control malonyl-CoA consumption and supply and significantly improved the fatty acid biosynthesis in *E. coli* (Xu et al. 2014). Also, modular co-culture engineering, which modularizes the biosynthesis pathway for expression in different strains or species in one culture, offers another perspective for pathway balancing and bioproduction enhancement (Zhang and Wang 2016). For this, there have been increasing efforts for the biosynthesis of various biomolecules using modular co-culture engineering, which lays a foundation for straight-chain aliphatic carboxylic acid production. Last but not least, large-scale cultivation using advanced bioreactor techniques can be utilized to improve the overall biosynthesis performance at the industrially relevant levels. This is especially important for establishing high substrate consumption, high cell density, and high product production processes. It is also a necessary step for the industrial application of any biosynthesis systems.

4 Concluding Remarks

Biosynthesis of straight-chain aliphatic carboxylic acids has been attracting significant research interest in recent years. This chapter highlights representative members of this group of biomolecules to demonstrate the potential of this vigorous field. Based on the above discussions, identifying the right microbe is a critical step for biosynthesis. In fact, many microbes natively have strong power for producing

certain straight-chain aliphatic carboxylic acids. These native producers have been extensively recruited in the past several decades, which has led to considerable application success. On the other hand, model organisms, such as *E. coli* and *S. cerevisiae*, can be genetically and metabolically engineered to accommodate heterologous pathways or a hybrid of native and heterologous pathways to achieve biosynthesis of the desired acid products.

Another key factor for successful biosynthesis is the establishment of proper pathways. In fact, the choice of feasible biosynthesis pathways is largely dependent upon the type of the end products. For the relatively small straight-chain acids, there are multiple independent pathways available for biosynthesis. For the relatively large acids, the choices of the feasible pathways are much limited. In fact, long-chain acids are almost entirely produced from the fatty acid biosynthesis pathway.

After the selection of the pathway, biosynthesis optimization needs to be pursued to meet the need of potential industrial production. To this end, enhancing the efficiency of biosynthesis pathways plays a critical role. Currently, many pathway enzymes (from same or different organisms) with desired functions can be identified from protein or genome databases and assembled in a selected microbial host for effective biosynthesis, which has resulted in tremendous success in the past. For future studies, the increasing availability of genome sequences of various organisms, fueled by DNA sequencing technologies, will make strong and positive impacts. Moreover, advances in protein engineering research, e.g., directed evolution, will promote the emergence of enzymes with needed diversity and catalytic activities.

On the other hand, balancing the distribution of metabolic resource between the target pathway, competing pathways, and the background metabolism, is essential for improving the biosynthesis performance. However, controlling such a balance is often times highly complicated, as suggested by previous studies (Jones et al. 2015). Therefore, innovative engineering strategies, such as biosensor-based dynamic control, need to be adopted to attract metabolic flux into the target pathway and reduce the byproduct formation. It should be noted that computational approaches, such as genome-scale modeling can also be utilized to facilitate biosynthesis system optimization. For example, there have been efforts for modeling the fatty acid biosynthesis behaviors, which provides important perspectives for engineering the global metabolism for supporting the desired biosynthesis (Youngquist et al. 2012). Collectively, innovations in various fields of science and technology will catalyze new and powerful development of straight-chain aliphatic carboxylic acid biosynthesis.

References

Abbott DA, Zelle RM, Pronk JT, Van Maris AJA (2009) Metabolic engineering of *Saccharomyces cerevisiae* for production of carboxylic acids: current status and challenges. FEMS Yeast Res 9: 1123–1136

Akawi L, Srirangan K, Liu X et al (2015) Engineering *Escherichia coli* for high-level production of propionate. J Ind Microbiol Biotechnol 42:1057–1072

Alissandratos A, Kim HK, Easton CJ (2014) Formate production through carbon dioxide hydrogenation with recombinant whole cell biocatalysts. Bioresour Technol 164:7–11
Ataei SA, Vasheghani-Farahani E (2008) *In situ* separation of lactic acid from fermentation broth using ion exchange resins. J Ind Microbiol Biotechnol 35:1229–1233
Beardslee T, Picataggio S (2012) Bio-based adipic acid from renewable oils. Lipid Technol 24:223–225
Blatti JL, Michaud J, Burkart MD (2013) Engineering fatty acid biosynthesis in microalgae for sustainable biodiesel. Curr Opin Chem Biol 17:496–505
Bowen CH, Bonin J, Kogler A et al (2016) Engineering *Escherichia coli* for conversion of glucose to medium-chain ω-hydroxy fatty acids and α,ω-dicarboxylic acids. ACS Synth Biol 5:200–206
Brenner K, You L, Arnold FH (2008) Engineering microbial consortia: a new frontier in synthetic biology. Trends Biotechnol 26:483–489
Cao Z, Gao H, Liu M, Jiao P (2006) Engineering the acetyl-CoA transportation system of *Candida tropicalis* enhances the production of dicarboxylic acid. Biotechnol J 1:68–74
Cavalcante WDA, Leitão RC, Gehring TA et al (2017) Anaerobic fermentation for n-caproic acid production: a review. Process Biochem 54:106–119
Cheong S, Clomburg JM, Gonzalez R (2016) Energy-and carbon-efficient synthesis of functionalized small molecules in bacteria using non-decarboxylative Claisen condensation reactions. Nat Biotechnol 34:556–561
Choe H, Joo JC, Cho DH et al (2014) Efficient CO2-reducing activity of NAD-dependent formate dehydrogenase from *Thiobacillus* sp. KNK65MA for formate production from CO_2 gas. PLoS One 9:e103111
Choi K, Jeon BS, Kim BC et al (2013) *In situ* biphasic extractive fermentation for hexanoic acid production from sucrose by *Megasphaera elsdenii* NCIMB 702410. Appl Biochem Biotechnol 171:1094–1107
De Brabander P, Uitterhaegen E, Verhoeven E et al (2021) *In situ* product recovery of bio-based industrial platform chemicals: a guideline to solvent selection. Fermentation 7:26
De Smit SM, De Leeuw KD, Buisman CJN, Strik DPBTB (2019) Continuous n-valerate formation from propionate and methanol in an anaerobic chain elongation open-culture bioreactor. Biotechnol Biofuels 12:132
Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R (2011) Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. Nature 476:355–359
Deng Y, Ma L, Mao Y (2016) Biological production of adipic acid from renewable substrates: current and future methods. Biochem Eng J 105:16–26
Deng Y, Mao Y (2015) Production of adipic acid by the native-occurring pathway in Thermobifida fusca B6. J Appl Microbiol 119:1057–1063
Deng Y, Mao Y, Zhang X (2015) Driving carbon flux through exogenous butyryl-CoA: acetate CoA-transferase to produce butyric acid at high titer in *Thermobifida fusca*. J Biotechnol 216: 151–157
Dhande YK, Xiong M, Zhang K (2012) Production of C5 carboxylic acids in engineered *Escherichia coli*. Process Biochem 47:1965–1971
Duncan SH, Barcenilla A, Stewart CS et al (2002) Acetate utilization and butyryl coenzyme A (CoA): acetate-CoA transferase in butyrate-producing bacteria from the human large intestine. Appl Environ Microbiol 68:5186–5190
Dwidar M, Park JY, Mitchell RJ, Sang BI (2012) The future of butyric acid in industry. Sci World J 2012:9
Eguchi SY, Nishio N, Nagai S (1985) Formic acid production from H2 and bicarbonate by a formateutilizing methanogen. Appl Microbiol Biotechnol 22:148–151
Eş I, Khaneghah AM, Hashemi SMB, Koubaa M (2017) Current advances in biological production of propionic acid. Biotechnol Lett 39:635–645
Fletcher E, Feizi A, Bisschops MMM et al (2017) Evolutionary engineering reveals divergent paths when yeast is adapted to different acidic environments. Metab Eng 39:19–28

Gemperlein K, Dietrich D, Kohlstedt M et al (2019) Polyunsaturated fatty acid production by *Yarrowia lipolytica* employing designed myxobacterial PUFA synthases. Nat Commun 10:4055

Ghogare R, Chen S, Xiong X (2020) Metabolic engineering of oleaginous yeast *Yarrowia lipolytica* for overproduction of fatty acids. Front Microbiol 11:1717

Gonzalez-Garcia RA, McCubbin T, Navone L et al (2017) Microbial propionic acid production. Fermentation 3:21

Gössi A, Burgener F, Kohler D et al (2020) *In-situ* recovery of carboxylic acids from fermentation broths through membrane supported reactive extraction using membrane modules with improved stability. Sep Purif Technol 241:116694

Grootscholten TIM, Steinbusch KJJ, Hamelers HVM, Buisman CJN (2013) High rate heptanoate production from propionate and ethanol using chain elongation. Bioresour Technol 136:715–718

Gullo M, Verzelloni E, Canonico M (2014) Aerobic submerged fermentation by acetic acid bacteria for vinegar production: process and biotechnological aspects. Process Biochem 49:1571–1579

Han L, Peng Y, Zhang Y et al (2017) Designing and creating a synthetic omega oxidation pathway in *Saccharomyces cerevisiae* enables production of medium-chain α, ω-dicarboxylic acids. Front Microbiol 8:2184

Haushalter RW, Phelan RM, Hoh KM et al (2017) Production of odd-carbon dicarboxylic acids in *Escherichia coli* using an engineered biotin-fatty acid biosynthetic pathway. J Am Chem Soc 139:4615–4618

Heil CS, Wehrheim SS, Paithankar KS, Grininger M (2019) Fatty acid biosynthesis: chain-length regulation and control. Chembiochem 20:2298–2321

Jawed K, Mattam AJ, Fatma Z et al (2016) Engineered production of short chain fatty acid in *Escherichia coli* using fatty acid synthesis pathway. PLoS One 11:e0160035

Jeon BS, Choi O, Um Y, Sang BI (2016) Production of medium-chain carboxylic acids by *Megasphaera* sp. MH with supplemental electron acceptors. Biotechnol Biofuels 9:129

Jiang L, Fu H, Yang HK et al (2018) Butyric acid: applications and recent advances in its bioproduction. Biotechnol Adv 36:2101–2117

Jones JA, Toparlak TD, Koffas MAG (2015) Metabolic pathway balancing and its role in the production of biofuels and chemicals. Curr Opin Biotechnol 33:52–59

Kandasamy V, Vaidyanathan H, Djurdjevic I et al (2013) Engineering *Escherichia coli* with acrylate pathway genes for propionic acid synthesis and its impact on mixed-acid fermentation. Appl Microbiol Biotechnol 97:1191–1200

Karekar SC, Srinivas K, Ahring BK (2020) Continuous *in-situ* extraction of acetic acid produced by *Acetobacterium woodii* during fermentation of hydrogen and carbon dioxide using Amberlite FPA53 ion exchange resins. Bioresour Technol Rep 12:100568

Kassab E, Fuchs M, Haack M et al (2019) Engineering *Escherichia coli* FAB system using synthetic plant genes for the production of long chain fatty acids. Microb Cell Factories 18:163

Kruyer NS, Peralta-Yahya P (2017) Metabolic engineering strategies to bio-adipic acid production. Curr Opin Biotechnol 45:136–143

Lang K, Zierow J, Buehler K, Schmid A (2014) Metabolic engineering of *Pseudomonas* sp. strain VLB120 as platform biocatalyst for the production of isobutyric acid and other secondary metabolites. Microb Cell Factories 13:2

Le QAT, Kim HG, Kim YH (2018) Electrochemical synthesis of formic acid from CO2 catalyzed by *Shewanella oneidensis* MR-1 whole-cell biocatalyst. Enzym Microb Technol 116:1–5

Leber C, Da Silva NA (2014) Engineering of *Saccharomyces cerevisiae* for the synthesis of short chain fatty acids. Biotechnol Bioeng 111:347–358

Ledesma-Amaro R, Nicaud JM (2016) *Yarrowia lipolytica* as a biotechnological chassis to produce usual and unusual fatty acids. Prog Lipid Res 61:40–50

Lee H, Han C, Lee HW et al (2018) Development of a promising microbial platform for the production of dicarboxylic acids from biorenewable resources. Biotechnol Biofuels 11:310

Lian J, Zhao H (2015) Reversal of the β-oxidation cycle in *Saccharomyces cerevisiae* for production of fuels and chemicals. ACS Synth Biol 4:332–341

Liu H, Song Y, Fan X et al (2021) *Yarrowia lipolytica* as an oleaginous platform for the production of value-added fatty acid-based bioproducts. Front Microbiol 11:608662
Liu P, Jarboe LR (2012) Metabolic engineering of biocatalysts for carboxylic acids production. Comput Struct Biotechnol J 3:e201210011
Lu X, Vora H, Khosla C (2008) Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. Metab Eng 10:333–339
Nakano S, Fukaya M, Horinouchi S (2004) Enhanced expression of aconitase raises acetic acid resistance in *Acetobacter aceti*. FEMS Microbiol Lett 235:315–322
Picataggio S, Rohrer T, Deanda K et al (1992) Metabolic engineering of *Candida tropicalis* for the production of long-chain dicarboxylic acids. Nat Biotechnol 10:893–899
Pituch A, Walkowiak J, Banaszkiewicz A (2013) Butyric acid in functional constipation. Prz Gastroenterol 8:295–298
Ragsdale SW, Pierce E (2008) Acetogenesis and the Wood-Ljungdahl pathway of CO_2 fixation. Biochim Biophys Acta Proteins Proteomics 1784:1873–1898
Ranaei V, Pilevar Z, Khaneghah AM, Hosseini H (2020) Propionic acid: method of production, current state and perspectives. Food Technol Biotechnol 58:115–127
Richter H, Qureshi N, Heger S et al (2012) Prolonged conversion of n-butyrate to n-butanol with *Clostridium saccharoperbutylacetonicum* in a two-stage continuous culture with *in-situ* product removal. Biotechnol Bioeng 109:913–921
Rigouin C, Gueroult M, Croux C et al (2017) Production of medium chain fatty acids by *Yarrowia lipolytica*: combining molecular design and TALEN to engineer the fatty acid synthase. ACS Synth Biol 6:1870–1879
Rutter CD, Zhang S, Rao CV (2015) Engineering *Yarrowia lipolytica* for production of medium-chain fatty acids. Appl Microbiol Biotechnol 99:7359–7368
Saini M, Wang ZW, Chiang CJ, Chao YP (2014) Metabolic engineering of *Escherichia coli* for production of butyric acid. J Agric Food Chem 62:4342–4348
Sarria S, Kruyer NS, Peralta-Yahya P (2017) Microbial synthesis of medium-chain chemicals from renewables. Nat Biotechnol 35:1158–1166
Sathesh-Prabu C, Lee SK (2015) Production of long-chain α,ω-dicarboxylic acids by engineered *Escherichia coli* from renewable fatty acids and plant oils. J Agric Food Chem 63:8199–8208
Schuchmann K, Müller V (2013) Direct and reversible hydrogenation of CO_2 to formate by a bacterial carbon dioxide reductase. Science 342:1382–1385
Song H, Eom MH, Lee S et al (2010) Modeling of batch experimental kinetics and application to fed-batch fermentation of *clostridium tyrobutyricum* for enhanced butyric acid production. Biochem Eng J 53:71–76
Song JW, Lee JH, Bornscheuer UT, Park JB (2014) Microbial synthesis of medium-chain α,-ω-dicarboxylic acids and ω-aminocarboxylic acids from renewable long-chain fatty acids. Adv Synth Catal 356:1782–1788
Song Y, Lee JS, Shin J et al (2020) Functional cooperation of the glycine synthasereductase and Wood-Ljungdahl pathways for autotrophic growth of *Clostridium drakei*. Proc Natl Acad Sci U S A 117:7516–7523
Tang DYY, Yew GY, Koyande AK et al (2020) Green technology for the industrial production of biofuels and bioproducts from microalgae: a review. Environ Chem Lett 18:1967–1985
Torella JP, Ford TJ, Kim SN et al (2013) Tailored fatty acid synthesis via dynamic control of fatty acid elongation. Proc Natl Acad Sci U S A 110:11290–11295
Tseng HC, Prather KLJ (2012) Controlled biosynthesis of odd-chain fuels and chemicals via engineered modular metabolic pathways. Proc Natl Acad Sci U S A 109:17925–17930
Vidra A, Németh Á (2018) Bio-produced acetic acid: a review. Period Polytech Chem Eng 62:245–256
Wang J, Lin M, Xu M, Yang ST (2016) Anaerobic fermentation for production of carboxylic acids as bulk chemicals from renewable biomass. In: Advances in Biochemical Engineering/Biotechnology. pp 323–363

Warnecke T, Gill RT (2005) Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. Microb Cell Factories 4:25

Wentzel A, Ellingsen TE, Kotlar HK et al (2007) Bacterial metabolism of long-chain n-alkanes. Appl Microbiol Biotechnol 76:1209–1221

Werpy T, Petersen G (2004) Top value added chemicals from biomass: volume I—Results of screening for potential candidates from sugars and synthesis gas. Office of Scientific and Technical Information (OSTI). Off Sci Tech Inf 69

Wu J, Zhang X, Xia X, Dong M (2017) A systematic optimization of medium chain fatty acid biosynthesis via the reverse beta-oxidation cycle in *Escherichia coli*. Metab Eng 41:115–124

Xiong M, Yu P, Wang J, Zhang K (2015) Improving engineered *Escherichia coli* strains for high-level biosynthesis of isobutyrate. AIMS Bioeng 2:60–74

Xu P, Li L, Zhang F et al (2014) Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. Proc Natl Acad Sci U S A 111:11299–11304

Xue J, Balamurugan S, Li DW et al (2017) Glucose-6-phosphate dehydrogenase as a target for highly efficient fatty acid biosynthesis in microalgae by enhancing NADPH supply. Metab Eng 41:212–221

Youngquist JT, Lennen RM, Ranatunga DR et al (2012) Kinetic modeling of free fatty acid production in *Escherichia coli* based on continuous cultivation of a plasmid free strain. Biotechnol Bioeng 109:1518–1527

Yu C, Cao Y, Zou H, Xian M (2011) Metabolic engineering of *Escherichia coli* for biotechnological production of high-value organic acids and alcohols. Appl Microbiol Biotechnol 89:573–583

Yuzbasheva EY, Mostova EB, Andreeva NI et al (2018) A metabolic engineering strategy for producing free fatty acids by the *Yarrowia lipolytica* yeast based on impairment of glycerol metabolism. Biotechnol Bioeng 115:433–443

Zhang A, Yang ST (2009) Engineering *Propionibacterium acidipropionici* for enhanced propionic acid tolerance and fermentation. Biotechnol Bioeng 104:766–773

Zhang C, Yang H, Yang F, Ma Y (2009) Current progress on butyric acid production by fermentation. Curr Microbiol 59:656–663

Zhang H, Wang X (2016) Modular co-culture engineering, a new approach for metabolic engineering. Metab Eng 37:114–121

Zhang K, Woodruff AP, Xiong M et al (2011) A synthetic metabolic pathway for production of the platform chemical isobutyric acid. ChemSusChem 4:1068–1070

Zhu X, Tao Y, Liang C et al (2015) The synthesis of n-caproate from lactate: a new efficient process for medium-chain carboxylates production. Sci Rep 5:14360

Zhu Y, Yang ST (2004) Effect of pH on metabolic pathway shift in fermentation of xylose by *Clostridium tyrobutyricum*. J Biotechnol 110:143–157

Zhu Z, Hu Y, Teixeira PG et al (2020) Multidimensional engineering of *Saccharomyces cerevisiae* for efficient synthesis of medium-chain fatty acids. Nat Catal 3:64–74

Microbial Production of Amines and Amino Acids by Fermentation

Volker F. Wendisch and Anastasia Kerbs

Contents

Abstract Amino acids and derived amines are widely produced and utilized industrially. Fermentation is the preferred route for amino acid production. Established fermentation processes are characterized by economies of scale. Therefore, producing strains are constantly improved with regard to titer, yield, and productivity as key performance indicators embracing the newest metabolic engineering technologies.

V. F. Wendisch (✉) · A. Kerbs
Genetics of Prokaryotes, Faculty of Biology and CeBiTec, Bielefeld University, Bielefeld, Germany
e-mail: volker.wendisch@uni-bielefeld.de

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_3

Furthermore, metabolic engineering enabled the production of non-proteinogenic amino acids, omega-amino acids, functionalized amino acids, and diamines. In this chapter, we will focus on strain development for the fermentative production of these nitrogenous compounds by bacteria.

1 Introduction

Amines contain at least one ammonium group and amino acids in addition at least one carboxyl substituent. In this chapter, we will cover the fermentative production of amines and amino acids by bacteria. The reader is referred to other chapters and reviews regarding the production of peptides, proteins, and other nitrogen-containing polymers as well as with respect to eukaryotic production hosts. The primary focus will be on strain development. Regarding process intensification, the reader is referred to other book chapters and reviews.

Amino acids are relevant in many markets. The food and feed amino acids such as L-glutamate and L-lysine are produced at the annual million-ton scale to meet the demand for human and animal nutrition including the expanding aquaculture sector. Amino acids are also precursors of active pharmaceutical ingredients such as L-dihydroxyphenylalanine (L-DOPA). The amino acid market is growing by about 5% (predicted compound annual growth rate for the period 2021–2026). In 2026, the market for animal feed amino acids will be worth about US$15.8 billion although growth was affected by COVID-19 and African swine fever (Global Amino Acids Market 2021).

Amino acids are synthesized chemically (e.g., glycine which lacks a stereocenter), obtained by extraction (e.g., L-cysteine from human hair), or by enzyme catalysis (e.g., L-aspartate from fumarate), but microbial fermentation is the dominating way of production, primarily using *Corynebacterium* or *Escherichia coli* strains (Wendisch 2020). Fermentation leads to the L-stereoisomer that can be used by animals and humans with highest efficiency (almost only L-lysine can be used, and eventhough methionine accepting D-amino acid oxidase and L-specific transaminase exist, pure L-methionine has 25% higher nutritional potency than the chemically produced racemate) (Koga et al. 2017). While currently methionine is produced as racemate primarily, a two-step fermentation plus biotransformation process yielding L-methionine has been commercialized by CJ and is operated at an annual capacity of 90,000 tons (Shim et al. 2017). In addition, a fermentation process directly yielding L-methionine using an *E. coli* strain has been developed by Metabolic Explorer and was acquired by Evonik (Evonik Industries 2016a). The annual global production of L-methionine and of other amino acids that are produced by fermentation is given in Table 1.

Table 1 Amino acids produced by fermentation and annual global production. Adapted and updated from (Global Amino Acids Market 2021; Wendisch 2020; Evonik Industries 2016a; Evonik Industries 2016b; Nikkei 2016; Wendisch et al. 2016a; Ajinomoto Co., Inc. 2021). Values for annual global production may include production by other means than fermentation (extraction, enzymatic, chemical)

Amino acid	Annual global production (metric tons)
L-arginine	1200
L-glutamate	3,240,000
L-glutamine	1300
L-histidine	400
L-isoleucine	400
L-leucine	500
L-lysine	2,600,000
L-methionine	90,000
L-phenylalanine	12,650
L-proline	350
L-serine	350
L-threonine	700,000
L-tryptophan	41,000
L-tyrosine	165
L-valine	500

2 New Technological Developments for Production of Proteinogenic Amino Acids

Fermentative production of bulk products such as the food and feed amino acids seeks to constantly improve titers, yields, and productivities. Thus, the development of amino acid overproducing strains was and is at the forefront of embracing and developing new metabolic engineering technologies (Wendisch 2020; Wendisch et al. 2006; Wendisch et al. 2016b; Lee and Wendisch 2017; Wolf et al. 2021). In synthetic and systems metabolic engineering of amino acid producers strategies involving the CRISPR systems (Schultenkämper et al. 2020; Wang et al. 2021), flux enforcement (Fig. 1) (von Kamp and Klamt 2017; Prell et al. 2021; Haupka et al. 2021), adaptive laboratory evolution (ALE) (Stella et al. 2019; Hennig et al. 2020; Graf et al. 2019; Yu et al. 2021), biosensors for screening, driving mutation or switching gene expression on demand (Jurischka et al. 2020; Tan et al. 2020; Henke et al. 2020; Della Corte et al. 2020), microfluidic droplet-based screening (Balasubramanian et al. 2021), synthetic consortia (Sgobba et al. 2018; Burmeister et al. 2018; Vortmann et al. 2021; Sgobba and Wendisch 2020; Burmeister et al. 2021), and strategies for coproduction of secreted and cell-bound target products (Henke et al. 2018; Ma et al. 2018) have been developed, tested, characterized, and applied.

A recent example of strain engineering may have relevance for new production processes, i.e. non-sterile L-lysine fermentation (Hirota et al. 2017). The enzyme PtxD was shown to oxidize phosphite to phosphate (Costas et al. 2001) and used as the dominant selection marker in yeast when phosphite was added to the growth medium and cells could only grow when PtxD was synthesized (Kanda et al. 2014). PtxD was used for regeneration of redox cofactor, e.g., for mannitol production

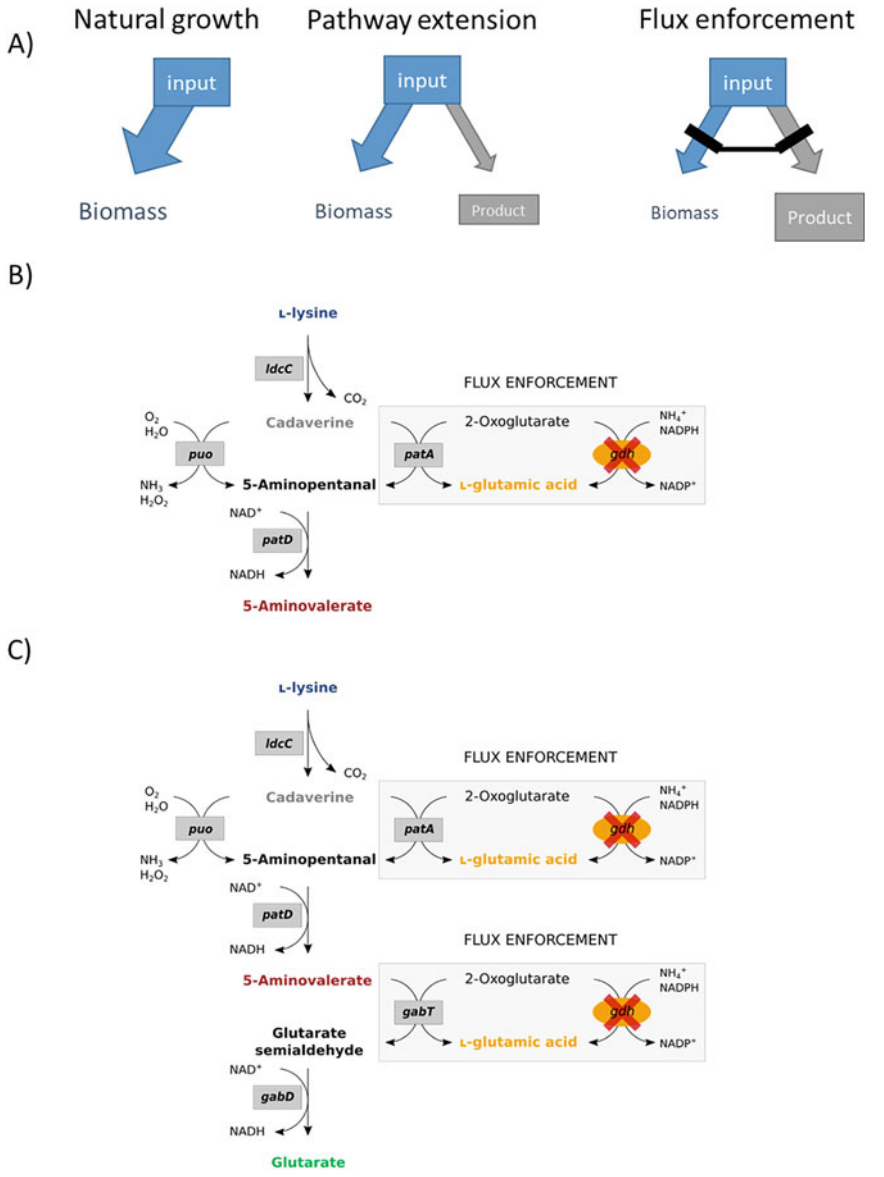


Fig. 1 Flux enforcement coupling glutamate generating transaminase(s) in production pathways to growth in strains lacking L-glutamic acid dehydrogenase. Flux enforcement couples the activity of a production pathway to growth (**a**). In production of 5-aminovalerate from lysine, PuO deaminates cadaverine using molecular oxygen without concomitant formation of L-glutamic acid (**b**, left), while PatA catalyzes a cadaverine dependent transamination reaction converting 2-oxoglutarate to L-glutamic acid (1 coupling site) (**b**, right). Transaminase GabT is involved in conversion of 5-aminovalerate to glutarate, thus, when the PuO pathway is extended to glutarate 1 coupling site for concomitant formation of L-glutamic acid by flux enforcement is available (**c**, left). When the PatA pathway is extended to glutarate, two transamination reactions (PatA and GabT) provide 2 coupling sites for concomitant formation of L-glutamic acid by flux enforcement is available (**c**, right)

(Reshamwala et al. 2014) or enzyme catalysis by Baeyer–Villiger monooxygenases (Rioz-Martínez et al. 2011). Later, PtxD was used to support growth with phosphite as the only source of phosphorus as shown for yeasts, *E. coli,* cyanobacteria*, Bacillus subtilis,* and *C. glutamicum* (Shaw et al. 2016; Ou et al. 2019; Motomura et al. 2018; Guo et al. 2020; Lei et al. 2021; Schwardmann et al. 2022). Under non-sterile conditions, using phosphite and PtxD from *Pseudomonas stutzeri* expressed instead of the chromosomal gene *exeR*, an L-lysine producing strain produced about 41 g L^{-1} L-lysine during batch fermentation for 60 h (Lei et al. 2021). This approach allowed to counter-select contaminants in a synthetic co-culture and flasks handled under non-sterile conditions; however, it remains to be seen if the developed strain would allow for stabile L-lysine production under non-sterile conditions in large bioreactors or even in open ponds.

3 Non-Proteinogenic Amino Acids

Production of the bulk amino acids L-glutamate and L-lysine occurs at the million-tons-scale; however, their low margins shifted the focus of major producers toward more profitable amino acids including non-proteinogenic and derivatized amino acids (Wendisch 2020). While proteinogenic amino acids serve the food and feed industries, non-proteinogenic amino acids are relevant for different markets. For example, the omega-amino acid 5-aminovalerate (5AVA) is a monomeric precursor for the plastic polyamide PA5, whereas the cyclic amino acid L-pipecolic acid is a chiral building block for the synthesis of therapeutics, and halogenated amino acids such as 7-Br-L-tryptophan are constituents of active pharmaceutical ingredients with antimicrobial or anticancer activities. Due to its favorable ecological footprint, we focus on the fermentative production of these specialty amino acids.

3.1 γ-Aminobutyrate (GABA)

GABA can be synthesized by two routes, either by decarboxylation of L-glutamate or in a three-step cascade from L-ornithine via putrescine (Fig. 2). GABA finds application due to its blood pressure-lowering activity, in functional foods and as a precursor for polyamide PA4. *C. glutamicum* with its known capacity of L-glutamate production is a suitable host once the genes for utilization of GABA as sole carbon and nitrogen source (*gabTDP* controlled by GabR) were deleted (Zhu et al. 2020). Since this bacterium lacks glutamate decarboxylase activity, heterologous expression of the L-glutamate decarboxylase genes *gadA* or *gadB* from *E. coli* was required and 8 g L^{-1} GABA was produced within 96 h (Takahashi et al. 2012). Upon expression of a gene encoding an L-glutamate decarboxylase variant with a broader pH optimum under control of the strong synthetic promoter H36 a product titer of about 38 g L^{-1} was reached (Choi et al. 2015). Recently, a CRISPR-Cas approach

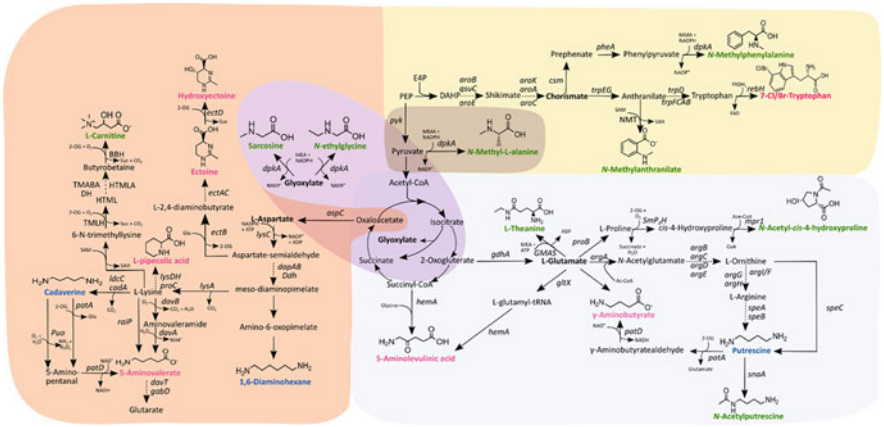


Fig. 2 Schematic overview of the synthesis of alkylated amino acids (green), non-proteinogenic amino acids (pink), and diamines (blue) by microbial fermentation. Metabolites derived from the amino acid L-aspartate are highlighted in orange, L-glutamate derived metabolites are highlighted in light blue. Synthesis of alkylated amino acids with glyoxylate as precursor is marked in purple and synthesis of modified aromatic amino acids is highlighted in yellow. Pyruvate derived synthesis of *N*-methyl-L-alanine is marked in brown. *2-OG* 2-oxoglutarate, *argA* amino acid *N*-acetyltransferase, *argB* N-acetyl-L-glutamate kinase, *argC* *N*-acetyl-γ-glutamyl-phosphate reductase, *argD* *N*-acetyl-L-ornithine aminotransferase, *argE* *N*-acetyl-L-ornithine deacetylase, *argF* L-ornithine carbamoyl transferase, *argH* argininosuccinate lyase, *argI* L-ornithine carbamoyl transferase, *aroA* 3-Phosphoshikimate 1-carboxyvinyltrnsferase, *aroB* 3-Dehydroquinate synthase, *aroC* chorismate synthase, *aroE* shikimate dehydrogenase, *aroK* shikimate kinase, *aspC* aspartate transaminase, *BBH* γ-butyrobetaine hydroxylase, *cadA*L-lysine decarboxylase, *csm* chorismate mutase, *dapA* dihydrodipicolinic acid synthase, *dapB* 4-hydroxy-tetrahydrodipicolinate reductase, *davA*L-lysine monooxygenase, *davB* γ-aminovaleramidase, *davT* 5-aminovalerate amino transferase, *Ddh* m-diaminopimelate dehydrogenase, *dpkA* *N*-alkylamino acid dehydrogenase, *E4P* Erythrose 4-phosphate, *ectA*L-2,4-diaminobutyrate acetyltransferase, *ectB*L-2,4-diaminobutyrate aminotransferase, *ectC* ectoine synthase, *ectD* ectoine hydroxylase, *gabD* succinate-semialdehyde dehydrogenase, *gdhA* glutamate dehydrogenase, *gltX* glutamate-tRNA ligase, *GMAS* γ-glutamylmethylamide synthetase, *hemA* 5-aminolevulinate synthase, *HTMLA* 3-hydroxy-L-N^{ε}-trimethyllysine aldolase, *ldcC* lysine decarboxylase, *lysA* diaminopimelate decarboxylase, *lysDH* lysine dehydrogenase, *MEA* monoethylamine, *MMA* monomethylamine, *mpr1* *N*-acetyltransferase, *NMT* N-methyltransferase, *patA* N-acetyltransferase, *patD* γ-aminobutyraldehyde dehydrogenase, *PEP* phosphoenolpyruvate, *pheA* chorismate mutase/prephenate dehydratase, *proB* γ-glutamyl kinase, *proC* pyrroline-5-carboxylate reductase, *puO* putrescine oxidase, *pyk* pyruvate kinase, *qsuC* 3-dehydroquinate dehydratase, *rebH* halogenase, *SAH* *S*-adenosylhomocysteine, *SAM* *S*-adenosylmethionine, *SmP$_4$H*L-proline cis-4-hydroxylase, *snaA* *N*-acetyltransferase, *speA* arginine decarboxylase, *speB* agmatinase, *speC* ornithine decarboxylase, *TMABA* 4-trimethylaminobutyraldehyde, *TMLH*L-N^{ε}-trimethyllysine hydroxylase, *trpAB* tryptophan synthase, *trpC* indole-glycerol phosphate synthase, *trpD* anthranilate phosphoribosyl transferase, *trpEG* anthranilate synthase, *trpF* phosphoribosyl anthranilate isomerase

was chosen to delete three endogenous genes (*yggB* for L-glutamate export, *gabD,* and *gabT*) and upon plasmid-borne expression of the heterologous *gadB2* gene from *Lactobacillus brevis* ATCC 367 about 28 g L^{-1} GABA was produced (Cho et al. 2017). GABA production via the three-step cascade comprising

L-ornithine cyclodeaminase Ocd from *P. putida* and the *E. coli* enzymes PatA (putrescine transaminase) and PatD (γ-aminobutyraldehyde dehydrogenase) in an L-ornithine overproducing *C. glutamicum* strain led to a product titer of about 5 g L^{-1} (Jorge et al. 2016).

3.2 5-Aminovalerate (5AVA)

The omega-amino acid 5-AVA is the monomeric precursor of polyamide PA5. Four different biochemical pathways have been used for fermentative 5AVA production from endogenous L-lysine (Fig. 2). Route 1 consists of oxidative deamination of L-lysine by molecular oxygen-dependent L-lysine-α-oxidase (RaiP) from *Scomber japonicus* followed by a spontaneous decarboxylation reaction (Cheng et al. 2020). Route 2 also requires molecular oxygen for oxidative decarboxylation catalyzed by L-lysine monooxygenase (DavA) from *P. putida,* while the subsequent deamidation by γ-aminovaleramidase (DavB) from *P. putida* only requires water (Adkins et al. 2013). Route 3 combines L-lysine decarboxylase from *E. coli* with molecular oxygen-dependent putrescine oxidase PuO from *Rhodococcus qingshengii* and NAD-dependent γ-aminobutyraldehyde dehydrogenase PatD from *E. coli* (Fig. 1) (Haupka et al. 2020). Route 4 is not dependent on molecular oxygen. This cascade combines L-lysine decarboxylase, 2-oxoglurate-dependent putrescine/cadaverine transaminase PatA, and NAD-dependent γ-aminobutyraldehyde dehydrogenase PatD from *E. coli* (Fig. 1) (Jorge et al. 2017). Since route 3 and route 4 only differ by PuO catalyzing oxidative desamination or PatA catalyzing transamination yielding L-glutamate, the effect of flux enforcement could be studied (Fig. 1). Flux enforcement by deletion of *gdh*, the gene for the major ammonium assimilating enzyme L-glutamic acid dehydrogenase (Haupka et al. 2020; Pérez-García et al. 2018) results in a metabolic setup in which growth requires the production of 5AVA. The enzyme pair glutamine synthetase (GS) and L-glutamic acid-2-oxoglutarate aminotransferase (GOGAT, also known as L-glutamic acid synthase) synthesize L-glutamic acid in an ATP-dependent manner during ammonium starvation at ammonium concentrations below 5 mM, thus, it cannot fully compensate for the lack of L-glutamic acid dehydrogenase (Kholy et al. 1993; Tesch et al. 1998).

3.3 Halogenated Tryptophan

About one-third of all agrochemicals are halogenated. Halogenated amino acids are among the more than 5000 naturally occurring halogenated compounds (Schnepel and Sewald 2017), e.g., in bioactives such as the antibiotics chloramphenicol and pyrroindomycin, the plant growth-regulating thienodolin, and the antifungal pyrrolnitrin (Hammer et al. 1997; van Pée and Hölzer 1999). Unlike chemical

synthesis, biosynthetic pathways do not lead to unwanted by-products due to excellent regioselectivity and they do not require hazardous conditions since halogenases use halide ions and molecular oxygen. The halogenase RebH from *Lechevalieria aerocolonigenes* is involved in biosynthesis of the DNA topoisomerase I inhibitor rebeccamycin (Nishizawa et al. 2005; Onaka et al. 2003). FADH-dependent RebH works in concert with NADH-dependent flavin reductase RebF (Nishizawa et al. 2005) with regard to recycling the reduced cofactor $FADH_2$. Cross-linked enzyme aggregates including RebH, RebF, and an alcohol dehydrogenase for isobutanol-driven NADH recycling catalyzed chlorination of L-tryptophan to 7-Cl-L-tryptophan at the gram scale (Frese and Sewald 2015; Schnepel et al. 2016). Extension of L-tryptophan biosynthesis by RebH and RebF in an L-tryptophan overproducing *C glutamicum* strain enabled fermentative production of 7-Cl-tryptophan from simple carbon and nitrogen sources in the presence of chloride ions (Veldmann et al. 2019a). Replacing chloride ions in the culture medium by bromide ions allowed for de novo production of 7-Br-tryptophan (Veldmann et al. 2019b) (Fig. 2). Halogenases with different regioselectivities such as ThaI from *Streptomyces albogriseolus* and engineered enzyme variants (Moritzer et al. 2019; Neubauer et al. 2018), fermentative production of tryptophans chlorinated or brominated at the 4, 5, 6, and 7 positions will be possible in the future.

3.4 L-Ornithine

L-Ornithine finds pharmaceutical applications due to its beneficial effects on the liver and the heart (Wu et al. 2019). L-Ornithine is an intermediate of L-arginine biosynthesis and this basic amino acid can replace basic arginine or lysine in cyanophycin, a polymer found in cyanobacteria (Wördemann et al. 2021). Since L-ornithine is derived from L-glutamate, overproducing *C. glutamicum* strains represent a good starting point for metabolic engineering. For example, such an L-glutamate producer was converted to a strain overproducing L-ornithine from mannitol (Sheng et al. 2021). The utilization of mannitol was improved based on knowledge on the regulation of its catabolic operon (Laslo et al. 2012) and on bottlenecks of fructose utilization (Kiefer et al. 2002; Georgi et al. 2005). Systems metabolic engineering comprised increased glycolytic and pentose phosphate flux by gene overexpression, avoiding loss of the intermediates acetyl-CoA and 2-oxoglutarate by deletion of genes for their conversion, transport engineering to avoid loss of the amino acids L-glutamate,L-valine, L-leucine, and L-isoleucine, and interrupting biosynthesis of L-proline and L-arginine (Zhang et al. 2019). This strain produced about 44 g L^{-1} from glucose, about 19 g L^{-1} from xylose, and about 54 g L^{-1} L-ornithine from mannitol (Sheng et al. 2021). Because L-ornithine is a precursor of L-citrulline and L-arginine and can be converted to L-proline by cyclodeaminase (Jensen and Wendisch 2013) and to putrescine by L-ornithine decarboxylase (Schneider and Wendisch 2010; Schneider et al. 2012) a chassis concept for the production of these target compounds was established (Jensen et al. 2015).

3.5 ε-Aminolevulinic Acid (5-ALA)

5-ALA is a non-proteinogenic amino acid and an intermediate in tetrapyrrole biosyntheses such as heme, chlorophyll, or vitamin B12, thus, it finds applications as a biostimulant for crop production and as a feed additive. *C. glutamicum* was engineered to produce 7.5 g L^{-1} 5-ALA by overexpression of *hemA* from *R. sphaeroides* and *rhtA* from *E. coli*, while precursor supply was optimized via deletion of *ppc, ldhA, pqo, cat, pta, ackA,* and *pbp1b* (Feng et al. 2016). Further strain development involving deletion of *succCD* and balanced expression of heterologous *hemA* and native *ppc* allowed to produce 18.5 g L^{-1} 5-ALA using enzymatically hydrolyzed cassava bagasse as feedstock (Chen et al. 2020). A similarly engineered *E. coli* strain produced about 7 g L^{-1} 5-ALA when CRISPRi was used to repress *hemB* expression, improving extracellular 5-ALA accumulation (Miscevic et al. 2021). The pathway has been extended for heme production. Although only titers below 1 g L^{-1} were obtained, this type of animal-free heme production may cater to vegan and vegetarian customers (Ko et al. 2021).

3.6 L-Pipecolic Acid

L-Pipecolic acid is a cell-protecting compatible solute (Pérez-García et al. 2019) also used as a chiral building block in the pharma industry (Tsuge and Matsuzawa 2021). This cyclic, non-proteinogenic amino acid can be synthesized from L-lysine. Thus, a synthetic pathway comprising oxidative deamination, dehydration, and reduction by two enzymatic steps catalyzed by L-lysine 6-dehydrogenase (deaminating) from *Silicibacter pomeroyi* and endogenous pyrroline 5-carboxylate reductase as well as a spontaneous reaction was embedded into L-lysine overproducing strain (Pérez-García et al. 2016) (Fig. 2). Transport engineering abolishing L-lysine export improved production of L-pipecolic acid to a titer of 14.4 g L^{-1} (Pérez-García et al. 2017a). Engineered *E. coli* strains produced about 5 g L^{-1} L-pipecolic acid based on L-lysine cyclodeaminase from *Streptomyces hygroscopicus* (Ying et al. 2017). Other routes have been used in the biotransformation of L-lysine, but not for de novo fermentative production by *E. coli* or *C. glutamicum*. A synthetic consortium consisting of a starch-utilizing *E. coli* strain and an L-pipecolic acid-producing *C. glutamicum* strain has also been used for starch-based production of L-pipecolic acid (Sgobba et al. 2018) and the concept was extended to chitin utilization by a related microbial consortium (Vortmann et al. 2021).

3.7 Ectoine and Hydroxyectoine

Ectoine is a cyclic, non-proteogenic amino acid, and an extremolyte (Gießelmann et al. 2019; Richter et al. 2019) used to protect the skin from cell damage and aging, to treat atopic dermatitis, lung inflammation, allergic rhinitis, or Alzheimer's disease (Gießelmann et al. 2019). In a process called "microbial milking," the halophilic *Halomonas elongata* is grown under high salt conditions (15% salt) for intracellular ectoine accumulation before being transferred to low salt conditions (3% salt) for release of ectoine elicited by the osmotic down-shift (Becker and Wittmann 2020). To avoid the associated costly process layout, corrosive damage to conventional stainless steel fermenters and connected devices due to the high salt conditions, other microbial hosts have been engineered for de novo production under low salt conditions. Heterologous expression of the *ectABCD* operon from *Pseudomonas stutzeri* in *C. glutamicum* strains engineered to provide sufficient supply of the precursor aspartate semialdehyde, an intermediate of L-lysine biosynthesis, enabled ectoine overproduction with some hydroxyectoine (Gießelmann et al. 2019; Czech et al. 2018; Pastor et al. 2010; Pérez-García et al. 2017b) (Fig. 2). By deletion of the lysine exporter gene *lysE* (Vrljic et al. 1996) accumulation of L-lysine as a by-product was abolished which improved ectoine production to about 5 g L^{-1} from glucose without the use of a high-salinity medium (Becker and Wittmann 2020). *C. glutamicum* strain Ecto-5 produced about 22 g L^{-1} ectoine (Pérez-García et al. 2017b). By using a monocistronic design to individually control the expression level of each of the three genes *ectA, ectB, and ectC* and screening a library of about 200.000 possible variants strain *C. glutamicum* ect^{opt} producing about 65 g L^{-1} ectoine in a fed-batch process was isolated (Gießelmann et al. 2019). Using a two-stage glucose feeding strategy during fermentation of an engineered *E. coli* strain (Ning et al. 2016) increased ectoine production from about 25 g L^{-1} to about 48 g L^{-1} ectoine (Dong et al. 2021).

3.8 L-Theanine

The ethylated L-glutamate derivative L-theanine is the most abundant free amino acid in green tea and is responsible for its typical taste. Due to its taste-enhancing property and its health benefits, L-theanine finds application in several industries (Williams et al. 2019; Nakagawa 1970). While in tea plants L-theanine is synthesized by L-theanine synthetase (L-glutamate:ethylamine ligase, EC 6.3.1.6) from L-glutamate and monoethylamine (MEA), microbial biosynthesis of L-theanine relies on the (side) activities of certain glutamyl-transferase or amide synthetase enzymes such as γ-glutamyltranspeptidase (EC 2.3.2.2) (Suzuki and Kumagai 2002), L-glutaminase (EC 3.5.1.2) (Takashi et al. 1996), L-glutamine synthetase (GS, EC 6.3.1.2) (Tachiki 1986), or γ-glutamylmethylamide synthetase (EC 6.3.4.12, GMAS) (Kimura et al.

1992). As compared to GS, GMAS shows a higher ligation activity toward MEA, and, thus is preferred (Kimura et al. 1992; Mindt et al. 2020).

Fermentative production of L-theanine has been described by metabolically engineered *E. coli*, *Pseudomonas putida*, and *C. glutamicum* (Hagihara et al. 2021; Ma et al. 2020; Benninghaus et al. 2021). Expression of *gmaS* from *Methylovorus mays* in an L-glutamate producing *C. glutamicum* strain enabled L-thenanine production if MEA was added to the medium (Fig. 2). Deletion of mechanosensitive channel gene *yggB* reduced loss of L-glutamate as intermediate and a titer of about 42 g L^{-1} was reached (Ma et al. 2020). Upon heterologous expression of *gmaS* from *Methylobacterium extorquens,* an L-glutamate producing, *P. putida* strain produced about 17 g L^{-1} L-theanine from glycerol and MEA, and if xylose utilization genes from *Caulobacter crescentus* were expressed in addition a product titer of about 21 g L^{-1} L-theanine was achieved (Benninghaus et al. 2021). Notably, an L-theanine producing *E. coli* strain has been developed, which does not require the addition of external MEA (Hagihara et al. 2021). MEA is synthesized by endogenous acetaldehyde dehydrogenase EutE yields acetaldehyde from the central intermediate of carbon metabolism acetyl-CoA. Acetaldehyde subsequently istransaminated to MEA by alanine-dependent ω-transaminase SpuC-II from P. putida, which was known to be active with aromatic aldehydes such as vanillin (Hagihara et al. 2021). This E. coli strain produced about 16 g L 1 from glucose, but without the requirement to add MEA to the growth medium (Hagihara et al. 2021).

4 *N*-Alkylation of Amino Acids

N-alkylated amino acids can naturally be found in bacteria and eukaryotes. The N^5-ethylated L-glutamate derivative L-theanine, for example, is responsible for the umami taste in green tea leaves (see Sect. 3.8) (Cartwright et al. 1954; Narukawa et al. 2008). Incorporated into proteins and peptides, they increase lipophilicity and render proteins less prone to proteolysis. Thus, used as building blocks of peptide-based drugs, these peptidomimetics enhance pharmacokinetic properties contributing to high bioavailability such as higher stability against proteolytic degradation, receptor selectivity (Cody et al. 1997), and better membrane permeability (Gordon et al. 2002). Alternatively to the chemical synthesis, the less hazardous fermentative production of *N*-alkylated amino acids using dehydrogenases, methyltransferases, and reductases has been developed in recent years (Hyslop et al. 2019).

4.1 DpkA Derived Alkylation

Independent fermentative processes for *N*-alkylated amino acids using the *N*-alkylamino acid dehydrogenase DpkA from *Pseudomonas putida* have been developed (Mindt et al. 2019a, 2018, 2019b; Kerbs et al. 2021). Natively, DpkA

functions in the catabolism of D-lysine and D-proline, where it catalyzes imine reduction of piperidine-2-carboxylate and pyroline-2-carboxylate to L-pipecolic acid and L-proline, respectively (Muramatsu et al. 2005). Moreover, DpkA catalyzes the reductive alkylamination of 2-oxo acids yielding the respective *N*-alkylated amino acid. The spontaneous imine formed between the 2-oxo acid and the alkylamine in an aqueous solution is reduced by DpkA (Kung and Wagner 1970; Muramatsu et al. 2004; Mihara et al. 2005). DpkA accepts a wide range of 2-oxo acids (Mihara et al. 2005), some of which are present in the bacterial cell, e.g., as intermediates in the TCA cycle, glycolysis, or amino acid biosynthesis pathways. However, simple introduction of DpkA into a bacterial host might not be sufficient for the efficient and selective production of a specific *N*-alkylamino acid, but metabolic engineering for high precursor supply and high expression of *dpkA* is required.

High titers of *N*-methyl-L-alanine were achieved by metabolically engineered *C. glutamicum* combining heterologous expression of *dpkA* from *P. putida* and a high supply of the precursor 2-oxo acid pyruvate. For this approach, the pyruvate overproducing strain ELB-P was chosen (Wieschalka et al. 2012). In addition, methylamine was supplied as a methyl donor (Mindt et al. 2018). Pyruvate accumulation up to 17.6 g L^{-1} in shake flasks was achieved by deletion of pyruvate dehydrogenase subunit E1p (*aceE*), the pyruvate-quinone oxidoreductase gene (*pqo*), the lactate dehydrogenase gene (*ldhA*), and the terminal regulatory domain of the acetohydroxyacid synthase gene (*ilvN*). Furthermore, genes encoding alanine aminotransferase (*alaT*) and alanine-valine aminotransferase (*avtA*) were deleted to prevent L-alanine formation. Balancing the carbon source concentration and media optimization in respect to ammonium and methylamine content was improved and enabled production of 32 g L^{-1} *N*-methyl-L-alanine with a yield of 0.7 g g^{-1} glucose and a volumetric productivity of 0.35 g L^{-1} h^{-1} in a fed-batch bioreactor (Mindt et al. 2018) (Table 2).

To transfer this concept, further microbial routes were engineered for *N*-alkylated amino acid synthesis in *C. glutamicum*(Mindt et al. 2019a,**b,** Kerbs et al. 2021). Characterization of DpkA enzyme revealed glyoxylate as a substrate for *N*-methylamination yielding sarcosine, a glycine derivative, which shows potential as antipsychotic drug (Werdehausen et al. 2012; Tsai et al. 2004). The metabolically engineered glyoxylate producing strain GLX, which lacks malate synthase and shows low isocitrate dehydrogenase activity due to initiation codon exchange from ATG to GTG, was enabled to produce 2.4 g L^{-1} sarcosine in a glucose-based fermentation by overexpression of *dpkA* and addition of monomethylamine. Sarcosine production was improved to a titer of 8.7 g L^{-1} with a yield of 0.25 g g^{-1} and volumetric productivity of 0.12 g L^{-1} h^{-1} (Table 2) by the application of the second-generation feedstock xylose as sole carbon source and media optimization (Mindt et al. 2019b). Semi-rational enzyme engineering was employed to obtain a more efficient mutant of DpkA (Mindt et al. 2019a). Increased specific activity of DpkA ($DpkA^{F117L}$) for reductive alkylamination of glyoxylate with either methylamine or ethylamine was achieved by the introduction of a single mutation in the substrate-binding site. Integration of $DpkA^{F117L}$ into the central

Table 2 Overview of alkylated amino acids and diamines produced by microbial fermentation

Product	Host strain	Carbon source	Titer (g L^{-1})	Yield (g g^{-1})	Productivity (g L^{-1} h^{-1})	Fermentation mode
Glutamate derived						
N-acetyl-*cis*-4-hydroxy- L-proline	*E. coli*	Glucose	*	n.g.	n.g.	Shake flask
N-acetylputrescine	*C. glutamicum*	Glucose	1.6	n.g.	n.g.	Shake flask
Putrescine	*E. coli*	Glucose	42	0.26	1.23	Fed-batch
	C. glutamicum	Glucose	19	0.16	0.55	Fed-batch
		Rice straw hydrolysate	0.091	n.g.	0.0013	Shake flask
		Glycerol	0.4	n.g.	n.g.	Shake flask
		Xylose	1.3	n.g.	0.027	Shake flask
1,6-diaminohexane	*E. coli*	Glucose	0.5	n.g.	n.g.	Shake flask
Pyruvate derived						
N-methylalanine	*C. glutamicum*	Glucose	32	0.7	0.35	Fed-batch
Aspartate derived						
L-Carnitine	*E. coli*	Glycerol	0.0025	n.g.	n.g.	Shake flask
Cadaverine	*E. coli*	Glucose	28	n.g.	2.7	Shake flask
		Cellobiose	0.6	n.g.	n.g.	Test flasks
		Galactose	8.8	0.17	0.3	Fed-batch
	C. glutamicum					
		Glucose	104	n.g.	1.5	Fed-batch
		Xylose	103	n.g.	1.5	Fed-batch
	B. methanolicus					
		Methanol	6.5	n.g.	n.g.	Fed-batch
1,3-diaminopropane	*E. coli*	Glucose	13	n.g.	0.13	Fed-batch
Glyoxylate derived						

(continued)

Table 2 (continued)

Product	Host strain	Carbon source	Titer (g L^{-1})	Yield (g g^{-1})	Productivity (g L^{-1} h^{-1})	Fermentation mode
Sarcosine	*C. glutamicum*	Glucose	2.4	n.g.	n.g.	Shake flask
	C. glutamicum	Xylose	8.7	0.25	0.12	Fed-batch
N-ethylglycine		Rice straw hydrolysate	6.1	0.17	0.11	Fed-batch
Aromatic						
N-methyL-*O*-methylanthranilate	*E. coli*	Glucose	0.0157	n.g.	n.g.	Shake flask
N-methylanthranilate	*E. coli*	Glucose	0.0029	n.g.	0.0012	Shake flask
	C. glutamicum	Glucose	0.5	0.005	0.01	Fed-batch
N-methylphenylalanine	*C. glutamicum*	Glucose	0.7	0.052	0.01	Duetz
		Xylose	0.6	0.05	0.008	Duetz

n.g. not given
*:27 μmol g^{-1} wet cell weight

metabolism of GLX about 9 g L^{-1} sarcosine with a yield of 0.26 g g^{-1} and higher volumetric productivity (0.16 g L^{-1} h^{-1}) of sarcosine production was achieved. Above that, fermentative production of *N*-ethylglycine to a titer of 6.1 g L^{-1} with a yield of 0.17 g g^{-1} and volumetric productivity of 0.11 g L^{-1} h^{-1} was established from rice straw hydrolysate (Mindt et al. 2019a) (Table 2).

Sustainable production of industrially significant aromatic compounds via microbial fermentation is gaining more and more interest. The synthesis of aromatic amino acids is derived from the shikimate pathway in plants, fungi, algae, and bacteria (Mir et al. 2015). One of the two direct precursors for the shikimate pathway is the glycolytic intermediate PEP (Herrmann 1995). PEP also acts as a phosphate donor for phosphorylation of glucose during uptake via the PTS. Thus, almost 50% of the synthesized PEP is consumed by the PTS system (Flores et al. 2002). Sufficient precursor supply like PEP or shikimate is one of the challenging problems for aromatic compounds synthesis. Taking this into consideration a new fermentative route was established for the production of the secondary metabolite *N*-methylphenylalanine, derived from phenylpyruvate and methylamine. For this approach, a shikimate accumulating strain (Walter et al. 2020) was engineered to phenylpyruvate accumulating strain by deletion of the anthranilate synthase genes *trpEG* and of the genes encoding branched-chain amino acid aminotransferase IlvE and phenylalanine aminotransferase AroT to abolish biosynthesis of L-tryptophan and L-phenylalanine (Kerbs et al. 2021) (Fig. 2). Introduction of DpkA containing the amino acid exchanges P262A and M141L ($DpkA^{P262A,M141L}$) showed comparable catalytic efficiencies with phenylpyruvate and pyruvate, whereas the wild-type enzyme preferred the latter substrate over the former. Finally, glucose-based *N*-methylphenylalanine production to a titer of 0.73 ± 0.05 g L^{-1}, volumetric productivity of 0.01 g L^{-1} h^{-1}, and a yield of 0.052 g g^{-1} glucose was achieved. Xylose as the sole carbon source supported the production of *N*-methylphenylalanine to a titer of 0.6 ± 0.04 g L^{-1} with a volumetric productivity of 0.008 g L^{-1} h^{-1} and a yield of 0.05 g g^{-1} xylose (Kerbs et al. 2021) (Table 2).

4.2 SAM-Dependent Methylation

As *S*-adenosylmethionine (SAM) is the methyl donor for almost all cellular methylation reactions occurring in nature, the level of SAM must be regulated in response to metabolic changes (Luka et al. 2009). The glycine *N*-methyltransferase is methylating glycine in a SAM-dependent manner in mammalian cells to generate sarcosine, which plays a critical role in SAM homeostasis (Ducker and Rabinowitz 2017). SAM-dependent methyltransferases represent one of the largest enzyme classes (>200 proteins) in mammalian systems. However, many of them are still uncharacterized. Despite the rudimentary knowledge, they have been implied in various disease states (Petrossian and Clarke 2011).

In plants, a SAM-dependent transfer of a methyl group to anthranilate yielding *N*-methylanthranilate initiates the biosynthesis of *N*-methylated acridone alkaloids

and avenacin (Mugford et al. 2013; Rohde et al. 2008). Acridone alkaloids and avenacin are used for pharmaceutical and therapeutical purposes, since they have anticancer, cytotoxic, and antimicrobial properties (Michael 2017). Only one *N*-methyltransferase (ANMT) was characterized in *Ruta graveolens*, a plant that exclusively accumulates *N*-methylated acridones (Rohde et al. 2008). Solely overexpression of ANMT in the production host *E. coli* resulted in low titers of *N*-methylanthranilate (0.0016 g L^{-1}). Improved titers were obtained by deletion of the genes *tyrR* and *metJ*, encoding the negative regulator of transcription of the aromatic amino acid biosynthesis and SAM biosynthesis, respectively (Lee et al. 2019). In addition, the precursor supply was enhanced by deletion of *trpD* gene encoding anthranilate phosphoribosyl transferase and overexpression of *trpEG* gene encoding anthranilate synthase. A final titer of 0.029 g L^{-1} of *N*-methylanthranilate was achieved by glucose-based fermentation (Lee et al. 2019) (Table 2). Further conversion of *N*-methylanthranilate to *O*-methyL-*N*-methylanthranilate by ANMT yielded to a final titer of 0.016 g L^{-1} (Lee et al. 2019). Recently, the genome-reduced and robust host for synthetic biology *C. glutamicum* chassis strain C1* (Baumgart et al. 2018) was metabolically engineered for fermentative production of *N*-methylanthranilate (Walter et al. 2020). By metabolic engineering of the central carbon metabolism, aromatic amino acids biosynthesis and overexpression of anthranilate synthase gene *trpE* from *E. coli* in a feed-back resistant variant, an anthranilate accumulating *C. glutamicum* strain was established (3.1 g L^{-1}). Embedding ANMT from *Ruta graveolens* into the metabolism resulted in 0.25 g L^{-1} *N*-methylanthranilate. With co-expression of *sahH* gene encoding *S*-adenosylhomocysteine hydrolase, an elegant cofactor regeneration system was established, yielding an improved titer of 0.34 g L^{-1} *N*-methylanthranilate. Transferring this strategy to a bioreactor process a final titer of 0.5 g L^{-1} of *N*-methylanthranilate with a volumetric productivity of 0.01 g L^{-1} h^{-1} and a product yield of about 0.005 g g^{-1} glucose was reached (Walter et al. 2020).

The quaternary amine L-carnitine is a central compound in the metabolism of eukaryotes, namely energy production and fatty acid metabolism (Bernal et al. 2016). L-carnitine synthesis starts with L-N^{ϵ}-trimethyllysine (TML), which is a product of lysosomal or post-translational degradation of proteins in eukaryotes (Strijbis et al. 2010). TML is hydroxylated to 3-hydroxy-TML, by TML hydroxylase (TMLH), which is then cleaved to glycine and 4-trimethylaminobutyraldehyde (TMABA). This step is catalyzed by 3-hydroxy-TML aldolase (HTMLA). The synthesis is preceded by oxidation of TMABA by 4-trimethylaminobutyraldehyde dehydrogenase (TMABADH) yielding γ-butyrobetaine (γ-BB). L-carnitine is finally synthesized by hydroxylation of γ-BB by γ-BB hydroxylase (γ-BBH) (Kugler et al. 2021). Fermentative production of L-carnitine was reached in *E. coli* by application of a biosensor system responding to externally added L-carnitine (Kugler et al. 2020). *E. coli* strains harboring the previously described L-carnitine pathway were able to produce 0.27 mg L^{-1} L-carnitine de novo. By the addition of 1 mM TML the strain reached a titer of 2.5 mg L^{-1} L-carnitine from glycerol (Kugler et al. 2021).

4.3 N-*Hydroxylation*

N-hydroxylation plays a role in many biosynthesis pathways, such as the biosynthesis of siderophores in plants and bacteria (Neilands 1993; Mügge et al. 2020), where hydroxylated diamines are the first precursors in siderophore biosynthesis. The marine bacterium *Shewanella putrefaciens*, for example, possesses a putrescine monooxygenase yielding *N*-hydroxyputrescine, which is further converted to the siderophore putrebactin (Soe et al. 2012). Hydroxylating enzymes are also known for enhancing antibiotic resistance (Liu et al. 2016).

The most known *N*-hydroxylating enzymes are the cytochrome P450 enzymes (CYPs) and flavin-dependent monooxygenases (FMO). CYPs play a crucial role in detoxification reactions of several endogenous compounds or xenobiotics (Guengerich et al. 2016; Zanger and Schwab 2013), and in the metabolism of carcinogens and drugs that temper cancer growth (Wahlang et al. 2015). Hydroxylation of the antibiotic rifampicin at N^2-position by an FMO causes higher resistance against this antibiotic in *Nocardia* isolates (Liu et al. 2016). However, microbial production of *N*-hydroxylated amines or diamines has not been established.

4.4 N-*Acetylation*

Nature provides many examples for *N*-acylation such as *N*-acetylation. Biosynthesis of melatonin, a hormone released by the pineal gland for regulation of the sleep–wake cycle (Jiao et al. 2016; Zhao et al. 2019) as well as the biosynthesis of siderophores containing a *N*-acetylated intermediate (Lamont et al. 2006). *N*-acetylation reactions require CoA-activated acryl-groups for the transfer to primary or secondary amino groups amines or proteins (Aksnes et al. 2019; VanDrisse and Escalante-Semerena 2019). Conserved in all domains of life, *N*-acetylation occurs also as a post-translational modification of proteins, where proteins are partly or fully acetylated at their N-terminus or as part of detoxification processes, such as antibiotic resistance. Bacterial cell walls consist of polymeric peptidoglycan containing *N*-acetyl-D-glucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc) (Aksnes et al. 2019; VanDrisse and Escalante-Semerena 2019; Seltmann and Holst 2013; Ahmed et al. 2009).

N-acetylation of putrescine was shown as a by-product in putrescine overproducing *C. glutamicum* strain by endogenous *N*-acetyltransferase SnaA. The engineered strain accumulated about 1.6 g L^{-1} *N*-acetylputrescine (Nguyen et al. 2015a) (Table 2).

Proline derivatives, such as *trans*-4-hydroxy-L-proline (THOP) and *cis*-4-hydroxy-L-proline (CHOP), play an important role in the pharmaceutical industry and are useful compounds in cosmetic applications (Bach et al. 2013). *N*-acetylation of THOP results in a compound (oxaceprol), which is used as an atypical inhibitor of inflammation and is applied in osteoarthritis disease treatment (Bach and Takagi

2013). The hydroxyproline analogue CHOP was found to be used as an anticancer drug and to have tumor growth prevention properties by preventing procollagen folding. (Kerr et al. 1987; Poiani et al. 1997). However, CHOP shows toxic effects on non-collagen proteins (Bach et al. 2013; Poiani et al. 1997). CHOP toxicity can be decreased by its *N*-alkylation (Bach et al. 2013). Fermentative production of *N*-actetyl CHOP was established by overexpression of L-proline *cis*-4-hydroxylase (SmP_4H) from Rhizobium *Sinorhizobium meliloti* and *N*-acetyltransferase (Mpr1) from yeast *Saccharomyces cerevisiae* in L-proline accumulating *E. coli* strain (Bach et al. 2013). After optimization of growth conditions, a final titer of 26.5 μmol g^{-1} wet cell weight after 42 h (in 15 mL) was achieved (Bach et al. 2013).

5 Diamines

Diamines are the nitrogenous monomeric precursors of polyamides and are cationic low molecular weight molecules, consisting of an aliphatic and saturated carbon chain and two primary amino groups (Wendisch et al. 2018). The diamines cadaverine and putrescine (1,4-diminobutane), the triamine spermidine, and the tetramine spermine are among the most abundant naturally occurring polyamines (Wendisch et al. 2018; Hamana and Matsuzaki 1992). In nature, diamines play an important role in the physiology of many organisms. For example, diamines occur as phytohormones in plants and they are used as modulators of various transport ion channels (Wendisch et al. 2018; Kurata et al. 2006). In bacteria, diamines may act as signals in cell differentiation (Sturgill and Rather 2004) or may be involved in pH homeostasis (Rhee et al. 2007; Takatsuma and Kamio 2004). In industry, diamines serve as platform molecules with important applications, such as the synthesis of pesticides, additives for motor oils and fuels, or metal ion chelating agents (Wang et al. 2020). The synthesis of polyamides using diamines is primarily performed chemically from petroleum or natural gas (Babu et al. 2013). A more sustainable and less hazardous way of diamine production was achieved by microbial fermentation.

5.1 Cadaverine (1,5-Diaminopentane)

Cadaverine is used as a platform chemical in the industrial production of polymers, polyurethane, and chelating agents (Qian et al. 2011). By polycondensation of bio-based cadaverine and fermentatively produced succinic acid the synthesis of polyamides 54 can be achieved, which is a bio-based alternative to petroleum-based polyamides (Qian et al. 2011). Recently, a green bio-based process for the polyamide 56 and antimicrobial nanofiber was established by using recombinant *E. coli* by whole-cell biotransformation of lysine (Xue et al. 2021).

The key step in the synthesis of the five-carbon diamine cadaverine is the decarboxylation of the six-carbon amino acid L-lysine by L-lysine decarboxylase

(CadA) (Fig. 2). Efficient lysine decarboxylation to cadaverine was reached by overexpression of *cadA* in a lysine overproducing *E. coli* strain. Deletion of genes related to cadaverine utilization like putrescine/cadaverine aminopropyltransferase *speE*, spermidine acetyltransferase *speG*, putrescine/cadaverine aminotransferase *ygjY,* and glutamate-putrescine/glutamate-cadaverine ligase *puuA* led to 9.6 g L^{-1} cadaverine with a volumetric productivity of 0.32 g L^{-1} in fed-batch cultivation (Qian et al. 2011). Repression of the UDP-*N*-acetylmuramoylalanyl-D-glutamate 2,6-diaminopimelate ligase gene *murE*, involved in peptidoglycan biosynthesis, by small regulatory RNAs led to 55% improved cadaverine production with a titer of 12.6 g L^{-1} in fed-batch cultivation (Na et al. 2013). In a new approach using clustered regularly interspaced short palindromic repeats interference (CRISPRi) for gene repression, an engineered *E.coli* strain produced 28 g L^{-1} cadaverine with a volumetric productivity of 2.67 g L^{-1} h^{-1}, when *ygjP* was targeted (Ting and Ng 2020).

Usage of cellobiose as alternative carbon sources for direct cadaverine production was achieved by co-overexpression of endogenous CadA and β-glucosidase from *Thermobifida fusca.* A final titer of 0.6 g L^{-1} was reached within 48 h (Ikeda et al. 2013). In a different approach the native galactose operon *galETKM*, the genes *pgm* encoding a phosphoglucomutase and *galP* encoding for galactose-proton symporter were replaced by a synthetic expression cassette, to remove catabolic carbon repression (Lim et al. 2013; Kwak et al. 2017). In combination with overexpression of the endogenous *cadA* gene, a titer of 2 g L^{-1} cadaverine was accomplished using galactose as carbon source (Kwak et al. 2017). Improved cadaverine production to 8.8 g L^{-1} within 30 h in fed-batch fermentation was reached by deletion of putrescine/cadaverine aminopropyl transferase gene *speE*, the spermidine acetyltransferase gene *speG*, the γ-glutamylputrescine synthetase gene *puuA*, the putrescine importer gene *puuP*, and the cadaverine aminotransferase gene *ygjG* (Kwak et al. 2017). Recently, soybean residue hydrolysate, an economical source of nitrogen, was investigated for cadaverine production. It was shown that raw soybean residue hydrolysate, a by-product from soybean protein food processing (Zhou et al. 2011), can completely replace yeast extract as the only nitrogen source in metabolically engineered *E. coli* strains and a final titer of 0.52 g L^{-1} h^{-1} was reached with a constant feeding strategy and a total amount of 30 g L^{-1} sugar concentration (Guo et al. 2021).

C. glutamicum is a logical choice for cadaverine production since it is well established for lysine production with 2,400,000 tons/year (Lee and Wendisch 2017). To establish cadaverine production, the *hom* gene encoding for L-homoserine dehydrogenase was replaced with the *cadA* gene from *E. coli* in *C. glutamicum* ATCC 13032. With this strategy a final titer of 2.6 g L^{-1} cadaverine was produced (Mimitsuka et al. 2007). Alternatively, lysine decarboxylase gene encoded by *ldcC* was heterologously overexpressed in *C. glutamicum (*Kind et al. 2010*)*. Previously, this strain was engineered for a higher precursor supply for lysine biosynthesis. Thus, modifications like an L-lysine feed-back resistant version of the aspartokinase encoded by $lysC^{T311I}$ (Kalinowski et al. 1991), the attenuation of homoserine dehydrogenase encoded by hom^{V59A} (Eikmanns et al. 1991), and enhancement of

anaplerosis via pyruvate carboxylase encoded by pyc^{P485S} (Peters-Wendisch et al. 2001) were performed. By external addition of the cofactor pyridoxal the strain produced 0.3 mol mol^{-1} cadaverine (Kind et al. 2010). In another approach, a lysine production *C. glutamicum* strain (120 g L^{-1}) with 12 defined genome-based changes was chosen as a platform strain (Becker et al. 2011). Improved cadaverine production was achieved by heterologous expression of a codon-optimized *ldcC* gene. Formation of the by-product *N*-acetylcadaverine was prevented by deletion of *snaA* encoding an *N*-acetyltransferase, which shows activity with the diamines putrescine, spermidine, and spermine (Nguyen et al. 2015a). Together with the deletion of the L-lysine exporter *lysE* and overexpression of *cgmA*, which encodes for the export system for L-arginine and the diamines putrescine and cadaverine (Nguyen et al. 2015a; Lubitz et al. 2016), a final titer of 88 g L^{-1} of cadaverine was produced in fed-batch cultivation (Table 2) (Lubitz et al. 2016). Engineered industrial *C. glutamicum* PKC lysine producing strain was further modified for cadaverine synthesis, by integration of *ldcC* into the locus of *lysE* and controlling its expression via a strong H30 promotor. Industrial glucose-based cadaverine production reached the highest titer of about 104 g L^{-1} in fed-batch culture with a volumetric productivity of about 1.5 g L^{-1} h^{-1} (Kim et al. 2018).

By the heterologous expression of the *xylA* and *xylB* genes, combined with overexpression of the *tkt* operon, downregulation of the *sucCD* and isocitrate dehydrogenase gene *icd*, deletion of the cg*1722* and *lysE* genes, and overexpression of the *fbp* gene, a xylose based cadaverine production was established with a titer of 103 g L^{-1} and a productivity of 1.5 g L^{-1} h^{-1} in fed-batch cultivation (Buschke et al. 2011; Buschke et al. 2013).

Cadaverine production was also shown in the methylotroph *B. methanolicus*. After metabolically engineering *B. methanolicus* strain was able to produce 6.5 g L^{-1} cadaverine in fed-batch cultivation from methanol, a reduced C1 compound that can be produced from CO_2 (Nærdal et al. 2015; Irla et al. 2019).

5.2 Putrescine (1,4-Diaminobutane)

The diamine putrescine consists of four carbon atoms and is used as a component in insecticides, medicines, and other chemicals (Chae et al. 2020). Many organisms have been reported to naturally synthesize putrescine, since it has an important role in cell growth (Tabor and Tabor 1985). Biological putrescine synthesis can be either achieved via multi-step conversion of arginine or via one-step ornithine decarboxylation (Schneider and Wendisch 2010; Chae et al. 2020). However, the latter pathway is dominantly used for fermentative putrescine production, because in the arginine derived pathway urea is generated as a by-product, which cannot further be utilized by some microorganisms (Schneider and Wendisch 2010; Qian et al. 2009).

E. coli natively synthesizes putrescine; however, the wild-type strain only produces a trace amount of putrescine (Sugiyama et al. 2016). Thus, metabolic engineering of *E. coli* was performed to redirect the biosynthetic flux toward putrescine

synthesis. First, accumulation of the precursor L-ornithine was established by overexpression of *N*-acetyl-L-glutamate kinase gene *argB*, *N*-acetyl-γ-glutamyl-phosphate reductase gene *argC*, *N*-acetyl-L-ornithine amino-transferase gene *argD*, *N*-acetyl-L-ornithine deacetylase gene *argE*, argininosuccinate lyase gene *argH*, and putrescine/L-ornithine antiporter gene *potE*. Conversion of ornithine to putrescine was enhanced by overexpression of biosynthetic ODC gene *speC* and degradative ODC gene *speF*. Competing pathways were disturbed by deletion of spermidine synthase gene *speE*, spermidine *N*-acetyltransferase gene *speG*, the putrescine importer gene *puuP*, γ-glutamate-putrescine ligase encoded by *puuA*, and L-ornithine carbamoyl transferase gene *argI* (Fig. 2). Additional deletion of *rpoS* gene, encoding a stress-responsive RNA polymerase sigma factor, yielded a final glucose-based production of 24 g L^{-1} putrescine with a productivity of 0.75 g L^{-1} h^{-1} in fed-batch cultivation (Qian et al. 2009). Deletion of both L-ornithine carbamoyl transferase genes *argI* and *argF* is causing undesirable L-arginine auxotrophy for large-scale fermentations. Instead of deletion, the expression level of *argF* gene was fine-tuned by the introduction of a synthetic small RNA targeting *argF*. In addition, L-glutamine synthase gene *glnA* was repressed, which resulted in a final titer of 43 g L^{-1} putrescine production (Noh et al. 2017).

In comparison to *E. coli*, *C. glutamicum* lacks putrescine biosynthesis and catabolism; however, it can tolerate up to 500 mM during fast growth (Schneider and Wendisch 2010). Thus, *C. glutamicum* is a suitable production host for engineering and improvement of effective putrescine production. Already little modifications, such as overexpression of *speC* gene, derived from *E. coli*, and deletion of the arginine repressor *argR* and L-ornithine carbamoyl transferase gene *argF* yielded in 6 g L^{-1} putrescine production with a volumetric productivity of 0.1 g L^{-1} h^{-1} (Schneider and Wendisch 2010). Further modifications were performed to a) overcome L-arginine auxotrophy, caused by deletion of *argF*, b) increase precursor supply, and c) prevent by-product formation. By decreasing *argF* activity through modifications of the promotor, the translational start codon (TSS), and the ribosome binding site (RBS), already an improved titer of 19 g L^{-1} putrescine was achieved in glucose-based fed-batch fermentation (Schneider et al. 2012). Due to the reduction of 2-oxoglutarate dehydrogenase activity by exchange of the TSS and overexpression of glyceraldehyde-3-phosphate dehydrogenase and pyruvate carboxylase genes the precursor supply was increased. By-product formation of *N*-acetylputrescine was prevented by deletion of putrescine/polyamine *N*-acetyltransferase of *snaA* gene. Further, a decrement of feed-back inhibition of acetylglutamate kinase by L-arginine was established (Nguyen et al. 2015a). Finally, *cgmA* deletion increased putrescine production up to 24%, presumably by abolishing the export of putrescine (Nguyen et al. 2015b; Kind et al. 2011). Recently, putrescine production was shown from hydrolyzed rice straw, containing a mixture of 23.7 g L^{-1} glucose and 13.6 g L^{-1} xylose. In this approach 0.091 g L^{-1} putrescine with a volumetric productivity of 0.001 g L^{-1} h^{-1} was produced by an engineered *C. glutamicum* strain overexpressing *xylA* gene from *Xanthomonas campestris* together with endogenous *xylB* (Sasikumar et al. 2021).

Utilization of alternative carbon sources was established by overexpression of the glycerol utilization operon *glpFKD* of *E. coli* in a putrescine production host (Meiswinkel et al. 2013a). A final titer of 0.4 g L^{-1} was reached from 20 g L^{-1} glycerol (Meiswinkel et al. 2013a). *C. glutamicum* strains were also engineered for utilization of the pentoses arabinose, by the introduction of *araBAD* genes from *E. coli (*Kawaguchi et al. 2008*)*, and xylose, by the overexpression of *xylA* and *xylB* genes from *E. coli* (Kawaguchi et al. 2006) or the *xylA* gene from *Xanthomonas campestris* together with *xylB* from *C. glutamicum* (Meiswinkel et al. 2013b). Metabolically engineered putrescine overproduction *C. glutamicum* strains were able to produce 1.3 g L^{-1} putrescine with a volumetric productivity of 27 mg L^{-1} h^{-1} from xylose (Meiswinkel et al. 2013b).

By pathway extension, a *C. glutamicum* strain overproducing putrescine was engineered for production of 4-amino-1-butanol, an intermediate for drugs and a precursor of degradable polymers (Prabowo et al. 2020). The extension comprised putrescine amino transferase yielding aminobutyraldehyde and aldehyde dehydrogenase for reduction of the aldehyde to the amino alcohol. Upon fine-tuning the overexpression of the respective genes *ygjG* and *yqhD* about 25 g L^{-1} 4-amino-1-butanol was produced (Prabowo et al. 2020).

5.3 1,3-Diaminopropane

Synthesis of the diamine 1,3-diaminopropane occurs in a few microorganisms, like *Actinobacteria* and *Pseudomonas* species. It can be synthesized via two pathways: the C_4 and the C_5 pathway (Tabor and Tabor 1985). It was shown that *Acinetobacter baumannii* can synthesize 1,3-diaminopropane via the C_4 pathway by conversion of L-aspartate semialdehyde by 2-oxoglutarate 4-aminotransferase Dat (Ikai and Yamamoto 1997) and L-2,4-diaminobutanoate decarboxylase Ddc (Ikai and Yamamoto 1994). As opposed to this, *Pseudomonas aeruginosa* can directly synthesize 1,3-diaminopropane in a C_5 pathway by one single dehydrogenation, performed by spermidine dehydrogenase (encodes by *spdH*) (Dasu et al. 2006). Heterologous constructed pathway for 1,3-diaminopropane synthesis was employed in engineered *E. coli* cells. In silico flux response analysis revealed that the production rate of 1,3-diaminopropane is higher via the C_4 pathway than via the C_5 pathway and also no additional co-factors are required. Therefore, the C_4 pathway including overexpression of codon-optimized *dat* and *ddc* genes was introduced to an *E. coli* strain that harbors feed-back resistant aspartokinases in combination with overexpressed *ppc* and *aspC* genes to increase the flux toward 1,3-diaminopropane (Chae et al. 2015). Additional deletion of ATP-dependent 6-phosphofructokinase isozyme 1 gene *pfkA* was beneficial for 1,3-diaminopropane production. The final engineered *E. coli* strain produced 13 g L^{-1} of 1,3-diaminopropane in a glucose minimal medium during fed-batch cultivation (Chae et al. 2015).

5.4 *1,6-Diaminohexane*

The six-carbon diamine 1,6-diaminohexane is highly interesting for the industry, since it can be used for the synthesis of polyamides PA6,6 and PA6,10 (Chae et al. 2020). Microbial production of 1,6-diaminohexane turned out to be challenging because of the necessity of many complicated steps (Burk et al. 2013; Lau 2018). Nevertheless, microbial production of 1,6-diaminohexane was achieved by heterologous expression of key enzymes glutamyl-coenzyme A transferase, β-keto thiolase, 2-oxoacid reductase, dehydratase, acyl-CoA dehydrogenase, dehydrogenase, amine dehydrogenase, and homo lysine decarboxylase (Fig. 2). With this strategy, 0.5 g L^{-1} 1,6-diaminohexane production was achieved (Table 2) (Lau 2018).

6 Concluding Remarks and Outlook

The traditional amino acid industry has recently seen exciting developments that were driven by low margins even though benefitting from economies of scale, on the one hand, and by technology push on the other hand. The new technologies for amino acid production strain development included, for example, CRISPR (Cleto et al. 2016; Schultenkämper et al. 2020; Wang et al. 2021), flux enforcement (von Kamp and Klamt 2017; Prell et al. 2021; Haupka et al. 2021), ALE (Stella et al. 2019; Hennig et al. 2020; Graf et al. 2019; Yu et al. 2021), and droplet-based screening (Balasubramanian et al. 2021). By using a CRISPR-Cas system for gene deletion and expression of heterologous *gadB2* gene from *Lactobacillus brevis* ATCC 367 high production of GABA was achieved. Repression of genes in *E. coli* was applied by using CRISPRi to produce high titers of the diamine cadaverine. Besides novel strain development technology, we foresee improved process control strategies including designed microbial consortia (Burmeister et al. 2018; Vortmann et al. 2021; Sgobba and Wendisch 2020) and their dynamic control (Pérez-García et al. 2021), open pond-fermentation based on xenogeneic nitrogen and phosphorus sources (Schwardmann et al. 2022; Guo et al. 2020), insight into population heterogeneity by single-cell analysis (Grünberger et al. 2014; Dusny and Grünberger 2020; Täuber et al. 2021).

At a large scale, production based on carbon sources without competing uses as food and feed becomes pivotal. A number of examples for the use of agricultural side streams for amino acid production have been published (Chen et al. 2020; Mindt et al. 2019a; Guo et al. 2021; Sasikumar et al. 2021). Contaminants and inhibitors present in these side streams still hamper process efficiency, both regarding cellular production and downstream processing. Solutions to be developed need to take into consideration also the seasonal qualitative variation and availability.

The fermentation processes described here outcompete chemical synthesis with respect to regio- and stereoselectivity. In this respect, fermentative production of L-theanine may serve as an example, since the most abundant free amino acid in green

tea could be produced by metabolically engineered *E. coli*, *P. putida,* and *C. glutamicum* strains. The fact that production of L-theanine became possible due to the (side) activities of certain glutamyl-transferases or amide synthetases (-γ-glutamyltranspeptidase, L-glutaminase, L-glutamine synthetase, or γ-glutamylmethylamide synthetase) indicates that it is well worthwhile to explore the substrate spectra of enzymes catalyzing related enzyme reactions when a new fermentative route is sought after. This can also be exemplified by the production of alkylated amino acids based on the side activity of the amino acid dehydrogenase DpkA from *P. putida* that readily alkylaminates 2-oxo acids. Enzyme engineering will play an increasing role in streamlining the substrate spectrum to fit the cellular supply of 2-oxo acids that can be adapted by complementary metabolic engineering.

We foresee that the technology push will drive further developments to produce a wider range of amines and amino acids by fermentative processes. Specifically, functionalization of nitrogenous compounds by alkylation, halogenation, and hydroxylation will be accessible via fermentation.

References

Adkins J, Jordan J, Nielsen DR (2013) Engineering *Escherichia coli* for renewable production of the 5-carbon polyamide building-blocks 5-aminovalerate and glutarate. Biotechnol Bioeng 110: 1726–1734

Ahmed AM, Younis EEA, Osman SA et al (2009) Genetic analysis of antimicrobial resistance in *Escherichia coli* isolated from diarrheic neonatal calves. Vet Microbiol 136:397–402

Aksnes H, Ree R, Arnesen T (2019) Co-translational, post-translational, and non-catalytic roles of *N*-terminal acetyltransferases. Mol Cell 73:1097–1114

Ajinomoto Co., Inc. (2021) https://www.ajinomoto.co.jp/company/en/ir/event/presentation/main/011117/teaserItems1/01/linkList/01/link/FY21Q1_Results_E.pdf Accessed 27 Sep 2021

Babu RP, O'Connor K, Seeram R (2013) Current progress on bio-based polymers and their future trends. Prog Biomater 2:8

Bach TMH, Hara R, Kino K et al (2013) Microbial production of *N*-acetyl cis-4-hydroxy-L-proline by coexpression of the rhizobium L-proline *cis*-4-hydroxylase and the yeast *N*-acetyltransferase Mpr1. Appl Microbiol Biotechnol 97:247–257

Bach TMH, Takagi H (2013) Properties, metabolisms, and applications of L-proline analogues. Appl Microbiol Biotechnol 97:6623–6634

Balasubramanian S, Chen J, Wigneswaran V et al (2021) Droplet-based microfluidic high throughput screening of *Corynebacterium glutamicum* for efficient heterologous protein production and secretion. Front Bioeng Biotechnol 9:668513

Baumgart M, Unthan S, Kloß R et al (2018) *Corynebacterium glutamicum* chassis C1*: building and testing a novel platform host for synthetic biology and industrial biotechnology. ACS Synth Biol 7:132–144

Becker J, Wittmann C (2020) Microbial production of extremolytes - high-value active ingredients for nutrition, health care, and well-being. Curr Opin Biotechnol 65:118–128

Becker J, Zelder O, Häfner S et al (2011) From zero to hero—design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. Metab Eng 13:159–168

Benninghaus L, Walter T, Mindt M et al (2021) Metabolic engineering of *Pseudomonas putida* for fermentative production of L-theanine. J Agric Food Chem 69:9849–9858

Bernal V, Arense P, Cánovas M (2016) L-carnitine, the vitamin BT : uses and production by the secondary metabolism of bacteria. In: Industrial biotechnology of vitamins, biopigments, and antioxidants. John Wiley & Sons, Ltd, pp 389–419

Burk MJ, Burgard AP, Osterhout RE, Pharkya P (2013) Microorganisms and methods for the biosynthesis of adipate, hexamethylenediamine and 6-aminocaproic acid. BRPI1011227A8

Burmeister A, Akhtar Q, Hollmann L et al (2021) (Optochemical) control of synthetic microbial coculture interactions on a microcolony level. ACS Synth Biol 10:1308–1319

Burmeister A, Hilgers F, Langner A et al (2018) A microfluidic co-cultivation platform to investigate microbial interactions at defined microenvironments. Lab Chip 19:98–110

Buschke N, Becker J, Schäfer R et al (2013) Systems metabolic engineering of xylose-utilizing *Corynebacterium glutamicum* for production of 1,5-diaminopentane. Biotechnol J 8:557–570

Buschke N, Schröder H, Wittmann C (2011) Metabolic engineering of *Corynebacterium glutamicum* for production of 1,5-diaminopentane from hemicellulose. Biotechnol J 6:306–317

Cartwright RA, Roberts EAH, Wood DJ (1954) Theanine, an amino-acid *n*-ethyl amide present in tea. J Sci Food Agric 5:597–599

Chae TU, Ahn JH, Ko Y-S et al (2020) Metabolic engineering for the production of dicarboxylic acids and diamines. Metab Eng 58:2–16

Chae TU, Kim WJ, Choi S et al (2015) Metabolic engineering of *Escherichia coli* for the production of 1,3-diaminopropane, a three carbon diamine. Sci Rep 5:13040

Chen J, Wang Y, Guo X et al (2020) Efficient bioproduction of 5-aminolevulinic acid, a promising biostimulant and nutrient, from renewable bioresources by engineered *Corynebacterium glutamicum*. Biotechnol Biofuels 13:41

Cheng J, Luo Q, Duan H et al (2020) Efficient whole-cell catalysis for 5-aminovalerate production from L-lysine by using engineered *Escherichia coli* with ethanol pretreatment. Sci Rep 10:990

Cho JS, Choi KR, Prabowo CPS et al (2017) CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. Metab Eng 42:157–167

Choi JW, Yim SS, Lee SH et al (2015) Enhanced production of gamma-aminobutyrate (GABA) in recombinant *Corynebacterium glutamicum* by expressing glutamate decarboxylase active in expanded pH range. Microb Cell Factories 14:21

Cleto S, Jensen JVK, Wendisch VF, Lu TK (2016) *Corynebacterium glutamicum* metabolic engineering with CRISPR interference (CRISPRi). ACS Synth Biol 5:375–385

Cody WL, He JX, Reily MD et al (1997) Design of a potent combined pseudopeptide endothelin-A/endothelin-B receptor antagonist, Ac-dBhg16-Leu-Asp-Ile-[NMe]Ile-Trp21 (PD 156252): examination of its pharmacokinetic and spectral properties. J Med Chem 40:2228–2240

Costas AM, White AK, Metcalf WW (2001) Purification and characterization of a novel phosphorus-oxidizing enzyme from *Pseudomonas stutzeri* WM88. J Biol Chem 276:17429–17436

Czech L, Poehl S, Hub P et al (2018) Tinkering with osmotically controlled transcription allows enhanced production and excretion of ectoine and hydroxyectoine from a microbial cell factory. Appl Environ Microbiol 84:e01772-17

Dasu VV, Nakada Y, Ohnishi-Kameyama M et al (2006) Characterization and a role of *Pseudomonas aeruginosa* spermidine dehydrogenase in polyamine catabolism. Microbiology 152:2265–2272

Della Corte D, van Beek HL, Syberg F et al (2020) Engineering and application of a biosensor with focused ligand specificity. Nat Commun 11:4851

Dong Y, Zhang H, Wang X et al (2021) Enhancing ectoine production by recombinant *Escherichia coli* through step-wise fermentation optimization strategy based on kinetic analysis. Bioprocess Biosyst Eng 44:1557–1566

Ducker GS, Rabinowitz JD (2017) One-carbon metabolism in health and disease. Cell Metab 25:27–42

Dusny C, Grünberger A (2020) Microfluidic single-cell analysis in biotechnology: from monitoring towards understanding. Curr Opin Biotechnol 63:26–33

Eikmanns BJ, Metzger M, Reinscheid D et al (1991) Amplification of three threonine biosynthesis genes in *Corynebacterium glutamicum* and its influence on carbon flux in different strains. Appl Microbiol Biotechnol 34:617–622

Evonik Industries (2016a) Evonik to acquire technology from METEX for the fermentative production of methionine—Evonik Industries. https://corporate.evonik.com/en/media/press-releases/nutrition-and-care/evonik-to-acquire-technology-from-metex-for-the-fermentative-production-of-methionine-106336.html. Accessed 6 Aug 2021

Evonik Industries (2016b) Evonik starts construction of second methionine complex in Singapore - Evonik Industries. https://corporate.evonik.com/en/investor-relations/evonik-starts-construction-of-second-methionine-complex-in-singapore-106948.html. Accessed 6 Aug 2021

Feng L, Zhang Y, Fu J et al (2016) Metabolic engineering of *Corynebacterium glutamicum* for efficient production of 5-aminolevulinic acid. Biotechnol Bioeng 113:1284–1293

Flores S, Gosset G, Flores N et al (2002) Analysis of carbon metabolism in *Escherichia coli* strains with an inactive phosphotransferase system by 13C labeling and NMR spectroscopy. Metab Eng 4:124–137

Frese M, Sewald N (2015) Enzymatic halogenation of tryptophan on a gram scale. Angew Chem Int Ed 54:298–301

Georgi T, Rittmann D, Wendisch VF (2005) Lysine and glutamate production by *Corynebacterium glutamicum* on glucose, fructose and sucrose: roles of malic enzyme and fructose-1,6-bisphosphatase. Metab Eng 7:291–301

Gießelmann G, Dietrich D, Jungmann L et al (2019) Metabolic engineering of *Corynebacterium glutamicum* for high-level ectoine production: design, combinatorial assembly, and implementation of a transcriptionally balanced heterologous ectoine pathway. Biotechnol J 14:e1800417

Global Amino Acids Market (2021) | COVID-19 Impact, Methionine, MSG, Valine. https://industry-experts.com/verticals/food-and-beverage/global-amino-acids-market-products-and-applications. Accessed 6 Aug 2021

Gordon DJ, Tappe R, Meredith SC (2002) Design and characterization of a membrane permeable *N*-methyl amino acid-containing peptide that inhibits Aβ1–40 fibrillogenesis. J Pept 60:37–55

Graf M, Haas T, Müller F et al (2019) Continuous adaptive evolution of a fast-growing *Corynebacterium glutamicum* strain independent of protocatechuate. Front Microbiol 10:1648

Grünberger A, Wiechert W, Kohlheyer D (2014) Single-cell microfluidics: opportunity for bioprocess development. Curr Opin Biotechnol 29:15–23

Guengerich FP, Waterman MR, Egli M (2016) Recent structural insights into cytochrome P450 function. Trends Pharmacol Sci 37:625–640

Guo X, Li M, Li H et al (2021) Enhanced cadaverine production by engineered *Escherichia coli* using soybean residue hydrolysate (SRH) as a sole nitrogen source. Appl Biochem Biotechnol 193:533–543

Guo Z-W, Ou X-Y, Liang S et al (2020) Recruiting a phosphite dehydrogenase/formamidase-driven antimicrobial contamination system in *Bacillus subtilis* for nonsterilized fermentation of acetoin. ACS Synth Biol 9:2537–2545

Hagihara R, Ohno S, Hayashi M et al (2021) Production of L-theanine by *Escherichia coli* in the absence of supplemental ethylamine. Appl Environ Microbiol 87:e00031–e00021

Hamana K, Matsuzaki S (1992) Polyamines as a chemotaxonomic marker in bacterial systematics. Crit Rev Microbiol 18:261–283

Hammer PE, Hill DS, Lam ST et al (1997) Four genes from *Pseudomonas fluorescens* that encode the biosynthesis of pyrrolnitrin. Appl Environ Microbiol 63:2147–2154

Haupka C, Brito LF, Busche T et al (2021) Genomic and transcriptomic investigation of the physiological response of the methylotroph *bacillus methanolicus* to 5-aminovalerate. Front Microbiol 12:664598

Haupka C, Delépine B, Irla M et al (2020) Flux enforcement for fermentative production of 5-Aminovalerate and Glutarate by *Corynebacterium glutamicum*. Catalysts 10:1065

Henke NA, Austermeier S, Grothaus IL et al (2020) *Corynebacterium glutamicum* CrtR and its orthologs in *actinobacteria*: conserved function and application as genetically encoded biosensor for detection of geranylgeranyl pyrophosphate. Int J Mol Sci 21:5482

Henke NA, Wiebe D, Pérez-García F et al (2018) Coproduction of cell-bound and secreted value-added compounds: simultaneous production of carotenoids and amino acids by *Corynebacterium glutamicum*. Bioresour Technol 247:744–752

Hennig G, Haupka C, Brito LF et al (2020) Methanol-essential growth of *Corynebacterium glutamicum*: adaptive laboratory evolution overcomes limitation due to methanethiol assimilation pathway. Int J Mol Sci 21:3617

Herrmann KM (1995) The shikimate pathway as an entry to aromatic secondary metabolism. Plant Physiol 107:7–12

Hirota R, Abe K, Katsuura Z-I et al (2017) A novel biocontainment strategy makes bacterial growth and survival dependent on phosphite. Sci Rep 7:44748

Hyslop JF, Lovelock SL, Watson AJB et al (2019) *N*-Alkyl-α-amino acids in nature and their biocatalytic preparation. J Biotechnol 293:56–65

Ikai H, Yamamoto S (1994) Cloning and expression in *Escherichia coli* of the gene encoding a novel L-2,4-diaminobutyrate decarboxylase of *Acinetobacter baumannii*. FEMS Microbiol Lett 124:225–228

Ikai H, Yamamoto S (1997) Identification and analysis of a gene encoding L-2,4-diaminobutyrate:2-ketoglutarate 4-aminotransferase involved in the 1,3-diaminopropane production pathway in *Acinetobacter baumannii*. J Bacteriol 179:5118–5125

Ikeda N, Miyamoto M, Adachi N et al (2013) Direct cadaverine production from cellobiose using β-glucosidase displaying *Escherichia coli*. AMB Express 3:67

Irla M, Heggeset TMB, Nærdal I et al (2019) Corrigendum: genome-based genetic tool development for *Bacillus methanolicus*: theta- and rolling circle-replicating plasmids for inducible gene expression and application to methanol-based cadaverine production. Front Microbiol 10:425

Jensen JVK, Eberhardt D, Wendisch VF (2015) Modular pathway engineering of *Corynebacterium glutamicum* for production of the glutamate-derived compounds ornithine, proline, putrescine, citrulline, and arginine. J Biotechnol 214:85–94

Jensen JVK, Wendisch VF (2013) Ornithine cyclodeaminase-based proline production by *Corynebacterium glutamicum*. Microb Cell Factories 12:1–10

Jiao J, Ma Y, Chen S et al (2016) Melatonin-producing endophytic bacteria from grapevine roots promote the abiotic stress-induced production of endogenous melatonin in their hosts. Front Plant Sci 7:1387

Jorge JMP, Leggewie C, Wendisch VF (2016) A new metabolic route for the production of gamma-aminobutyric acid by *Corynebacterium glutamicum* from glucose. Amino Acids 48:2519–2531

Jorge JMP, Pérez-García F, Wendisch VF (2017) A new metabolic route for the fermentative production of 5-aminovalerate from glucose and alternative carbon sources. Bioresour Technol 245:1701–1709

Jurischka S, Bida A, Dohmen-Olma D et al (2020) A secretion biosensor for monitoring sec-dependent protein export in *Corynebacterium glutamicum*. Microb Cell Factories 19:11

Kalinowski J, Cremer J, Bachmann B et al (1991) Genetic and biochemical analysis of the aspartokinase from *Corynebacterium glutamicum*. Mol Microbiol 5:1197–1204

Kanda K, Ishida T, Hirota R et al (2014) Application of a phosphite dehydrogenase gene as a novel dominant selection marker for yeasts. J Biotechnol 182–183:68–73

Kawaguchi H, Sasaki M, Vertès AA et al (2008) Engineering of an L-arabinose metabolic pathway in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 77:1053–1062

Kawaguchi H, Vertès AA, Okino S et al (2006) Engineering of a xylose metabolic pathway in *Corynebacterium glutamicum*. Appl Environ Microbiol 72:3418–3428

Kerbs A, Mindt M, Schwardmann L, Wendisch VF (2021) Sustainable production of *N*-methylphenylalanine by reductive methylamination of phenylpyruvate using engineered *Corynebacterium glutamicum*. Microorganisms 9:824

Kerr JS, Ruppert CL, Tozzi CA et al (1987) Reduction of chronic hypoxic pulmonary hypertension in the rat by an inhibitor of collagen production. Am Rev Respir Dis 135:300–306
Kholy ER, Eikmanns BJ, Gutmann M, Sahm H (1993) Glutamate dehydrogenase is not essential for glutamate formation by *Corynebacterium glutamicum*. Appl Environ Microbiol 59:2329–2331
Kiefer P, Heinzle E, Wittmann C (2002) Influence of glucose, fructose and sucrose as carbon sources on kinetics and stoichiometry of lysine production by *Corynebacterium glutamicum*. J Ind Microbiol Biotechnol 28:338–343
Kim HT, Baritugo K-A, Oh YH et al (2018) Metabolic engineering of *Corynebacterium glutamicum* for the high-level production of cadaverine that can be used for the synthesis of biopolyamide 510. ACS Sustainable Chem Eng 6:5296–5305
Kimura T, Sugahara I, Hanai K, Tonomura Y (1992) Purification and characterization of γ-glutamylmethylamide synthetase from *Methylophaga* sp. AA-30. Biosci Biotechnol Biochem 56:708–711
Kind S, Jeong WK, Schröder H, Wittmann C (2010) Systems-wide metabolic pathway engineering in *Corynebacterium glutamicum* for bio-based production of diaminopentane. Metab Eng 12: 341–351
Kind S, Kreye S, Wittmann C (2011) Metabolic engineering of cellular transport for overproduction of the platform chemical 1,5-diaminopentane in *Corynebacterium glutamicum*. Metab Eng 13: 617–627
Ko YJ, Kim M, You SK et al (2021) Animal-free heme production for artificial meat in *Corynebacterium glutamicum* via systems metabolic and membrane engineering. Metab Eng 66:217–228
Koga R, Miyoshi Y, Sakaue H et al (2017) Mouse D-amino-acid oxidase: distribution and physiological substrates. Front Mol Biosci 4:82
Kugler P, Fröhlich D, Wendisch VF (2020) Development of a biosensor for crotonobetaine-CoA ligase screening based on the elucidation of *Escherichia coli* carnitine metabolism. ACS Synth Biol 9:2460–2471
Kugler P, Trumm M, Frese M, Wendisch VF (2021) L-Carnitine production through biosensor-guided construction of the *Neurospora crassa* biosynthesis pathway in *Escherichia coli*. Front Bioeng Biotechnol 9:671321
Kung HF, Wagner C (1970) The enzymatic synthesis of *N*-methylalanine. Biochim Biophys Acta 201:513–516
Kurata HT, Marton LJ, Nichols CG (2006) The polyamine binding site in inward rectifier K+ channels. J Gen Physiol 127:467–480
Kwak DH, Lim HG, Yang J et al (2017) Synthetic redesign of *Escherichia coli* for cadaverine production from galactose. Biotechnol Biofuels 10:20
Lamont IL, Martin LW, Sims T et al (2006) Characterization of a gene encoding an acetylase required for pyoverdine synthesis in *Pseudomonas aeruginosa*. J Bacteriol 188:3149–3152
Laslo T, von Zaluskowski P, Gabris C et al (2012) Arabitol metabolism of *Corynebacterium glutamicum* and its regulation by AtlR. J Bacteriol 194:941–955
Lau MK (2018) Recombinant bacterial cells producing (S)-2-amino-6-hydroxypimelate. US9890405B2
Lee HL, Kim S-Y, Kim EJ et al (2019) Synthesis of methylated anthranilate derivatives using engineered strains of *Escherichia coli*. J Microbiol Biotechnol 29:839–844
Lee J-H, Wendisch VF (2017) Production of amino acids—genetic and metabolic engineering approaches. Bioresour Technol 245:1575–1587
Lei M, Peng X, Sun W et al (2021) Nonsterile L-lysine fermentation using engineered phosphite-grown *Corynebacterium glutamicum*. ACS Omega 6:10160–10167
Lim HG, Seo SW, Jung GY (2013) Engineered *Escherichia coli* for simultaneous utilization of galactose and glucose. Bioresour Technol 135:564–567
Liu L-K, Abdelwahab H, Martin Del Campo JS et al (2016) The structure of the antibiotic deactivating, *N*-hydroxylating rifampicin monooxygenase*. J Biol Chem 291:21553–21562

Lubitz D, Jorge JMP, Pérez-García F et al (2016) Roles of export genes *cgmA* and *lysE* for the production of L-arginine and L-citrulline by *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 100:8465–8474

Luka Z, Mudd SH, Wagner C (2009) Glycine *N*-methyltransferase and regulation of *S*-adenosylmethionine levels. J Biol Chem 284:22507–22511

Ma H, Fan X, Cai N et al (2020) Efficient fermentative production of L-theanine by *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 104:119–130

Ma W, Wang J, Li Y et al (2018) Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) co-produced with L-isoleucine in *Corynebacterium glutamicum* WM001. Microb Cell Factories 17:93

Meiswinkel TM, Gopinath V, Lindner SN et al (2013b) Accelerated pentose utilization by *Corynebacterium glutamicum* for accelerated production of lysine, glutamate, ornithine and putrescine. Microb Biotechnol 6:131–140

Meiswinkel TM, Rittmann D, Lindner SN, Wendisch VF (2013a) Crude glycerol-based production of amino acids and putrescine by *Corynebacterium glutamicum*. Bioresour Technol 145:254–258

Michael JP (2017) Chapter one—acridone alkaloids. In: Knölker H-J (ed) The alkaloids: chemistry and biology. Academic Press, pp 1–108

Mihara H, Muramatsu H, Kakutani R et al (2005) *N*-methyl-L-amino acid dehydrogenase from *Pseudomonas putida*. FEBS J 272:1117–1123

Mimitsuka T, Sawai H, Hatsu M, Yamada K (2007) Metabolic engineering of *Corynebacterium glutamicum* for cadaverine fermentation. Biosci Biotechnol Biochem 71:2130–2135

Mindt M, Hannibal S, Heuser M et al (2019a) Fermentative production of *N*-alkylated glycine derivatives by recombinant *Corynebacterium glutamicum* using a mutant of imine reductase DpkA from *Pseudomonas putida*. Front Bioeng Biotechnol 7:232

Mindt M, Heuser M, Wendisch VF (2019b) Xylose as preferred substrate for sarcosine production by recombinant *Corynebacterium glutamicum*. Bioresour Technol 281:135–142

Mindt M, Risse JM, Gruß H et al (2018) One-step process for production of *N*-methylated amino acids from sugars and methylamine using recombinant *Corynebacterium glutamicum* as biocatalyst. Sci Rep 8:12895

Mindt M, Walter T, Kugler P, Wendisch VF (2020) Microbial engineering for production of *N*-functionalized amino acids and amines. Biotechnol J 15:1900451

Mir R, Jallu S, Singh TP (2015) The shikimate pathway: review of amino acid sequence, function and three-dimensional structures of the enzymes. Crit Rev Microbiol 41:172–189

Miscevic D, Mao J-Y, Kefale T et al (2021) Strain engineering for high-level 5-aminolevulinic acid production in *Escherichia coli*. Biotechnol Bioeng 118:30–42

Moritzer A-C, Minges H, Prior T et al (2019) Structure-based switch of regioselectivity in the flavin-dependent tryptophan 6-halogenase Thal. J Biol Chem 294:2529–2542

Motomura K, Sano K, Watanabe S et al (2018) Synthetic phosphorus metabolic pathway for biosafety and contamination management of cyanobacterial cultivation. ACS Synth Biol 7:2189–2198

Mugford ST, Louveau T, Melton R et al (2013) Modularity of plant metabolic gene clusters: a trio of linked genes that are collectively required for acylation of triterpenes in oat. Plant Cell 25:1078–1092

Mügge C, Heine T, Baraibar AG et al (2020) Flavin-dependent N-hydroxylating enzymes: distribution and application. Appl Microbiol Biotechnol 104:6481–6499

Muramatsu H, Mihara H, Goto M et al (2005) A new family of NAD(P)H-dependent oxidoreductases distinct from conventional Rossmann-fold proteins. J Biosci Bioeng 99:541–547

Muramatsu H, Mihara H, Kakutani R et al (2004) Enzymatic synthesis of *N*-methyl-L-phenylalanine by a novel enzyme, *N*-methyl-L-amino acid dehydrogenase, from *Pseudomonas putida*. Tetrahedron Asymmetry 15:2841–2843

Na D, Yoo SM, Chung H et al (2013) Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. Nat Biotechnol 31:170–174

Nærdal I, Pfeifenschneider J, Brautaset T, Wendisch VF (2015) Methanol-based cadaverine production by genetically engineered *bacillus methanolicus* strains. Microb Biotechnol 8: 342–350

Nakagawa M (1970) Constituents in tea leaf and their contribution to the taste of green tea liquor. Jpn Agric Res Q 5:43–47

Narukawa M, Morita K, Hayashi Y (2008) L-theanine elicits an umami taste with inosine 5-′-monophosphate. Biosci Biotechnol Biochem 72:3015–3017

Neilands JB (1993) Siderophores. Arch Biochem Biophys 302:1–3

Neubauer PR, Widmann C, Wibberg D et al (2018) A flavin-dependent halogenase from metagenomic analysis prefers bromination over chlorination. PLoS One 13:e0196797

Nguyen AQD, Schneider J, Reddy GK, Wendisch VF (2015b) Fermentative production of the diamine putrescine: system metabolic engineering of *Corynebacterium glutamicum*. Meta 5: 211–231

Nguyen AQD, Schneider J, Wendisch VF (2015a) Elimination of polyamine *N*-acetylation and regulatory engineering improved putrescine production by *Corynebacterium glutamicum*. J Biotechnol 201:75–85

Nikkei (2016) Mitsui unit to beef up US output of amino acid for feed. In: Nikkei Asia. https://asia.nikkei.com/Business/Mitsui-unit-to-beef-up-US-output-of-amino-acid-for-feed. Accessed 6 Aug 2021

Ning Y, Wu X, Zhang C et al (2016) Pathway construction and metabolic engineering for fermentative production of ectoine in *Escherichia coli*. Metab Eng 36:10–18

Nishizawa T, Aldrich CC, Sherman DH (2005) Molecular analysis of the rebeccamycin L-amino acid oxidase from *Lechevalieria aerocolonigenes* ATCC 39243. J Bacteriol 187:2084–2092

Noh M, Yoo SM, Kim WJ, Lee SY (2017) Gene expression knockdown by modulating synthetic small RNA expression in *Escherichia coli*. Cell Syst 5:418–426.e4

Onaka H, Taniguchi S, Igarashi Y, Furumai T (2003) Characterization of the biosynthetic gene cluster of rebeccamycin from *Lechevalieria aerocolonigenes* ATCC 39243. Biosci Biotechnol Biochem 67:127–138

Ou X-Y, Wu X-L, Peng F et al (2019) Metabolic engineering of a robust *Escherichia coli* strain with a dual protection system. Biotechnol Bioeng 116:3333–3348

Pastor JM, Salvador M, Argandoña M et al (2010) Ectoines in cell stress protection: uses and biotechnological production. Biotechnol Adv 28:782–801

Pérez-García F, Brito LF, Wendisch VF (2019) Function of L-pipecolic acid as compatible solute in *Corynebacterium glutamicum* as basis for its production under hyperosmolar conditions. Front Microbiol 10:340

Pérez-García F, Burgardt A, Kallman DR et al (2021) Dynamic co-cultivation process of *Corynebacterium glutamicum* strains for the fermentative production of riboflavin. Fermentation 7:11

Pérez-García F, Jorge JMP, Dreyszas A et al (2018) Efficient production of the dicarboxylic acid glutarate by *Corynebacterium glutamicum* via a novel synthetic pathway. Front Microbiol 9: 2589

Pérez-García F, Peters-Wendisch P, Wendisch VF (2016) Engineering *Corynebacterium glutamicum* for fast production of L-lysine and L-pipecolic acid. Appl Microbiol Biotechnol 100:8075–8090

Pérez-García F, Risse JM, Friehs K, Wendisch VF (2017a) Fermentative production of L-pipecolic acid from glucose and alternative carbon sources. Biotechnol J 12:1600646

Pérez-García F, Ziert C, Risse JM, Wendisch VF (2017b) Improved fermentative production of the compatible solute ectoine by *Corynebacterium glutamicum* from glucose and alternative carbon sources. J Biotechnol 258:59–68

Peters-Wendisch PG, Schiel B, Wendisch VF et al (2001) Pyruvate carboxylase is a major bottleneck for glutamate and lysine production by *Corynebacterium glutamicum*. J Mol Microbiol Biotechnol 3:295–300

Petrossian TC, Clarke SG (2011) Uncovering the human methyltransferasome. Mol Cell Proteomics 10:M110.000976
Poiani GJ, Kemnitzer JE, Fox JD et al (1997) Polymeric carrier of proline analogue with antifibrotic effect in pulmonary vascular remodeling. Am J Respir Crit Care Med 155:1384–1390
Prabowo CPS, Shin JH, Cho JS et al (2020) Microbial production of 4-amino-1-butanol, a four-carbon amino alcohol. Biotechnol Bioeng 117:2771–2780
Prell C, Busche T, Rückert C et al (2021) Adaptive laboratory evolution accelerated glutarate production by *Corynebacterium glutamicum*. Microb Cell Factories 20:97
Qian Z-G, Xia X-X, Lee SY (2009) Metabolic engineering of *Escherichia coli* for the production of putrescine: a four carbon diamine. Biotechnol Bioeng 104:651–662
Qian Z-G, Xia X-X, Lee SY (2011) Metabolic engineering of *Escherichia coli* for the production of cadaverine: a five carbon diamine. Biotechnol Bioeng 108:93–103
Reshamwala SMS, Pagar SK, Velhal VS et al (2014) Construction of an efficient *Escherichia coli* whole-cell biocatalyst for D-mannitol production. J Biosci Bioeng 118:628–631
Rhee HJ, Kim E-J, Lee JK (2007) Physiological polyamines: simple primordial stress molecules. J Cell Mol Med 11:685–703
Richter AA, Mais C-N, Czech L et al (2019) Biosynthesis of the stress-protectant and chemical chaperon ectoine: biochemistry of the transaminase EctB. Front Microbiol 10:2811
Rioz-Martínez A, Kopacz M, de Gonzalo G et al (2011) Exploring the biocatalytic scope of a bacterial flavin-containing monooxygenase. Org Biomol Chem 9:1337–1341
Rohde B, Hans J, Martens S et al (2008) Anthranilate *N*-methyltransferase, a branch-point enzyme of acridone biosynthesis. Plant J 53:541–553
Sasikumar K, Hannibal S, Wendisch VF, Nampoothiri KM (2021) Production of biopolyamide precursors 5-amino valeric acid and putrescine from rice straw hydrolysate by engineered *Corynebacterium glutamicum*. Front Bioeng Biotechnol 9:635509
Schneider J, Eberhardt D, Wendisch VF (2012) Improving putrescine production by *Corynebacterium glutamicum* by fine-tuning ornithine transcarbamoylase activity using a plasmid addiction system. Appl Microbiol Biotechnol 95:169–178
Schneider J, Wendisch VF (2010) Putrescine production by engineered *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 88:859–868
Schnepel C, Minges H, Frese M, Sewald N (2016) A high-throughput fluorescence assay to determine the activity of tryptophan halogenases. Angew Chem Int Ed 55:14159–14163
Schnepel C, Sewald N (2017) Enzymatic halogenation: a timely strategy for regioselective C−H activation. Chemistry 23:12064–12086
Schultenkämper K, Brito LF, Wendisch VF (2020) Impact of CRISPR interference on strain development in biotechnology. Biotechnol Appl Biochem 67:7–21
Schwardmann LS, Dransfeld AK, Schäffer T, Wendisch VF (2022) Metabolic engineering of *Corynebacterium glutamicum* for sustainable production of the aromatic dicarboxylic acid dipicolinic acid. Microorganisms 10:730
Seltmann G, Holst O (2013) The bacterial cell wall. Springer Science & Business Media
Sgobba E, Stumpf AK, Vortmann M et al (2018) Synthetic *Escherichia coli-Corynebacterium glutamicum* consortia for L-lysine production from starch and sucrose. Bioresour Technol 260: 302–310
Sgobba E, Wendisch VF (2020) Synthetic microbial consortia for small molecule production. Curr Opin Biotechnol 62:72–79
Shaw AJ, Lam FH, Hamilton M et al (2016) Metabolic engineering of microbial competitive advantage for industrial fermentation processes. Science 353:583–586
Sheng Q, Wu X, Jiang Y et al (2021) Highly efficient biosynthesis of L-ornithine from mannitol by using recombinant *Corynebacterium glutamicum*. Bioresour Technol 327:124799
Shim J, Shin Y, Lee I, Kim SY (2017) L-Methionine production. Adv Biochem Eng Biotechnol 159: 153–177

Soe CZ, Pakchung AAH, Codd R (2012) Directing the biosynthesis of putrebactin or desferrioxamine B in *Shewanella putrefaciens* through the upstream inhibition of ornithine decarboxylase. Chem Biodivers 9:1880–1890

Stella RG, Wiechert J, Noack S, Frunzke J (2019) Evolutionary engineering of *Corynebacterium glutamicum*. Biotechnol J 14:e1800444

Strijbis K, Vaz FM, Distel B (2010) Enzymology of the carnitine biosynthesis pathway. IUBMB Life 62:357–362

Sturgill G, Rather PN (2004) Evidence that putrescine acts as an extracellular signal required for swarming in *Proteus mirabilis*. Mol Microbiol 51:437–446

Sugiyama Y, Nakamura A, Matsumoto M et al (2016) A novel putrescine exporter SapBCDF of *Escherichia coli*. J Biol Chem 291:26343–26351

Suzuki H, Kumagai H (2002) Autocatalytic processing of gamma-glutamyltranspeptidase. J Biol Chem 277:43536–43543

Tabor CW, Tabor H (1985) Polyamines in microorganisms. Microbiol Rev 49:81–99

Tachiki T (1986) Glutamine synthetase of some fermentation bacteria: function and application. Korean J Microbiol Biotechnol:506–508

Takahashi C, Shirakawa J, Tsuchidate T et al (2012) Robust production of gamma-amino butyric acid using recombinant *Corynebacterium glutamicum* expressing glutamate decarboxylase from *Escherichia coli*. Enzym Microb Technol 51:171–176

Takashi T, Takeshi Y, Masashi U et al (1996) Purification and some properties of glutaminase from *Pseudomonas nitroreducens* IFO 12694. Biosci Biotechnol Biochem 60:1160–1164

Takatsuma Y, Kamio Y (2004) Molecular dissection of the *Selenomonas ruminantium* cell envelope and lysine decarboxylase involved in the biosynthesis of a polyamine covalently linked to the cell wall peptidoglycan layer. Biosci Biotechnol Biochem 68:1–19

Tan S, Shi F, Liu H et al (2020) Dynamic control of 4-hydroxyisoleucine biosynthesis by modified L-isoleucine biosensor in recombinant *Corynebacterium glutamicum*. ACS Synth Biol 9:2378–2389

Täuber S, Blöbaum L, Wendisch VF, Grünberger A (2021) Growth response and recovery of *Corynebacterium glutamicum* colonies on single-cell level upon defined pH stress pulses. Front Microbiol 12:711893

Tesch M, Eikmanns B, de Graaf A, Sahm H (1998) Ammonia assimilation in *Corynebacterium glutamicum* and a glutamate dehydrogenase-deficient mutant. Biotechnol Lett 20:953–957

Ting W-W, Ng I-S (2020) Metabolic manipulation through CRISPRi and gene deletion to enhance cadaverine production in *Escherichia coli*. J Biosci Bioeng 130:553–562

Tsai G, Lane H-Y, Yang P et al (2004) Glycine transporter I inhibitor, *N*-methylglycine (sarcosine), added to antipsychotics for the treatment of schizophrenia. Biol Psychiatry 55:452–456

Tsuge Y, Matsuzawa H (2021) Recent progress in production of amino acid-derived chemicals using *Corynebacterium glutamicum*. World J Microbiol Biotechnol 37:49

van Pée KH, Hölzer M (1999) Specific enzymatic chlorination of tryptophan and tryptophan derivatives. Adv Exp Med Biol 467:603–609

VanDrisse CM, Escalante-Semerena JC (2019) Protein acetylation in bacteria. Annu Rev Microbiol 73:111–132

Veldmann KH, Dachwitz S, Risse JM et al (2019b) Bromination of L-tryptophan in a fermentative process with *Corynebacterium glutamicum*. Front Bioeng Biotechnol 7:219

Veldmann KH, Minges H, Sewald N et al (2019a) Metabolic engineering of *Corynebacterium glutamicum* for the fermentative production of halogenated tryptophan. J Biotechnol 291:7–16

von Kamp A, Klamt S (2017) Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. Nat Commun 8:15956

Vortmann M, Stumpf AK, Sgobba E et al (2021) A bottom-up approach towards a bacterial consortium for the biotechnological conversion of chitin to L-lysine. Appl Microbiol Biotechnol 105:1547–1561

Vrljic M, Sahm H, Eggeling L (1996) A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. Mol Microbiol 22:815–826
Wahlang B, Falkner KC, Cave MC, Prough RA (2015) Role of cytochrome P450 monooxygenase in carcinogen and chemotherapeutic drug metabolism. Adv Pharmacol 74:1–33
Walter T, Al Medani N, Burgardt A et al (2020) Fermentative *N*-Methylanthranilate production by engineered *Corynebacterium glutamicum*. Microorganisms 8:866
Wang L, Li G, Deng Y (2020) Diamine biosynthesis: research progress and application prospects. Appl Environ Microbiol 86:e01972-20
Wang Q, Zhang J, Al Makishah NH et al (2021) Advances and perspectives for genome editing tools of *Corynebacterium glutamicum*. Front Microbiol 12:654058
Wendisch VF (2020) Metabolic engineering advances and prospects for amino acid production. Metab Eng 58:17–34
Wendisch VF, Bott M, Eikmanns BJ (2006) Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for biotechnological production of organic acids and amino acids. Curr Opin Microbiol 9:268–274
Wendisch VF, Eberhardt D, Herbst M, Vold Korgaard Jensesn J (2016a) Biotechnological production of amino acids and nucleotides. In: Bicas J, Maróstica M Jr, Pastore G (eds) Biotechnological production of natural ingredients for food industry. Bentham Science Publishers, pp 60–163
Wendisch VF, Jorge JMP, Pérez-García F, Sgobba E (2016b) Updates on industrial production of amino acids using *Corynebacterium glutamicum*. World J Microbiol Biotechnol 32:105
Wendisch VF, Mindt M, Pérez-García F (2018) Biotechnological production of mono- and diamines using bacteria: recent progress, applications, and perspectives. Appl Microbiol Biotechnol 102:3583–3594
Werdehausen R, Kremer D, Brandenburger T et al (2012) Lidocaine metabolites inhibit glycine transporter 1: a novel mechanism for the analgesic action of systemic lidocaine? Anesthesiology 116:147–158
Wieschalka S, Blombach B, Eikmanns BJ (2012) Engineering *Corynebacterium glutamicum* for the production of pyruvate. Appl Microbiol Biotechnol 94:449–459
Williams J, Sergi D, McKune AJ et al (2019) The beneficial health effects of green tea amino acid L-theanine in animal models: promises and prospects for human trials. Phytother Res 33:571–583
Wolf S, Becker J, Tsuge Y et al (2021) Advances in metabolic engineering of *Corynebacterium glutamicum* to produce high-value active ingredients for food, feed, human health, and well-being. Essays Biochem 65:197–212
Wördemann R, Wiefel L, Wendisch VF, Steinbüchel A (2021) Incorporation of alternative amino acids into cyanophycin by different cyanophycin synthetases heterologously expressed in *Corynebacterium glutamicum*. AMB Express 11:55
Wu X-Y, Guo X-Y, Zhang B et al (2019) Recent advances of L-ornithine biosynthesis in metabolically engineered *Corynebacterium glutamicum*. Front Bioeng Biotechnol 7:440
Xue C, Hsu K-M, Chiu C-Y et al (2021) Fabrication of bio-based polyamide 56 and antibacterial nanofiber membrane from cadaverine. Chemosphere 266:128967
Ying H, Tao S, Wang J et al (2017) Expanding metabolic pathway for de novo biosynthesis of the chiral pharmaceutical intermediate L-pipecolic acid in *Escherichia coli*. Microb Cell Factories 16:52
Yu X, Shi F, Liu H et al (2021) Programming adaptive laboratory evolution of 4-hydroxyisoleucine production driven by a lysine biosensor in *Corynebacterium glutamicum*. AMB Express 11:66
Zanger UM, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacol Ther 138:103–141

Zhang B, Gao G, Chu X-H, Ye B-C (2019) Metabolic engineering of *Corynebacterium glutamicum* S9114 to enhance the production of L-ornithine driven by glucose and xylose. Bioresour Technol 284:204–213

Zhao D, Yu Y, Shen Y et al (2019) Melatonin synthesis and function: evolutionary history in animals and plants. Front Endocrinol 10:249

Zhou X, Wang F, Hu H et al (2011) Assessment of sustainable biomass resource for energy use in China. Biomass Bioenergy 35:1–11

Zhu L, Mack C, Wirtz A et al (2020) Regulation of γ-aminobutyrate (GABA) utilization in *Corynebacterium glutamicum* by the PucR-type transcriptional regulator GabR and by alternative nitrogen and carbon sources. Front Microbiol 11:544045

Strategies for Improving Biotherapeutic Protein Production in Microbial Cell Factories

Priyanka Priyanka, Somesh Mishra, and Anurag S. Rathore

Contents

Abstract Recombinant protein manufacturing is experiencing a global transformation across multiple verticals for achieving sustainable production. Since there is no single host that is perfect for expression of each protein, different expression systems like yeast, bacteria, and mammalian cells have been used. The selection of a protein expression system is a key for protein manufacturing. If glycosylation and post-translational modification are desired in the manufacture of proteins, then a eukaryotic or mammalian expression system is preferred, or else bacteria are preferred hosts due to their simplicity, low cost, and shorter processing time. Thus, considering the advantages of microbial cell factories, the biopharmaceutical industry is always looking for ways to improve process efficiency, accelerate drug development, and lower production costs. Varied molecular biology approaches, ranging from clone development, cell line engineering, and process development all have a role to play in this endeavor. This chapter aims to address the topic of manufacturing

P. Priyanka · S. Mishra · A. S. Rathore (✉)
Department of Chemical Engineering, Indian Institute of Technology, Hauz Khas, New Delhi, India
e-mail: asrathore@biotechcmz.com

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_4

biotherapeutics via microbial cell factories. Activities from cloning to upstream process development are described.

1 Introduction

Over the past few decades, biotherapeutics have successfully treated numerous life-threatening diseases and shown major clinical successes. They can be broadly classified as proteins and peptides that have revolutionized the treatment of otherwise difficult to cure diseases, such as cancer, rare genetic disorder, and autoimmune diseases (Balbás and Lorence 2004; Kesik-Brodacka 2018; Das et al. 2020). Different types of biotherapeutics include enzymes, non-immune proteins, monoclonal antibodies (mAb), crystallizable antibody fragment (Fc), and antigen-binding fragment (Fab). The "Big 6" mAbs, namely Humira, Enbrel, Remicade, Avastin, Rituxan, and Herceptin, have propelled the biotech industry over the past decade (Zhong et al. 2011). The total number of biotherapeutic products approved in 1990 was around 9, and the number rose to 112 in 2018 (Walsh 2018). The pharmaceuticals global market is expected to increase by $389 billion by 2024 from around $237.2 billion in 2018. From 2019 to 2024, a staggering compound annual growth rate (CAGR) of 8.59% is expected (O'Flaherty et al. 2020; Reh 2020). However, the high cost with low production efficiency imposes a significant challenge for devising an economically acceptable manufacturing process (Palomares et al. 2004).

Mostly, FDA- and EMA-approved biotherapeutics are recombinant proteins produced via microbial hosts, either yeast or bacteria. The microbial platform for protein production is preferred due to many factors like the ease in manufacturing, low production cost, well-known genome (in case of *E. coli* and *S. cerevisiae*), and ease of genome editing or recombination (Jozala et al. 2016; Tripathi and Shrivastava 2019). Additionally, the genetically engineered host offers the advantages of enhanced solubility, stability, and post-translation modification attributes that are absent in wild-type microbes. Increased interest in these protein drugs is due to their specificity, potency, stability, without generation of any harmful waste (Zhong et al. 2011; Ueda 2014).

After establishing a cloning strategy and selection of a recombinant host, upstream process development and optimization are performed for improving product yield (Ramalingam et al. 2007). The product titer and overall process efficiency can be increased by utilizing recombinant DNA technology, modifying host strain and media formulations, and applying process controls. These strategies have been successfully employed to increase protein production yield with minimal impact on production costs (Gronemeyer et al. 2014). Prokaryotic or eukaryotic cells are cultivated in bioreactors to produce biotherapeutic products. A spectrum of operational variables, such as pH, temperature, substrate concentration, dissolved oxygen (DO), agitation rate, and system homogeneity, collectively control the final productivity, quality, and consistency of the product. The manufacturing process is known to impact the heterogeneity of the final product (Harms et al. 2002; Joshi and Rathore 2020).

2 Heterologous Protein Expression

Recombinant proteins are expressed in various hosts, such as bacteria, yeast, mammalian cells, insect cells, transgenic animals, and transgenic plants. Among these, *E. coli* is preferred and is one of the most exploited prokaryotic hosts for manufacturing of recombinant proteins, in view of considerations related to the cost, time, convenience, and safety (Tripathi and Shrivastava 2019). Despite its many benefits, the expression of complex biotherapeutic in *E. coli* poses many issues, such as glycosylation and insufficiency in proteolytic processing, protein maturation, and disulfide bonds formation. The general cloning and expression flow chart for heterologous protein is shown in Fig. 1.

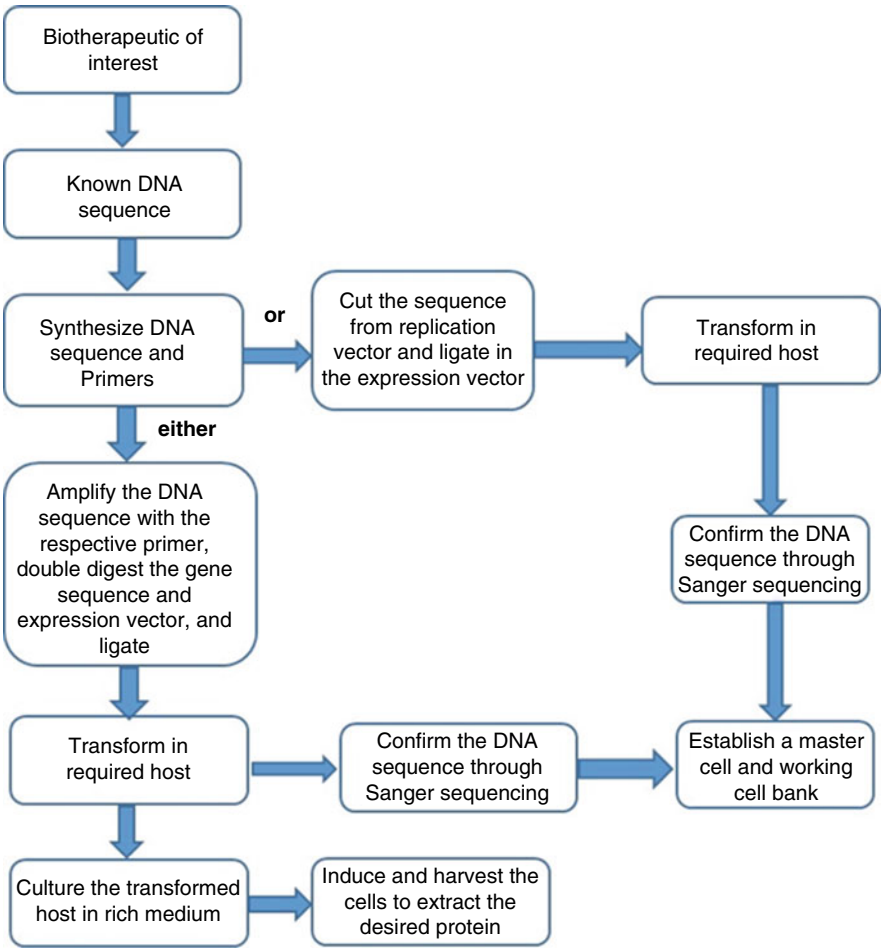


Fig. 1 General protocol for heterologous protein expression (adapted from Walsh 2014)

3 Role of Molecular Biology

Recombinant protein production via microbial hosts has revolutionized the therapeutic protein market. The upstream consideration is presented in Fig. 2. After the cloning strategy is finalized, appropriate host and vector are selected (Rosano and Ceccarelli 2014; Rosano et al. 2019). Both transcriptional and translational aspects of protein production are affected by optimal configuration of a series of genetic elements in a well-designed expression vector as depicted in Fig. 2. Additionally, the inclusion of the origin of replication determines the number of vector copy and selectable markers, i.e., antibiotic resistance marker or substrate, with enzymatic auxotrophy provides a phenotypic selection. There are, however, several concerns related to the extensive use of plasmid-based overexpression, like genetic instability, resulting from segregational instability, structural instability, and allele segregation of plasmid-based overexpression. To circumvent the limitations of plasmid-based overexpression, integrating the target genes into the host chromosome is a better technique. To achieve chromosomal integration in *E. coli*, homologous recombination, site-specific recombination, and transposon-mediated gene transposition are frequently utilized (Gu et al. 2015). Furthermore, antibiotics or other selective agents must be used to maintain the presence of plasmids in host cells, which increases the overall bioprocess cost and raises environmental problems. In the last decade, auxotrophic strains have been used to alleviate such issues. As a consequence, auxotrophs that require the inclusion of non-standard or artificial amino acids in order to produce essential proteins have been devised.

3.1 Cloning Strategies

Though gene cloning is straightforward in microbial systems, cloning of complex biotherapeutics, such as mAb, Fab, F(ab)2, scFv, membrane proteins, and other high molecular weight proteins, is non-trivial. If the protein is a humanized Fab or a mAb, then either multiple gene cloning under a single promoter with internal ribosome entry site (IRES) and poly adenine (PA) sites or the vector with more than one promoter is helpful (Fig. 3) (Gupta and Shukla 2017). Hence, the nature of the protein decides the design of the complete expression unit. In some cases, changing orientation of the light and heavy chain and cloning under unique promoter can facilitate enhanced and equal expression of the antibody fragment (Priyanka and Rathore 2021). Researchers have tried to increase the number of target gene copies in the plasmid itself to enhance protein synthesis. For instance, LTNF, which is just an eleven amino acid ribosomal peptide, has been cloned as an eight unit concatemer under a single promotor with tryptophan at the C terminal to generate monomers (Komives et al. 2017). If the protein is humanized and its starting codon is not methionine, then methionine removal strategy is of paramount importance. Table 1 summarizes all reported strategies for making an innate N terminus.

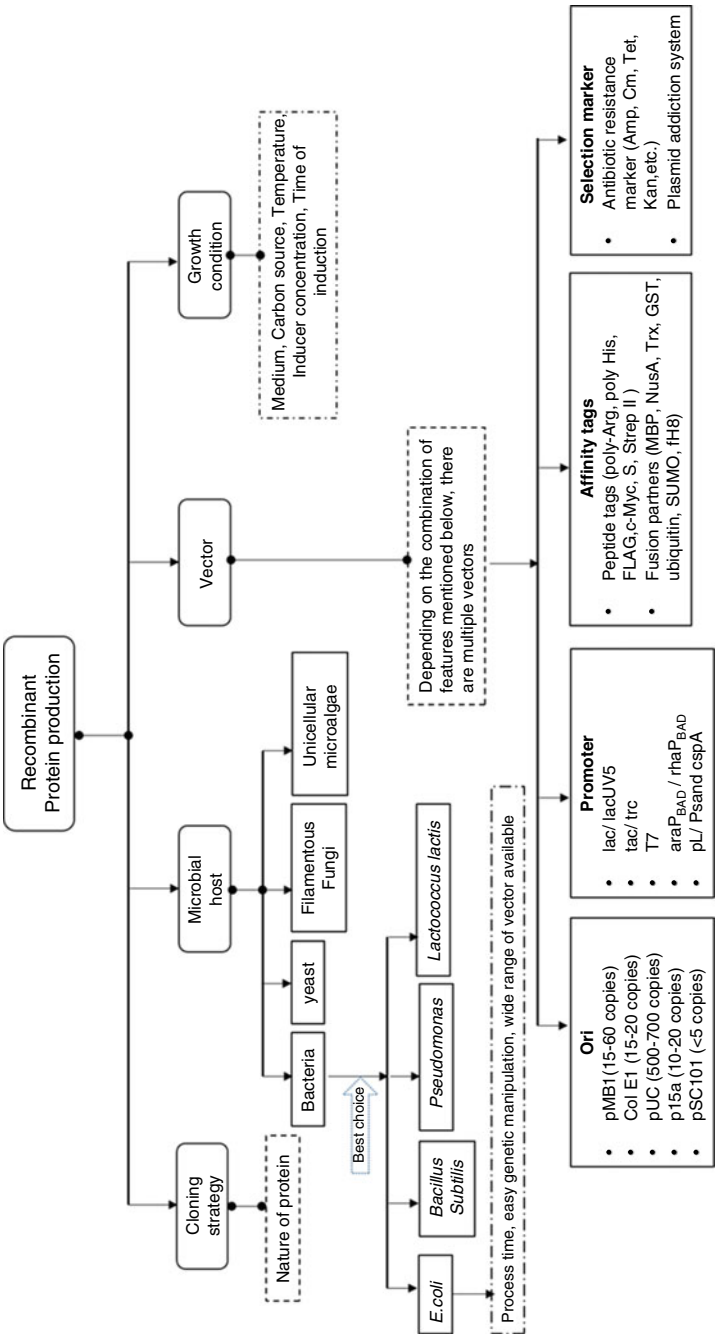


Fig. 2 Flowchart depicting upstream considerations for selecting and optimizing host cell, vector, and growth condition to produce recombinant proteins via microbial expression system (adapted from Sørensen and Mortensen 2005a, b)

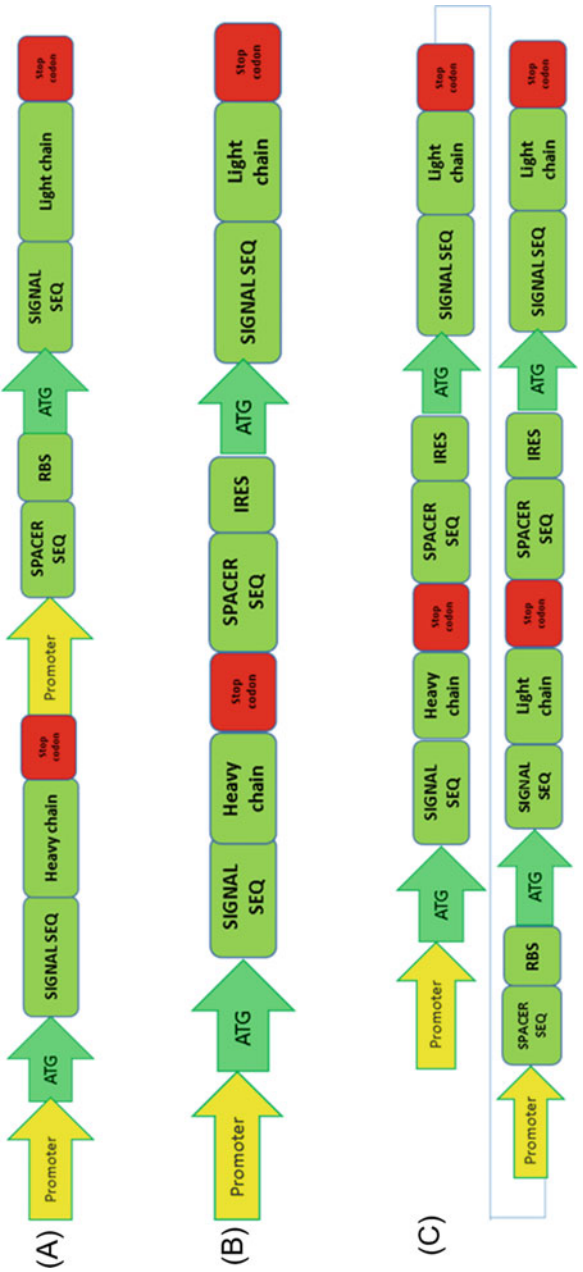


Fig. 3 Antibody fragment expression cassettes. (**a**) dual promotor strategy, (**b**) cloning under single promotor, and (**c**) multiple gene copy cloning

Table 1 Strategies for N terminal methionine removal

Strategies for methionine removal	Approaches	Prerequisites	References
Cyanogen bromide	Cleaves methionine in highly acidic condition	Method is restricted to proteins having zero internal methionine residues	Boix et al. (1996)
Protease specific sites	Enzymatic removal	N terminal protease specific sites like factor Xa, enterokinase, and cathepsin C	Belagaje et al. (1997)
Aminopeptidase	Enzymatic removal	N terminal protease specific methionine removal	Notomista et al. (1999), Shapiro et al. (1988)
Secretory signal sequences	Signal sequence peptidase	N terminal secretory signal sequence	Huang et al. (1998)

3.2 Codon Optimization

Once the expression unit is set, the next important step is to optimize the DNA sequence codon as per the host choice via algorithms, such as GeneOptimizer and IDT codon optimizer (Raab et al. 2010). The difference in codon usage may result in translational errors that include amino acid substitution, frameshift mutation, premature translational termination, or low expression (Rosano and Ceccarelli 2009). In the case of *E. coli*, for heterologous protein expression, some of the rare codons are AGA, AGG, CGA, CGG (arginine), GGA (glycine), AAG (lysine), AUA (isoleucine), CUA (leucine), and CCC (proline). Researchers have found that the presence of CGG rare codon in p27 protease realm of herpes simplex virus resulted in the production of the target protein having a molecular weight with 3 kDa higher in *E. coli* due to frameshift mutation (Macauley-Patrick et al. 2005). In the same study, glutamine amino acid substitution has been observed instead of arginine due to misinterpretation of CGG as CAG. In yet another study, it has been established that codon optimization resulted in several-fold enhancement of protein expression (Wu et al. 2004). Not just codon optimization, researchers have also tried co-expressing plasmid carrying tRNA for the rare codon. Rosetta and BL21 codon plus are among those expression hosts, which could allow high efficiency and perfect expression of rare codon carrying protein (Kumar et al. 2005; Tegel et al. 2010). Combinatorial effect of multiparameter RNA stabilization and codon optimization has been proven effective in enhancing recombinant protein expression (Fath et al. 2011; Leppek et al. 2021).

3.3 Host Selection

E. coli is one of the widely chosen, conventional platforms for producing recombinant proteins. As it is well characterized and highly studied, the doubling time of

E.coli is <20 minutes, and there is an ease in doing genetic manipulation and studying them. Further, molecular biology tools available for *E. coli* are well documented, which makes it one of the preferred hosts. However, *E. coli* can often yield incorrect folding when it comes to proteins requiring post-translational modification (Jozala et al. 2016; Tripathi and Shrivastava 2019). Also, the complex protein production via *E. coli* in its active form may require laborious processing and have low process yields. To overcome these problems, researchers have tried to genetically modify the wild-type *E. coli* strains (Table 2). Specifically, the synthetic biology approaches have been used to design, build, and test genome-reduced microbial chassis that display desirable characteristics for heterologous protein expression (Kim et al. 2020).

Advance microbiology methods have made it possible to clone and express complex biotherapeutics using microbial platform. These advancements include multiple gene cloning, strong promoter, novel vector features like strong ribosomal binding site (RBS), secretion signals, protein tags, and induction methods. A list of complex biotherapeutics from microbial origin is presented in Table 3.

It should be noted that overproduced recombinant polypeptides in *E. coli* frequently form inclusion bodies, insoluble clumps of denaturated proteins, in the cytoplasm or periplasm. Under these circumstances, protein recovery from *E. coli* cells necessitates lyzing the cells and solubilizing the proteins aggregates with strong chaotropic chemicals, then refolds the polypeptide under a suitable condition. Additional challenge is the presence of endotoxin, a toxic lipopolysaccharide found in Gram-negative bacteria's cell walls that is difficult to remove during purification. Other systems are presently being investigated as a result of the concerns with *E. coli*. The systems that are being sought are those that secrete proteins in high yield in a soluble, physiologically active form in the culture medium, avoiding inclusion bodies and associated refolding issues. Gram positive bacteria, such as Bacillus and Streptomyces, are promising hosts (Anné et al. 2014). They are commonly used in industry to produce homologous proteins for secreting extracellularly in huge quantities. However, there have been numerous bottlenecks in expressing eukaryotic protein in these hosts. Impaired translation, ineffective translocation across the plasma membrane, delayed secreted protein release from the cell wall, improper folding, and extracellular protease cleavage of the recombinant protein are all examples of issues that are faced when attempting to secrete target proteins.

3.4 Plasmids

Plasmids are essential extra-chromosomal DNA that is responsible for transferring genetic material among expression hosts. Their classification is based on replicative or expression, low copy number or high copy number, self-induced or chemically induced, and finally on the choice of the host cell—*E.coli*, yeast, or mammalian. In host-based plasmids, the replicative plasmid is the same among all hosts, but the

Table 2 Commercially available engineered *E.coli* strains for recombinant protein production

Hosts	Features	References
BL21 DE3	It is a widely used expression host, deficient in lon and omp-t proteases. This strain was constructed by inserting T7 RNA polymerase gene in a bacterial genome and is inducible by IPTG and lactose.	Raj et al. (2008)
BL21 DE3 PLyS	This is a BL21 derivative. Its additional feature is to avoid leaky expression. The pLysS plasmid, which has chloramphenicol selectable marker, produces T7 lysozyme to diminish base level expression of the gene of interest.	Miroux and Walker (1996)
BL21 Star-pLysS	This strain is known for improved mRNA stability since it has mutated RNase E, decreasing mRNA degradation and increasing protein expression. The involvement of pLysS allows tight control of expression.	Rosano and Ceccarelli (2014)
BL21-SI	This strain has salt-inducible proU promoter. Protein production is induced with 0.1–0.5 M of NaCl. It helps in improving protein solubility.	Maldonado et al. (2007)
BL21-AI	This strain is also a BL21 derivative, additionally, it has arabinose inducible promoter araBAD promoter.	Amid et al. (2011)
Tuner pLysS	This strain has pLysS plasmid, which provides tighter control in addition to the lac permease mutation, which in turn allows a homogenous level of induction.	Pell (2004)
Origami	Origami host strains are K-12 derivatives which have oxidizing environment for favoring disulfide formation since it is deficient in thioredoxin reductase (trxB) and glutathione reductase (gor) genes.	Wang et al. (2010)
Origami B	This strain is derived from a lacZY mutant of BL21 and in additional carries the same trxB/gor mutations as the original Origami strains. Hence, this strain pool offers the desirable characteristics of BL21, Tuner, and Origami hosts in one strain.	Rabhi-Essafi et al. (2007)
Rosetta	This strain has BL21 lacZY (Tuner & trade) derivatives, which are intended to cater for codon bias. It enhances the expression of eukaryotic proteins that contain *E. coli* rare codons. It has tRNAs for the codons AUA, AGG, AGA, CUA, CCC, and GGA on a compatible plasmid.	Tegel et al. (2011)
Rosetta pLysS	This strain is an improved version of Rosetta since it has tighter control overexpression.	Verma and Verma (2012)
Rosetta-gami-pLysS	Rosetta-gami host strains are Origami derivatives that combine the enhanced disulfide bond formation resulting from trxB/gor mutations along with enhanced expression of eukaryotic proteins that contain codons rarely used in *E. coli*.	Verma and Verma (2012)
Arctic Express	These cells are derived from the Stratagene BL21-Gold and help in improving the solubility. Arctic Express (DE3) also has the same features.	De Luca et al. (2015)
BL21 Codon Plus	BL21-CodonPlus-RIL is an expression host which carries extra copies of the genes that recognize the AGA/AGG (arginine), AUA (isoleucine), and CUA (leucine) rare codons in *E.coli*.	Kumar et al. (2005)

(continued)

Table 2 (continued)

Hosts	Features	References
AD494	These cells are K 12 derivative and thioredoxin reductase (trxB) mutants, which enable disulfide bond formation in the cytoplasm, enabling the potential to produce properly folded active proteins.	Lin et al. (2009)
BL21trxB	Thioredoxin reductase (trxB) mutant in protease deficient BL21 background.	Yang et al. (2004)
NovaBlue (DE3)	This is K-12 derivative strain with additional benefits like high transformation efficiency and recA and endA mutation. It results in high yields of excellent quality plasmid DNA. T7 polymerase assures high level of expression. Leaky expression is rendered due to the presence of lacIq repressor.	Dieryck et al. (2003)
BLR	This increases plasmid monomer yields and helps with stabilization of target plasmids containing repetitive sequences.	Chart et al. (2000)
C41(DE3)	This is a BL21 DE3 derivative that effectively expresses toxic and membrane proteins across all organism classes.	Sørensen and Mortensen (2005b)
C43(DE3)	BL21 DE3 derivative that is efficient in expressing toxic and membrane proteins across all classes of organisms.	Tian et al. (2009)
Lemo21 (DE3)	Tunable expression is achieved by varying the level of lysozyme (lysY) that is adjusted by the addition of L-rhamnose in the culture expressing the protein of interest. Tuning the expression level may also result in more soluble, properly folded protein for difficult soluble proteins.	Schlegel et al. (2012)
SHuffle T7	*E. coli* K12 cells derivative that expresses constitutively a chromosomal copy of the disulfide bond isomerase DsbC. DsbC promotes proper folding of proteins.	Lobstein et al. (2012)

expression plasmid is due to multiple cloning sites or polylinker sites (MCS), ribosome binding sites (RBSs), and the promoter. The general schematic expression of a vector is shown in Fig. 4.

3.4.1 Origin of Replication

It signifies a specific DNA sequence at which replication initiates. Origin of replication and plasmid size are controlling the copy number within a host cell. Figure 2 describes the list of origin of replication, based upon copy number, DnaA boxes, AT rich sequences, plasmid specific replication initiation factor (Rep), and repeating units (iterons, which are classical bacterial replicons) (Bora and Ahmed 2019).

3.4.2 Promoter

The promoter is a critical stretch of nucleotide DNA sequences specific to sigma subunit of RNA polymerase in the case of *E.coli*, in which the -10 and -35 elements

Table 3 Lists of microbial origin complex biotherapeutics

Biopharmaceutical product	Clinical indications	Approval year and innovator
rh insulin (Humulin)	Diabetes	1982 US, Eli Lilly
Interferon α2a (Roferon)	Leukemia	1986 US, Hoffmann-La-Roche
Interferon α2b (Intron)	Hepatitis, cancer, and genital warts	1986 US, Schering-Plough
Somatotropin rh growth hormone (Humatrope)	Growth hormone deficiency	1987 US, Eli Lilly
Filgrastim (Neupogen)	Neutropenia	1991 US, Amgen Inc.
Interferon β-1b (Betaferon)	Multiple sclerosis	1993 US, Schering AG
Fast acting insulin (Lispro)	Diabetes	1996 US, Eli Lilly
Reteplase (Retavase)	Acute myocardial infarction	1996 US, Roche
Interferon alfacon-1 (Infergen)	Chronic Hepatitis C	1997 US, Amgen
Interleukin 11 (Neumega: oprelvekin)	Platelet deficiency during chemotherapy: avoid thrombocytopenia	1997 US, Genetic Institute with Pfizer
OspA (Outer surface protein A fragment: LYMerix)	Lyme disease vaccine	1998 US, SmithKline Beecham
Glucagon	Hypoglycemia	1998 US, Eli Lilly
Denileukin diftitox (Ontak, combination of diphtheria toxin and interleukin 2)	T cell lymphoma	1999 US, Seragen Inc.
TNFα tasonermin (Beromun)	Soft tissue sarcoma	1999 EU, Boehringer Ingelheim
Glargine (long acting insulin: Lantus)	Diabetes	2000, US Aventis
Anakinra (modified interleukin 1 receptor antagonist: Kineret)	Rheumatoid arthritis	2001 US, Amgen
Nesiritide (Natrecor: humen B type natriuretic peptide)	Congestive heart failure	2001 US, Scios Inc.
Teriparatide (Parathyroid hormone: Forteo)	Chronic osteoporosis treatment	2002 US, Eli Lilly
Pegvisomant (Somavert: growth hormone receptor antagonist)	Acromegaly	2003 US, Pharmacia NV
Palifermin (modified keratinocyte growth factor: Kepivance)	Oral mucositis for patients undergoing chemotherapy	2004 US, Amgen, Inc.
Recombinant calcitonin salmon (Calcitonin)	Osteoporosis	2005 US, Upsher-Smith Lab
Ranibizumab (Lucentis)	Wet age related macular degeneration	2006 US, Novartis 2007 EU, Genentech

(continued)

Table 3 (continued)

Biopharmaceutical product	Clinical indications	Approval year and innovator
Human parathyroid hormone (Preotact)	Osteoporosis	2006 EU, Nycomed Denmark
Somatropin (Accretropin)	Growth hormone deficiency	2008 US, Cangene
Certolizumab pegol (Cimzia)	Crohn's disease	2008 US, 2009 EU-UCB
Nplate (Romiplostim)	Thrombocytopenia	2009 EU, Amgen EU
Rh urate oxidase, PEGylated (Krystexxa)	Chronic gout	2010 US, Savient
Rh GCSF, Filgrastim (Nivestim)	Neutropenia	2010 US, Hospira (Pfizer parent organisation)
Glucarpidase (Voraxaze)	Lowering of toxic level of methotrexate conc. in patients with impaired renal function	2012 UK, BTG International
Parathyroid hormone (Preos)	Osteoporosis, hypoparathyroidism	2013 EU, NPS Pharmaceuticals
Afrezza (Rh insulin)	Diabetes mellitus	2014 US, Mankind
Toujeo (insulin glargine)	Diabetes mellitus	2015 US, Sanofi
Abasaglar (biosimilar to previously known Lantus in EU, Abasria in EU, Basaglar in USA)	Diabetes mellitus	2015 US, Eli Lilly 2014 EU, Boehringer Ingelheim
Oncaspar (Pegaspargase: rh asparaginase)	Lymphoblastic leukemia, lymphoma	2016 EU, Baxalta Innovations
Lusduna (insulin glargine, biosimilar to Lantus)	Diabetes mellitus	2017 EU, Merck Sharp & Dohme
Admelog (rapid acting human insulin analog)	Diabetes mellitus	2017 US, Sanofi
Oxervate (cenegermin: rh nerve growth factor)	Neurotrophic keratitis	2017 EU, Dompé Farmaceutici
Myalepta (Rh leptin analog)	Some forms of lipodystrophy	2018 EU, Aegerion Pharmaceuticals
Fulphila (pegfilgrastim: pegylated rh Gcsf, biosimilar to Neulasta)	Neutropenia	2018 US, Mylan
Nivestym (filgrastim-aafi in USA, Nivestim in EU)	Neutropenia	2018 Pfizer

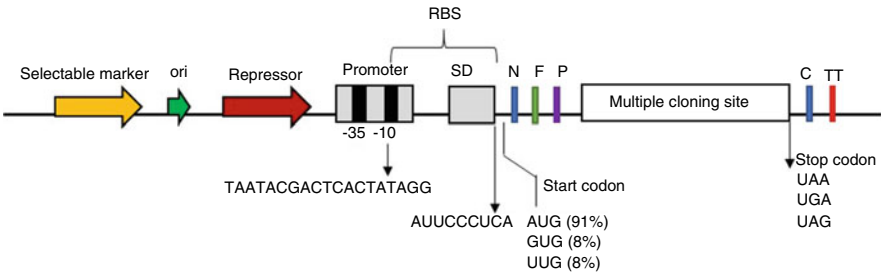


Fig. 4 Schematic representation of a prokaryotic expression vector salient features (information adapted from Stanbury et al. 2017). The example shown is of T7 promoter (URL 1). The RBS consists of Shine–Dalgarno sequence (SD) followed by AT rich 8 bp optimal spacer sequence. The 16 s ribosome 3′ end interacts with the SD sequence for initiating translation. All three start codons along with their codon biasness in *E.coli* are shown. The shown N (N terminal tag), F (Fusion tag), P (protease cleavage sites), C (C terminal tags) details are provided in Fig. 2 and Table 1. The TT (transcription terminator) main goal is to reduce leaky expression and also vector stabilization. Leaky expression denotes some basal level expression in the absence of inducer. The repressor encodes a regulatory gene that can alter the activity of the promoter. The origin of replication determines the copy number and selectable marker is for phenotypic selection (Hannig and Makrides 1998)

are named for their relative positions to the transcription initiation site, which is specified as +1 (Cooper 2000). The consensus sequence of -35 region of T7 promotor is TTGACA and TATAAT for -10 region. The nucleotides most commonly found in distinct promoters are represented by the consensus sequences. Lac promoter, a key component of lac operon acts as a staple of promoter-related research (Germán L. Rosano and Ceccarelli 2014). Lactose acts as an inducer of protein production. However, the presence of a radially metabolizable carbon source results in poor induction (such as glucose). Lac UV5 is a mutant variant of the lac promoter which can induce in the presence of glucose, though it is not entirely insensitive to catabolite repression. The most commonly used T7 promoter sequence is "TAATACGACTCACTATAGG" and -35 region consists of "TTGACA" and -10 is "TATAAT" (URL, 1). Commercially, several plasmids with weak and strongest promoters are available. Few of the promoters used in microbial plasmids are described in Fig. 2. T7 is a widely used potent promoter in microbial plasmids. An advanced and stable version of T7 promoter, also known as T7C p/p, has been explored for eliminating nonproducing cells from the culture broth. T7 promotor can be induced in the presence of IPTG or autoinduced by lactose. List of commonly used promotor is shown in Table 4 (the information is adapted from Germán L. Rosano and Ceccarelli 2014; Walsh 2014, Singh et al. 2017). In synthetic biology, promoter sequences have been used to design advanced biosensors that are able to respond to various physiological parameters such as pH, hypoxia, temperature, lactate and inflammation inducible promoter for making therapeutic bacteria (Ozdemir et al. 2018; Chowdhury et al. 2019; Chien et al. 2022).

Table 4 The most commonly used promoters and their inducers for producing recombinant proteins in *E. coli*

Promotors	Inducers	Features
T7	Isopropyl-β-d-thiogalactoside (IPTG), lactose	Well characterized, high expression
Trp	Indoleacrylic acid, tryptone starvation	Well characterized, high expression but leads to leaky expression
Tac	IPTG, lactose	A functional hybrid of Trp and lac, well characterized but leaky
lac	IPTG, lactose	Relatively low expression but strongly regulated in the presence of glucose
phoA	Phosphate starvation	Its limitation is medium availability
proU	Salt	Consumption of salts by the cells is the major issue
cspA	Reduced temperature	Growth limitation due to reduced temperature
cadA	Acidic pH	High expression but low pH could be toxic to cells though very convenient and cost effective
recA	Nalidixic acid	Non titerable
araBAD	L-Arabinose	Strongly regulated in the presence of glucose and rapidly inducible
λcI^{857}	Temperature shift to 40–42 °C	Industrially advantageous, cost effective, less downstream processing, good amount of expression
λpL	High temperature	High temperature is not suitable for cell growth
spA	Downshift in temperature to 15 degree C	Cost effective, less downstream processing, good amount of expression though insoluble in few cases

3.4.3 Ribosome Binding Site (RBS)

RBS, also referred as Shine–Dalgarno sequence in prokaryotes and as Kozak sequence in eukaryotes, is essential to initiate translation (Laursen et al. 2005). Researchers have established the relationship between translation efficiency and the secondary structure of RBS (De Smit and Van Duin 1990). Translation efficiency is enhanced by the richness of adenine and thymine (Laursen et al. 2002). Internal ribosome entry site plays a crucial role in initiating translation during bicistronic expression under a single promoter. *E. Coli* RBS, which is AAGGAGG and initiation codon, is separated by a spacer sequence of around 5–7 nucleotides that is mostly poly adenine (Chen et al. 1994). The 5′ untranslated region (UTR), which consists of RBS and spacer sequence, results in improved translational efficiency toward high-level recombinant protein expression (Park et al. 2007; Le et al. 2020). Other researchers have engineered various 5' UTR through site-specific mutations to analyze the effect on the protein expression (Park et al. 2007). A considerable change in protein expression is achieved by slight modification in 5' UTR (Seo et al. 2014).

3.4.4 Protein Tags, Molecular Chaperone, Affinity Tags, and Signal Sequences

Inclusion body generation is the most common issue faced during heterologous protein expression in microbial platforms. Protein tags and molecular chaperones are co-expressed to improve solubility. Molecular chaperones like GroEL, GroES, and DnaA-DnaK-DnaJ are energy-dependent systems that help properly fold and restrain aggregation issues (Gupta and Shukla 2017). Solubility can also be enhanced by fusing the targeted protein with the tags, including N utilization substance A (NusA), maltose binding proteins (MBP), glutathione S transferase (GST), or metal-binding proteins such as SmbA and CusF (Ahmad et al. 2018). SUMO tags and Halo tag7 have proven their effectiveness in improving the expression and solubility of around 23 human proteins in *E.coli* (Gupta and Shukla 2016). Affinity tags like His6, Strep-II, FLAG, Fh8, and Trx are mostly being utilized in detection and purification (Jia and Jeon 2016). Some of the strains derived from BL21 DE3, like C41 and C43 (Table 3) are found effective in improving solubility and expression of toxic proteins (Laursen et al. 2005; Tian et al. 2009). In expression hosts, proteins are mainly directed toward either periplasmic or extracellular translocon with the help of specific signal sequences (Fig. 2). A cytoplasmic expression followed by lysis is also a common practice. Also, it has been reported that optimization of signal sequences can help in improving the secretion to expression ratio (Humphreys et al. 2000). A novel technology called ESTEC has been developed for extracellular expression in *E. coli.* This method has proven its effectiveness in obtaining a significantly large quantity of native and correctly folded protein with a reduced load on downstream processing (Wacker et al. 2012).

3.4.5 Selectable Marker

A selectable marker must be incorporated into the plasmid backbone to select for transformants/transconjugants. In the *E.coli* system, antibiotic resistance genes are widely utilized to serve the purpose. Earlier, the ampicillin resistance gene is habitually employed in the expressions along with replication plasmids. The *bla* gene expresses periplasmic enzyme that deactivates the β-lactam ring of β-lactam antibiotics and confers resistance to ampicillin. However, with continuous secretion of β-lactamase degradation antibiotic occurs, and as a result, ampicillin is almost exhausted in a few hours (Germán L. Rosano and Ceccarelli 2014). Therefore, it is proven that antibiotics such as tetracycline, whose resistance depends on active efflux from resistant cells, are very stable (Korpimäki et al. 2003). Auxotrophic strains have also been explored and encouraged as selectable markers. Using enzymes involved in the production of a specific substrate and then restoring them into the media as a selection marker that does not rely on antibiotic resistance genes is an effective selection marker. Metal resistance genes have also been investigated as selection markers like cadmium resistance gene, arsenic resistance gene, and

chromium resistance gene (Yazdankhah et al. 2018). Most commonly explored antibiotic resistance free plasmid systems are (a) toxin/antitoxin-based systems, (b) metabolism-based systems, and (c) operator repressor titration systems. Significant efforts have also been made toward the biocontainment of genetically modified organisms to prevent their proliferation in the ecosystem (Mandell et al. 2015). Researchers are leaning toward using the plasmid addiction technique, owing to the high costs of scaling up due to antibiotics and the rise in multiple drug resistance in the population. Also, in the plasmid addiction system, synthetic auxotrophs are created in which essential enzymes are computationally redesigned to show metabolic dependency on synthetic amino acids for biocontainment approaches (Mandell et al. 2015; Hirota et al. 2017).

4 Upstream Process Development Approach

In order to develop and optimize upstream processes for achieving enhanced production, researchers have explored clone selection, media screening, process parameter optimization (temperature, pH, inducer concentration, induction phase), impact of various media components/additives on the protein concentration, and effect of different operational parameters (batch, fed-batch, and continuous cultivation) (Rathore et al. 2021; Shojaosadati et al. 2008; Gronemeyer et al. 2014).

4.1 Medium Screening and Optimization

Medium screening and optimization are known to significantly impact productivity, while marginally affecting the overall cost of manufacturing. Thus, due to the availability of the various media components/additives and their direct impact on metabolic pathways, medium screening and optimization are routinely performed (Priyanka et al. 2022; Kumar et al. 2019; Maity and Mishra 2019). Complex medium is popular as it is economical, nutritionally rich, and gives the maximum possible cell growth. However, it results in batch-to-batch variability and significant foaming at a commercial scale. On the contrary, a defined medium assures batch-to-batch consistency but results in reduced biomass yield and lower protein production (Kangwa et al. 2015). One study has shown that a defined medium supplemented with yeast extract and casamino acid resulted in two-fold enhanced protein production (Jeong et al. 2004). Different medium additives, such as amino acids and sugars, serve as building blocks for the production of proteins and nucleotides. However, their excess quantity can generate undesirable by-products. On the contrary, if present in low quantity, cell starvation is likely, resulting in a reduction in cell viability and product yield. Thus, optimization of each media component is of considerable importance for maximizing production. Further, careful consideration and monitoring of medium components can reduce impurities at the process stage.

A wide variety of techniques, including one factor at a time (OFAT), modern statistical and mathematical techniques, such as an artificial neural network (ANN) and genetic algorithm (GA), have been employed to serve the purpose of media optimization (Fig. 5). OFAT is a widely used conventional approach to optimize medium components involving testing of one factor at a time, in lieu of multiple factors all at once (Jordan et al. 2013). Figure 6 shows general strategy for implementing design of experiment. Design of experiment (Plackett–Burman design, Taguchi, central composite, such as response surface methodology (RSM), and Box–Behnken design) and high throughput methods (ANN, GA, and fuzzy

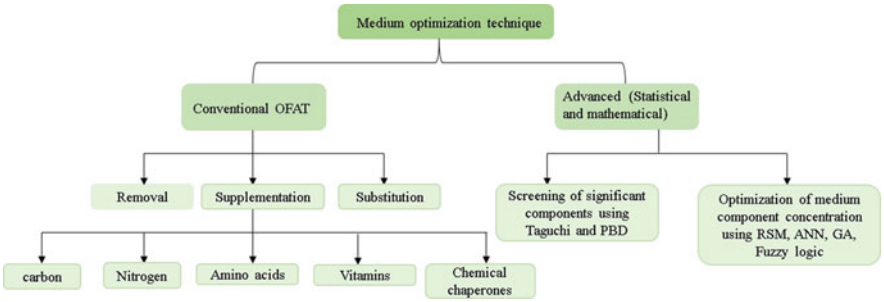


Fig. 5 Schematic representation of various medium optimization strategies (adapted from Singh et al. 2017)

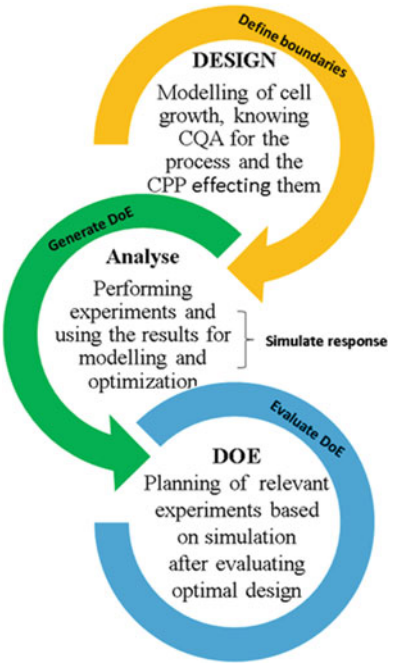


Fig. 6 Schematic representation of design of experiment strategy. Critical Quality Attributes (CQA); Critical Process Parameters (CPP); Design of Experiments (DoE)

logic) are commonly employed to reduce the number of experiments required for medium optimization (Maity and Mishra 2019).

4.2 *Process Parameter Optimization*

Process parameters, such as temperature, pH, inducer concentration, induction phase, agitation, aeration, and culture volume, also play vital role in improving overall process yield (Krause et al. 2016; Uhoraningoga et al. 2018). Figure 7 lists the process parameters which have major impact on the cellular environment.

4.2.1 Temperature

Temperature is an important parameter to optimize as it impacts both cell growth and product formation. At high temperatures, the thermal death rate surpasses the growth rate and hence results in a net decrease in viable cell concentration. The temperature may become the rate-limiting step in the fermentation as the bioreaction rate is found to increase with the increase in temperature when compared to nutrient diffusion rate, and at this point, diffusion becomes rate-limiting. Therefore, at high temperatures, diffusion limits must be carefully considered (Liu 2017). High temperature may encourage cell growth, yet it may be detrimental to protein expression since a

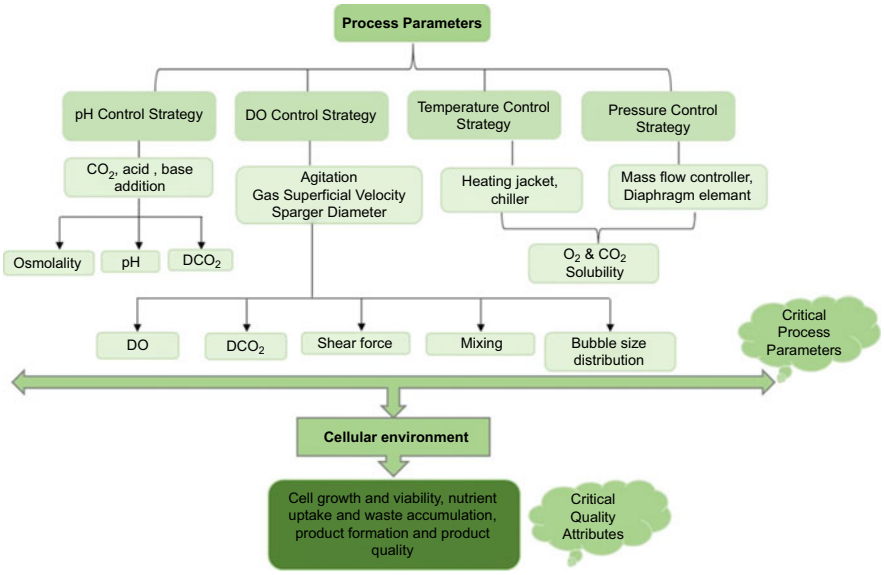


Fig. 7 Schematic showing the various operating parameters of a microbial process that affect process performance and product quality (adapted from Li et al. 2010)

high growth rate can result in a higher probability of plasmid loss (Papaneophytou and Kontopidis 2014).

Further, as the growth phase and production phases are separate, temperature has to be optimized for each step while retaining nutrient properties. Also, the temperature is reported to affect the rate-limiting steps during fermentation and for achieving a higher amount of soluble protein (Shuler and Kargi 2002). Different studies have confirmed the impact of fermentation temperature on the concentration of product and productivity of recombinant protein production in microbial cell factories (Yilmaztekin et al. 2013). Degradation of mRNA was found to follow first-order kinetics, and it is reported to decrease with the temperature that ultimately increases protein production (Shin et al. 1997; Neymotin et al. 2014; Kim and Jacobs-Wagner 2018; Baudrimont et al. 2019).

4.2.2 pH

pH is another critical physiological parameter, which affects metabolic enzyme activity and hence, cellular growth. The optimal pH of cells to grow is different from the product formation. Also, during growth, the pH may vary significantly, depending on the nature of nitrogen source. For example, if sole nitrogen source is ammonium, then increment in hydrogen ions levels is observed in the culture broth, which leads to pH reduction. On the contrary, if the sole nitrogen source is nitrate, then hydrogen is withdrawn from the medium by reducing nitrate to ammonia and results in a pH increase. During cellular growth, the pH can vary due to organic acid production, CO_2 supply, and amino acid generation from small peptides (Manderson et al. 2006). Different studies carried out by researchers confirm that pH influences enzyme activity and cellular growth during fermentation (Soni et al. 2008; Kumar et al. 2019). A set point of pH is typically defined and maintained around the cell environment by adding acid or base. In a few cases, a defined trajectory of pH, balance of optimum pH for growth, and protein production has been followed (Shuler and Kargi 2002; Junker and Wang 2006). The pH is mostly measured using a combined glass reference electrode, which can withstand sterilization temperature and pressure of 120 °C and 138 kN/m^2 (Najafpour 2015).

4.2.3 Inducer Concentration and Induction Phase

The amount of inducer varies proportionally to culture biomass. It depends on promoter strength, repressor gene, gene's cellular location, nature of the protein, and inducer type (Donovan et al. 1996; Monika Cserjan-Puschmann et al. 2002; Marbach and Bettenbrock 2012). Researchers have tried different concentrations of inducer solely based on cell mass. For most cases, 1 mM IPTG/L is a reasonable starting point with 0.1–1 mM IPTG/L suitable for secretory proteins to avoid insolubility, cell lysis, and growth inhibition (Shin et al. 1997; Couto et al. 2017; Silaban et al. 2019). Induction can be performed in the early log, mid-log, and late

log phases (San-Miguel et al. 2013). Often, the high cell density is achieved during late log phase that results in high product levels (Ou et al. 2004). Mid-log phase has proven effective in other cases (Donovan et al. 1996; Galloway et al. 2003). Additionally, strategies like supplying the inducer continuously along with the feed, prolonging the duration of the induction phase, changing the temperature during induction, using lactose as inducer for diauxic growth have shown to help in improving the process yield (Shojaosadati et al. 2008; Kopp et al. 2019).

4.2.4 Dissolved Oxygen (DO)

Apart from a carbon source, dissolved oxygen (DO) is also an essential substrate for aerobic organism growth. It can act as a limiting substrate due to its limited solubility in the water. The amount of DO in the medium is introduced and improved by air supply, pure oxygen supply, and agitation (Priyanka et al. 2019a, b). Often, liquid film surrounding the gas bubbles provides the resistance to interphase oxygen transfer and is given by the oxygen transfer rate (OTR) Eq. 1.

$$N_{O2} = K_L a(C^* - C_L) = OTR \tag{1}$$

where N_{O2} is the oxygen transfer flux (g/m^2.s), K_L is the oxygen transfer coefficient (m/s), a is the gas–liquid interfacial area (m^2/m^3), C^* is saturated DO concentration (g/L), and C_L is the actual DO concentration in the broth (g/L).

It is essential to maintain an optimal DO concentration inside a reactor since by-products like acetate, lactate, succinate, and ethanol are produced under critical oxygen concentration. Also, at high DO levels, free oxidants are produced which are detrimental to cellular growth and productivity (Shojaosadati et al. 2008). The partial pressure of oxygen (O_2) in the headspace provides the amount of dissolved oxygen in the medium and whose value is measured via DO probe (Winters 2009; Najafpour 2015). In order to get accurate O_2 levels, the probe is installed in the well-mixed region away from spargers or walls of the bioreactor. Mostly, DO probes are based on either galvanic or polarographic, where galvanic DO probes are preferred over polarographic ones (Winters 2009).

4.2.5 Dissolved Carbon Dioxide (CO_2)

Excess carbon dioxide (CO_2) gas in the growth medium adversely affects cellular metabolism by lowering its internal pH. It results in unfavorable cellular growth, product formation, and glycosylation patterns (Liu 2017). The level of dissolved CO_2 in the culture medium is usually measured by observing CO_2 partial pressure (pCO_2) (Pattison et al. 2000). At an industrial scale, a higher level of dissolved CO_2 is observed in the growth medium due to cellular respiration. The lower mass transfer flux between liquid and gaseous phases is the reason for this observation (Gerlach et al. 2019). In-line gas analyzers are theoretically used to measure pCO_2.

In the case of acid-producing culture, pCO_2 measurement can be done indirectly by pH measurement. Recently, sterilization friendly pCO_2 probes have been developed and used in bioreactors to measure the CO_2 concentrations (Winters 2009; Gerlach et al. 2019).

4.2.6 Pressure

Positive pressure is maintained inside the bioreactor to ensure aseptic conditions. The bioreactor pressure is influenced by the sum of partial pressures of all dissolved gases. High pressure inside the bioreactor increases the driving force for O_2 solubility and results in increased cell density culture. Also, the atmospheric pressure measurement of inlet and outlet gas is essential, especially for the cultures where oxygen consumption rate is meager and with little difference between inlet and outlet gas composition. Even a 1% decrease in pressure can lead to about 1% change in the oxygen concentration reading in paramagnetic oxygen gas analyzers. The pressure of inlet and outlet gas is generally measured using paramagnetic analyzers in laboratory reactors, and with diaphragm sensors in commercial bioreactors (Winters 2009).

4.2.7 Redox Potential

The rate and extent of much oxidative–reductive reaction are affected by redox potential, making it an important parameter (Prade et al. 2020). A typical fermentation medium's redox potential is a dynamic feature of ionic strength of the medium, DO, and pH. It is usually expressed in millivolts. The redox of a medium can be reduced either by sparging N_2 or by adding reducing agents like N_2S or cysteine HCl, and it can also be increased by adding some oxidizing agents. Redox probes are used for inline measurement of redox potential during fermentation (Winters 2009).

4.2.8 Ionic Strength

The transport of various nutrients in and out of the cells, such as carbon source, micronutrients, gas uptake, and thus the cells metabolic activity is affected by the ionic strength of the medium. Typically, the ionic strength can be expressed as the following equation:

$$I = 0.5 \sum C_i Z_i^2$$

where I is the medium ionic strength (expressed in either Molar, mol/L of solution, or molal, mol/kg of solvent), Z_i is the ionic charge, and C_i is the molar concentration of any ionic species (M, mol/L). Higher concentration above

stoichiometric requirement is generally known as inhibitory substrate concentration to cellular functions. The inhibitory concentration of a few of the essential components during cellular growth is >200 g/L of glucose, leading to a reduction in the water activity and ethanol formation in yeast. While >40 g/L of NaCl result in very high osmotic pressure. Some other compounds like methanol during *Pichia pastoris* cultures are inhibitory at even lower concentration, e.g., 1 g/L. The ionic strength of a medium can be measured based on conductivity and osmolality (Shuler and Kargi 2002; Doran 1995).

4.2.9 Culture Volume

The measurement of conductivity, capacitance, and differential pressure is used to determine the bioreactor volume. Weight-based sensors, such as load cell or balance, may also be used. However, the connections between bioreactor and load cell have to be flexible for the correct measurement (Shuler and Kargi 2002). The difference in timing of death phase, oxidative stress, and mutational frequency was observed based on the culture volume and type of vessel used (Kram and Finkel 2014).

4.2.10 Foam

Foam formation during the fermentation process can cause serious operational difficulties, particularly in high cell density fermentation. As cell density increases, an increase in foam formation is observed due to cell lysis, which results in the accumulation of host cell proteins in the medium. During the fermentation, the foam levels can interfere with level sensors, and in case it escapes, it can disrupt exhaust filter operations, which may also result in contamination. In manufacturing, numerous methods have been employed to reduce the rate of foaming. Foaming inside a bioreactor can be minimized by carefully designed impeller blades, better positioning of additives inlet ports, and the use of surface-active chemical agents. Though foam breaking chemicals are widely used, their application can result in lower overall mass transfer coefficient (k_La, mmoL of O_2/mL of medium), resulting in reduced gas supply and inhibited cell growth. As of now, foam formation is not very well understood. Yet, it has been reported to increase by use of the complex medium as well as the concentration of the extracellular product. Generally, capacitance and conductance probes are used for measurement of foam levels (Shuler and Kargi 2002; Shojaosadati et al. 2008).

5 Conclusions and Future Perspective

As far as recombinant expression is concerned, *E.coli* is always the first choice as the microbial cell factory. It is the most well-known expression hosts among all prokaryotes for expressing even large membrane proteins or humanized antibody

fragments, although it remains challenging to produce small peptides (Rosano and Ceccarelli 2014; Komives et al. 2017; Bhambure and Gani 2018). As defined in Table 2 and Fig. 2, special strains and vectors can come in handy for complex proteins. Large-scale protein expression studies reveal that *E. coli* can express less than 50% and 15% of bacterial and non-bacterial proteins, respectively. It indicates the versatility of the *E. coli* system (Rosano et al. 2019). It can be concluded that skillful use of the vast genetic toolbox with notable consideration toward mRNA stability, codon bias, and inclusion body formation can play a decisive role in manufacturing of biotherapeutic products.

The associated process complexity and batch inconsistency in manufacturing therapeutic drugs (such as non-glycosylated mAb, Fabs, scFvs) raise serious concerns to deal with conventional manufacturing practices. Thus, it is difficult to propose a universal strategy for biotherapeutic development that can resolve all issues. There is always scope for further innovation and improvement. As it is essential to evaluate respective cultivation strategies for every biotherapeutic expression, this chapter aims to offer a comprehensive perspective of biotherapeutic production in microbial hosts.

Continuous improvement in protein expression has contributed significantly to modern biotechnology. Modern molecular biology techniques such as CRISPR/cas9, metabolic engineering, and system biology for strain engineering are at the forefront for enhancing the quality and quantity of recombinant protein production. Glyco-engineering strategies in vitro and in vivo, site-directed mutagenesis, and directed evolution help improve production of recombinant protein with enhanced activity and stability. Continuous bioprocessing, high throughput process development, cell-free expression systems, and single-use disposable bioreactor are also seen as enormously important expansions. Process optimization using advanced data analytics is also becoming popular in biopharmaceutical production. Despite the fact that 40 years have passed since the first human protein was expressed in *E. coli*, many challenges remain and innovations will continue to occur.

References

Ahmad I, Nawaz N, Darwesh NM et al (2018) Overcoming challenges for amplified expression of recombinant proteins using Escherichia coli. Protein Expr Purif 144:12–18

Amid A, Ismail NA, Yusof F, Salleh HM (2011) Expression, purification, and characterization of a recombinant stem bromelain from Ananas comosus. Process Biochem 46:2232–2239

Anné J, Vrancken K, van Mellaert L et al (2014) Protein secretion biotechnology in Gram-positive bacteria with special emphasis on Streptomyces lividans. Biochim Biophys Acta Mol Cell Res 1843:1750–1761

Balbás P, Lorence A (2004) Recombinant gene expression: reviews and protocols. Methods in molecular biology, 2nd edn. Springer

Baudrimont A, Jaquet V, Wallerich S et al (2019) Contribution of RNA degradation to intrinsic and extrinsic noise in gene expression. Cell Rep 26:3752–3761.e5

Belagaje RM, Reams SG, Stan CLY, Prouty WF (1997) Increased production of low molecular weight recombinant proteins in *Escherichia coli*. Protein Sci 6:1953–1962

Bhambure RS, Gani KM (2018) A method for producing refolded recombinant humanized ranibizumab. WO 2018/211529 Al

Boix E, Wu YN, Vasandani VM et al (1996) Role of the N terminus in RNase a homologues: differences in catalytic activity, ribonuclease inhibitor interaction and cytotoxicity. J Mol Biol 257:992–1007

Bora AMA-HRS, Ahmed MMM (2019) Plasmids for optimizing expression of recombinant proteins in *E. coli*. In: Gull M (ed) Plasmids. IntechOpen

Chart H, Smith HR, La Ragione RM, Woodward MJ (2000) An investigation into the pathogenic properties of *Escherichia coli* strains BLR, BL21, DH5?? And EQ1. J Appl Microbiol 89:1048–1058

Chen H, Bjerknes M, Kumar R, Jay E (1994) Determination of the optimal aligned spacing between the Shine - Dalgarno sequence and the translation initiation codon of *Escherichia coli* m RNAs. Nucleic Acids Res 22:4953–4957

Chien T, Harimoto T, Kepecs B, Gray K, Coker C, Hou N, Pu K, Azad T, Nolasco A, Pavlicova M, Danino T (2022) Enhancing the tropism of bacteria via genetically programmed biosensors. Nat Biomed Eng 6(1):94–104

Chowdhury S, Castro S, Coker C, Hinchliffe TE, Arpaia N, Danino T (2019) Programmable bacteria induce durable tumor regression and systemic antitumor immunity. Nat Med 25 (7):1057–1063

Cooper GM (2000) Transcription in prokaryotes RNA polymerase and transcription. In: The cell: a molecular approach, 2nd edn. NCBI, Sunderland, pp 1–6

Couto MR, Rodrigues JL, Rodrigues LR (2017) Optimization of fermentation conditions for the production of curcumin by engineered *Escherichia coli*. J R Soc Interface 14:1–8

Cserjan-Puschmann M, Grabherr R, Striedner G et al (2002) Optimizing recombinant microbial fermentation processes. BioPharm 15:26–34

Das TK, Narhi LO, Sreedhara A et al (2020) Stress factors in mAb drug substance production processes: critical assessment of impact on product quality and control strategy. J Pharm Sci 109:116–133

De Luca V, Vullo D, Del Prete S et al (2015) Cloning, characterization and anion inhibition studies of a new γ-carbonic anhydrase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*. Bioorganic Med Chem 23:4405–4409

De Smit MH, Van Duin J (1990) Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc Natl Acad Sci U S A 87:7668–7672

Dieryck W, Noubhani AM, Coulon D, Santarelli X (2003) Cloning, expression and two-step purification of recombinant His-tag enhanced green fluorescent protein over-expressed in *Escherichia coli*. J Chromatogr B Anal Technol Biomed Life Sci 786:153–159

Donovan RS, Robinson CW, Click BR (1996) Review: optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. J Ind Microbiol 16:145–154

Doran PM (1995) Bioprocess engineering principles. Elsevier

Fath S, Bauer AP, Liss M et al (2011) Multiparameter RNA and codon optimization: a standardized tool to assess and enhance autologous mammalian gene expression. PLoS One 6(3):e17596

Galloway CA, Sowden MP, Smith HC (2003) Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. BioTechniques 34:524–530

Gerlach G, Guth U, Oelßner W et al (2019) CO_2 Measurement in biotechnology and industrial processes. In: Carbon dioxide sensing: fundamentals, principles, and applications, p 349

Gronemeyer P, Ditz R, Strube J (2014) Trends in upstream and downstream process development for antibody manufacturing. Bioengineering 1:188–212

Gu P, Yang F, Su T et al (2015) A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. Sci Rep 5:1–9

Gupta SK, Shukla P (2016) Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. Crit Rev Biotechnol 36:1089–1098

Gupta SK, Shukla P (2017) Microbial platform technology for recombinant antibody fragment production: a review. Crit Rev Microbiol 43:31–42

Hannig G, Makrides SC (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. Trends Biotechnol 16:89–92
Harms P, Kostov Y, Rao G (2002) Bioprocess monitoring. Curr Opin Biotechnol 13:124–127
Hirota R, Abe K, Katsuura ZI et al (2017) A novel biocontainment strategy makes bacterial growth and survival dependent on phosphite. Sci Rep 7:1–10
Huang HC, Wang SC, Leu YJ et al (1998) The Rana catesbeiana rcr gene encoding a cytotoxic ribonuclease. Tissue distribution, cloning, purification, cytotoxicity, and active residues for RNase activity. J Biol Chem 273:6395–6401
Humphreys DP, Sehdev M, Chapman AP et al (2000) High-level periplasmic expression in *Escherichia coli* using a eukaryotic signal peptide: importance of codon usage at the 5′ end of the coding sequence. Protein Expr Purif 20:252–264
Jeong KJ, Choi JH, Yoo WM et al (2004) Constitutive production of human leptin by fed-batch culture of recombinant rpoS-Escherichia coli. Protein Expr Purif 36:150–156
Jia B, Jeon CO (2016) High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. Open Biol 6:1–17
Jordan M, Voisard D, Berthoud A et al (2013) Cell culture medium improvement by rigorous shuffling of components using media blending. Cytotechnology 65:31–40
Joshi S, Rathore AS (2020) Assessment of structural and functional comparability of biosimilar products: trastuzumab as a case study. BioDrugs 34:209–223
Jozala AF, Geraldes DC, Tundisi LL et al (2016) Biopharmaceuticals from microorganisms: from production to purification. Brazilian J Microbiol 47:51–63
Junker BH, Wang HY (2006) Bioprocess monitoring and computer control: key roots of the current PAT initiative. Biotechnol Bioeng 95:226–261
Kangwa M, Yelemane V, Polat AN et al (2015) High-level fed-batch fermentative expression of an engineered staphylococcal protein a based ligand in *E. coli*: purification and characterization. AMB Express 5:70
Kesik-Brodacka M (2018) Progress in biopharmaceutical development. Biotechnol Appl Biochem 65:306–322
Kim S, Jacobs-Wagner C (2018) Effects of mRNA degradation and site-specific transcriptional pausing on protein expression noise. Biophys J 114:1718–1729
Kim K, Choe D, Lee DH, Cho BK (2020) Engineering biology to construct microbial chassis for the production of difficult-to-express proteins. Int J Mol Sci 21(3):990
Komives CF, Sanchez EE, Rathore AS et al (2017) Opossum peptide that can neutralize rattlesnake venom is expressed in *Escherichia coli*. Biotechnol Prog 33:81–86
Kopp J, Kolkmann AM, Veleenturf PG et al (2019) Boosting recombinant inclusion body production—from classical fed-batch approach to continuous cultivation. Front Bioeng Biotechnol 7: 1–12
Korpimäki T, Kurittu J, Karp M (2003) Surprisingly fast disappearance of β-lactam selection pressure in cultivation as detected with novel biosensing approaches. J Microbiol Methods 53:37–42
Kram KE, Finkel SE (2014) Culture volume and vessel affect long-term survival, mutation frequency, and oxidative stress of *Escherichia coli*. Appl Environ Microbiol 80:1732–1738
Krause M, Neubauer A, Neubauer P (2016) The fed-batch principle for the molecular biology lab: Controlled nutrient diets in ready-made media improve production of recombinant proteins in *Escherichia coli*. Microb Cell Fact:15
Kumar D, Batra J, Komives C, Rathore A (2019) QbD based media development for the production of fab fragments in *E. coli*. Bioengineering 6:29
Kumar S, Birah A, Chaudhary B et al (2005) Plant codon optimized cry genes of bacillus thuringiensis can be expressed as soluble proteins in *Escherichia coli* BL21 codon plus strain as NusA-cry protein fusions. J Invertebr Pathol 88:83–86
Laursen BS, Søren SA, Hedegaard J et al (2002) Structural requirements of the mRNA for intracistronic translation initiation of the enterobacterial infB gene. Genes Cells 7:901–910
Laursen BS, Sorensen HP, Mortensen KK, Sperling Petersen U.S. (2005) Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev 69:101–123

Le SB, Onsager I, Lorentzen JA, Lale R (2020) Dual UTR-A novel 5' untranslated region design for synthetic biology applications. Synth Biol 5:1–11

Leppek K, Byeon GW, Kladwang W et al (2021) Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. bioRxiv:1–74

Li F, Vijayasankaran N, Shen A et al (2010) Cell culture processes for monoclonal antibody production. MAbs 2:466–479

Lin HH, Yin LJ, Jiang ST (2009) Cloning, expression, and purification of pseudomonas aeruginosa keratinase in *Escherichia coli* AD494(DE3)pLysS expression system. J Agric Food Chem 57: 3506–3511

Liu S (2017) How cells grow. In: Bioprocess engineering: kinetics, sustainability, and reactor design. Elsevier, Amsterdam, pp 629–697

Lobstein J, Emrich CA, Jeans C et al (2012) SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Factories 11:753

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM (2005) Heterologous protein production using the Pichia pastoris expression system. Yeast 22(4):249–270

Maity N, Mishra S (2019) Statistically designed medium reveals interactions between metabolism and genetic information processing for production of stable human serum albumin in *Pichia pastoris*. Biomol Ther 9:568

Maldonado LMTP, Hernández VEB, Rivero EM et al (2007) Optimization of culture conditions for a synthetic gene expression in *Escherichia coli* using response surface methodology: the case of human interferon beta. Biomol Eng 24:217–222

Mandell DJ, Lajoie MJ, Mee MT et al (2015) Biocontainment of genetically modified organisms by synthetic protein design. Nature 518:55–60

Manderson D, Dempster R, Chisti Y (2006) A recombinant vaccine against hydatidosis: production of the antigen in Escherichia coli. J Ind Microbiol Biotechnol 33:173–182

Marbach A, Bettenbrock K (2012) Lac operon induction in *Escherichia coli*: systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. J Biotechnol 157:82–88

Miroux B, Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260: 289–298

Najafpour GD (2015) Fermentation process control. In: Biochemical engineering and biotechnology, 2nd edn. Elsevier, Amsterdam, pp 597–630

Neymotin B, Athanasiadou R, Gresham D (2014) Determination of in vivo RNA kinetics using RATE-seq. RNA 20:1645–1652

Notomista E, Cafaro V, Fusiello R et al (1999) Effective expression and purification of recombinant onconase, an antitumor protein. FEBS Lett 463:211–215

O'Flaherty R, Bergin A, Flampouri E et al (2020) Mammalian cell culture for production of recombinant proteins: a review of the critical steps in their biomanufacturing. Biotechnol Adv 43:107552

Ou J, Wang L, Ding X et al (2004) Stationary phase protein overproduction is a fundamental capability of *Escherichia coli*. Biochem Biophys Res Commun 314:174–180

Ozdemir T, Fedorec AJ, Danino T, Barnes CP (2018) Synthetic biology and engineered live biotherapeutics: toward increasing system complexity. Cell Syst 7(1):5–16

Palomares LA, Estrada-Mondaca S, Ramírez OT (2004) Production of recombinant proteins: challenges and solutions. Methods Mol Biol 267:15–52

Papaneophytou CP, Kontopidis G (2014) Statistical approaches to maximize recombinant protein expression in *Escherichia coli*: a general review. Protein Expr Purif 94:22–32

Park YS, Seo SW, Hwang S et al (2007) Design of 5′-untranslated region variants for tunable expression in *Escherichia coli*. Biochem Biophys Res Commun 356:136–141

Pattison RN, Swamy J, Mendenhall B et al (2000) Measurement and control of dissolved carbon dioxide in mammalian cell culture processes using an in situ fiber optic chemical sensor. Biotechnol Prog 16:769–774

Pell G (2004) Structural and biochemical analysis of *Cellvibrio japonicus* xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family gh-10 xylanases. J Biol Chem 279:11777–11788

Prade E, Zeck A, Stiefel F et al (2020) Cysteine in cell culture media induces acidic IgG1 species by disrupting the disulfide bond network. Biotechnol Bioeng:1091–1104

Priyanka P, Rathore AS (2021) A novel strategy for efficient expression of an antibody fragment in Escherichia coli: ranibizumab as a case study. J Chem Technol Biotechnol

Priyanka KJ, Gomes J, Rathore AS (2019a) Implementing process analytical technology for the production of recombinant proteins in *Escherichia coli* using an advanced controller scheme. Biotechnol J 14:1–10

Priyanka RS, Chopda V et al (2019b) Comparison and implementation of different control strategies for improving production of rHSA using *Pichia pastoris*. J Biotechnol 290:33–43

Priyanka P, Patil RS, Meshram P et al (2022) Ethanol as additive enhances expression of Ranibizumab in Escherichia coli: impact on cellular physiology and transcriptome. Process Biochem 112:167–176

Raab D, Graf M, Notka F et al (2010) The GeneOptimizer algorithm: using a sliding window approach to cope with the vast sequence space in multiparameter DNA sequence optimization. Syst Synth Biol 4:215–225

Rabhi-Essafi I, Sadok A, Khalaf N, Fathallah DM (2007) A strategy for high-level expression of soluble and functional human interferon as a GST-fusion protein in *E.coli*. Protein Eng Des Sel 20:201–209

Raj SM, Rathnasingh C, Jo JE, Park S (2008) Production of 3-hydroxypropionic acid from glycerol by a novel recombinant *Escherichia coli* BL21 strain. Process Biochem 43:1440–1446

Ramalingam S, Gautam P, Mukherjee KJ, Jayaraman G (2007) Effects of post-induction feed strategies on secretory production of recombinant streptokinase in *Escherichia coli*. Biochem Eng J 33:34–41

Rathore AS, Mishra S, Nikita S et al (2021) Bioprocess control: current progress and future perspectives. Curr Prog Future Perspect Life 11:1–21

Reh G (2020) 2020 Global Life Sciences Outlook. Deloitte Insights 2020

Rosano GL, Ceccarelli EA (2009) Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. Microb Cell Factories 8:1–9

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:1–17

Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: a 5-year update. Protein Sci 28:1412–1422

San-Miguel T, Pérez-Bermúdez P, Gavidia I (2013) Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. Springerplus 2:89

Schlegel S, Löfblom J, Lee C et al (2012) Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). J Mol Biol 423:648–659

Seo SW, Yang JS, Cho HS et al (2014) Predictive combinatorial design of mRNA translation initiation regions for systematic optimization of gene expression levels. Sci Rep 4:1–7

Shapiro R, Harper JW, Fox EA, Jansen HW, Friedrich H, Uhlmann E (1988) Expression of met-(−1) angiogenin in *Escherichia coli*: conversion to the authentic Glu-1 protein. Anal Biochem 175:450–461

Shin CS, Hong MS, Bae CS, Lee J (1997) Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]. Biotechnol Prog 13:249–257

Shojaosadati SA, Varedi Kolaeil SM, Babaeipour V, Farnoud AM (2008) Recent advances in high cell density cultivation for production of recombinant protein. Iran J Biotechnol 6:63–84

Shuler ML, Kargi F (2002) Bioprocess engineering basic concepts Second Edition

Silaban S, Gaffar S, Simorangkir M et al (2019) Effect of IPTG concentration on recombinant human prethrombin-2 expression in *Escherichia coli* BL21(DE3) ArcticExpress. IOP Conf Ser Earth Environ Sci 217:012039

Singh V, Haque S, Niwas R et al (2017) Strategies for fermentation medium optimization: an in-depth review. Front Microbiol 7:2087

Soni P, Kansal H, Banerjee UC (2008) Optimization of process parameters for the production of carbonyl reductase by *Candida viswanathii* in a laboratory-scale fermentor. J Ind Microbiol Biotechnol 35:167–173

Sørensen HP, Mortensen KK (2005a) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J Biotechnol 115:113–128

Sørensen HP, Mortensen KK (2005b) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. Microb Cell Factories 4:1

Stanbury PF, Whitaker A, Hall SJ (2017) The production of heterologous proteins. Princ Ferment Technol 1:725–775

Tegel H, Tourle S, Ottosson J, Persson A (2010) Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). Protein Expr Purif 69:159–167. https://doi.org/10.1016/j.pep.2009.08.017

Tegel H, Ottosson J, Hober S (2011) Enhancing the protein production levels in *Escherichia coli* with a strong promoter. FEBS J 278:729–739

Tian ZG, Dong TT, Yang YL et al (2009) Expression of antimicrobial peptide LH multimers in *Escherichia coli* C43(DE3). Appl Microbiol Biotechnol 83:143–149

Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:420

Ueda T (2014) Next-generation optimized biotherapeutics—a review and preclinical study. Biochim Biophys Acta—Proteins Proteomics 1844:2053–2057

Uhoraningoga A, Kinsella GK, Henehan GT, Ryan BJ (2018) The goldilocks approach: a review of employing design of experiments in prokaryotic recombinant protein production. Bioengineering 5:89

Verma MK, Verma YK (2012) New generation expression host system—aiming high for commercial production of recombinant protein. IOSR J Pharm Biol Sci 1:09–16

Wacker A, Chemi Leonhartsberger S, Candussio A, Schmid G (2012) Signal peptide for the production of recombinant proteins. US 8,148,494 B2 Apr. 3, 2012

Walsh G (2014) Biopharmaceutical benchmarks 2014. Nat Biotechnol 32:992–1000

Walsh G (2018) Biopharmaceutical benchmarks 2018. Nat Biotechnol 36:1136–1145

Wang A, Su Y, Wang S et al (2010) High efficiency preparation of bioactive human-defensin 6 in *Escherichia coli* origami(DE3)pLysS by soluble fusion expression. Appl Microbiol Biotechnol 87:1935–1942

Winters J (2009) Instrumentation and control. In: Principles of fermentation technology, 2nd edn, pp 215–241

Wu X, Jörnvall H, Berndt KD, Oppermann U (2004) Codon optimization reveals critical factors for high level expression of two rare codon genes in Escherichia coli: RNA stability and secondary structure but not tRNA abundance. Biochem Biophys Res Commun 313(1):89–96

Yang YH, Zheng GG, Li G et al (2004) Expression of bioactive recombinant GSLL-39, a variant of human antimicrobial peptide LL-37, in *Escherichia coli*. Protein Expr Purif 37:229–235

Yazdankhah S, Skjerve E, Wasteson Y (2018) Antimicrobial resistance due to the content of potentially toxic metals in soil and fertilizing products. Microb Ecol Health Dis 29:1548248

Yilmaztekin M, Cabaroglu T, Erten H (2013) Effects of fermentation temperature and aeration on production of natural isoamyl acetate by *Williopsis saturnus* var. saturnus. Biomed Res Int 2013:870802

Zhong X, Neumann P, Corbo M, Loh E (2011) Recent advances in biotherapeutics drug discovery and development. In: Drug discovery and development–present and future, pp 363–378

Current Trends and Prospects in Antimicrobial Peptide Bioprocessing

Kamila Botelho Sampaio de Oliveira, Michel Lopes Leite, Gisele Regina Rodrigues, Nicolau Brito da Cunha, Simoni Campos Dias, and Octavio Luiz Franco

Contents

K. B. S. de Oliveira · G. R. Rodrigues · N. B. da Cunha
Centro de Análises Proteômicas e Bioquímicas, Pós-graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Brazil

M. L. Leite
Centro de Análises Proteômicas e Bioquímicas, Pós-graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Brazil

Departamento de Biologia Molecular, Instituto de Ciências Biológicas, Campus Darcy Ribeiro, Universidade de Brasília, Brasília, Distrito Federal, Brazil

S. C. Dias
Centro de Análises Proteômicas e Bioquímicas, Pós-graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Brazil

Universidade de Brasília, Pós-graduação em Biologia Animal, Campus Darcy Ribeiro, Brasília, Brazil

O. L. Franco (✉)
Centro de Análises Proteômicas e Bioquímicas, Pós-graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, Brasília, Brazil

Universidade de Brasília, Pós-graduação em Patologia Molecular, Campus Darcy Ribeiro, Brasília, Brazil

S-Inova Biotech, Pós-graduação em Biotecnologia, Universidade Católica Dom Bosco, Campo Grande, Mato Grosso do Sul, Brazil

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_5

Abstract The increase in resistance to conventional antimicrobials in recent years has boosted the search for new antibiotics to treat serious infectious diseases, especially those generated by multi-resistant bacteria. In this context, antimicrobial peptides (AMPs) are alternative molecules for use as new therapeutic agents. AMPs are small bioactive proteins commonly produced by all living organisms, and they can be part of innate immunity. Due to their broad-spectrum antibacterial potential and other activities, including immunomodulatory and antitumor, they are of great interest to the pharmaceutical industry's production of biopharmaceuticals. Among the technological platforms applied in the process of development and manufacturing of AMPs, recombinant DNA technology has enabled the production of such molecules using bacterial and yeast cells as expression host systems on a laboratory scale and in large-scale environments. Furthermore, different bioprocessing strategies can be used for peptide industrial production, aiming to optimize the yield, make cultures more robust and significantly increase cell density. In this chapter, we will address recent developments and future directions in AMPs bioprocessing, including microbial expression systems, as well as bioprocessing and purification technologies. Here we also describe successful cases in this field and emphasize the prospects and challenges related to AMPs bioengineering.

1 Introduction

The increase in microbial resistance to antibiotics is a major public health problem around the world. Methicillin-resistant *Staphylococcus aureus* (MRSA) and β-lactamase-resistant *Escherichia coli* (ESBL) are bacterial strains that recur in many hospitals (Assis et al. 2017; Wang et al. 2019), mainly infecting patients whose immune system is compromised by other diseases or invasive therapies (Leite et al. 2019). Bacteria such as *Pseudomonas aeruginosa, Enterococcus faecium*, *Acinetobacter baumannii*, *Klebsiella pneumoniae,* and *Enterobacter* species are also commonly responsible for nosocomial infections exhibiting high microbial resistance to currently available drugs (Laws et al. 2019).

The emergence of multi-resistant microorganisms such as those mentioned above may occur due to the excessive and/or incorrect use of antimicrobial drugs by humans and other animals (Lombardi et al. 2019; Bhopale 2020; Dijksteel et al. 2021). As a result, many drugs commonly used to fight those pathogens are no longer effective (Leite et al. 2019). About 700,000 people die annually from resistant

bacterial infections. If measures are not taken to change the current scenario, it is estimated that by 2050 the number of deaths will be ten million per year (Mishra et al. 2017; Laws et al. 2019). The development of new, broad-spectrum, and low-toxicity drugs as an alternative therapy to conventional antibiotics is highly desirable. Both academia and industry are working on prospecting candidate molecules and producing new therapies to combat infectious diseases (Andersson et al. 2016; Leite et al. 2019).

In this context, antimicrobial peptides (AMPs) have emerged as a new approach for replacing or complementing traditional therapeutic compounds in combating resistant microorganisms (da Costa et al. 2015; Chen and Lu 2020). They are naturally produced by various living organisms, such as microorganisms, plants, and animals, acting as important immune system components against exogenous pathogens (Wang et al. 2019; Leite et al. 2019). Generally, they present activity against several microorganisms and can perform more than one function in some cases, including acting as an immunomodulator or antitumor, so they are considered promiscuous molecules (Franco 2011; Haney et al. 2017; Leite et al. 2019).

Due to the different functions that they can perform simultaneously, AMPs are molecules that seem to have great potential as candidates for new drugs, especially antibiotics (Tornesello et al. 2020; Liscano et al. 2020). The development of new AMP-based drugs needs these molecules to be obtained in large quantities for biotechnological studies and pharmaceutical evaluation (Parachin et al. 2012). There have been notable improvements in producing AMPs on an industrial scale during the last few years, due to several new technologies (Gupta and Shukla 2017). Innovations in the upstream bioprocessing steps, such as selection and development of the cell line, optimization of cell culture parameters, and the use of feeding strategies, have the ultimate goal of large-scale production (Gronemeyer et al. 2014; Tripathi and Shrivastava 2019). The application of recombinant systems in cell line construction is opportune in terms of yields and costs for AMPs production (Parachin et al. 2012).

Among these recombinant systems, microorganism hosts are generally the most used, because they have characteristics that enable high levels of synthesis, modification, and secretion of heterologous AMPs. Various recombinant proteins are already available on the market, including AMPs produced by these hosts (Gupta and Shukla 2017). *E. coli* bacteria is the most widely used microbial cell factory, since its genetic and biological processes are well known, exhibiting fast growth rate, high yield and simple upstream process (Khow and Suntrarachun 2012; Briand et al. 2016; Kaur et al. 2018). Otherwise, among the yeast systems, *Saccharomyces cerevisiae* is the most established. It is a robust, stress-tolerant yeast and uses simple nutrients (Öztürk et al. 2017).

After developing the expression system, the next step consists of producing the small-scale recombinant protein to screen and select transforming clones. At this point, it is essential to monitor the cultivation conditions, as they directly affect the expression of the recombinant AMP. Optimizing the temperature, pH, aeration, agitation, media composition, the concentration of inducers, induction time and the feeding strategies make the development of bioprocesses more effective (Kaur

et al. 2018). Continuing the process, bioreactor systems and bioprocess strategies such as batch, fed-batch, and continuous culture are employed for the production of large-scale AMPs as biopharmaceuticals (Tripathi and Shrivastava 2019). Subsequently to large-scale production, a key purpose is the recovery of the target biomolecule and removal of the impurities present in the culture medium (Gupta and Shukla 2017).

Known as the downstream process, this phase has innovative technologies for handling large volumes in production processes, as well as for the recovery of the recombinant biomolecule of interest (Gupta and Shukla 2017). Filtration removes cell biomass, providing culture medium clarification, and affinity chromatography captures the recombinant protein using specific separation resins, promoting a high degree of purity for the proteins of interest (Singh et al. 2013; Kimple et al. 2013; Arora et al. 2017). In summary, this chapter addresses innovative approaches to upstream and downstream processes for AMPs yielding biopharmaceuticals. Selection of the proper expression hosts, development of bioprocesses, recent strategies in bioprocessing, techniques related to purification, and ways to achieve lower production costs, while boosting manufacturing flexibility and final product quality, are here described.

2 Antimicrobial Peptides

2.1 Properties

AMPs are small bioactive proteins that are effective against several species of Gram-positive and -negative bacteria, fungi, parasites and viruses (enveloped and non-enveloped) (Bhopale 2020; Moretta et al. 2021). These molecules are also known a host defense peptides (Boto et al. 2018; Liang and Diana 2020), and they are essential elements of the innate immune system of living organisms, acting as a first defense line against microbial actions displaying microbicidal, bacteriostatic, and cytolytic properties (Sinha and Shukla 2019; Moretta et al. 2021). AMPs are the primary defense line against pathogenic microorganisms for plants and insects, as these do not present an adaptive immune system. In bacteria and other microorganisms, AMPs act in defense of their environmental niche (Browne et al. 2020). The bacteria *Paenibacillus polymyxa*, for instance, which develops in plant roots, produces the antibiotic polymyxin, capable of breaking down *P. aeruginosa* or *S. aureus* biofilms (Quinn et al. 2012).

AMPs are evolutionarily conserved molecules that are extremely diverse in composition and length, with different sequences, structures, and sources, but they have some typical features (Bhopale 2020). Most AMPs are generally characterized as short molecules containing fewer than 100 amino acid residues and have molecular masses ranging from 1 to 10 kDa. These peptides commonly possess a positive net charge ranging from +2 to +11, attributed to the presence of positively charged residues such as lysine and arginine residues, thus being characterized as cationic

molecules (Mahlapuu et al. 2020). These molecules usually have a considerable amount of hydrophobic residues (typically 50%) such as valine, leucine, isoleucine, alanine, methionine, phenylalanine, tyrosine, and tryptophan in the peptide sequence (Lee et al. 2017; Thapa et al. 2020; Dijksteel et al. 2021).

2.2 Mechanisms of Action and Targets

Generally, the first AMP interaction with the bacterial cell membrane occurs through hydrophobic and electrostatic interactions between cationic residues and anionic components of the microorganism's membrane (Moravej et al. 2018). The negative microbial surface charge may be provided by phospholipid head groups such as phosphatidylglycerol, cardiolipin, or phosphatidylserine (Boto et al. 2018). Eukaryotic membranes have zwitterionic phospholipids such as phosphatidylethanolamine, phosphatidylcholine, and sphingomyelin. These phospholipids contribute to membrane net charge at physiological pH. In addition to zwitterionic phospholipids, the presence of cholesterol molecules along the membrane may also contribute to a reduction in membrane fluidity and flexibility, which can generally reduce AMP activity (Boto et al. 2018; Dijksteel et al. 2021). Therefore, AMPs are generally non-toxic to mammalian cells, which is an attractive feature for their therapeutic use (Silva et al. 2011; Browne et al. 2020). Although AMPs' mechanism of action is still largely unknown, the electrostatic attraction between negatively charged cells components and positively charged AMPs is understood to be important as a first step, resulting in a strong interaction and further target cell membranes disruption (Tornesello et al. 2020).

After initial electrostatic interactions, self-assembly peptide accumulation on the bacterial membrane surface may cause membrane integrity loss and intracellular component leakage after reaching certain concentrations (Mirski et al. 2017; Dijksteel et al. 2021). Furthermore, there are a several widely accepted models of action that usually involve bacterial cytoplasmic membrane integrity being disrupted in many ways. Among them are the barrel-stave pore, detergent micellization, toroidal pore, disordered toroidal pore, membrane thinning/thickening charged lipid clustering, non-bilayer intermediate formation, oxidized phospholipid targeting, anion carrier, and non-lytic membrane depolarization, among others (Nguyen et al. 2011; Mahlapuu et al. 2020).

In the barrel-stave model, AMPs can self-organize into cylindrical bundles, which insert themselves in a perpendicular way into the membranes. Aqueous pore lumen formation may occur with the orientation of hydrophobic portions towards the hydrophobic bilayer interior. The pore may cause cytomembrane permeabilization, osmotic imbalance and further cell death (Seyfi et al. 2020). In the toroidal model, the pore formation occurs when the AMPs insert themselves perpendicularly into the bacterial plasmatic membranes, a mechanism similar to the barrel-stave model. However, in this model, the nonpolar AMPs amino acid residues interact with lipid head groups, causing membrane deflection and, further, generating torus

pores (Lee et al. 2017; Moretta et al. 2021). Another possibility is the carpet model, in which the AMPs are accumulated and arranged parallel to the cell membrane, covering it completely. At the same time, micelles are formed with the initial ruptured membranes, through the hydrophilic amino acids interactions of polar phospholipid heads, causing cytomembrane disruption (Deng et al. 2017; Lee et al. 2017; Zandsalimi et al. 2020). Moreover, peptides are capable of interacting with several targets, including proteins and carbohydrates, acting not only on the surface, but also inside the cell (Silva et al. 2011; Kumar et al. 2018). In addition to destabilizing bacterial membranes and causing their rupture (membrane-acting peptides), AMPs can also cross the membrane, destabilizing normal cellular processes, such as cell division, protein, nucleic acid and cell wall synthesis, being classified as "non-membrane acting peptides." AMPs can act against pathogens by causing different stresses at the same time until the combined action causes cell death (Boto et al. 2018). AMPs that act through membrane destruction are also able to act through non-destructive membrane, in addition to acting independently or synergistically with other AMPs (Dijksteel et al. 2021).

Nonetheless, AMPs have a therapeutic potential that goes beyond their antimicrobial activity. Some AMPs can also exhibit cytotoxic activity against tumor cells (Wang et al. 2019). Most tumor cells also have a surface negative charge on the membrane due to anionic overexpression of molecules such as phosphatidylserine. This characteristic allows the interaction with cationic AMPs (Wang et al. 2019). AMPs are also capable of inhibiting inflammatory responses and stimulating the proliferation of immune system cells, acting as immune modulators (Drayton et al. 2020). They are capable of recruiting and stimulating the proliferation of macrophages, neutrophils, eosinophils, activation of T lymphocytes and differentiation of dendritic cells, in addition to inducing or modulating pro-inflammatory cytokines and producing chemokines, chemotaxis, causing apoptosis and inhibiting the inflammatory response. This supports their use as potential therapeutic molecules against immune-related diseases (Liang and Diana 2020).

AMPs are thus able to act on several cell targets and are considered as promiscuous molecules in certain cases (Franco 2011). Multiple functions can be related to a single peptide structure, contrary to what was supposed years ago, which held that peptides have an unconditional structure directly associated with a particular function. Knowledge about AMPs promiscuity has been gaining ground in several fields of research, such as in antibiotic development (Silva et al. 2011; Franco 2011). Therefore, natural AMPs have enormous potential as an alternative approach for the development of new therapies, acting alone or in synergy with conventional drugs (de Oliveira et al. 2020; León-Buitimea et al. 2020).

Several studies have already demonstrated AMPs' therapeutic efficacy (Vilas Boas et al. 2017; Lima et al. 2017; Fensterseifer et al. 2019; León-Buitimea et al. 2020; Almeida et al. 2020). Another important feature is that AMPs are less susceptible to microbial resistance since they act on evolutionarily conserved cell membrane components. Therefore, bacterial cells would need different mutations over an extended period to completely redesign the structure of their cell membranes (Mercer et al. 2020; Mahlapuu et al. 2020). Thus, AMPs seem to have great potential

for antibiotic adjuvants, making it possible to reduce or circumvent the occurrence of antibiotic resistance. The synergy between AMPs and antibiotics, at lower dosages, can provide a reduction in the toxicity or adverse side effects of a drug (Browne et al. 2020).

Despite being less common, some anionic AMPs have been reported, such as the anionic AMP maximin-H5, isolated from amphibians, and dermcidin, secreted by the human eccrine sweat glands (Rios et al. 2016; Boparai and Sharma 2019). They are made up of negatively charged glutamic and aspartic acid residues, with a net negative charge ranging from −1 to −7. Precisely because of its negative charge, its mode of action differs from cationic AMPs. However, some anionic AMPs can be disruptive to bacterial cell membranes. These peptides are capable of using metal ions to form cationic salt bridges with the negative microbial membrane constituents, which enables cell penetration. When they reach the cytoplasm, they can bind to intracellular components such as ribosomes or inhibit ribonuclease activity, inducing cell death (Wang et al. 2019; Boparai and Sharma 2019; Moretta et al. 2021).

2.3 Classification

AMPs can be classified according to their structural aspects into three major subgroups: α-helical, β-sheet, and extended peptides (Tornesello et al. 2020; Moretta et al. 2021). The α-helical peptides, when in interaction with bacterial membranes, can be organized into a flexible amphipathic structure (Koo and Seo 2019; Cardoso et al. 2021). Examples of α-helical AMPs include magainin, temporins and melittin (Ge et al. 1999; Raja et al. 2017; Ramirez et al. 2019). Peptides in β-sheets may have greater structural stability in solution, due to the cysteine residues that are conserved and form disulfide bonds, which minimize proteinases degradation. This class includes AMPs such as protegrins from the cathelicidin family, defensins and tachyplesins (Kumar et al. 2018; Seyfi et al. 2020). Extended peptides are composed of a large proportion of proline and glycine residues and have no specific secondary structure, but when in contact with membranes, they generally form an amphipathic helical structure (Koo and Seo 2019). Proline-rich short-chain extended peptides, such as indolicidin and tritrpticin, can be isolated from mammals and from insects, such as apidaecin. Glycine-rich extended peptides can be isolated from insects and have sizes ranging from 8 to 30 kDa (Wang et al. 2019).

3 Upstream Process Development

Large-scale AMPs production is definitively a challenging task. The direct isolation and purification of these molecules from natural sources is normally extremely labor-intensive, in addition to resulting in low yields (Wibowo and Zhao 2019). Chemical synthesis, including Fmoc and other methodologies, is one of the main methods

currently used to obtain purified AMPs with high biological activity. Despite providing a high yield and level of purity, the high manufacturing cost of this technique is a limiting factor for the development of AMPs as biopharmaceuticals, particularly for peptides with more than 35 residues and that have post-translational modifications (Deng et al. 2017; Wibowo and Zhao 2019). Peptide synthesis by SPPS (Solid Phase Peptide Synthesis) is a complex and expensive production technique, despite being efficient. AMPs manufacture by SDDS is estimated to cost around US$50–400 per gram of amino acid produced (Moretta et al. 2020). Alternative approaches that increase production and make the AMPs development process cheaper are necessary (Sinha and Shukla 2019).

Upstream bioprocessing consists of several stages to achieve high yield and final product quality. Through recombinant DNA technology, scalable, economical, and sustainable AMPs production is possible. This strategy allows the cloning of foreign genes into specific vectors for expression in host systems, such as bacteria and yeast (Jozala et al. 2016; Wibowo and Zhao 2019; de Oliveira et al. 2020). The positive screening and selection of clones is performed, as well as small-scale assays to evaluate cell growth and protein product levels. Lastly, bioprocesses are conducted in bioreactor systems for large-scale production, and batch, fed-batch, and continuous strategies are applied for the mass production of the recombinant AMPs of biopharmaceutical interest (Gronemeyer et al. 2014; Jozala et al. 2016; Tripathi and Shrivastava 2019). Figure 1 exemplifies the upstream bioprocess of AMPs biopharmaceutical manufacturing.

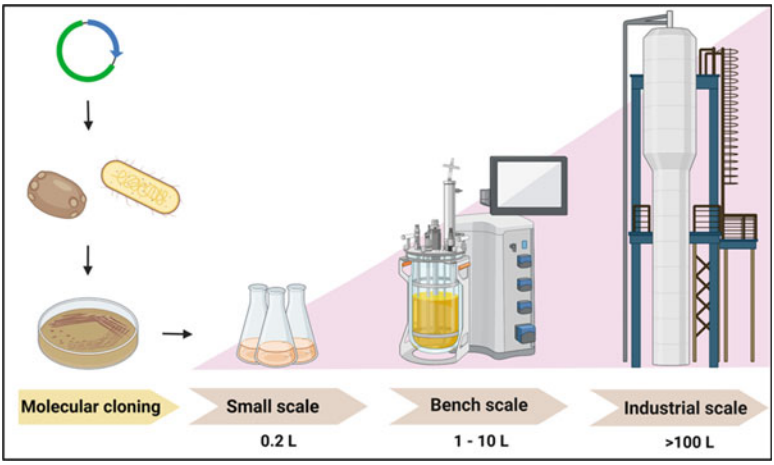


Fig. 1 Upstream bioprocess of AMPs biopharmaceutical manufacturing: molecular cloning, small-scale assays and scale-up of bioprocess. The figure was created by BioRender

3.1 Recombinant AMP Production in Microbial Cells

Bacterial expression systems are the type most used for heterologous gene expression since they are capable of producing a high level of recombinant molecules, show rapid multiplication, and have a simple media requirement (Gomes et al. 2016). Among them, the Gram-negative bacterium *E. coli* is the most extensively used for amplifying recombinant peptides' gene expression (Terpe 2006; Ahmad et al. 2018). In this regard, ~99% of proteins and peptides deposited in the Protein Data Bank (PDB) were produced in the *E. coli* expression system (de Oliveira et al. 2020). Under ideal culture conditions, bacterial cell number practically doubles every 20 min (Sezonov et al. 2007). Additionally, this host is also attractive due to its well-established genetic and expression protocols, wide availability of commercial expression vectors and cost-effectiveness (Sinha and Shukla 2019). There are several reports in the literature about the use of *E. coli* cells for recombinant AMPs expression. Recently, the proline-arginine-rich cationic peptide PR-39 was expressed fused to the SUMO or intein-chitin binding domain (CBD) in *E. coli* BL21 (DE3) pLysS, aiming to compare the expression level. Comparing the fusion protein concentration and single-step purification aspects, the intein system was better, even though similar amounts of pure PR-39 were recovered (250 μg L^{-1} of SUMO and 280 μg L^{-1} of intein) (Azari et al. 2020).

The *E. coli* (Rosetta DE3) system was also chosen to express more than 250 μg mL^{-1} of the human β-defensin 118 (DEFB118). This defensin is related to epididymal innate immunity, protecting the sperm against microorganism attacks in both male and female reproductive tracts. Antimicrobial assays demonstrated that DEFB118 activity occurred against Gram-negative and -positive bacteria at a minimum inhibitory concentration (MIC) of 4 μg mL^{-1} (Lin et al. 2020). Also using *E. coli* BL21 (DE3), a group of scientists created a modular system to express the lanthipeptide mersacidin, produced by *Bacillus amyloliquefaciens*. Mersacidin has bactericidal activity against several Gram-bacteria species, including methicillin-resistant *S. aureus* (Viel et al. 2021).

In recent years, *Bacillus subtilis* has been used as an alternative to the *E. coli* expression system for recombinant AMP production. *B. subtilis* is an endospore-forming Gram-negative soil bacterium and is generally recognized as safe (GRAS) because it lacks endotoxins, so it has been exploited as a production host for aquaculture industries (Pan et al. 2016; Cui et al. 2018). One of the greatest advantages of this bacterium is the possibility of naturally secreting the recombinant peptide (simplifying the downstream process), reducing hydrolysis of cell-associated proteins. Moreover, it simplifies the detection and purification processes of the target molecule (Zhang et al. 2020). Besides, *B. subtilis* does not produce lipopolysaccharides (LPS), preventing some degenerative disorders in humans and animals (Gomes et al. 2016).

Although bacterial cell-based expression systems are mostly used for the production of recombinant molecules, they are more prone to the degradation of cationic peptides (Li 2011). In addition, *E. coli* cells can produce inclusion bodies formed by

insoluble protein aggregates that hinder the recombinant AMP extraction and purification process (Gomes et al. 2016). Although engineered *E. coli* expresses periplasmic disulfide bond isomerase (DsbC) in the cytoplasm or through exporting the periplasm recombinant peptide (Abbas et al. 2013), this bacterium does not carry out other complex post-translational modifications (Parachin et al. 2012). Thus, for cases where there is a need for post-translational modifications, other systems should be chosen, such as yeasts or bacterial strains engineered for this purpose.

In addition to post-translational modifications, yeasts can carry out the correct protein folding. However, they have as a disadvantage the hyperglycosylation mechanisms and the need for aerobic fermentation, which reduces the growth rate, resulting in a lower recombinant protein yield (Juturu and Wu 2018). *S. cerevisiae* (baker's yeast) is a model organism for heterologous expression because its cell biology, genetics and biochemistry are well described (Gomes et al. 2016). This yeast is a suitable expression system for recombinant peptide production, as demonstrated by a group of scientists (Jiang et al. 2021). They expressed the cecropin P1, a positively charged α-helical peptide, isolated from the nematode *Ascaris suum*, which shows activity against *E. coli*, *Salmonella* sp., *Shigella* sp., and *Pasteurella* sp. In addition, this peptide also has antiviral activity against the PRRSV NADC30-like strain (Jiang et al. 2021). Defensins have also been successfully expressed in *S. cerevisiae*. By using the *MET17* promoter, Møller and colleagues expressed β-defensin-2 in *S. cerevisiae* cells (Møller et al. 2017).

In recent decades, the yeast *Pichia pastoris* (reclassified as *Komagataella phaffii)* (Naumov et al. 2018) has been extensively used for heterologous production of peptides and proteins, as it is capable of performing disulfide bridges, O- and N-glycosylation, and the correct processing of signal sequences (Wibowo and Zhao 2019). Several peptides have been expressed in a *K. phaffii* system (Wang et al. 2009; Basanta et al. 2010; Zhao et al. 2015; Zhang et al. 2018). Tachyplesin I (TP-I), a cationic peptide isolated from the Japanese horse crab (*Tachypleus tridentatus*) hemocytes, which inhibits the bacterial lipopolysaccharide (Li et al. 2019), was expressed in this system. Another example is the expression of the immunomodulatory and anti-inflammatory hybrid peptide (IAHP) LL-37Tα1. This peptide was efficiently produced in *K. phaffii* cells, demonstrating their ability to produce recombinant bioactive peptides (Ahmad et al. 2019).

Even though the choice of expression system must be made considering recombinant molecule properties, systems based on prokaryotic and eukaryotic cells have both strengths and weaknesses. The disadvantages can be overcome with multiple strategies, aimed at increasing recombinant molecule production. Codon optimization for specific organisms or engineered strains, the use of strong promoters and multimeric AMP expression in tandem or fused to a higher molecular mass protein are strategies that allow for a stable and high level of recombinant production (Fig. 2) (Deng et al. 2017). Below, we will address all these strategies to improve recombinant AMPs production.

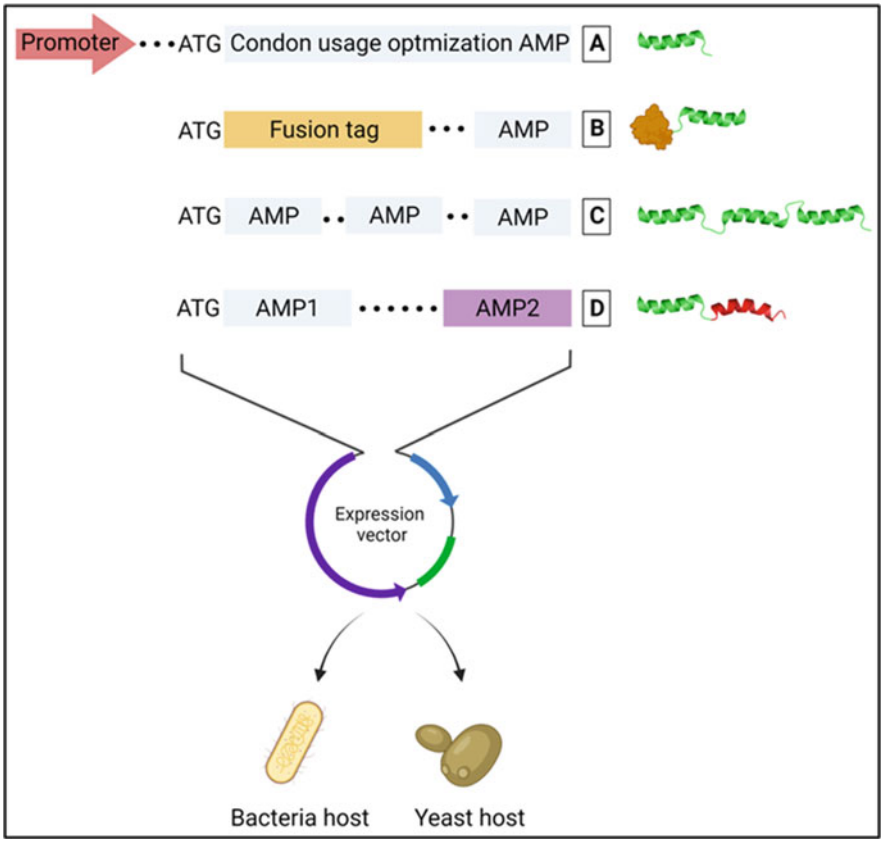


Fig. 2 Multiple strategies to increase the production of recombinant AMPs. (A) The use of strong promoters and codon optimization, (B) fusion to a higher molecular mass protein, (C) multimeric AMP expression in tandem and (D) hybridization of different AMPs. The figure was created by BioRender

3.2 Strategies to Enhance the Heterologous Expression Level

3.2.1 Cell Line Engineering and Host Strain

The choice of strain affects recombinant production success. At the end of the twentieth century, various potentially advantageous characteristics for heterologous production were tested in *E. coli* strains, generating line B, of which BL21(DE3) is the favorite host (Rosano et al. 2019). *E. coli* BL21 (DE3), and other parental B strains, have a deficiency in Lon protease, preventing exogenous protein degradation, an advantage to this system. In addition, the gene encoding the outer membrane protease OmpT is absent from the *E. coli* BL21 (DE3) genome, preventing the degradation of extracellular proteins (Rosano and Ceccarelli 2014). Another

advantage of the BL21 strain is that it expresses genes under the RNA polymerase promoters' control, such as lac, tac, trc, ParaBAD, PrhaBAD, T5, and T7 (de Oliveira et al. 2020). BL21(DE3) is able to provide a high AMP expression level, as observed in a study in which the yield of peptide P-113, derived from human saliva protein histatin 5, was 4 mg L^{-1} of bacterial suspension (Cheng et al. 2018).

Moreover, by transforming the *B. subtilis* (WB800) strain with a recombinant plasmid pHT-CI-CBF, scientists expressed cathelicidin-BF (CBF), a peptide isolated from snake venom (*Bungarus fasciatus*), fused with intein (Vogt et al. 2016). Data demonstrated that the intein expression system is a safe and efficient method by which to produce recombinant proteins in *B. subtilis,* since the yield of fusion CBF secreted reached ~0.5 mg L^{-1} (He et al. 2015). In addition to secretory capacity, this system can be exploited to produce recombinant molecules exposed on biofilm surface. Vogt and colleagues selected peptidic regions from tropomyosin and paramyosin, from the parasite *Echinococcus granulosus*, producing the peptides EgTrp and EgA31, respectively. They fused these peptides directly to the TsaA C-terminus, an important matrix protein (Vogt et al. 2016).

B. subtilis cells were genetically modified by the addition of a small tobacco etch virus (TEV) protease and the two-cistron expression vector gene of abaecin, previously isolated from *Apis mellifera* and further fused to TEV (Li et al. 2017). It has also been demonstrated that an expression system based on *B. subtilis* cells is capable of expressing a eukaryotic non-ribosomal peptide synthetase (esyn) gene that codes for the biosynthesis of the enniatin molecule (Zobel et al. 2015). Such examples reinforce the versatility of this system to produce recombinant peptides. In *K. phaffii*, recombinant peptides and proteins can be degraded through proteolysis mechanisms during transport or even in the extracellular space. To circumvent this limitation, protease-deficient strains such as SMD1163 (pep4 prb1 his4), SMD1165 (prb1 his4), and SMD1168 (his4 pep4) can be used. Although they have reduced proteolytic activity, due to gene silencing, these strains grow more slowly than the wild type, in addition to having low transformation efficiency and short cell viability (Daly and Hearn 2005; Ahmad et al. 2014). Even so, *K. phaffii* recombinant peptides' yield can be higher, as in the case of defensin VpDef, isolated from the mollusk *Venerupis philippinarum*, which was produced at a concentration of 60 μg mL^{-1} of culture medium (Meng et al. 2018). Another strategy to improve recombinant AMP yield is the use of both strong constitutive and inducible promoters.

3.2.2 Promoters and Codon Usage Optimization Strategies

The DNA region responsible for given gene transcription, or promoter, is another important component for recombinant AMPs production. Certain hosts have compatibility with different types of promoters. There is currently a plethora of commercially available promoters. Promoter T7, commonly found in pET series expression vectors, is a strong promoter most used for AMPs expression in *E. coli* cells (Deng et al. 2017). Promoters lac, tac, lacUV5, and T5 are also widely used

today, and research for the development of new and more efficient promoters continues to be developed. Inducible promoters were developed and derived from the constitutive phage promoters T5 (T5N25) and A1 (T7A1), capable of being recognized by RNAP of *E. coli* σ^{70}. Promoters not only decreased the basal expression but also increased the production of the recombinant protein (Schuller et al. 2020).

Another strategy to increase gene expression levels in bacterial systems is the use of dual promoters. Unlike eukaryotic systems, which have two separate expression cassettes, the dual promoters used in prokaryotic organisms are chimeric and are located, in tandem, upstream from the gene to be expressed (Öztürk et al. 2017). Using an optimized dual-promoter (P_{HpaII}–$P_{amyQ'}$) system, a group of scientists has demonstrated the increasing extracellular expression of β-CGTase, pullulanase, and α-CGTase in the *B. subtilis* (CCTCC M 2016536) strain. This strain was genetically manipulated to delete the *srfC*, *spoIIAC*, *nprE*, *aprE,* and *amyE* genes (Zhang et al. 2017).

In yeasts, the use of constitutive and inducible promoters is common. When compared to constitutive promoters, the inducible promoters are more used, because they allow a certain control over gene expression, which ends up resulting in a higher yield of recombinant peptides. Constitutive TEF1 and GPD promoters are the ones most used for *S. cerevisiae* expression. Although they can cause aggregation of folded proteins, making their secretion difficult, the promoters ADH1, GAPDH, PGK1, TPI, ENO, PYK1 also present an alternative (de Oliveira et al. 2020). Another factor that can influence the efficiency of promoters is polymorphism, which is already being evaluated (de Paiva et al. 2018).

Although the use of strong promoters could increase the heterologous expression levels, alternative strategies such as codon usage optimization can be used to boost the recombinant peptide expression. The codon usage bias refers to the availability of certain codons over others in the genome of organisms (Hanson and Coller 2018), and it can be a challenge for heterologous expression. Different species have a frequency of codons in a DNA sequence that is positively related to the corresponding tRNA, and the tRNA concentration in a cell is decisive for the number of amino acids accessible for protein translation extension (Fu et al. 2020).

The codon optimization approach had been used, in synthetic biology and the metabolic and cellular engineering fields, as an alternative by which to enhance heterologous gene expression levels (Lanza et al. 2014). In other words, codon optimization is suggested as a crucial factor in gene expression, since it consists of modifying synonymous codons through genetic engineering techniques, resulting in increased protein up-regulation and RNA levels (Mauro and Chappell 2014; Zhoua et al. 2016). In addition to contributing to mRNA stability, codon optimization also can impact ribosome translocation, connecting the processes of translation, elongation, and decay (Presnyak et al. 2015). Due to its importance to heterologous gene expression, scientists have developed, in recent years, new tools to facilitate codon optimization of synthetic genes. Such approaches explore the self-learning capacity of artificial intelligence (Tian et al. 2017; Fu et al. 2020), computational procedures (Chung and Lee 2012), and mathematical algorithms (Taneda and Asai 2020; Sen

et al. 2020). Organisms such as *B. subtilis* have another advantage, since the bias in codon usage is not a determining factor for heterologous expression. Additionally, transcription, translation, folding, and protein secretion processes, as well as methods of genetic manipulation and large-scale bioprocesses of this organism, are well described in the literature (Gomes et al. 2016).

Just changing a single codon synonym is enough to increase protein expression. The introduction of single codon synonym mutation (TCT → AGT) in the gene encoding the mrTNF-PADRE recombinant vaccine resulted in their enhanced production (~30% of total *E. coli* proteins) (Chu et al. 2018). It is also possible to improve translation efficiency in *S. cerevisiae* using the "condition-specific codon optimization" approach, as demonstrated by Lanza and colleagues (Lanza et al. 2014).

3.2.3 Tandem Multimeric Expression and Fusion Proteins

Another strategy commonly used for enhancing recombinant AMPs production is the tandem multimeric expression (Deng et al. 2017). Tandem peptide expression is an approach used for both prokaryotic and eukaryotic (yeasts) systems (Zhou et al. 2005; Fida et al. 2009; Wang et al. 2012). The expression of peptide LfcinB15-W4,10, a bovine lactoferricin, in four tandem repeats is more effective than monomers and the other repeats (2–8). At the end of the purification process, 10 mg of the tetramer with 99% purity was achieved (Tian et al. 2007). As demonstrated in the previous work, although it is possible to improve AMPs expression through multimeric tandem expression, the synthesis efficiency is not, unfortunately, proportional to copy number (Lee et al. 2002).

However, another strategy to overcome the low yield of recombinant AMPs is by expressing them fused to other proteins called fusion proteins which have been widely used in the expression of heterologous AMPs (Beaulieu et al. 2007; Liu et al. 2011; Mulder et al. 2015; Sousa et al. 2016; Xiao et al. 2017; Kaur et al. 2020). In addition to increasing the production level, these proteins facilitate the solubilization and purification of recombinant molecules (Costa et al. 2014). The small ubiquitin modifying (SUMO) fusion protein has an important function in AMPs' solubilization (Butt et al. 2005). Like SUMO, thioredoxin (Trx), a fusion protein from *E. coli*, also facilitates the solubilization of recombinant AMPs. However, unlike the former, Trx allows an increased expression rate (LaVallie et al. 1993; Costa et al. 2014). Another strategy is the use of polyhistidine ($_6$His) tag in the C- or N-terminal regions of a recombinant peptide. Although this tag is not a fusion protein, it facilitates the purification and detection by the western blot technique (Li 2011; Tavares et al. 2012; Belguesmia et al. 2020; Costa Ramos et al. 2021; Zhan et al. 2021).

3.2.4 AMP Hybridization

Hybridization, a combination of two native peptides, or even derivatives of hybrid peptides, is a method that has been used to produce new hybrid AMPs, in order to increase their antibacterial action, with reduced cytotoxicity (Wu et al. 2014; Klubthawee et al. 2020). A new hybrid AMP combined the α-helical fragments from peptides BMAP-27 and OP-145, producing the peptide H4, with the aim of maintaining its potent antimicrobial activity and reducing the cytotoxic profile against mammalian cells. The peptide H4 demonstrated activity against a broad spectrum of both Gram-negative and -positive bacteria, including the multidrug-resistant bacterial strains, in the range of 2.5–25 μM (Almaaytah et al. 2018).

Furthermore, the combination of two peptides can increase the plasma membrane permeabilization of the "parent" peptide. In order to evaluate the potential increase in activity of the hybrid peptides, hybrids were developed by joining membrane permeabilizing peptides (parasin or magainin 2) with membrane translocating peptides (DesHDAP1 or BF2). The results suggest that the permeabilizing activity is increased when the parent permeabilizing peptide is placed at the N-terminus, and through the addition of an alanine spacer between the sequences of the two parent peptides (Wade et al. 2019). Hybrid AMPs can also increase both the selectivity and stability of the molecule when compared to naturally occurring peptides (Yang et al. 2020). Through bioinformatic analysis, the 3.35 kDa hybrid magainin-thaumatin (MT) peptide was designed and expressed in *E. coli* BL21 (DE3) cells. Recombinant MT showed an inhibitory effect against *S. aureus*, *E. coli* DH5α, and *B. subtilis* at the MIC of 6.5, 20, and 9 μM, respectively (Tian et al. 2019).

Some hybrid peptides have strong antimicrobial activity against bacterial hosts, requiring the use of other expression systems. In this context, from hybridized plantaricin E (PlnE) and plantaricin (PlnF), type IIb bacteriocins, a ~5 kDa EF-1 hybrid peptide was developed (Li et al. 2020). This peptide can directly induce cell membrane permeabilization of *E. coli* cells. Expressing the recombinant EF-1 in the *K. phaffii* host, they recovered a yield of 32.65 mg L^{-1} with a purity of 94.9%. In addition to the bactericidal activity against enterohemorrhagic *E. coli* (EHEC) (MIC = 6.25 μM) and *E. coli* K88 (MIC = 3.125 μM) cells, recombinant EF-1 has no hemolytic activity (Li et al. 2020). The hybridization approach has the great potential to overcome some drawbacks, and it has been widely used (Jin et al. 2006; Xu et al. 2007; Arbulu et al. 2019; Agbale et al. 2019). Even though these strategies may have limitations, their use, alone or in combination, can increase the yield of recombinant AMPs, revolutionizing the production of biomolecules of medical and pharmaceutical interest.

4 Scale-Up from Small- to Large-Scale Fermentation

Bioprocess engineering has made considerable progress due to the high market demand for new biopharmaceuticals. New technologies related to bioprocessing techniques have been acquired for the large-scale production of proteins and peptides with biopharmaceutical potential for the treatment of various diseases. Economically viable production systems, which allow high yields to be achieved while maintaining the desired product quality, are now of great interest to bioprocesses industries (Potvin et al. 2012; Love et al. 2018; Tripathi and Shrivastava 2019).

An important step in the development process is the selection of a proper cell clone for the final production, which needs to fit the product quality requirements, processability and volumetric productivity. Clones can be selected for cell-specific and volumetric productivity, glycosylation profiles, aggregate formation, protein sequence heterogeneity and clone stability, among others. Cell culture conditions are decisive in productivity and product quality (Gronemeyer et al. 2014). AMPs biomanufacturing initially occurs on a small-scale, using shake flasks for expression system development. After this step, bioreactors are used to increase cell density and, consequently, the yield of the recombinant molecule of interest (Wibowo and Zhao 2019).

A defensin-like-peptide-P2 was successfully produced in *K. phaffii* using shake flasks. The induction of the recombinant peptide was performed with 0.5% methanol (v/v) every 24 h during the 120 h induction period, obtaining an expression level of 108.05 mg L^{-1}. Then, the process was scaled up, cultivation was carried out in a 5 L bioreactor, and a total of 1.69 g L^{-1} of peptide P2 was achieved. P2 exhibited bacterial reduction activity of 80–97% against multi-resistant *S. aureus* in RAW264.7 macrophages, among other activities (Yang et al. 2019).

Bacterial-based cell systems have also been used to express recombinant AMPs in shake flasks. One example is fowlicidin-2 expressed in *E. coli* BL21 (DE3). The induced expression of the peptide occurred by the addition of IPTG at a final concentration of 0.3 mM for 4 h in Luria-Bertani (LB) medium, at a temperature of 37 °C, under agitation. The results indicate a yield of 202 mg L^{-1} of the peptide of interest. The recombinant peptide demonstrated significant antimicrobial activity against a wide range of Gram-negative and -positive bacteria (Feng et al. 2015). Production of several other peptides was achieved on a small-scale using shake flasks (Sang et al. 2017; Meng et al. 2019).

Controlling certain parameters during production processes can make a big difference in AMPs yield. Monitoring critical operating parameters, including agitation, aeration, dissolved oxygen (DO), temperature, pH, and feed, is important and can be controlled in bioreactors. By monitoring the available oxygen rate, for example, it is possible to increase the availability of oxygen when it is low due to high cell density (Wibowo and Zhao 2019; Tripathi and Shrivastava 2019). The stability of these parameters allows high cell density and greater specific yield with a quality product to be achieved. In view of this, the physiological characterization of

production strains is essential for the proper development of a bioprocess (Gupta and Shukla 2017; Tripathi and Shrivastava 2019).

The composition of the culture medium also significantly influences cell growth and protein yield. The optimal selection of sources of carbon, nitrogen, salts, minerals and some growth factors and their proper concentrations are essential to achieve higher cell density and a higher level of recombinant proteins. The different carbon sources that are used as the main components of the cultivation medium in bioproduction, for instance, significantly affect cell metabolism, protein production and quality (García-Ortega et al. 2019). Complex, chemically defined or even semi-defined culture media can provide a nutrient-rich environment (Wibowo and Zhao 2019).

The parameters for obtaining a greater expression level of the hybrid magainin-thanatin (MT) recombinant AMP, produced in *E. coli*, were evaluated in shake flasks. Induction time (0, 1, 2, 3, 4, 5, and 6 h), temperature (32, 35, 37, 39 and 42 °C), IPTG concentration (0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, and 1.2 mM) and the culture medium were evaluated separately in different cultures. Maximum production of the MT was observed by cultivation in TB medium at 37 °C and induction with 0.8 mM IPTG for 5 h (Tian et al. 2019). In another work, the recombinant AMP UBI18-35, derived from the natural human AMP ubiquicidin, was also expressed in *E. coli* Rosetta (DE3) pLysS. SDS-PAGE electropherogram processing by densitometry demonstrated that the ideal conditions for the best production of AMP UBI18-35 include expression at 28 °C for 4 h after induction with 0.5 mM IPTG (Ashcheulova et al. 2018).

After carrying out the cultivation in shake flasks, cultures on larger scales are desired to obtain a superior recombinant AMPs yield. In this context, different bioprocess techniques can be employed. One example is a simple batch mode of cultivation in which the supply of essential nutrients only occurs at the beginning of bioprocess. Otherwise, in a fed-batch mode, nutrients are added at specific rates throughout the process. Commercial production of some antibiotics, such as penicillin, occurs in fed-batch. In continuous cultivations, the provision of nutrients to microorganisms constantly occurs through the addition of a fresh culture medium, and part of the culture is removed to collect the product simultaneously (Li et al. 2014b; Tripathi and Shrivastava 2019).

4.1 Batch Processes

Simple discontinuous batches can be useful in initial bioprocessing applications, for studies of the physiology and kinetic parameters of the microorganism. During this phase, which lasts between 24 and 30 h, the addition of the substrate occurs only at the beginning of the cultivation, which extends until it reaches a certain concentration of cells, or until the initial substrate is completely consumed during the process. As there is no other substrate addition during the procedure, cell density and yield are limited. Another disadvantage is that during this phase the accumulation of

secondary metabolites can occur, which can be toxic to the host microorganism, and may cause a limitation in obtaining the product of interest (Ÿztürk et al. 2016; Mears et al. 2017; Blunt et al. 2018).

Large-scale production of the recombinant AMP plantaricin (PlnE) expressed in *E. coli* BL21 (DE3) was performed in a bioreactor with a volume of 12 L of LB. Cultivation was carried out in a simple batch mode at 30 and 25 °C. The expression induction was done with IPTG (0.5 mM) for 3 h when OD600 reached 0.4. According to data presented, the yield of purified plantaricin E was 140 and 180 mg at 30 and 25 °C, respectively, in 12 L of LB after 3 h of induction (Pal and Srivastava 2015). To improve the expression level of the ABP-dHC-cecropin A in *E. coli*, large-scale bioprocess was performed. The production was carried out in a culture medium with 10 g L^{-1} of glucose in a 5 L bioreactor, at a stirring rate of 300 rpm, temperature at 37 °C, dissolved oxygen at 50% air saturation by cascade agitation between 300 and 800 rpm and pH 7.0. After 4 h of cultivation, the recombinant AMP production was performed using 0.5 mM IPTG and maintained for another 3 h. According to the results, 47.3 g L^{-1} of biomass was achieved in the batch, which is 14× more than the biomass obtained in culture flasks with LB medium. The amount of soluble protein recovered was 8643 mg L^{-1} (Zhang et al. 2016).

4.2 Fed-Batch Processes

The discontinuous fed-batch process allows a high biomass concentration to be obtained, and the product of interest is due to the possibility of extending process time, as well as the gradual feeding of the selected substrate in precise quantities (García-Ortega et al. 2019). Generally, it starts with a low concentration of substrate and, when its complete consumption occurs, more substrate is added to keep the fermentation process going without exceeding the ideal level. Product collection only happens at the end of the process, allowing for good sterilization conditions for the procedure (Li 2011). This cultivation strategy is advantageous for many microorganisms such as *E. coli*, *S. cerevisiae*, *B. subtilis*, and *K. phaffii* (Philip et al. 2017). In fed-batches, a simple batch phase is performed before feeding the system, in which the consumption of available substrate and accumulation of biomass will occur (Looser et al. 2015). When the initial carbon source is finished, the batch type is changed to fed-batch, in which the production of the recombinant protein will take place. Usually, the substrates used are the carbon source, together with minerals and trace elements needed (Ÿztürk et al. 2016). The feeding methods commonly used are continuous or constant, exponential and pulsed (García-Ortega et al. 2019).

During continuous and exponential feeding techniques, substrate supply is carried out constantly after simple batch bioprocess. Continuous feeding is based on nutrient addition at a continuous rate throughout the entire process. Exponential feeding allows the specific growth rate maintenance at a pre-defined level. The substrate amount needed for cells to reach the target concentration can be calculated

beforehand, and substrates are constantly supplied throughout the process (Ÿztürk et al. 2016). Due to nutrient addition during the procedure, there is a volume increase inside the bioreactor, causing a dilution of cell concentration. Thus, feeding is exponential due to exponential cell growth (D'Anjou and Daugulis 2000). Feedback feeding happens by monitoring indirect variables, such as oxygen or carbon dioxide dissolved in the system, or by the concentration of the carbon source, such as methanol or ethanol (Ÿztürk et al. 2016). Therefore, feeding can occur through pulses, according to dissolved oxygen increase rate, which indicates a metabolism reduction due to substrate depletion (Zhao et al. 2008; Looser et al. 2015).

The peptide NZX, a plectasin derivative, was expressed in *K. phaffii* (X-33) in a 5 L bioreactor, aiming to boost NZX production. After cell growth on glucose, the induction of NZW expression was performed in a fed-batch with methanol for 120 h. According to the results, the total biomass concentration obtained was 268 g $L^{-1,}$ and secreted protein was 2820 mg L^{-1}. The recombinant peptide showed high stability and low cytotoxicity, in addition to antimicrobial activity against *Staphylococcus hyicus,* both in in vitro and in in vivo experiments (Liu et al. 2020). Another group of scientists also used the yeast *K. phaffii* as a host for AMPs production. According to the data, the AMP clavanin MO was successfully expressed in *K. phaffii*. Therefore, the cultivation of the recombinant strain was then carried out in a bioreactor using BMGY medium with 40 g L^{-1} of glycerol in the discontinuous batch phase. After the total consumption of glycerol, the fed-batch carried out with 0.5% methanol added every 12 h during 72 h was initiated. At the end of the bioprocess, 5 mg mL^{-1} of the recombinant clavanin MO was recovered (Mulder et al. 2015).

Piscidin 4 (TP4), derived from tilapia *Oreochromis niloticus*, was produced in a 500 L bioreactor with a culture medium of basal salts with trace elements. In the fed-batch method, recombinant yeast *K. phaffii* growth was stimulated by adding glycerol as a carbon source until it ran out. Then, the discontinuous phase, also fed with glycerol, was conducted to obtain a higher cell density. Finally, the induction phase with 100% methanol was carried out for 65 h to induce the production of the recombinant TP4. Fed-batch bioprocess in recombinant strain *K. phaffii* (KM71) XS10 was also performed for the production of the recombinant bovine lactoferrampin–lactoferricin (LFA–LFC). The batch was kept for 12 h in a 10 L bioreactor with basal salts medium containing trace elements. After this step, a glycerol feed (50% (v/w)) was started at 10 mL/h for 12 h. Six hours later, peptide induction was initiated by the addition of 0.5% methanol (v/v) every 24 h. According to the results, it was possible to obtain about 1.025 $\pm$ 169 mg L^{-1} of LFA-LFC in the culture supernatant and about 53 $\pm$ 4 mg L^{-1} of the purified LFA-LFC (Tang et al. 2012).

4.3 Continuous Processes

Although relevant studies have already been published about discontinuous fed-batches, they can be less robust, more laborious and time-consuming. While the continuous mode allows for stable culture conditions, it provides a similar physiological state for all cells in the culture medium and, for these reasons, it has been one of the most widely used strategies for obtaining physiological data (García-Ortega et al. 2019). In the continuous mode, the target compound can be produced in large quantities due to the system maintenance at stationary phase. Fresh culture medium is added to the bioreactor, and part of the culture is continuously removed at a constant value: the dilution rate (D). This bioprocess is also known as chemostat (chemical and static environment), since the conditions inside the reactor (substrate, cell and product concentrations) can be stable (Mears et al. 2017; Blunt et al. 2018). Other parameters such as pH value, oxygen rate, working volume and nutrient supply are also kept constant (Koller 2018).

Continuous processes can keep the growth rate moderate and constant for long periods, which avoids the non-productive time spent on harvesting, cleaning, preparing a new medium, sterilizing and cooling the culture in a discontinuous process, generating higher average yield (Blunt et al. 2018). Among other advantages, it allows knowledge about bioprocess physiology, making it less laborious for the operation and maintenance of the process when it reaches a steady state (Blunt et al. 2018). Furthermore, the inhibition risk by substrate, or by-products, is lower as the final product is continuously collected throughout the process (Li et al. 2014a).

The interest in continuous cultures is not restricted to studies of microbial physiology and process development. It is also of great interest for the manufacture of recombinant proteins, as it allows cells to be kept in production states for a longer time, and consequently, the yield of the process significantly increases, while costs fall (Peebo and Neubauer 2018; Khanal and Lenhoff 2021). Several examples of continuous processes to produce recombinant proteins in different microorganisms have been reported in the scientific literature. However, some shortcomings associated with continuous processes make their commercialization limited (Rathore et al. 2015; Peebo and Neubauer 2018).

Both the stability and sterility during a long period of cultivation, as well as the lack of flexibility in the short term (caused by the need for long periods of execution), associated with the genetic inability of the cells can be cited as some limitations of this method of bioprocess (Rathore et al. 2015; Peebo and Neubauer 2018). Mainly due to these issues, continuous production of recombinant proteins is used particularly for the manufacturing of high-demand biopharmaceuticals. Recombinant insulin, produced in *S. cerevisiae* in the 1990s, is the only known example of a continuous industrial recombination process using microorganisms (Diers et al. 1991).

5 Purification of AMPs: Downstream Process Development

The downstream process consists of the recovery and purification of recombinant peptides, aiming to reduce costs (Clarke 2013). In this regard, innovative approaches have been applied in purification techniques of the downstream pharmaceutical industry (Gupta and Shukla 2017). These techniques are divided into three different stages: (a) an initial recovery (extraction or isolation), (b) purification (removal of most contaminants), and (c) polishing for removal of specified contaminants (Zydney 2016; Gupta and Shukla 2017; Tripathi and Shrivastava 2019). Taken together, all these steps improved the search, production and application of therapeutic peptides (Agyei et al. 2017). The purification process for recombinant AMPs is demonstrated in Fig. 3.

5.1 Recovery

At the recovery stage, the most common techniques applied in the industries are centrifugation, tangential flow microfiltration (MF-TFF), or depth filtration. The objective is to remove cells, fine particulates, colloids and soluble impurities prior to the initial purification steps (Pieracci et al. 2018). Generally, organisms such as yeasts export the recombinant peptides to the extracellular space, while in bacterial systems, the heterologous molecules are sent to the periplasmic space. To recover the extracellular recombinant peptide, it is necessary to concentrate it by a centrifugation

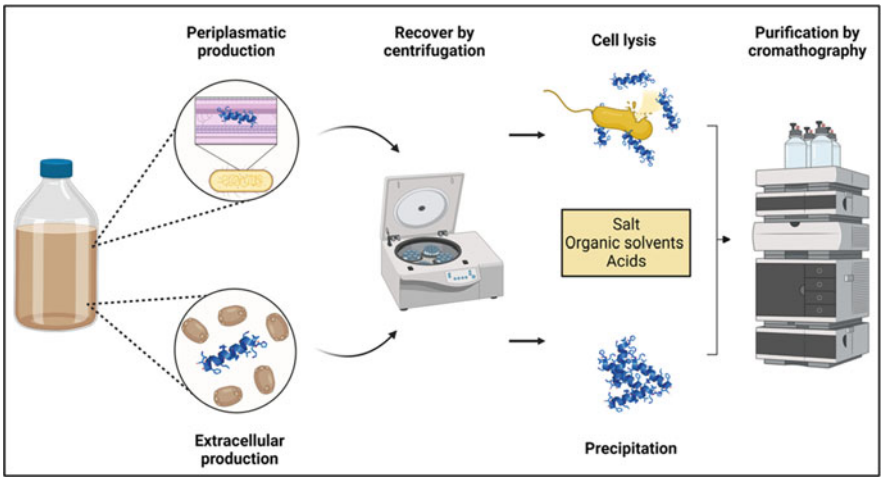


Fig. 3 Purification process of recombinant AMPs. The process addressed to both the periplasmic and extracellular spaces, through a signal peptide, can initially be recovered by centrifugation and, after this, from the cell lysis process (periplasm) or by precipitation (extracellular). The final step consists of the purification by chromatographic techniques. The figure was created by BioRender

or ultracentrifugation process. In addition to chromatography, the precipitation of the samples can be done to improve their concentration (Jozala et al. 2016).

The presence of a signal peptide located in the C or N-terminus region allows the export of the recombinant peptide to the extracellular (in *K. phaffii*) or periplasmic (in *E. coli*) space (Weinacker et al. 2013). For purified peptides in periplasmic space, the cells are submitted to lysis (sonication, high-pressure homogenizer, passing through mills, etc.) and clarification can be used to remove the cell debris. Therefore, the clarified product is purified using precipitation and/or chromatography. In some cases, recombinant peptides produced in periplasmic space can be accumulated as inclusion bodies (IBs), which are inactive, making an extra step necessary to refold the peptides to their native conformation (Ehgartner et al. 2017; Tripathi and Shrivastava 2019).

5.2 Purification

After the recombinant peptide recovery, the next step is the purification process by chromatography techniques (Tripathi and Shrivastava 2019). There are several types of chromatography, such as size exclusion chromatography, ion-exchange chromatography, low-pressure hydrophobic interaction chromatography, countercurrent distribution, partition chromatography, and reverse-phase high-performance liquid chromatography (RP-HPLC) (Tripathi and Shrivastava 2019). Although the chromatography techniques are well-established processes, they are more expensive, when compared to other purification approaches, as well as presenting limitations in throughput and scalability (Tripathi and Shrivastava 2019). To overcome some of those limitations, it is important that the purification strategy must be evaluated for every single case, and it can change according to the particular application (Carta and Jungbauer 2010). In this context, several AMPs have already been purified using chromatography techniques.

The host defense peptides (HDPs) IDR1, MX226, LL37, CRAMP, HHC-10, E5, and E6 were purified in a simplified 2-step purification method. After recombinant production in high yields, using *E. coli* strain BL21 with SUMO fusion, the peptides were first purified using a Ni-NTA Sepharose column. The recombinant peptide IDR-1 was successfully purified. From a yield of about 1.5 kg of wet biomass, 6% of the total protein produced is equivalent to the SUMO-IDR-1, with an estimated yield of 0.48 g/L of fermentation. Subsequently, the SUMO-specific protease sumoase was used to release the peptide from the fusion. Following the cleavage, reversed-phase chromatography was used to separate SUMO and sumoase from the peptide and recover the homogenously pure peptide (Bommarius et al. 2010).

6 Optimization of the Industrial Processes

Recently, different methods have been used in order to optimize the industrial process, such as high-throughput devices (HTD), design of experiments (DoE), and process analytical technology (PAT). The HTD method is divided into screening (HTS), process development and experimentation (HTPD), and experimentation (HTE). These processes together seek to obtain data quickly and reliably, doing tests on a small-scale, producing multiple data points to be generated with low usage of laboratory space, in addition to saving time (Silva et al. 2021).

HTS is responsible for producing hundreds of thousands of data points per day, using microtiter plates with 384 and 1536 wells (Mayr and Bojanic 2009). Another option is microfluidics, which use very small amounts of liquids, with quick analysis generating multiple data points and thus reducing time and laboratory space (Whitesides 2006). The goal of HTS is optimization of the process and delivering the final product actively and quickly (Shukla et al. 2017; Silva et al. 2021). HTPD makes use of complex computational tools to seek to establish a better design and understanding of the process (Bhambure et al. 2011). HTE is the union of HTS and HTPD to provide an overview of the miniaturization and automation process, enabling applicability in the biopharmaceutical industry (Silva et al. 2021; São Pedro et al. 2021).

In this regard, a group of scientists reported the combination of two different methods (high-throughput (HTP) processes and flow-mediated synthesis). They rapidly obtained a large amount of information about new peptides. They tested a library of potential AMPs against *P. aeruginosa*, and the results allowed them to identify the peptide with the highest antimicrobial activity. The workflow contributes to the discovery and optimization of peptide structures as new antimicrobial agents for biomedical applications (Judzewitsch et al. 2020). The suitability of both the diffusion self-interaction parameters (kDa) and osmotic second virial coefficients (B22) was investigated by using high-throughput screening, aiming to seek the peptides' formulation aggregation risk. In this study, thermal stress effects on colloidal stability were evaluated, as well as six buffering systems at two-selected pH values, four tonicity agents and a common preservative. Acetate and succinate buffer at pH 4.5, combined with glycerol or mannitol, demonstrated greater stability of peptides. Data suggest that HT methods display important information about the optimization of colloidal stability during the early development of peptide-based liquid formulations (Dauer et al. 2021).

In addition to the HTS method, DoE also allows bioprocesses to be optimized through monitoring and control of the experiment. DoE enables simultaneous evaluation of a large number of variables. Thus, this methodology obtains maximum information about the process, saving time and financial resources to obtain a high-quality product (Kasemiire et al. 2021). This methodology was used to optimize the AMP human β-defensin 2 (HBD2) expression yields in *E. coli* BL21 (DE3). The authors used 24-factorial design of experiments (DoE) to evaluate the following variables, namely cell density, temperature, induction period, and inducer

concentration. They got 19 different combinations, and the best condition was pre-induction temperature of 37 °C, a cell density of 1.0 U (600 nm), an induction temperature of 20 °C and a 0.1 mM of gene expression inducer (IPTG) over 4 h. Researchers concluded that these conditions produced the HBD2 peptide in a higher proportion than previously tested (Corrales-García et al. 2020).

Process Analytical Technology (PAT), initially from the US Food and Drug Administration (FDA), is responsible for the regulation of biopharmaceutical production, dealing with measuring, analyzing, monitoring, and ultimately controlling all substantial characteristics of a bioprocess (Scott and Wilcock 2006; Kornecki and Strube 2018). It is important to remember that the transition from up- to downstream in a bioprocess is always challenging. Any small change can affect final product quality. PAT was used for monitoring and controlling specific reactions during the manufacturing process of the recombinant lethal toxin-neutralizing factor (rLTNF), which acts against rattlesnake venom. Monitoring was performed in three critical production stages (solubilization of the IBs, enzymatic cleavage with α-chymotrypsin and quenching of the reaction at the optimal time). The authors reported that the process with PAT tools through various batches of rLTNF production would allow real-time analysis of quality and production control (Hebbi et al. 2020). Therefore, the methods described above are of great importance, as they seek to find the best operating conditions, in addition to determining the operating windows. These features are extremely useful in minimizing the impact of batch-to-batch variations and human error during the bioprocess.

7 Conclusions and Future Directions

Bioprocessing technology is a great opportunity for developing recombinant therapeutic AMPs on a large-scale. The promiscuity of these molecules makes them promising alternatives in the development of new biopharmaceuticals against a number of infectious diseases. The use of microorganisms and different fermentation processes allows large-scale high-quality AMPs production. Therefore, optimization of process conditions seems to be extremely important to generate biologically active and stable molecules. Different strategies can be applied for cell line engineering, such as promoters and codon usage optimization, tandem multimeric expression and fusion protein production, and AMP hybridization.

General optimization of AMPs bioprocessing seeks the best expression system and conditions for production and recovery of compounds. However, it must be emphasized that there is no ideal production system and that they all have advantages and disadvantages. Analysis of AMPs physicochemical characteristics is essential to find the most useful conditions to produce stable and active molecules, as well as in large quantity. The same occurs during the recombinant strain cultivation on a small or large-scale. Cultivation conditions and bioprocess strategies must be carefully analyzed, as they directly influence the recombinant AMPs expression.

The choice of an appropriate recovery method for an AMP after its production is also very important, as it can allow AMPs with a high degree of purity to be obtained.

In summary, optimization methods for industrial process are important developments for this field. Through best operating conditions, these optimized methods increase the production rate and reduce costs in both upstream and downstream process development. Such improvements are important because they can allow us to produce AMPs in high quantities and purity, and these will then be applied in the pharmaceutical industry, research, and pre-clinical and clinical trials. Furthermore, bioprocessing automation can reduce the impact related to variations during bioprocesses, as well as human error. All approaches related to upstream and downstream processes in the recombinant production of AMPs mentioned here generate lower production costs, manufacturing diversification and recombinant AMPs quality.

Acknowledgments This work was supported by the Coordination of Improvement of Higher Education Personnel (CAPES); National Council of Technological and Scientific Development (CNPq), Federal District Research Support Foundation (FAPDF) and Support Foundation for the Development of Education, Science and Technology of the State of Mato Grosso do Sul (FUNDECT).

References

Abbas A, Plattner S, Shah KH, Bohlmann H (2013) Comparison of periplasmic and intracellular expression of Arabidopsis thionin proproteins in *E. coli*. Biotechnol Lett 35:1085–1091

Agbale CM, Sarfo JK, Galyuon IK et al (2019) Antimicrobial and antibiofilm activities of helical antimicrobial peptide sequences incorporating metal-binding motifs. Biochemistry 58:3802–3812

Agyei D, Ahmed I, Akram Z et al (2017) Protein and peptide biopharmaceuticals: an overview. Protein Pept Lett 24:94–101

Ahmad M, Hirz M, Pichler H, Schwab H (2014) Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol 98:5301–5317

Ahmad I, Nawaz N, Darwesh NM et al (2018) Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*. Protein Expr Purif 144:12–18

Ahmad B, Hanif Q, Xubiao W et al (2019) Expression and purification of hybrid ll-37tα1 peptide in *Pichia pastoris* and evaluation of its immunomodulatory and anti-inflammatory activities by LPS neutralization. Front Immunol 10:1–12

Almaaytah A, Qaoud MT, Abualhaijaa A et al (2018) Hybridization and antibiotic synergism as a tool for reducing the cytotoxicity of antimicrobial peptides. Infect Drug Resist 11:835–847

Almeida LHO, de Oliveira CFR, Rodrigues MS et al (2020) Adepamycin: design, synthesis and biological properties of a new peptide with antimicrobial properties. Arch Biochem Biophys 691:108487

Andersson DI, Hughes D, Kubicek-Sutherland JZ (2016) Mechanisms and consequences of bacterial resistance to antimicrobial peptides. Drug Resist Updat 26:43–57

Arbulu S, Jiménez JJ, Gútiez L et al (2019) Cloning and expression of synthetic genes encoding native, hybrid- and bacteriocin-derived chimeras from mature class IIa bacteriocins, by *Pichia pastoris* (syn. *Komagataella* spp.). Food Res Int 121:888–899

Arora S, Saxena V, Ayyar BV (2017) Affinity chromatography: a versatile technique for antibody purification. Methods 116:84–94

Ashcheulova DO, Efimova LV, Lushchyk AY et al (2018) Production of the recombinant antimicrobial peptide UBI 18-35 in *Escherichia coli*. Protein Expr Purif 143:38–44

Assis LM, Nedeljković M, Dessen A (2017) New strategies for targeting and treatment of multidrug resistant *Staphylococcus aureus*. Drug Resist Updat 31:1–14

Azari M, Asad S, Mehrnia MR (2020) Heterologous production of porcine derived antimicrobial peptide PR-39 in *Escherichia coli* using SUMO and intein fusion systems. Protein Expr Purif 169:105568

Basanta A, Gómez-Sala B, Sánchez J et al (2010) Use of the yeast *Pichia pastoris* as an expression host for secretion of enterocin L50, a leaderless two-peptide (L50A and L50B) bacteriocin from *Enterococcus faecium* L50▽. Appl Environ Microbiol 76:3314–3324

Beaulieu L, Tolkatchev D, Jetté JF et al (2007) Production of active pediocin PA-1 in *Escherichia coli* using a thioredoxin gene fusion expression approach: cloning, expression, purification, and characterization. Can J Microbiol 53:1246–1258

Belguesmia Y, Bendjeddou K, Kempf I et al (2020) Heterologous biosynthesis of five new class II bacteriocins from *Lactobacillus paracasei* CNCM I-5369 with antagonistic activity against pathogenic *Escherichia coli* strains. Front Microbiol 11:1–9

Bhambure R, Kumar K, Rathore AS (2011) High-throughput process development for biopharmaceutical drug substances. Trends Biotechnol 29:127–135. https://doi.org/10.1016/j.tibtech.2010.12.001

Bhopale GM (2020) Antimicrobial peptides: a promising avenue for human healthcare. Curr Pharm Biotechnol 21:90–96

Blunt W, Levin DB, Cicek N (2018) Bioreactor operating strategies for improved polyhydroxyalkanoate (PHA) productivity. Polymers (Basel) 10:1197

Bommarius B, Jenssen H, Elliott M et al (2010) Cost-effective expression and purification of antimicrobial and host defense peptides in *Escherichia coli*. Peptides 31:1957–1965

Boparai JK, Sharma PK (2019) Mini review on antimicrobial peptides, sources, mechanism and recent applications. Protein Pept Lett 27:4–16

Boto A, Pérez de la Lastra J, González C (2018) The road from host-defense peptides to a new generation of antimicrobial drugs. Molecules 23:311

Briand L, Marcion G, Kriznik A et al (2016) A self-inducible heterologous protein expression system in *Escherichia coli*. Sci Rep 6:33037

Browne K, Chakraborty S, Chen R et al (2020) A new era of antibiotics: the clinical potential of antimicrobial peptides. Int J Mol Sci 21:7047

Butt TR, Edavettal SC, Hall JP, Mattern MR (2005) SUMO fusion technology for difficult-to-express proteins. Protein Expr Purif 43:1–9

Cardoso P, Glossop H, Meikle TG et al (2021) Molecular engineering of antimicrobial peptides: microbial targets, peptide motifs and translation opportunities. Biophys Rev 13:35–69

Carta G, Jungbauer A (2010) Protein chromatography. Wiley, Chichester

Chen CH, Lu TK (2020) Development and challenges of antimicrobial peptides for therapeutic applications. Antibiotics 9:24

Cheng KT, Wu CL, Yip BS et al (2018) High level expression and purification of the clinically active antimicrobial peptide P-113 in *Escherichia coli*. Molecules 23:800

Chu C, Zhang W, Li J et al (2018) A single codon optimization enhances recombinant human TNF-α vaccine expression in *Escherichia coli*. Biomed Res Int 2018:1–8

Chung BKS, Lee DY (2012) Computational codon optimization of synthetic gene for protein expression. BMC Syst Biol 6:134

Clarke KG (2013) Downstream processing. In: Bioprocess engineering. Elsevier, Amsterdam, pp 209–234

Corrales-García LL, Serrano-Carreón L, Corzo G (2020) Improving the heterologous expression of human β-defensin 2 (HBD2) using an experimental design. Protein Expr Purif 167:105539

Costa Ramos LF, Rangel J, Andrade G et al (2021) Identification and recombinant expression of an antimicrobial peptide (cecropin B-like) from soybean pest *Anticarsia gemmatalis*. J Venom Anim Toxins Incl Trop Dis 27:1–12

Costa S, Almeida A, Castro A, Domingues L (2014) Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. Front Microbiol 5:1–20

Cui W, Han L, Suo F et al (2018) Exploitation of *Bacillus subtilis* as a robust workhorse for production of heterologous proteins and beyond. World J Microbiol Biotechnol 34:145

D'Anjou MC, Daugulis AJ (2000) Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast. Biotechnol Lett 22:341–346

da Costa JP, Cova M, Ferreira R, Vitorino R (2015) Antimicrobial peptides: an alternative for innovative medicines? Appl Microbiol Biotechnol 99:2023–2040

Daly R, Hearn MTW (2005) Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J Mol Recognit 18:119–138

Dauer K, Pfeiffer-Marek S, Kamm W, Wagner KG (2021) Microwell plate-based dynamic light scattering as a high-throughput characterization tool in biopharmaceutical development. Pharmaceutics 13:172

de Oliveira KBS, Leite ML, Rodrigues GR et al (2020) Strategies for recombinant production of antimicrobial peptides with pharmacological potential. Expert Rev Clin Pharmacol 13:367–390

de Paiva DP, Rocha TB, Rubini MR et al (2018) A study on the use of strain-specific and homologous promoters for heterologous expression in industrial *Saccharomyces cerevisiae* strains. AMB Express 8:82

Deng T, Ge H, He H et al (2017) The heterologous expression strategies of antimicrobial peptides in microbial systems. Protein Expr Purif 140:52–59

Diers IV, Rasmussen E, Larsen PH, Kjaersig I (1991) Yeast fermentation processes for insulin production. Bioprocess Technol 13:166–176

Dijksteel GS, Ulrich MMW, Middelkoop E, Boekema BKHL (2021) Review: lessons learned from clinical trials using antimicrobial peptides (AMPs). Front Microbiol 12:616979

Drayton M, Kizhakkedathu JN, Straus SK (2020) Towards robust delivery of antimicrobial peptides to combat bacterial resistance. Molecules 25:3048

Ehgartner D, Sagmeister P, Langemann T et al (2017) A novel method to recover inclusion body protein from recombinant *E. coli* fed-batch processes based on phage ΦX174-derived lysis protein E. Appl Microbiol Biotechnol 101:5603–5614

Feng X, Xu W, Qu P et al (2015) High-yield recombinant expression of the chicken antimicrobial peptide fowlicidin-2 in *Escherichia coli*. Biotechnol Prog 31:369–374

Fensterseifer ICM, Felício MR, Alves ESF et al (2019) Selective antibacterial activity of the cationic peptide PaDBS1R6 against Gram-negative bacteria. Biochim Biophys Acta Biomembr 1861:1375–1387

Fida HM, Kumada Y, Terashima M et al (2009) Tandem multimer expression of angiotensin I-converting enzyme inhibitory peptide in *Escherichia coli*. Biotechnol J 4:1345–1356

Franco OL (2011) Peptide promiscuity: an evolutionary concept for plant defense. FEBS Lett 585: 995–1000

Fu H, Liang Y, Zhong X et al (2020) Codon optimization with deep learning to enhance protein expression. Sci Rep 10:17617

García-Ortega X, Cámara E, Ferrer P et al (2019) Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (*Komagataella phaffii*) using the methanol-free GAP promoter. Where do we stand? New Biotechnol 53:24–34

Ge Y, MacDonald DL, Holroyd KJ et al (1999) In vitro antibacterial properties of pexiganan, an analog of magainin. Antimicrob Agents Chemother 43:782–788

Gomes AR, Byregowda SM, Veeregowda BM, Balamurugan V (2016) An overview of heterologous expression host systems for the production of recombinant proteins. Adv Anim Vet Sci 4: 346–356

Gronemeyer P, Ditz R, Strube J (2014) Trends in upstream and downstream process development for antibody manufacturing. Bioengineering 1:188–212

Gupta SK, Shukla P (2017) Sophisticated cloning, fermentation, and purification technologies for an enhanced therapeutic protein production: a review. Front Pharmacol 8:1–17
Haney EF, Mansour SC, Hancock REW (2017) Antimicrobial peptides. Springer, New York
Hanson G, Coller J (2018) Codon optimality, bias and usage in translation and mRNA decay. Nat Rev Mol Cell Biol 19:20–30
He Q, Fu A, Li T (2015) Expression and one-step purification of the antimicrobial peptide cathelicidin-BF using the intein system in *Bacillus subtilis*. J Ind Microbiol Biotechnol 42: 647–653
Hebbi V, Kumar D, Rathore AS (2020) Process analytical technology implementation for peptide manufacturing: cleavage reaction of recombinant lethal toxin neutralizing factor concatemer as a case study. Anal Chem 92:5676–5681
Jiang R, Zhang P, Wu X et al (2021) Expression of antimicrobial peptide Cecropin P1 in *Saccharomyces cerevisiae* and its antibacterial and antiviral activity in vitro. Electron J Biotechnol 50:16–22
Jin F, Xu X, Wang L et al (2006) Expression of recombinant hybrid peptide cecropinA(1–8)–magainin2(1–12) in *Pichia pastoris*: purification and characterization. Protein Expr Purif 50: 147–156
Jozala AF, Geraldes DC, Tundisi LL et al (2016) Biopharmaceuticals from microorganisms: from production to purification. Braz J Microbiol 47:51–63
Judzewitsch PR, Corrigan N, Trujillo F, et al (2020) High-throughput process for the discovery of antimicrobial polymers and their upscaled production via flow polymerization. Macromolecules 53:631–639. https://doi.org/10.1021/acs.macromol.9b02207
Juturu V, Wu JC (2018) Heterologous protein expression in *Pichia pastoris*: latest research progress and applications. Chembiochem 19:7–21
Kasemiire A, Avohou HT, De Bleye C et al (2021) Design of experiments and design space approaches in the pharmaceutical bioprocess optimization. Eur J Pharm Biopharm 166:144–154
Kaur J, Kumar A, Kaur J (2018) Strategies for optimization of heterologous protein expression in *E. coli*: roadblocks and reinforcements. Int J Biol Macromol 106:803–822
Kaur N, Dilawari R, Kaur A et al (2020) Recombinant expression, purification and PEGylation of Paneth cell peptide (cryptdin-2) with value added attributes against *Staphylococcus aureus*. Sci Rep 10:12164
Khanal O, Lenhoff AM (2021) Developments and opportunities in continuous biopharmaceutical manufacturing. MAbs 13:1903664
Khow O, Suntrarachun S (2012) Strategies for production of active eukaryotic proteins in bacterial expression system. Asian Pac J Trop Biomed 2:159–162
Kimple ME, Brill AL, Pasker RL (2013) Overview of affinity tags for protein purification. Curr Protoc Protein Sci 73:9.9.1–9.9.23
Klubthawee N, Adisakwattana P, Hanpithakpong W et al (2020) A novel, rationally designed, hybrid antimicrobial peptide, inspired by cathelicidin and aurein, exhibits membrane-active mechanisms against *Pseudomonas aeruginosa*. Sci Rep 10:9117
Koller M (2018) A review on established and emerging fermentation schemes for microbial production of polyhydroxyalkanoate (PHA) biopolyesters. Fermentation 4:1–30
Koo HB, Seo J (2019) Antimicrobial peptides under clinical investigation. Pept Sci 111:1–15
Kornecki M, Strube J (2018) Process analytical technology for advanced process control in biologics manufacturing with the aid of macroscopic kinetic modeling. Bioengineering 5:25
Kumar P, Kizhakkedathu J, Straus S (2018) Antimicrobial peptides: diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. Biomol Ther 8:4
Lanza AM, Curran KA, Rey LG, Alper HS (2014) A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. BMC Syst Biol 8:33
LaVallie ER, DiBlasio EA, Kovacic S et al (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Bio/Technology 11:187–193

Laws M, Shaaban A, Rahman KM (2019) Antibiotic resistance breakers: current approaches and future directions. FEMS Microbiol Rev 43:490–516

Lee J, Kim M, Cho J, Kim S (2002) Enhanced expression of tandem multimers of the antimicrobial peptide buforin II in *Escherichia coli* by the DEAD-box protein and trxB mutant. Appl Microbiol Biotechnol 58:790–796

Lee EY, Lee MW, Fulan BM et al (2017) What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? Interface Focus 7:20160153

Leite ML, Sampaio KB, Costa FF et al (2019) Molecular farming of antimicrobial peptides: available platforms and strategies for improving protein biosynthesis using modified virus vectors. An Acad Bras Cienc 91:1–23

León-Buitimea A, Garza-Cárdenas CR, Garza-Cervantes JA et al (2020) The demand for new antibiotics: antimicrobial peptides, nanoparticles, and combinatorial therapies as future strategies in antibacterial agent design. Front Microbiol 11:1–10

Li Y (2011) Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. Protein Expr Purif 80:260–267

Li T, Bin CX, Chen JC et al (2014a) Open and continuous fermentation: products, conditions and bioprocess economy. Biotechnol J 9:1503–1511

Li Y, Wang J, Yang J et al (2014b) Recombinant expression, purification and characterization of antimicrobial peptide ORBK in *Escherichia coli*. Protein Expr Purif 95:182–187

Li L, Mu L, Wang X et al (2017) A novel expression vector for the secretion of abaecin in *Bacillus subtilis*. Braz J Microbiol 48:809–814

Li H, Ali Z, Liu X et al (2019) Expression of recombinant tachyplesin I in *Pichia pastoris*. Protein Expr Purif 157:50–56

Li Z, Cheng Q, Guo H et al (2020) Expression of hybrid peptide EF-1 in *Pichia pastoris*, its purification, and antimicrobial characterization. Molecules 25:5538

Liang W, Diana J (2020) The dual role of antimicrobial peptides in autoimmunity. Front Immunol 11:1–9

Lima SMF, Freire MS, Gomes ALO et al (2017) Antimicrobial and immunomodulatory activity of host defense peptides, clavanins and LL-37, in vitro: an endodontic perspective. Peptides 95:16–24

Lin Q, Xie K, Chen D et al (2020) Expression and functional characterization of a novel antimicrobial peptide: human beta-defensin 118. Biomed Res Int 2020:1–10

Liscano Y, Oñate-Garzón J, Delgado JP (2020) Peptides with dual antimicrobial–anticancer activity: strategies to overcome peptide limitations and rational design of anticancer peptides. Molecules 25:4245

Liu SN, Han Y, Zhou ZJ (2011) Fusion expression of pedA gene to obtain biologically active pediocin PA-1 in *Escherichia coli*. J Zhejiang Univ Sci B 12:65–71

Liu H, Yang N, Mao R et al (2020) A new high-yielding antimicrobial peptide NZX and its antibacterial activity against *Staphylococcus hyicus* in vitro/vivo. Appl Microbiol Biotechnol 104:1555–1568

Lombardi L, Shi Y, Falanga A et al (2019) Enhancing the potency of antimicrobial peptides through molecular engineering and self-assembly. Biomacromolecules 20:1362–1374

Looser V, Bruhlmann B, Bumbak F et al (2015) Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. Biotechnol Adv 33:1177–1193

Love KR, Dalvie NC, Love JC (2018) The yeast stands alone: the future of protein biologic production. Curr Opin Biotechnol 53:50–58

Mahlapuu M, Björn C, Ekblom J (2020) Antimicrobial peptides as therapeutic agents: opportunities and challenges. Crit Rev Biotechnol 40:978–992

Mauro VP, Chappell SA (2014) A critical analysis of codon optimization in human therapeutics. Trends Mol Med 20:604–613

Mayr LM, Bojanic D (2009) Novel trends in high-throughput screening. Curr Opin Pharmacol 9: 580–588

Mears L, Stocks SM, Sin G, Gernaey KV (2017) A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. J Biotechnol 245:34–46

Meng DM, Lv YJ, Zhao JF et al (2018) Efficient production of a recombinant Venerupis philippinarum defensin (VpDef) in *Pichia pastoris* and characterization of its antibacterial activity and stability. Protein Expr Purif 147:78–84

Meng D-M, Li W-J, Shi L-Y et al (2019) Expression, purification and characterization of a recombinant antimicrobial peptide Hispidalin in *Pichia pastoris*. Protein Expr Purif 160:19–27

Mercer DK, Torres MDT, Duay SS et al (2020) Antimicrobial susceptibility testing of antimicrobial peptides to better predict efficacy. Front Cell Infect Microbiol 10:1–34

Mirski T, Niemcewicz M, Bartoszcze M et al (2017) Utilisation of peptides against microbial infections—a review. Ann Agric Environ Med 25:205–210

Mishra B, Reiling S, Zarena D, Wang G (2017) Host defense antimicrobial peptides as antibiotics: design and application strategies. Curr Opin Chem Biol 38:87–96

Møller TSB, Hay J, Saxton MJ et al (2017) Human β-defensin-2 production from *S. cerevisiae* using the repressible MET17 promoter. Microb Cell Factories 16:11

Moravej H, Moravej Z, Yazdanparast M et al (2018) Antimicrobial peptides: features, action, and their resistance mechanisms in bacteria. Microb Drug Resist 24:747–767

Moretta A, Salvia R, Scieuzo C et al (2020) A bioinformatic study of antimicrobial peptides identified in the Black Soldier Fly (BSF) *Hermetia illucens* (Diptera: Stratiomyidae). Sci Rep 10:16875

Moretta A, Scieuzo C, Petrone AM et al (2021) Antimicrobial peptides: a new hope in biomedical and pharmaceutical fields. Front Cell Infect Microbiol 11:1–26

Mulder KC, de Lima LA, Aguiar PS et al (2015) Production of a modified peptide clavanin in *Pichia pastoris*: cloning, expression, purification and in vitro activities. AMB Express 5:46

Naumov GI, Naumova ES, Boundy-Mills KL (2018) Description of *Komagataella mondaviorum* sp. nov., a new sibling species of *Komagataella* (*Pichia*) *pastoris*. Antonie van Leeuwenhoek 111:1197–1207

Nguyen LT, Haney EF, Vogel HJ (2011) The expanding scope of antimicrobial peptide structures and their modes of action. Trends Biotechnol 29:464–472

Öztürk S, Ergün BG, Çalık P (2017) Double promoter expression systems for recombinant protein production by industrial microorganisms. Appl Microbiol Biotechnol 101:7459–7475

Pal G, Srivastava S (2015) Scaling up the production of recombinant antimicrobial Plantaricin E from a heterologous host, *Escherichia coli*. Probiotics Antimicrob Proteins 7:216–221

Pan X, Yang Y, Liu X et al (2016) Secretory expression of a heterologous protein, Aiio-AIO6BS, in *Bacillus subtilis* via a non-classical secretion pathway. Biochem Biophys Res Commun 478: 881–886

Parachin NS, Mulder KC, Viana AAB et al (2012) Expression systems for heterologous production of antimicrobial peptides. Peptides 38:446–456

Peebo K, Neubauer P (2018) Application of continuous culture methods to recombinant protein production in microorganisms. Microorganisms 6:56

Philip P, Meier K, Kern D et al (2017) Systematic evaluation of characteristics of the membrane-based fed-batch shake flask. Microb Cell Factories 16:122

Pieracci JP, Armando JW, Westoby M, Thommes J (2018) Industry review of cell separation and product harvesting methods. In: Biopharmaceutical processing. Elsevier, Amsterdam, pp 165–206

Potvin G, Ahmad A, Zhang Z (2012) Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: a review. Biochem Eng J 64:91–105

Presnyak V, Alhusaini N, Chen YH et al (2015) Codon optimality is a major determinant of mRNA stability. Cell 160:1111–1124

Quinn GA, Maloy AP, McClean S et al (2012) Lipopeptide biosurfactants from *Paenibacillus polymyxa* inhibit single and mixed species biofilms. Biofouling 28:1151–1166

Raja Z, André S, Abbassi F et al (2017) Insight into the mechanism of action of temporin-SHa, a new broad-spectrum antiparasitic and antibacterial agent. PLoS One 12:e0174024

Ramirez LS, Pande J, Shekhtman A (2019) Helical structure of recombinant melittin. J Phys Chem B 123:356–368

Rathore AS, Agarwal H, Sharma AK et al (2015) Continuous processing for production of biopharmaceuticals. Prep Biochem Biotechnol 45:836–849

Rios AC, Moutinho CG, Pinto FC et al (2016) Alternatives to overcoming bacterial resistances: state-of-the-art. Microbiol Res 191:51–80

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:1–17

Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: a 5-year update. Protein Sci 28:1412–1422

Sang M, Wei H, Zhang J et al (2017) Expression and characterization of the antimicrobial peptide ABP-dHC-cecropin A in the methylotrophic yeast *Pichia pastoris*. Protein Expr Purif 140:44–51

São Pedro MN, Silva TC, Patil R, Ottens M (2021) White paper on high-throughput process development for integrated continuous biomanufacturing. Biotechnol Bioeng 118:3275–3286

Schuller A, Cserjan-Puschmann M, Tauer C et al (2020) *Escherichia coli* σ70 promoters allow expression rate control at the cellular level in genome-integrated expression systems. Microb Cell Factories 19:58

Scott B, Wilcock A (2006) Process analytical technology in the pharmaceutical industry: a toolkit for continuous improvement. PDA J Pharm Sci Technol 60:17–53

Sen A, Kargar K, Akgün E, Plnar MC (2020) Codon optimization: a mathematical programing approach. Bioinformatics 36:4012–4020

Seyfi R, Kahaki FA, Ebrahimi T et al (2020) Antimicrobial peptides (AMPs): roles, functions and mechanism of action. Int J Pept Res Ther 26:1451–1463

Sezonov G, Joseleau-Petit D, D'Ari R (2007) *Escherichia coli* physiology in Luria-Bertani Broth. J Bacteriol 189:8746–8749

Shukla AA, Rameez S, Wolfe LS, Oien N (2017) High-throughput process development for biopharmaceuticals. In: Advances in biochemical engineering/biotechnology. Springer Nature, Berlin, pp 401–441

Silva ON, Mulder KCL, Barbosa AEAD et al (2011) Exploring the pharmacological potential of promiscuous host-defense peptides: from natural screenings to biotechnological applications. Front Microbiol 2:1–14

Silva TC, Eppink M, Ottens M (2021) Automation and miniaturization: enabling tools for fast, high-throughput process development in integrated continuous biomanufacturing. J Chem Technol Biotechnol jctb6792

Singh N, Pizzelli K, Romero JK et al (2013) Clarification of recombinant proteins from high cell density mammalian cell culture systems using new improved depth filters. Biotechnol Bioeng 110:1964–1972

Sinha R, Shukla P (2019) Antimicrobial peptides: recent insights on biotechnological interventions and future perspectives. Protein Pept Lett 26:79–87

Sousa DA, Mulder KCL, Nobre KS et al (2016) Production of a polar fish antimicrobial peptide in *Escherichia coli* using an ELP-intein tag. J Biotechnol 234:83–89

Taneda A, Asai K (2020) COSMO: a dynamic programming algorithm for multicriteria codon optimization. Comput Struct Biotechnol J 18:1811–1818

Tang X-S, Shao H, Li T-J et al (2012) Dietary supplementation with bovine lactoferrampin–lactoferricin produced by *Pichia pastoris* fed-batch fermentation affects intestinal microflora in weaned piglets. Appl Biochem Biotechnol 168:887–898

Tavares LS, Rettore JV, Freitas RM et al (2012) Antimicrobial activity of recombinant Pg-AMP1, a glycine-rich peptide from guava seeds. Peptides 37:294–300

Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72:211–222

Thapa RK, Diep DB, Tønnesen HH (2020) Topical antimicrobial peptide formulations for wound healing: current developments and future prospects. Acta Biomater 103:52–67

Tian ZG, Teng D, Yang YL et al (2007) Multimerization and fusion expression of bovine lactoferricin derivative LfcinB15-W4,10 in *Escherichia coli*. Appl Microbiol Biotechnol 75: 117–124

Tian J, Yan Y, Yue Q et al (2017) Predicting synonymous codon usage and optimizing the heterologous gene for expression in *E. coli*. Sci Rep 7:9926

Tian L, Zhang D, Su P et al (2019) Design, recombinant expression, and antibacterial activity of a novel hybrid magainin–thanatin antimicrobial peptide. Prep Biochem Biotechnol 49:427–434

Tornesello AL, Borrelli A, Buonaguro L et al (2020) Antimicrobial peptides as anticancer agents: functional properties and biological activities. Molecules 25:2850

Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:1–35

Viel JH, Jaarsma AH, Kuipers OP (2021) Heterologous expression of mersacidin in *Escherichia coli* elucidates the mode of leader processing. ACS Synth Biol 10:600–608

Vilas Boas LCP, de Lima LMP, Migliolo L et al (2017) Linear antimicrobial peptides with activity against herpes simplex virus 1 and Aichi virus. Pept Sci 108:e22871

Vogt CM, Schraner EM, Aguilar C, Eichwald C (2016) Heterologous expression of antigenic peptides in *Bacillus subtilis* biofilms. Microb Cell Factories 15:137

Wade HM, Darling LEO, Elmore DE (2019) Hybrids made from antimicrobial peptides with different mechanisms of action show enhanced membrane permeabilization. Biochim Biophys Acta Biomembr 1861:182980

Wang A, Wang S, Shen M et al (2009) High level expression and purification of bioactive human α-defensin 5 mature peptide in *Pichia pastoris*. Appl Microbiol Biotechnol 84:877–884

Wang FJ, Song HL, Wang XM et al (2012) Tandem multimer expression and preparation of hypoglycemic peptide MC6 from *Momordica charantia* in *Escherichia coli*. Appl Biochem Biotechnol 166:612–619

Wang J, Dou X, Song J et al (2019) Antimicrobial peptides: promising alternatives in the post feeding antibiotic era. Med Res Rev 39:831–859

Weinacker D, Rabert C, Zepeda AB et al (2013) Applications of recombinant *Pichia pastoris* in the healthcare industry. Braz J Microbiol 44:1043–1048

Whitesides GM (2006) The origins and the future of microfluidics. Nature 442:368–373

Wibowo D, Zhao C-X (2019) Recent achievements and perspectives for large-scale recombinant production of antimicrobial peptides. Appl Microbiol Biotechnol 103:659–671

Wu R, Wang Q, Zheng Z et al (2014) Design, characterization and expression of a novel hybrid peptides melittin (1-13)-LL37 (17-30). Mol Biol Rep 41:4163–4169

Xiao S, Gao Y, Wang X et al (2017) Peroxisome-targeted and tandem repeat multimer expressions of human antimicrobial peptide LL37 in *Pichia pastoris*. Prep Biochem Biotechnol 47:229–235

Xu X, Jin F, Yu X et al (2007) High-level expression of the recombinant hybrid peptide cecropinA (1-8)-magainin2(1-12) with an ubiquitin fusion partner in *Escherichia coli*. Protein Expr Purif 55:175–182

Yang N, Teng D, Mao R et al (2019) A recombinant fungal defensin-like peptide-P2 combats multidrug-resistant *Staphylococcus aureus* and biofilms. Appl Microbiol Biotechnol 103:5193–5213

Yang Y, Wu D, Wang C et al (2020) Hybridization with insect cecropin a (1–8) improve the stability and selectivity of naturally occurring peptides. Int J Mol Sci 21:1470

Ÿztürk S, Ÿalık P, Ÿzdamar TH (2016) Fed-batch biomolecule production by *Bacillus subtilis*: a state of the art review. Trends Biotechnol 34:329–345

Zandsalimi F, Talaei S, Noormohammad Ahari M et al (2020) Antimicrobial peptides: a promising strategy for lung cancer drug discovery? Expert Opin Drug Discov 15:1343–1354

Zhan N, Zhang L, Yang H et al (2021) Design and heterologous expression of a novel dimeric LL37 variant in *Pichia pastoris*. Microb Cell Factories 20:143

Zhang J, Movahedi A, Wei Z et al (2016) High-level SUMO-mediated fusion expression of ABP-dHC-cecropin A from multiple joined genes in *Escherichia coli*. Anal Biochem 509:15–23

Zhang K, Su L, Duan X et al (2017) High-level extracellular protein production in *Bacillus subtilis* using an optimized dual-promoter expression system. Microb Cell Factories 16:32

Zhang X, Jiang A, Qi B et al (2018) Secretion expression of human neutrophil peptide 1 (HNP1) in *Pichia pastoris* and its functional analysis against antibiotic-resistant *Helicobacter pylori*. Appl Microbiol Biotechnol 102:4817–4827

Zhang L, Wei D, Zhan N et al (2020) Heterologous expression of the novel α-helical hybrid peptide PR-FO in *Bacillus subtilis*. Bioprocess Biosyst Eng 43:1619–1627

Zhao W, Wang J, Deng R, Wang X (2008) Scale-up fermentation of recombinant *Candida rugosa* lipase expressed in *Pichia pastoris* using the GAP promoter. J Ind Microbiol Biotechnol 35: 189–195

Zhao H, Tang J, Cao L et al (2015) Characterization of bioactive recombinant antimicrobial peptide parasin I fused with human lysozyme expressed in the yeast *Pichia pastoris* system. Enzym Microb Technol 77:61–67

Zhou Y, Cao W, Wang J et al (2005) Comparison of expression of monomeric and multimeric adenoregulin genes in *Escherichia coli* and *Pichia pastoris*. Protein Pept Lett 12:349–355

Zhoua Z, Danga Y, Zhou M et al (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci U S A 113: E6117–E6125

Zobel S, Kumpfmüller J, Süssmuth RD, Schweder T (2015) *Bacillus subtilis* as heterologous host for the secretory production of the non-ribosomal cyclodepsipeptide enniatin. Appl Microbiol Biotechnol 99:681–691

Zydney AL (2016) Continuous downstream processing for high value biological products: a review. Biotechnol Bioeng 113:465–475

Bioproduction of Cyclic Disulfide-Rich Peptides for Drug Modalities

Kuok Yap, Conan K. Wang, David J. Craik, and Linda H. L. Lua

Contents

Abstract Cyclic disulfide-rich peptides have extensive applications in drug development. These peptides are widely produced using solid-phase peptide synthesis which generates substantial amounts of toxic and hazardous waste. Recombinant bioproduction platforms for cyclic disulfide-rich peptides offer a more environmentally sustainable alternative. This chapter highlights a recently established microbial-based bioproduction platform that utilizes both *Pichia pastoris* and *Escherichia coli* for the production of cyclic disulfide-rich peptides.

K. Yap
Protein Expression Facility, The University of Queensland, Brisbane, Australia

Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, Australia

Institute for Molecular Bioscience, Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Queensland, Brisbane, Australia

C. K. Wang · D. J. Craik
Institute for Molecular Bioscience, Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, The University of Queensland, Brisbane, Australia

L. H. L. Lua (✉)
Protein Expression Facility, The University of Queensland, Brisbane, Australia
e-mail: l.lua@uq.edu.au

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_6

1 Introduction

Cyclic disulfide-rich peptides (CDRPs) have exceptional thermal stability and are resistant to proteolytic degradation because of their unique structures that comprise a cyclic backbone cross-linked by the presence of one or more disulfide bonds (Colgrave and Craik 2004; Colgrave et al. 2010; Cheneval et al. 2014). They occupy a niche in the pharmaceutical market by combining the advantages of small-molecule and protein-based drugs. Being larger than small-molecule drugs, CDRPs display high target selectivity, which minimizes adverse side effects (Craik et al. 2013) while being smaller than protein-based biologics, CDRPs have a better chance of achieving oral bioavailability than biologics (Wong et al. 2012). Figure 1 highlights examples of CDRPs found in plants and animals or others that were synthetically engineered cyclic (Craik et al. 1999). One of the smallest CDRPs is the sunflower trypsin inhibitor (SFTI-1) with one disulfide-bond (Fig. 1a). Larger CDRPs include a naturally linear peptide with two disulfide bonds from the cone snail *Conus victoriae* that was engineered to be cyclic called cVc1.1 (Fig. 1b), an antimicrobial peptide with three disulfide bonds called Rhesus theta defensin-1 (RTD-1) (Fig. 1c) and a cyclotide with three disulfide bonds found in the Vietnamese Gấc plant called *Momordica cochinchinensis* trypsin-inhibitor II (MCoTI-II) (Fig. 1d).

With the prospect of CDRPs being approved for use in the clinic in the coming years, addressing the large-scale production challenges is becoming important. Currently, most CDRPs are produced synthetically using solid-phase peptide synthesis (SPPS). This process generates substantial waste and is not ideal for industrial-scale production (Cheneval et al. 2014; Merrifield 1963). Recombinant-based bioproduction methods are environmentally sustainable alternatives and are actively

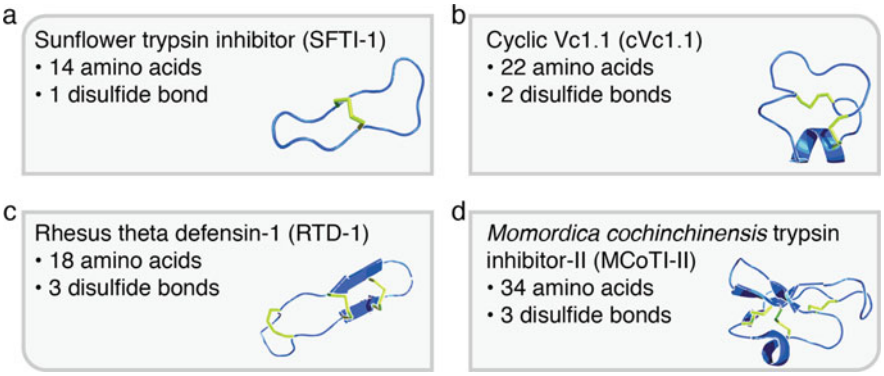


Fig. 1 Examples of CDRPs: (**a**) SFTI-1, (**b**) cVc1.1 (**c**) RTD-1, and (**d**) MCoTI-II are shown in order of increasing number of disulfide bonds. Each panel shows the amino acid length and disulfide content of the respective cyclic peptide. Their 3D structures are illustrated in blue with disulfide bonds colored in yellow. (PDB ID for SFTI-1 is 1JBL; cVc1.1 is 4TTL; RTD-1 is 1HVZ; and MCoTI-II is1HA9)

being considered for CDRP biomanufacturing. Of these, microbial systems are favored as the expression hosts due to their fast growth rates, relatively inexpensive media, and simple genetic manipulation (Tripathi and Shrivastava 2019). These advantages were evident when insulin, the first microbially produced peptide drug, was approved for market in 1982 (Johnson 1983). In recent years, plant expression systems have emerged as a viable option for producing recombinant therapeutics because they can properly fold complex proteins (Dirisala et al. 2017; Yao et al. 2015). After Elelyso, the first plant-derived pharmaceutical drug to be approved by the US Food and Drug Administration (FDA) in 2012 for the treatment of Gaucher's disease (Fox 2012) substantial research efforts have demonstrated using plants as bio-factories for the bioproduction of antibodies (Arntzen 2015; Bally et al. 2018), vaccines (Streatfield and Howard 2003), complex carbohydrates (Roberts et al. 2018), as well as, CDRPs (Jackson et al. 2019; Poon et al. 2018). With a range of bioproduction platforms available, the challenge is determining which one is most suitable for CDRP drug manufacturing. This chapter explores recent progress examining microbial and plant-based bioproduction systems utilized to generate CDRPs and provides an outlook for the future production of CDRPs based on the knowledge acquired.

2 Potential High-Value Applications of CDRPs

CDRPs are valued for pharmaceutical applications as molecular scaffolds because of their exceptional stability due to their ability to withstand heat (Sable et al. 2016), harsh acid conditions (Wang et al. 2014), and proteolysis from a wide range of proteases (Wang et al. 2014; Chan et al. 2011). A desired biological function can be introduced by inserting bioactive amino acid sequences onto these extremely stable CDRP scaffolds, a concept known as molecular grafting. Grafted peptides gain non-native biological activities while retaining the inherent stability attributes of native CDRP scaffolds. Numerous studies have engineered CDRPs as scaffolds for various medical applications, with examples shown in Table 1. The most promising candidate based on a point mutation of a prototypical cyclotide, kalata B1 (kB1), is scheduled for clinical trials for the treatment of multiple sclerosis (Gründemann et al. 2019). Based on the evidence described briefly in this section, it is clear that CDRPs have attractive pharmaceutical properties with immense potential to be developed into high-value drug modalities.

Table 1 Medical applications of selected CDRPs

Scaffold	Activity	High-value application	References
SFTI-1	Anti-angiogenesis	Inhibiting tumor progression	Chan et al. (2015)
	Chymase inhibitor	Cancer	Li et al. (2019)
	Angiogenic	Cardiovascular and wound healing	Chan et al. (2011)
	Tau aggregation inhibitor	Alzheimer's disease	Wang et al. (2016)
cVc1.1	Cyclization	Neuropathic pain	Clark et al. (2010)
kB1	Immunomodulation	Multiple sclerosis	Wang et al. (2016)
	Melanocortin receptor 4 agonist	Obesity	Eliasen et al. (2012)
	Lymphocyte proliferation inhibitor	Multiple sclerosis	Thell et al. (2016)
MCoTI	Factor XIIa inhibitor	Cardiovascular disease	Swedberg et al. (2016)
	Cytokine receptor CXCR4 antagonist	Anti-HIV	Aboye et al. (2012)
	Matriptase inhibitor	Anti-tumor	Quimbar et al. (2013)
	Tyrosine kinase inhibitor	Chronic myeloid leukemia	Huang et al. (2015)

3 Production of CDRPs

3.1 Synthetic Peptide Synthesis

CDRPs are commonly produced using SPPS (Merrifield 1963). The process requires multiple processing steps after assembly of a side-chain protected peptide chain on solid resin support. A typical workflow (Cheneval et al. 2014) is illustrated in Fig. 2. After assembly, the peptide chain is cleaved from the resin by mild acid cleavage to expose the C-terminus (Fig. 2a) for intramolecular cyclization (Fig. 2b). The side-chain protected cyclic peptide (Fig. 2c) is then treated with strong acid to remove the remaining protecting groups, liberating the sulfhydryl groups of Cys residues for disulfide bond formation under oxidative folding conditions (Fig. 2d). Finally, the cyclic oxidized peptide (Fig. 2e) is purified using reverse-phase high-pressure liquid chromatography.

According to Green Chemistry metrics, SPPS-based production of CDRPs is not an environmentally friendly process (Jad et al. 2019). For every kilogram of peptide produced on a commercial scale, it generates hazardous waste in the multi-ton range (Ritter 2017). The types of hazardous waste produced during the assembly and handling of CDRPs include dichloromethane (DCM), *N,N*-dimethylformamide (DMF), trifluoroacetic acid (TFA), triisopropylsilane (TIPS), and piperidine (Isidro-Llobet et al. 2019) (Fig. 2). The production of disulfide-rich peptides requires the disposal of large volumes of folding waste because the folding reactions are carried out at low concentrations to prevent oligomerization (Cheneval et al. 2014). With increasing environmental awareness amongst government, industry, and the

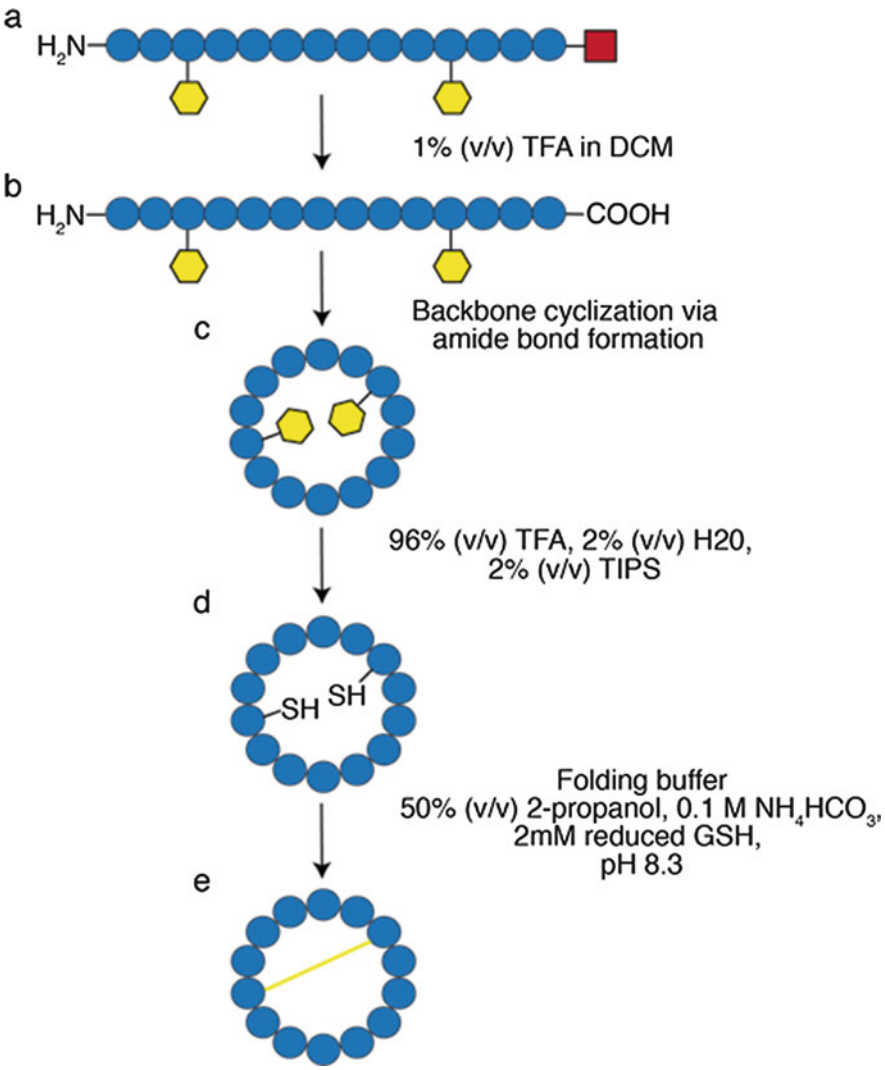


Fig. 2 Schematic workflow of a chemical production process for CDRPs, which includes peptide-chain synthesis, cyclization, and folding (Cheneval et al. 2014). (**a**) Linear side-chain (yellow hexagon) and C-terminal (red square) protected peptide chain (blue circles). The entire sequence is side-chained protected, but only Cys is shown for simplicity. (**b**) Termini-free linear side-chain protected peptide precursor after mild acid hydrolysis of C-terminal protecting group. (**c**) Cyclic side-chain protected peptide due to backbone cyclization via amide bond formation. (**d**) Cyclic reduced peptide after harsh acid liberation of side-chain protecting groups. (**e**) Oxidized CDRP after the formation of the disulfide bonds. A folding buffer containing an organic solvent that is used for the oxidation of kB1 is shown as an example

public, there is a need for more sustainable and environmentally friendly alternatives.

3.2 Recombinant Bioproduction of CDRPs

Recombinant technologies offer a pathway toward the sustainable production of CDRPs. The bioproduction of CDRPs has been explored using microbial and plant-based expression systems. Backbone cyclization is easily achieved in chemical synthesis (Cheneval et al. 2014) but the main challenge in bioproduction methods for CDRPs comes from the need for endogenous machinery for post-translational cyclization. We highlight two strategies that have been used to overcome this cyclization bottleneck. The first is an in-vivo intein-mediated cyclization approach and the second an in-vitro or in-vivo asparaginyl endopeptidases (AEPs)-mediated cyclization approach.

3.2.1 Cyclization Strategies in Bioproduction of CDRPs

The intein-mediated backbone cyclization approach, also known as expressed protein ligation (EPL), utilizes an intramolecular ligation strategy based on native chemical ligation (NCL). NCL is a chemoselective ligation reaction carried out under aqueous conditions to ligate two unprotected peptides, one containing an N-terminal cysteine and another containing a C-terminal α-thioester group (Dawson et al. 1994). The chemical ligation can be carried out intramolecularly by incorporating both reactive groups on the same peptide chain, which results in backbone cyclization as illustrated in Fig. 3a. The first step requires the removal of the N-terminal methionine to generate the α-Cys on the N-terminus of the peptide. This reaction can be performed in-vivo by endogenous host proteases such as methionyl aminopeptidases (Camarero et al. 2001). The α-thioesters are generated by fusing modified inteins onto the C-terminus of linear peptides or proteins (Camarero and Muir 1999). Hence, the presence of an N-terminus α-Cys and C-terminus α-thioesters on a linear peptide facilitates backbone cyclization.

The second approach relies on a class of protein ligases called AEPs. These enzymes have provided biochemists with exciting new tools to explore a range of applications based on amide-bond formation, including backbone cyclization (Nguyen et al. 2014; Rehm et al. 2019; Rehm et al. 2020). Most AEPs in higher organisms function as hydrolases, but a handful of AEPs from plants catalyzes amide-bond formation, a function crucial for the cyclization of plant-derived cyclic precursors (CPs) (Du et al. 2020; James et al. 2018). For intramolecular cyclization, a simple tripeptide sequence, Asx-Xaa-Yaa (P1-P1'-P2'), at the C-terminal is required, where Asx is an Asn or Asp at P1 followed by any small amino acid at P1' and an aliphatic or hydrophobic amino acid at P2'. Processing occurs between P1 and P1' as illustrated in Fig. 4. A transpeptidation reaction occurs after

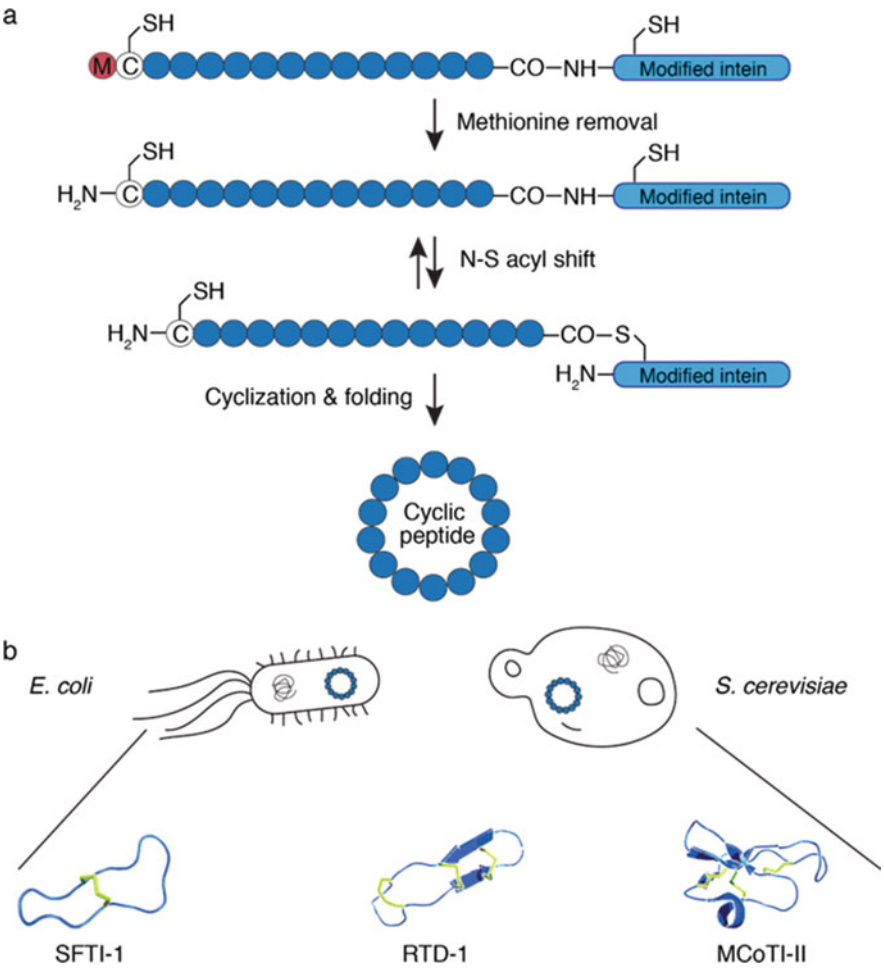


Fig. 3 Overview of the bioproduction of cyclic disulfide-rich peptides in *E. coli* and *S. cerevisiae* using the EPL strategy. (**a**) The in-vivo multiple-step intein-based backbone cyclization of a linear peptide. (**b**) *E. coli* or *S. cerevisiae* have both been used to recombinantly produce SFTI-1, RTD-1, and MCoTI-based peptides using the EPL approach

processing, whereby an incoming nucleophile sequence of P1"-P2" is accepted, forming the P1-P1" amide-bond. All proteinogenic amino acids except for Pro are tolerated at the P1" position, and Cys, Ile, Leu and Val are preferred at the P2" position (Nguyen et al. 2016). The combination of residues at the positions described paired with respective AEPs used for recombinant production of cyclic disulfide-rich peptides is illustrated in Fig. 4 (Poon et al. 2018; Yap et al. 2020). Ligation or cyclization with AEPs can be carried out with minimal 'scaring' of the native sequence, leaving only the Asx residue, compared to other protein ligases, which require longer, more rigid recognition sequences that are not amenable to change.

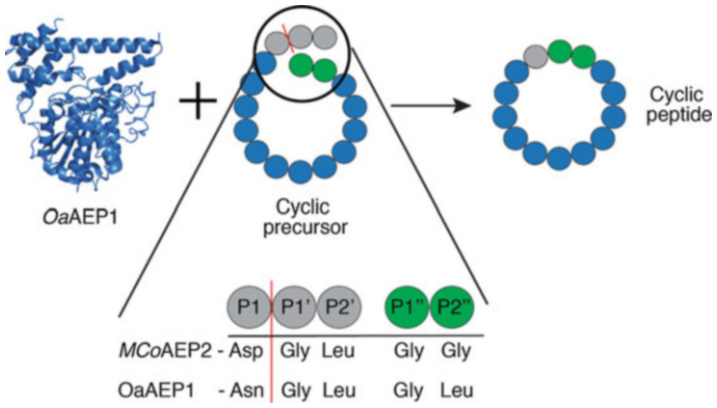


Fig. 4 AEP-mediated cyclization of CPs and key residues for intramolecular cyclization. A couple of examples showing the residues used for recombinant production. The red line indicates the AEP processing position between P1 and P1'

3.2.2 Recombinant Systems for Bioproduction of CDRPs

3.2.2.1 Intein-Mediated Microbial Bioproduction

Early efforts in developing microbial systems for producing CDRPs relied on the in-cell EPL backbone cyclization approach in *E. coli,* as illustrated in Fig. 3a (Gould et al. 2012; Li et al. 2016). MCoTI-II was the first correctly folded cyclotide produced recombinantly and structurally validated by 1D and 2D ^{1}H Nuclear Magnetic Resonance (NMR) spectroscopy (Camarero et al. 2007). The ability to produce correctly folded cyclotide, MCoTI-II, in its native fold was a major accomplishment as a previous attempt at producing another cyclotide, kB1, resulted in misfolded peptides (Kimura et al. 2006). In subsequent studies, SFTI-1 (Li et al. 2016) and RTD-1 (Gould et al. 2012) were produced in *E. coli*, and MCoTI-I, a paralog of MCoTI-II, was produced in *Saccharomyces cerevisiae* using the same EPL backbone cyclization approach (Jagadish et al. 2015). These studies were trailblazers for the recombinant production of CDRPs by producing a range of structurally distinct CDRPs as shown in Fig. 3b. However, the primary aim of these studies was to develop cell-based cyclic peptide libraries for rapid screening of biological activity. There are several limitations associated with these studies from a commercial bioproduction perspective, including the need for a reduced cysteine at the N-terminus and the reported yields. The EPL approach is limited to peptides that have reduced cysteine on the N-terminus. The reported yields of 10–180 μg L^{-1} culture is not commercially viable because the yields widely achieved for industrial-scale production are in the g L^{-1} scale (Spohner et al. 2015).

3.2.2.2 AEP-Mediated Plant Bioproduction

Plant-based bioproduction is considered for the production of CDRPs because a large number of CDRPs such as kB1 and MCoTI-II originate from plants (Gran 1970). The first study to demonstrate plant-based CDRP production used a model experimental plant, *Arabidopsis thaliana* to stably produce SFTI-1 (Mylne et al. 2011). *A. thaliana* is not an ideal candidate for commercial bioproduction due to its small stature. *Nicotiana benthamiana* is more suited for industrial-scale production because it has been used to produce a range of potential therapeutic candidates (Schillberg and Finnern 2021). A grafted-SFTI-1 and kB1 cyclic peptides were both produced in the leaves of *N. benthamiana* when co-expressed with an AEP (Jackson et al. 2019; Poon et al. 2018). These plant-derived CDRPs were validated to have the same monoisotopic mass and were structurally equivalent to their native counterparts using high-resolution mass spectrometry (HRMS) and NMR-based structural characterization, respectively. Although promising, there are several limitations to using plants as bio-factories for the large-scale manufacturing of CDRPs. The highest yield reported for a plant-produced cyclic peptide is 199 μg g^{-1} dry weight (Poon et al. 2018) which is significantly lower compared to other CDRP production platforms (Yap et al. 2020; Yap et al. 2021). The downstream purification issues related to plant-based production include the tedious harvesting of transformed tissues while sieving away non-transformed tissues, multiple organic solvent extractions, and the complex purification of CDRPs from crude plant lysate inundated with endogenous plant proteins (Jackson et al. 2019; Poon et al. 2018).

3.2.2.3 AEP-Mediated *P. pastoris*-Based Bioproduction

P. pastoris has all the desirable attributes as an expression host and has been awarded GRAS (generally recognized as safe) status by the FDA. Additionally, *P. pastoris* is a proven expression host in the space of pharmaceutical manufacturing with numerous high-value drugs manufactured by this platform, including Kalbitor®, a kallikrein inhibitor (Walsh 2010; Ciofalo et al. 2006) and Jetrea®, for the treatment of vitreomacular adhesion (Mullard 2013). With advances in synthetic biology and engineering, *P. pastoris*-derived drugs with human-like glycosylation patterns to lower host immunogenicity are now achievable (Choi et al. 2003; Laukens et al. 2015). Another strong attribute of *P. pastoris* is its fast growth rate in relatively inexpensive media compared to other eukaryotic expression systems. The ability of *P. pastoris* to secreted recombinant proteins into the growth media at yields of up to 22 g L^{-1} is a major advantage from a downstream processing perspective because it mitigates the need for complex cell lysis and purification from crude lysate (Katrolia et al. 2011).

The limitations of both the EPL-mediated CDRP production in microbial systems and CDRP production in plants are the complex downstream purification requirements and the low CDRP yield (Jackson et al. 2019; Poon et al. 2018; Gould et al. 2012). A recent study selected *P. pastoris* to produce CPs because of all the

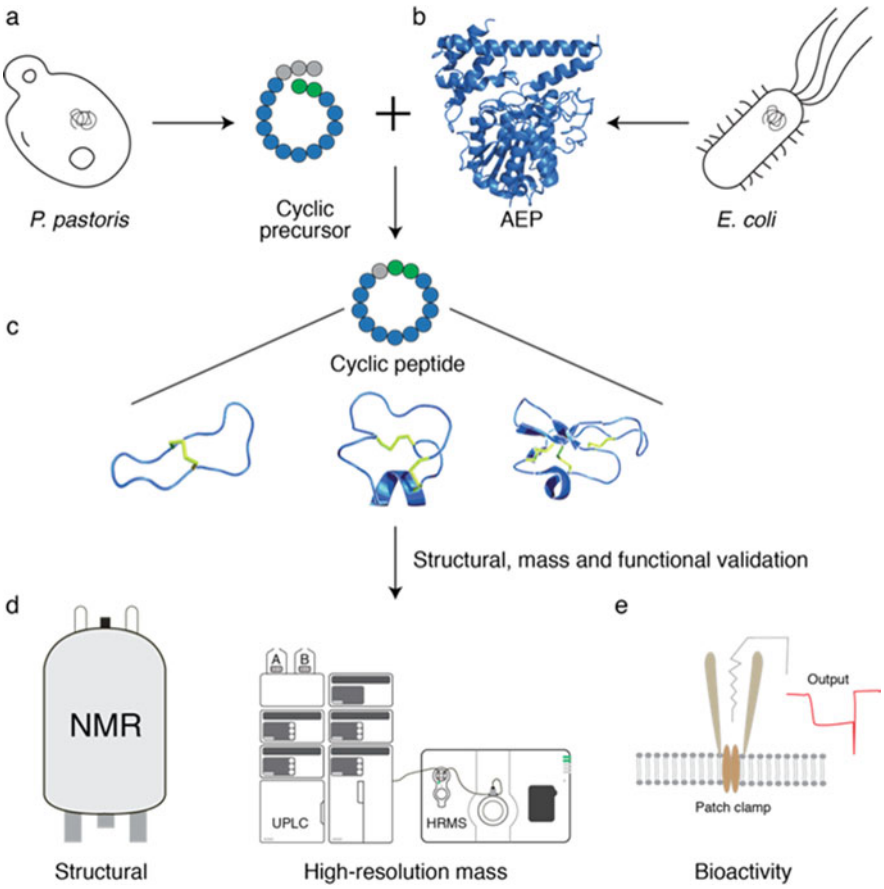


Fig. 5 Diagrammatic representation of the *P. pastoris*-based bioproduction platform for CDRPs. (**a**) *P. pastoris*-secreted CPs and (**b**) recombinant AEPs produced in *E. coli*. (**c**) The bioproduction of three structurally distinct CDRPs as a result of in-vitro cyclization of *P. pastoris*-derived CPs by *E. coli*-derived AEPs. (**d**) Structural, mass, and (**e**) bioactivity validation of recombinant CDRPs

attributes described (Yap et al. 2020). The study demonstrated the potential of a *P. pastoris*-based CDRP production platform to overcome both yield and downstream processing limitations. CPs were secreted into the growth media by attaching an α-mating factor signal peptide onto the N-terminus of CPs (Fig. 5a), bypassing the need for complex downstream processing. The secreted CPs achieved yield improvements of 2–3 orders of magnitude compared to previous recombinant attempts (Yap et al. 2020; Yap et al. 2021). Three CPs have been successfully secreted into growth media: a kallikrein-related protease inhibitor based on the SFTI-1 scaffold, cVc1.1, and MCoTI-II (Yap et al. 2020). After single affinity purification, these CPs were enriched and concentrated, decreasing the working volume for downstream in-vitro cyclization (Yap et al. 2020).

Prior to in-vitro cyclization, AEP zymogens were recombinantly produced in *E. coli* followed by acid-dependent activation into their active form (Fig. 5b) (Du et al. 2020; Yang et al. 2017). These secreted and properly folded CPs were then cyclized in-vitro by active AEPs into mature CDRPs (Fig. 5c). With little or no alteration to the workflow, this versatile production platform produced CDRPs from three structurally distinct families on a small scale. Laboratory scale-up production using a bioreactor was performed on cVc1.1 and MCoTI-II. MCoAEP was selected for large-scale cyclization of MCoTI-II CPs and [C247A]OaAEP1 was used for the cyclization of cVc1.1 (Yap et al. 2021). A 5 L laboratory scale-up production of cVc1.1 and MCoTI-II reported yields of 85–97 mg L^{-1} culture, the highest ever reported in any recombinant production of CDRPs and an improvement of 2–3 order of magnitude compared to previous studies, in a preliminary unoptimized bioreactor (Yap et al. 2020; Yap et al. 2021).

Establishing reliable and robust validation checkpoints is integral to any successful therapeutic production. CDRPs with more than one disulfide bond may fold into multiple isomers with different disulfide bond connectivities, resulting in different 3D structures with different activities compared to the native fold (El Hamdaoui et al. 2019). Structural analysis by NMR spectroscopy coupled with HRMS analysis was used to validate recombinantly produced CDRPs (Fig. 5d). *P. pastoris*-derived AEP-cyclized MCoTI-II and [G22N]cVc1.1 were equivalent to their chemically-derived counterparts based on mass validation using HRMS and further supported by structural validation using 1D and 2D ^{1}H NMR spectroscopy (Yap et al. 2020; Yap et al. 2021). Both recombinant CDRPs were found to be equipotent compared to their respective chemically-derived counterparts in biological assays, a trypsin inhibition assay for MCoTI-II and a human G protein-coupled $GABA_B$ receptor-mediated inhibition assay of $Ca_V2.2$ (N-type) voltage-gated calcium channel for [G22N]cVc1.1 (Fig. 5e) (Yap et al. 2020; Yap et al. 2021). These validation checkpoints prove that recombinantly produced CDRPs were structurally equivalent and functionally equipotent compared to the chemically synthesized counterparts.

In summary, our *P. pastoris*-based cyclic peptide production platform has demonstrated versatility, scalability, suited for easy downstream purification, and achieved yields heading in the right direction for commercial production. More importantly, this platform is more environmentally sustainable or "greener" than SPPS (Yap et al. 2020) because it does not rely on hazardous solvents or generate toxic waste.

4 Future Directions and Conclusions

This *P. pastoris*-based production platform for CDRPs on a non-optimized laboratory-scale bioreactor production run yields about 100 mg L^{-1} culture, representing a 2–3 order of magnitude improvement compared to yields from previous bioproduction systems. Further optimization of the current platform to achieve g L^{-1} yields is feasible. One such process optimization is utilizing minimal growth

media instead of complex media where individual chemical components are further optimized. Another strategy is exploring different methanol feeding strategies. These optimization processes can drastically increase *P. pastoris* biomass, in turn increasing yields of CPs. A study producing rhamnosidase demonstrated that by switching to a minimal growth media and optimizing the methanol feeding strategy, dry cell weight of *P. pastoris* culture increases significantly from 8.3 g L^{-1} to 60 g L^{-1}, in turn, achieving production yields of 2 g L^{-1} (Markošová et al. 2015).

The current *P. pastoris*-based CDRPs production platform uses highly flammable methanol to induce expression which may not be amenable to commercial manufacturing. Methanol induction on a commercial scale presents significant manufacturing challenges. For example, a 10,000 L bioreactor production would require 50 L of methanol at 0.5% methanol induction every 12-24 hours, which poses significant flammable risks. A focus on methanol-free *P. pastoris* production is trending in this research space (Shen et al. 2016). Adapting the current platform towards methanol independent expression would be a significant step toward commercial manufacturing of CDRPs.

The capability of the yeast-based bioproduction platform to fold structurally more complex CDRPs and graft non-native cyclotides will need to be evaluated in future work. Cyclic chlorotoxin, a cyclic peptide with four disulfide bonds, is an ideal candidate to test the folding capability of the yeast-based bioproduction platform on more complex peptides (Akcan et al. 2011). Currently, the yeast-based bioproduction platform has produced only native cyclotide, MCoTI-II (Yap et al. 2020; Yap et al. 2021). Future endeavors will determine if grafted cyclotides with major amino acid sequence changes can be properly folded, secreted, and cyclized using the *P. pastoris*-based bioproduction platform.

The field of CDRPs is still in its infancy yet has garnered immense research interest as drug modalities for medical applications (Craik et al. 1999). The issues of industrial manufacturing of potential CDRP-based drugs in an economically and environmentally viable manner remain unsolved. These high-value molecules have in the past been produced exclusively by chemical synthesis which is not readily scalable and not environmentally sustainable. Previous attempts at producing CDRPs in bacterial or plant-based systems have yet to approach the realms of commercial viability. The *P. pastoris*-based bioproduction platform is more FDA friendly and has been used to produce FDA-approved drugs. The *P. pastoris*-based bioproduction platform for CDRPs described in this chapter is the first bioproduction platform moving in the right direction towards commercial-scale manufacturing of CDRP-based drugs in a sustainable and economically viable manner.

Acknowledgments Work in our laboratories is supported by access to the facilities of the Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science (CE200100012).

References

Aboye TL et al (2012) Design of a novel cyclotide-based CXCR4 antagonist with anti-human immunodeficiency virus (HIV)-1 activity. J Med Chem 55:10729–10734

Akcan M et al (2011) Chemical re-engineering of chlorotoxin improves bioconjugation properties for tumor imaging and targeted therapy. J Med Chem 54:782–787

Arntzen C (2015) Plant-made pharmaceuticals: from 'edible vaccines' to Ebola therapeutics. Plant Biotechnol J 13:1013

Bally J et al (2018) The rise and rise of *Nicotiana benthamiana*: a plant for all reasons. Annu Rev Phytopathol 56:405–426

Camarero JA, Fushman D, Cowburn D, Muir TW (2001) Peptide chemical ligation inside living cells: in vivo generation of a circular protein domain. Bioorg Med Chem 9:2479–2484

Camarero JA, Kimura RH, Woo YH, Shekhtman A, Cantor J (2007) Biosynthesis of a fully functional cyclotide inside living bacterial cells. Chembiochem 8:1363–1366

Camarero JA, Muir TW (1999) Biosynthesis of a head-to-tail cyclized protein with improved biological activity. J Am Chem Soc 121:5597–5598

Chan LY, Craik DJ, Daly NL (2015) Cyclic thrombospondin-1 mimetics: grafting of a thrombospondin sequence into circular disulfide-rich frameworks to inhibit endothelial cell migration. Biosci Rep 35:e00270

Chan LY et al (2011) Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds. Blood 118:6709–6717

Cheneval O et al (2014) Fmoc-based synthesis of disulfide-rich cyclic peptides. J Org Chem 79: 5538–5544

Choi B-K et al (2003) Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. Proc Natl Acad Sci U S A 100:5022–5027

Ciofalo V, Barton N, Kreps J, Coats I, Shanahan D (2006) Safety evaluation of a lipase enzyme preparation, expressed in *Pichia pastoris*, intended for use in the degumming of edible vegetable oil. Regul Toxicol Pharmacol 45:1–8

Clark RJ et al (2010) The engineering of an orally active conotoxin for the treatment of neuropathic pain. Angew Chem Int Ed 49:6545–6548

Colgrave ML, Craik DJ (2004) Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot. Biochemistry 43:5965–5975

Colgrave ML, Korsinczky MJ, Clark RJ, Foley F, Craik DJ (2010) Sunflower trypsin inhibitor-1, proteolytic studies on a trypsin inhibitor peptide and its analogs. Biopolymers 94:665–672

Craik DJ, Daly NL, Bond T, Waine C (1999) Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. J Mol Biol 294:1327–1336

Craik DJ, Fairlie DP, Liras S, Price D (2013) The future of peptide-based drugs. Chem Biol Drug Des 81:136–147

Dawson PE, Muir TW, Clark-Lewis I, Kent SB (1994) Synthesis of proteins by native chemical ligation. Science 266:776–779

Dirisala VR et al (2017) Recombinant pharmaceutical protein production in plants: unraveling the therapeutic potential of molecular pharming. Acta Physiol Plant 39:1–9

Du J et al (2020) A bifunctional asparaginyl endopeptidase efficiently catalyzes both cleavage and cyclization of cyclic trypsin inhibitors. Nat Commun 11:1575

El Hamdaoui Y et al (2019) Periplasmic expression of 4/7 α-conotoxin TxIA analogues in *E. coli* favours ribbon isomer formation-suggestion of a binding mode at the α7 nAChR. Front Pharmacol 10:577

Eliasen R et al (2012) Design, synthesis, structural and functional characterization of novel melanocortin agonists based on the cyclotide kalata B1. J Biol Chem 287:40493–40501

Fox JL (2012) Peptide therapeutics: current status and future directions. Nat Biotechnol 30:472

Gould A et al (2012) Recombinant production of rhesus θ-defensin-1 (RTD-1) using a bacterial expression system. Mol BioSyst 8:1359–1365

Gran L (1970) An oxytocic principle found in *Oldenlandia affinis* DC. Medd Nor Farm Selsk 12:80

Gründemann C, Stenberg KG, Gruber CW (2019) T20K: an immunomodulatory cyclotide on its way to the clinic. Int J Pept Res Ther 25:9–13
Huang Y-H et al (2015) Design of substrate-based BCR-ABL kinase inhibitors using the cyclotide scaffold. Sci Rep 5:1–15
Isidro-Llobet A et al (2019) Sustainability challenges in peptide synthesis and purification: from R & D to production. J Org Chem 84:4615–4628
Jackson MA et al (2019) Rapid and scalable plant-based production of a potent plasmin inhibitor peptide. Front Plant Sci 10:602
Jad YE, Kumar A, El-Faham A, de la Torre BG, Albericio F (2019) Green transformation of solid-phase peptide synthesis. ACS Sustain Chem Eng 7:3671–3683
Jagadish K et al (2015) Recombinant expression and phenotypic screening of a bioactive cyclotide against α-synuclein-induced cytotoxicity in baker′s yeast. Angew Chem Int Ed 54:8390–8394
James AM, Haywood J, Mylne JS (2018) Macrocyclization by asparaginyl endopeptidases. New Phytol 218:923–928
Johnson IS (1983) Human insulin from recombinant DNA technology. Science 219:632–637
Katrolia P et al (2011) Molecular cloning and high-level expression of a β-galactosidase gene from *Paecilomyces aerugineus* in *Pichia pastoris*. J Mol Catal B Enzym 69:112–119
Kimura RH, Tran AT, Camarero JA (2006) Biosynthesis of the cyclotide kalata B1 by using protein splicing. Angew Chem Int Ed 45:973–976
Laukens B, De Wachter C, Callewaert N (2015) Engineering the *Pichia pastoris* N-glycosylation pathway using the GlycoSwitch technology in *Glyco-engineering*. Springer, pp 103–122
Li CY, Yap K, Swedberg JE, Craik DJ, de Veer SJ (2019) Binding loop substitutions in the cyclic peptide SFTI-1 generate potent and selective chymase inhibitors. J Med Chem 63:816–826
Li Y, Aboye T, Breindel L, Shekhtman A, Camarero JA (2016) Efficient recombinant expression of SFTI-1 in bacterial cells using intein-mediated protein trans-splicing. Pept Sci 106:818–824
Markošová K, Weignerová L, Rosenberg M, Křen V, Rebroš M (2015) Upscale of recombinant α-L-rhamnosidase production by *Pichia pastoris* MutS strain. Front Microbiol 6:1140
Merrifield RB (1963) Solid phase peptide synthesis. I. the synthesis of a tetrapeptide. J Am Chem Soc 85:2149–2154
Mullard A (2013) 2012 FDA drug approvals. Nat Rev Drug Discov 12:87
Mylne JS et al (2011) Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. Nat Chem Biol 7:257–259
Nguyen GK, Hemu X, Quek JP, Tam JP (2016) Butelase-mediated macrocyclization of d-amino-acid-containing peptides. Angew Chem Int Ed 55:12802–12806
Nguyen GK et al (2014) Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. Nat Chem Biol 10:732–738
Poon S et al (2018) Co-expression of a cyclizing asparaginyl endopeptidase enables efficient production of cyclic peptides in planta. J Exp Bot 69:633–641
Quimbar P et al (2013) High-affinity cyclic peptide matriptase inhibitors. J Biol Chem 288:13885–13896
Rehm FBH, Tyler TJ, Yap K, Durek T, Craik DJ (2020) Improved asparaginyl ligase-catalyzed transpeptidation via selective nucleophile quenching. Angew Chem Int Ed 60:4004–4008
Rehm FBH et al (2019) Site-specific sequential protein labeling catalyzed by a single recombinant ligase. J Am Chem Soc 141:17388–17393
Ritter SK (2017) Five green chemistry success stories. Chem Eng News 95:16–20
Roberts AW et al (2018) Functional characterization of a glycosyltransferase from the moss *Physcomitrella patens* involved in the biosynthesis of a novel cell wall arabinoglucan. Plant Cell 30:1293–1308
Sable R et al (2016) Constrained cyclic peptides as immunomodulatory inhibitors of the CD2: CD58 protein–protein interaction. ACS Chem Biol 11:2366–2374
Schillberg S, Finnern R (2021) Plant molecular farming for the production of valuable proteins–critical evaluation of achievements and future challenges. J Plant Physiol 258:153359

Shen W et al (2016) A novel methanol-free *Pichia pastoris* system for recombinant protein expression. Microb Cell Factories 15:1–11

Spohner SC, Müller H, Quitmann H, Czermak P (2015) Expression of enzymes for the usage in food and feed industry with *Pichia pastoris*. J Biotechnol 202:118–134

Streatfield SJ, Howard JA (2003) Plant production systems for vaccines. Expert Rev Vaccines 2: 763–775

Swedberg JE et al (2016) Substrate-guided design of selective FXIIa inhibitors based on the plant-derived *Momordica cochinchinensis* trypsin inhibitor-II (MCoTI-II) scaffold. J Med Chem 59: 7287–7292

Thell K et al (2016) Oral activity of a nature-derived cyclic peptide for the treatment of multiple sclerosis. Proc Natl Acad Sci U S A 113:3960–3965

Tripathi NK, Shrivastava A (2019) Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. Front Bioeng Biotechnol 7:420

Walsh G (2010) Biopharmaceutical benchmarks 2010. Nat Biotechnol 28:917–924

Wang CK, Northfield SE, Huang Y-H, Ramos MC, Craik DJ (2016) Inhibition of tau aggregation using a naturally-occurring cyclic peptide scaffold. Eur J Med Chem 109:342–349

Wang CK et al (2014) Molecular grafting onto a stable framework yields novel cyclic peptides for the treatment of multiple sclerosis. ACS Chem Biol 9:156–163

Wong CT et al (2012) Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment. Angew Chem Int Ed 51:5620–5624

Yang R et al (2017) Engineering a catalytically efficient recombinant protein ligase. J Am Chem Soc 139:5351–5358

Yao J, Weng Y, Dickey A, Wang KY (2015) Plants as factories for human pharmaceuticals: applications and challenges. Int J Mol Sci 16:28549–28565

Yap K et al (2020) An environmentally sustainable biomimetic production of cyclic disulfide-rich peptides. Green Chem 22:5002–5016

Yap K et al (2021) Yeast-based bioproduction of disulfide-rich peptides and their cyclization via asparaginyl endopeptidases. Nat Protoc 16:1740–1760

Hyaluronic Acid (Hyaluronan)

Meliawati Meliawati, Moritz Gansbiller, and Jochen Schmid

Contents

Abstract Hyaluronic acid is one of the most valuable polysaccharides due to its enormous and unique biofunctionality in the human body, which renders it of highest interest for pharmaceutical and cosmeceutical applications. Within this chapter, we will give an overview of the development of hyaluronic acid as a commercial product, including the various origins, recombinant production, optimization of the production process as well as purification strategies. The main scope will be on microbial production of hyaluronic acid, the biosynthetic pathway, the different fermentation processes and strategies for overproduction, as well as optimized downstream processing. We will also give an overview of the commercial producers and their current production processes as well as the patents which are currently active in that field. In addition, we will present and discuss applications in the field of cosmetics, pharmaceuticals as well as material science. In sum, we will give a current and comprehensive overview of hyaluronic acid production in the year 2021.

M. Meliawati · M. Gansbiller · J. Schmid (✉)
Institute for Molecular Microbiology and Biotechnology, University of Münster, Münster, Germany
e-mail: jochen.schmid@uni-muenster.de

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_7

1 Introduction

As native polysaccharide, hyaluronan or hyaluronic acid (HA) is present in various parts of the animal and human body, but can also be found in the surrounding environment of microbial strains. Based on its nowadays known enormous health-promoting properties and biofunctionality, such as the unique water-binding capacity, it arose a high interest even from the beginning on. From the historical point of view, hyaluronic acid was first mentioned as an unusual uronic acid-containing carbohydrate polymer with an extremely high molecular weight as extracted from the vitreous of bovine eyes in the year 1934, which also resulted in the name, based on hyalos (stands for glass in the Greek language) and uran (an abbreviation for uronic acid) (Meyer and Palmer 1934). Soon after its discovery, the unique properties of this new biopolymer were described to be different from other glycosaminoglycans (GAGs), and its monomer composition was described to be composed of uronic acids and amino sugars, as well as traces of pentose (Meyer and Palmer 1934).

Over the next 10 years, HA was isolated from various animal organs, such as joint fluid, the umbilical cord, and nowadays it is known that HA can be extracted from almost all vertebrate tissues (Cowman et al. 2015a). In the year 1937, HA was for the first time extracted from the capsules of Streptococci groups A and C, which proved that it is also a microbial polysaccharide (Kendall et al. 1937). In the year 1948, the first report of the kinetics of enzymatic hydrolysis of HA was published (Dorfman 1948) and 3 years later, the first data about the structure of HA in aqueous solutions was published, describing the relationship between viscosity and velocity gradients to be dependent on higher concentrations of HA (Ogston and Stanier 1951). A random coil (irregular helical) confirmation of HA was determined via light scattering in the year 1955 for the first time (Laurent and Gergely 1955). In 1954, the disaccharide repeating unit of HA was reported, based on enzymatic cleavage experiments by the use of hyaluronidase from the *Streptococcus*, revealing that the disaccharide unit is composed of glucuronic acid and *N*-acetyl-glucosamine, without the presence of a pentose (Linker and Meyer 1954). In addition, the molecular weight (M_w) of the repeating disaccharide unit was determined to consist of 397 Da, which reinforced the formerly hypothesized disaccharide structure. Unlike sulfated polysaccharides, some of the initial proof of HA's ability to interact with living cells came from the observation that HA accelerates cell growth. It has also been observed that HA initiates some cell aggregation. This was the first indication of a unique binding of the polysaccharide to the cell surface and thus its highly valuable pharmacological as well as cosmeceutical properties. In addition, the different observed viscosities of HA solutions in presence of different inorganic salts caused scientific interest. In contrast to many other polysaccharides, the highest viscosity was observed in distilled water, and that phenomenon was proposed to be related to pH and ionic strength of the solution. This kind of behavior has been described for the first time by Fuoss et al. for polyelectrolytes (Fuoss 1948) and is still now a state of common knowledge. Fundamental research on the

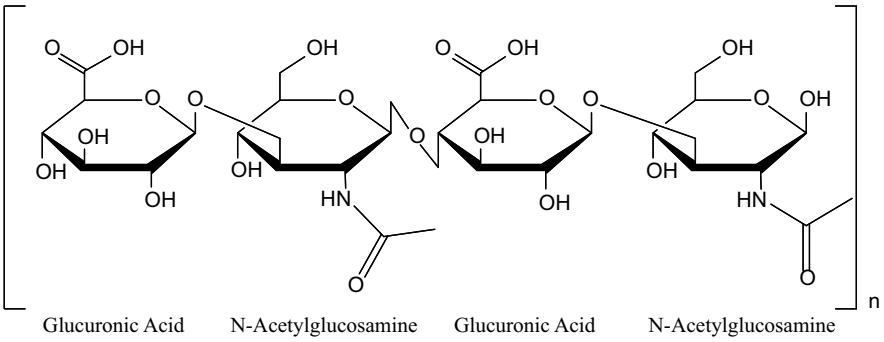


Fig. 1 Chemical structure of hyaluronic acid which is composed of disaccharide repeats of glucuronic acid and *N*-acetylglucosamine linked via β-(1–4) and β-(1–3)-glycosidic bonds

physicochemical properties of HA has already begun in the year 1951 with the study of Balazs et al. (Balazs 1979; Balazs and Laurent 1951). It also has been observed quite early that the high viscosity of HA solutions can be completely destroyed by illumination with UV light (Hvidberg et al. 1959) or exposure to X-rays in 1957 (Caputo 1957), indicating the quite unstable chemical structure of HA towards various external influences or stresses. Complete degradation of the chemical structure of HA has, later on, been described for gamma radiation exposure as well as low initial levels of ionization radiation in general (Kim et al. 2008b).

The native function as a highly valuable GAG in various parts of the human and animal body was elucidated to play diverse functions in structural and physiological maintenance of tissues, as well as the mediation of cell behaviors. It was shown to contribute to keeping tissue homeostasis and to function as a cell-signaling molecule via interaction with a variety of binding proteins. Based on its remarkable viscoelastic behavior, it also functions as a lubricant for the joints (Abatangelo et al. 2020; Dicker et al. 2014; Laurent and Fraser 1992). By that, this polysaccharide shows health-promoting properties and a high biofunctionality as well as an enormous water-binding capacity. Based on these highly valuable properties, HA is applied in medical as well as cosmetical applications, making it one of the most valuable and costly cosmeceutical polysaccharides. From the molecular structure (Donati et al. 2001), HA is a high molecular weight polymer of linear glycosaminoglycan which consists of glucuronic acid (GlcA) and *N*-acetylglucosamine (GlcNAc) repeats linked via β-(1–4) and β-(1–3)-glycosidic bonds (Fig. 1). By that, it is highly hydrophilic and charged, which renders its remarkable osmotic swelling capacity (Chong et al. 2005). At the beginning of its commercial production, HA was mainly obtained from the extraction of animal tissues and fluids, such as rooster combs, cattle vitreous humor, and bovine synovial fluid (Boeriu et al. 2013). Amongst most natural resources, rooster combs contain the highest content of HA with about 7.5 g/L (Laurent and Fraser 1992). Despite the high yields, animal-derived HA raised and experienced some concerns due to the risk of contaminations that can occur in the form of protein, nucleic acid, bacterial, or viral material. By that, intensive

purification is essential to achieve a certain level of HA purity and, as consequence, HA production costs are much higher compared to many other microbial polysaccharides and polymers (Freitas et al. 2011). Therefore, the development of safer and more sustainable sources and production routes is of high importance. Over the years, HA production has gradually shifted towards animal-free methods by means of microbial cell factories. Microbially, HA is naturally produced by some pathogens like *Streptococcus* sp. and *Pasteurella multocida* (DeAngelis and Achyuthan 1996; Sze et al. 2016). These bacteria use HA for encapsulating their cells by what they could escape the host's immune system, as HA masquerades the cells and protects them from inducing the immune response (Cress et al. 2014). From an industrial point of view, commercial HA production from these bacteria is complicated due to their pathogenic nature, which categorizes them as risk class 2 organisms. By that, large-scale industrial production is hampered by high regulatory hurdles and cost-intensive safety issues as well as purification procedures, which results in the high price of currently available HA products (Freitas et al. 2011).

2 Biosynthesis Pathway

Despite the unique properties, HA itself is a more "simple" polymer in regard to its molecular structure when compared to the other highly diverse microbial polysaccharides. Its biosynthesis follows the synthase-based pathway, which normally is used for the production of homopolysaccharides, which only consist of one type of carbohydrate monomers in their polymeric structure, such as cellulose or alginate (Schmid et al. 2015). Cellulose is a real homopolysaccharide that only consists of β-(1–4) linked glucose units and can be produced by eukaryotes (algae, plants) as well as prokaryotic bacteria via the synthase-based biosynthetic pathway (Rehm 2010). In the case of alginate biosynthesis, the initial polysaccharide is formed by polymerization of solely mannuronic acid (M) residues via the alginate synthase (Alg8 in *Pseudomonas aeruginosa*), from which some of them are converted towards guluronic acid (G) by the action of various epimerases, thus finally resulting in the typical M:G ratio of the different alginate types (Rehm and Valla 1997). In the case of HA synthesis, the HA synthase directly uses the two building blocks β-D-glucuronic acid and β-D-*N*-acetyl-glucosamine and alternating links them via β-(1–4) and β-(1–3)-glycosidic bonds.

Two classes of HA synthase (HAS) have been described today (Table 1). Most of the known HAS belong to Class I (Weigel 2015), while the HAS from *P. multocida* is the only member of the Class II (Weigel and DeAngelis 2007). Furthermore, the latter is the only known HAS coming from Gram-negative bacteria. In contrast to the Class I, Class II HAS is not an integral membrane protein but a membrane-anchored protein instead and has two protein domains which each has a glycosyltransferase activity. On the other hand, The Class I HAS is defined as a single domain integral membrane protein (Weigel 2015). Class I bacterial HAS belongs to the glycosyltransferases and is integrated into the membrane

Table 1 Comparison of Class I and Class II HAS

Properties	Class I	Class II
Source organisms	Streptococci, vertebrate, virus	*Pasteurella multocida*
Topology	Integral membrane protein	Peripheral protein
Protein domains	1	2
HA chain growth	Reducing end: Streptococci, mouse, human	Non-reducing end
	Non-reducing end: *Xenopus laevis*	
Expression as soluble active protein	Not possible	Possible ($PmHAS^{1-703}$)

polymerizing the precursor molecules by adding new moieties to the reducing end of the polysaccharide chain (Weigel and DeAngelis 2007). It contain a core of four transmembrane helices which are connected to at least one intracellular loop, in which the consensus sequence of processive glycosyltransferases is included. The same structure is found in cellulose and curdlan synthases (Heldermon et al. 2001; Saxena et al. 1995). A combined glycosyltransferase and translocase activity (Thomas and Brown 2010; Tlapak-Simmons et al. 1999; Weigel and DeAngelis 2007) or inclusion of an HA secreting ABC transporter (Ouskova et al. 2004; Schulz et al. 2007) is hypothesized as a reaction mechanism. But in contrast, cellulose synthase polymerizes at the non-reducing end, which is the main difference to the bacterial HAS. However, the HAS from the frog claw *Xenopus laevis* was shown to catalyze the polymerization at the non-reducing end (Bodevin-Authelet et al. 2005). The Streptococcal HAS is described to contain six membrane regions, of which four are integral and two are amphipathic, with an additional cytoplasmic domain (Heldermon et al. 2001). The integral membrane regions translocate the growing HA strands by forming a pore in the membrane, and the cytosolic part has the function of a glycosyltransferase domain that is expected to have several activities. These activities are described to include UDP-substrate binding and binding of the disaccharide repeating unit (HA-GlcA and HA-GlcNAc) as well as the two transferase activities to add GlcA and GlcNAc to the growing polymer chain (Weigel 2015).

For the microbial production strains, *Streptococcus zooepidemicus* represents the best examined organism which provided the most important insights on HA biosynthesis (Sze et al. 2016). The corresponding operon, namely *has*, consists of five genes: *hasA, hasB, hasC, hasD*, and *hasE*. The *hasA* gene encodes for HAS, the key enzyme in HA biosynthesis. The *hasB* and *hasC* genes encode for UDP-glucose dehydrogenase and UDP-glucose pyrophosphorylase, which generate UDP-GlcA. The *hasD* and *hasE* genes encode for pyrophosphorylase and phosphoglucoisomerase, which play an essential role in the generation of UDP-GlcNAc (Blank et al. 2008). Different compositions of the *has*-operon in bacteria have been described. For example, only the *hasA, hasB, hasC* genes are present in the *has*-operon of *Streptococcus pyogenes,* while the precursor encoding

genes for UDP-GlcNAc is located in another genomic region (DeAngelis et al. 1993). *Streptococcus uberis,* for example, carries only *hasA* and *hasB* located in the operon while a *hasC* homolog is present at a distinct locus (Ward et al. 2001). Overall, the biosynthesis of HA follows the synthase-dependent pathway in which the HAS catalyzes not only polymerization, but also translocation and secretion (Weigel 2015). Up to now the complete mechanism of HAS is not clarified based on the complex structure and the integral membrane domains. Just recently, the first three-dimensional atomic-scale model was presented, by which it was able to identify nine HAS-specific sub-structural elements and to elucidate their roles in HA biosynthesis (Agarwal et al. 2019). A combination of in silico modelling and mutation experiments suggested a three-step molecular mechanism for the growing HA chain from the reducing end in combination with overlapping binding sites of UDP-GlcNAc and UDP-GlcA. This 3-step mechanism can be briefly summarized to involve (a) Release of the bound UDP-substrate from the polymer (results in a glycosyl enzyme intermediate), (b) Release of the two bases with catalytic activity caused by the conformational change from UDP-release, (c) Glycosyltransfer reaction ending up in a glycosidic linkage and complete release of the second catalytic base 2. By that insights, a targeted engineering of HAS concerning substrate composition and the M_w comes very close (Fig. 2).

3 Rheological Properties

The structure of HA (Fig. 1) allows the formation of several hydrogen bonds within the molecule, which leads to a rod-like extended structure with high rigidity (Khabarov et al. 2014; Scott and Heatley 1999). These structures may form secondary and tertiary structures like twisted ribbons and double helices with other molecule strands (Scott et al. 1991; Scott and Heatley 1999). Depending on both concentration and M_w, these molecules in solution exhibit different viscoelastic properties, from solutions with low viscosity to high viscosity solutions with shear thinning behavior. These effects can be simply explained by the overlapping of the molecules, to form structural networks, leading to these properties. By that, it becomes evident, that smaller molecule chains (low M_w) require higher concentrations for overlapping compared to larger molecules (high M_w) (Dodero et al. 2019). The concentrations at which polymer chains interactions may just occur are called critical overlap concentration c^*. For hyaluronan with M_w between 0.9 and 6 MDa values between 10.9 and 0.32 mg mL^{-1} are reported (Cowman et al. 2015b). These effects lead to great variations in the flow behavior of HA solutions. For example, a 1% solution of high M_w HA (1.1–4.3 MDa) is described as shear thinning with zero shear viscosities of 1500–220,000 mPa s, while these values increase 10.5–12-fold by doubling the concentration to 2% (Bothner and Wik 1987). A more recent study shows similar zero shear viscosities of 900–117,000 mPa s of 1% solutions of HA between 1.1 and 4.0 MDa, while the concentration dependency increases with increasing M_w (Dodero et al. 2019). On the other hand, low M_w HA (150 kDa)

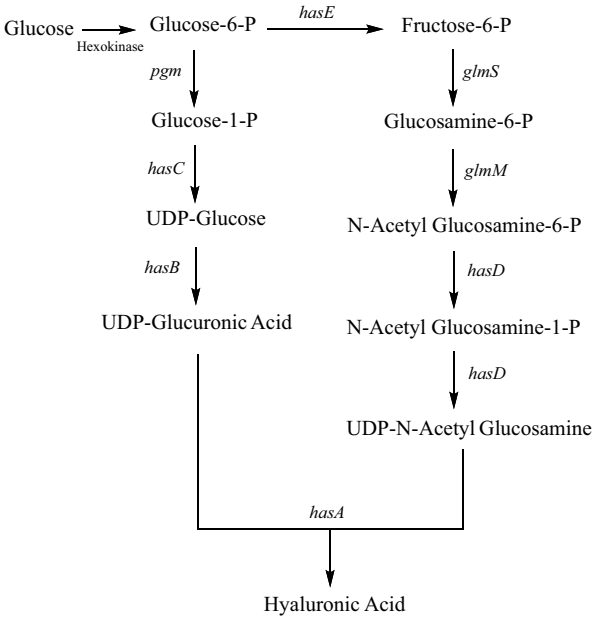


Fig. 2 Overview of HA biosynthesis pathway in *S. zooepidemicus.* In the cell, glucose is converted to glucose-6-phosphate by hexokinase activity, then finally into UDP-glucuronic acid by reaction steps which involve phosphoglucomutase as encoded by the *pgm* gene, UDP-glucose pyrophosphorylase (*hasC*) and UDP-glucose dehydrogenase (*hasB*). On the other hand, phosphoglucoisomerase (*hasE*) converts the glucose-6-phosphate to fructose-6-phosphate, which finally will be converted into UDP-*N*-acetyl glucosamine by the amidotransferase (*glmS*), mutase (*glmM*), and acetyltransferase-pyrophosphorylase (*hasD*). The two precursors will be polymerized by the hyaluronic acid synthase (*hasA*) to finally form hyaluronic acid

exhibits almost Newtonian (i.e., shear rate independent) flow behavior with viscosities of about 40–100 mPa s (Ambrosio et al. 1999). As a viscoelastic material, however, the viscoelastic properties of HA solutions are more important compared to its flow behavior. Due to the entanglements as already described, above certain M_w and concentrations, HA forms pseudo-gel structures, which differ from real gel structures, where the molecules interact via stronger intermolecular interactions, such as ion-mediated interactions or covalent bonds. Considering the properties of pure HA in solution, some studies have been conducted in the form of oscillatory shear experiments, starting as early as 1968. HA shows the behavior of a Maxwell fluid, exhibiting liquid behavior at lower frequencies and gel-like properties at high frequencies, which are separated by a crossover frequency (Gibbs et al. 1968). These first investigations by custom-designed oscillatory Couette type rheometers in 1968 showed a pH-dependency of this crossover frequency, where the lowest crossover frequency was observed at pH 2.5 and attributed to increased chain stiffness at this pH and was recently extended by combining modern oscillatory shear experiments with micro-rheology (dynamic light scattering. This allowed an extension of the

initial frequency range of 10^{-1}–10^{1} rad s^{-1} to 10^{-2}–10^{5} rad s^{-1}, and the observation of the intermediate rubbery region and a second crossover frequency, which are typical for non-crosslinked entangled polymer solutions (Dodero et al. 2019). Despite the importance of understanding the rheological properties of HA at different concentrations and M_w and the availability of highly sensitive and versatile rheometry, there is still a lack of further rheological behaviors like time- and temperature-dependent behavior as most studies focus on the frequency dependency. Another cause of this lack of basic characterization might be found in the recent focus on the characterization of modified HA, HA blends and formulations, as research of HA gets more application-focused and product-driven.

4 Fermentative Production

4.1 Natural Producers

Based on the history of HA production, the native microbial producers are of high interest to replace animal-derived HA production. Most research focuses on their exploitation to produce HA in high yields with certain purity level and M_w, which highly affects its rheological properties and thus applications. Major microbial HA production is obtained from the genus *Streptococcus*. Since they are known pathogens, high safety measures are required for handling the bacteria, and the HA obtained thereof during fermentation and downstream processing.

Streptococci are facultative anaerobic bacteria that mainly produce lactate as the main fermentation product. The composition of fermentation products changes depending on the conditions applied and therefore optimization of process parameters is essential to intensify HA production (Chong et al. 2005). For this, the effect of different process parameters like aeration, agitation, pH, and temperature have been intensively investigated (Liu et al. 2018). HA biosynthesis, as most exopolysaccharides (EPS) pathways, is a metabolical energy-intensive process where five ATPs are required to form one disaccharide unit (Armstrong et al. 1997). By that, aerobic conditions are much more favorable than microaerophilic or anaerobic conditions since more energy for HA synthesis can be provided. Chong and Nielsen (2003) reported that aerobic conditions resulted in 50% increase of HA productivity and doubled the M_w (Fong Chong and Nielsen 2003). On the other hand, HA titer and M_w require different optimal pH and temperature therefore highest titer and M_w might not be realized at the same time. Highest HA titer was achieved at pH 7.0, while highest M_w at pH 8.0. Similarly, highest HA titer was obtained at 37 °C, but lower temperature contributed to higher M_w. For this reason, a two-stage fermentation represents a promising strategy for the optimal production of HA with high M_w (Liu et al. 2018). Furthermore, as for most microbial EPS production processes, one of the main issues of HA production is the increasing viscosity of the fermentation broth, which finally limits the oxygen and mass transfer rates in the bioreactor. The titers of microbial HA production vary depending on

different factors such as the production strain, media composition as well as bioreactor design, but the achievable titers from *Streptococcus* sp. fermentation are up to 6–7 g/L (Lu et al. 2016; Zhang et al. 2006). While various carbon sources can be used, glucose is the typical substrate and delivers the highest HA yields (Im et al. 2009). The product titer and M_w can be increased by different approaches like temperature switches as well as the addition of intermediates and precursors for HA biosynthesis (Jagannath and Ramachandran 2010). In addition, the process variant can also strongly affect HA production and batch fermentation is commonly used due to its simplicity and product flexibility for industrial producers, but fed-batch processes are revealed to increase yields of HA production in combination with reduced process times (Liu et al. 2011). Continuous processes for microbial EPS production on an industrial scale are quite difficult to realize due to the high dilution rates which are required for cell removal and, in general, can cause genetic instability of the production strain what renders them unfavorable for commercial application on microbial processes (Blank et al. 2005; Chong et al. 2005).

From the molecular biology point of view, most studies suggested that HAS is the key enzyme as well as the limiting factor in HA production. Overexpression of the endogenous *hasA* is described to increase HA titers of *Streptococcus* sp. (Izawa et al. 2011; Zakeri et al. 2017). Competition of HA biosynthesis and biomass formation for precursors needs fine-tuning of cell growth and HA production to obtain maximal product titers. Elimination of competing pathways is a pivotal approach for optimizing HA production. For this, typical overflow metabolites of *Streptococcus* sp. such as lactate, acetate, formate, and ethanol were successfully targeted to channel the carbon flux towards HA biosynthesis (Chong and Nielsen 2003). Furthermore, inactivation of the hyaluronidase-encoding gene is important to prevent degradation of the HA produced. Deletion of hyaluronidase encoding gene in *S. zooepidemicus* was successfully applied to improve the titer and M_w of produced HA (Pourzardosht and Rasaee 2017). In general, most research for HA production by *Streptococcus* sp. focus on optimization of the fermentation process much more than on engineering of the production strain. This, on the one hand, can be attributed to the limited genetic tools which are available for Streptococci and, on the other hand, to the customers demand for green, biobased and not "genetically engineered" products. But next to process optimization, strain engineering is an essential and highly efficient way to further boost HA production towards economic production, which gets more and more accepted by the customers and end-users.

4.2 Recombinant Production

Today, the majority of commercial HA is produced by the natural producer *Streptococcus* sp. However, the pathogenic nature of *Streptococci* also raises some concerns based on the presence of endotoxins which requires extensive purification and limits the applications of HA in medical sectors (Liu et al. 2011). By that, many efforts have been conducted to realize heterologous HA production by

non-pathogenic microorganisms to reduce the complexity of the production process. However, initial recombinant expression resulted in quite low product titers, which rendered native HA production the most successful process for a quite long time (Liu et al. 2011). By means of sophisticated metabolic engineering approaches and especially synthetic biology-driven engineering, recombinant HA production reached a competitive and even superior approach during the last years. Compared to other important EPS, the HA biosynthesis comprises a minimalistic encoding operon. For instance, the size of the *has*-operon in *S. zooepidemicus* is 6.8 kb which is much smaller than the xanthan operon of *Xanthomonas campestris,* which is 14.5 kb in size, and even much less in size compared to the welan operon of *Sphingomonas* sp. with 30.5 kb (Schmid et al. 2015). Interestingly, only *hasA* is missing while the homologs of other *has* genes are readily available in the genome of the recombinant hosts. By that, heterologous expression of *hasA*, supported by overexpression of the native *has* homologs, is usually sufficient to direct the engineered strain into HA production.

Recombinant HA production has been demonstrated in microorganisms, such as *Escherichia coli* (Mao et al. 2009; Woo et al. 2019)*, Corynebacterium glutamicum* (Cheng et al. 2019; Hoffmann and Altenbuchner 2014; Wang et al. 2020)*, Lactococcus lactis* (Jeeva et al. 2019)*, Bacillus subtilis* (Li et al. 2019; Widner et al. 2005)*, and Pichia pastoris* (Jeong et al. 2014). For this, the key enzyme of HA biosynthesis, HAS, is isolated from the natural producer and heterologously expressed along with overexpression of the native genes encoding for the precursors formation of the selected host. Most studies utilized the Class I HAS from *Streptococcus* sp., whereas other studies utilized the Class II HAS from *P. multocida*. The product titer and M_w of the resulting HA vary depending on the engineered strains and process parameters (Table 2). Early recombinant production often suffered from low product titers, but just recently *C. glutamicum* was developed towards a highly promising recombinant HA producer. By a series of metabolic engineering approaches, Wang et al. have successfully designed an outstanding production strain with HA titers of 74 g/L (Wang et al. 2020). This is much higher compared to native production in *Streptococcus* sp. with 6–7 g/L (Lu et al. 2016; Zhang et al. 2006). However, such a high titer has to be compensated with low M_w. The average M_w of HA produced by the engineered strain was 53 kDa, which is significantly lower in comparison to the one produced by *Streptococcus* sp. with 2 MDa. It is hypothesized that a high HA titer could be achieved by rapid polymerization by the HAS (Cheng et al. 2019).

Today, industrial-scale HA production by means of recombinant cell factories has been conducted by Novozyme via an engineered *B. subtilis* strain. The related study described the introduction of the *hasA* gene from *S. equisimilis* together with overexpression of the endogenous *tuaD*, *gcaD*, and *gtaB* genes, which are homologs of *hasB*, *hasC*, and *hasD*, respectively. Overexpression of *tuaD* resulted in significantly increased HA yields, while additional overexpression of *gcaD* or *gtaB* showed only a minor impact on the production, indicating that UDP-GlcUA as a precursor is the limiting factor for HA biosynthesis in *B. subtilis*. This is in contrast to *S. zooepidemicus* where it was suggested that UDP-GlcNAc was the limiting

Table 2 Overview of HA titers and molecular weight from different sources, production method, and cultivation conditions

Organism source	Strain number	Genetic modifications	Culture mode	Working volume	Culture media	HA titer	M_w	References
Human or animal-derived								
Rooster combs					Chemical extraction	1 mg/g	1.2 MDa	Kang et al. (2010)
Human umbilical cord residual					Chemical extraction	n.a[a]	0.5 MDa	Lago et al. (2005)
Native bacterial producer								
S. zooepidemicus	IBRC-M 10919	Deletion of hyaluronidase	Shake flask	200 mL	TSB	~9 g/L	3.8 MDa	Pourzardosht and Rasaee (2017)
S. zooepidemicus	NJUST01		Shake flask	100 mL	Media with glucose and yeast extract	6.7 g/L	1.5–2.1 MDa	Zhang et al. (2006)
S. zooepidemicus	R42	Overexpression of bacterial hemoglobin from *Vitreoscilla* sp.	Batch	7 L	Media with sucrose, maltose, glutamic acid	6.7 g/L	0.4 MDa	Lu et al. (2016)
S. equi	RSKK 677		Shake flask	100 mL	Media with sucrose, peptone, and yeast extract	12 g/L	79.4 kDa	Güngör et al. (2019)
S. equisimilis	CVCC55116		Shake flask	20 mL	Media with glucose	175 mg/L	n.a.	Chen et al. (2012)
S. thermophilus	YIT 2084	Overexpression of *sthasA* and *sthasB*	Batch	1 L	Skim milk	1.2 g/L	1 MDa	Izawa et al. (2011)
Recombinant hosts								
Bacillus subtilis	168	Chromosomal integration of s*ehasA-bstuaD-bsgtaB*	Fed-batch	3 L	Minimal media with sucrose	n.a	1.1–1.2 MDa	Widner et al. (2005)

(continued)

Table 2 (continued)

Organism source	Strain number	Genetic modifications	Culture mode	Working volume	Culture media	HA titer	M_w	References
Bacillus subtilis	WB600	Overexpression of s*uhasA-bshasB-bshasC*	Batch	5 L	Minimal media with sucrose	3.65 g/L	0.4–7 MDa	Li et al. (2019)
Corynebacterium glutamicum	ATCC 13032	Overexpression of s*ehasA-cghasB*	Fed-batch	2 L	Minimal media with glucose	28.7 g/L	20–40 kDa	Cheng et al. (2019)
Corynebacterium glutamicum	ATCC 13032	Overexpression of s*zhasA*	Shake flask	25 mL	Minimal media with glucose	1.2 g/L	<270 kDa	Hoffmann and Altenbuchner (2014)
Corynebacterium glutamicum	ATCC 13032	Overexpression of *sphasA-cgugdA2-ppglmS*	Fed-batch	2.5 L	Glucose-corn steep powder medium	74.1 g/L	53 kDa	Wang et al. (2020)
Lactococcus lactis	NZ9000	Overexpression of s*zhasA-szhasB-szhasE-llpgmA*	Fed-batch	1.2 L	M17 media with glucose	~1 g/L	0.4–1.4 MDa	Jeeva et al. (2019)
Kluyveromyces lactis	GG799	Chromosomal integration of *pmhasA* and *xlhasB*	Batch	1 L	Yeast nitrogen base with glucose	1.89 g/L	2 MDa	Gomes et al. (2019)
Escherichia coli	JM109	Overexpression of p*mhasA* and e*ckfiD*	Fed-batch	500 mL	Terrific Broth	2.0–3.8 g/L	1.5 MDa	Mao et al. (2009)
Escherichia coli	K12W3310	Overexpression of s*zhasA* and *ecgalU-ecugd*	Shake flask	100 mL	LB with 3 g/L glucose and 3 g/L galactose	29.98 mg/L	1.4 kDa	Woo et al. (2019)
Streptomyces albulus	CRM003	Overexpression of s*zhasA-saudgA-saglmU-sagtaB*	Fed-batch	1.5 L	M3G medium with 60 g/L glucose	6.2 g/L	2 MDa	Yoshimura et al. (2015)
Agrobacterium sp.	ATCC 31749	Overexpression of p*mhasA* and e*ckfiD*	Shake flask	50 mL	Minimal media with sucrose, lactose, and glycerol	0.3 g/L	0.7–2 MDa	Mao and Chen (2007)

Pichia pastoris	GS115	Overexpression of *xlhasA2-xlhasB* and *ppahasC-ppahasD-ppahasE*	Fed-batch	1 L	Media with glucose and yeast extract	0.8–1.7 g/L	1.2–2.5 MDa	Jeong et al. (2014)
In vitro synthesis								
		$PmHAS^{1–703}$		5 mL	UDP-GlcA and UDP-GlcNAc	2.7 g/L	1.49 MDa	Gottschalk et al. (2019)
		$PmHAS^{1–703}$		300 µL	Sucrose and GlcNAc	4 g/L	2.3 MDa	Eisele et al. (2018)

bs: *B. subtilis*, cg: *C. glutamicum*, ec: *E. coli*, ll: *L. lactis,* pm: *P. multocida*, pp.: *Pseudomonas putida*, ppa: *P. pastoris*, sa: *Streptomyces avermitilis*, sp.: *S. pyogenes,* st: *S. thermophilus*, su: *S. uberis*, sz: *S. zooepidemicus*, xl: *X. laevis*

[a] n.a.: no values available

precursor (Chen et al. 2009). Finally, the recombinant HA had a M_w in the range of 1.1–1.2 MDa with a polydispersity index of 1.5 (Widner et al. 2005). Recombinant HA production in eukaryotic microbes is rarely described and has only been demonstrated for *Pichia Pastoris* (Jeong et al. 2014). By that approach, a HA titer up to 1.7 g/L with a M_w of 1.2 MDa by heterologous expression of *hasA* gene from *X. laevis* along with the native *hasC*, *hasD*, *hasE* homologs could be achieved. Similar to *B. subtilis,* UDP-GlcA is also the limiting precursor in *P. pastoris.* Furthermore, it was demonstrated that expression of *hasA* under the regulation of a weak promoter resulted in HA with higher M_w in comparison to expression by the use of a strong promoter. This was due to high level of HAS resulted in the rapid exhaustion of the intracellular HA precusors, which led to a low ratio between the precursors and HAS and might demonstrate the high processivity of the HAS. For this, fine-tuning of the expression level of the HAS as well as the precursors genes are essential to achieve optimal titers and desired M_w.

4.3 In Vitro Production

Apart from microbial fermentation, several studies have also explored the potential of HA production via in vitro synthesis. This method is relatively new developed and some studies reported the production of HA in vitro in small-scale reactions. In principle, HA synthesis can be achieved by mixing the sugar nucleotide precursors, cofactors, and the HAS protein. Due to the nature of membrane-spanning Class I HAS, the research was mostly conducted on membrane fractions (Boeriu et al. 2013). Although in vitro HA production is possible, high costs of membranes, sugar nucleotides and cofactors hampers its feasibility at a commercial scale (Sze et al. 2016). For this, utilization of Class II HAS from *P. multocida* (PmHAS) is a better option since, unlike the class I HAS from *Streptococcus*, PmHAS is a peripheral protein and therefore must not be bound to a membrane to perform its function (DeAngelis 1996; Sze et al. 2016). In the year 2000, Jing and De Angelis created a mutant of the peripheral $PmHAS^{1-703}$, which showed high solubility in combination with high activity (Jing and DeAngelis 2000). One-pot synthesis of in vitro HA production includes the six enzymes: glucuronic acid kinase, UDP–sugar pyrophosphorylase and pyrophosphatase for generation of UDP-GlcA; GlcNAc-1-phosphate kinase, UDP–GlcNAc pyrophosphorylases and pyrophosphatase for UDP-GlcNAc; as well as HAS for polymerization and resulted in a titer of 2.7 g/L by use of GlcA and GlcNAc as substrates (Gottschalk et al. 2019). Furthermore, Eisele et al. (2018) demonstrated in vitro HA synthesis, which utilized sucrose and GlcNAc as the main substrates (Eisele et al. 2018). The strategy involved several enzyme module systems which allowed in situ regeneration of nucleotide sugars. For this, different recombinant enzymes have to be cloned, produced, and purified before being used in the enzymatic synthesis. In addition to the substrates and enzymes, additional components like NAD^+ and UDP also have to be added to the reaction mixtures. Finally, this approach achieved a final titer of 4 g/

L with M_w of 2.3 MDa. Further research is required to analyze the feasibility of in vitro HA synthesis, especially in scaling up experiments. Efficient production of the enzymes, substrates reutilization, as well as product separation, need to be addressed to produce HA with high yield and purity (Boeriu et al. 2013).

5 Extraction and Purification

As HA is often applied in medical or pharmaceutical applications, the purification process must guarantee that the final product is free from endotoxins and proteins. In the case of microbial production by S. *zooepidemicus*, the removal of streptolysin and proteins is of high importance, as these compounds might lead to lysis of red blood cells or cause strong immune responses, respectively (Bitterman-Deutsch et al. 2015; Franz et al. 2011). HA from animal tissue must be purified from protein and carries the inherent risk of contaminations with animal viruses or prions. Among other regulatory criteria of medically applied HA (Huerta-Angeles et al. 2016), recent sources state a required purity of a protein content below 0.3–0.1%, depending on the application (Ferreira et al. 2021; Hyaluronate 2016). An ultrapure HA with a protein content below 0.5% was described to be non-inflammatory at a concentration of 1% (Balazs 1979). Endotoxin levels must be below 0.05 I.U. mg^{-1}, and residual ethanol or other organic solvents $\leq$0.5% and bacterial contamination $\leq$100 cfu mg^{-1} (Ferreira et al. 2021; Hyaluronate 2016). These high demands on purity, especially for injection purposes, require extensive purification processes compared to other commercial EPS, resulting in the high market prices of HA.

In the case of HA production from animal sources, extraction steps precede the purification process. The first described extraction was done by acetone precipitation from bovine vitreous humor (Meyer and Palmer 1934). Later developed methods include extraction using organic solvents, isopropanol or sodium acetate, as well as quaternary ammonium salts. Enzymatic treatments using Alcalase, Papain, Pronase (a mixture of endo- and exopeptidases), and Trypsin have been described as well (Abdallah et al. 2020).

Following extraction from animal tissue or cell separation in the case of microbial production, several purification methods and combinations thereof have been investigated to obtain the required purities ($\geq$99.5%) for pharmaceutical and medical applications. Purification methods include precipitation with ethanol or 2-propanol (isopropanol), tangential flow filtration and diafiltration, silica gel adsorption, protein electrodeposition, dialysis, and anion exchange chromatography (Abdallah et al. 2020; Cavalcanti et al. 2020). The main goal of the HA purification process is the removal of protein and other compounds causing inflammation. From microbially produced HA, the protein content after precipitation ranges from 12 to 14%, which requires further purification steps. Mainly purification is applied by filtration, as this is quite cheap compared to other methods such as enzymatic treatment, adsorption methods or ion exchange chromatography (Ferreira et al. 2021).

Highest reported theoretical purities from *S. zooepidemicus* fermentations are described for protein concentration as low as 0.06% by combining precipitation, silica gel adsorption, charcoal filtration followed by 5× diafiltration and sterile filtration with a yield of 50% in respect to the unpurified precipitate (Rangaswamy and Jain 2008). Another study reports an equally pure product with a protein content of 0.07% by a combination of microfiltration and ultrafiltration in a diafiltration setup, yielding 89% starting from a 1 g/L HA solution. However, these purification steps require either very long treatment times (up to 74 h) or multiple processing steps, which renders HA production very expensive or impractical for industrial scale. A recently published study based on process simulation suggests an increase in production costs of 30–78% for the production of ultrapure HA suitable for injection compared to HA for topological applications, based on fermentative production by *Streptococcus* sp. (Ferreira et al. 2021). Their approach, however, targets the simultaneous production of both lower and higher purity grades with 90% and 10% production ratios, respectively. A comprehensive overview of sources, processing methods and purities of HA is shown in Table 3.

6 Commercial Producers

According to Grand View Research, the worldwide HA market size will reach USD 16.6 billion by 2027 and is projected to grow by an average annual growth rate of 8.1% for the coming years (Grand View Research 2020). HA market is anticipated to rise with an increasing aging population. The global pandemic Covid-19 also has a positive impact on the market, especially for injectable HA. Most HA product is commercialized as sodium hyaluronate, but other derivatives such as hyaluronan oligosaccharides or oxidized hyaluronic acid are also available. Contipro, one of the key players in HA market, provides HA in different forms, including fibers, hydrogen, films, or micelles. Shiseido is the first company to conduct large-scale HA production from non-animal sources. Today, most companies produce HA from microbial fermentations, by means of *S. zooepidemicus* (Table 4). Novozymes is the only company that is running commercial scale recombinant HA production from non-Streptococci strain, by utilizing *B. subtilis*. Founded in 2000, Hyalose is a USA-based company that focuses on the commercialization of technologies for HA production, from conventional animal extraction, microbial fermentation, to enzymatic HA synthesis.

7 Patent

Following its first discovery in the 1930s (Meyer and Palmer 1934), HA has garnered much attentions due to its impressive characteristics. Since then, many investigations focused on the exploitation of HA and its derivatives which results in

Table 3 Overview over animal and microbial HA sources and downstream operations, including their yields and purities of the final product. Adapted from Abdallah et al. (2020), Cavalcanti et al. (2020)

Animal sources	Downstream operations	Product yield	Purity	Source
Swordfish eyeball	Centrifugation	101.1 mg from 2.5 L raw extract	99.73% (0.009 mg/mL protein)	Murado et al. (2012)
	Ultrafiltration–diafiltration			
	Electrodisposition			
Shark eyeball	Alcoholic precipitation w/ 0.5 M NaCl	481.8 mg from 2.5 L raw extract	99.85% (0.007 mg/mL protein)	Murado et al. (2012)
	Alkaline process			
Mussel	Acetone treatment (removal of fat)	6.1 mg/g tissue dry weight	n.a.[a]	Volpi and Maccari (2003)
	Papain treatment			
	Sodium acetate/ethanol precipitation			
	Ion exchange chromatography			
Stingray liver	Acetone treatment (removal of fat)	0.81 mg/g tissue dry weight	n.a.	Sadhasivam et al. (2013)
	Papain treatment			
	Sodium acetate/ethanol precipitation			
	Ion exchange chromatography			
Yellowfin Tuna eyeball	Acetone precipitation	n.a.	n.a.	Mizuno et al. (1991)
	Actinase E treatment			
	Dialysation			
Bigeye Tuna eyeball	Cetylpyridinium chloride precipitation	0.42 mg/g dry weight	n.a.	Amagai et al. (2009)
	KAc/ethanol precipitation			
	Acetone precipitation			
	Dialysis			
Roster wattle	Papain treatment	17.9 mg/g	n.a.	Nakano et al. (1994)
	Dialysis			
	Cellulose acetate electrophoresis			
Rooster comb	Papain treatment	39.8 mg/g	n.a.	Nakano et al. (1994)
	Dialysis			
	Cellulose acetate electrophoresis			
	Pronase treatment	n.a.	n.a.	Swann (1968)
	Chloroform treatment			
	Ion exchange chromatography			
	Organic solvents/aetate extraction	n.a.	n.a.	Swann (1968)
	Chloroform treatment			
	Sodium acetate extraction	1 mg/g	n.a.	Kang et al. (2010)
	Dialysis	Frozen tissue		

(continued)

Table 3 (continued)

Chicken comb	Papain treatment Ethanol purification Centrifugation	15 mg (hexUA)/g dry tissue	n.a.	Rosa et al. (2012)
Bovine eyeball	Use of organic sodium salt Dialysis	469.9 mg/L vitreous humor	n.a.	Gherezghiher et al. (1987)
Bovine synovial fluid	Use of quaternary ammonium salt Dialysis	250 mg/L synovial fluid	n.a.	Matsumura et al. (1963)
	Trypsin and pronase treatment Chloroform treatment Filtration	n.a.	n.a.	Cullis-Hill (1989)
Pig eyeball	Ultrafiltration–diafiltration Protein electrodeposition	0.04 g/L vitreous humor	n.a.	Murado et al. (2012)
Pig synovial fluid	Trypsin and pronase treatment Chloroform treatment and filtration	n.a.	n.a.	Cullis-Hill (1989)
Sheep synovial fluid	Trypsin and pronase treatment Chloroform treatment and filtration	n.a.	n.a.	Cullis-Hill (1989)
Owl monkey eyeball	Organic solvents Chloroform treatment	n.a.	n.a.	Balazs (1979)
	Organic sodium salt Dialysis	291.8 mg/L vitreous humor	n.a.	Gherezghiher et al. (1987)
Eggshell membrane	Papain treatment	39.02 mg/g eggshell	n.a.	Khanmohammadi et al. (2014)
	Trypsin treatment	44.82 mg/g eggshell	n.a.	
	Isopropanol and sodium acetate Silica gel and activated carbon	5.3 mg/g eggshell	n.a.	
Microbial sources	**Downstream processing**	**Product recovery**	**Purity**	**Source**
Streptococcus zooepidemicus	Precipitation: 2-propanol (3:1 v/v); filtration: 2 filtration steps; ultrafiltration in diafiltration mode; adsorption: charcoal (1–2% w/v)	72.2%	99.2%	Jagadeeswara Reddy (2013)
	Precipitation: ethanol (3:2 to 3:1 v/v); filtration; adsorption: aromatic resin and activated charcoal carbon (3% w/v each)	78–82%	99%	Han et al. (2004)

(continued)

Table 3 (continued)

	Precipitation: isopropanol, ethanol, or acetone (1–1.5% v/v) and cetylpyridinium chloride (10% solution); adsorption: activated charcoal (0.1% w/v)	60–70%	87–92%	Nimrod et al. (1988)
	Precipitation: ethanol (1:1 v/v); filtration (diafiltration mode); adsorption: charcoal (2% w/v) and gamma-alumina (1% w/v)	n.a.	99.2%	Won et al. (2008)
	Precipitation: organic solvents (1:1 to 3:1 v/v); adsorption: bentonite (1% w/v) and activated carbon (3% w/v); ion exchange chromatography	99.3%	99.94%	Hemant et al. (2013)
Streptococcus pyogenes	Precipitation: ethanol (3:1 v/v) and CTAB 0.32%; tangential filtration (diafiltration mode)	n.a.	99.84%	Bracke and Thacker (1985)
Streptococcus equi	Precipitation: isopropyl alcohol (2:1 v/v); filtration	14.3–50%	99.8%	Brown et al. (1994)
	Precipitation: ethanol (1–2:1 v/v); enzymatic treatment: protease	n.a.	90%	Kim et al. (2008a)

[a] n.a.: no values available

an increasing number of patents concerning HA. First patent applications on HA were filled in the 1940s, and the number of patents keep growing over the years, notably in the past 20 years (Fig. 3). According to the European Patent Office, more than 100,000 HA-relevant patents have been submitted that cover various fields of synthesis, processing, and applications. The USA is the country with the most contribution in patent applications, followed by China and Japan. Shiseido holds most numbers of the patents with more than 1841 patents. Pharmaceutical companies like Sanofi, Pfizer, and Rohto Pharma also have ample numbers of HA patents. Furthermore, many universities are also among the major patent holders. More recent patents encompass HA formulations with other substances for applications in pharmaceutical, medical, and cosmetics fields. For example, WO2021020950A1 describes hydrogels that comprised of HA and pluronic for prevention and treatment of articular and cartilage injury.

8 Conclusion and Future Outlook

In conclusion, HA is a highly valuable polysaccharide with a broad range of medical applications. Based on relatively low stability of HA compared to other polysaccharides, the downstream processing is the main cost driving factor and must be

Table 4 Some major companies in HA production and technology

Company	Country	HA source	Commercial brands	Application fields
Contipro	Czech Republic	*Streptococcus zooepidemicus*	n.a.[a]	Research, medical, pharmaceutical, cosmetics, nutrition, veterinary
Shiseido	Japan	*Streptococcus zooepidemicus*	n.a.	Cosmetics and pharmaceutical
Lifecore	USA	*Streptococcus zooepidemicus*	Corgel	Research, pharmaceutical
Anika Therapeutics	USA	*Streptococcus zooepidemicus*	Orthovisc	Therapeutic
LG Chem	South Korea	*Streptococcus zooepidemicus*	Hyruan, Yvoire	Pharmaceutical, cosmetic, veterinary
Galderma	Switzerland	*Streptococcus zooepidemicus*	Restylane	Cosmetics
Allergan	USA	*Streptococcus zooepidemicus*	Juvéderm	Cosmetics
Sanofi	French	Rooster combs	Hyalgan	Therapeutic
Novozymes	Denmark	*Bacillus subtilis*	Hyasis, HyaCare	Pharmaceutical
Hyalose[b]	USA	In vitro *synthesis*	Select-HA	Process technology

[a] n.a. means no information available
[b] The company focuses on commercialization of technology for HA production

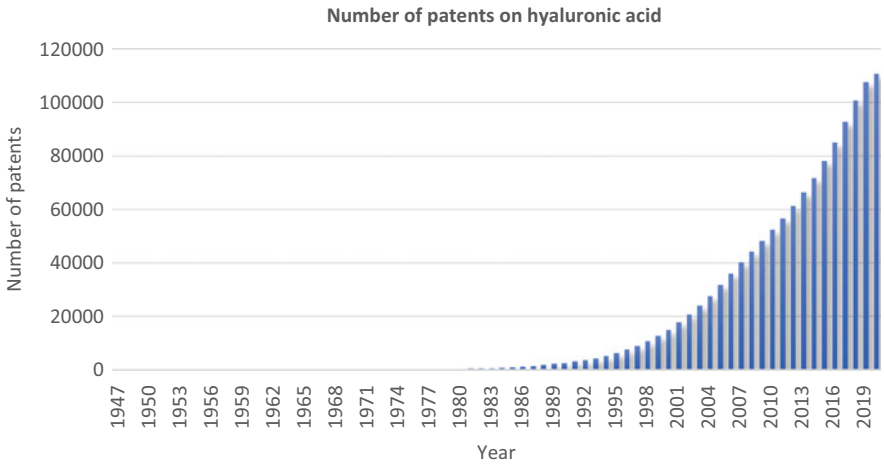


Fig. 3 Overview of number of patents related to HA over the years. (Data source: European Patent Office)

optimized in the future. As seen by the progress in heterologous expression, much higher product titer can be obtained than native production strains or HA of animal origins. By that, it can be expected that future large-scale production will be carried out via optimized production strains to improve yields and thus the economy of the whole production process. The in vitro synthesis of HA is not yet ready for large-scale and economical production due to the high cost of the substrates and enzymes production and purification. However, the progress in understanding and modelling enzymes by new and highly sophisticated tools such as AlphaFold in combination with targeted enzyme engineering might also improve the activity and stability of the HAS soon. By that, the in vitro approach might be massively improved.

Nevertheless, the production of the various enzymes still has to be considered when comparing the economics of the different process variants. An essential aspect that might define the future production processes is the desired M_w and purity for the different applications. Here, the selection of the production strain, in vitro synthesis by engineered enzymes or adapted downstream processing might be applied for HA production with specific M_w. In sum, the production of HA is of growing interest for the cosmeceutical industry, and many reports on the progress in HA production can be expected within the next years. Especially, the utilization of the latest molecular biology tools and approaches based on synthetic biology will massively accelerate the progress in HA research, thus making it a valuable polysaccharide of high functionality.

References

Abatangelo G, Vindigni V, Avruscio G, Pandis L, Brun P (2020) Hyaluronic acid: redefining its role. Cells 9:1–19

Abdallah MM, Fernández N, Matias AA, Bronze MR (2020) Hyaluronic acid and chondroitin sulfate from marine and terrestrial sources: extraction and purification methods. Carbohydr Polym 243:116441

Agarwal G, Krishnan KV, Prasad SB, Bhaduri A, Jayaraman G (2019) Biosynthesis of hyaluronic acid polymer: dissecting the role of sub structural elements of hyaluronan synthase. Sci Rep 9: 1–12

Amagai I, Tashiro Y, Ogawa H (2009) Improvement of the extraction procedure for hyaluronan from fish eyeball and the molecular characterization. Fish Sci 75:805–810

Ambrosio L, Bortacchiello A, Netti PA, Nicolais L (1999) Rheological study on hyaluronic acid and its derivative solutions. J Macromol Sci A 36:991–1000

Armstrong DC, Cooney MJ, Johns MR (1997) Growth and amino acid requirements of hyaluronic-acid-producing *Streptococcus zooepidemicus*. Appl Microbiol Biotechnol 47:309–312

Balazs EA (1979) Ultrapure hyaluronic acid and the use thereof. US Patent US4141973A

Balazs EA, Laurent TC (1951) Viscosity function of hyaluronic acid as a polyelectrolyte. J Polym Sci 6:665–667. https://doi.org/10.1002/pol.1951.120060517

Bitterman-Deutsch O, Kogan L, Nasser F (2015) Delayed immune mediated adverse effects to hyaluronic acid fillers: report of five cases and review of the literature. Dermatol Rep 7:5851

Blank LM, McLaughlin RL, Nielsen LK (2005) Stable production of hyaluronic acid in *Streptococcus zooepidemicus* chemostats operated at high dilution rate. Biotechnol Bioeng 90:685–693

Blank LM, Hugenholtz P, Nielsen LK (2008) Evolution of the hyaluronic acid synthesis (has) operon in *Streptococcus zooepidemicus* and other pathogenic Streptococci. J Mol Evol 67:13–22

Bodevin-Authelet S, Kusche-Gullberg M, Pummill PE, DeAngelis PL, Lindahl U (2005) Biosynthesis of hyaluronan. J Biol Chem 280:8813–8818

Boeriu CG, Springer J, Kooy FK, van den Broek LAM, Eggink G (2013) Production methods for hyaluronan. Int J Carbohydr Chem 2013:1–14

Bothner H, Wik O (1987) Rheology of hyaluronate. Acta Otolaryngol Suppl 442:25–30

Bracke JW, Thacker K (1985) Hyaluronic acid from bacterial culture. US Patent 4517295

Brown KK, Ruiz LLC, van de Wijn I (1994) Ultrapure hyaluronic acid and method of making it. US Patent 4782046A

Caputo A (1957) Depolymerization of hyaluronic acid by X-rays. Nature 179:1133–1134. https://doi.org/10.1038/1791133a0

Cavalcanti ADD, Melo BAG, Ferreira BAM, Santana MHA (2020) Performance of the main downstream operations on hyaluronic acid purification. Process Biochem 99:160–170

Chen S-J, Chen J-L, Huang W-C, Chen H-L (2009) Fermentation process development for hyaluronic acid production by *Streptococcus zooepidemicus* ATCC 39920. Korean J Chem Eng 26:428–432

Chong B, Nielsen L (2003) Amplifying the cellular reduction potential of Streptococcus zooepidemicus. J Biotechnol 100:33–41. https://doi.org/10.1016/s0168-1656(02)00239-0

Chen YH, Li J, Liu L, Liu HZ, Wang Q (2012) Optimization of flask culture medium and conditions for hyaluronic acid production by a *Streptococcus equisimilis* mutant NC2168. Braz J Microbiol 43:1553–1561

Cheng F, Yu H, Stephanopoulos G (2019) Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid. Metab Eng 55:276–289

Chong BF, Blank LM, Mclaughlin R, Nielsen LK (2005) Microbial hyaluronic acid production. Appl Microbiol Biotechnol 66:341–351

Cowman MK, Lee HG, Schwertfeger KL, McCarthy JB, Turley EA (2015a) The content and size of hyaluronan in biological fluids and tissues. Front Immunol 6:1–8

Cowman MK, Schmidt TA, Raghavan P, Stecco A (2015b) Viscoelastic properties of hyaluronan in physiological conditions. F1000Res 4:622

Cress BF, Englaender JA, He W, Kasper D, Linhardt RJ, Koffas MAG (2014) Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules. FEMS Microbiol Rev 38:660–697

Cullis-Hill D (1989) Preparation of hyaluronic acid from synovial fluid. US Patent 4879375

DeAngelis PL (1996) Enzymological characterization of the *Pasteurella multocida* hyaluronic acid synthase. Biochemistry 35:9768–9771

DeAngelis PL, Achyuthan AM (1996) Yeast-derived recombinant DG42 protein of *Xenopus* can synthesize hyaluronan in vitro. J Biol Chem 271:23657–23660

DeAngelis PL, Papaconstantinou J, Weigel PH (1993) Molecular cloning, identification, and sequence of the hyaluronan synthase gene from group A *Streptococcus pyogenes*. J Biol Chem 268:19181–19184

Dicker KT, Gurski LA, Pradhan-Bhatt S, Witt RL, Farach-Carson MC, Jia X (2014) Hyaluronan: a simple polysaccharide with diverse biological functions. Acta Biomater 10:1558–1570

Dodero A, Williams R, Gagliardi S, Vicini S, Alloisio M, Castellano M (2019) A micro-rheological and rheological study of biopolymers solutions: hyaluronic acid. Carbohydr Polym 203:349–355

Donati A, Magnani A, Bonechi C, Barbucci R, Rossi C (2001) Solution structure of hyaluronic acid oligomers by experimental and theoretical NMR, and molecular dynamics simulation. Biopolymers 59:434–445

Dorfman A (1948) The kinetics of the enzymatic hydrolysis of hyaluronic acid. J Biol Chem 172:377–387

Eisele A, Zaun H, Kuballa J, Elling L (2018) In vitro one-pot enzymatic synthesis of hyaluronic acid from sucrose and N-acetylglucosamine: optimization of the enzyme module system and nucleotide sugar regeneration. ChemCatChem 10:2969–2981

Ferreira RG, Azzoni AR, Santana MHA, Petrides D (2021) Techno-economic analysis of a hyaluronic acid production process utilizing *Streptococcal* fermentation. Processes 9:241

Fong Chong B, Nielsen LK (2003) Aerobic cultivation of *Streptococcus zooepidemicus* and the role of NADH oxidase. Biochem Eng J 16:153–162

Franz S, Rammelt S, Scharnweber D, Simon JC (2011) Immune responses to implants—a review of the implications for the design of immunomodulatory biomaterials. Biomaterials 32:6692–6709

Freitas F, Alves VD, Reis MAM (2011) Advances in bacterial exopolysaccharides: from production to biotechnological applications. Trends Biotechnol 29:388–398

Fuoss RM (1948) Viscosity function for polyelectrolytes. J Polym Sci 3:603–604. https://doi.org/10.1002/pol.1948.120030414

Gherezghiher T, Koss MC, Nordquist RE, Wilkinson CP (1987) Analysis of vitreous and aqueous levels of hyaluronic acid: application of high-performance liquid chromatography. Exp Eye Res 45:347–349

Gibbs DA, Merrill EW, Smith KA, Balazs EA (1968) Rheology of hyaluronic acid. Biopolymers 6:777–791

Gomes AMV, Netto JHCM, Carvalho LS, Parachin NS (2019) Heterologous hyaluronic acid production in *Kluyveromyces lactis*. Microorganisms 7:294

Gottschalk J, Zaun H, Eisele A, Kuballa J, Elling L (2019) Key factors for a one-pot enzyme cascade synthesis of high molecular weight hyaluronic acid. Int J Mol Sci 20(22):5664

Grand View Research (2020) Hyaluronic acid market size worth 16.6 billion by 2027 | CAGR: 8.1%. https://www.grandviewresearch.com/press-release/global-hyaluronic-acid-market

Güngör G, Gedikli S, Toptaş Y, Akgün DE, Demirbilek M, Yazıhan N, Aytar Çelik P, Denkbaş EB, Çabuk A (2019) Bacterial hyaluronic acid production through an alternative extraction method and its characterization. J Chem Technol Biotechnol 94:1843–1852

Han HY, Jang SH, Kim EC, Park JK, Han YJ, Lee CC, Park SH, Kim YC, Park HJ (2004) Microorganism producing hyaluronic acid and purification method of hyaluronic acid. Patent WO2004016771

Heldermon C, DeAngelis PL, Weigel PH (2001) Topological organization of the hyaluronan synthase from *Streptococcus pyogenes*. J Biol Chem 276:2037–2046

Hemant PN, Sonal T, Bondalakunta R (2013) Process for the purification of hyaluronic acid salts (HA) from fermentation broth. Patent WO2013132506

Hoffmann J, Altenbuchner J (2014) Hyaluronic acid production with *Corynebacterium glutamicum*: effect of media composition on yield and molecular weight. J Appl Microbiol 117:663–678

Huerta-Angeles G, Brandejsová M, Kulhánek J, Pavlík V, Šmejkalová D, Vágnerová H, Velebný V (2016) Linolenic acid grafted hyaluronan: process development, structural characterization, biological assessing, and stability studies. Carbohydr Polym 152:815–824

Hvidberg E, Kvorning SA, Schmidt A, Schou J (1959) Effect of ultraviolet irradiation on hyaluronic acid in vitro. Acta Pharmacol Toxicol 15:356–364. https://doi.org/10.1111/j.1600-0773.1959.tb00304.x

Hyaluronate S (2016) European pharmacopoeia, vol 2, 9th edn. European Directorate for Quality of Medicines & Health Care, Strasbourg

Im J-H, Song J-M, Kang J-H, Kang D-J (2009) Optimization of medium components for high-molecular-weight hyaluronic acid production by *Streptococcus* sp. ID9102 via a statistical approach. J Ind Microbiol Biotechnol 36:1337–1344

Izawa N, Serata M, Sone T, Omasa T, Ohtake H (2011) Hyaluronic acid production by recombinant *Streptococcus thermophilus*. J Biosci Bioeng 111:665–670

Jagannath S, Ramachandran KB (2010) Influence of competing metabolic processes on the molecular weight of hyaluronic acid synthesized by *Streptococcus zooepidemicus*. Biochem Eng J 48:148–158

Jeeva P, Shanmuga Doss S, Sundaram V, Jayaraman G (2019) Production of controlled molecular weight hyaluronic acid by glucostat strategy using recombinant *Lactococcus lactis* cultures. Appl Microbiol Biotechnol 103:4363–4375

Jeong E, Shim WY, Kim JH (2014) Metabolic engineering of *Pichia pastoris* for production of hyaluronic acid with high molecular weight. J Biotechnol 185:28–36

Jing W, DeAngelis PL (2000) Dissection of the two transferase activities of the *Pasteurella multocida* hyaluronan synthase: two active sites exist in one polypeptide. Glycobiology 10: 883–889

Kang DY, Kim WS, Heo IS, Park YH, Lee S (2010) Extraction of hyaluronic acid (HA) from rooster comb and characterization using flow field-flow fractionation (FlFFF) coupled with multiangle light scattering (MALS). J Sep Sci 33:3530–3536

Kendall FE, Heidelberger M, Dawson MH (1937) A serologically inactive polysaccharide elaborated by mucoid strains of group A hemolytic Streptococcus. J Biol Chem 118:61–69

Khabarov VN, Polyak F, Boykov PY, Selyanin MA (2014) Hyaluronic acid: production, properties, application in biology and medicine, 1st edn. Wiley, Chichester

Khanmohammadi M, Khoshfetrat AB, Eskandarnezhad S, Sani NF, Ebrahimi S (2014) Sequential optimization strategy for hyaluronic acid extraction from eggshell and its partial characterization. J Ind Eng Chem 20:4371–4376

Kim TH, Kim HL, Park SY, Jang DK (2008a) Method for purifying hyaluronic acid using calcium salt and phosphate salt, or calcium phosphate salt. Patent US20080194810

Kim JK, Srinivasan P, Kim JH, Choi JI, Park HJ, Byun MW, Lee JW (2008b) Structural and antioxidant properties of gamma irradiated hyaluronic acid. Food Chem 109:763–770. https://doi.org/10.1016/j.foodchem.2008.01.038

Lago G, Oruña L, Cremata JA, Pérez C, Coto G, Lauzan E, Kennedy JF (2005) Isolation, purification and characterization of hyaluronan from human umbilical cord residues. Carbohydr Polym 62:321–326

Laurent TC, Fraser JRE (1992) Hyaluronan 1. FASEB J 6:2397–2404

Laurent TC, Gergely J (1955) Light scattering studies on hyaluronic acid. J Biol Chem 212:325–333

Li Y, Li G, Zhao X, Shao Y, Wu M, Ma T (2019) Regulation of hyaluronic acid molecular weight and titer by temperature in engineered *Bacillus subtilis*. 3 Biotech 9:1–9

Linker A, Meyer K (1954) Production of unsaturated uronides by bacterial hyaluronidases. Nature 174:1192–1194

Liu L, Liu Y, Li J, Du G, Chen J (2011) Microbial production of hyaluronic acid: current state, challenges, and perspectives. Microb Cell Fact 10:1–9

Liu J, Wang Y, Li Z, Ren Y, Zhao Y, Zhao G (2018) Efficient production of high-molecular-weight hyaluronic acid with a two-stage fermentation. RSC Adv 8:36167–36171

Lu JF, Zhu Y, Sun HL, Liang S, Leng FF, Li HY (2016) Highly efficient production of hyaluronic acid by *Streptococcus zooepidemicus* R42 derived from heterologous expression of bacterial haemoglobin and mutant selection. Lett Appl Microbiol 62:316–322

Mao Z, Chen RR (2007) Recombinant synthesis of hyaluronan by *Agrobacterium* sp. Biotechnol Prog 23:1038–1042

Mao Z, Shin H-D, Chen R (2009) A recombinant *E. coli* bioprocess for hyaluronan synthesis. Appl Microbiol Biotechnol 84:63–69

Matsumura G, de Salegui M, Herp A, Pigman W (1963) The preparation of hyaluronic acid from bovine synovial fluid. Biochim Biophys Acta 69:574–576

Meyer K, Palmer JW (1934) The polysaccharide of the vitreous humor. J Biol Chem 107:629–634

Mizuno H, Iso N, Saito T, Ogawa H, Sawairi H, Saito M (1991) Characterization of hyaluronic acid of yellowfin tuna eyeball. Nippon Suisan Gakkaishi 57:517–519

Murado MA, Montemayor MI, Cabo ML, Vázquez JA, González MP (2012) Optimization of extraction and purification process of hyaluronic acid from fish eyeball. Food Bioprod Process 90:491–498

Nakano T, Nakano K, Sim JS (1994) A simple rapid method to estimate hyaluronic acid concentrations in rooster comb and wattle using cellulose acetate electrophoresis. J Agric Food Chem 42:2766–2768

Nimrod A, Greenman B, Kanner D, Landsberg M, Beck Y (1988) Method of producing high molecular weight sodium hyaluronate by fermentation of *Streptococcus*. US Patent 4780414

Ogston AG, Stanier JI (1951) The dimensions of the particle of hyaluronic acid complex in sinovial fluid. Biochem J 49:585–599

Ouskova G, Spellerberg B, Prehm P (2004) Hyaluronan release from *Streptococcus pyogenes*: export by an ABC transporter. Glycobiology 14:931–938

Pourzardosht N, Rasaee MJ (2017) Improved yield of high molecular weight hyaluronic acid production in a stable strain of *Streptococcus zooepidemicus* via the elimination of the hyaluronidase-encoding gene. Mol Biotechnol 59:192–199

Rangaswamy V, Jain D (2008) An efficient process for production and purification of hyaluronic acid from *Streptococcus equi* subsp. *zooepidemicus*. Biotechnol Lett 30:493–496

Reddy J (2013) Purification and characterization of hyaluronic acid produced by *Streptococcus zooepidemicus* strain 3523-7. J Biosci Biotechnol Discov 2:173

Rehm BHA (2010) Bacterial polymers: biosynthesis, modifications and applications. Nat Rev Microbiol 8:578–592

Rehm BHA, Valla S (1997) Bacterial alginates: biosynthesis and applications. Appl Microbiol Biotechnol 48:281–288

Rosa CS, Tovar AF, Mourão P, Pereira R, Barreto P, Beirão LH (2012) Purification and characterization of hyaluronic acid from chicken combs. Cienc Rural 42:1682–1687

Sadhasivam G, Muthuvel A, Pachaiyappan A, Thangavel B (2013) Isolation and characterization of hyaluronic acid from the liver of marine stingray *Aetobatus narinari*. Int J Biol Macromol 54: 84–89

Saxena IM, Brown RM, Fevre M, Geremia RA, Henrissat B (1995) Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. J Bacteriol 177:1419–1424

Schmid J, Sieber V, Rehm B (2015) Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. Front Microbiol 6:1–24

Schulz T, Schumacher U, Prehm P (2007) Hyaluronan export by the ABC transporter MRP5 and its modulation by intracellular cGMP. J Biol Chem 282:20999–21004

Scott JE, Heatley F (1999) Hyaluronan forms specific stable tertiary structures in aqueous solution: a 13C NMR study. Proc Natl Acad Sci U S A 96:4850–4855

Scott JE, Cummings C, Brass A, Chen Y (1991) Secondary and tertiary structures of hyaluronan in aqueous solution, investigated by rotary shadowing-electron microscopy and computer simulation. Hyaluronan is a very efficient network-forming polymer. Biochem J 274:699–705

Swann DA (1968) Studies on hyaluronic acid. Biochim Biophys Acta Gen Subj 156:17–30

Sze JH, Brownlie JC, Love CA (2016) Biotechnological production of hyaluronic acid: a mini review. 3 Biotech 6:67

Thomas NK, Brown TJ (2010) ABC transporters do not contribute to extracellular translocation of hyaluronan in human breast cancer in vitro. Exp Cell Res 316:1241–1253

Tlapak-Simmons VL, Baggenstoss BA, Kumari K, Heldermon C, Weigel PH (1999) Kinetic characterization of the recombinant hyaluronan synthases from *Streptococcus pyogenes* and *Streptococcus equisimilis*. J Biol Chem 274:4246–4253

Volpi N, Maccari F (2003) Purification and characterization of hyaluronic acid from the mollusc bivalve *Mytilus galloprovincialis*. Biochimie 85:619–625

Wang Y, Hu L, Huang H, Wang H, Zhang T, Chen J, Du G, Kang Z (2020) Eliminating the capsule-like layer to promote glucose uptake for hyaluronan production by engineered *Corynebacterium glutamicum*. Nat Commun 11:3120

Ward PN, Field TR, Ditcham WGF, Maguin E, Leigh JA (2001) Identification and disruption of two discrete loci encoding hyaluronic acid capsule biosynthesis genes *hasA*, *hasB*, and *hasC* in *Streptococcus uberis*. Infect Immun 69:392–399

Weigel PH (2015) Hyaluronan synthase: the mechanism of initiation at the reducing end and a pendulum model for polysaccharide translocation to the cell exterior. Int J Cell Biol 2015: 367579

Weigel PH, DeAngelis PL (2007) Hyaluronan synthases: a decade-plus of novel glycosyltransferases. J Biol Chem 282:36777–36781

Widner B, Behr R, von Dollen S, Tang M, Heu T, Sloma A, Sternberg D, DeAngelis PL, Weigel PH, Brown S (2005) Hyaluronic acid production in *Bacillus subtilis*. Appl Environ Microbiol 71:3747–3752

Won TY, Lee C, Seo SH (2008) Method for purifying hyaluronic acid. Patent WO2008062998

Woo JE, Seong HJ, Lee SY, Jang Y-S (2019) Metabolic engineering of *Escherichia coli* for the production of hyaluronic acid from glucose and galactose. Front Bioeng Biotechnol 7:351

Yoshimura T, Shibata N, Hamano Y, Yamanaka K (2015) Heterologous production of hyaluronic acid in an epsilon-poly-L-lysine producer, *Streptomyces albulus*. Appl Environ Microbiol 81: 3631–3640

Zakeri A, Rasaee MJ, Pourzardosht N (2017) Enhanced hyluronic acid production in *Streptococcus zooepidemicus* by over expressing HasA and molecular weight control with niscin and glucose. Biotechnol Rep (Amst) 16:65–70

Zhang J, Ding X, Yang L, Kong Z (2006) A serum-free medium for colony growth and hyaluronic acid production by *Streptococcus zooepidemicus* NJUST01. Appl Microbiol Biotechnol 72: 168–172

Polyhydroxyalkanoates (PHA): Microbial Synthesis of Natural Polyesters

Martin Koller, Anindya Mukherjee, Stanislav Obruca, and Manfred Zinn

Contents

M. Koller (✉)
Office of Research Management and Service, c/o Institute of Chemistry, NAWI Graz, University of Graz, Graz, Austria

ARENA - Association for Resource Efficient and Sustainable Technologies, Graz, Austria
e-mail: martin.koller@uni-graz.at

A. Mukherjee
Global Organization for PHA (GO!PHA), Amsterdam, The Netherlands

CEO, PHAXTEC, Inc., Wake Forest, NC, USA

S. Obruca
Faculty of Chemistry, Brno University of Technology, Brno, Czech Republic

M. Zinn
Institute of Life Sciences, University of Applied Sciences and Arts Western Switzerland, Sion, Switzerland

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_8

Abstract Among materials emulating fossil plastics in functionality and processability, polyhydroxyalkanoates (PHA) stand out as the sole group that is completely integrated into nature's closed loop material cycle. Being biobased, biosynthesized, biodegradable, home and industrial compostable, and biocompatible, PHA biopolymers outperform competing polymeric materials labelled with "bio" attributes claiming sustainability. PHA biopolymers exhibit versatile material characteristics mimicking fossil plastics and are the most auspicious candidates to replace established fossil plastics, resins, and fibers.

Roughly 40% of all prokaryotic strains accumulate PHA biopolymers, and more than 150 different hydroxyalkanoate (HA) monomers that make up PHA biopolymers have been described, making PHA the most versatile family of biopolymers known to humankind. Commercially relevant PHA homo- and heteropolyesters include 3-hydroxybutyrate, 3-hydroxyvalerate, 3-hydroxyhexanoate, and 4-hydroxybutyrate monomers. However, numerous other PHA heteropolyesters having higher number of carbon atoms in the individual building blocks have been studied and found to have reasonably relevant functionalities. Furthermore, PHA biopolymers can also be used in numerous non-plastic applications such as being the source of optically pure chemicals or biologically active substances. Therefore, we currently stand only at the beginning of PHA discovery and industrialization.

While reducing plastic pollution, greenhouse gas emissions and climate change are the current drivers for intensified exploration and commercialization of PHA biopolymers, they have a far greater role to play than just being exploited due to their renewable nature and intrinsic biodegradability which have also been reviewed here.

As consumers, brand owners, converters, waste managers, and policy makers conceive and acknowledge the beneficial attributes of PHA biopolymers in this current wave of commercialization, the next wave consisting of PHA biopolymers for durable applications would irreversibly reduce fossil plastics use helping us to make a quantum leap in reducing plastic pollution, greenhouse gas emissions, and climate change-related to fossil plastics use.

1 Introduction

By affecting our environment and the quality of our life, pollution from plastics has become an urgent issue of our times, ranking right behind Climate Change. Innovation in recycling—mechanical or chemical—continues to make headway; however, recycling rates have remained at around only 9% of all plastics discarded for over a decade, thus clearly showing that current policies and pure market economics are insufficient in increasing recycling rates. It is obvious that mitigating plastic pollution would require a paradigm shift; simply redesigning our use of plastics to reduce consumption or dramatically improving recovery and recycling is insufficient. It will also require significantly expanding waste plastics collection and processing. One solution is using materials that mimic plastics in processing, applicability, and functionality but, at the same time, having environmentally friendly end-of-life options. Therefore, materials are needed that can not only be recycled and reused; they must also have the inherent ability to biodegrade in our environment and not harm our ecosystem if they leak into soil or the aquatic environment (Koller 2019a).

One such material that is ubiquitous in our lives is cellulose, a biopolymer, which is extracted from wood and cotton and is used as viscose, rayon, paper, and other important articles. While the processes to arrive at such utilitarian materials are complex and sometimes involve toxic chemicals, humankind has learned to control them well. Cellulose leaking into the environment biodegrades readily. It also composts, can be recycled, and is renewable. Exactly the same way, polyhydroxyalkanoates (PHA), a versatile class of natural and microbially produced biopolymers (Rehm 2010), act like cellulose in the environment; they biodegrade, they compost, they can be recycled, they are biosynthesized, and they are renewable. However, in contrast to cellulose and other well-known biopolymers, they also have plastic-like functions and processing features, many of which we have learned to appreciate and rely on. Moreover, if erroneously entering thermic recycling systems (incineration), PHA do not harm the environment, in contrast to, e.g., poly(vinyl chloride) (PVC) (Koller and Mukherjee 2020).

2 PHA: General Aspects

2.1 Early Discovery of PHA

PHA's discovery and commercialization has a history similar to many common substances we derive today from fossil carbon, although the protocols for their biosynthesis via fermentation were developed already a century ago. For example, the famous *Clostridia*-based Weizmann process for production of acetone, isopropanol, and 1-butanol starting from renewable feedstocks was used on large scale already during World War I. The discovery of PHA even dates back as far as

1888, when the renowned Dutch microbiologist Martinus Willem Beijerinck observed light-refractive inclusions inside microorganisms that differed from endospores under a light microscope. However, unable to imagine what waves these inclusion bodies will make in the scientific community many decades later, he did not further examine them. Only decades later, Beijerinck's experiments were repeated, and the inclusions were identified as PHA (reviewed by Chowdhury 1963). This was almost 40 years after Beijerinck's observations, when in 1923 Maurice Lemoigne from the Institute Pasteur, the most frequently cited pioneer of PHA, described the excretion of 3-hydroxybutyrate (3HB) by resting *Bacillus* "M" cells under anaerobic conditions (Lemoigne 1923). In 1925, Lemoigne performed quantitative studies of this product and theorized 3HB to be a degradation product of an intracellular polymer (Lemoigne 1925). Indeed, his assumption was correct, and in 1927, he succeeded in isolating solid poly(3-hydroxybutyrate) (P3HB) by chloroform extraction and elucidated its empirical chemical formula $(C_4H_6O_2)_n$ (Lemoigne 1927).

Nevertheless, PHA, despite their beneficial properties, went off the radar of science due to the rapid development of fossil plastics. Interest in PHA started rising again in the 1950s with the elucidation of their predominant biological role as energy and carbon storage in cells of *Bacillus megaterium* and *Bacillus cereus*, their intracellular circularity and their thermoplastic-like properties (Macrae and Wilkinson 1958). An original citation from the abstract of their article postulated for the first time: "The evidence that poly-β-hydroxybutyrate is a reserve carbon and energy source is discussed." In this key study, the authors claimed for the first time the fundamentals of PHA biosynthesis and intracellular degradation:

1. "It (PHA; note by authors) should best be biosynthesized in an environment containing excess external carbon and energy source."
2. "The supposed carbon and/or energy source (PHA; note by authors) should be capable of being broken down in the absence of an external carbon and energy source."
3. "The products of breakdown of the storage compound (PHA; note by authors) is capable of being used as a source of carbon and/or energy to prevent cell autolysis and death."

The recognition that PHA biopolyesters are "materials with plastics-like properties from living organisms" was again almost dismissed as being of merely academic relevance until the work of Prof. Gerhart Braunegg in the 1970s considered one of the pioneers of the modern PHA efforts (and the doctoral advisor of one of the authors of this review) and many other innovative researchers at that time. Prof. Braunegg carried on his pioneering work despite being told as a young scientist at a scientific conference that "plastics are produced from petroleum, not by bacteria" (personal communication G. Braunegg). However, his conviction and labor were rewarded when Imperial Chemical Industries (ICI), UK, in 1989 introduced the first commercial PHA biopolymer and Unilever started making bottles from it to fill in shampoo. Industrial economics, however, won the battle at that time. Given the low price of fossil fuels and plastics thereof in comparison to the costs necessary for

producing PHA biopolymers, ICI stopped their PHA production. It was not until the 1990s when the level of plastic pollution and the beneficial properties of PHA could no longer be ignored, thus triggering companies like PHB Industrial S.A., Monsanto, Metabolix, Proctor and Gamble, or Kaneka to start industrial PHA biopolymer research and production.

The first PHA discovered was the polymer of the chiral repeat unit (monomer) (*R*)-3-hydroxybutyrate (3HB). This poly(3-hydroxybutyrate) (abbreviated in the literature as PHB or P(3HB)) is the homopolyoxoester of the monomer 3HB. P (3HB) is a natural biopolymer just like nucleic acids, proteins, cellulose, starch, or chitin. The 3HB monomer also occurs in the human metabolism as one of the ketone bodies; also, its di- and trimers are rapidly metabolized in vivo and can therefore be used as a carbon and energy source for severely injured persons suffering from hyperglycemia and could replace glucose infusions especially in patients suffering from diabetes (Tasaki et al. 1999; Chen and Wu 2005). In animals and humans, 3HB is not only an intermediate metabolite but also plays an important regulatory role by influencing gene expression, lipid metabolism, neuronal functions, and the overall metabolic rate as recently reviewed by Mierziak et al. (2021). Oligomeric P(3HB) has been found also in many organisms (Seebach et al. 1994) where they can form complexes with polyphosphate and act as a calcium transport channel in membranes (Reusch et al. 1995; Seebach et al. 1996).

2.2 PHA Are Biosynthesized

2.2.1 "Biopolymer" versus "Bioplastic"

We used the term "biopolymer" above which is defined as any polymeric substance that is formed inside a living organism, hence, "biosynthesized". Such biosynthesized materials are typically also biodegradable. Polymers that are both biobased and can biodegrade are defined by Merriam-Webster as a "bioplastic," thereby implying that biopolymers and bioplastics are one and the same. Webster also defines plastics as materials of chemoorganic synthetic origin that can be formed into different shapes by application of heat. Therefore, in the narrowest sense, the word "bioplastics" is wrongly defined, since plastics are synthetically produced, and they do not biodegrade in nature. The expression "bioplastics" was first used in 1989; however, since then that definition has degenerated into including all materials that are either biodegradable or renewable or both. The Swiss Academy of Technical Sciences elaborated a factsheet on bioplastics illustrating the problematic terms used in industry (online resource 1 n.d.). The authors of present chapter, therefore, have restricted themselves to the terms "biopolymers" (not "bioplastics") and "plastics," respectively, to differentiate between naturally synthesized and biodegradable polymers and synthetic fossil polymers that do not biodegrade.

For illustration, materials like polylactic acid (PLA) and poly(butylene succinate) (PBS), and sometimes poly(ε-caprolactone) (PCL) and poly(butylene adipate terephthalate) (PBAT) are also described as "bioplastics," while none of these are either

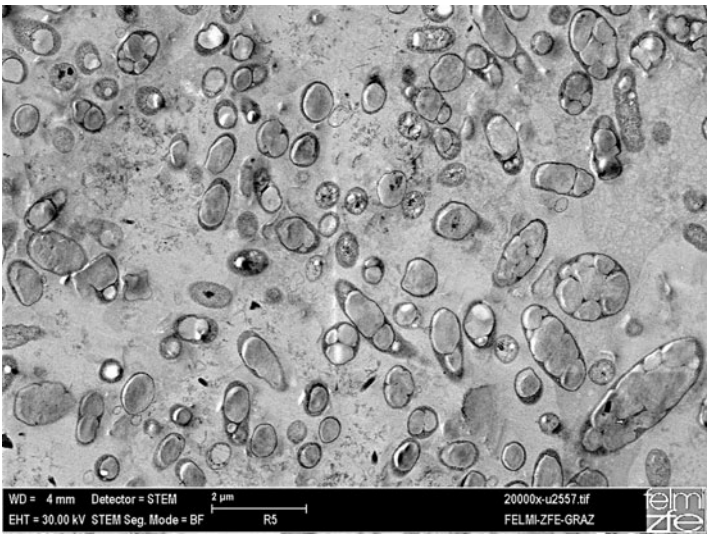


Fig. 1 PHA carbonosomes observed in *Cupriavidus necator* cells as bright inclusions. Magnification: 20,000×. (STEM picture prepared and provided with courtesy by E. Ingolić, FELMI-ZFE Graz)

biopolymers or bioplastics *sensu stricto* (Koller and Mukherjee 2020). This degeneration of definitions has created confusion among consumers, in the industry, and in academia. For example, PLA is synthetically polymerized from lactic acid, and lactic acid is a natural product. Nevertheless, PLA is not a natural material (it is not biosynthesized), nor is it a biopolymer (it does not occur in nature). PCL and PBAT are synthetic polymers (production based on fossil feedstocks), while PCL biodegrades in nature, and PBAT can be composted in industrial composters. PBS is partially renewable and biodegrades in certain environments. These subtle yet important differences between these polymers, which are often thoughtlessly and carelessly summarized as "bioplastics," and the human tendency to generalize have given rise to ambiguous definitions such as "bioplastics." Indeed, no single category that can exactly and unambiguously define PLA, PBS, PBAT, and PCL exists. Therefore, the use of the term "bioplastics" will be avoided in this chapter.

Having set the confusions with definitions straight, the revision of PHA can continue. PHA are biopolymers since they are formed inside living organisms and are therefore biodegradable; this is the most important aspect in differentiating PHA biopolyesters from PLA, PCL, PBAT, PBS, etc. In the case of PHA, microbes organize not only the conversion of renewable feedstocks to monomers (hydroxyalkanoates) but also polymerize them to synthesize the PHA biopolyesters in granular organelle-like cell components, which are about 200–500 nm in size and easily visible during microscopic observation (Fig. 1). More recently, these inclusions, having a hydrophobic PHA core and a hydrophilic proteinaceous surface layer, are termed "carbonosomes" (Jendrossek 2009). The most recognized model describing their intracellular formation, the so-called scaffold model, hypothesizes

that complexes consisting of P(3HB) synthase and the protein PhaM are linked to the bacterial nucleoid, forming the PHA granule initiation complex (Jendrossek and Pfeiffer 2014).

In 1944, Lemoigne and co-workers demonstrated that the microscopically visible lipophilic inclusions already observed more than 50 years ago by Martinus Willem Beijerinck indeed contain the polymeric materials that were extractable from the cells using chloroform (Lemoigne et al. 1944). The final proof for the nature of these inclusions was provided by Weibull in 1953, when he studied isolated PHA granules prepared by dissolving the *Bacillus megaterium* cell wall using lysozyme (Weibull 1953). Inside living cells, PHA inclusions reside in neither solid nor liquid form but rather in a mobile amorphous elastomeric state and do not resemble crystalline structures seen after extraction and having gone through a heat history; this was already assumed in the 1980s (Barnard and Sanders 1988) and early 1990s (de Koning and Lemstra 1992). Actually, there had been two hypotheses explaining the presence of P(3HB)'s thermodynamically unfavorable amorphous state in native granules. The "kinetic model" predicted the low crystallization rate of polymer in small native granules, while the "plasticizer model" relied on the presence of unknown molecules in native granules, which prevented granules from crystallization (Bonthrone et al. 1992). In 2019, it was experimentally proven that both theories are correct and that the amorphous state of intracellular PHA is kinetically (a) stabilized by the low rate of crystallization in limited volume in small PHA granules, and (b) that water present in PHA granules acts as a plasticizer, thus preventing the polymer from transition into the thermodynamically favorable crystalline state (Sedlacek et al. 2019a).

As mentioned before, polymers generally consist of many repeat units, such as 3HB in the case of P(3HB). This homopolymer is a linear (polymer chains do not entangle), waterproof, microwaveable, absolutely isotactic (stereoisomeric; chains easily crystallize) polyester with thermoplastic properties and a glass transition temperature (T_g) of around 0 °C. Hence, polymer chains keep on moving and crystallizing even at room temperature. As mentioned above, P(3HB) was the first described member of the PHA family, discovered already in 1888 by Beijerinck (reviewed by Chowdhury 1963).

Catalysts driving the conversion of raw materials to PHA are exclusively intracellular biocatalysts, namely, the enzymatic toolbox catabolizing heterotrophic or autotrophic substrates to acetyl-CoA and other PHA precursors. The enzymes, converting acetyl-CoA and related compounds to PHA, viz., 3-ketothiolase (a.k.a. acetyl-CoA-acetyltransferase, β-ketoacyl-CoA thiolase, or thiolase II, EC 2.3.1.9), catalyze the condensation of the C2-compound acetyl-CoA to the C4-compound acetoacetyl-CoA and are encoded on the *phaA* gene. Other enzymes include NADPH-dependent 3-ketoacyl reductase (EC 1.1.1.36) encoded on the *phaB* gene, which enables the "pseudofermentative" reduction of acetoacetyl-CoA to (*R*)-3-hydroxybutyryl-CoA to regenerate $NADP^+$, and PHA synthases (a.k.a. PHA polymerases, EC 2.3.1.-) encoded on *phaC* genes, which are responsible for the polycondensation of (*R*)-3-hydroxybutyryl-CoA to P(3HB). Based on their substrate specificity and structure of subunits, PHA synthases were originally grouped in three types or classes, namely, type I, II, and III. Types I and III PHA

synthases are responsible for the biosynthesis of short-chain-length PHA (*scl*-PHA) like P(3HB) or poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) (P(3HB-*co*-3HV)) in most technologically relevant PHA production strains (Rehm 2003). Later, PHA synthases found in *Bacillus* sp. were recategorized as a new PHA synthase type IV due to considerable structural differences to type III synthases; such type IV synthases preferably polymerize *scl*-monomers such as 3HB and 3HV but can also utilize *mcl*-monomers as minor building blocks (Tsuge et al. 2015).

In the early 1980s, a second class of PHAs had been found by Witholt and co-workers at University of Groningen, the Netherlands (de Smet et al. 1983; Karmann et al. 2019). This so-called medium-chain-length PHA (*mcl*-PHA) consists of monomers having carbon units between 6 and 14 units. The material properties are found to range from thermoplastic elastomers to fluidoplastics but with the suitable functional fatty acid as growth substrate; also thermosets and rubber-like polymers had been produced with wild-type strains mainly belonging to *Pseudomonas* sp. Interestingly, these *mcl*-PHA-accumulating strains were also isolated from the environment (e.g., *Pseudomonas guezennei*, Simon-Colin et al. 2008) and appeared to be part of microbial mats as it is the case of P(3HB) producing bacteria as shown by the pioneering research by D.C. White at the University of Tennessee, USA (Findlay et al. 1990).

Organisms tested on lab or pilot scale for PHA production on gaseous substrates (CO_2, CO, CH_4) possess the enzymatic toolbox to convert these gases to organic building blocks, instead of or in addition to the enzymes needed for conversion of sugars (enzymes involved in glycolysis (Emden-Meyerhof-Parnas pathway), the KDPG pathway (Entner-Doudoroff pathway), or hydrolases for conversion of di- and polysaccharides to convertible monosaccharides, fats (lipases, enzymes needed for β-oxidation), etc.).

Intracellular PHA depolymerases (encoded on *phaZ* genes), in turn, degrade the stored PHA to carbon and energy substrates under conditions of starvation, thus serving for the "cyclic nature of PHA metabolism" (Doi et al. 1990). In addition, extracellular PHA depolymerases (EC 3.1.1.75 and EC 3.1.1.76), belonging to the group of carboxylesterases, are excreted by many bacteria or fungi, are less specific than their intracellular counterparts, and serve for biodegradation of spent PHA articles when composted. Hence, in the environment where microorganisms secrete PHA depolymerases, those microorganisms are basically predating on PHA originated from lysed cells (reviewed by Jendrossek and Handrick 2002).

Figure 2 visualizes the principle metabolism of PHA biosynthesis starting from model substrates, indicating the reactions catalyzed by enzymes running PHA biosynthesis and intracellular degradation.

2.3 PHA Play Multifaceted Roles in Nature

PHA are important constituents of the organisms that produce them. These carbonosomes (Jendrossek 2009) where the PHA are stored primarily act as carbon

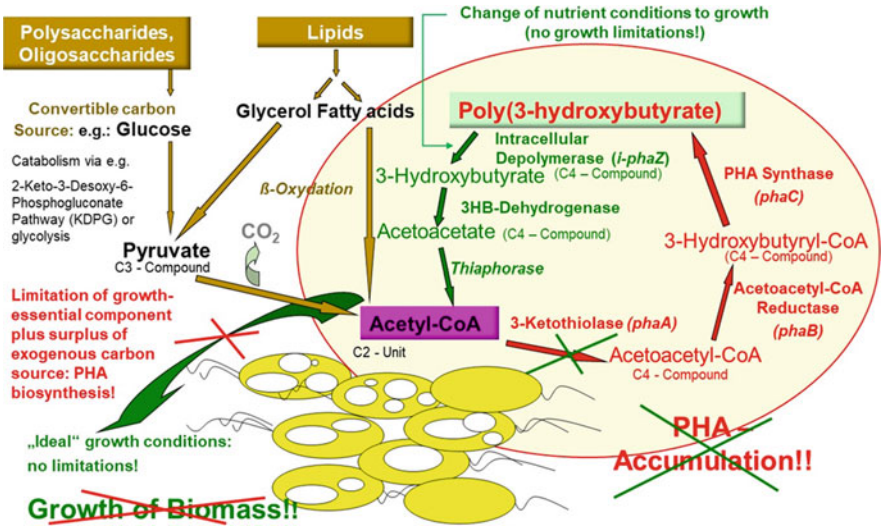


Fig. 2 Schematic of intracellular P(3HB) homopolyester biosynthesis and degradation. *Brown arrows:* catabolism of substrates to acetyl-CoA, the central metabolite for both biomass formation and P(3HB) biosynthesis. *Green arrows* indicate balanced microbial growth under not nutritionally limited (balanced) cultivation conditions and intracellular mobilization of PHA reserves, respectively; *red crosses* symbolize stop of biomass growth by the onset of limitation of growth-essential nutrient. *Red arrows* indicate P(3HB) biosynthesis under nutritional stress (e.g., limitation of nitrogen or phosphate source); *green crosses*, in turn, symbolize inhibition of P(3HB) formation when nutrients are balanced again

and energy storage compounds to support the cells during periods of starvation (Macrae and Wilkinson 1958). Moreover, in many PHA-producing sporulating organisms like *Bacillus cereus* or *Bacillus megaterium*, the onset of endospore formation depends on intracellular carbon and energy supply accomplished by the degradation of the PHA pool (Sadykov et al. 2017; Valappil et al., 2008), while in *Azotobacter* sp., PHA serve as an electron sink to regulate local oxygen levels (Dawes and Senior 1973).

In addition, PHA granules also act as important protectants by supporting the cells when they are exposed to challenging stress conditions, such as sudden osmotic imbalance, excessive UV radiation, oxidative, or thermal (both heat and frost) stress. During these challenging events, cells harboring PHA have a survival advantage in comparison to PHA-free cells, as it was demonstrated by exerting both PHA-accumulating wild-type strains and their PHA-free mutant strains to abovementioned stress factors. This complex interrelation of PHA to the SOS response of cells to challenging conditions has only recently been expansively studied, and review articles on the multifaceted roles of PHA in the survival of microorganisms, which goes far beyond avoiding starvation during lack of exogenous carbon supply, have been written during the last few years (Obruca et al. 2018, 2020, 2021; Müller-Santos et al. 2020).

In addition to their roles as carbon and energy stocks and stress protectants, the biological functions of PHA encompass their importance as an electron source to drive reductive sulfate respiration in purple and green sulfur bacteria, such as *Desulfovibrio*, *Desulfobacterium*, *Desulfococcus*, *Chromatium*, or *Amoebobacter* (today: *Thiocapsa*), which enables their anoxygenic photosynthesis (Hai et al. 2004). Moreover, in nitrogen-fixing (diazotrophic) microbes, PHA biosynthesis serves for maintaining energy production and NAD^+ recovery; hence, they perform a "pseudofermentative" reaction for the necessary oxidation of redox equivalents (Encarnación et al. 2002).

Interestingly, PHA also play pivotal roles in several symbiotic or syntrophic relationships between bacteria and other organisms. For example, PHA metabolism plays an important role between different bacterial species, as demonstrated for microbial mats by the syntrophic action of heterotrophic PHA-producing species (e.g., *Labrenzia* sp., *Halomonas* sp.), which, during the day, convert organic compounds excreted by photosynthetic cyanobacteria into PHA. These cyanobacteria, in turn, generate these organic compounds at night by converting CO_2 generated by oxidative PHA degradation in the heterotrophic bacteria, when no exogenous carbon sources are available for them. Hence, each species maintains the metabolism of the other running at different times of the day and night (Villanueva et al. 2007). Moreover, the role of PHA metabolism of the betaproteobacterium *Herbaspirillum seropedicae* in promoting the growth of the millet *Setaria viridis* was described, showing the significance of PHA biosynthesis in symbiosis between bacteria and plants (Alves et al. 2019). In addition, in 2013, Kim et al. demonstrated that PHA biosynthesis and accumulation in microbial symbiont cells are required for normal symbiotic association with and beneficial effects for the insects that symbiotically host the microbes due to the environmental conditions within the host, which are, to a certain extent, stressful for the bacteria (Kim et al. 2013).

Besides high-molecular-mass PHA, it was reviewed by Jendrossek and Pfeiffer (2014) that oligo-P(3HB), consisting of about 100–200 3HB monomers, as well as P(3HB) consisting of about 30 3HB units covalently bound to proteins ("conjungated P(3HB)", cPHB) presumably exist in all organisms, are supposed to have functions different to storage P(3HB), and might be of vital importance, but their biological functions are not fully elucidated yet (Table 1).

2.4 PHA Production Strains: Bacteria and Archaea as Cell Factories for Biopolymer Production

Natural PHA production strains are found in environments as diverse as soil in the temperate zone as well as polar regions and desserts, the rhizosphere of plants, marine water and their sediments, salt lakes, freshwater lakes, estuaries, activated sludge, the vicinity of submarine volcanoes, and many more. Many of them are mesophilic organisms concerning their optimal pH-value, temperature, halophilicity,

Table 1 The multifaceted role of high-molecular-mass PHA for microbes producing them

Role	Reference
Carbon and energy storage	Macrae and Wilkinson (1958)
Electron sink for intracellular oxygen regulation	Dawes and Senior (1973)
Electron source for sulfate respiration in purple and green sulfur bacteria	Hai et al. (2004)
Substrate for the onset of sporulation in Bacilli	Valappil et al. (2008)
Protectant against hypertonic salinity	Obruca et al. (2017)
Protectant against sudden osmotic imbalances	Sedlacek et al. (2019b)
Oxidation protectant	Obruca et al. (2010)
Heat protectant	Obruca et al. (2016a)
Freeze/thawing protectant	Obruca et al. (2016b)
UV protectant	Slaninova et al. (2018)
Enabling symbiosis between bacteria and other microbes	Villanueva et al. (2007)
Enabling symbiosis between bacteria and plants	Alves et al. (2019)
Enabling symbiosis between bacteria and insects	Kim et al. (2013)
"Pseudofermentative" effects by maintaining energy production and NAD^+ recovery of nitrogen-fixing bacteria	Encarnación et al. (2002)

and substrate concentration. The currently most used organisms for industrial-scale PHA production are, as far as disclosed by the manufacturers, mesophilic soil bacteria, namely, *Cupriavidus necator*, *Azohydromonas australica* (former *Alcaligenes latus*), *Pseudomonas* sp., *Aeromonas* sp., or *Paraburkholderia sacchari* (formerly *Burkholderia sacchari*). Only recently, transformant strains of the halophile *Halomonas* sp. are used by new companies established in PR China to produce P(3HB) and its copolyesters with 3-hydroxyvalerate (3HV), 4-hydroxybutyrate (4HB), and 3-hydroxyhexanoate (3HHx). However, the number of excellent wild-type PHA production strains well adapted to heavily challenging environmental conditions, such as salinity exceeding 5 M NaCl (Koller 2019b), temperatures around the freezing point of water (*Pseudomonas extremaustralis* thrives well at temperatures between 4 and 37 °C (López et al. 2009)) or above 50 °C (*Schlegelella thermodepolymerans* has its temperature optimum at 55 °C (Kourilova et al. 2020); *Thermus thermophilus* showed PHA biosynthesis at 75 °C (Pantazaki et al. 2003)), and pH-values below 4 (acidophiles; *Bacillus cereus* accumulates increased amounts of PHA at pH 4.5–5.8 (Valappil et al. 2008)) and above 11 (alkaliphiles; for *Spirulina platensis*, the optimum pH-value for PHA production is 9–11 (Jau et al. 2005)), is steadily growing in the scientific literature (reviewed by Karthikeyan and Mehariya 2021).

From the microbiological point of view, microbes capable of PHA biosynthesis are found in both prokaryotic domains *Bacteria* and *Archaea*. Among bacteria, PHA biosynthesis is described for both Gram-negative and Gram-positive strains; for

obligate chemoheterotrophs; for diverse oxygenic or anoxygenic photoautotrophic species like members of the cyanobacteria (prominent genera: *Spirulina*, *Nostoc*, *Synechocystis, Synechococcus*, *Anabaena*, *Chlorogloea*, *Aulosira*, etc.; Singh et al. 2017), sulfur bacteria, or purple non-sulfur bacteria (*Rhodospirilli*; Klask et al. 2015); and for some type II methanotrophs (e.g., *Methylomonas* sp.; Pérez et al. 2019). Regarding PHA-producing representatives from the *Archaea* domain, these organisms are typically extreme extremophiles, such as the salt-requiring members of the haloarchaea group, which encompasses several potent PHA-producing genera, such as *Haloferax*, *Halorubrum*, *Halogeometricum*, *Haloarcula*, or *Natrinema* (Koller 2019b; Karray et al. 2021). Besides the ample prokaryotes producing PHA, there are a limited number of reports in the literature describing PHA production by eukaryotic wild-type strains, such as yeasts (Abd-El-Haleem 2009) and microalgae (García et al. 2021); PHA production by genetically engineered plants (Poirier and Brumbley 2010) and other higher organisms is out of scope of the present text. Although PHA biosynthesis typically is an aerobic process, some reports are available describing anaerobic PHA accumulation by members of the genus *Clostridium*, which, in the future, might attract interest in such processes for the co-production of PHA and well-known clostridial metabolites like solvents and protein-rich biomass (Flüchter et al. 2019).

Table 2 provides an overview of most important PHA production strains; feedstocks used by them for production of the industrially important PHA types P(3HB), poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) (P(3HB-*co*-3HV)), poly (3-hydroxybutyrate-*co*-4-hydroxybutyrate) (P(3HB-*co*-4HB)), and poly (3-hydroxybutyrate-*co*-3-hydroxyhexanoate) (P(3HB-*co*-3HHx)); and natural environments of their original isolation. Species of current or near-future industrial significance for PHA production are presented in bold.

2.5 Renewable Resources as Feedstocks for PHA Production

The diverse group of natural PHA production strains described above all convert renewable raw materials into biomass and PHA. Carbohydrates are by far the most popular and suitable carbon substrates for PHA biosynthesis; monosaccharides, especially glucose and fructose, are converted by the majority of heterotrophic PHA producers, while the number of strains able to convert disaccharides (lactose, sucrose, maltose, etc.) are considerably lower. For example, lactose (present in concentrations of about 4 wt% in the surplus product whey) is readily converted without the need for prior hydrolysis to PHA by strains expressing the enzyme β-galactosidase (EC 3.2.1.23) like *Hydrogenophaga pseudoflava* (Koller et al. 2007a) or *Bacillus megaterium* (Obruca et al. 2011), while for direct sucrose conversion, the enzyme invertase (EC 3.2.1.26) is needed, present in PHA producers *Hfx. mediterranei* (Koller et al. 2012), *P. sacchari* (Miranda de Sousa Dias et al. 2017), and *Azotobacter vinelandii* (Page 1992). For direct use of polysaccharides, such as starch, the production strain needs to display α-amylase (EC 3.2.1.1) activity,

Table 2 Overview of most important PHA production strains; feedstocks used by them for production of the industrially important PHA types P(3HB), poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate) (P(3HB-*co*-3HV)), poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate) (P(3HB-*co*-4HB)), and poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate) (P(3HB-*co*-3HHx)); and natural environments of their first isolation

Type of PHA	Feedstocks	Natural microbial production strains				
		Gram-negative bacteria (selection)	Gram-positive bacteria (selection)	Cyanobacteria (selection)	Archaea (selection)	Eukaryotes
P(3HB)	Fructose, glucose, glycerol; sucrose (beet and cane; molasses, green syrup), waste cooking oil; CO_2; CH_4; biodiesel; hydrolyzed whey lactose; starch; syngas; spent coffee grounds	*Cupriavidus necator* (*former Ralstonia eutropha*); *Azohydromoas australica* (*A. lata*; former *Alcaligenes latus*); *Paraburkholderia sacchari*; *Halomonas* sp.; *Methylobacterium* sp., *Methylomonas* sp.; *Paraburkholderia fungorum*; *Hydrogenophaga pseudoflava*; *Saccharophagus degradans*; *Rhodospirillum rubrum*	*Bacillus megaterium*, *Bacillus cereus*; *Clostridium beijerinckii* ASU10	*Synechocystis* sp. PCC 6803 and MA19; *Chlorogloea fritschii*; *Gloeothece* sp. PCC 6909; *Oscillatoria limosa*; *Nostoc muscorum*; *Spirulina platensis*; *Spirulina subsalsa*; *Anabaena cylindrica*; *Aulosira fertilissima*; *Phormidium* sp. TISTR 8640	*Haloarcula marismortui*; *Haloarcula* sp. IRU1; *Halogeometricum borinquense* TN9; *Halococcus saccharolyticus*; *Haloterringena hispanica*; *Halobiforma haloterrestris*; *Natronobacterium gregoryi*; *Haloquadratum walsbyi*	*Scenedesmus* sp. (microalga)

(continued)

Table 2 (continued)

Type of PHA	Feedstocks	Natural microbial production strains				
		Gram-negative bacteria (selection)	Gram-positive bacteria (selection)	Cyanobacteria (selection)	Archaea (selection)	Eukaryotes
Location of strain isolation	Bacteria:					
	– *C. necator:* soil bacterium; first ancestor isolated as *Hydrogenomonas eutropha*; originally *isolated from a lawn soil* at the US-East Coast, probably state of NY, by researchers from the University of California (Schatz and Bovell 1952). Makkar and Casida (1987) isolated an *Alcaligenes eutrophus* strain (obtained from soil in the vicinity of University Park, PA, USA), which was later reclassified to *Cupriavidus necator* (Vandamme and Coenye 2004)					
	– *Azohydromonas australica*: isolated as "two strains (H-1 and H-4; today DSM 1122 and 1123) of the new organism . . . were isolated from soil samples collected from two widely separated sites of the Berkeley campus of the University of California" (Palleroni and Palleroni 1978); termed as *Alcaligenes latus* in this publication, today *Azomonas lata* in DSM strain collection. These strains showed high similarity with a strain isolated from a soil sample collected at Orara Reserve, Manly Vale, Australia (DSM 1124; basis for new species name *Azomonas australica*). Today, it is assumed that these organisms belong to the same species (Xie and Yokota 2005). Strains DSM 1123 and 1124 were comprehensively compared many years ago for their PHA accumulation performance on the inexpensive agricultural raw materials starch hydrolysate (maltose), green syrup (intermediate of sucrose production), and beet molasses (Braunegg et al. 1999)					
	– *Paraburkholderia sacchari* (DSM 17165): isolated as *Burkholderia sacchari* IPT 101 by Gomez et al. (1996) (researchers from IPT—Instituto de Pesquisas Tecnológicas) in Brazil from soil of sugar cane plantation					
	Halomonas sp. (ancestor of the currently utilized engineered *Halomonas bluephagenesis* strains developed by the group pf George Chen): isolated as *Halomonas* TD01 from solid and water samples collected from Aydingkol salt lake in Xinjiang, PR China (Tan et al. 2011; team of George Chen)					
	– *Bacillus megaterium*: isolated as soil bacterium with various subspecies (e.g., DSM 90, DSM 333, etc.) and from other environments (e.g., petroleum effluents); strain *Bacillus megaterium* uyuni S20 (well-described PHA producer) was originally isolated from a Bolivian salt lake (Rodriguez-Contreras et al. 2013). Origin of other important *Bacillus megaterium* ssp. (e.g., *B. megaterium* KM or Maurice Lemoigne's strain, by which PHA was discovered) remains unclear					
	Bacillus cereus: isolated from soil samples, milk or whey (Pirttijärvi et al. 1998)					
	Cyanobacteria:					
	– *Synechocystis* sp. MA19: isolated from a wet volcanic rock in Japan (Miyake et al. 1996). *Synechocystis* sp. PCC 6803: isolated from "natural samples". Origin of other *Synechocystis* sp. according to DSMZ: "sediments", "water column", etc.					
	– *Chlorogloea fritschii*: isolated from Indian soil samples (Mitra 1950)					

	– *Nostoc muscorum* ("inhabits both terrestrial and freshwater aquatic environments" (Cameron 1960))					
	– *Spirulina platensis*: natural water organism (according to DSMZ)					
	– *Spirulina subsalsa* (according to DSMZ: "Isolated from sediments from the Netherlands Schiermonnikoog")					
	Archaea (reviewed by Koller 2019b):					
	– *Haloarcula marismortui*: isolated from Dead Sea					
	– *Haloarcula* sp. IRU1: isolated from hypersaline Lake Urmia, Iran					
	– *Halogeometricum borinquense*: isolated from solar salterns of Marakkanam in Tamil Nadu, India					
	– *Halococcus saccharolyticus*: isolated from salt in Cadiz, Spain					
	– *Halobiforma haloterrestris*: isolated from the surface of hypersaline soil collected in Aswan (Egypt)					
	– *Natronobacterium gregoryi*: isolated from soda salt lake liquors from the East African Magadi soda lake					
	– *Haloquadratum walsbyi*: isolated at the Egyptian Sinai Peninsula and from saltern crystallizers in Australia and Spain					
P (3HB-*co*-3HV)	Fructose, glucose, glycerol; sucrose (beet and cane; molasses); hydrolyzed whey lactose; plus 3HV-precursor for all bacterial strains	*Cupriavidus necator*; *Azohydromonas australica*; *Halomonas* sp.; *Paraburkholderia fungorum*; *Hydrogenophaga pseudoflava*		*Oscillatoria limosa*; *Anabaena cylindrica*; *Phormidium* sp. TISTR 8640; *Aulosira fertilissima*; *Anabaena spiroides* TISTR 8075 (no 3HV-precursors needed; copolyester production from CO_2 as sole carbon source; Tarawat et al. 2020)	*Hfx. mediterranei*; *Haloarcula hispanica*; *Halogeometricum borinquense* E3; *Halobacterium noricense*; *Halococcus dombrowskii*; *Hcc. hamelinensis*; *Hcc. morrhuae*; *Hcc. qingdaonensis*; *Hcc. salifodinae*; *Halorubrum chaviator*; *Hrr. coriense*; *Natrinema ajinwuensis* (=*altunense*); *Nnm. palladium*; *Natronococcus occultus*	*Rhodotorula minuta* Y4 (yeast)
	Note: Archaea listed in this row: copolyester production without 3HV-precursor addition on biodiesel; hydrolyzed whey lactose; extruded rice bran; stillage; hydrolyzed sugarcane bagasse, hydrolyzed cassava bagasse					

(continued)

Table 2 (continued)

Type of PHA	Feedstocks	Natural microbial production strains				
		Gram-negative bacteria (selection)	Gram-positive bacteria (selection)	Cyanobacteria (selection)	Archaea (selection)	Eukaryotes
Location of strain isolation	– *Paraburkholderia fungorum* (DSM 1749): before known as *Pseudomonas hydrogenovora*; isolated as "strain 9–5" by Kodama et al. (1975) at Department of Agricultural Chemistry, University of Tokyo from "samples of soil, river water or sewage" (1 of 100 strains isolated from about 200 natural samples). Other subspecies: *P. fungorum* DSM 17061 was isolated from fungus *Phanerochaete chrysosporium* (according to DSMZ; country unknown)					
	Archaea (reviewed by Koller 2019b):					
	– *Haloferax mediterranei*: isolated from evaporation ponds of solar salterns at the Spanish coast near Alicante as "strain Q4" by Rodríguez-Valera et al. (1980)					
	– *Halogeometricum borinquense* E3: isolated from solar salterns of Marakkanam in Tamil Nadu, India					
	Halobacterium noricense: isolated from a bore core of an Austrian Permian salt deposit					
	– *Halococcus dombrowskii*: isolated from dry rock salt from Austrian alpine salt mine					
	– *Hcc. hamelinensis:* isolated from stromatolites from the Hamelin Pool in the Australian Shark Bay					
	– *Hcc. morrhuae*: Dead Sea isolate					
	– *Hcc. qingdaonensis:* isolated from a crude sea-salt sample collected near Qingdao, PR China					
	– *Hcc. salifodinae*: isolated from alpine rock salt in Austria					
	– *Halorubrum chaviator*: isolated from sea salt in Baja California, Mexico, Western Australia, and Greece					
	– *Hrr. coriense*: isolated from the Dead Sea; *Natrinema ajinwuensis* (=*altunense*): isolated from salt production pans, India					
	– *Natrinema palladium*: isolated from Kayacik saltern, Turkey					
	– *Ncc. occultus*: isolated from the Lake Magadi, Kenya					
P (3HB-*co*-4HB)	Sucrose + 4HB-precursor; glucose + precursor; glycerol + precursor	*Cupriavidus necator*; *Paraburkholderia sacchari* (former *Burkholderia sacchari*); *Delftia acidovorans*; *Haloferax mediterranei*; *Aneurinibacillus* sp. H1	–	–	*Hfx. mediterranei*	–

Location of strain isolation	– *Delftia acidovorans*: isolated by Y. Doi's group as *Comamonas acidovorans* DS-17 from "activated sludge from the municipal wastewater treatment plant of Narashino-shi, Japan" (Saito and Doi 1994). Renamed to *D. acidovorans* by Wen et al. (1999)					
	– *Aneurinibacillus* sp. H1: isolated by the team of one of the co-authors of this chapter (S.O.) from compost in Brno, Czech Republic, by using an osmoselection protocol (Pernicova et al. 2020)					
P (3HB-*co*-3HHx)	Inexpensive vegetable oils (canola or soy); waste cooking oil	*Aeromonas caviae*; *Aeromonas hydrophila*	–	–	–	–
Location of strain isolation	– *Aeromonas caviae*: isolated from epizootic of young guinea pigs (DSM 7323), USA; influent of a municipal wastewater treatment plant in Bangkok (DSM 29415), minced meat (DSM 30025; country unknown), or used oil emulsions (DSM 30188; according to DSMZ, country unknown). *Aeromonas hydrophila*: isolated from surface water (according to DSMZ; country unknown; DSM 30016), tin of milk with a fishy odor (DSM 30187; country unknown) (all information from DSMZ strain collection). The PHA production strain of Lee et al. (2000a) was isolated from "raw sewage samples." Both *Aeromonas hydrophila* and *A. caviae* were also isolated from grocery's products ("leave parts of vegetables"; Callister and Agger 1987)					

which is the case, e.g., in *Hfx. mediterranei* (Huang et al. 2006). Reports on direct cellulose utilization for PHA biosynthesis based on the catalytic action of the endohydrolase cellulase (EC 3.2.1.4) activity are scarce in literature. However, similar to starch, cellulose can be hydrolyzed during upstream processing to its monomer (glucose), which is then converted by most industrially relevant PHA producers.

Apart from easily water-miscible hydrophilic heterotrophic substrates, many lipophilic substrates are also converted by microbes for PHA biosynthesis (Walsh et al. 2015). This encompasses long-chain fatty acids (Chee et al. 2010), triacylglycerides (Basnett et al. 2018), or fatty acid methyl esters (Koller and Braunegg 2015a). In the case of such lipophilic substrates, we often observe higher substrate-to-PHA yields than when starting from carbohydrates. This is caused by the stepwise degradation of the fatty acid chains to a pool of acetyl-CoA during the β-oxidation pathway. Acetyl-CoA, in turn, acts as the direct precursor of PHA under conditions favoring PHA biosynthesis (typically ample supply with an exogenous carbon source, combined with limited availability of other growth-essential nutrients, such as nitrogen or phosphate). Technologically, lipophilic substrates pose challenges for the cultivation process concerning the necessary fine distribution of the substrates in aqueous cultivation media. Substrate droplets must be as tiny as possible to realize a maximum lipid-to-water surface area. Often, emulsifiers are needed for this purpose (Muhr et al. 2013). However, some organisms like *Pseudomonas putida*, which are excellent PHA producers from lipophilic substrates, produce emulsifiers (rhamnolipids) that improve the availability of the lipophilic substrates during the cultivation (Bonilla et al. 2005).

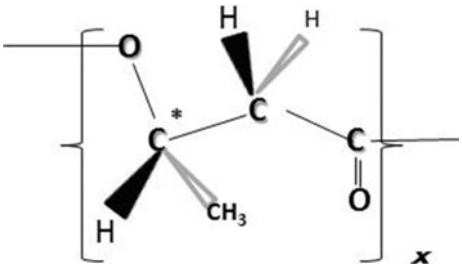
A raw material of increasing interest for PHA production is methane, the major component of biogas and of natural gas (Khosravi-Darani et al. 2019). Biogas, in turn, can be generated by anaerobic decomposition of various organic waste and surplus materials, such as kitchen and gastro-waste, agro-waste, waste of animal-rearing (Achinas et al. 2017), and even spent items made of fossil plastics and biopolymers like PHA, which ultimately closes the carbon cycle of PHA (Vu et al. 2020). Biogas formation provides a biotechnological substrate of high homogeneity and purity in comparison to the direct use of heterogenic organic waste streams for PHA production, which are generally characterized by a fluctuation in their chemical composition (Saratale et al. 2021). Constant composition of the feedstock, in turn, facilitates the reproduction of PHA quality in subsequent cultivation setups. In principle, microbes are able to convert methane to PHA need, in addition to the PHA synthesis enzymes and the enzymatic machinery of the serine pathway, which allows the assimilation of methane to cellular molecules (Strong et al. 2016).

Another fascinating approach is the photoautotrophic conversion of CO_2 via the Calvin cycle to biomass and PHA by cyanobacteria and eukaryotic microalgae. In addition to PHA, cyanobacteria convert CO_2 to produce other marketable bioproducts, such as bioactives or phycobiliproteins (a class of pigments used as food colorants, immunofluorescence markers, etc.), which makes these organisms versatile cellular factories. Here, robust cyanobacterial wild-type strains are available, which can directly utilize CO_2 from abundantly available industrial effluent gas. This strategy combines mitigation of a greenhouse gas (CO_2) with the

generation of value-added bioproducts. If one develops this thought further, it is easily imaginable to use effluent gas directly stemming from fossil plastic incineration for PHA bioplastic manufacturing, thus turning a petrochemical product into a biopolymer (Kamravamanesh et al. 2018).

Another waste-derived feedstock for PHA production of tomorrow is syngas, a mixture of CO_2, CO, and H_2 (Amstutz and Zinn 2020). Technically, syngas can be generated by gasification of a variety of organic waste and surplus materials, such as switchgrass, lignocelluloses, and other organic materials. *Rhodospirillum rubrum*, a non-sulfur purple bacterium, is currently best described as a producer of P(3HB) homopolyester from syngas. Such organisms not only assimilate CO_2 as carbon feedstock but also convert CO, which is toxic for most organisms even in low concentrations via the so-called water-shift reaction to CO_2; the entire metabolic pathway needed is the so-called Wood-Ljungdahl pathway (reductive acetyl-CoA pathway), which serves microbes to generate energy and in autotrophic carbon assimilation. In the methyl branch of this pathway, CO_2 is reduced to formic acid, which, under the consumption of ATP, generates a methyl molecule bound on the corrinoid iron-sulfur protein. H_2 present in syngas serves as hydrogen donor, while CO_2 acts as electron acceptor and biomass building block. The reaction between CO_2 (converted to a methyl group in the methyl branch of the pathway) and CO (generated in the carbonyl branch or consumed directly from syngas), catalyzed by the acetyl-CoA synthase complex, generates acetyl-CoA, which finally acts as a precursor for P(3HB) biosynthesis (Karmann et al. 2019). Similar to biogas, syngas is a homogenous carbon source, which facilitates the reproducible biosynthesis of homogenous PHA. However, safety precautions need to be observed due to the risks that come along with working with syngas. This challenge was addressed by Karmann et al. (2017) who developed a process analytical technology (PAT) platform for safe handling of syngas for biotechnological purposes in general and explicitly for PHA production.

3 P(3HB) Homopolyester



3.1 P(3HB)'s History

As disclosed above, P(3HB) is the homopolyoxoester of the monomer 3HB. P(3HB) is a polymer with thermoplastic properties and a glass transition temperature (T_g) around 0 °C (above that temperature, polymer chains keep on moving and crystallizing, even at room temperature). P(3HB) is for sure the best-studied and most frequently referred PHA biopolymer from this family of natural biopolymers. Despite being easily produced, P(3HB) homopolyester reveals drawbacks during melt processing. Its remarkably high degree of crystallinity, high melting point, and the overlap of its melting range and onset of degradation impede its processing in its pristine form (Modi et al. 2011). Moreover, P(3HB) is thermosensitive. When processing P(3HB), it is often observed that crystallization occurs very slowly. Biodegradation of highly crystalline P(3HB) occurs typically at lower rates compared to PHA copolyester specimens of similar molecular mass and shape when subjected to the same environmental conditions. However, this high crystallinity gives P(3HB) excellent creep resistance. It is reported that such specimens retain stable properties for 5 years when stored in a temperature range between −40 and +60 °C. They also have excellent UV stability over time, outperforming PP in this aspect (Hänggi 2018).

3.2 P(3HB) Properties

Despite being easily produced (no precursor substrates needed for incorporation of building blocks other than 3HB), P(3HB) homopolyester has drawbacks in processing it as a thermoplastic:

- Its remarkably high degree of crystallinity (X_c; about 60–70%), high melting point (T_m; about 170–180 °C), and the overlap of its melting range and the onset of degradation temperature (T_d; typically below 200 °C, often about only 180 °C) impede its processing as a sole component without additives and/or blending materials (Modi et al. 2011).
- P(3HB) is thermosensitive; at melting temperature, splitting of P(3HB) chains can easily occur via intermediate formation of hexa-bonds by the β-ester bonds and subsequent random chain scission, thus leading to reduced molecular mass. This process is catalyzed especially by the presence of small quantities of oxides of bivalent metals (Csomorová et al. 1994).
- When processing P(3HB), it is often observed that crystallization occurs too slowly, which is needed for the production of hard and creep-resistant specimens. This process can take a very long time: at room temperature as long as about 400 days, in comparison to up to 200 days for P(3HB-*co*-3HV) copolyester (Hänggi 2018). Only after this time, the properties of injection molded specimens change from ductile to tough and even to brittle. Importantly, crystallization time can be considerably decreased by reducing the size of polymer crystals, also

referred to as spherulites. Without additives, nucleation occurs spontaneously, resulting in large spherulite sizes of up to 2 mm. Barham and colleagues recognized in 1984 these large spherulite formations during the crystallization of highly pure P(3HB) and stated: "When foreign particles are added the nucleation rate is modified." (Barham et al. 1984). Indeed, crystalline boron nitride or other nucleating agents like nano-clay or natural fibers dramatically reduce the size of the crystalline spherulites to about 20 μm, and this, in turn, reduced the crystallization time to about 3 days, and the microstructure of the P3HB became more homogenous (Hänggi 2018). Barham et al. (1984) also noticed that crystallization temperature is decisive for the crystallization kinetics; for P(3HB), molten specimens should rapidly be cooled down to 90 °C, but not below, which would just freeze the material, making it unstable for months until the thermodynamic optimum is reached. In addition, friction between P(3HB) chains occurs when cooling down the material; this can be avoided by adding plasticizers (e.g., sorbitol, glycerol, etc.), which are molecules compatible with both the crystalline and the amorphous domains of the polymers.

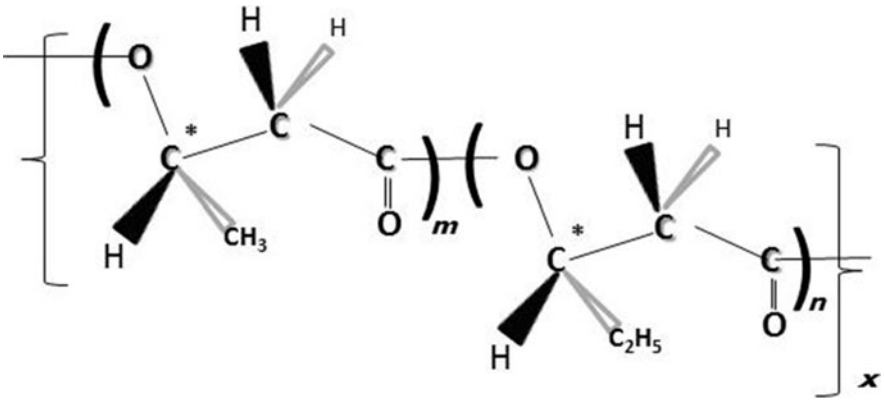
- Moreover, biodegradation of highly crystalline P(3HB), despite being in principle biodegradable and compostable, occurs typically at lower rates compared to PHA copolyester specimens of similar molecular mass and shape when subjected to the same environmental degradation conditions.
- Brittleness of P(3HB) is a typical issue, even when compounding and processing the material appropriately. This is due to the transformation from the amorphous to the crystalline state, which is accompanied by a change in density: crystals are of higher density than the amorphous domains. The polymer molecules leave the amorphous domains toward the more dense crystalline part; now, voids are generated between the spherulites, occurring as cracks through the parts. This issue can be overcome by adding plasticizers or compatible polymers to fill the voids between the crystals or to link the crystals (spherulites) with molecules (Hänggi 2018).

However, there are also notable advantages in using P(3HB):

- P(3HB) is beneficial for producing hard, creep-resistant items. This is due to the fact that P(3HB) does not stop crystallizing until the thermodynamic optimum is reached, that is, until all polymer molecules are fixed in crystals. As soon as this stable status is reached, the specimens made of P(3HB) remain there (no more amorphous domains to post-crystallize) and do not alter their properties anymore for a long time. It is reported that such specimens keep their stable properties for 5 years when stored in a wide temperature range between −40 and +60 °C (Hänggi 2018).
- P(3HB) has the beneficial property that, due to the linearity of the polymer chains unable to entangle, its melt viscosity can be adjusted as required for specific processing. In the molten state, the polymer chains are highly mobile, while when cooling down they get a degree of viscosity according to the current temperature. As a rule of thumb, a change of 10 °C alters viscosity by about 40 times. This allows the fine-tuning the viscosity directly at the processing machine according

to specific requirements: higher speed molding requires lower temperature because crystallization onset occurs earlier. In contrast, when aiming at filling the finest cavities or complex structures in a given mold, the temperature needs to be increased, and crystallization of P(3HB) only starts after having filled all mold cavities (Hänggi 2018).
– With respect to mechanical and UV stability over time, P(3HB) is reported to even outperform fossil plastics like PP (Hänggi 2018).

4 P(3HB-*co*-3HV) Copolyester, the Best Researched PHA Heteropolyester



4.1 The First Discovery of PHA Heteropolyesters

While the chemical structure of P(3HB) homopolyester was elucidated in the 1920s by Lemoigne, it wasn't until the 1970s that P(3HB-*co*-3HV) was discovered when Wallen and Davis isolated PHA from dried activated sludge by chloroform extraction and then precipitating the material by ether addition and cooling. Prepared PHA samples were treated with hot ethanol, and surprisingly the ethanol-insoluble fraction exhibited the infrared spectrum published for P(3HB) homopolyester and featured a high melting point of about 170 °C, also typical for P(3HB). In contrast, the ethanol-soluble fraction precipitated when cooling the solution; the obtained material had a T_m of only about 100 °C. Wallen and Davis correctly hypothesized that this polymer contained 3HB building blocks but had a different structure than P(3HB) homopolyester (Wallen and Davis 1972). Wallen and Rohwedder, after further research, revealed in 1974 the isolation of new microbial strains accumulating PHA from carbon sources present in effluent water. Based on GC-MS analyses, it was confirmed that the polymer contained, besides 3HB, also 3HV and some 3HHx units. This was the first unambiguous proof that PHA monomers other than 3HB existed (Wallen and Rohwedder 1974). In 1981, Morikawa and Marchessault

discovered that pyrolysis of such 3HB- and 3HV-containing microbial PHA generates unsaturated compounds (crotonic acid and pentenoic acid, respectively), which were recognized as valuable chemical synthons (Morikawa and Marchessault 1981). Using GC-MS, Findlay and White (1983) discovered a total of 11 different 3-hydroxyalkanoates (3HAs) in polymers extracted from marine sediments and 6 different 3HAs in PHA accumulated in *Bacillus megaterium*.

4.2 Biosynthesis of P(3HB-co-3HV)

The more popular PHA producers like *C. necator* or *A. lata* biosynthesize P(3HB) homopolyester from simple carbon sources like sugars, while P(3HB-*co*-3HV) copolyester biosynthesis requires feeding them co-substrates chemically related to 3HV. Typically, fatty acids with an odd number of carbon atoms, such as propionic or valeric acid, can be used as 3HV-precursors. They are metabolized by the strains to propionyl-CoA, which undergoes condensation with acetyl-CoA, yielding 3-hydroxyvaleryl-CoA (3HV-CoA). 3HV-CoA, in turn, undergoes PHA synthase-catalyzed polymerization reaction to prolong the growing PHA chains (leading to the incorporation of 3HV monomers), analogous to 3-hydroxybutyryl-CoA, which leads to 3HB production. This principle of producing P(3HB-*co*-3HV) copolyesters by feeding appropriate 3HV-precursor feeding was patented by the inventors Holmes, Wright, and Collins for ICI at the very beginning of the 1980s (EP0052459A1), where it is claimed that "The copolymers are made microbiologically: for part of the cultivation the micro-organism is under conditions of limitation of a nutrient, e.g., nitrogen source, required for growth but not polyester accumulation. For at least part of this period of growth limitation the substrate is an acid or a salt thereof that tives the comonomer units." They concluded that "propionic acid was the preferred acid".

Additionally, there exists a number of PHA producers from the haloarchaea branch, which produce P(3HB-*co*-3HV) from simple, 3HV-structurally unrelated substrates like sugars or glycerol (see Table 2; reviewed by Koller 2019a, b). Among these haloarchaea, P(3HB-*co*-3HV) production from unrelated substrates is best described for *Hfx. mediterranei*, where it was elucidated that multiple constitutively active propionyl-CoA-generating pathways are responsible for the permanent availability of the 3HV-precursor propionyl-CoA (Han et al. 2013).

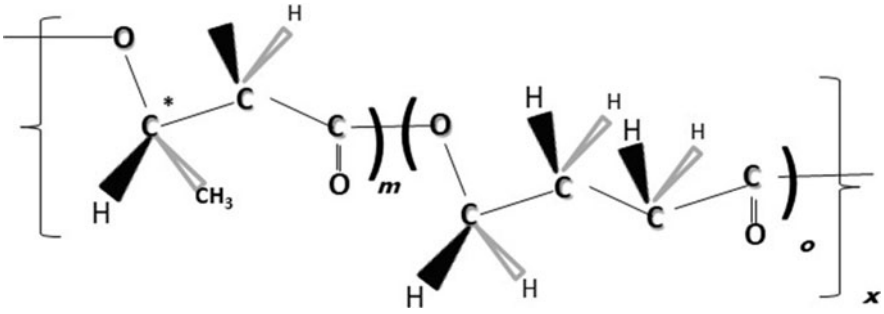
4.3 Properties of P(3HB-co-3HV)

In 1992, Luzier summarized for the first time the advantageous material features of P(3HB-*co*-3HV) copolyesters in comparison to P(3HB) homopolyester in properties like melting temperature, crystallinity, tensile strength, flexural modulus, elongation at break, and impact strength for different grades of PHA commercially produced at that time by Imperial Chemical Industries (ICI), UK. In this study, Luzier

investigated P(3HB) homopolyester and P(3HB-*co*-3HV) with 10 or 20 mol% 3HV. With increasing 3HV content, melting point T_m decreased from 170 °C for the homopolyester P(3HB) to 140 and 130 °C for the two P(3HB-*co*-3HV) copolyesters; degree of crystallinity (X_c) decreased from 80% to 60% and 35%, respectively. Tensile strength (40, 25, and 20 MPa) and flexural modulus (3.5, 1.2, and 0.8 GPa) also decreased, while elongation at break (8, 20, and 50%) and impact strength (60, 110 and 350 J/m) increased with increasing 3HV content in PHA.

The utilization of a continuous cultivation technique with dual carbon (butyric and valeric acids) and nitrogen (ammonium) limited growth, and *C. necator* produced P(3HB-*co*-3HV) with 52 mol% 3HV content resulting in a very low T_m of around 75 °C (Zinn et al. 2003). Hence, P(3HB-*co*-3HV) copolyesters exhibited better processability due to a broader processing temperature window between melting point T_m and onset of thermal decomposition T_d, which is about the same (~180 °C) for P(3HB-*co*-3HV) copolyester and its P(3HB) homopolyester counterpart. The P(3HB-*co*-3HV) copolyesters being more flexible and tougher than P(3HB) allowed them to be processed via extrusion, injection, and blow molding, thus enabled them to be used for manufacturing bottles, extruded sheets, films, fibers, and P(3HB-*co*-3HV)-coated paper (Luzier 1992).

5 P(3HB-*co*-4HB) Copolyester



4-hydroxybutyrate (4HB) is the only well-studied achiral monomer found in natural PHA. 4HB was first described as a PHA building block by the team of Yoshiharu Doi in Japan, who discovered this novel monomer in PHA samples produced by two strains of *C. necator* when being supplied with butyric acid and 4HB-precursors such as 4-hydroxybutyric acid or 4-chlorobutyric acid. Feeding these organisms with butyric acid alone resulted in accumulation of pure P(3HB) homopolyester. Depending on the ratio between butyric acid and 4-hydroxybutyric acid (or 4-chlorobutyric acid) fed to the organisms and the type of strain selected, up to 49 mol% 4HB was incorporated into the poly(3HB-*co*-4HB) copolyesters. The presence of 4HB in PHA was confirmed by NMR. A significant decrease in crystallinity at increasing 4HB fraction in PHA was reported, with the copolyester

with 49 mol% 4HB being almost completely amorphous without any detectable crystalline regions (Doi et al. 1988). Since then, as comprehensively reviewed by Utsunomia et al. (2020a), these copolyesters resemble, depending on the 4HB fraction, thermoplastic polymers at low 4HB fractions or elastomers at high 4HB fractions.

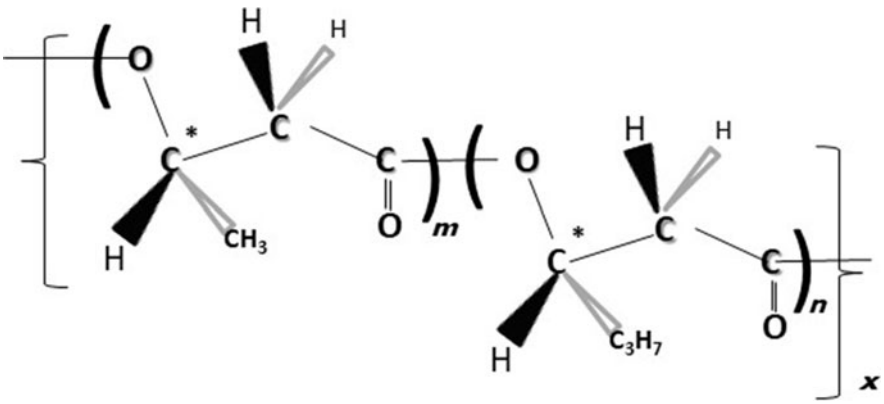
Other strain, carbon source, and 4HB-precursor combinations have been described in the literature for the production of 4HB-containing PHA heteropolyesters. Miranda de Sousa Dias et al. (2017) carried out fed-batch cultivation of the excellent sucrose converter *P. sacchari* (in the original article, the old name *Burkholderia sacchari* was used) for high-productivity (almost 2 g/(L h)) P(3HB-*co*-4HB) copolyester biosynthesis on sugar cane sucrose as the main carbon source, where γ-butyrolactone (GBL) was used as the 4HB precursor (Miranda de Sousa Dias et al. (2017)). Kucera et al. (2019) used the strain *Cupriavidus malaysiensis* USMAA2-4 (DSM 19379), isolated from the Malaysian Sg. Pinang River, for production of poly(3HB-*co*-4HB) copolyesters from different 4HB-related precursors like GBL, 1,4-butanediol, ε-caprolactone, and 1,6-hexandiol; highest 4HB shares in P(3HB-*co*-4HB) were reported for the use of GBL and 1,4-butandiol. Co-feeding of GBL or 1,4-butandiol and 3HV-related precursors (propionic or valeric acid) resulted in biosynthesis of P(3HB-*co*-3HV-*co*-4HB) terpolyesters. Another important P(3HB-*co*-4HB) copolyester producer is the bacterium *Delftia acidovorans*. This strain was used by Mothes and Ackermann (2005) to produce P(3HB-*co*-4HB) copolyesters with tailored 4HB content in a two-stage continuous chemostat cultivation setup when feeding defined mixtures of acetic acid and GBL. Surprising findings were made in 2007 by Lee et al., who noticed that increasing magnesium concentrations in cultivation medium drastically increases the 4HB in P(3HB-*co*-4HB) copolyesters when supplied with mixtures of glucose and 1,4-butandiol as the 4HB precursor. Using a glucose/precursor mix of 4/1, only 10 mol% 4HB were found in the copolyester when using a low $MgSO_4*7H_2O$ concentration of 0.05 mM. This 4HB content was boosted to 52 mol% when increasing the $MgSO_4*7H_2O$ concentration to 0.30 mM. Authors explained this effect by magnesium ions negatively impacting the uptake of the main carbon source glucose, probably due to the disturbance of the cellular transmembrane transport caused by bivalent cations like Mg^{2+}(Lee et al. 2007).

Apart from the use of bacteria as biocatalysts, Hermann-Krauss et al. (2013) described the production of P(3HB-*co*-3HV-*co*-4HB) terpolyester by the haloarchaeon *Hfx. mediterranei* when feeding inexpensive crude glycerol phase from biodiesel production and GBL as 4HB-precursor. These experiments were based on previous findings by Koller et al. (2007b), who used the same strain for P(3HB-*co*-3HV-*co*-4HB) production on hydrolyzed whey lactose, another abundantly available raw material along with GBL.

Regarding the precursors needed for incorporation of 4HB building blocks into respective P(3HB-*co*-4HB) copolyesters, one might argue that these 4HB precursors are not natural products; therefore, can P(3HB-*co*-4HB) copolyesters be deemed "natural"? Indeed, 4-hydroxybutyric acid is typically produced as its sodium salt by alkaline saponification of γ-butyrolactone, which, in turn, is a typical petrochemical

product. It is industrially produced via dehydrocyclization of 1,4-butanediol at a temperature of 180–300 °C, catalyzed by copper. This means that the three major 4HB-precursors for biosynthesis of P(3HB-*co*-4HB) copolyesters can be chemically converted into each other (1,4-butanediol → GBL → 4-hydroxybutyrate). 1,4-butandiol, in turn, is industrially produced from the alkane butane, again a fossil product, via oxidation with molecular oxygen plus tellurium oxide-containing catalysts. However, during the last decade, a biosynthetic route for production of 1,4-butanediol by recombinant bacteria was developed, in addition to the ample reports on 2,3-butandiole biosynthesis. Burgard et al. (2016) has suggested that their systems biology approach could enable the development of recombinant *Escherichia coli* (*E. coli*) expressing a cascade of heterologous enzymes, and this recombinant strain is able to convert glucose via succinyl-CoA (citric acid cycle intermediate), succinate semialdehyde, 4-hydroxybutyrate, 4-hydroxybutyryl-CoA, and 4-hydroxybutanal to 1,4-butanediol. Besides glucose, this recombinant strain was also able to convert other hexoses, pentoses, and sucrose to 1,4-butanediol. In addition, it should be emphasized that the 4HB-precursor 4-hydroxybutric acid is not exclusively produced chemically in the above-described sequence (1,4-butanediol → GBL → 4-hydroxybutyrate) but also occurs in living organisms such as in the human brain, where it acts as an important neurotransmitter (Cash 1994).

6 P(3HB-*co*-3HHx) Copolyester



Hybrid short-chain-length-medium-chain-length PHA (*scl-mcl*-PHA) copolyesters consisting of 3HB and a relatively small amount of *mcl*-PHA building blocks such as 3-hydroxyhexanoate (3HHx), 3-hydroxyoctanoate (3HO), or 3-hydroxydecanoate (3HD) were originally developed and patented by Isao Noda and co-workers. These PHA copolymers were shown to overcome problems associated with well-established P(3HB-*co*-3HV) copolyesters called "isodimorphism," whereby the 3HV units in the copolyester poly(3HB-*co*-3HV) are easily incorporated into the

crystal 3HB-lattice and vice versa, thus preventing efficient disruption of the highly crystalline matrix of P(3HB) homopolymers, therefore requiring the presence of very high percentage of the 3HV comonomer in the respective copolymer to improve their processability and subsequent commercial application. In contrast, *mcl* building blocks described above disturb the 3HB matrix even more efficiently than the achiral building block 4HB does (Noda et al. 2010). Especially poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate) (P(3HB-*co*-3HHx)) having 10–17 mol% 3HHx feature excellent flexibility, expressed by exceptionally high elongation at break of up to 850%, which outperforms commercially available BIOPOL™ P(3HB-*co*-3HV) with 20 mol% 3HV (Chen et al. 2001).

The first report on the production of such *scl-mcl*-hybrid-type P(3HB-*co*-3HHx) copolyesters was provided by Kobayashi et al. (1994), who demonstrated production of P(3HB-*co*-3HHx) by natural *Aeromonas* spp. on fats and oils. This was a scientific breakthrough at that time, as prior to this, it was set in stone that microbes produce either *scl*-PHA or *mcl*-PHA, simply depending on the type of their PHA synthase. These novel findings resulted in the first patent on such types of PHA by Shiotani and Kobayashi for Kanegafuchi Chemical Industry Co Ltd. (US Patent 5,292,860, 1994), which already claimed the use of *Aeromonas caviae* as production strain. Later, patents for production of this type of PHA were filed by Kaneka and Procter & Gamble for the inventions in this field by several researchers including Isao Noda (US Patent 5,498,692, 1996; US Patent 5,990,271, 1999), in addition to a patent on halogen-free process for recovery for such *scl-mcl*-PHA hybrid copolyesters (US Patent 5,942,597).

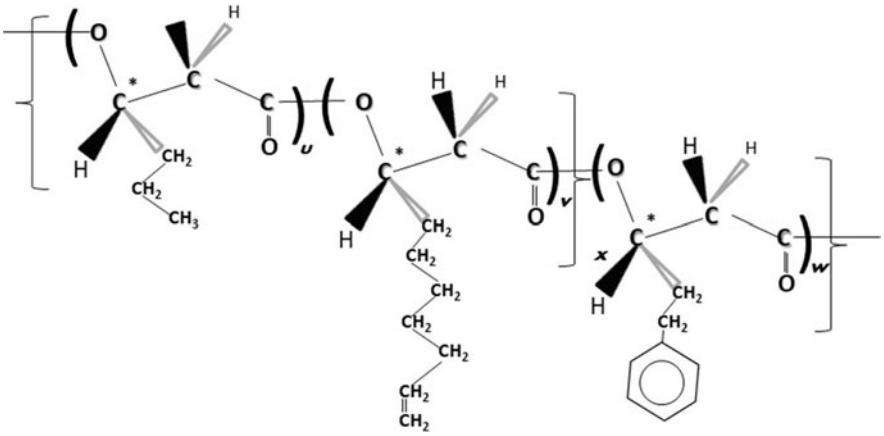
On a bioreactor scale (2 L working volume), such materials were for the first time produced by the wild-type strain *Aeromonas hydrophila* cultivated on glucose during the phase of balanced growth, followed by feeding lauric acid under phosphate limitation during the PHA accumulation phase. This strain was isolated from raw sewage samples by Lee et al. (2000a) and shown to produce P(3HB-*co*-3HHx) from long-chain fatty acids lauric acid and oleic acid in high cell density (96 g/L) cultivations; intracellular P(3HB-*co*-3HHx) fractions in cell mass amounted to 45 wt %, with 17 mol% 3HHx found in P(3HB-*co*-3HHx). Volumetric productivity for poly(3HB-*co*-3HHx) was calculated to amount to slightly more than 1 g/(L h). With glucose or sodium gluconate as substrates, P(3HB) homopolyester was produced instead of P(3HB-*co*-3HHx) copolyester by the same organisms (Lee et al. 2000a).

The industrial-scale process was described by Chen et al., who achieved a volumetric productivity of 0.54 g/(L h) for P(3HB-*co*-3HHx) with 11 mol% randomly distributed 3HHx in a two-stage batch process (glucose for growth phase, lauric acid for phosphate-limited PHA accumulation phase) on a 20 m^3 scale (working volume 10 m^3), corresponding to the production of 100 g/L biomass containing 50% P(3HB-*co*-3HHx) within 46 h of cultivation. Importantly, the authors underlined in this study that the batch cultivation mode was applied due to complications observed when running this process with *A. hydrophila* 4AK4 in fed-batch mode (Chen et al. 2001).

Besides wild-type *A. caviae*, genetically modified *C. necator* (in publication: *Ralstonia eutropha*) expressing the *Pseudomonas fluorescens* GK-13 synthase gene

is also reported to accumulate *scl-mcl*-PHA hybrid copolyesters. Dependent on the substrate provided (salts of different fatty acids), the generated copolyesters contained, besides 3HB, *mcl*-PHA building blocks such as 3HHx, 3HO, 3HD, or 3-hydroxydodecanoate (3HDD) (Noda et al. 2005). Later, Riedel et al. (2015) developed a process for *scl-mcl*-PHA hybrid copolyester production by rec. *R. eutropha* from industrial rendered animal waste fats of high melting temperature and high free fatty acid content; on a 5-L bioreactor scale, a cell dry mass of 45 g/L, containing 60 wt% P(3HB-*co*-3HHx) with 19 mol% 3HHx, was obtained. For recovery of P(3HB-*co*-3HHx) produced in this process from biomass, the authors only recently presented a two-stage extraction process based on the biogenic, nontoxic solvents acetone (solvent for P(3HB-*co*-3HHx)), and isopropanol (precipitant of P(3HB-*co*-3HHx) from the acetone solution) (Bartels et al. 2020).

7 Other PHA Copolyesters



Besides the abovementioned PHA copolymers that have 4 distinct comonomers (building blocks; 3HB, 3HV, 4HB, and 3HHx), to date some 150 different comonomers for PHA have been discovered (Hazer and Steinbüchel 2007; Chen 2009), and comprehensive reviews on these building blocks have been composed (Zinn et al. 2001, 2016; Koller 2018a). Generally, such copolymers are either *scl*-PHA (building blocks with no more than five carbon atoms; polymerization catalyzed by Class I or Class III PHA synthases) or *mcl*-PHA (building blocks with 6–14 carbon atoms, e.g., 3HHx, 3-hydroxyheptanoate (3HHp), 3HO, 3-hydroxynonanoate (3HN), 3HD, 3HDD, etc.). Exceptions in this context are the above-described "Nodax™-type" *scl-mcl*-hybrid PHAs like P(3HB-*co*-3HHx). For *mcl*-PHA, copolyesters with saturated (unbranched and branched) and unsaturated (unbranched only) and side chain-functionalized monomers are reported, typically characterized by low glass transition temperature T_g, low crystallinity, and low melting temperature T_m (Zinn 2010). Functionalized monomers include halogenated

(bromo-, chloro-, and fluoro-3-hydroxyalkanoates) representatives, cyano-hydroxyalkanoates, and aromatic building blocks (e.g., 3-hydroxy-5-phenylvalerate, 3-hydroxy-5*p*-methylphenylvalerate) (Hany et al. 2009; Hanik et al. 2019; reviewed by Zinn et al. 2001; Koller 2018a).

Post-synthetic chemical modification of *mcl*-PHA with terminally unsaturated monomers (obtained by feeding ω-unsaturated fatty acids as precursors to *Pseudomonas putida*) can generate PHA bearing epoxide groups, which allows chemical cross-linking of the materials for enhancing the elastomeric properties by reducing crystallinity, increasing water resistance, and, at the same time, not compromising the material's biodegradability (Park et al. 1998). Alternatively, such pendant vinyl groups in PHA building blocks can be chemically oxidized to terminal diol groups, which leads to highly hydrophilic PHA (Lee et al. 2000b). By oxidation of poly (3-hydroxyoctanoate-*co*-3-hydroxyundec-10-enoate) with $KMnO_4$, Lee and Park (2000) succeeded for the first time in producing PHA with pendant carboxylate groups (about half of the vinyl groups were converted to carboxylate groups). These materials had hydrophilicity high enough to make them even water-soluble, while they were no more soluble in the typical hydrophobic PHA solvent chloroform. As an unintended side effect, these harsh oxidation conditions degraded the backbone of the PHA biopolymer, resulting in substantially reduced molecular mass. This problem was overcome in 2003 by Stigers and Tew, who used OsO_4 and ozone and obtained carboxylated poly(3-hydroxyoctanoate-*co*-3-hydroxyundec-10-enoate) with 100% conversion of the pendant vinyl groups. These materials showed only little molecular mass degradation and high solubility in polar solvents, such as acetone and acetone/water and THF/water mixtures (Stigers and Tew 2003). Olefinic PHA was further modified with amino acids rendering them to interesting hybrid polymers for medical applications (Bassas-Galià et al. 2015).

Table 3 provides an overview on material characteristics of the "bulk PHA" polyesters P(3HB), P(3HB-*co*-3HV), P(3HB-*co*-4HB), P(4HB), P(3HB-*co*-3HHx), and selected types of *mcl*-PHA.

8 Bioreactors, Cultivation Regimes, and Product Formation Conditions for PHA

8.1 Principle Aspects of PHA Cultivations

PHA bioproduction typically occurs under controlled conditions such as temperature, pH value, and dissolved oxygen tension (DOT) in aerated bioreactors. These bioreactors are typically stirred tank reactors (STRs) developed for some of the well-established aerobic cultivation processes known in biotechnology, namely, production of yeast biomass (autocatalytic process), penicillin (*Penicillium chrysogenum*), or citric acid (*Aspergillus niger*). In 2005, Lenz and Marchessault illustratively

Table 3 Typical values reported for the material properties of different types of PHA

Type of PHA	Melting point T_m [°C]	Glass transition temperature T_d [°C]	Degree of crystallinity X_c [%]	Elongation at break ε [%]	Tensile strength σ [MPa]
P(3HB)	170–180[a,b,c]	~3–5[b,d]	55–80[a,e]	3[f], 4[c], 8[b]	36[f], 40–50[b, g,h]
P(3HB-*co*-10%-3HV)	~140[b]	~0–1[i,g]	60[b]	20[b]	25[b]
P(3HB-*co*-20%-3HV)	~130[b]	~−15	35[b]	50[b]	20[b]
P(3HB-*co*-4HB)	160[j] (11% 4HB)	−2 (7% 4HB; [h])	~20[k] to 32[j]	242 (10% 4HB)[h]	24[h] (11% 4HB)
	152 (17% 4HB)[j]	−7 (16% 4HB[h])	~13[k] to 30 (17% 4HB[a,b,k])	444 (16% 4HB)[h]	26[h] (20% 4HB)
P(4HB)	60[f,g]	~−50[f,g,h]	34[h]	1000[g,f]	50[f]
P(3HB-*co*-3HHx)	~170 (for 12–68% 3HHx)[l]; 164–156 (for 4–12% 3HHx)[d]	−1 to −12 (for 4–12% 3HHx)[d]	56–21 (for 12–68% 3HHx)[l]	4–177 (for 12–68% 3HHx)[l]	18–7 (for 12–68% 3HHx)[l]
	151 (for 12% 3HHx)[m, p]	−1 (for 12% 3HHx)[m, p]	29 (for 12% 3HHx)[m, p]	67.3 (for 12% 3HHx)[m, p]	11.3 (for 12% 3HHx)[m, p]
mcl-PHA	~ 40 to 60[n,o]	~−30 to −60[i,n,o]	~10 to 20[n,o]	Some 100%[i,o]	Typically below 10
	53 and 65 (*mcl*-PHA with 3HHx, 3HO 3HD, and 15% or 39% 3HDD)[m, p]	−44 and −43 (*mcl*-PHA with 3HHx, 3HO 3HD, and 15% or 39% 3HDD)[m, p]	12 and 19 (*mcl*-PHA with 3HHx, 3HO 3HD, and 15% or 39% 3HDD)[m, p]	189 and 125 (*mcl*-PHA with 3HHx, 3HO 3HD, and 15% or 39% 3HDD)[m, p]	8.7 and 11.3 (*mcl*-PHA with 3HHx, 3HO 3HD, and 15% or 39% 3HDD)[m, p]
	58 (*mcl*-PHA with 3% 3HHx, 19% 3HO, 27% 3HD, 20% 3HDD, and 31% or 49% 3HTD)[m, p]	−40 (*mcl*-PHA with 3% 3HHx, 19% 3HO, 27% 3HD, 20% 3HDD, and 31% or 49% 3HTD)[m, p]	17.5 (*mcl*-PHA with 3% 3HHx, 19% 3HO, 27% 3HD, 20% 3HDD, and 31% or 49% 3HTD)[m, p]	275 (*mcl*-PHA with 3% 3HHx, 19% 3HO, 27% 3HD, 20% 3HDD, and 31% or 49% 3HTD)[m, p]	7.6 (*mcl*-PHA with 3% 3HHx, 19% 3HO, 27% 3HD, 20% 3HDD, and 31% or 49% 3HTD)[m, p]
	67 (*mcl*-PHA with 2% 3HHx, 11% 3HO, 22% 3HD, 16% 3HDD, and 49% 3HTD)[m]	−40 (*mcl*-PHA with 2% 3HHx, 11% 3HO, 22% 3HD, 16% 3HDD, and 49% 3HTD)[m, p]	17 (*mcl*-PHA with 2% 3HHx, 11% 3HO, 22% 3HD, 16% 3HDD, and 49% 3HTD)[m, p]	108 (*mcl*-PHA with 2% 3HHx, 11% 3HO, 22% 3HD, 16% 3HDD, and 49% 3HTD)[m, p]	3.2 (*mcl*-PHA with 2% 3HHx, 11% 3HO, 22% 3HD, 16% 3HDD, and 49% 3HTD)[m, p]

(continued)

Table 3 (continued)

Type of PHA	Melting point T_m [°C]	Glass transition temperature T_d [°C]	Degree of crystallinity X_c [%]	Elongation at break ε [%]	Tensile strength σ [MPa]
	Poly(HP-*co*-HA-*co*-HE): No crystalline melting endotherm[c]	−40 to −6: poly(HP-*co*-HA-*co*-HE) with 0–60% aromatic building blocks[c]			

3HB: 3-hydroxybutyrate; 3HV: 3-hydroxyvalerate; 3HHx: 3-hydroxyhexanoate; 3HO: 3-hydroxyoctanoate; 3HD: 3-hydroxydecanoate; 3HDD: 3-hydroxydodecanoate; 3HTD: 3-hydroxytetradecanoate; 4HB: 3-hydroxybutyrate
Poly(HP-*co*-HA-*co*-HE): *mcl*-PHA with saturated (3HHx, 3HO), unsaturated (3-hydroxyheptenoate, 3-hydroxynonenoate, 3-hydroxyundecenoate) and aromatic (3-hydroxyphenylvalerate) building blocks
[a]Depending on degree on polymerization (molecular mass)
[b]Luzier (1992)
[c]Hartmann et al. (2004)
[d]Murugan et al. (2017)
[e]Mitomo et al. (2001)
[f]Martin and Williams (2003)
[g]Reviewed by Utsunomia et al. (2020a)
[h]Saito et al. (1996)
[i]Reviewed by Koller (2018a)
[j]Kunioka and Doi (1990)
[k]Cong et al. (2008)
[l]Volova et al. (2016)
[m]Liu and Chen (2007)
[n]Muhr et al. (2013)
[o]Strongly dependent on exact monomeric composition (saturated, unsaturated, and aromatic building blocks)
[p]Zinn (2010)

remarked that PHA production is performed "by a largescale fermentation process not unlike the brewing of beer" (Lenz and Marchessault 2005).

Indeed, especially the processes for penicillin and citric acid production are regarded as templates for PHA production: these two bioproducts are produced in a similar two-stage process as PHA is as well. In the first stage, a nutritionally balanced cultivation medium is used to generate a high concentration of catalytically active biomass. After that, a switch in the cultivation conditions toward unbalanced growth conditions provokes changes in the metabolism, characterized by the stopping of the biomass formation (no more cell propagation) and the formation of the secondary bioproducts penicillin, citric acid, or PHA, respectively. In the case of penicillin, this switch from mycelium formation to predominant antibiotic excretion

occurs below a critical oxygen level and depletion of complex nutrients (e.g., corn steep liquor) (Zangirolami et al. 1997), while citric acid formation is typically initiated at the time point when accessible iron ions are no more available for the fungal cells in sufficient quantities; now, *A. niger* cells excrete and "send out" citric acid to scavenge remaining iron ions by forming citrate-iron complexes, which can be metabolized by the cells for building important cellular components, such as cytochromes, and expression of the lyase enzyme aconitase (EC 4.2.1.3), which is pivotal to keep the citric acid cycle (TCC) running (Max et al. 2010). In the case of PHA biopolyesters, the switch from the phase of balanced microbial growth to product formation is typically accomplished by an extreme increase of the ratio of carbon to nitrogen source. Hence, nitrogen limitation stops the formation of active prokaryotic biomass but shifts the carbon flux toward PHA accumulation (Macrae and Wilkinson 1958). For some industrially important PHA production strains, a certain amount of PHA is already synthesized during the logarithmic (exponential) phase of bacterial growth ("partially growth-associated product formation"; observed, e.g., for *Cupriavidus necator*). However, high productivity for PHA is not observed before the onset of nutritionally challenging conditions (Braunegg et al. 1998). Besides nitrogen limitation, also phosphate limitation was described as a factor provoking PHA biosynthesis, because it prevents further biomass growth due to the impossibility to generate important phosphate-containing cell components like nucleic acids, cofactors, or adenosine phosphates. However, phosphate limitation is rather scarcely used on an industrial scale due to the decreased pH buffer capacity of phosphate-free medium. Oxygen limitation also provokes PHA biosynthesis due to the interruption of the TCC, which is based on the impossibility of regeneration of the oxidized form of redox equivalents (NAD^+, $NADP^+$) when molecular oxygen as terminal electron acceptor is missing. However, such oxygen-limited bioprocesses drastically reduce volumetric product formation rates, which impedes industrial implementation. This makes nitrogen limitation still the method of choice for high-throughput PHA production processes on a larger scale. On a lab scale, other growth-limiting factors have also been described to favor PHA production rates when present in suboptimal quantities, such as potassium (potassium limitation was described to result in PHA of extraordinary high molecular mass by a methane-utilizing mixed culture), sulfur, or iron (Helm et al. 2008).

On the industrial scale, PHA production in said bioreactors typically occurs in fed-batch mode; hence, the substrate is added according to its conversion by the microbes. This is analogous to other established processes: during yeast production, substrate pulses need to be added carefully to avoid switching the metabolism from biomass formation to ethanolic fermentation, despite the ample availability of oxygen ("crabtree effect"); excessive carbon source concentrations during yeast propagation need to be avoided (Pfeiffer and Morley 2014). During penicillin and citric acid production, too, carbon source needs to be supplied well in accordance with the consumption rate by the cells in order to avoid the formation of conidia and to favor mycelium growth instead of premature onset of product formation. Especially in the case of penicillin production, excessive loads of carbon sources would stop fungal biomass growth too early and provoke the start of penicillin excretion at

biomass concentration too low to reach sufficient volumetric productivity for the antibiotic (Birol et al. 2002).

8.2 Continuous Cultivation

The future of production of PHA of predefined composition and reproducible molecular mass and polydispersity might be the switch to continuous cultivation processes. This is characterized by the productivities of continuous cultivations that outperform discontinuous batch and fed-batch processes, thereby drastically increasing operation periods per cultivation batch at maximum productivity, minimizing the number of interruption of operation for starting up, shutting down, and cleaning the equipment, and reduced personal involvement during the operation. In contrast to autocatalytic processes like yeast propagation or production of primary metabolites like lactic acid, one-stage continuous biotechnological processes are only to some extent suitable for high-throughput production of secondary metabolites like PHA. Nevertheless, single-stage cultivations were successfully implemented to tailor the monomeric unit compositions of P(3HB-*co*-3HV) (Zinn et al. 2003) and *mcl*-PHA (Amstutz et al. 2019; Hartmann et al. 2004, 2006). This was possible due to the fact that the chemostat continuous operation mode allows nutrient limitations in contrast to nutrient starvations as it is the case in batch and fed-batch cultivations. All nutrients are efficiently used converted, and thus a tailored composition can be achieved (Zinn et al. 2003).

Higher volumetric productivities than in single-stage chemostats can be achieved when two continuously operated bioreactors are linked in series. The active biomass is formed under continuous supply of a nutritionally balanced medium, while in a second vessel, only traces of nitrogen source from the first bioreactor are present. Carbon source is supplied at a feed rate according to the consumption rate by the cells. This enables considerably higher intracellular PHA fractions compared to single-stage continuous setups. Moreover, two-stage continuous cultures better address problems associated with different conversion rates for different substrates needed to produce copolyesters (reviews by Koller and Muhr 2014; Koller and Braunegg 2015b). Such two-stage continuous PHA production processes were successfully demonstrated in the past, e.g., for tailored P(3HB-*co*-4HB) production by *Delftia acidovorans* cultivated on acetic acid and GBL (Mothes and Ackermann 2005). Recently, a modified two-stage approach has been demonstrated, where the culture harvest of the first chemostat was fed to four chemostats at the second stage. This allowed the study of different growth conditions in the second stage to control the content of the 3-hydroxyphenylvalerate in the *mcl*-PHA accumulated in a recombinant *P. putida* strain (Hanik et al. 2019).

Beyond this two-stage process engineering setup, a combination of a continuously operated STR for autocatalytic biomass propagation and a tubular plug flow reactor for PHA accumulation even better matches the fundamentally different kinetic characteristics of biomass formation (autocatalytic) on the one hand, and

Fig. 3 Multistage bioreactor cascade for high-throughput PHA biosynthesis, financed by BASF SE, in operation at Graz University of Technology, Austria. (Own picture M. Koller)

PHA accumulation in not multiplying cells (first-order kinetics) on the other hand. Based on theoretical considerations, this was already postulated by Braunegg and associates in 1995 (Braunegg et al. 1995), while it took far more than one decade until the proof of concept for these ideas was provided by Atlić et al. (2011), who produced high cell concentrations of *C. necator* biomass on glucose in a continuous STR, linked to a cascade of four additional continuously operated STRs, which acted as a process engineering substitute for the mentioned tubular reactor and were supplied with carbon source only. This approach can be understood by the fact that such STR cascades, with an increasing number of STRs, approach the plug flow characteristics of a tubular reactor. Indeed, high productivity of P(3HB) homopolyester were obtained in this setup, which was operated for many days under steady-state conditions; moreover, P(3HB) quality (molecular mass and polydispersity) were highly uniform and reproducible. A photograph of this cascade setup is provided in Fig. 3.

In the future, such multistage continuous PHA production processes might enable the formation of PHA of tailored microstructure, e.g., blocky structured PHA copolyesters with alternating segments of 3HB oligomers and other building blocks such as 3HV 4HB, or 33HHx (Utsunomia et al. 2020b). Such blocky structured heteropolyesters enable more stringent control of polymer properties by increasing pliability, at the same time maintaining the structural rigidity provided by certain block regions (Ashby et al. 2018). This would be possible by fine-tuned supply of the individual cascade bioreactors with specific carbon sources acting as precursors

of individual PHA building blocks: while supply with a main carbon source (e.g., sugars) generates blocks of 3HB-oligomers, supply of short fatty acids with an odd number of carbon atoms (propionic acid, valeric acid) generates segments with high 3HV content; supply of compounds like γ-butyrolactone, 1,4-butanediol, or sodium 4-hydroxybutyrate results in 4HB blocks. This might enable the high-throughput production of high-quality PHA heteropolyesters of a microstructure of unprecedented reproducibility, which is not accessible by established fed-batch cultivation regimes. On the downside, continuous bioprocesses, especially on multistage, are more prone to disturbations and contaminations by alien microflora than discontinuous setups. This could partly be overcome by the application of robust, extremophile production strains which can be cultivated monoseptically even under strongly reduced sterility precautions, as recently postulated by the concept of "next-generation industrial biotechnology" (NGIB; Chen and Jiang 2018). Such concepts would further contribute to the reduction of PHA production costs.

9 PHA Recovery

As an important difference between the above-described three bioprocesses, citric acid and penicillin are excreted as extracellular products, while PHA gets stored intracellularly as granular inclusion bodies. This poses additional challenges in PHA recovery after the cultivation process compared to the recovery of the extracellular bioproducts like citric acid (precipitation as calcium citrate at elevated temperature) or penicillin (extraction of the easily soluble sodium salt). For the release of PHA granules stored in the interior of microbial cells, the cells need to be disrupted to set PHA granules free (reviewed by Koller 2020).

Remarkably, PHA-producing wild-type strains are characterized by rather high robustness of their cell wall, which in turn challenges PHA recovery. This is in contrast to recombinant PHA producers, derived from wild-type strains that are not naturally equipped to carry excessive loads of storage materials, as demonstrated for recombinant *E. coli*, a nonnatural PHA synthesizing strain, transformed with PHA biosynthesis genes. These recombinant organisms were indeed able to store excessive loads of PHA, which makes the cells fragile; stirring this biomass in diluted NaOH solution resulted in cell disintegration and release of PHA granules (Choi and Lee 1999). Mechanic cell disintegration for PHA recovery was demonstrated by high-pressure homogenization as it is also used for the recovery of other intracellular bioproducts like many enzymes. After mechanical break up of cells, released PHA granules can be separated from cell debris in the liquid phase by dissolved air floatation (Koller et al. 2013).

More frequently, extraction methods are used which, on the one hand, weaken the cell membrane and, on the other hand, solubilize the PHA biopolyesters. Best established PHA extraction solvents like chloroform, although giving high product purity, extraction yields, and extraction efficiency, are halogenated and toxic and stem from petroleum chemistry. The use of such solvents in biotechnology,

especially for large-scale use, should be avoided. Current trends in PHA recovery, therefore, aim to replace these traditional methods by "greener" approaches, such as biogenic solvents (alcohols, lactic acid esters, acetone, etc. (reviewed by Koller 2020)), digestion of non-PHA cell mass by enzyme cocktails (Yasotha et al. 2006), or even in the intestine of insects (Chee et al. 2019), use of supercritical solvents (Hejazi et al. 2003), nontoxic surfactants like soaps (Pospisilova et al. 2021), or ionic liquids (Kobayashi et al. 2015).

10 Commercialization of PHA

It has to be emphasized that current PHA production on an industrial scale is embryonic compared to the global fossil plastic production, which amounts to about 400 million tons per year, excluding fibers for textiles such as polyesters and polyamides and elastomers, such as for car tires. Including those take the total synthetic polymer production to over 550 million tons per year. PHA production did not exceed 10 kt in 2020 although companies claim production capacity estimated to approximately 20–30 kt per year. In 2020, the total production volume of renewable polymers was about 4.2 million tons, or about 1% of the total fossil plastics production. Several PHA-producing companies have announced significant volume projections and capacity announcements exceeding one million tons over the next 5 years to 2025 (Koller and Mukherjee 2022). One primary driver behind such large-scale volume announcements is the acceptance of the financial markets on the viability of PHA as a material that can replace single-use plastics packaging (Jost 2018). This is reflected by the recent listing of Danimer Scientific on NASDAQ, one of the US stock exchanges. However, the European Union is about to announce their Single-Use Plastics Directive, which bans fossil plastics as well as renewable polymers such as PLA, PBS, and PHA. However, legislation in PR China and in the United States mostly favor the use of compostable and biodegradable materials which should help the cause of large-scale production and acceptance PHA as an alternate to single-use petrol-based plastics.

Indeed, we currently witness annual growth rates for biopolymers of about 8%, with PHA being part of this upward trend. In any case, the global PHA market is projected to grow from US-$57 mio. in 2019 to US-$98 mio. in only 5 years, which corresponds to a compound annual growth rate (CAGR) of 11.2%. It is expected that PR China will keep the lead in global PHA production also during the next years, while Europe currently produces about 25% of all PHA. However, European PHA production is expected to increase by 27% from 2019 to 2023 (reported by Kourmentza et al. 2020).

Industrial-scale PHA biopolymers currently being produced are *scl*-PHA including P(3HB), P(3HB-*co*-3HV), and P(3HB-*co*-4HB) and one *scl*-*mcl*-PHA, namely, P(3HB-*co*-3HHx) (Koller and Mukherjee 2022). There is one company commercializing medium-chain-length PHA (*mcl*-PHA) where the medium-chain-length part is higher than six carbon atoms on a reasonable scale: PolyFerm Canada sells

VersaMer® PHA as "irregular pieces, pellets, latex" consisting, *inter alia*, of P(3HO-*co*-3HHx), P(3HN-*co*-3HHp), and *mcl*-PHA containing also unsaturated building blocks. These materials have outstanding low T_g values (−45 to −35 °C), low T_m (45–65 °C), low molecular mass typical for *mcl*-PHA (100–150 kDa), and very high elongation at break of 1200–1400% (online resource 2 n.d.). In 2016, PolyFerm Canada has licensed its technology to TerraVerdae Bioproducts, an industrial biotechnology company developing advanced bioplastics and other biomaterials from C1-feedstocks (methanol), which are scaling up (online resource 2 n.d. and personal communication Bruce Ramsay).

The majority of companies commercializing PHA, however, is focusing on the "top-selling" *scl*-PHA bulk products, namely, brand names of commercialized polyesters in square brackets below:

- P(3HB): PHB Industrial S.A, Brazil [Biocycle®], Bio-On, Italy [Minerv-PHA®] (currently not producing), Biomer, Germany [Biomer®], Mango Materials, USA [YOPP®], COFCO, PR China, Newlight Technologies LLC, USA [AirCarbon® PHA], NAFIGATE Corporation [HYDAL® PHA]; in the past: Imperial Chemical Industries (ICI), UK, until end of 1990s [BIOPOL®]; Chemie Linz/PCD Polymere GmbH, Austria, in the late 1980s
- P(3HB-*co*-3HV): TianAn Biologic Materials Co. [ENMAT®], PHB Industrial S.A, Brazil [Biocycle®], Bio-On, Italy [Minerv-PHA®], PhaBuilder, PR China; in the past: Telles (joint-venture Metabolix & ADM from 2009 to 2012 [Mirel®], Imperial Chemical Industries, UK [BIOPOL®])
- P(3HB-*co*-4HB): Tianjin GreenBio Materials Co., Ltd., PR China [SoGreen®], Shenzhen Ecomann Biotechnology Co., Ltd., PR China [AmBio®], Tepha Medical Devices, Inc., USA [TephaELAST®], CJ, Republic of Korea (technology sold from Metabolix) [Yield10®], PhaBuilder, PR China [copolyester "mP34HB 10"], Medpha, PR China [Medpha®], in addition to Tepha's surgically important homopolyester P(4HB) [TephaFLEX®]
- P(3HB-*co*-3HHx): Danimer Scientific, USA (former Meredian Holdings Group Inc. and MHG; technology originally from Proctor & Gamble) [Nodax®], Kanegafuchi Chemical Industry Co Ltd. (Kaneka), Japan [PHBM®], Bluepha Co., Ltd., PR China [Bluepha PHA], RWDC Industries, Ltd., USA [Solon®]

11 Spent PHA Is Naturally Degraded

Although PLA undergoes composting to CO_2, water, and biomass in industrial composting facilities under elevated temperature, it is not compostable under home-composting conditions and does not biodegrade in the environment. This high recalcitrance of PLA was demonstrated a long time ago, e.g., by Hyon et al. (1984), who showed that PLA fibers remained intact for half a year in phosphate buffer saline at 37 °C and pH value 7.4; only when elevating the temperature to 100 °C, 50% mass loss was observed. This is in contrast to PHA types that are being commercialized today and discussed above (Ong et al. 2017). These PHA grades are

all biodegradable and compostable, both under industrial and home-composting conditions, which is confirmed by the strict certification processes many commercially available types of PHA were subjected to (Koller and Mukherjee 2020). Subsequent paragraphs provide examples for degradability of different types of commercial PHA.

11.1 P(3HB)

Even the highly crystalline P(3HB) homopolyester is biodegradable and compostable. In aerobic environments, the end products of biodegradation are carbon dioxide, water, and humus, while under anaerobic conditions, CH_4 is also produced, with no harmful intermediates or by-products being generated. In this context, the company Biomer claims that their products (P(3HB)) are "fully biodegradable" and compostable (online resource 3 n.d.). A study by Kim et al. compared biodegradability of P(3HB) homopolyester from ICI with the thermoplastic starch-based composite Mater-Bi® from Novamont and Sky-Green®, a chemosynthetic biodegradable polymer consisting of succinic acid, adipic acid, butanediol, and ethylene glycol. Biodegradation experiments were carried out in forest soil, sandy soil, activated sludge soil, and in farm soil at different temperatures (28, 37, and 60 °C) over a period of 28 days. P(3HB) homopolyester showed complete degradation (90% conversion to gases and water, the rest being humus is defined as complete biodegradation) in all conditions, including activated sludge soil at 37 °C. None of the aforementioned polymers tested showed complete biodegradation at any conditions and soil. Some showed modest degradation in sandy soil and forest soil. In farm soil as well, P(3HB) again outperformed the other polymers at 37 °C (Kim et al. 2000). It was shown long time ago by Kumagai et al. (1992) that degradability of P(3HB) strongly depends on the polymer's crystallinity; when incubated in solution of *Alcaligenes faecalis* depolymerase at 37 °C and pH value 7.4, faster degradation was observed for polymer samples (produced either by solvent casting or by crystallization from the melt) of lower crystallinity, and molecular mass remained mainly unchanged during degradation. The authors concluded that the depolymerase enzyme first hydrolyzes P(3HB) chains in an amorphous state on the film surface and subsequently degrades P(3HB) chains in the crystalline state. Importantly, this study also showed that spherulite size has hardly any impact on the degradation rate. In 1992, Luzier noticed that no degradation of P(3HB) was observed in environments free of biocatalytic activity, such as in dry or humid air, which in turns makes products made of P(3HB) stable in non-biotic environments, e.g., during the use of packaging, as a formed product or article where they do not come in contact with microorganisms that can consume them as food or substrate.

11.2 P(3HB-co-3HV)

Luzier (1992) studied the biodegradability of P(3HB-*co*-3HV) copolyesters in the early 1990s. He demonstrated that ICI BIOPOL® products (made from P(3HB-*co*-3HV)) exhibited the slowest degradation in seawater (350 weeks for complete degradation of standard injection-molded specimens), while the highest degradation rates were reported in anaerobic sewage (only 6 weeks for complete degradation). Degradation rates in estuarine sediments, aerobic sewage, and soil were in between these two extremes. Similar to P(3HB) homopolyester, no degradation was observed in humid or dry air without the presence of microflora. In their biodegradability tests, Rosa et al. (2004) demonstrated that P(3HB) having 72% degree of crystallinity and poly(3HB-*co*-3HV) having 50% degree of crystallinity (3HV fraction not disclosed) from PHB/ISA demonstrated weight loss over 10 months. Both polymers showed similar biodegradation in soil composting medium at 46 and at 24 °C in a soil simulator containing *inter alia* manure.

ENMAT's Y1000P P(3HB-*co*-3HV) has been certified as "compostable" by the US-Biodegradable Products Institute (BPI) in 2008 and is listed as a Food Contact Material ("FCM") substance No. 744 in Table 1 of Annex I of the Plastics Regulation of the EU and EU REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) compliant since 2008 (online resource 4 n.d.).

P(3HB-*co*-3HV) has been demonstrated to hold high promise for biomedical application due to its high biocompatibility. In contrast to competing materials used in biomedical applications such as PLA and PLGA, P(3HB-*co*-3HV)'s biodegradation products are less bioactive in vivo and result in less tissue acidification. Moreover, in vivo degradation of P(3HB-*co*-3HV) occurs at slower rates than for PLA, which makes the material more useful for implant manufacturing for bone repair and bone regeneration. Especially when produced by Gram-positive microbes like representatives of the genera *Bacillus*, *Streptomyces*, or *Rhodococcus*, P(3HB-*co*-3HV) is reported to have biocompatibility superior to PHA from Gram-negatives due to the lacking lipopolysaccharides, a group of endotoxins from the cell wall of Gram-negatives, which are typically co-extracted with PHA during the recovery process and cause inflammatory reactions (Zinn et al. 2001; Koller 2018b; Bucsella et al. 2020). The expedient biocompatibility of P(3HB-*co*-3HV) was only recently substantiated again by Mohandas et al. (2021), who used cast films made of P(3HB-*co*-3HV) from the marine bacterium *Bacillus cereus* MCCB 281 to test attachment, viability, and proliferation of fibroblast cells on them; high compatibility for these cells on the films was shown. Moreover, by hemolysis, in vitro platelet adhesion, and coagulation assays, the excellent blood compatibility of these P(3HB-*co*-3HV) films for use as blood contact graft materials was shown.

11.3 P(3HB-co-4HB)

Already in 1989, Kunioka et al. noticed that rate of biodegradation of P(3HB-*co*-4HB) film samples with 0–37 mol% 4HB in soil and activated sludge was enhanced by the presence of 4HB units (Kunioka et al. 1989). In a study by Nakamura et al. (1992), the impact of 4HB fraction in P(3HB-*co*-4HB) films on biodegradability was studied in more details; in the aqueous solution of an extracellular *A. faecalis* P (3HB) depolymerase, films with 0, 6, 10, 28, 85, and 94% 4HB were subjected toward enzymatic degradation at 37 °C and pH value of 7.5. The rate of degradation, expressed by mass loss, strongly increased with increasing 4HB fraction op to 28%; already the copolyester with 6% 4HB was degraded much faster than the P(3HB) homopolyester, which correlates well with decreasing crystallinity. However, samples with very high 4HB content were degraded considerably slower than P(3HB) due to the hard susceptibility of the depolymerase enzyme to the 4HB moieties (Nakamura et al. 1992). In marine environment, degradability of P(3HB), P(3HB-*co*-3HV) (4, 21 and 61% 3HV) and poly(3HB-*co*-4HB) (10% 4HB) samples (films obtained by solvent casting) was compared during an incubation period of 1 year. It was shown that all samples were degraded via surface erosion, while surface erosion was most pronounced for the P(3HB-*co*-3HV) sample with 21% 3HV. Interestingly, the speed of surface erosion was not too much dependent on PHA's composition but highly on the temperature of seawater (Doi et al. 1992). Similar to the results for P (3HB) described by Kumagai et al. (1992), no change in molecular mass during degradation was observed for any of the samples (Doi et al. 1992).

For currently commercially available P(3HB-*co*-4HB), GreenBio's product "Sogreen-00X" can be "completely degraded into carbon dioxide and water at the environment of soil, rivers, sewage and marine water in 3–6 months" (manufacturer information; online resource 5 n.d.), while Shenzhen Ecomann's Ambio® P (3HB-*co*-4HB) is TÜV Austria certified for industrial and home composting and also FDA approved (online resource 6 n.d.).

11.4 P(4HB)

P(4HB) homopolyester is not accessible by wild-type organisms due to the fact that the microbial metabolism of 4HB precursors intrinsically always yields, besides 4HB, also some 3HB. However, its in vivo degradation yields 4HB as the sole product, which is a naturally occurring, biocompatible compound in human tissue, including various organs, muscle, and brown fat. Therefore, P(4HB) is FDA approved for medical application as a suture material since 2007 and still the only PHA with FDA clearances for clinical usage. Currently, Tepha Inc. is the globally only company producing P(4HB) on a commercial scale. They produce the material from inexpensive feedstocks by recombinant *E. coli*, based on a patent originally filed by Metabolix (reviewed by Utsunomia et al. 2020a). Regarding in vivo

degradability of TephaFLEX® poly(4HB), its absorption rate is reported with only 8–52 weeks, which is considerably faster than reported for P(3HB). With an in vivo half-life of only about 27 min, 4HB is metabolized very quickly in the human body to CO_2 and water (online resource 7 n.d.).

11.5 P(3HB-co-3HHx)

PHBH®, the "100% plant-derived biopolymer developed by Kaneka, is certified to biodegrade in seawater," and received the label "OK Biodegradable MARINE" by TÜV Austria (formerly Vinçotte). This means that "the degree of biodegradation should be 90% or more within 6 months in seawater (30 °C), and disintegration to less than 10% remnants in a 2 mm sieve after 12 weeks at 30 °C. Acquired certification in September 2017 from Vinçotte (now TÜV Austria Belgium), an international certification organization based in Belgium." This degradation under standardized marine conditions (23 °C) was successfully demonstrated for prototype PHBH®specimens like drinking straws, bottles, and knives. According to comparative experiments, PHBH®was faster degraded under marine conditions than PCL, PHSA, PBAT, PBS, and PLA (not degraded in seawater at all after 28 days). Toxicity tests were also carried out to assess the biocompatibility of PHBM®for marine organisms; no acute toxicity was observed for fish, plankton, and shellfish. Further labels for Kaneka's PHBH®are "OK Compost Industrial," "OK Compost Home," "OK Biodegradable Soil" (certification in progress), and "OK Biobased" according to TÜV Austria, the "Biobased" certification for Japan, and the "Industrial Compostable" certification for Japan and USA. Eric Lepoudre, Business Manager Biopolymers at the Green Polymer Division at Kaneka, reported that according to ISO 14855 tests, PHBH®shows even better aerobic biodegradability in compost than cellulose and almost identical anaerobic biodegradation in aqueous phase according to ISO 14853 than cellulose (note: PLA was not degraded at all under these conditions during a 28 days test span), while under anaerobic conditions in solid phase according to ISO 15985, degradation was only insignificantly slower than for cellulose and much faster than for PLA (Lepoudre 2018) [online resource 8]).

Solon® P(3HB-*co*-3HHx) is thoroughly tested for its biodegradability, compostability, and natural origin; the materials are certified for biodegradability in soil ("OK biodegradable SOIL", TÜV Austria), fresh water ("OK biodegradable WATER"), and marine water ("OK biodegradable MARINE"), home and industrial composting ("OK compost HOME", "OK compost"), and with "OK biobased (online resource 9).

The same biodegradability was ascertained for Danimer Scientific's Nodax® P (3HB-*co*-3HHx), which is certificated, besides its "biobased" nature (ASTM D 6866) regarding its anaerobic and aerobic degradability in soil (TÜV Austria, ASTM D5988), freshwater (TÜV Austria, ASTM D5271, EN 29408), and marine water (TÜV Austria, ASTM D6691) and its suitability for industrial (TÜV Austria,

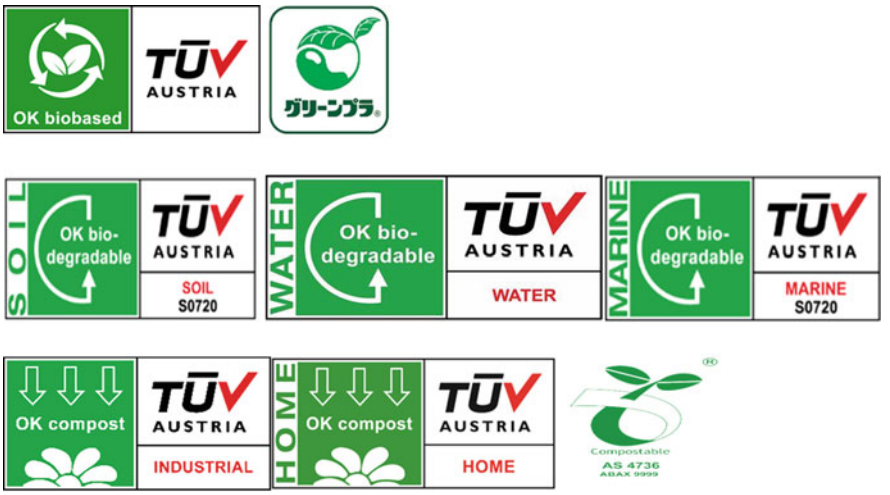


Fig. 4 Selected certification labels awarded to different commercially available PHA products. Upper row: "OK biobased" by TÜV Austria and Japanese "biobased" label; middle row: "OK biodegradables SOIL," "OK biodegradable Water" (for degradation in freshwater), and "OK biodegradable MARINE" by TÜV Austria. Third raw: TÜV Austria labels for industrial and home composting and "compostable" certification according to Australian Standard (AS) 4736-2006

ASTM D6400, EN 13432) and home composting (TÜV Austria, ASTM D6400, EN 13432) (online resource 10 n.d.).

In 2014, Meredian Inc. (joint later Danimer Scientific) received a Food and Substance Contact Notification approval from FDA, certifying that their P (3HB-*co*-3HHx) biopolyesters are safe to use for food contact and can be classified as "nonhazardous waste" after disposal (online resource 11 n.d.). Figure 4 shows the most important certificates for currently commercialized PHA.

12 Conclusions

Fossil plastics have exhibited their beneficial role in improving our quality of life in food and medical applications. Plastics in automobiles have allowed for many new technological and safety improvements. Hence, there is little doubt that plastics with their ubiquitous presence in our lives have made our society better. However, the end-of-life issues with plastics, especially their lack of appropriate collection systems and leakage into the environment, may have already risen to epidemic propositions. A UN study concluded that microplastics are the next major environmental epidemic that is yet to be appropriately measured. These and issue of greenhouse gas emissions due to their manufacture has raised significant awareness among consumers and to a degree among policy makers and the industry. While regulations are

being put in place to reduce plastics use, not all of these measures are necessarily beneficial to the environment and practical from a convenience standpoint. The benefits that plastics provide do not need to be sacrificed due to their negative environmental impacts. PHA is the perfect example of alternatives that can bridge the benefits of plastics without the environmental damage they cause.

PHA have been established as macromolecules occurring in nature, being produced by microorganisms and therefore being biodegraded by them. Moreover, they originate from natural substrates. They play a leading and important role in the metabolism of numerous microorganisms. Their manufacture on large scale requires the use of the same microorganisms as found in nature. PHA are nontoxic and biocompatible to humans and other living organisms; therefore, they are metabolized if ingested by living organisms. The benefits exhibited by fossil plastics are also available through the use of PHA. The limitations of fossil plastics have opened the door for PHA to play a leading role to continue to benefit humanity while maintaining nature's cycle of circularity and sustainability due to nature's ability to not only produce them but also to biodegrade them.

As shown, we currently witness considerable activities in different global regions toward commercialization of the most important "bulk PHA" such as P(3HB), P(3HB-*co*-3HV), P(3HB-*co*-4HB), P(3HB-*co*-3HHx), and, to a minor extent, P(4HB) and *mcl*-PHA copolyesters. The next years will draw a clearer picture about which concepts of marketing of PHA are indeed future-fit, be it in terms of market requirements, customer acceptance, sustainability, or economic feasibility. However, it is quite clear that long-term success of commercial PHA needs to be grounded on some defined solid fundaments: robust and powerful microbial production strains, optimized and simple cultivation facilities, amply available renewable feedstocks, and sustainable and inexpensive downstream processing technologies.

References

Abd-El-Haleem DA (2009) Biosynthesis of polyhydroxyalkanotes in wild type yeasts. Pol J Microbiol 58(1):37–41

Achinas S, Achinas V, Euverink GJW (2017) A technological overview of biogas production from biowaste. Engineering 3(3):299–307

Alves LPS, do Amaral FP, Kim D, Bom MT, Gavídia MP, Teixeira CS, Holthman F, de Oliveira Pedrosa F, Maltempi de Souza E, Satie Chubatsu L, Müller-Santos M, Stacey G (2019) Importance of poly-3-hydroxybutyrate metabolism to the ability of *Herbaspirillum seropedicae* to promote plant growth. Appl Environ Microbiol 85(6):e02586-18

Amstutz V, Zinn M (2020) Syngas as a sustainable carbon source for PHA production. In: Koller M (ed) The handbook of polyhydroxyalkanoates, Microbial biosynthesis and feedstocks, vol 1. CRC Press, Boca Raton, pp 377–416

Amstutz V, Hanik N, Pott J, Utsunomia C, Zinn M (2019) Tailored biosynthesis of polyhydroxyalkanoates in chemostat cultures. Methods Enzymol 627:99–123

Ashby RD, Solaiman DK, Nuñez A, Strahan GD, Johnston DB (2018) *Burkholderia sacchari* DSM 17165: a source of compositionally-tunable block-copolymeric short-chain poly (hydroxyalkanoates) from xylose and levulinic acid. Bioresour Technol 253:333–342

Atlić A, Koller M, Scherzer D, Kutschera C, Grillo-Fernandes E, Horvat P, Chiellini E, Braunegg G (2011) Continuous production of poly([*R*]-3-hydroxybutyrate) by *Cupriavidus necator* in a multistage bioreactor cascade. Appl Microbiol Biotechnol 91(2):295–304

Barham PJ, Keller A, Otun EL, Holmes PA (1984) Crystallization and morphology of a bacterial thermoplastic: poly-3-hydroxybutyrate. J Mater Sci 19(9):2781–2794

Barnard GN, Sanders JK (1988) Observation of mobile poly(β-hydroxybutyrate) in the storage granules of *Methylobacterium* AM1 by in vivo ^{13}C-NMR spectroscopy. FEBS Lett 231(1):16–18

Bartels M, Gutschmann B, Widmer T, Grimm T, Neubauer P, Riedel SL (2020) Recovery of the PHA copolymer P(HB-*co*-HHx) with non-halogenated solvents: influences on molecular weight and HHx-content. Front Bioeng Biotechnol 8:944

Basnett P, Marcello E, Lukasiewicz B, Panchal B, Nigmatullin R, Knowles JC, Roy I (2018) Biosynthesis and characterization of a novel, biocompatible medium chain length polyhydroxyalkanoate by *Pseudomonas mendocina* CH50 using coconut oil as the carbon source. J Mater Sci Mater Med 29(12):1–11

Bassas-Galià M, Gonzalez A, Micaux F, Gaillard V, Piantini U, Schintke S, Zinn M, Mathieu M (2015) Chemical modification of polyhydroxyalkanoates (PHAs) for the preparation of hybrid biomaterials. CHIMIA 69:627–630

Birol G, Ündey C, Cinar A (2002) A modular simulation package for fed-batch fermentation: penicillin production. Comput Chem Eng 26(11):1553–1565

Bonilla M, Olivaro C, Corona M, Vazquez A, Soubes M (2005) Production and characterization of a new bioemulsifier from *Pseudomonas putida* ML2. J Appl Microbiol 98(2):456–463

Bonthrone KM, Clauss J, Horowitz DM, Hunter BK, Sanders JKM (1992) The biological and physical chemistry of polyhydroxyalkanoates as seen by NMR spectroscopy. FEMS Microbiol Lett 103(2–4):269–277

Braunegg G, Lefebvre G, Renner G, Zeiser A, Haage G, Loidl-Lanthaler K (1995) Kinetics as a tool for polyhydroxyalkanoate production optimization. Can J Microbiol 41(13):239–248

Braunegg G, Lefebvre G, Genser KF (1998) Polyhydroxyalkanoates, biopolyesters from renewable resources: physiological and engineering aspects. J Biotechnol 65(2–3):127–161

Braunegg G, Genser K, Bona R, Haage G, Schellauf F, Winkler E (1999) Production of PHAs from agricultural waste material. Macromol Symp 144(1):375–383

Bucsella B, Hoffmann A, Zollinger M, Stephan F, Pattky M, Daumke R, Heiligtag FJ, Frank B, Bassa-Galia M, Zinn M, Kalman F (2020) Novel RP-HPLC based assay for selective and sensitive endotoxin quantification. Anal Methods 12(38):4621–4634

Burgard A, Burk MJ, Osterhout R, Van Dien S, Yim H (2016) Development of a commercial scale process for production of 1,4-butanediol from sugar. Curr Opin Biotechnol 42:118–125

Callister SM, Agger WA (1987) Enumeration and characterization of *Aeromonas hydrophila* and *Aeromonas caviae* isolated from grocery store produce. Appl Environ Microbiol 53(2):249–253

Cameron RE (1960) Communities of soil algae occurring in the Sonoran Desert in Arizona. J Ariz Acad Sci 1(3):85–88

Cash CD (1994) Gammahydroxybutyrate: an overview of the pros and cons for it being a neurotransmitter and/or a useful therapeutic agent. Neurosci Biobehav Rev 18(2):291–304

Chee JY, Tan Y, Samian MR, Sudesh K (2010) Isolation and characterization of a *Burkholderia* sp. USM (JCM15050) capable of producing polyhydroxyalkanoate (PHA) from triglycerides, fatty acids and glycerols. J Polym Environ 18(4):584–592

Chee JY, Lakshmanan M, Jeepery IF, Hairudin NHM, Sudesh K (2019) The potential application of *Cupriavidus necator* as polyhydroxyalkanoates producer and single cell protein: a review on scientific, cultural and religious perspectives. Appl Food Biotechnol 6(1):19–34

Chen GQ (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. Chem Soc Rev 38(8):2434–2446

Chen GQ, Jiang XR (2018) Next generation industrial biotechnology based on extremophilic bacteria. Curr Opin Biotechnol 50:94–100

Chen GQ, Wu Q (2005) Microbial production and applications of chiral hydroxyalkanoates. Appl Microbiol Biotechnol 67(5):592–599

Chen G, Zhang G, Park S, Lee S (2001) Industrial scale production of poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate). Appl Microbiol Biotechnol 57(1):50–55

Choi JI, Lee SY (1999) Efficient and economical recovery of poly(3-hydroxybutyrate) from recombinant *Escherichia coli* by simple digestion with chemicals. Biotechnol Bioeng 62(5): 546–553

Chowdhury AA (1963) Poly-β-hydroxybuttersäure abbauende Bakterien und Exoenzym. Arch Microbiol 47(2):167–200

Cong C, Zhang S, Xu R, Lu W, Yu D (2008) The influence of 4HB content on the properties of poly (3-hydroxylbutyrate-*co*-4-hydroxylbutyrate) based on melt molded sheets. J Appl Polym Sci 109:1962–1967

Csomorová K, Rychlý J, Bakoš D, Janigova I (1994) The effect of inorganic additives on the decomposition of poly(beta-hydroxybutyrate) into volatile products. Polym Degrad Stab 43(3): 441–446

Dawes EA, Senior PJ (1973) The role and regulation of energy reserve polymers in micro-organisms. Adv Microb Physiol 10:135–266

de Koning GJM, Lemstra PJ (1992) The amorphous state of bacterial poly[(*R*)-3-hydroxyalkanoate] in vivo. Polymer 33:3292–3294

de Smet MJ, Eggink G, Witholt B, Kingma J, Wynberg H (1983) Characterization of intracellular inclusions formed by *Pseudomonas oleovorans* during growth on octane. J Bacteriol 154:870–878

Doi Y, Kunioka M, Nakamura Y, Soga K (1988) Nuclear magnetic resonance studies on unusual bacterial copolyesters of 3-hydroxybutyrate and 4-hydroxybutyrate. Macromolecules 21(9): 2722–2727

Doi Y, Segawa A, Kawaguchi Y, Kunioka M (1990) Cyclic nature of poly (3-hydroxyalkanoate) metabolism in *Alcaligenes eutrophus*. FEMS Microbiol Lett 67(1–2):165–169

Doi Y, Kanesawa Y, Tanahashi N, Kumagai Y (1992) Biodegradation of microbial polyesters in the marine environment. Polym Degrad Stab 36(2):173–177

Encarnación S, del Carmen Vargas M, Dunn MF, Dávalos A, Mendoza G, Mora Y, Mora J (2002) AniA regulates reserve polymer accumulation and global protein expression in *Rhizobium etli*. J Bacteriol 184(8):2287–2295

Findlay RH, White DC (1983) Polymeric beta-hydroxyalkanoates from environmental samples and *Bacillus megaterium*. Appl Environ Microbiol 45(1):71–78

Findlay RH, Trexler MB, Guckert JB, White DC (1990) Laboratory study of disturbance in marine sediments: response of a microbial community. Mar Ecol Prog Ser 62:121–133

Flüchter S, Follonier S, Schiel-Bengelsdorf B, Bengelsdorf FR, Zinn M, Dürre P (2019) Anaerobic production of poly (3-hydroxybutyrate) and its precursor 3-hydroxybutyrate from synthesis gas by autotrophic clostridia. Biomacromolecules 20(9):3271–3282

García G, Sosa-Hernández JE, Rodas-Zuluaga LI, Castillo-Zacarías C, Iqbal H, Parra-Saldívar R (2021) Accumulation of PHA in the microalgae *Scenedesmus* sp. under nutrient-deficient conditions. Polymers 13(1):131

Gomez JGC, Rodrigues MFA, Alli RCP, Torres BB, Netto CB, Oliveira MS, Da Silva LF (1996) Evaluation of soil gram-negative bacteria yielding polyhydroxyalkanoic acids from carbohydrates and propionic acid. Appl Microbiol Biotechnol 45(6):785–791

Hai T, Lange D, Rabus R, Steinbüchel A (2004) Polyhydroxyalkanoate (PHA) accumulation in sulfate-reducing bacteria and identification of a class III PHA synthase (PhaEC) in *Desulfococcus multivorans*. Appl Environ Microbiol 70(8):4440–4448

Han J, Hou J, Zhang F, Ai G, Li M, Cai S, Zhao D, Zhou J, Xiang H (2013) Multiple propionyl coenzyme A-supplying pathways for production of the bioplastic poly (3-hydroxybutyrate-*co*-3-hydroxyvalerate) in *Haloferax mediterranei*. Appl Environ Microbiol 79(9):2922–2931

Hänggi U (2018) Virgin PHB has thermoplastic properties, but is not a thermoplast presentation at the PHA congress 2018 in Cologne

Hanik N, Utsunomia C, Arai S, Matsumoto KI, Zinn M (2019) Influence of unusual co-substrates on the biosynthesis of medium-chain-length polyhydroxyalkanoates produced in multistage chemostat. Front Bioeng Biotechnol 7:301

Hany R, Brinkmann M, Ferri D, Hartmann R, Pletscher E, Rentsch D, Zinn M (2009) Crystallization of an aromatic biopolyester. Macromolecules 42:6322–6326

Hartmann R, Hany R, Geiger T, Egli T, Witholt B, Zinn M (2004) Tailored biosynthesis of olefinic medium-chain-length Poly[(*R*)-3-hydroxyalkanoates] in *Pseudomonas putida* GPo1 with improved thermal properties. Macromolecules 37(18):6780–6785

Hartmann R, Hany R, Pletscher E, Ritter A, Witholt B, Zinn M (2006) Tailor-made olefinic medium-chain-length poly[(*R*)-3-hydroxyalkanoates] by *Pseudomonas putida* GPo1: batch versus chemostat production. Biotechnol Bioeng 93(4):737–746

Hazer B, Steinbüchel A (2007) Increased diversification of polyhydroxyalkanoates by modification reactions for industrial and medical applications. Appl Microbiol Biotechnol 74(1):1–12

Hejazi P, Vasheghani-Farahani E, Yamini Y (2003) Supercritical fluid disruption of *Ralstonia eutropha* for poly(β-hydroxybutyrate) recovery. Biotechnol Prog 19(5):1519–1523

Helm J, Wendlandt KD, Jechorek M, Stottmeister U (2008) Potassium deficiency results in accumulation of ultra-high molecular weight poly-β-hydroxybutyrate in a methane-utilizing mixed culture. J Appl Microbiol 105(4):1054–1061

Hermann-Krauss C, Koller M, Muhr A, Fasl H, Stelzer F, Braunegg G (2013) Archaeal production of polyhydroxyalkanoate (PHA) co- and terpolyesters from biodiesel industry-derived by-products. Archaea 2013:129268

Huang TY, Duan KJ, Huang SY, Chen CW (2006) Production of polyhydroxyalkanoates from inexpensive extruded rice bran and starch by *Haloferax mediterranei*. J Ind Microbiol Biotechnol 33(8):701–706

Hyon SH, Jamshidi K, Ikada Y (1984) Polymers as biomaterials. ACS Symp Ser 51:51–65

Jau MH, Yew SP, Toh PS (2005) Biosynthesis and mobilization of poly(3-hydroxybutyrate) [P (3HB)] by *Spirulina platensis*. Int J Biol Macromol 36(3):144–151

Jendrossek D (2009) Polyhydroxyalkanoate granules are complex subcellular organelles (carbonosomes). J Bacteriol 191(10):3195–3202

Jendrossek D, Handrick R (2002) Microbial degradation of polyhydroxyalkanoates. Annu Rev Microbiol 56(1):403–432

Jendrossek D, Pfeiffer D (2014) New insights in the formation of polyhydroxyalkanoate granules (carbonosomes) and novel functions of poly (3-hydroxybutyrate). Environ Microbiol 16(8): 2357–2373

Jost V (2018) Packaging related properties of commercially available biopolymers—an overview of the status quo. Express Polym Lett 12(5):429–435

Kamravamanesh D, Lackner M, Herwig C (2018) Bioprocess engineering aspects of sustainable polyhydroxyalkanoate production in cyanobacteria. Bioengineering 5(4):111

Karmann S, Follonier S, Egger D, Hebel D, Panke S, Zinn M (2017) Tailor-made PAT platform for safe syngas fermentations in batch, fed-batch and chemostat mode with *Rhodospirillum rubrum*. Microb Biotechnol 10(6):1365–1375

Karmann S, Panke S, Zinn M (2019) Fed-batch cultivations of *Rhodospirillum rubrum* under multiple nutrient-limited growth conditions on syngas as a novel option to produce poly (3-hydroxybutyrate) (PHB). Front Bioeng Biotechnol 7:59

Karray F, Abdallah MB, Baccar N, Zaghden H, Sayadi S (2021) Production of poly (3-hydroxybutyrate) by *Haloarcula*, *Halorubrum*, and *Natrinema* haloarchaeal genera using starch as a carbon source. Archaea 2021:8888712

Karthikeyan OP, Mehariya S (2021) Polyhydroxyalkanoates from extremophiles: a review. Bioresour Technol 325:124653

Khosravi-Darani K, Yazdian F, Babapour F, Amirsadeghi AR (2019) Poly(3-hydroxybutyrate) production from natural gas by a methanotroph native bacterium in a bubble column bioreactor. Chem Biochem Eng Q 33(1):69–77

Kim MN, Lee AR, Yoon JS, Chin IJ (2000) Biodegradation of poly(3-hydroxybutyrate), Sky-Green® and Mater-Bi® by fungi isolated from soils. Eur Polym J 36(8):1677–1685

Kim JK, Won YJ, Nikoh N, Nakayama H, Han SH, Kikuchi Y, Ha Rhee Y, Park HY, Kwon JY, Kurokawa K, Dohmae N, Fukatsu T, Lee BL (2013) Polyester synthesis genes associated with stress resistance are involved in an insect-bacterium symbiosis. Proc Natl Acad Sci 110(26): E2381–E2389

Klask C, Raberg M, Heinrich D, Steinbüchel A (2015) Heterologous expression of various PHA synthase genes in *Rhodospirillum rubrum*. Chem Biochem Eng Q 29(2):75–85

Kobayashi G, Shiotani T, Shima Y, Doi Y (1994) Biosynthesis and characterization of poly (3-hydroxybutyrate-co-3-hydroxy-hexanoate) from oils and fats by *Aeromonas* sp. OL-338 and Aeromonas sp. FA440. In: Doi Y, Fukuda K (eds) Biodegradable plastics and polymers. Elsevier, Amsterdam, pp 410–416

Kobayashi D, Fujita K, Nakamura N, Ohno H (2015) A simple recovery process for biodegradable plastics accumulated in cyanobacteria treated with ionic liquids. Appl Microbiol Biotechnol 99(4):1647–1653

Kodama T, Igarashi Y, Minoda Y (1975) Isolation and culture conditions of a bacterium grown on hydrogen and carbon dioxide. Agric Biol Chem 39(1):77–82

Koller M (2018a) Chemical and biochemical engineering approaches in manufacturing polyhydroxyalkanoate (PHA) biopolyesters of tailored structure with focus on the diversity of building blocks. Chem Biochem Eng Q 32(4):413–438

Koller M (2018b) Biodegradable and biocompatible polyhydroxy-alkanoates (PHA): auspicious microbial macromolecules for pharmaceutical and therapeutic applications. Molecules 23(2):362

Koller M (2019a) Switching from petro-plastics to microbial polyhydroxyalkanoates (PHA): the biotechnological escape route of choice out of the plastic predicament? EuroBiotech J 3(1): 32–44

Koller M (2019b) Polyhydroxyalkanoate biosynthesis at the edge of water activity-haloarchaea as biopolyester factories. Bioengineering 6(2):34

Koller M (2020) Established and advanced approaches for recovery of microbial polyhydroxyalkanoate (PHA) biopolyesters from surrounding microbial biomass. EuroBiotech J 4(3):113–126

Koller M, Braunegg G (2015a) Biomediated production of structurally diverse poly (hydroxyalkanoates) from surplus streams of the animal processing industry. Polimery 60(5): P298–P308

Koller M, Braunegg G (2015b) Potential and prospects of continuous polyhydroxyalkanoate (PHA) production. Bioengineering 2(2):94–121

Koller M, Muhr A (2014) Continuous production mode as a viable process-engineering tool for efficient poly(hydroxyalkanoate) (PHA) bio-production. Chem Biochem Eng Q 28(1):65–77

Koller M, Mukherjee A (2020) Polyhydroxyalkanoates—linking properties, applications, and end-of-life options. Chem Biochem Eng Q 34(3):115–129

Koller M, Mukherjee A (2022) A new wave of industrialization of PHA biopolyesters. Bioengineering 9(2):74

Koller M, Hesse P, Bona R, Kutschera C, Atlić A, Braunegg G (2007a) Potential of various archae- and eubacterial strains as industrial polyhydroxyalkanoate producers from whey. Macromol Biosci 7(2):218–226

Koller M, Hesse P, Bona R, Kutschera C, Atlić A, Braunegg G (2007b) Biosynthesis of high quality polyhydroxyalkanoate co- and terpolyesters for potential medical application by the archaeon *Haloferax mediterranei*. Macromol Symp 253(1):33–39

Koller M, Salerno A, Reiterer A, Malli H, Malli K, Kettl KH, Narodoslawsky M, Schnitzer H, Chiellini E, Braunegg G (2012) Sugarcane as feedstock for biomediated polymer production. In: Goncalves JF, Correia KD (eds) Sugarcane: production, cultivation and uses, Agriculture issues and policies. Nova Science, Hauppauge, pp 105–136

Koller M, Niebelschütz H, Braunegg G (2013) Strategies for recovery and purification of poly[(*R*)-3-hydroxyalkanoates] (PHA) biopolyesters from surrounding biomass. Eng Life Sci 13(6):549–562

Kourilova X, Pernicova I, Sedlar K, Musilova J, Sedlacek P, Kalina M, Koller M, Obruca S (2020) Production of polyhydroxyalkanoates (PHA) by a thermophilic strain of *Schlegelella thermodepolymerans* from xylose rich substrates. Bioresour Technol 315:123885

Kourmentza K, Kachrimanidou V, Psaki O, Pateraki C, Ladakis D, Koutinas A (2020) Competitive advantage and market introduction of PHA polymers and potential use of PHA monomers. In: Koller M (ed) The handbook of polyhydroxyalkanoates, Postsynthetic treatment, processing and application, vol 3. CRC Press, Taylor & Francis, Boca Raton, pp 177–201

Kucera D, Novackova I, Pernicova I, Sedlacek P, Obruca S (2019) Biotechnological production of poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate-*co*-3-hydroxyvalerate) terpolymer by *Cupriavidus* sp. DSM 19379. Bioengineering 6(3):74

Kumagai Y, Kanesawa Y, Doi Y (1992) Enzymatic degradation of microbial poly (3-hydroxybutyrate) films. Die Makromolekulare Chemie 193(1):53–57

Kunioka M, Doi Y (1990) Thermal degradation of microbial copolyesters: poly(3-hydroxybutyrate-co-3-hydroxyvalerate) and poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate). Macromolecules 23(7):1933–1936

Kunioka M, Kawaguchi Y, Doi Y (1989) Production of biodegradable copolyesters of 3-hydroxybutyrate and 4-hydroxybutyrate by *Alcaligenes eutrophus*. Appl Microbiol Biotechnol 30(6):569–573

Lee MY, Park WH (2000) Preparation of bacterial copolyesters with improved hydrophilicity by carboxylation. Macromol Chem Phys 201(18):2771–2774

Lee SH, Oh DH, Ahn WS, Lee Y, Choi JI, Lee SY (2000a) Production of poly (3-hydroxybutyrate-*co*-3-hydroxyhexanoate) by high-cell-density cultivation of *Aeromonas hydrophila*. Biotechnol Bioeng 67(2):240–244

Lee MY, Park WH, Lenz RW (2000b) Hydrophilic bacterial polyesters modified with pendant hydroxyl groups. Polymer 41(5):1703–1709

Lee WH, Azizan MNM, Sudesh K (2007) Magnesium affects poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate) content and composition by affecting glucose uptake in *Delftia acidovorans*. Malays J Microbiol 3(1):31–34

Lemoigne M (1923) Production d'acide β-oxybutyrique par certaines bacteries du groupe du *Bacillus subtilis*. Comptes rendus de l'Académie des sciences 176:1761

Lemoigne M (1925) The origin of b-hydroxybutyric acid obtained by bacterial process. Comptes rendus de l'Académie des sciences 180:1539–1541

Lemoigne M (1927) Études sur l'autolyse microbienne origine de l'acide β-oxybutyrique formé par autolyse. Annales de l'Institut Pasteur 41:148

Lemoigne M, Delaporte B, Croson M (1944) Contribution a l'etude botanique et biochimique des bacteries du genre *Bacillus*. II. Valeur du test des lipides β-hydroxylbutyriques pour la caracterisation des especes. Annales de l'Institut Pasteur 70:224–235

Lenz RW, Marchessault RH (2005) Bacterial polyesters: biosynthesis, biodegradable plastics and biotechnology. Biomacromolecules 6(1):1–8

Liu W, Chen GQ (2007) Production and characterization of medium-chain-length polyhydroxyalkanoate with high 3-hydroxytetradecanoate monomer content by fadB and fadA knockout mutant of *Pseudomonas putida* KT2442. Appl Microbiol Biotechnol 76(5): 1153–1159

López NI, Pettinari MJ, Stackebrandt E, Tribelli PM, Pötter M, Steinbüchel A, Méndez BS (2009) *Pseudomonas extremaustralis* sp. nov., a poly(3-hydroxybutyrate) producer isolated from an Antarctic environment. Curr Microbiol 59(5):514–519

Luzier WD (1992) Materials derived from biomass/biodegradable materials. Proc Natl Acad Sci U S A 89(3):839–842

Macrae RM, Wilkinson JF (1958) Poly-β-hyroxybutyrate metabolism in washed suspensions of *Bacillus cereus* and *Bacillus megaterium*. Microbiology 19(1):210–222

Makkar NS, Casida LE Jr (1987) *Cupriavidus necator* gen. nov., sp. nov.; a nonobligate bacterial predator of bacteria in soil. Int J Syst Evol Microbiol 37(4):323–326

Martin DP, Williams SF (2003) Medical applications of poly-4-hydroxybutyrate: a strong flexible absorbable biomaterial. Biochem Eng J 16(2):97–105

Max B, Salgado JM, Rodríguez N, Cortés S, Converti A, Domínguez JM (2010) Biotechnological production of citric acid. Braz J Microbiol 41(4):862–875

Mierziak J, Burgberger M, Wojtasik W (2021) 3-Hydroxybutyrate as a metabolite and a signal molecule regulating processes of living organisms. Biomolecules 11(3):402

Miranda de Sousa Dias M, Koller M, Puppi D, Morelli A, Chiellini F, Braunegg G (2017) Fed-batch synthesis of poly(3-hydroxybutyrate) and poly(3-hydroxybutyrate-co-4-hydroxybutyrate) from sucrose and 4-hydroxybutyrate precursors by *Burkholderia sacchari* strain DSM 17165. Bioengineering 4(2):36

Mitomo H, Hsieh WC, Nishiwaki K, Kasuya K, Doi Y (2001) Poly(3-hydroxybutyrate-co-4-hydroxybutyrate) produced by *Comamonas acidovorans*. Polymer 42(8):3455–3461

Mitra AK (1950) Two new algae from Indian soils. Ann Bot 14(56):457–464

Miyake M, Erata M, Asada Y (1996) A thermophilic cyanobacterium, *Synechococcus* sp. MA19, capable of accumulating poly-β-hydroxybutyrate. J Ferment Bioeng 82(5):512–514

Modi S, Koelling K, Vodovotz Y (2011) Assessment of PHB with varying hydroxyvalerate content for potential packaging applications. Eur Polym J 47(2):179–186

Mohandas SP, Balan L, Gopi J, Anoop BS, Mohan PS, Philip R, Cubelio SS, Singh IB (2021) Biocompatibility of polyhydroxybutyrate-*co*-hydroxyvalerate films generated from *Bacillus cereus* MCCB 281 for medical applications. Int J Biol Macromol 176:244–252

Morikawa H, Marchessault RH (1981) Pyrolysis of bacterial polyalkanoates. Can J Chem 59(15): 2306–2313

Mothes G, Ackermann JU (2005) Synthesis of poly(3-hydroxybutyrate-*co*-4-hydrobutyrate) with a target mole fraction of 4-hydroxybutyric acid units by two-stage continuous cultivation of *Delftia acidovorans* P4a. Eng Life Sci 5(1):58–62

Muhr A, Rechberger EM, Salerno A, Reiterer A, Schiller M, Kwiecień M, Adamus G, Kowalczuk M, Strohmeier K, Schober S, Mittelbach M, Koller M (2013) Biodegradable latexes from animal-derived waste: biosynthesis and characterization of *mcl*-PHA accumulated by *Ps. citronellolis*. React Funct Polym 73(10):1391–1398

Müller-Santos M, Koskimäki JJ, Alves LPS, de Souza EM, Jendrossek D, Pirttilä AM (2020) The protective role of PHB and its degradation products against stress situations in bacteria. FEMS Microbiol Rev 45(3):fuaa058

Murugan P, Gan CY, Sudesh K (2017) Biosynthesis of P(3HB-*co*-3HHx) with improved molecular weights from a mixture of palm olein and fructose by *Cupriavidus necator* Re2058/pCB113. Int J Biol Macromol 102:1112–1119

Nakamura S, Doi Y, Scandola M (1992) Microbial synthesis and characterization of poly (3-hydroxybutyrate-*co*-4-hydroxybutyrate). Macromolecules 25(17):4237–4241

Noda I, Green PR, Satkowski MM, Schechtman LA (2005) Preparation and properties of a novel class of polyhydroxyalkanoate copolymers. Biomacromolecules 6(2):580–586

Noda I, Lindsey SB, Caraway D (2010) Nodax™ class PHA copolymers: their properties and applications. In: Chen GGQ (ed) Plastics from bacteria. Springer, Berlin, pp 237–255

Obruca S, Marova I, Svoboda Z, Mikulikova R (2010) Use of controlled exogenous stress for improvement of poly(3-hydroxybutyrate) production in *Cupriavidus necator*. Folia Microbiol 55(1):17–22

Obruca S, Marova I, Melusova S, Mravcova L (2011) Production of polyhydroxyalkanoates from cheese whey employing *Bacillus megaterium* CCM 2037. Ann Microbiol 61(4):947–953

Obruca S, Sedlacek P, Mravec F, Samek O, Marova I (2016a) Evaluation of 3-hydroxybutyrate as an enzyme-protective agent against heating and oxidative damage and its potential role in stress response of poly(3-hydroxybutyrate) accumulating cells. Appl Microbiol Biotechnol 100(3): 1365–1376

Obruca S, Sedlacek P, Krzyzanek V, Mravec F, Hrubanova K, Samek O, Kucera D, Benesova P, Marova I (2016b) Accumulation of poly(3-hydroxybutyrate) helps bacterial cells to survive freezing. PLoS One 11(6):e0157778

Obruca S, Sedlacek P, Mravec F, Krzyzanek V, Nebesarova J, Samek O, Kucera D, Benesova P, Hrubanova K, Milerova M, Marova I (2017) The presence of PHB granules in cytoplasm protects non-halophilic bacterial cells against the harmful impact of hypertonic environments. N Biotechnol 39:68–80

Obruca S, Sedlacek P, Koller M, Kucera D, Pernicova I (2018) Involvement of polyhydroxyalkanoates in stress resistance of microbial cells: biotechnological consequences and applications. Biotechnol Adv 36(3):856–870

Obruca S, Sedlacek P, Slaninova E, Fritz I, Daffert C, Meixner K, Sedrlova Z, Koller M (2020) Novel unexpected functions of PHA granules. Appl Microbiol Biotechnol 104(11):4795–4810

Obruca S, Sedlacek P, Koller M (2021) The underexplored role of diverse stress factors in microbial biopolymer synthesis. Bioresour Technol 326:124767

Ong SY, Chee JY, Sudesh K (2017) Degradation of polyhydroxyalkanoate (PHA): a review. J Siberian Federal Univ Biol 10(2):211–225

Online resource 1 (n.d.). https://www.satw.ch/fileadmin/user_upload/documents/02_Themen/06_Rohstoffe/Factsheet_Biopolymere__EN.pdf. Last accessed 26 May 2021

Online resource 2 (n.d.). PolyFerm Canada. Last accessed 19 May 2021
Online resource 3 (n.d.). http://www.biomer.de/. Last accessed 19 May 2021
Online resource 4 (n.d.). http://www.tianan-enmat.com/index.html. Last accessed 19 May 2021
Online resource 5 (n.d.). http://www.tjgreenbio.com/en/Product.aspx?cid=54&title=Product%20Description. Last accessed 20 May 2021
Online resource 6 (n.d.). OK Compost Home biodegradable t-shirt bags - Shenzhen Ecomann Biotechnology Co.,Ltd (ecplaza.net). Last accessed 20 May 2021
Online resource 7 (n.d.). https://www.tepha.com/technology/overview/. Last accessed 19 May 2021
Online resource 8 (n.d.). Lepoudre, Erwin Kopie (gbs2018.com). Last accessed 19 May 2021
Online resource 9 (n.d.). https://www.solon.eco/products-and-applications. Last accessed 19 May 2021
Online resource 10 (n.d.). https://danimerscientific.com/. Last accessed 19 May 2021
Online resource 11 (n.d.). Meredian's PHA Ready for Commercial Scale Production - bioplastics MAGAZINE. Last accessed 19 May 2021
Page WJ (1992) Production of polyhydroxyalkanoates by *Azotobacter vinelandii* UWD in beet molasses culture. FEMS Microbiol Rev 9(2–4):149–157
Palleroni NJ, Palleroni AV (1978) *Alcaligenes latus*, a new species of hydrogen-utilizing bacteria. Int J Syst Evol Microbiol 28(3):416–424
Pantazaki AA, Tambaka MG, Langlois V, Guerin P, Kyriakidis DA (2003) Polyhydroxyalkanoate (PHA) biosynthesis in *Thermus thermophilus*: purification and biochemical properties of PHA synthase. Mol Cell Biochem 254(1):173–183
Park WH, Lenz RW, Goodwin S (1998) Epoxidation of bacterial polyesters with unsaturated side chains. III. Crosslinking of epoxidized polymers. J Polym Sci Part A Polym Chem 36(13): 2389–2396
Pérez R, Casal J, Muñoz R, Lebrero R (2019) Polyhydroxyalkanoates production from methane emissions in *Sphagnum* mosses: assessing the effect of temperature and phosphorus limitation. Sci Total Environ 688:684–690
Pernicova I, Novackova I, Sedlacek P, Kourilova X, Kalina M, Kovalcik A, Koller M, Nebesarova J, Krzyzanek V, Hrubanova K, Masilko J, Slaninova E, Obruca S (2020) Introducing the newly isolated bacterium *Aneurinibacillus* sp. H1 as an auspicious thermophilic producer of various polyhydroxyalkanoates (PHA) copolymers-1. Isolation and characterization of the bacterium. Polymers 12(6):1235
Pfeiffer T, Morley A (2014) An evolutionary perspective on the Crabtree effect. Front Mol Biosci 1:17
Pirttijärvi TSM, Ahonen LM, Maunuksela LM, Salkinoja-Salonen MS (1998) *Bacillus cereus* in a whey process. Int J Food Microbiol 44:31–41
Poirier Y, Brumbley SM (2010) Metabolic engineering of plants for the synthesis of polyhydroxyalkanoates. In: Chen G-Q (ed) Plastics from bacteria. Springer, Berlin, pp 187–211. https://doi.org/10.1007/978-3-642-03287-5_8. ISBN: 978-3-642-03287-5 (online), 978-3-642-03286-8 (print)
Pospisilova A, Novackova I, Prikryl R (2021) Isolation of poly(3-hydroxybutyrate) from bacterial biomass using soap made of waste cooking oil. Bioresour Technol 326(124):683
Rehm BH (2003) Polyester synthases: natural catalysts for plastics. Biochem J 376(1):15–33
Rehm BH (2010) Bacterial polymers: biosynthesis, modifications and applications. Nat Rev Microbiol 8(8):578–592
Reusch RN, Huang RP, Bramble LL (1995) Poly-3-hydroxybutyrate/polyphosphate complexes from voltage activated Ca^{2+} channels in the plasma membranes of *Escherichia coli*. Biophys J 69(3):754–766
Riedel SL, Jahns S, Koenig S, Bock MC, Brigham CJ, Bader J, Stahl U (2015) Polyhydroxyalkanoates production with *Ralstonia eutropha* from low quality waste animal fats. J Biotechnol 214:119–127
Rodriguez-Contreras A, Koller M, de Sousa M, Dias M, Calafell M, Braunegg G, Marqués-Calvo MS (2013) Novel poly[(*R*)-3-hydroxybutyrate]-producing bacterium isolated from a Bolivian hypersaline lake. Food Technol Biotechnol 51(1):123–130
Rodríguez-Valera F, Ruiz-Berraquero F, Ramos-Cormenzana A (1980) Behaviour of mixed populations of halophilic bacteria in continuous cultures. Can J Microbiol 26:1259–1263

Rosa DS, Lotto NT, Lopes DR, Guedes CGF (2004) The use of roughness for evaluating the biodegradation of poly-β-(hydroxybutyrate) and poly-β-(hydroxybutyrate-*co*-β-valerate). Polym Test 23(1):3–8

Sadykov MR, Ahn JS, Widhelm TJ, Eckrich VM, Endres JL, Driks A, Rutkowski GE, Wingerd KL, Bayles KW (2017) Poly(3-hydroxybutyrate) fuels the tricarboxylic acid cycle and de novo lipid biosynthesis during *Bacillus anthracis* sporulation. Mol Microbiol 104(5):793–803

Saito Y, Doi Y (1994) Microbial synthesis and properties of poly(3-hydroxybutyrate-*co*-4-hydroxybutyrate) in *Comamonas acidovorans*. Int J Biol Macromol 16(2):99–104

Saito Y, Nakamura S, Hiramitsu M, Doi Y (1996) Microbial synthesis and properties of poly (3-hydroxybutyrate-*co*-4-hydroxybutyrate). Polym Int 39(3):169–174

Saratale RG, Cho SK, Saratale GD, Kadam AA, Ghodake GS, Kumar M, Bharagava RN, Kumar G, Kim DS, Mulla SI, Shin HS (2021) A comprehensive overview and recent advances on polyhydroxyalkanoates (PHA) production using various organic waste streams. Bioresour Technol 124:685

Schatz A, Bovell C Jr (1952) Growth and hydrogenase activity of a new bacterium, *Hydrogenomonas facilis*. J Bacteriol 63(1):87

Sedlacek P, Slaninova E, Enev V, Koller M, Nebesarova J, Marova I, Hrubanova K, Krzyzanek V, Samek O, Obruca S (2019a) What keeps polyhydroxyalkanoates in bacterial cells amorphous? A derivation from stress exposure experiments. Appl Microbiol Biotechnol 103(4):1905–1917

Sedlacek P, Slaninova E, Koller M, Nebesarova J, Marova I, Krzyzanek V, Obruca S (2019b) PHA granules help bacterial cells to preserve cell integrity when exposed to sudden osmotic imbalances. N Biotechnol 49:129–136

Seebach D, Brunner A, Burger HM, Schneider J, Reusch RN (1994) Isolation and ^{1}H-NMR spectroscopic identification of poly(3-hydroxybutanoate) from prokaryotic and eukaryotic organisms—determination of the absolute configuration (*R*) of the monomeric unit 3-hydroxybutanoic acid from *Escherichia coli* and spinach. Eur J Biochem 224:317–328

Seebach D, Brunner A, Burger HM, Reusch RN, Bramble LL (1996) Channel-forming activity of 3-hydroxybutanoic-acid oligomers in planar lipid bilayers. Helv Chim Acta 79:507–517

Simon-Colin C, Alain K, Colin S, Cozien J, Costa B, Guezennec JG, Raguenes GHC (2008) A novel *mcl* PHA-producing bacterium, *Pseudomonas guezennei* sp. nov., isolated from a 'kopara' mat located in Rangiroa, an atoll of French Polynesia. J Appl Microbiol 104:581–586

Singh AK, Sharma L, Mallick N, Mala J (2017) Progress and challenges in producing polyhydroxyalkanoate biopolymers from cyanobacteria. J Appl Phycol 29(3):1213–1232

Slaninova E, Sedlacek P, Mravec F, Mullerova L, Samek O, Koller M, Hesko O, Kucera D, Marova I, Obruca S (2018) Light scattering on PHA granules protects bacterial cells against the harmful effects of UV radiation. Appl Microbiol Biotechnol 102(4):1923–1931

Stigers DJ, Tew GN (2003) Poly(3-hydroxyalkanoate)s functionalized with carboxylic acid groups in the side chain. Biomacromolecules 4(2):193–195

Strong PJ, Laycock B, Mahamud SNS, Jensen PD, Lant PA, Tyson G, Pratt S (2016) The opportunity for high-performance biomaterials from methane. Microorganisms 4(1):11

Tan D, Xue YS, Aibaidula G, Chen GGQ (2011) Unsterile and continuous production of polyhydroxybutyrate *by Halomonas* TD01. Bioresour Technol 102(17):8130–8136

Tarawat S, Incharoensakdi A, Monshupanee T (2020) Cyanobacterial production of poly (3-hydroxybutyrate-*co*-3-hydroxyvalerate) from carbon dioxide or a single organic substrate: improved polymer elongation with an extremely high 3-hydroxyvalerate mole proportion. J Appl Phycol 32(2):1095–1102

Tasaki O, Hiraide A, Shiozaki T, Yamamura H, Ninomiya N, Sugimoto H (1999) The dimer and trimer of 3-hydroxybutyrate oligomer as a precursor of ketone bodies for nutritional care. JPEN J Parenter Enteral Nutr 23(6):321–325

Tsuge T, Hyakutake M, Mizuno K (2015) Class IV polyhydroxyalkanoate (PHA) synthases and PHA-producing *Bacillus*. Appl Microbiol Biotechnol 99(15):6231–6240

Utsunomia C, Ren Q, Zinn M (2020a) Poly(4-hydroxybutyrate): current state and perspectives. Front Bioeng Biotechnol 8:257

Utsunomia C, Hanik N, Zinn M (2020b) Biosynthesis and sequence control of *scl*-PHA and *mcl*-PHA. In: Koller M (ed) The handbook of polyhydroxyalkanoates, Microbial biosynthesis and feedstocks, vol 1. CRC Press, Taylor & Francis, Boca Raton, pp 167–200

Valappil SP, Rai R, Bucke C, Roy I (2008) Polyhydroxyalkanoate biosynthesis in *Bacillus cereus* SPV under varied limiting conditions and an insight into the biosynthetic genes involved. J Appl Microbiol 104(6):1624–1635

Vandamme P, Coenye T (2004) Taxonomy of the genus *Cupriavidus*: a tale of lost and found. Int J Syst Evol Microbiol 54(6):2285–2289

Villanueva L, Navarrete A, Urmeneta J, Geyer R, White DC, Guerrero R (2007) Monitoring diel variations of physiological status and bacterial diversity in an estuarine microbial mat: an integrated biomarker analysis. Microb Ecol 54:523–531

Volova TG, Syrvacheva DA, Zhila NO, Sukovatiy AG (2016) Synthesis of P(3HB-*co*-3HHx) copolymers containing high molar fraction of 3-hydroxyhexanoate monomer by *Cupriavidus eutrophus* B10646. J Chem Technol Biotechnol 91(2):416–425

Vu DH, Åkesson D, Taherzadeh MJ, Ferreira JA (2020) Recycling strategies for polyhydroxyalkanoate-based waste materials: an overview. Bioresour Technol 298:122393

Wallen LL, Davis EN (1972) Biopolymers of activated sludge. Environmen Sci Technol 6(2): 161–164

Wallen LL, Rohwedder WK (1974) Poly-beta-hydroxyalkanoate from activated sludge. Environ Sci Technol 8(6):576–579

Walsh M, O'Connor K, Babu R, Woods T, Kenny S (2015) Plant oils and products of their hydrolysis as substrates for polyhydroxyalkanoate synthesis. Chem Biochem Eng Q 29(2): 123–133

Weibull C (1953) Characterization of the protoplasmic constituents of *Bacillus megaterium*. J Bacteriol 66(6):696

Wen A, Fegan M, Hayward C, Chakraborty S, Sly LI (1999) Phylogenetic relationships among members of the Comamonadaceae, and description of *Delftia acidovorans* (den Dooren de Jong 1926 and Tamaoka et al. 1987) gen. nov., comb. nov. Int J Syst Evol Microbiol 49(2):567–576

Xie CH, Yokota A (2005) Reclassification of *Alcaligenes latus* strains IAM 12599T and IAM 12664 and *Pseudomonas saccharophila* as *Azohydromonas lata* gen. nov., comb. nov., *Azohydromonas australica* sp. nov. and *Pelomonas saccharophila* gen. nov., comb. nov., respectively. Int J Syst Evol Microbiol 55(6):2419–2425

Yasotha K, Aroua MK, Ramachandran KB, Tan IKP (2006) Recovery of medium-chain-length polyhydroxyalkanoates (PHAs) through enzymatic digestion treatments and ultrafiltration. Biochem Eng J 30(3):260–268

Zangirolami TC, Johansen CL, Nielsen J, Jørgensen SB (1997) Simulation of penicillin production in fed-batch cultivations using a morphologically structured model. Biotechnol Bioeng 56(6): 593–604

Zinn M (2010) Biosynthesis of medium-chain-length poly[(R)-3-hydroxyalkanoates]. In: Chen GGQ (ed) Plastics from bacteria. Springer, Berlin, pp 213–236

Zinn M, Witholt B, Egli T (2001) Occurrence, synthesis and medical application of bacterial polyhydroxyalkanoate. Adv Drug Deliv Rev. 53(1):5–21

Zinn M, Weilenmann HU, Hany R, Schmid M, Egli T (2003) Tailored synthesis of poly([*R*]-3-hydroxybutyrate-co-3-hydroxyvalerate) (PHB/HV) in *Ralstonia eutropha* DSM 428. Acta Biotechnol 23:309–316

Zinn M, Lee SY, Chen GGQ (2016) In memoriam of Prof. Bernard Witholt. Biotechnol J 11:195–196

Recent Advances in Poly-(γ-Glutamic Acid) Production by Microbial Fermentation

Sha Li, Yibin Qiu, Hong Xu, Rui Wang, and Peng Lei

Contents

Abstract Poly (γ-glutamic acid) (γ-PGA) is a microbial biopolymer composed of D- and L-glutamic acid monomers connected by γ-amide linkage. γ-PGA is considered a promising biopolymer and has broad applications in food, medicine, agriculture, daily chemicals, and environmental protection due to its water-soluble, biodegradable, and nontoxic properties. The production of γ-PGA has already been established on an industrial scale. This chapter provides updated information about strain breeding, biosynthesis, fermentation, purification, and application of

S. Li (✉) · Y. Qiu · H. Xu · R. Wang · P. Lei
State Key Laboratory of Materials-Oriented Chemical Engineering, College of Food Science and Light Industry, Nanjing University of Technology, Nanjing, China
e-mail: lisha@njtech.edu.cn

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_9

γ-PGA. The biosynthetic mechanism of γ-PGA and regulation of molecular weight by synthetic biology are covered in detail. The current and potential applications of γ-PGA have also been reviewed. Finally, future outlooks of microbial γ-PGA production and application are discussed in recent progress, challenges, and trends in this field. This chapter contributes to the further understanding of efficient production of diversified γ-PGA and will serve as valuable references for reducing the cost of production and further development of commercial-scale applications of γ-PGA.

1 Research Progress of Poly-(γ-Glutamic Acid)-Producing Strains

1.1 Screening and Classification of Poly-(γ-Glutamic Acid)-Producing Strains

Since the 1990s, as a new biopolymer, poly-(γ-glutamic acid) (γ-PGA) has attracted global attention. In 1937, Ivanovics first isolated γ-D-PGA from the capsular component of the disease-causing Gram-positive bacterium *Bacillus anthracis* (Ivanovics and Bruckner 1937). This compositional synthetic γ-D-PGA is embedded in the membrane, which can enhance the virulence of the strain and help bacteria in different environments to enhance the resistance of the cell, to prevent the damage of adverse factors (Schneerson et al. 2003). In 1942, Bovarnick et al. found that the γ-PGA synthesized by *Bacillus subtilis* could be transported to the outside of the cell membrane, which belongs to the secretory mode. Most of the current strains applied in fermentation production belong to this type, which opened a new era of the synthesis of γ-PGA by microbial fermentation (Bovarnick 1942). Until now, the industrial production of γ-PGA mainly depends on microbial synthesis. In addition to screening and identification of γ-PGA-producing strains directly through microbial high viscosity phenotypes and physiological characteristics, it can also carry out high-throughput screening based on the electrostatic interaction between basic dyes (neutral red) and γ-PGA polymers which were reported by Zeng et al. (2013). At present, more and more microorganisms have been reported to be able to synthesize γ-PGA, such as *Bacillus sp.*, *Fusobacterium nucleatum*, archaea and eukaryotes, etc. (Candela et al. 2009; Weber 1989; Hezayen et al. 2001). Among them, the wild-type strains used for γ-PGA production are mainly focus on *Bacillus* sp., including *Bacillus subtilis*, *Bacillus licheniformis*, and *Bacillus amyloliquefaciens*.

Table 1 summarizes the main strains currently used for γ-PGA fermentation. Based on the differences of glutamic acid precursor requirement, γ-PGA-producing strains can be divided into two categories: the exogenous glutamic acid-dependent strain (type I) and the exogenous glutamic acid-independent strain (type II). The concentration of γ-PGA synthesized by glutamic acid-dependent strains was relatively high and accounted for most of the reported strains. These strains require glutamic acid in the culture medium to produce γ-PGA. Generally, the synthesis

Table 1 An overview of different γ-PGA producing strains

Strain	Main components of medium (g/L)	Culture conditions	γ-PGA (g/L)	Stereochemical composition (D:L)	Molecular weight ($\times 10^5$ Da)
Glutamate dependent strain					
B. licheniformis ATCC 9945A (Ko and Gross 1998)	Citric acid 12, glycerol 80, L-glutamic acid 20, NH_4Cl 7	37 °C 2–3 days	5–20.5	(44–85): (55–15)	2–8
B. Licheniformis NCIM2324 (Bajaj and Singhal 2009b)	Glycerol 62.4, citric acid 15.2, L-glutamic acid 20, ammonium sulfate 8	37 °C 96 h	35.75	–	–
B. subtilis NX-2 (Zhang et al. 2012c; Wu et al. 2006b)	Cane molasses 60, MGWL 40, $(NH_4)_2SO_4$ 5	32 °C 96 h	52.1	77:23	14–24
B. subtilis IFO 3335 (Kunioka 1997)	Citric acid 20–50, L-glutamic acid 30, $(NH_4)_2SO_4$ 5–10	37 °C 2 days	10–20	80:20	10
B. subtilis F-2-01 (Kuboat et al. 1993)	Glucose 80, peptone 15, yeast extract 5, urea 3, L-glutamic acid 70	37 °C 2–3 days	25–50	69:31	5
B. subtilis chungkookjang (Ashiuchi et al. 2001a)	Sucrose 50, L-glutamic acid 20, $(NH_4)_2SO_4$ 20	30 °C 5 days	13.5–16.5	–	–
B. subtilis MR-141 (Ogawa et al. 1997)	Maltose 60, soy sause 70, sodium L-glutamate 30	40 °C 3–4 days	35	–	–
Glutamate independent strain					
B. subtilis TAM-4 (Ito et al. 1996)	Glucose 75, NH_4Cl 18	37 °C 4 days	20	78:22	2
B. subtilis C1(Shih et al. 2005)	Citric acid 22, glycerol 170, NH_4Cl 7	37 °C 6 days	21.4	97:3	7.3–7.9
B. Licheniformis A35 (Cheng et al. 1989)	Glucose 75, NH_4Cl 18	30 °C 3–5 days	8–12	(50–80): (50–20)	–
B. Amyloliquefaciens LL3(Cao et al. 2011)	Sucrose 50, $(NH_4)2SO_4$ 2	37 °C 48 h	4.36	1.53:98.47	4.7

(continued)

Table 1 (continued)

Strain	Main components of medium (g/L)	Culture conditions	γ-PGA (g/L)	Stereochemical composition (D:L)	Molecular weight ($\times 10^5$ Da)
B. Amyloliquefaciens NX-2S(Qiu et al. 2017; Sha et al. 2019c)	Inulin 60, $(NH_4)2SO_4$ 6	32 °C 60 h	6.85	(60–85): (40–15)	14

efficiency is high, and the product concentration can reach 20–50 g/L. However, due to the high cost caused by the addition of a large amount of glutamate in the medium, γ-PGA is still faced with the limitation of application cost in some fields, such as agriculture and feed.

Glutamate-independent production strain is a new research hotspot in the field of γ-PGA microbiosynthesis because it does not need to add extra glutamate, which greatly reduces the fermentation cost. Zhang et al. isolated a glutamate-independent γ-PGA producer *Bacillus subtilis* C10 from sauce products, which could use glucose as a carbon source and produce 3.73 g/L of γ-PGA (Zhang et al. 2012a). *Bacillus amyloliquefaciens* LL3, an independent glutamate-producing bacterium isolated from fermented food, was obtained 4.36 g/L of γ-PGA in a 200 L fermentor when used sucrose as the substrate (Cao et al. 2011). Peng et al. isolated a strain of *Bacillus methylotrophicus* SK19.001 using glycerol as carbon source from the soil; the γ-PGA production reached 14 g/L without the addition of amino acid precursor (Peng et al. 2015). A novel strain named *Bacillus amyloliquefaciens* NX-2S was isolated from soil samples of Jerusalem artichoke tubers by Qiu et al. The strain preferred to assimilate inulin as the sole carbon source and could produce 6.85 g/L of γ-PGA without glutamic acid addition, which provides a new strategy for the γ-PGA production from low-cost renewable non-food resources (Qiu et al. 2017). However, the γ-PGA productivity of these strains is low, and there has been no report for industrial production and application so far.

The mechanism of substrate glutamate dependence between the two types of γ-PGA-producing strains above has also been studied. Genomic analysis of glutamate-dependent strains *Bacillus subtilis* GXA-28 and glutamate-independent strains *Bacillus subtilis* GXA-5 showed that the genes related to the differences in glutamate dependence were mainly involved in sugar transport metabolism and amino acid metabolism and 13 genes related to γ-PGA biosynthesis (Zeng et al. 2017). In order to study the mechanism of glutamate dependence of *Bacillus subtilis* NX-2, Sha et al. conducted transcriptome analysis of cultured strains with or without addition of glutamate. Turns out that the overexpression of *gltA*, *gltB*, *putM*, and *rocA* genes can obviously promote the accumulation of γ-PGA, suggesting that intracellular glutamate synthesis plays a key role in the regulation of γ-PGA production in glutamate-dependent strains (Sha et al. 2019a).

1.2 Mutagenic Breeding of γ-PGA-Producing Strains

In order to select highly efficient γ-PGA-producing strains to meet the needs of industrial production, traditional breeding techniques have been successfully applied to improve the production efficiency of γ-PGA. Using atmospheric and room temperature plasma (ARTP), Qiu et al. successfully improved the γ-PGA productivity of *Bacillus amyloliquefaciens* NX-2S. Compared with the wild-type strain, the γ-PGA yield of the mutant strain NX-2S154 increased by 58% when used inulin crude extract as the substrate (Qiu et al. 2019). By UV mutagenesis, directed evolution and three rounds of genome shuffling, Zhang et al. obtained a γ-PGA high-yielding strain C2 with high glucose tolerance. Compared with the parent strain W14, the mutant strain increased the γ-PGA production by 3.5 times and the biomass by 2.3 times, respectively (Zhang et al. 2017). However, due to the blindness of traditional mutagenesis breeding and the lack of a fast and effective high-throughput quantitative detection method for γ-PGA concentration in the fermentation broth, the traditional breeding process of γ-PGA strains is still very inefficient.

1.3 Construction and Metabolic Regulation of Efficient Engineered γ-PGA Strains

1.3.1 Research Progress on Engineering of Wild-Strain γ-PGA-Producing Strains

With the development of molecular biology and the establishment of genetic manipulation platform in *Bacillus* species, it has become a new focus on improving γ-PGA production by using genetic engineering technology. At present, genetic engineering of the wild-type strains to enhance the γ-PGA synthesis mainly revolves around the following aspects:

1. *Enhance the substrate utilization pathway*: Qiu et al. used the self-developed CRISPR-Cas9 nickase genomic trace-free editing system to modify the substrate inulin hydrolysis module of *B. amyloliquefaciens* strain by combining the inulinase action module and successfully realized the conversion of substrate inulin and the increase of γ-PGA production (Qiu et al. 2020). Feng et al. transformed the PTS system of the original strain into a non-PTS system in *B. amyloliquefaciens* LL3 and constructed an energy-saving sucrose metabolism pathway to reduce the energy consumption of carbon source metabolism, ultimately achieving a 38.5% increase in γ-PGA production (Feng et al. 2017). For metabolic regulation by improving utilization rate of industry by-product glycerol in *B. licheniformis* WX-02, Zhan et al. used the type of promoter (P43), *ytzE* promoter (*PytzE*), and bacABC operon promoter (*PbacA*) to replace natural *glpFK* promoter, and the glycerol consumption in corresponding mutant strains WX02 P43glpFK, WX02 PytzEglpFK, and WX02-PbacAglpFK is improved by

30.9%, 26.42%, and 18.8%, respectively. And the γ-PGA concentration of the three mutant strains was 33.71%, 23.39%, and 30.05% higher than that of the initial strain WX-02, respectively (Zhan et al. 2017).

2. *Enhance the precursor glutamate synthesis pathway*: Glutamate is a precursor of γ-PGA synthesis. In *B. amyloliquefaciens* LL3, Feng et al. knocked out the glutamate-degrading enzyme gene *rocG*, *gudB*, and glutamate-inhibiting protein gene *rocR*. Results showed that both *rocG* and *gudB* knockout could increase γ-PGA production by about 38% (Zhang et al. 2015). Furthermore, the glutamate utilization pathway was inhibited by synthetic sRNA against glutamate-degrading gene *rocG* and glutamine synthetase *glnA*. The results showed that the yield of anti-*glnA* sRNA expressing strain NK-anti-*glnA* and NK-anti-*glnA*-*rocG* γ-PGA decreased by 55.9% and 44.3%, respectively, compared with the control NK-E11. Compared with the control NK-E11 strain, the γ-PGA yield of anti-*rocG* sRNA was increased by 18.5% to 11.04 g/L, and the final γ-PGA yield reached 20.3 g/L by batch fermentation of NK-anti-*rocG* strain (Feng et al. 2015). In *B. amyloliquefaciens* NBCSO, Sha et al. optimized γ-PGA synthetic precursor pathway by expression of relevant key genes (citrate synthase *citA*, glutamate synthetase gltA, proline dehydrogenase *ycgM*, and Δ1-pyrrolin-5-carboxylic acid dehydrogenase *ycgN*). At the same time, the shunt pathway of α-ketoglutarate and glutamic acid was weakened, and the final yield of γ-PGA was increased to 22.62 ± 0.41 g/L. These results indicate that glutamate precursors play an important role in further increasing the concentration of γ-PGA (Sha et al. 2020a).
3. *Overexpression of γ-PGA synthase and knockdown of γ-PGA degradation pathway*: Ashiuchi et al. deleted the γ-PGA synthase gene *pgsBCA* by using the pKPSD plasmid in *B. subtilis* ISW 1214 and then replaced *pgsBCA* under the control of the xylose-inducible promoter. The genetically engineered strain could produce a large amount of γ-PGA in both L-glutamate- and D-glutamate-rich medium (Ashiuchi et al. 2006). Yeh et al. introduced a synthetic expression control sequence (SECS) upstream of *pgsBCA* gene in *B. subtilis* DB430 strain to produce 28 g/L γ-PGA in glutamate-free medium (Yeh et al. 2010). In most γ-PGA-producing strains, there existed specific γ-PGA-degrading enzymes which are responsible for the degradation of γ-PGA. The presence of degrading enzyme PgdS can degrade and reduce the γ-PGA production. Therefore, researchers began to try to increase γ-PGA production by knocking out the degrading enzyme gene. Mitsui et al. studied the effects of the knocking out of γ-PGA exonuclease gene *ggt*, γ-PGA endonuclease gene *pgdS*, and DL-polypeptide endonuclease family gene on γ-PGA synthesis in *B. subtilis* (natto). The results showed that only the *cwlo* gene knockout strain of DL-endopeptidase increased the γ-PGA production by two times compared to the original strain (Mitsui et al. 2011). Scoffone et al. found that double knockout of pgdS and ggt resulted in a twofold increase in γ-PGA production in the strain compared to the *Bacillus subtilis* strain (Scoffone et al. 2013), while Kimura et al. found that ggt single gene knockout had no effect on γ-PGA synthesis (Kimura et al. 2004a).
4. *Removing the by-product synthesis pathway*: Most γ-PGA synthetic strains, in addition to the synthesis of γ-PGA, will also synthesize a large number of EPS polysaccharide, levan (fructose oligosaccharides), and other by-products. The

existence of these by-products will not only affect the production of γ-PGA by competition of substrate and energy but also interfere with the subsequent separation and purification of γ-PGA. In *B. amyloliquefaciens* C06, the production of γ-PGA was increased from 3.2 g/L to 6.8 g/L by deleting the polysaccharide biosynthetic gene *epsA* (Liu et al. 2011). Feng et al. studied the effects of knockout the gene clusters *itu*, *bae*, *srf*, and *fen* encoding four antibiotic substances synthesis genes on the synthesis of γ-PGA in *B. amyloliquefaciens* LL3. The γ-PGA yield from the strain with double knockout of *itu* and *srf* gene clusters was increased from 3.3 to 4.5 g/L (Gao et al. 2016).

5. *Enhanced synthesis of ATP and NADHP cofactors*: In *B. licheniformis* WX-02, overexpression of the glucose-6-phosphate dehydrogenase gene *zwf* increased the activity of *zwf* by 9.28 times, thus increasing the production of NADPH and reducing the accumulation of acetone and 2,3-butanediol by products. Finally, the maximum concentration of γ-PGA reached 9.13 g/L, which was 35% higher than that of the original strain (Cai et al. 2017). The supply of ATP plays a vital role in the biosynthesis of γ-PGA. Cai et al. engineered the electron respiration chain by deleting the cytochrome bd oxidase branch, which increased the γ-PGA production by 19.27% in *Bacillus licheniformis*. Furthermore, the ATP content was increased to 3.53 μmol/g DCW by overexpression of ATP biosynthesis gene *adK*, *resD* gene, and *Vitreoscilla* hemoglobin gene *vgb* gene, and finally the γ-PGA production of mutant strain WX-BCVAR was 38.64% higher than that of wild-type strain WX-02, reaching 43.81 g/L (Cai et al. 2018).

1.3.2 Application of Synthetic Biology Techniques in γ-PGA-Producing Strains

A great deal of work has been done in γ-PGA high-yielding wild strains. With the rapid development of synthetic biology techniques, some researchers also try to use new model hosts for heterogeneous synthesis of γ-PGA. Ashiuchi et al. heterologously expressed γ-PGA synthase pgsBCA in *Escherichia coli* and obtained recombinant strains that yield γ-PGA less than 1 g/L (Ashiuchi et al. 1999a). In order to provide sufficient D-glutamate substrate donors, Cao et al. co-express γ-PGA synthase *pgsBCA* and glutamate racemase *racE* and finally obtained only 0.645 ± 0.016 g/L of γ-PGA (Cao et al. 2013). After optimizing the expression of γ-PGA synthase by constitutive and inductive regulation, the content of γ-PGA in the recombinant strain reached 3.7 g/L after batch feeding (Jiang et al. 2006). In conclusion, the yield of heterogeneous production of γ-PGA in *Escherichia coli* is low.

Researchers explore a new chassis for γ-PGA production. *Corynebacterium glutamicum* can produce high concentrations of glutamate precursor and save the addition of exogenous glutamate, which has been considered as an ideal host for γ-PGA production. The researchers introduced the γ-PGA synthase into *Corynebacterium glutamicum* 13,032 and obtained 0.5 g/L of γ-PGA (Cao et al. 2010). Xu et al. optimized and induced the expression of γ-PGA synthase gene in a high-yield glutamate-producing strain of glutamate, *C. glutamicum* F343; 11.4 g/L of γ-PGA

was obtained in the engineered strain. By further introduction of the glutamate racemase gene *racE* from *Bacillus subtilis*, the yield of γ-PGA (2000–4000 kDa) reached to 21.3 g/L. This is the highest reported yield for the synthesis of γ-PGA by genetically engineered bacteria so far (Xu et al. 2019a).

γ-PGA is a kind of ultrahigh-molecular-weight biological polymer, which makes it has good plasticity in the degree of polymerization. Like most polysaccharides, polypeptides, and other biological polymers, γ-PGA in different degrees of polymerization tend to present different physical and chemical properties and biological activity. With the continuous expansion of γ-PGA application field, the demand for specific molecular weight biopolymer products has been increasingly urgent. Therefore, it is a valuable research direction to find a way to synthesize biopolymers with a controlled molecular weight of γ-PGA. At present, industrial regulation of the molecular weight of γ-PGA depends on acid-base hydrolysis, poor control, and pollution to the environment. So, how to achieve the efficient biosynthesis of different molecular weights γ-PGA is an urgent problem to be solved. The analysis of the synthesis mechanism of γ-PGA polymerase is still in the preliminary stage. The structure characterization of the complex enzyme has not been successfully obtained, due to inability to grasp its enzymatic properties and catalytic mechanism. Therefore, the regulation mechanism of γ-PGA polymerase on the molecular weight of γ-PGA has not been clearly understood, and the regulation of the molecular weight of γ-PGA by γ-PGA polymerase needs further study in the future.

A specific degrading enzyme gene *pgdS/ywtD* was also found on the γ-PGA synthase gene cluster responsible for the degradation of γ-PGA, which provides a new idea for regulating the molecular weight of γ-PGA products. At present, the research on γ-PGA-degrading enzymes is mainly focused on the production of γ-PGA bacteria. According to enzyme cleavage mechanism, it can be divided into endonucleated and exonucleated degrading enzymes. Suzuki and Tahara cloned a gene encoding a γ-PGA-degrading enzyme, named *ywtD*, from strain *B. subtilis* IFO 16449 (Suzuki and Tahara 2003). The gene located on downstream of the γ-PGA synthase gene cluster, *ywtABC*, is partially in line with the gene encoding D/L-endonuclease, a glycopeptide-degrading enzyme. Heterologous expression of the gene *ywtD* in *E. coli* further confirmed that the enzyme has γ-PGA degradation enzyme activity and only specific hydrolysis of the γ-glutamic acid bond between D-glutamic acid and L-glutamic acid in γ-PGA. The resulting hydrolysates were high-molecular-weight γ-PGA (490 kDa, containing 100% L-glutamic acid) and low-molecular-weight γ-PGA (11 kDa, containing 80% D-glutamic acid and 20% L-glutamic acid). Ashiuchi et al. studied the γ-PGA-degrading enzyme gene *pgdS* from *B. subtilis* (Chungkookjang) and found that its sequence was basically the same as that of *ywtD* gene, which was an endonucleotide degrading enzyme. When different configurations of γ-PGA were used as the substrate for hydrolysis, it was found that the γ-PGA specifically degraded the γ-glutamyl bond between D- and L-glutamic acid. The above results indicated that different *B. subtilis* sources of γ-PGA degradation enzyme hydrolysis substrate have different configurations (Ashiuchi and Misono 2002a). Yao et al. expressed the *ywtD* gene of γ-PGA-degrading enzyme derived from *B. subtilis* NX-2 in *E. coli* and studied its enzymatic properties.

γ-PGA with molecular weight distribution in the range of 20–1000 kDa was obtained through enzymatic hydrolysis experiment (Yao et al. 2009). Kimura et al. identified γ-glutamyl transferase GGT from *B. subtilis* NAFM5, which has exonuclease hydrolytic activity. At the later stage of fermentation, γ-PGA is hydrolyzed from the amino terminal to generate free D- and L-glutamate monomers, which can be used as nutrient components for the growth of bacteria (Kimura and Itoh 2003). King et al. studied the hydrolysis characteristics of γ-PGA-degrading enzyme from *B. licheniformis* ATCC 9945A and found that the enzyme was tightly bound to γ-PGA and could be activated by metal ions Zn^{2+} and Ca^{2+}, belonging to the endonucleated γ-PGA-degrading enzyme (King et al. 2001).

In *Bacillus anthracis*, the gene *capD*-encoding γ-PGA-degrading enzyme is not only participates in the hydrolysis of γ-PGA but also is responsible for binding γ-PGA to the peptidoglycan on the cell wall. Unlike other γ-PGA-degrading enzymes, *capD* is immobilized on the surface of spore for subsequent catalytic reactions. DNA sequence alignment showed that *capD* has low homology with *B. subtilis*-degrading enzyme YWTD (<15%) and high homology with γ-glutamyl transferase GGT (Candela and Fouet 2005).

γ-PGA-degrading enzyme activity was also found in some microorganisms that did not synthesize γ-PGA. Tanaka et al. isolated and purified a γ-PGA hydrolase from the fermentation liquid of Myrothecium sp.TM-4222. The enzyme has endonuclease activity to γ-PGA and can hydrolyze the γ-amide bond between L-glutamic acid and L-glutamic acid (Tanaka et al. 1993). In addition, degrading enzymes with γ-PGA exonuclease activity were found in the culture medium of *Flavobacterium* and *Micromonospora melanosporea (*Volcani and Margalith 1957*;* Muro et al. 1990*)*. Although there are some reports about γ-PGA-degrading enzymes at present, there is still a lack of commercial γ-PGA-degrading enzymes with high enzyme activity and good stability, and the production cost is high by separating and purifying the enzyme for in vitro enzymatic hydrolysis of γ-PGA. In the process of γ-PGA fermentation, the establishment of "side formation-side degradation" process through rational regulation of degradation enzymes can realize the substrate economy and efficiently obtain γ-PGA with desirable molecular weights, which lays the foundation for industrial application. Sha et al. screened γ-PGA-degrading enzymes from different bacterial sources in *B. amyloliquefaciens* NB and found that the γ-PGA-degrading enzyme PgdS divided from *B. subtilis* NX-2 was the most suitable for the synthesis of low-molecular-weight γ-PGA (LMW-γ-PGA) by the NB strain (Sha et al. 2019b).

For further implementation of γ-PGA synthesis with various molecular weight control, the Xu group optimized heterogenous expression of γ-PGA-degrading enzyme in the pre-built γ-PGA efficient synthesis *B. amyloliquefaciens* cell factory and revealed the correlation between the expression level of NX-2-derived degradation enzyme PgdS and the molecular weight of γ-PGA. On this basis, the CRISPRi system was further designed to regulate the expression level of PgdS, and the dynamic regulation of PgdS was realized by the strategy of multiple sgRNA combinations. Different molecular weights of γ-PGA were obtained in a strain with high (>800 kDa)-, medium (400–600 kDa)-, and low (50–100 kDa)-

molecular-weight, and the yield reached 25–27 g/L (Sha et al. 2020a). In addition, the molecular weight of γ-PGA corresponds to the activity of γ-PGA-degrading enzyme, and the stereochemistry of γ-PGA also affects the subsequent molecular weight regulation. In previous reports, glutamate racemase was confirmed to be involved in the regulation of this process (Halmschlag et al. 2019). In order to achieve the regulation of the stereochemical configuration of γ-PGA, Sha et al. further regulated the stereochemical configuration of γ-PGA by using the racemic glutamate gene *racE* and γ-PGA synthase gene *pgsA* by coupling γ-PGA-degrading enzyme, realizing that the current industrial production accuracy is difficult to meet the low molecular weight (<10 kDa) γ-PGA (Sha et al. 2020b).

2 Occurrence and Biosynthetic Mechanism of γ-PGA

In recent years, some progress has been made in studies on the metabolic pathway and key enzymes of γ-PGA. For the mechanism of γ-PGA synthesis, the metabolic pathway and related enzymes of γ-PGA synthesis were also different in different strains. Ogawa et al. studied the relationship between glutamate supply of γ-PGA and its precursors, indicating that glutamate required by microorganisms for production of γ-PGA came from their own metabolism or the external environment (Ogawa et al. 1997). At present, it is generally believed that glutamate-independent strains provide L-glutamate monomer through their own metabolic pathways. However, for glutamate-dependent strains, there are different ways to obtain L-glutamate monomer supply; γ-PGA monomer is not all sourced from the conversion of exogenous glutamate.

The supply of exogenous glutamate also has different effects. Briefly, the biosynthesis pathways of γ-PGA can be roughly defined as follows. One is the synthesis of glutamic acid monomers, including glycolysis, tricarboxylic acid cycle, and glutamic acid synthesis pathways. The second is the polymerization pathway of glutamic acid monomer to generate γ-PGA (Fig. 1). Wu et al. analyzed the pathway of γ-PGA glutamate-dependent synthesis by ^{13}C isotope tracer method in a γ-PGA strain *B. subtilis* NX-2, and the results showed that glucose synthesized glutamate precursors through glycolysis, pentose phosphate pathway, tricarboxylic acid cycle, and other pathways, and then the synthesized glutamate precursors were polymerized to γ-PGA. When glucose concentration was 40 g/L in the medium, about 9% of the γ-PGA carbon skeleton was entered by glucose. These results indicated that glucose was mainly used as a growth-limiting substrate for the growth and energy metabolism, which was only a small part of the carbon skeleton of γ-PGA molecules. However, L-glutamic acid was the main source of the carbon skeleton of γ-PGA molecules (Wu et al. 2008).

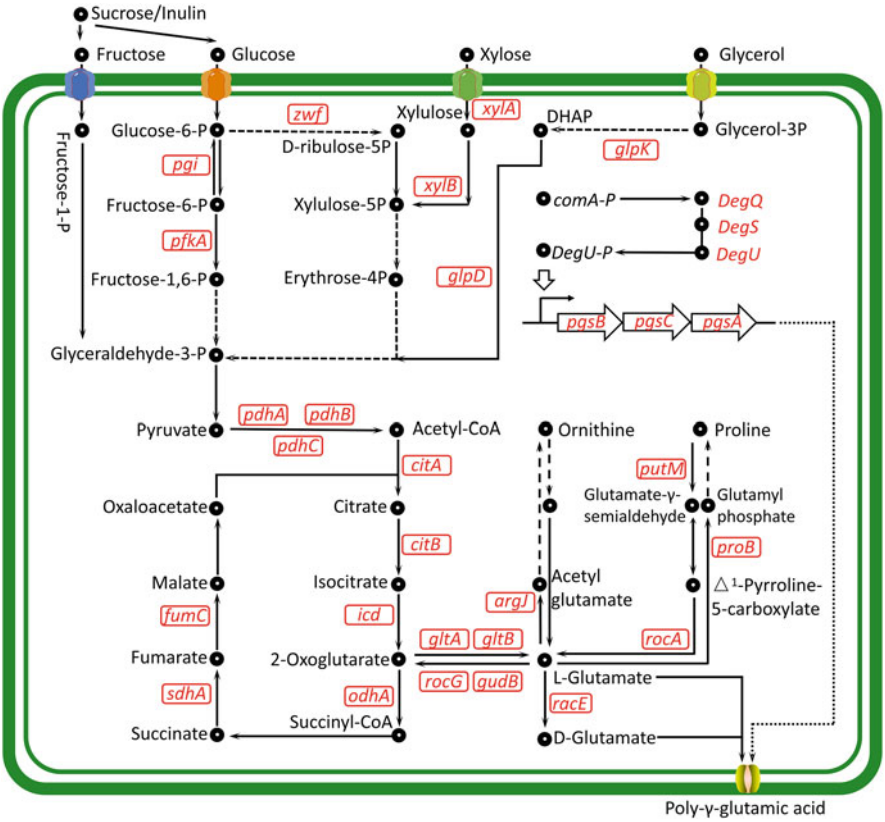


Fig. 1 Schematic diagram of γ-PGA biosynthesis pathway with different substrates

2.1 *Synthesis of γ-PGA Precursors*

2.1.1 Synthesis of L-Glutamate

It is well-known that glutamic acid is the precursor of γ-PGA synthesis. L-glutamic acid is mainly synthesized by anabolism and catabolism in *Bacillus subtilis*. Usually, one molecule of L-glutamine and one molecule of α-ketoglutarate generate two molecules of L-glutamic acid in the coupled reaction system of L-glutamine synthase and L-glutamic synthase (Belitsky et al. 2000). In the presence of glutamine, L-glutamate is formed from α-ketoglutarate through the action of glutamate dehydrogenase. In addition, L-aspartic acid transaminase can also convert L-aspartic acid and α-ketoglutarate to oxaloacetic acid and L-glutamic acid. L-arginine and L-proline can form L-glutamic acid through the catabolic pathway (Belitsky and Sonenshein 1998). Since γ-PGA-producing bacteria can be classified as glutamate-dependent and non-glutamate-dependent, the study of intracellular L-glutamate synthesis is of

great significance for the understanding of the synthesis mechanism in γ-PGA-producing strains.

2.1.2 Synthesis of D-Glutamate

For the synthesis of D-glutamate, there are two main confirmed pathways, involving D-aminotransaminase and glutamate racemase. The activity of D-aminotransaminase was higher in γ-D-PGA or γ-DL-PGA-producing bacteria with higher proportion of D-type monomers. The researchers initially detected D-aminotransaminase activity in *B. anthracis* and *B. licheniformis* ATCC 9945A (Ashiuchi and Misono 2002b). Although it is generally believed that D-glutamate is catalyzed by D-aminotransferase, the activity of this enzyme was not detected in the cell lysate of *B. subtilis* IFO3336, and only the highly active glutamate racemase was detected (Ashiuchi et al. 1998). D-aminotransaminase and glutamate racemase activity were studied in *B. subtilis* subsp. chungkookjang, and the results show that the D-aminotransaminase and glutamate racemase coexist in logarithmic phase. During the stable phase, the activity of D-aminotransaminase decreased sharply, while the activity of glutamate racemase increased (Ashiuchi et al. 2001a). The results indicated that D-aminotransaminase in *B. subtilis* subsp. chungkookjang probably did not participate in the supply of D-glutamic acid in γ-PGA production and D-glutamic acid was all catalyzed by glutamate racemase.

In general, the activity of glutamate racemase in most bacteria is very low, such as glutamate racemase Mur I in *E. coli*. If the expression of this enzyme is increased, it will lead to mitotic aberration and cell proliferation inhibition (Balikó and Venetianer 1993). Obviously, the glutamate racemic enzyme in *B. subtilis* has different properties. This enzyme has high catalytic efficiency and low affinity for D-glutamate, which is very important for maintaining high concentration of D-glutamate when L-glutamate is accumulated in large amounts in the cell. After purifying the glutamate racemase protein (RacE) (Ashiuchi et al. 1998), Ashiuchi et al. isolated and purified the RacE isoenzyme YrpC (Ashiuchi et al. 1999b). YrpC is similar to RacE in its primary structure, substrate specificity, cofactor independence, and other enzymatic characteristics, but its kinetic parameters are completely different. YrpC is similar to the glutamate racemic enzyme Mur I in *E. coli*, where overexpression of this enzyme also inhibits bacterial growth. Unlike RacE, YrpC has low catalytic efficiency and high affinity for L-glutamate, which is an advantage for providing the limited D-glutamate necessary for cell growth at low intracellular concentrations of L-glutamate. When knockout of gene *racE*, the growth of the cells does not need to add additional D-glutamic acid. Furthermore, for strains with double knockout of *racE* and *yrpC*, it was found exogenous D-glutamic acid must be added to maintain the normal growth of cells. The relationship between *yrpC* and *racE* is complementary, which also indicates that in *B. subtilis*, glutamate racemase provides D-glutamate for the synthesis of γ-PGA (Kimura et al. 2004b).

2.2 Polymerization of γ-PGA Precursors

The γ-PGA synthase gene has been of great interest to the γ-PGA research groups for many years. With the research and development, many strains of γ-PGA synthase operons have been cloned and verified. Ashiuchi and team (Belitsky et al. 2000; Candela and Fouet 2006) cloned and verified γ-PGA synthase genes for the first time, including *pgsB*, *pgsC*, and *pgsA* genes from *B. subtilis* natto chromosome genome (Ashiuchi et al. 2001a; Ashiuchi et al. 2001b). The *pgsBCA* knockout strain could not synthesize γ-PGA, but all other traits were the same as those of the wild strain. The results showed that *pgsBCA* was necessary for the synthesis of γ-PGA. Using newly synthesized protein kinetics data from an in vitro transcription and translation system, the researchers summarized the functions of individual proteins of γ-PGA synthase. The results showed that the binding between *pgsB* and *pgsC* was very tight, while the binding between *pgsBC* and *pgsA* was relatively loose, and there was little difference in the affinity between *pgsBC* and *pgsBCA* for glutamate. However, due to the instability of γ-PGA synthase, the researchers were unable to isolate and purify the enzyme, although the function of each protein in the γ-PGA polymerase system can be elucidated through some experimental data: *pgsB* is considered to be the main catalytic protein in the γ-PGA polymerase system. The structural characteristics of an amide ligase were found in this enzyme, and the previously identified ligases were all cytoplasmic enzymes. This is the first example of a membrane ligase. In addition, it was found that *pgsB* can catalyze the hydrolysis of ATP in the presence of glutamic acid monomer, providing energy for the extension of γ-PGA peptide chain (Kimura et al. 2009; Tomosho et al. 2008).

The N-acetyltransferase domain similar to that of N-acetylglutamate synthase was found in *pgsC*. Interestingly, three series similar N-acetyltransferase domains have been found in the C-terminal of ε-PL synthase, and this domain is important for the enzyme-catalyzed reaction (Yamanaka et al. 2008). At present, *pgsC* has been found only in the strains that can synthesize γ-PGA, indicating that *pgsC* plays an important role in γ-PGA synthesis (Vetting et al. 2005; Min et al. 2009).

There are membrane-anchoring regions in pgsA that are responsible for positioning the pgsBCA complex on the membrane (Ashiuchi and Misono 2002b). Homologous sequences of this protein are found in a variety of biological genes, belonging to the cytosolute protein serine/threonine phosphatase with divalent cation binding sites, including Zn^{2+}, Mn^{2+}, Fe^{2+}, and Ca^{2+}. PgsA may be responsible for the extracellular transport of γ-PGA (Nordlund and Eklund 1995; Rusnak and Mertz 2000). In addition, PgsA had high homology with the isomerase MslH, and the strain could synthesize γ-PGA with high L-glutamate ratio after knockout of *pgsA* gene in *B. subtilis*. Therefore, it is speculated that PgsA may affect the configuration of glutamic acid monomer during γ-PGA synthesis (Ogasawara et al. 2019; Sawada et al. 2018).

PgsE exists downstream of the *pgsBCA* gene, but its function remains unclear. The researchers hypothesized that this protein is functionally similar to the CapE protein in *B. anthracis*. Recent studies by Ashiuchi and Yamashiro et al. have shown

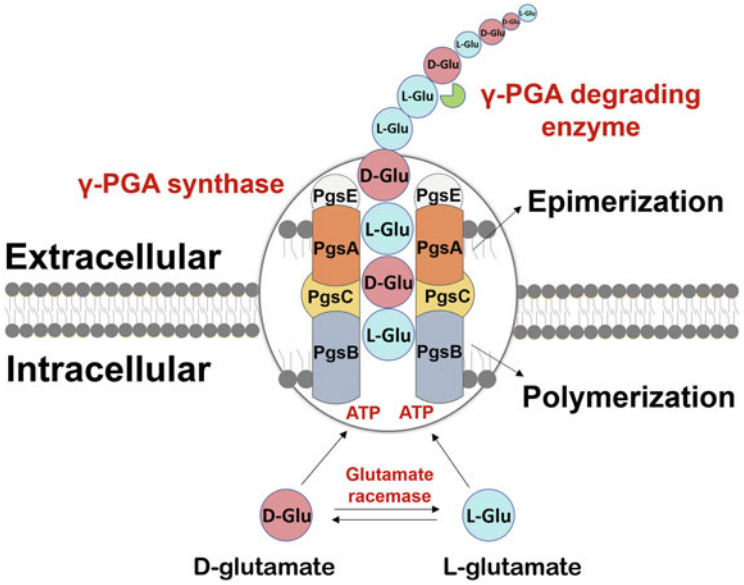


Fig. 2 Schematic representation of γ-PGA biosynthesis by membrane-bound proteins. An illustration depicts the gamma-PGA biosynthesis. Substrate glutamate isomerized under the action of glutamate racemase (RacE). D-glutamate and L-glutamate plus ATP in the presence of gamma-PGA synthase with the help of PGSA, PGSB, PGSC, PGSE forms a biopolymer gamma-PGA which is secreted into cell culture medium. In this process, there are still gamma-PGA degrading enzyme that can degrade gamma-PGA

that this protein plays an important role in some plasmid-containing γ-PGA-producing bacteria (Ashiuchi and Yamashiro 2009). Recent studies have reported that the expression of PgsE can promote the increase of the molecular weight and yield of γ-PGA. PgsE was similar to the N-terminus of the key protein responsible for the assembly of ubiquitin synthetase. It is speculated that PgsE may stabilize the assembly of the PgsBCA complex, thus promoting the increase of the molecular weight and yield of γ-PGA. In the future, the biological functions and related catalytic mechanisms of PgsE remain to be further explored (Fujita et al. 2021). The catalysis of γ-PGA synthesis by these membrane-bound proteins has not been fully understood. Candela et al. concluded that γ-PGA synthase must perform two functions: polymerization of γ-PGA and transport of γ-PGA. According to the function of each protein, a γ-PGA synthesis model based on PgsB, PgsC, PgsA, and PgsE polymerase systems was also established (Candela and Fouet 2006) (Fig. 2).

3 Fermentation Engineering for γ-PGA Production

3.1 Fermentation Medium for Producing γ-PGA

3.1.1 The Effect of Carbon Source on γ-PGA Production

As one of the nutrients necessary for the growth of microorganisms, carbon source plays an important role in the growth and development of microorganisms. It is a carbon nutrient used by microorganisms to build the carbon skeleton of bacteria and metabolites and to provide energy for life. In the process of microbial fermentation to produce γ-PGA, it commonly used glycerol, fructose, lactose, galactose, glucose, sucrose, maltose, and starch as the carbon sources. Different strains require different carbon sources, and even those using the same carbon source, the amount of the carbon source would be different. Ju et al. found that starch is the best carbon source for the fermentation of *B. subtilis* MJ80 to produce γ-PGA (Ju et al. 2014). Shi et al. used the response surface method to study the fermentation process of γ-PGA by *B. subtilis* ZJU-7 and found that using sucrose as a carbon source can maximize the yield of γ-PGA as high as 58.2 g/L (Shi et al. 2006). Jung et al. used *B. subtilis* RKY3 to produce γ-PGA; glycerol was the best carbon source (Jung et al. 2005). Jiang et al. found that when glucose is used as the carbon source for the fermentation of *B. subtilis* NX-2 to produce γ-PGA, the maximum yield of γ-PGA can reach 70.9 g/L (Jiang et al. 2016).

3.1.2 The Effect of Nitrogen Source on γ-PGA Production

As an indispensable component in the growth and reproduction process of microorganisms, nitrogen sources can provide nitrogen nutrition for cell life activities. Generally, the available nitrogen sources of microorganisms are divided into organic nitrogen sources and inorganic nitrogen sources. Organic nitrogen sources are beef extract, casein peptone, malt extract, yeast extract, etc., while inorganic nitrogen sources include urea, NH_4Cl, NH_4NO_3, $(NH_4)_2SO_4$, etc. Different microbial strains have different utilization abilities to these two types of nitrogen sources. Ju et al. found that *B. subtilis* MJ80 has a strong ability to utilize inorganic nitrogen source urea and can use urea as the best nitrogen source (Ju et al. 2014). Shi et al. used tryptone and cheap corn flour as nitrogen source to produce γ-PGA by *B. subtilis* ZJU-7 (Shi et al. 2006).

3.1.3 The Effect of Metal Ions on the Yield of γ-PGA

In the process of γ-PGA production by microbial fermentation, K^+, Mg^{2+}, Na+, Mn^{2+}, and other metal ions have the effect on γ-PGA yield. Ju et al. found that K_2HPO_4 and $MnSO_4$ added to the fermentation of *Bacillus* RKY3 influenced the

yield of γ-PGA produced. K_2HPO_4 can promote the growth of bacterial cells, while Mn^{2+} can promote the synthesis of γ-PGA. Ashiuchi et al. found that adding 0.5% NaCl to the fermentation system of strain *B. subtilis* (chungkookjang) can promote the production of γ-PGA per unit volume to reach 13.5 g/L (Ashiuchi et al. 2001a).

3.2 Fermentation Factors for Producing γ-PGA

Temperature, pH, dissolved oxygen, stirring method, and other factors of the fermentation system can all affect the final yield of γ-PGA. Here we mainly discuss dissolved oxygen. The production of γ-PGA by microbial fermentation is generally carried out under aerobic conditions. Oxygen is one of the necessary conditions for the production of γ-PGA. Generally, oxygen can affect the biosynthesis and energy metabolism of microbial cells during fermentation, and the regulation of dissolved oxygen directly affects the yield of γ-PGA.

There are many studies on optimizing dissolved oxygen. Su et al. inserted the *vgb* (*Vitreoscilla* hemoglobin gene) into the genome of the γ-PGA-producing strain *B. subtilis* S18–3. It can increase the ability of *B. subtilis* S18–3 to carry more oxygen to achieve the purpose of increasing the dissolved oxygen in the fermentation system (Su et al. 2010). Qun et al. adopted batch fermentation method and changed the stirring speed in stages; within 24 h in the fermentation process, the fermenter speed was adjusted to 600 rpm and dropped to 400 rpm after 24 h. Compared with the previous experimental method that did not change the speed (maintained at 400 rpm), the method significantly increased the oxygen content in the fermentation system during the first 24 hours of cell growth (Wu et al. 2006a). Xu et al. developed a method to increase the oxygen mass transfer coefficient by adding oxygen carriers to the fermentation system, which increased the yield of γ-PGA to 39.4 ± 0.19 g/L (Zhang et al. 2012b). Our team designed and used an aerobic plant cellulose bed experiment to increase the amount of dissolved oxygen in the fermentation system and ultimately increase the yield of γ-PGA (Feng et al. 2016).

3.3 Fermentation Method for Producing γ-PGA

At present, the production of γ-PGA by microbial fermentation mainly includes liquid fermentation and solid fermentation. Among them, the most widely used fermentation method is submerged fermentation. Bajaj et al. reported that *B. licheniformis* ATCC 9945 contained grains (Bajaj and Singhal 2011). The yield of γ-PGA could reach up to 23 g/L in 48 h in the medium that contains acetic acid, glycerol and citric acid. The yield of γ-PGA increased to 35 g/L when glutamic acid was added in the medium (Bajaj and Singhal 2011). When glutamic acid and citric

acid were used as the fermentation substrates for *B. subtilis* IFO 3335 growth, the yield of γ-PGA was only 20 g/L (Bajaj and Singhal 2011).

B. licheniformis CCRC12826, *B. licheniformis* WBL-3, *B. subtilis* R23, *B. subtilis* NX-2, and *B. subtilis* TAM-4 are also the primary strains used for submerged fermentation production of γ-PGA (Table 1). However, the production of γ-PGA by submerged fermentation is costly and time-consuming. As γ-PGA is a highly viscous, in the later stage of submerged fermentation, the fermentation broth usually becomes extremely viscous, resulting in insufficient dissolved oxygen, which in turn affects the growth and metabolism of the bacteria and ultimately affects the yield of γ-PGA. Therefore, in order to solve the above problems, some researchers have utilized solid-state fermentation method to produce γ-PGA due to its couple of advantages, such as low energy consumption and low pollution. In addition, solid-state fermentation can also use solid waste produced in industry and agriculture as the fermentation substrate. This fermentation method can not only increase the yield of γ-PGA but can also be used for solid waste utilization hence promoting environmental protection (Liu et al. 2013). At present, the strains *B. subtilis* CCTCC 202048, *B. subtilis* B6–1, *B. subtilis* ME714, *B. licheniformis* NCIM 2324, *B. subtilis* natto, *B. subtilis* NX-2, *B. subtilis* 3–10, and *B. licheniformis* C1 that can be used to produce γ-PGA by solid-state fermentation have been reported (Table 1) (Junnan et al. 2018).

4 Separation and Purification of γ-PGA

For high-end fields, such as food, cosmetics, and medical materials, high-grade γ-PGA products are often needed to meet the requirements of industry standards. To obtain such high-quality product, separation and purification of γ-PGA from fermentation broth are required. However, most of the studies on the fermentation production of γ-PGA have focused on the screening of strains, the optimization of medium and culture conditions, and the improvement of γ-PGA yields. The extraction and purification of γ-PGA from fermentation broth research is very few.

For most fermentation products, removal of the bacteria from the culture media is the first step for subsequent purification. In general, filtration and centrifugation are the common methods. However, it is not trivial to separate the cells from the fermentation broth containing γ-PGA product. γ-PGA is usually firmly attached to the surfaces of γ-PGA-producing strains as an extracellular capsule in the early stages of growth (Ogunleye et al. 2015). These γ-PGA-coated cells show electronegativity due to the ionization effect under neutral pH conditions, making the cells stable due to the electrostatic repulsion, which in turn makes it difficult to directly achieve the solid-liquid separation through filtration or centrifugation without pre-treatment (Do et al. 2001). Because γ-PGA is a polymer with a relatively high molecular weight, the fermentation broth tends to be very viscous, which slows down the sedimentation rate of centrifugation and therefore requires higher centrifugal force for bacterial separation (Shih and Van 2001).

Do et al. (2001) found that acidification of fermentation broth before centrifugation could significantly improve the process efficiency and energy consumption. By lowering the pH of the fermentation broth to 3, the viscosity and the ξ potential value of the cells of the fermentation broth with 2.5% γ-PGA at 35 °C were reduced to 1/6 and 1/3 of the original value, respectively. After acidification, the bacteria can be removed by centrifugation under 5000 × *g* centrifugal force for 30 mins, while 22,000 × *g* for 60 min was required if acidification is not performed. Due to the simple process and low cost, acidification has also become the main step to obtain cell-free clear liquid from γ-PGA fermentation broth in industry. The commonly used acidification reagents include inorganic acids, including trichloroacetic acid (TCA), HCl, and H_2SO_4. The acidification degree is generally adjusted to about 2–3 pH of fermentation broth.

The next step after obtaining the cell-free broth is to extract the γ-PGA from the solution, which there are three main strategies that are alcohol sedimentation, metal ion-induced sedimentation, and quaternary ammonium salt sedimentation.

Recovery of γ-PGA from cell-free broth by alcohol precipitation is the most commonly used technology (Goto and Kunioka 1992). γ-PGA is heteropeptide in structure, whose molecular status in aqueous solution is similar to that of protein, DNA, polysaccharide, and other water-soluble polymers. Adding intermediate polar organic solvents, such as ethanol, to γ-PGA solution can reduce the activity of water molecules, reduce the dielectric constant of the solution, and repel water molecules from the γ-PGA around. Through the interaction of polar groups, γ-PGA molecules coagulated and precipitated under the action of van der Waals force. The solvents used for γ-PGA precipitation recovery have been reported as ethanol, methanol, and propanol, among which ethanol is the most widely used due to its cheapness, easy availability, and high safety (Do et al. 2001; Goto and Kunioka 1992; Birrer et al. 1994; Kubota et al. 1993). For cell-free broth with 1–2% γ-PGA content, ethanol with a final concentration of 75–80% is usually used. The higher the concentration of γ-PGA, the lesser the amount of ethanol can be used. Therefore, the cost of ethanol can be reduced by concentrating the γ-PGA broth. However, using alcohol as a precipitator has several disadvantages. Firstly, alcohol will lead to the coprecipitation of proteins and nucleic acids in the broth, which would reduce the purity of the product. Secondly, alcohol precipitation method is not enough to recover γ-PGA completely, usually resulting in the loss of 15%. Finally, the use of large quantities of organic solvents may face environmental regulatory issues (Manocha and Margaritis 2010). Therefore, despite the easy operation of alcohol precipitation, many researchers are still searching for more efficient and low-cost γ-PGA recovery processes.

Manocha and Margaritis (2010) proposed a method to recover γ-PGA from cell-free fermentation broth based on metal-ion affinity. Studies have found that γ-PGA side chain is rich of carboxyl groups that can chelate with many metal cations, such as Cu^{2+}, Mn^{2+}, Al^{3+}, Cr^{3+}, Fe^{3+}, etc. When γ-PGA dissolves, its long chain is extended due to the repulsion force that form the negatively charged carboxyl groups. When large metal cations enter the solution, the negative carboxyl groups become neutral due to the chelation, and the long-chain molecules lose the

electrostatic repulsion, form flocculates with the bound metal ions, and then settle down. This is also the main mechanism of γ-PGA as a flocculant in the field of water treatment (Salehizadeh and Shojaosadati 2001). Manocha and Margaritis (Manocha and Margaritis 2010) make use of this principle. The cell-free fermentation broth was treated with divalent copper ($CuSO_4$), resulting in γ-PGA precipitation, which was collected as precipitation by centrifugation. The precipitate was then dissolved and dialyzed with deionized water to obtain pure γ-PGA. The study compared the recovery efficiency and selectivity of γ-PGA with ethanol precipitation method. It was found that 85% of the γ-PGA in the broth was recovered by the $CuSO_4$ method and 82% by the ethanol method. In addition, ethanol method detected 48% of the protein in the broth, while $CuSO_4$ method detected only 3% of the protein in the final purified PGA, indicating that later method had better selectivity than former method. Interestingly, the $CuSO_4$ method did not result in the residual Cu in the purified γ-PGA. Therefore, copper-induced precipitation of γ-PGA offers an effective recovery technique that can selectively precipitate γ-PGA from the fermentation broth without the use of expensive organic solvents.

Liu et al. (2018) also proposed a scheme for γ-PGA recovery using Gemini quaternary ammonium salt-induced precipitation. This strategy is similar in principle to the metal ion-induced γ-PGA deposition, which uses quaternary ammonium cation to bind the carboxylic acid group of γ-PGA forming water-insoluble complexes. A Gemini quaternary ammonium salt called DOBP 10 (1,4-bis-[3,3-′-(1-decylpyridinium) methyloxy] butane dibromide) was used in the study. 10% DOBP 10 was added to the fourfold diluted cell-free fermentation broth to produce complex precipitation. The precipitate was dissolved in ethanol again, and then 30% NaCl was added to the solution to separate DOBP 10 and precipitate γ-PGA. The recovery rate of γ-PGA was more than 90%. However, compared with alcohol and metal deposition, this process is more complicated.

5 Applications of γ-Polyglutamic Acid

5.1 Agricultural Planting

In recent years, γ-PGA, as a plant biostimulant, has shown great success in the field of agriculture, especially in China. In 1994, Donlar Corporation reported for the first time that γ-PGA could improve the absorption of nitrogen, phosphorus, and potassium by plants and promoted crop growth (Kinnersley et al. 1994). In 2000, Hara (2000) coated the modified γ-PGA on the surface of seeds and spread it on arid sandy land for germination experiments. It was found that seeds could germinate and grow normally, while the control could not. Therefore, γ-PGA is believed to have a good application prospect in barren mountains and desert reconstruction. Wang et al. (Wang et al. 2008) found that γ-PGA had effects on biological control and fertilizer efficiency enhancement. Xu et al. (2014) reported that γ-PGA can improve the nitrogen use efficiency of wheat, rape, and other crops and promote crop yield

increase and believed that γ-PGA can regulate the key enzyme activity of plant nitrogen metabolism through Ca^{2+}/CaM signaling pathway, thus improving the absorption and assimilation of nitrogen and thus promoting plant growth. Lei et al. (2015) studied the effects of γ-PGA on the growth, physiological, and biochemical characteristics of rapeseed seedlings under low temperature (4 °C). The results showed that the fresh weight and chlorophyll content of rapeseed seedlings induced by γ-PGA increased by 24.5% and 50.9%, respectively, after 144 h of stress. Similarly, under NaCl stress, γ-PGA significantly increased K^+/Na^+ value, proline content, and antioxidant enzyme activity of rape seedlings (Lei et al. 2016). The content of malondialdehyde was significantly reduced, indicating that it enhanced the tolerance of rape seedlings to salt stress. Guo et al. (2017) also found the same effect on wheat.

Due to the polycarboxylic acid properties, γ-PGA also has a potential role in mitigating plant heavy metal toxicity. Pang et al. (2018) found that under Pb and Cd stresses, γ-PGA could effectively alleviate the growth inhibition of cucumber seedlings by heavy metals, and with the increase of γ-PGA concentration, the contents of chlorophyll a and b also increased. After γ-PGA treatment, the contents of Pb and Cd in cucumber seedlings decreased by 74.13% and 38.65%, respectively. What is more interesting is that γ-PGA simultaneously mitigates the adverse effects of Cd and Pb on soil microbial community.

Lei and Xu et al. systematically studied the mechanism of γ-PGA in assisting abiotic stress in plant environment (Lei et al. 2017; Xu et al. 2016; Xu et al. 2017a). Fluorescent labeling showed that γ-PGA did not enter the cytoplasm but instead attached to the surface of root protoplasm. Here, it triggered a burst of H_2O_2 in roots by enhancing the transcription of RbohD and RbohF, and the elicited H_2O_2 further activated an influx of Ca^{2+} into root cells. Ca^{2+} signaling was transmitted via the stem from roots to leaves, where it elicited a fresh burst of H_2O_2 by Ca^{2+}-binding proteins CBL9, CPK4, and CPK5. The H_2O_2 signal promoted brassinolide and jasmonic acid biosynthesis by upregulating key genes (*dwf4* and *lox2*, respectively) for synthesizing these compounds. Lastly, brassinolide and jasmonic acid increased H_2O_2 which promoted proline accumulation and total antioxidant capacity (T-AOC) improvement, resulting in improving the resilience of plants. Unfortunately, some questions remain unanswered. Is there a γ-PGA-specific binding protein on the cell membrane? If so, how does the receptor protein bind to γ-PGA and mediate the production of these signals? These problems will be the last barrier to solve the relationship between γ-PGA and plant stress resistance in the future.

5.2 Food

γ-PGA, also known as natto bacteria gum, was first found in "natto" – fermented beans. Thus, it is food-borne substances with food safety. It has been widely used in food additives, antifreeze agents, dietary supplements, etc.

5.2.1 Food Additives

Food quality is often determined by flavor and shelf life. γ-PGA of high molecular weight can effectively embed the functional components of food, which reduce the damage of harmful factors in food processing and transportation (e.g., temperature, bacteria, etc.). Besides, it has a promoting effect on the stability and improvement of food quality. Previous studies have shown that the addition of γ-PGA to one or more bitter substances (e.g., amino acids, peptides, quinine, caffeine, minerals, etc.) can significantly reduce bitterness (Katsuragi et al. 1998).

γ-PGA can be added in the production of starch foods (mainly baked goods and noodles) to improve quality and extend shelf life. The addition of γ-PGA in the production of wheat bread can improve the water retention ability of bread and significantly improve the rheological properties and thermal properties of wheat dough. After adding γ-PGA, the gelatinization temperature of flour increased to further expand the dough. In the storage process, the addition of PGA makes the wheat bread soft, that is, the bread hardness is reduced, thus effectively delaying the bread aging (Shyu et al. 2008).

γ-PGA can also be used to modify emulsified food and improve its stability. Shyu and Sung (2010) studied the effects of different concentrations of γ-PGA (0.05, 0.1, and 0.5 g/kg w/w) on the viscosity, foam stability, and emulsification properties of sponge cake paste. The viscosity, foam stability, and emulsion stability of the cake were significantly improved when 0.5 g/kg γ-PGA was added. The cross-linking of triglyceride with the polyhydroxyl groups of γ-PGA enhances the stability of the emulsion. During the storage process, γ-PGA reduces the hardness and mastic ability of sponge cake, which maintains the cohesion of sponge cake, and has the effect of preventing cake spoilage.

Further γ-PGA function is its utility as a healthy functional oil reducer in fried food. Lim et al. (Lim et al. 2012) studied the effects of γ-PGA on the oil absorption and water loss of donuts during frying. The results showed that γ-PGA improved the water retention ability of donuts. After frying for 4 min, the oil absorption of the control donut was about 0.7 g/g dough, while the oil absorption of the PGA treated donut was about 0.2 g/g dough. When the PGA concentration was increased from 0.25 g/100 g of dough to 1 g/100 g of dough, the amount of oil absorbed by the doughnut was reduced by a factor of five.

5.2.2 Antifreeze Agents

In addition to being used as a food additive, γ-PGA of low molecular weight is an excellent antifreeze agent. It can effectively inhibit the growth of ice crystals and reduce the amount of water that can be frozen in food, which avoid the destruction of food structure by ice crystals (Jia et al. 2019). The antifreeze activity of γ-PGA in the molecular weight range of less than 20,000 Da is higher than that of glucose, which is called a highly antifreeze substance. Its antifreeze activity tends to decrease with

the increase of molecular weight (Mitsuiki et al. 1998). In addition, compared with the commonly used small molecule antifreeze agents, such as glucose and inorganic salts, γ-PGA of low molecular weight has less impact on food quality due to its D-glutamic acid component which has a lighter taste (Park et al. 2005).

The survival ability of probiotics can be also improved by adding γ-PGA in the process of freeze-drying. Bhat et al. (2013) found that the protective effect of 10% γ-PGA on *Lactobacillus casei* was significantly better than that of 10% sucrose in the freeze-drying process. Compared with trehalose, sorbitol, and NATA (bacterial cellulose produced by *Acetobacter xylostae*), γ-PGA can provide better protection for lactic acid bacteria during freeze-drying.

5.2.3 Food Nutritions

γ-PGA is also a good dietary fiber, which is usually processed into health products. It can increase gastrointestinal peristalsis, remove the body garbage, as well as maintain the human digestive. Additionally, its final decomposition product is free glutamic acid, which is beneficial to human health (Wang et al. 2016). As γ-PGA contains γ-amide bond, it can reduce the sensitivity of the gastrointestinal enzyme and is often used as a nutritional aid. It can promote the absorption of calcium ions in the human body by increasing calcium solubility and intestinal absorption; thus it can be used for treatment towards osteoporosis. The increase in calcium solubility is attributed to the large number of anions in γ-PGA which inhibit the production of insoluble calcium phosphate. In addition, γ-PGA also promotes intestinal calcium absorption by increasing passive calcium transport in the small intestine (Tanimoto et al. 2001). Tanimoto et al. (2007) found that postmenopausal women who took a single dose of γ-PGA had a better intestinal calcium absorption and individuals with a lower basal absorption capacity than average benefited more from γ-PGA consumption. In recent years, some researchers have added γ-PGA to the animal feed to promote the absorption and utilization of minerals in poultry and improve the quality of meat.

5.3 Daily Chemical Products

γ-PGA is widely used for cosmetics because of its super water absorption, moisture retention, and cell affinity. Since the application process of cosmetics involves the cellular osmotic absorption of molecules, its application effect is closely related to molecular weight. Generally, γ-PGA of low molecular weight can be absorbed by the skin and achieve deep moisturizing effect, and, on the other hand, γ-PGA of high molecular weight can form a membrane structure on the skin surface conducive to protecting skin moisture (Xu et al. 2017b).

γ-PGA, also named natto gum (nattogum) in the cosmetics pharmacopoeia international name, is often added to facial masks, morning and night cream, and

other skincare products as highly effective moisturizing ingredients. It plays a significant role in maintaining epidermal moisture and reducing the corneous layer of water loss. When applied in shampooing and hair care products, γ-PGA can lock the moisture of hair surface and hair scales, maintaining hair quality, nourishing scalp, as well as reducing dandruff generation and hair boring, etc.

When used in skincare ointment, γ-PGA can reduce skin allergy, roughness, and skin damage caused by lack of water, restore cell function, and improve skin immunity and metabolism (Wang et al. 2016).

5.3.1 Antioxidant

γ-PGA of low molecular weight often has excellent antioxidant activity. Wang et al. (2020) examined γ-PGA having different molecular weights for removing ultra-oxygen anion free radical ($O_2^{\cdot-}$), hydroxyl radical (·OH), and 1,1-diphenyl-2-trinitrobenzene hydrazine (DPPH). The results showed that all the γ-PGA indicated a certain degree of radical scavenging activity with the clearance of $O_2^{\cdot-}$ ·OH, and DPPH reached 31.36%, 97.55%, and 74.64% respectively. In addition, by building a UV-induced mice skin fibroblasts (L929) oxidative damage light aging model, it was observed that γ-PGA with different molecular weights could repair the damage of photoaging cells to a certain extent and reduce the increase of oxygen free radical (ROS) and nitrogen free radical (NO) in the photoaging cells with the γ-PGA having a molecular weight of 3.0×10^5 Da had the best effect.

γ-PGA can also be used as an active component of hyaluronidase inhibitors to maintain skin elasticity and improve skin allergy by inhibiting the permeability of inflammatory cells (Sung et al. 2014). γ-PGA maintains skin moisture and elasticity by inhibiting the activity of hyaluronidase, an enzyme that degrades hyaluronic acid in the dermis. γ-PGA have also been widely used in the production of wet wipes, baby diapers, and other sanitary products (Zhang 2020).

5.3.2 Skin Protection

γ-PGA of high molecular weight has outstanding moisture retention and water absorption, forming a protective film on the skin. It can inhibit the loss of skin moisture and penetrate the deep skin, restoring the skin self-moisturizing system (Liu et al. 2015). Based on γ-PGA as the precursor material of the hydrogel, Wang et al. (2019) constructed a new double-network hydrogel film material ("the second skin ") similar to epidermal skin tissue through covalent cross-linking and molecular self-assembly technology which combined with the polyphenolic structural molecule tannic acid (TA). The hydrogel demonstrated broad-spectrum UV resistance in the range of 275–360 nm, excellent bioadhesion and water resistance (under sweating and dynamic physiological conditions), without invading skin tissues. In addition, when the surface structure of the hydrogel is damaged, it can realize perfect self-repair function within 1 min. On the one hand, the excellent performance is due

to the γ-PGA containing many carboxylic acid groups in molecular structure, enhancing the skin cell affinity adaptation ability. On the other hand, the double network formed by the TA molecular assembly with the hydrogel membrane on the skin surface to construct a stable 3D network structure interact with the skin surface to form multiple, strong integration and intelligent self-healing effect.

5.4 Tissue Engineering, Regenerative Medicine, and Drug Delivery

γ-PGA, as a natural polyamino acid produced by microbial fermentation, has great water solubility, biocompatibility, and degradability. Besides, γ-PGA contains many active carboxyl groups in the molecular structure, which are conducive to functional modification. Therefore, γ-PGA has been widely used in tissue engineering scaffolds, regenerative medicine, and drug carriers.

5.4.1 Tissue Engineering and Regenerative Medicine Materials

Because of the excellent water solubility, γ-PGA has been used to develop hydrogel materials in recent years. Studies have shown that wound of human tissue under wet conditions has a faster repair rate and the reduced rate of scar relative to the dry wound. Therefore, the bionic materials prepared with γ-PGA mainly as the wound dressings and hemostatic materials, which are similar with human tissue (such as skin, cartilage, etc.) in their structure and functions, have shown rapid development. Based on γ-PGA as the primary material, Xu et al. constructed a series of wound repair materials by means of enzymatic in situ catalysis (Chen et al. 2017) and chemical cross-linking (Xu et al. 2019b) via molecular modification. Among them, γ-PGA-based injectable hydrogel material had shown excellent function of promoting wound repair and regeneration, wound healing, and hemostasis with cell compatibility. Furthermore, γ-PGA nanofiber materials based on microfluidic spinning technology could effectively promote wound repair and reduce scar generation. Sun et al. (2020) modified the double bond of the γ-PGA and composited with antibacterial polylysine through photopolymerization using visible light to develop the antibacterial and biocompatible hydrogel, achieving the repair and regeneration of skin infections. Hsieh et al. (2005) prepared a γ-PGA/chitosan composite biomaterial and demonstrated that the composite had better hydrophilicity, biocompatibility, and mechanical properties as compared to the conventional chitosan matrix, which was a promising tissue engineering biomaterial. γ-PGA/chitosan composite biomaterial could also be used as wound dressing. Tsao et al. (2011) designed a chitosan/γ-PGA polyelectrolyte complex (PEC) as an easily peeled material to treat wounds. The results showed that chitosan/γ-PGA PECs contain a suitable moisture content and exhibited good mechanical properties, both of which were beneficial for

the wound dressing as it can be easily detached from the wound surface without damaging the newly regenerated tissue.

5.4.2 Drug Carrier

γ-PGA is also an ideal drug carrier material, mainly because of the large amount of highly active carboxyl groups in its molecular structure which is suitable for functional modification. Most anticancer drugs (e.g., paclitaxel (TXL), camptothecin (CPT), etc.) are water-insoluble in clinical practice, which greatly limits the clinical application. Moreover, the increase of drug dose not only increases the cost but also increases the drug toxicity. The efficient utilization of controlled/sustained release and targeted delivery of drugs by using water-soluble γ-PGA molecules carrying hydrophobic drugs have attracted extensive attention of drug delivery researchers. Li et al. (1998) developed a water-soluble complex by covalent bonding of paclitaxel drug to γ-PGA (PG-TXL), which demonstrated a more significant antitumor efficiency than TXL. Clinical experiments showed that the uptake of PG-TXL by tumor cells was about five times higher than that of regular TXL at the same dose. The side effects were reduced, with the antitumor activity and the maximum tolerated dose (MTD) were significantly improved in animal models. Besides, PG-TXL has unique tumor-targeting capabilities. The polymer was targeted into the cell and able to decompose to deliver high doses of paclitaxel to the tumor, increasing the antitumor efficiency. Bhatt et al. (2002) formed γ-PGA complex with camptothecin drug (CPT-PGA) by combining the carboxyl group of the γ-PGA with the functional group (20S-hydroxyl) of camptothecin. This way the water solubility and therapeutic effect of the drug can be significantly improved.

γ-PGA can also be used in combination with other anticancer drugs to reduce toxicity. Cisplatin anticancer drugs (CDDP) have been widely used to treat a variety of cancers. However, the inhibition of bone marrow growth, severe nephrotoxicity, and other defects limit the clinical application of CDDP. Zhang et al. (2018) constructed PGA-Asp-maleimide-cisplatin-peptide complex (PAMCP). The results showed that PAMCP reduced the toxicity of CDDP in vitro and in vivo with the increasing antitumor efficiency.

5.5 Environmental Protection

Industrial and agricultural sewage has caused a severe burden to the environment, damaging animal and plant health. At present, the flocculation sedimentation method has been frequently used to achieve efficient sewage treatment. There are many kinds of flocculants, the quality of which directly affects the treatment effect. γ-PGA is a biodegradable polymer material. Compared with traditional flocculants, γ-PGA is environmentally friendly with outstanding water absorption, adhesion, and adsorption bridging effect, which has a broad application prospect.

Flocculation is a commonly used and effective method to remove suspended solids and metal ions in wastewater treatment (Deng et al. 2003). γ-PGA with relatively large molecular weight (more than 1.5 million) can form colloids in water, which has adsorption bridging effect and biodegradability. It can be used as an environment-friendly flocculant to replace the traditional synthetic flocculant in wastewater pollution treatment. Bajaj and Singhal (2009a) studied the flocculation activity of γ-PGA and optimized the flocculation conditions. The results showed that the maximum flocculation activity of the γ-PGA (6.2×10^6 Da) produced by *Bacillus subtilis* R 23 fermentation was up to 34.7 L /OD with the optimal flocculation pH and coagulant ion Ca^{2+}.

γ-PGA has the excellent chelating ability of metal ions (Ni^{2+}, Cr^{3+}, Cu^{2+}, and Pb^{2+}) due to a good deal of carboxylic anion groups, which can be used for the removal and recovery of heavy metals. Hajdu et al. (2012) combined γ-PGA-Ni complexed nanoparticles with membrane separation technology to remove toxic lead ions from an aqueous solution. More than 99.8% of lead ions in water can be removed by the convenient low-pressure ultrafiltration technology, making the permeate meet the drinking water standard recommended by WHO. Chang et al. (2013) developed γ-PGA-coated superparamagnetic iron oxide NPs (γ-PGA/Fe_3O_4 NPs) by coprecipitation method for the removal of heavy metal ions. The results showed that γ-PGA/Fe_3O_4 NPs had a higher specific surface area and removal efficiency for all metal ions, as compared to Fe_3O_4 NPS or γ-PGA. Specifically, γ-PGA/Fe_3O_4 NPs showed a better removal activity at the solution pH higher than 6.0, removing more than 99% of Cr^{3+}, Cu^{2+}, Pb^{2+}, and 77% of Ni^{2+}. Sakamoto and Kawase (2016) used both water-insoluble γ-PGA and water-soluble sodium salt form of γ-PGA (γ-PGANa) as a low-cost, safe, and environmentally friendly biological adsorbent to recover cesium from radioactive wastewater. The adsorption principle of γ-PGA and γ-PGANa on cesium in radioactive wastewater is due to the electrostatic interaction between carboxylic acid ions ($-COO^-$) and cesium cation (Cs^+) via chemical reaction. The maximum adsorption capacities at equilibrium of γ-PGA and γ-PGANa for Cs were 345 mg-Cs(g-adsorbent)$^{-1}$ at pH 6.0 and 290 mg-Cs(g-adsorbent)$^{-1}$ at pH 9.0, respectively.

6 Conclusion and Future Outlook

As a biopolymer, γ-PGA has attracted worldwide attention. At present, remarkable progress has been made in screening and modifying production strains, optimizing fermentation conditions, and purifying γ-PGA. In particular, the industrial applications of γ-PGA have been widely recognized, which has resulted in the in-depth research of its metabolic pathways and essential enzymes.

The biological macromolecular polymer structure regulation and assembling mechanism is a significant problem, even though many studies have explored the key enzymes for γ-PGA synthesis. Understanding the molecular regulatory mechanisms of γ-PGA biosynthesis and control of stereoisomers would undoubtedly prove

valuable. However, due to the lack of relevant crystal structure data, the mechanism of γ-PGA polymerization and degradation has not been fully elucidated. As different γ-PGA synthases have differences in stability, activity, and molecular weight of products, future research should analyze its protein structure and regulation mechanism. Such work would result in feasible solutions for the production of γ-PGA of variable structures.

With the increasing trend in using biomass as a carbon source for fermentation processes, much research of γ-PGA bioproduction has aimed at improving the cost-effectiveness and the efficiency of recovery. At present, most of the reported production strains are glutamic acid-dependent, which require large amounts of glutamic acid in the medium. Therefore, some glutamic acid-independent strains, reducing the cost of fermentation by saving the cost of glutamate, have become the new research hotspot for the biosynthesis of γ-PGA. However, their low productivity remains the main bottleneck in industrial application. Further efforts are needed to carry out the production of γ-PGA with the use of cost-effective methods like harnessing the potential of agricultural waste in the production along with the construction of efficient microbial strains which can reduce the cost of fermentative production. Genetic manipulation could also be exploited to develop novel γ-PGA-producing strains, such as thermo- and salt-tolerant bacterial extremophiles. Finally, the establishment of γ-PGA separation processes for different downstream applications could be decisive in improving the cost-effectiveness of production and will help expand the application field of γ-PGA. Therefore, the systematic approach combining synthetic biology, metabolic engineering, and fermentation regulation technology may significantly improve the biosynthetic efficiency and product diversity of γ-PGA, which lay a foundation for further expanding its downstream application.

References

Ashiuchi M, Kamei T, Baek DH et al (2001a) Isolation of *Bacillus subtilis* (chungkookjang), a poly-γ-glutamate producer with high genetic competence. Appl Micobiol Biotechnol 57:764–769

Ashiuchi M, Nawa C, Kamei T et al (2001b) Physiological and biochemical characteristics of poly-γ-glutamate synthetase complex of *Bacillus subtilis*. Eur J Biochem 268:5321–5328

Ashiuchi M, Misono H (2002a) Biochemistry and molecular genetics of poly-γ-glutamate synthesis. Appl Microbiol Biotechnol 59(1):9–14

Ashiuchi M, Misono H (2002b) Poly-γ-glutamic acid. In: Fahnestock SR, Steinbüchel A (eds) Biopolymers. Wiley, Weinheim

Ashiuchi M, Shimanouchi K, Horiuchi T et al (2006) Genetically engineered poly-γ-glutamate producer from *Bacillus subtilis* ISW1214. Biosci Biotechnol Biochem 70:1794–1797

Ashiuchi M, Soda K, Misono H (1999a) A poly-γ-glutamate synthetic system of *Bacillus subtilis* IFO 3336: gene cloning and biochemical analysis of poly-γ-glutamate produced by *Escherichia coli* clone cells. Biochem Biophys Res Commun 263:6–12

Ashiuchi M, Soda K, Misono H (1999b) Characterization of *yrpC* gene product of *Bacillus subtilis* IFO 3336 as glutamate racemase isozyme. Biosci Biotechnol Biochem 63:792–798

Ashiuchi M, Tani K, Soda K et al (1998) Properties of glutamate racemase from *Bacillus subtilis* IFO 3336 producing poly-γ-glutamate. J Biochem 123:1156–1163

Ashiuchi M, Yamashiro D. 2009. Moonlighting function of the *pgsE* gene product, a novel member of the membranous enzyme complex responsible for the synthesis of D-glutamate-containing poly-γ-glutamate. The first international conference of D-amino acid research. Awaji

Bajaj I, Singhal R (2011) Poly (glutamic acid)-an emerging biopolymer of commercial interest. Bioresour Technol 102(10):5551–5561

Bajaj IB, Singhal RS (2009a) Flocculation properties of poly(γ-glutamic acid) produced from *Bacillus subtilis* isolate. Food Bioprocess Technol 4(5):745–752

Bajaj IB, Singhal RS (2009b) Enhanced production of poly(γ-glutamic acid) from *Bacillus licheniformis* NCIM 2324 by using metabolic precursors. Appl Biochem Biotechnol 159(1): 133–141

Balikó G, Venetianer P (1993) An *Escherichia coli* gene in search of a function: phenotypic effects of the gene recently identified as *murI*. J Bacteriol 175:6571–6577

Belitsky BR, Sonenshein AL (1998) Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. J Bacteriol 180:6298–6305

Belitsky BR, Wray L Jr, Fisher SH et al (2000) Role of *tnrA* in nitrogen source-dependent repression of *Bacillus subtilis* glutamate synthase gene expression. J Bacteriol 182:5939–5947

Bhat AR, Irorere VU, Bartlett T et al (2013) *Bacillus subtilis* natto: a non-toxic source of poly-γ-glutamic acid that could be used as a cryoprotectant for probiotic bacteria. AMB Expr 3(1):1–9

Bhatt R, Vries P, Klein J, et al. 2002. Polyglutamic acid-camptothecin conjugates and methods of preparation. WO2001070275A3

Birrer GA, Cromwick AM, Gross RA (1994) Gamma-poly(glutamic acid) formation by *Bacillus licheniformis* 9945a: physiological and biochemical studies. Int J Biol Macromol 16:265–275

Bovarnick M (1942) The formation of extracellular D(−)-glutamic acid polypeptide by *Bacillus subtillis*. J Biol Chem 145:415–424

Cai D, Chen Y, He P et al (2018) Enhanced production of poly-γ-glutamic acid by improving ATP supply in metabolically engineered *Bacillus licheniformis*. Biotechnol Bioeng 115(10): 2541–2553

Cai D, He P, Lu X et al (2017) A novel approach to improve poly-γ-glutamic acid production by NADPH regeneration in *Bacillus licheniformis* WX-02. Sci Rep 7:43404

Candela T, Fouet A (2005) Bacillus anthracis CapD, belonging to the γ-glutamyltranspeptidase family, is required for the covalent anchoring of capsule to peptidoglycan. Mol Microbiol 57(3): 717–726

Candela T, Fouet A (2006) Poly-gamma-glutamate in bacteria. Mol Microbiol 60:1091–1098

Candela T, Moya M, Haustant M et al (2009) *Fusobacterium nucleatum*, the first gram-negative bacterium demonstrated to produce polyglutamate. Can J Microbiol 55:627–632

Cao M, Geng W, Liu L et al (2011) Glutamic acid independent production of poly-γ-glutamic acid by *bacillus amyloliquefaciens* LL3 and cloning of *pgsBCA* genes. Bioresour Technol 102(5): 4251–4257

Cao M, Geng W, Zhang W et al (2013) Engineering of recombinant *Escherichia coli* cells co-expressing poly-γ-glutamic acid (γ-PGA) synthetase and glutamate racemase for differential yielding of γ-PGA. Microb Biotechnol 6(6):675–684

Cao M, Song C, Jin Y et al (2010) Synthesis of poly (γ-glutamic acid) and heterologous expression of *pgsBCA* genes. J Mol Catal B Enzym 67(1–2):111–116

Chang J, Zhong Z, Xu H et al (2013) Fabrication of poly(γ-glutamic acid)-coated Fe_3O_4 magnetic nanoparticles and their application in heavy metal removal. Chin J Chem Eng 21(11): 1244–1250

Chen W, Wang R, Xu T et al (2017) A mussel-inspired poly (γ-glutamic acid) tissue adhesive with high wet strength for wound closure. J Mater Chem B 5(28):5668–5678

Cheng C, Asada Y, Aida T (1989) Production of γ-polyglutamic acid by *Bacillus subtilis* A35 under denitrifying conditions. Agric Biol Chem 53:2369–2375

Deng S, Bai R, Hu X et al (2003) Characteristics of a bioflocculant produced by *bacillus mucilaginosus* and its use in starch wastewater treatment. Appl Microbiol Biotechnol 60(5): 588–593

Do JH, Chang HN, Lee SY (2001) Efficient recovery of poly (glutamic acid) from highly viscous culture broth. Biotechnol Bioeng 76(3):219–223

Feng J, Gu Y, Quan Y et al (2015) Improved poly-γ-glutamic acid production in *bacillus amyloliquefaciens* by modular pathway engineering. Metab Eng 32:106–115

Feng J, Gu Y, Quan Y et al (2017) Construction of energy-conserving sucrose utilization pathways for improving poly-γ-glutamic acid production in *bacillus amyloliquefaciens*. Microb Cell Factories 16:98

Feng X, Tang B, Jiang Y et al (2016) Efficient production of poly-γ-glutamic acid from cane molasses by *Bacillus subtilis* NX-2 immobilized on chemically modified sugarcane bagasse. J Chem Technol Biotechnol 91(7):2085–2093

Fujita KI, Tomiyama T, Inoi T et al (2021) Effect of pgsE expression on the molecular weight of poly(γ-glutamic acid) in fermentative production. Polym J 53:409–414

Gao W, Liu F, Zhang W et al (2016) Mutations in genes encoding antibiotic substances increase the synthesis of poly-γ-glutamic acid in *bacillus amyloliquefaciens* LL3. MicrobiologyOpen 6(1): e00398

Goto A, Kunioka M (1992) Biosynthesis and hydrolysis of poly(γ-glutamic acid) from *Bacillus subtilis* IFO3335. Biosci Biotechnol Biochem 56:1031–1035

Guo Z, Yang N, Zhu C et al (2017) Exogenously applied poly-γ-glutamic acid alleviates salt stress in wheat seedlings by modulating ion balance and the antioxidant system. Environ Sci Pollut Res 24:6592–6598

Hajdu I, Bodnár M, Csikós Z et al (2012) Combined nano-membrane technology for removal of lead ions. J Membr Sci 409-410:44–53

Halmschlag B, Steurer X, Putri SP et al (2019) Tailor-made poly-γ-glutamic acid production. Metab Eng 55:239–248

Hara T (2000) Desert greening: greening by utilization of microbial macromolecules. Kobunshi 49: 367–370

Hezayen FF, Rehm BHA, Tindall BJ et al (2001) Transfer of *Natrialba asiatica* B1T to *Natrialba taiwanensis sp.* nov. and description of *Natrialba aegyptiaca sp.* nov., a novel extremely halophilic, aerobic, non-pigmented member of the archaea from Egypt that produces extracellular poly(glutamic acid). Int J Syst Evol Microbiol 51:1133–1142

Hsieh CY, Tsai SP, Wang DM et al (2005) Preparation of gamma-PGA/chitosan composite tissue engineering matrices. Biomaterials 26(28):5617–5623

Ito Y, Tanaka T, Ohmachi T (1996) Glutamic acid independent production of poly(γ-glutamic acid) by *Bacillus subtilis* TAM-4. Biosci Biotechnol Biochem 60:1239–1242

Ivanovics G, Bruckner V (1937) The chemical nature of the immuno-specific capsule substance of anthrax *bacillus*. Naturwissenschaften 25:250

Jia WS, Zhi YY, Zhong Y et al (2019) Effect of γ-glutamic acid on the antifreeze protection of frozen dough and noodles. Food Ind Sci Tech 7

Jiang H, Shang L, Yoon SH et al (2006) Optimal production of poly-γ-glutamic acid by metabolically engineered *Escherichia coli*. Biotechnol Lett 28:1241–1246

Jiang Y, Tang B, Xu Z et al (2016) Improvement of poly-gamma-glutamic acid biosynthesis in a moving bed biofilm reactor by *Bacillus subtilis* NX-2. Bioresour Technol 218:360–366

Ju WT, Song YS, Jung WJ, Park RD (2014) Enhanced production of poly-gamma-glutamic acid by a newly-isolated *Bacillus subtilis*. Biotechnol Lett 36(11):2319–2324

Jung DY, Jung S, Yun JS et al (2005) Influences of cultural medium component on the production of poly (γ-glutamic acid) by *bacillus* sp. RKY3. Biotechnol Bioproc Eng 10(4):289

Junnan F, Juan L, Lishan XU et al (2018) Recent advances in poly-γ-glutamic acid production by microbial fermentation. Chin J Appl Environ Biol 24(5):1041–1049

Katsuragi Y, Sugiura Y, Ono S, et al. 1998. Bitterness-relieving agent. WO2000021390A1

Kimura K, Itoh Y (2003) Characterization of poly-γ-glutamate hydrolase encoded by a bacteriophage genome: possible role in phage infection of *Bacillus subtilis* encapsulated with poly-γ-glutamate. Appl Environ Microbiol 69(5):2491–2497

Kimura K, Tran LSP, Do TH et al (2009) Expression of the pgsB encoding the poly-gamma-DL-glutamate synthetase of *Bacillus subtilis* (natto). Biosci Biotechnol Biochem 73:1149–1155

Kimura K, Tran LSP, Uchida I et al (2004a) Characterization of *Bacillus subtilis* γ-glutamyltransferase and its involvement in the degradation of capsule poly-γ-glutamate. Microbiology 150(12):4115–4123

Kimura K, Tran LSP, Itoh Y (2004b) Roles and regulation of the glutamate racemase isogenes, *racE* and *yrpC*, in *Bacillus subtilis*. Microbiology 150:2911–2920

King EC, Blacker AJ, Bugg TD (2001) Enzymatic breakdown of poly-γ-D-glutamic acid in *bacillus licheniformis*: identification of a polyglutamyl γ-hydrolase enzyme. Biomacromolecules 1(1): 75–83

Kinnersley A M, Koskan L P, Meah A R Y, et al. 1994. Composition and method for enhanced fertilizer uptake by plants. CA2148798A1

Ko HY, Gross AR (1998) Effects of glucose and glycerol on γ-poly(glutamic acid) formation by *Bacillus licheniformis* ATCC 9945A. Biotechnol Bioeng 57:430–437

Kuboat H, Matsunobu T, Uotani K (1993) Production of poly(γ-glutamic acid) by *Bacillus subtilis* F-2-01. Biosci Biotech Biochem 57:1212–1213

Kubota H, Matsunobu T, Uotani K et al (1993) Production of poly(γ-glutamic acid) by *Bacillus subtilis* F-2-01. Biosci Biotech Biochem 57:1212–1213

Kunioka K (1997) Biosynthesis and chemical reactions of poly(amino acid)s from microorganisms. Appl Microbiol Biotechnol 47:469–475

Lei P, Pang X, Feng X et al (2017) The microbe-secreted isopeptide poly-γ-glutamic acid induces stress tolerance in Brassica napus L. seedlings by activating crosstalk between H_2O_2 and Ca^{2+}. Sci Rep 7:41618

Lei P, Xu Z, Ding Y et al (2015) Effect of poly (γ-glutamic acid) on the physiological responses and calcium signaling of rape seedlings (*Brassica napus* L.) under cold stress. J Agric Food Chem 63:10399–10406

Lei P, Xu Z, Liang J et al (2016) Poly (γ-glutamic acid) enhanced tolerance to salt stress by promoting proline accumulation in Brassica napus L. Plant Growth Regul 78:233–241

Li C, Yu DF, Newman RA et al (1998) Complete regression of well-established tumors using a novel water-soluble poly(L-glutamic acid)-paclitaxel conjugate. Cancer Res 58(11):2404–2409

Lim SM, Kim J, Shim JY et al (2012) Effect of poly-γ-glutamic acids (PGA) on oil uptake and sensory quality in doughnuts. Food Sci Biotechnol 21(1):247–252

Liu F, Tian W, Li L et al (2013) Optimization of solid-state fermentation conditions for antagonistic *Bacillus subtilis* SQR9 producing bio-organic fertilizer. Chin J Appl Environ Biol 19(1):90–95

Liu J, Ma X, Wang Y et al (2011) Depressed biofilm production in *bacillus amyloliquefaciens* C06 causes γ-polyglutamic acid (γ-PGA) overproduction. Curr Microbiol 62:235–241

Liu T, Nobeshima H, Ojima Y, Azuma M (2018) A new method to purify poly-γ-glutamic acid using gemini quaternary ammonium salts and characterization of its ionic complex. J Chem Eng Jpn 51(5):431–437

Liu X, Liu F, Liu SY et al (2015) Moist efficacy and safety evaluation of polyglutamate. Chem Ind Daily Use 45(05):275–278

Manocha B, Margaritis A (2010) A novel method for the selective recovery and purification of γ-polyglutamic acid from *Bacillus licheniformis* fermentation broth. Biotechnol Prog 26(3): 734–742

Min L, Jin Z, Caldovic L et al (2009) Mechanism of allosteric inhibition of N-acetyl-L-glutamate synthase by L-arginine. J Biol Chem 284:4873–4880

Mitsui N, Murasawa H, Sekiguchi J (2011) Disruption of the cell wall lytic enzyme CwlO affects the amount and molecular size of poly-γ-glutamic acid produced by *Bacillus subtilis* (natto). J Gen Appl Microbiol 57(1):35–43

Mitsuiki M, Mizuno A, Tanimoto H et al (1998) Relationship between the antifreeze activities and the chemical structures of oligo- and poly(glutamic acid). J Agric Food Chem 46(3):891–895

Muro T, Nagamori Y, Okada S et al (1990) Some properties and action of poly(glutamic acid) hydrolase II from *Micromonospora melanosporea* IFO 12515. J Agric Chem Soc Jpn 54(4): 1065–1067

Nordlund P, Eklund H (1995) Di-iron-carboxylate proteins. Curr Opin Struct Biol 5:758–766

Ogasawara Y, Shigematsu M, Sato S et al (2019) Involvement of peptide epimerization in poly-γ-glutamic acid biosynthesis. Org Lett 21(11):3972–3975

Ogawa Y, Yamaguchi F, Yuasa K et al (1997) Efficient production of γ-polyglutamic acid by *Bacillus subtilis* (natto) in jar fermenters. Biosci Biotech Biochem 61:1684–1687

Ogunleye A, Bhat A, Irorere VU et al (2015) Poly-γ-glutamic acid: production, properties and applications. Microbiology 161(1):1–17

Pang X, Lei P, Feng X et al (2018) Poly-γ-glutamic acid, a bio-chelator, alleviates the toxicity of cd and Pb in the soil and promotes the establishment of healthy *Cucumis sativus* L. seedling. Environ Sci Pollut Res 25:19975–19988

Park C, Choi JC, Choi YH et al (2005) Synthesis of super-high-molecular-weight poly-γ-glutamic acid by *Bacillus subtilis* subsp. chungkookjang. J Mol Catal B Enzym 35(4–6):128–133

Peng Y, Jiang B, Zhang T et al (2015) High-level production of poly(γ-glutamic acid) by a newly isolated glutamate-independent strain, *bacillus methylotrophicus*. Process Biochem 50(3): 329–335

Qiu Y, Sha Y, Zhang Y et al (2017) Development of Jerusalem artichoke resource for efficient one-step fermentation of poly-(γ-glutamic acid) using a novel strain *bacillus amyloliquefaciens* NX-2S. Bioresour Technol 239:197–203

Qiu Y, Zhang Y, Zhu Y, Sha Y, Xu Z, Feng X, Li S, Xu H (2019) Improving poly-(γ-glutamic acid) production from a glutamic acid-independent strain from inulin substrate by consolidated bioprocessing. Bioprocess Biosyst Eng 42(10):1711–1720

Qiu Y, Zhu Y, Sha Y et al (2020) Development of a robust *bacillus amyloliquefaciens* cell factory for efficient poly(γ-glutamic acid) production from Jerusalem artichoke. ACS Sustain Chem Eng 8(26):9763–9774

Rusnak F, Mertz P (2000) Calcineurin: form and function. Physiol Rev 80:1483–1521

Sakamoto S, Kawase Y (2016) Adsorption capacities of poly-gamma-glutamic acid and its sodium salt for cesium removal from radioactive wastewaters. J Environ Radioact 165:151–158

Salehizadeh H, Shojaosadati SA (2001) Extracellular biopolymeric flocculants: recent trends and biotechnological importance. Biotechnol Adv 19(5):371–385

Sawada K, Araki H, Takimura Y et al (2018) Poly-L-gamma-glutamic acid production by recombinant *Bacillus subtilis* without *pgsA* gene. AMB Express 8(1):110

Schneerson R, Kubler-Kielb J et al (2003) Poly(γ-D-glutamic acid) protein conjugates induce IgG antibodies in mice to the capsule of *bacillus anthracis*: a potential addition to the anthrax vaccine. Proc Natl Acad Sci U S A 100(15):8945–8950

Scoffone V, Dondi D, Biino G et al (2013) Knockout of *pgdS* and *ggt* genes improves γ-PGA yield in *B. subtilis*. Biotechnol Bioeng 110(7):2006–2012

Sha Y, Qiu Y, Zhu Y et al (2020a) CRISPRi-based dynamic regulation of hydrolase for synthesis of poly-γ-glutamic acid with variable molecular weights. ACS Synth Biol 9(9):2450–2459

Sha Y, Huang Y, Zhu Y et al (2020b) Efficient biosynthesis of low-molecular-weight poly-γ-glutamic acid based on stereochemistry regulation in *bacillus amyloliquefaciens*. ACS Synth Biol 9(6):1395–1405

Sha Y, Sun T, Qiu Y et al (2019a) Investigation of glutamate dependence mechanism for poly-γ-glutamic acid production in *Bacillus subtilis* based on transcriptome analysis. J Agric Food Chem 67(22):6263–6274

Sha Y, Zhang Y, Qiu Y et al (2019b) Efficient biosynthesis of low-molecular-weight poly-γ-glutamic acid by stable overexpression of *PgdS* hydrolase in *bacillus amyloliquefaciens* NB. J Agric Food Chem 67(1):282–290

Sha Y, Tao S, Qiu Y et al (2019c) Investigation of glutamate dependence mechanism for poly-γ-glutamic acid production in *Bacillus subtilis* on the basis of transcriptome analysis. J Agric Food Chem 67(5):6263–6274

Shi F, Xu Z, Cen P (2006) Optimization of γ-polyglutamic acid production by *Bacillus subtilis* ZJU-7 using a surface-response methodology. Biotechnol Bioprocess Eng 11(3):251–257

Shih I, Wu P, Shieh C (2005) Microbial production of a poly(γ-glutamic acid) derivative by *Bacillus subtilis*. Process Biochem 40(8):2827–2832

Shih L, Van YT (2001) The production of poly-(γ-glutamic acid) from microorganisms and its various applications. Bioresour Technol 79(3):207–225

Shyu YS, Hwang JY, Hsu CK (2008) Improving the rheological and thermal properties of wheat dough by the addition of γ-polyglutamic acid. LWT - Food Sci Technol 41(6):982–987

Shyu YS, Sung WC (2010) Improving the emulsion stability of sponge cake by the addition of γ-polyglutamic acid. J Mar Sci Technol 18(6):895–900

Su Y, Li X, Liu Q et al (2010) Improved poly-γ-glutamic acid production by chromosomal integration of the *Vitreoscilla* hemoglobin gene (*vgb*) in *Bacillus subtilis*. Bioresour Technol 101(12):4733–4736

Sun A, He X, Li L et al (2020) An injectable photopolymerized hydrogel with antimicrobial and biocompatible properties for infected skin regeneration. NPG Asia Mater 12(1):1–11

Sung MH, Park C, Choi JC, et al. 2014. Hyaluronidase inhibitor containing poly-gamma-glutamic acid as an effective component. U8916141B2

Suzuki T, Tahara Y (2003) Characterization of the *Bacillus subtilis ywtD* gene, whose product is involved in γ-polyglutamic acid degradation. J Bacteriol 185(7):2379–2382

Tanaka T, Hiruta O, Futamura T et al (1993) Purification and characterization of poly (γ-glutamic acid) hydrolase from a filamentous fungus, *Myrothecium* sp. TM-4222. Biosci Biotechnol Biochem 57(12):2148–2153

Tanimoto H, Fox T, Eagles J et al (2007) Acute effect of poly-gamma-glutamic acid on calcium absorption in post-menopausal women. J Am Coll Nutr 26(6):645–649

Tanimoto H, Mori M, Motoki M et al (2001) Natto mucilage containing poly-gamma-glutamic acid increases soluble calcium in the rat small intestine. Biosci Biotechnol Biochem 65(3):516–521

Tomosho JW, Moran RG, Coward JK (2008) Concentration-dependent processivity of multiple glutamate ligations catalyzed by folylpoly-γ-glutamate synthetase. Biochemistry 47:9040–9050

Tsao CT, Chang CH, Lin YY et al (2011) Evaluation of chitosan/γ-poly(glutamic acid) polyelectrolyte complex for wound dressing materials. Carbohydr Polym 84(2):812–819

Vetting MW, De Carvalho LPS, Yu M et al (2005) Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys 433:212–226

Volcani BE, Margalith P (1957) A new species (*Flavobacterium polyglutamicum*) which hydrolyzes the gamma-L-glutamyl bond in polypeptides. J Bacteriol 74(5):646–655

Wang Q, Chen S, Zhang J et al (2008) Co-producing lipopeptides and poly-γ-glutamic acid by solid-state fermentation of *Bacillus subtilis* using soybean and sweet potato residues and its biocontrol and fertilizer synergistic effects. Bioresour Technol 99:3318–3323

Wang R, Wang X, Zhan Y et al (2019) A dual network hydrogel sunscreen based on poly-γ-glutamic acid/tannic acid demonstrates excellent anti-UV, self-recovery, and skin-integration capacities. ACS Appl Mater Interf 11(41):37502–37512

Wang WG, Wang W, Zhao YL et al (2016) Research and application progress of γ-polyglutamic acid. J Henan Univ Technol 2

Wang XX, Li S, Ren ZK et al (2020) Antioxidant effect of γ-poly (glutamate) on skin photoaging induced by UV radiation. Bioprocessing 18(05):561–566

Weber J (1989) Poly(γ-glutamic acid)s are the major constituent of nematocysts in hydra (hydrozoa, Cnidaria). J Biol Chem 265(17):9664–9669

Wu Q, Xu H, Shi N et al (2008) Improvement of poly (γ-glutamic acid) biosynthesis and redistribution of metabolic flux with the presence of different additives in *Bacillus subtilis* CGMCC 0833. Appl Microbiol Biotechnol 79(4):527

Wu Q, Xu H, Zhang L et al (2006a) Production, purifcation and properties of γ-glutamyltranspeptidase from a newly isolated *Bacillus subtilis* NX-2. J Mol Cat B Enzym 43(1–4):113–117

Wu Q, Xu H, Xu L, Ouyang P (2006b) Biosynthesis of poly(γ-glutamic acid) in *Bacillus subtilis* NX-2: regulation of stereochemical composition of poly(γ-glutamic acid). Process Biochem 41: 1650–1655

Xu G, Zha J, Cheng H et al (2019a) Engineering *Corynebacterium glutamicum* for the de novo biosynthesis of tailored poly-γ-glutamic acid. Metab Eng 56:39–49

Xu T, Yang R, Ma X et al (2019b) Bionic poly(γ-glutamic acid) electrospun fibrous scaffolds for preventing hypertrophic scars. Adv Healthc Mater 8(13):1900123

Xu Z, Lei P, Feng X et al (2014) Calcium involved in the poly (γ-glutamic acid)-mediated promotion of Chinese cabbage nitrogen metabolism. Plant Physiol Biochem 80:144–152

Xu Z, Lei P, Feng X et al (2016) Analysis of the metabolic pathways affected by poly(γ-glutamic acid) in *Arabidopsis thaliana* based on genechip microarray. J Agric Food Chem 64:6257–6266

Xu Z, Lei P, Pang X et al (2017a) Exogenous application of poly-γ-glutamic acid enhances stress defense in *Brassica napus* L. seedlings by inducing cross-talks between Ca^{2+}, H_2O_2, brassinolide, and jasmonic acid in leaves. Plant Physiol Biochem 118:460–470

Xu H, Feng XH, Xu DL et al (2017b) Development and application prospect of polyamino acid functional polymers. Biotechnology 6

Yamanaka K, Maruyama C, Takagi H et al (2008) ε-Poly-L-lysine dispersity is controlled by a highly unusual nonribosomal peptide synthetase. Nat Chem Biol 4:766–772

Yao J, Jing J, Xu H et al (2009) Investigation on enzymatic degradation of γ-polyglutamic acid from *Bacillus subtilis* NX-2. J Mol Catal B Enzym 56(2–3):158–164

Yeh CM, Wang JP, Lo SC et al (2010) Chromosomal integration of a synthetic expression control sequence achieves poly-γ-glutamate production in a *Bacillus subtilis* strain. Biotechnol Prog 26(4):1001–1007

Zeng W, Chen G, Guo Y et al (2017) Production of poly-γ-glutamic acid by a thermotolerant glutamate-independent strain and comparative analysis of the glutamate dependent difference. AMB Expr 7(1):213

Zeng W, Lin Y, Qi Z et al (2013) An integrated high-throughput strategy for rapid screening of poly (γ-glutamic acid)-producing bacteria. Appl Microbiol Biotechnol 97(5):2163–2172

Zhan Y, Zhu C et al (2017) Improvement of glycerol catabolism in *Bacillus licheniformis* for production of poly-γ-glutamic acid. Appl Microbiol Biotechnol 101(19):7155–7164

Zhang C, Wu DJ, Zheng XJ (2017) Generation of a high γ-polyglutamic acid-producing, high glucose-tolerant *Bacillus subtilis* strain via genome shuffling. J Biobaased Mater Bioenergy 11(1):73–77

Zhang H, Zhu J, Zhu X et al (2012a) High-level exogenous glutamic acid-independent production of poly-(γ-glutamic acid) with organic acid addition in a new isolated *Bacillus subtilis* C10. Bioresour Technol 116:241–246

Zhang D, Feng X, Li S et al (2012b) Effects of oxygen vectors on the synthesis and molecular weight of poly (γ-glutamic acid) and the metabolic characterization of *Bacillus subtilis* NX-2. Process Biochem 47(12):2103–2109

Zhang D, Feng X, Zhou Z et al (2012c) Economical production of poly(γ-glutamic acid) using untreated cane molasses and monosodium glutamate waste liquor by *Bacillus subtilis* NX-2. Bioresour Technol 114:583–588

Zhang JX (2020) Study on synthesis and application of γ-polyglutamic acid. Green Sci Tech 24: 212–215

Zhang L, Zhu X, Wu S et al (2018) Fabrication and evaluation of a γ-PGA-based self-assembly transferrin receptor-targeting anticancer drug carrier. Int J Nanomedicine 13:7873

Zhang W, He Y, Gao W et al (2015) Deletion of genes involved in glutamate metabolism to improve poly-γ-glutamic acid production in *B. amyloliquefaciens* LL3. J Ind Microbiol Biotechnol 42:297–305

Bioengineering and Bioprocessing of Virus-Like Particle Vaccines in *Escherichia coli*

Rufika S. Abidin and Frank Sainsbury

Contents

Abstract Virus-like particles (VLPs) are macromolecular assemblies of recombinant viral structural proteins. Self-assembling with high precision, the repetitive nature of VLPs, their size, and particulate form result in the capacity for effective stimulation of humoral and cellular immune responses. VLP vaccines have achieved commercial success for protection against cognate viruses. However, there is a rich field of research dedicated to harnessing the immune-stimulatory properties of VLPs for the presentation of heterologous antigens for protection and treatment of chronic and infectious diseases. Peptides, proteins, carbohydrates, and small molecules have been converted into effective immunogens via presentation on various VLP platforms. In each case, unique bioengineering challenges must be overcome to allow

R. S. Abidin
Hasanuddin University Medical Research Center, Hasanuddin University, Makassar, Indonesia

F. Sainsbury (✉)
Centre for Cell Factories and Biopolymers, Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD, Australia
e-mail: f.sainsbury@griffith.edu.au

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_10

the recovery, assembly, and structural fidelity of the VLP platform, as well as the immunogenicity of the antigen. This review highlights some of the biomolecular engineering approaches that are being employed to effectively present diverse biological molecules and chemical moieties on VLP platforms and bioprocessing strategies for their efficient recovery from the prokaryotic expression host *Escherichia coli*.

1 Introduction

Virus-like particles are generally derived from the structural proteins of viruses, which self-assemble into spherical or rod-shaped structures, mimicking the original virus. A VLP does not contain other viral components such as pathogenic proteins or genetic material and is thus nonpathogenic and noninfectious (Teunissen et al. 2013; Lua et al. 2014; Rodriguez-Limas et al. 2013). VLPs can be constructed from the capsid proteins of enveloped or non-enveloped viruses. In the case of enveloped viruses, a lipid bilayer coats the capsid supporting the presentation of glycoproteins embedded in the membrane (Kushnir et al. 2012). The existence of naturally occurring VLPs was first reported in 1965 (Blumberg et al. 1965), and their ability to induce cellular and humoral immunity was subsequently identified in 1968 (Bayer et al. 1968). The development of genetic engineering further allowed the expression and purification of synthetic VLPs (Hagensee et al. 1993; Kirnbauer et al. 1992; Li et al. 1997), which enabled characterization and understanding of their assembly (Li et al. 1997; Salunke et al. 1986). Over the course of more than 4 decades, at least 110 VLPs from 75 different viral families have been constructed and evaluated (Zeltins 2013; Yan et al. 2015). VLPs are deemed an ideal biological vehicle due to their biocompatibility, solubility, uptake efficiency, and their capability for targeted delivery and drug loading (Yan et al. 2015). Amenable to both genetic engineering and chemical modification, recombinant VLPs have been developed for various uses such as antigen carriers (Babin et al. 2013; Kawano et al. 2014; Wibowo et al. 2012) or as carriers of other cargos such as drugs (Glasgow and Tullman-Ercek 2014), proteins (Abbing et al. 2004), and both oligonucleotides and plasmid DNA (Braun et al. 1999).

To date, there are three VLP-based vaccines that have been officially approved for clinical use (Kushnir et al. 2012). The first to be approved was the recombinant hepatitis B (HB) VLP expressed in yeast (Recombivax HB, Merck), approved by the Food and Drug Administration (FDA) for use as a vaccine in 1986. The second was recombinant HPV VLP vaccine produced either in yeast (Gardasil, Merck), approved by the FDA in 2006, or in insect cells (Cervarix, GlaxoSmithKline), approved by the FDA in 2009. The third was a recombinant hepatitis E (HE) VLP vaccine produced in *Escherichia coli* (Hecolin, Xiamen Innovax Biotech), approved by Chinese FDA in 2011 (Rodriguez-Limas et al. 2013). Of note, is the fact that these VLP vaccines are all active against the cognate virus. The inherent immunogenicity of VLPs, related to their size and structure, suggests that they also make

promising platforms for the presentation of heterologous antigens either by genetic fusion, chemical conjugation, or encapsidation. Such bioengineering efforts to expand the use of VLP platforms produced in *E. coli* as vaccines for both infectious and chronic diseases are the focus of this review.

2 VLPs and Immunogenicity

VLP-based vaccines are engineered to carry heterologous antigens with the goal of stimulating one or both branches of adaptive immunity: humoral or cellular immunity. VLP vaccine candidates carrying non-proteinaceous antigens primarily aim to stimulate an antibody response. Humoral immunity relies on antibody production, which is initiated by B-cell receptor (BCR) recognition of foreign antigen (Murphy et al. 2008). BCR binds specifically and sterically to a surface-exposed region of the antigen known as the antigenic determinant or B-cell epitope. In proteins, the antigenic determinant may be a conformational epitope, in that it relies on the tertiary or quaternary structure of the antigen. Or, an epitope could simply be a string of amino acids, which is thus known as a continuous or linear epitope. Heterologous peptide antigens displayed on a VLP carrier usually come in the form of identified conformational or linear epitopes fused or conjugated to the VLP (Lee et al. 2016). The display of conformational epitopes poses a significant protein engineering challenge (Lua et al. 2014) and will be discussed in later sections.

Cellular immunity, on the other hand, relies on the activation of T cells by T-cell receptor (TCR) recognition of the antigenic determinant displayed on the surface of antigen-presenting cells (APCs). In this case, the antigenic determinant is a combination of a major histocompatibility complex (MHC) molecule and a string of amino acids derived from the antigen, the T-cell epitope (Murphy et al. 2008). Cytotoxic T-cell (Tc) epitopes are processed via cytosolic proteolysis as a part of the MHC Class I pathway, while helper T-cell (Th) epitopes are processed via lysosomal proteolysis as a part of the MHC Class II pathway. Tc epitopes are usually 8–10 amino acids long, while Th epitopes are usually longer (10–15 amino acids), and both are often hydrophobic (Mitic et al. 2014). For the delivery of heterologous Tc or Th epitopes by VLP vaccines, the peptides can either be displayed on the VLP surface or encapsidated within the interior of a VLP. However, the hydrophobicity of T-cell epitopes imposes stringent constraints on the location and context of epitope insertion to permit efficient expression and recovery of recombinant subunits and on the assembly and stability of VLPs.

The versatility of VLPs as heterologous antigen vaccine carriers is enhanced by the inherent properties of the particles that determine their immunostimulatory capacity. Principally, the repetitive nature of VLP structures activates innate as well as adaptive immunity by recognition of repetitive patterns as a characteristic of foreign entities (Bachmann and Jennings 2010). The repetitive structure of VLPs is due to the assembly of identical VLP subunits composed of one or more proteins. Therefore, VLP subunits can be manipulated to display foreign antigens in a

repetitive array that mirrors the structure on which it is presented. The multivalent display of heterologous antigens on the surface of a VLP augments humoral immunity by BCR cross-linking, which has been shown to strengthen B-cell activation and, in the case of self-antigen presentation, break B-cell tolerance (Bachmann and Jennings 2010; Schiller and Chackerian 2014). Furthermore, optimal BCR-mediated B-cell activation depends on antigen spacing. A distance of 5–10 nm between antigens is optimal for B-cell activation (Jegerlehner et al. 2002; Bachmann and Zinkernagel 1997), and, while this is not achievable with non-repetitive soluble antigens, the geometry of VLPs generally directs such spacing. In addition, there is some evidence to say that the repetitive pattern also permits complement activation, which can also enhance B-cell responses (Barrington et al. 2001).

The size of VLPs (20–200 nm) also contributes to their inherent immunogenicity. On one hand, their size allows them to migrate directly to secondary lymphoid organs where they may present intact antigens with a native configuration to B cells in the germinal center, which is a prerequisite for the stimulation of antibody production (Swartz 2001; Reddy et al. 2006; Oussoren et al. 1997). On the other hand, the size of VLPs also enables efficient uptake by professional antigen-presenting cells (APCs), especially dendritic cells. APC stimulation activates the MHC Class II pathway to further support the humoral response and antibody production by B cells. However, APC stimulation also activates the MHC Class I pathway, either directly or via cross-presentation, to support the cytotoxic activity of T cells (Bachmann and Jennings 2010). Recombinant icosahedral capsids may also have an empty cavity, which opens the possibility of loading immunogenic cargos such as polypeptides and nucleic acids. Uptake of VLPs can result in the intracellular release of their cargo, which is essential for cross-presentation and the stimulation of cellular immunity.

3 Bioengineering Strategies for Surface Presentation

Bioengineering approaches to presenting pathogenic antigens on heterologous VLP platforms must consider several factors such as immunogenicity, structural preservation, upstream and downstream bioprocess development, scalability, cost of production, and stability/shelf life. To improve humoral immunogenicity, antigens should be presented on the surface of the VLPs and their conformation preserved in the case of discontinuous epitopes. To improve cellular immunogenicity, vaccine design and VLP selection must ensure that the VLPs are taken up by APCs and are correctly processed by the MHC Class I or Class II pathways. Furthermore, VLP-based vaccine development must ensure that foreign antigen coupling does not interfere with assembly competency of the VLP and does not cause aggregation. In this review, we will focus on bioengineering design to improve immunogenicity and structural preservation while touching on the other factors briefly.

3.1 Genetic Fusion

VLPs subjected to bioengineering via genetic means generally consist of capsid protein monomers or assembly subunits, called capsomeres, composed of homo-multimers. For stimulation of antibody production, antigenic sequences are fused to surface-exposed regions of the capsid monomer. To achieve this, the fusion of peptide or subunit antigens can be located at surface-exposed loops, the N-terminal domain, or the C-terminal domain of the capsid monomer, depending on the structure of the VLP platform used.

3.1.1 Genetic Fusion to Surface-Exposed Loops

Antigen insertion on heterologous VLPs is frequently done by replacing an immunodominant region of a surface-exposed loop with the antigenic sequence (Chackerian et al. 1999; Cheong et al. 2009), or simply by inserting the antigen without native amino acid removal (Kawano et al. 2014; Tumban et al. 2011; Zamora et al. 2006; Anggraeni et al. 2013; Rivera-Hernandez et al. 2013; Abidin et al. 2015; Tekewe et al. 2015; Pattinson et al. 2019). This approach often poses a bioengineering challenge as the removal of native regions of the capsid monomer or the addition of foreign antigens has been shown to interfere with the assembly VLPs. For example, while the removal of different immunodominant regions in the bovine papilloma virus (BPV) L1 protein has been reported not to interfere with assembly (Chackerian et al. 1999), replacing these regions with the heterologous CCR5 extracellular loop resulted in the loss of assembly competency in three out of four constructs depending on which native immunodominant region was replaced. Conversely, removal of hepatitis B virus surface antigen (Hbs-Ag) Tc epitopes resulted in a loss of assembly in three out of four constructs, while replacement with an influenza M1 Tc epitope could restore assembly for one of these constructs. Moreover, the one Hbs-Ag construct that retained the ability to assemble after native Tc epitope removal subsequently lost its ability to assemble following foreign epitope insertion (Cheong et al. 2009). Nevertheless, insertion of nine amino acid N-terminus of Aβ protein into an immunodominant loop of BPV L1 (Zamora et al. 2006) and insertion of conserved regions of HPV L2 protein into the AB loop of PP7 bacteriophage (Tumban et al. 2011) did not seem to interfere with assembly. These examples underline that while removal of capsid regions and replacing them with foreign sequences are likely to affect capsid protein assembly, such an impact is not predictable, and insertion without removal of native sequences is often less likely to interfere with assembly.

Bioengineering strategies to maintain assembly competency and to maintain antigen conformational structure have been reported in the literature. For example, Anggraeni et al. (Anggraeni et al. 2013) compared two bioengineering approaches when inserting the H190 loop of the influenza hemagglutinin (HA) onto the HI loop of the murine polyomavirus (MPyV) VP1 major capsid protein. Both computational

modeling and biochemical analysis confirmed that the presentation of a dual tandem repeat of H190 on VP1 surface-exposed loop produced more stably assembled VLPs in comparison to antigen presentation with flanking linker sequences. The dual tandem repeat approach also showed better antibody reactivity, which indicated that this approach might have better preserved the H190 antigen conformation. The introduction of flexible peptide linker sequences is a common strategy to reduced interference of the inserted sequence with capsid protein structure. Flanking GGG linker sequences may have helped the insertion of a hydrophobic influenza M1 Tc epitope on the surface-exposed DE and HI loops of the SV40 capsid protein, as the recombinant VLP assembled well when expressed in insect cells (Kawano et al. 2014). However, insertion of the same epitope on HI loop on the closely related MPyV VP1, expressed in a bacterial host, resulted in aggregation of the capsid subunit (Abidin et al. 2015). The different outcomes of the modification of these closely related VLP platforms may be due to the difference in the expression host. While the modified SV40 VP1 was expressed in insect cells and assembly occurred in vivo, the modified MPyV VP1 was expressed in a bacterial host and purified as capsomeres for further assembly in vitro. This underlines the important role of expression host and downstream processes in recombinant VLP bioengineering. In this case, to overcome aggregation of purified capsomeres in vitro, the inserted hydrophobic Tc epitope was flanked with double charged amino acid (DD) (Abidin et al. 2015).

3.1.2 Genetic Fusion to the N-Terminus or C-Terminus

Antigen fusion to free termini of a capsid monomer is possible and, in the case of longer sequences or structured domain insertions, preferable for antibody stimulation if either terminal domain is surface-exposed. Reported VLP platforms with surface-exposed N-terminal domain include the alfalfa mosaic virus (AIMV) (Yusibov et al. 1997), potato X virus (PVX) (Marusic et al. 2001), and the MS2 bacteriophage (Tumban et al. 2012). Fusion of the V3 loop of an HIV-1 protein or a rabies virus derived glycoprotein to the N-terminus of AIMV, and insertion of a linear B-cell epitope of the HIV gp41 protein onto the N-terminus of PVX, both expressed in plants, did not seem to interfere with assembly and stimulated considerable antibody production against the respective antigens. On the other hand, of four different HPV L2 epitopes separately fused to the N-terminus of the bacteriophage MS2 coat protein, three resulted in more regular recombinant VLPs, while one did not form VLP at all (Tumban et al. 2012). The higher proportion of constructs that form stable VLPs may have been aided by the application of a particular bioengineering strategy where the coding sequences of two capsid monomers were connected as a dual tandem repeat. This was based on structural studies which demonstrated that the C-terminus of one monomer is in close proximity with the N-terminus of an adjacent monomer in an assembled MS2 VLP. To avoid disruption in assembly by steric hindrance, the heterologous antigen was inserted only at the N-terminus of the upstream capsid monomer.

The C-terminal domains of VLP components have also been modified to stimulate either the humoral or cellular immunity. For humoral immunity, the C-terminal domain needs to be externally presented or at least available for the presentation of fused antigenic sequences on the surface. Li et al. (2021) inserted model linear T- and B-cell epitopes on the surface-exposed C-terminus of the P22 capsid protein. P22 VLPs assemble in vivo during expression in *E. coli* into stable VLPs, and the vaccine candidates produced stimulated high antibody titer and allowed T-cell antigen cross-presentation for a strong T-cell response.

Comparison of genetic fusion at different regions of the same capsid protein has provided more insights on the bioengineering of antigens on heterologous VLP platforms. Research reported in the literature once more highlight that the formation of stable recombinant VLP vaccines depends very much on the compatibility of the inserted antigens with the site of insertion and their effect on the overall structure of the recombinant VLP. Babin et al. (2013) compared influenza NP Tc epitope fusion to the N-terminal or the C-terminal domains of the papaya mosaic virus (PapMV) capsid protein. Fusion to the N-terminus yielded higher expression levels of the recombinant VLPs produced than fusion to the C-terminus. Interestingly, since fusion into the N-terminus domain results in a loss of the localization signal needed for effective expression of the original plant expression host, expression host was changed *E. coli* host where in vivo assembly was retained and expression yield was improved.

3.2 Chemical Conjugation

In addition to biological approach via genetic fusion, presentation of antigens on the surface of heterologous VLP platforms can be achieved by chemical conjugation. The advantage of chemical conjugation is the ability to not only conjugate protein antigens but also antigens of other biochemical nature such as alkaloids and polysaccharides. There are many ways to conjugate a heterologous antigen to the surface of VLPs depending on the nature of the antigen (Pokorski and Steinmetz 2011). It is most frequently carried out via common bioconjugation chemistries exploiting addressable, or surface-exposed, cysteine or lysine residues. N-hydroxysuccinimide (NHS) and maleimide reactive groups can form covalent linkages with side chains of lysine or cysteine residues, respectively, under aqueous and generally benign conditions that preserve antigenicity.

The versatility of the chemical conjugation approach has been shown in examples using the Qβ bacteriophage VLP. The alkaloid nicotine was conjugated via a succinimate linker, leading to an anti-smoking vaccine candidate that reached clinical trials (Maurer et al. 2005). The success of this approach was a result of the high density of nicotine presented on the Qβ VLP surface (565 per VLP) and subsequently high antibody response. A variation of this approach is to first introduce or modify chemical functionality, permitted further flexibility and control. Yin et al. (2013) conjugated the tumor-associated carbohydrate antigen (TACA) to

introduce alkyne groups on the VLP surface via the copper (I)-catalyzed azide-alkyne cycloaddition reaction (CuAAC), or "click" chemistry. Using this approach they were able to conjugate TACA to Qβ VLPs at different densities and demonstrate that rather than the amount of the antigen, it is the density and organized display of the antigen that play an important role in eliciting antibody production and IgG isotype switching.

Chemical conjugation can also be a useful strategy for peptide and protein display. Pastori et al. (2012) used a hetero-bifunctional linker to conjugate several alpha-helix regions of the HIV gp41 protein with the A205 bacteriophage VLP. The linker contains an NHS-ester on one end designed to bind with an introduced lysine on the surface of the A205 bacteriophage VLP, while on the other end contains a maleimide designed to bind with a terminal cysteine of different variants of the alpha-helical antigen. The conjugated VLP vaccine candidate was formed via successive mixing of the vaccine candidate components. However, in this case the number of antigens and their density on the surface of the VLP was dependent on the length and the amino acid composition of the conjugated antigen, resulting in limited control over the number and density of surface antigens. In another example, a malaria antigen, circumsporozoite protein (CSP), was conjugated as an entire protein domain to Qβ VLPs (Khan et al. 2015), resulting in a much higher antibody response than that against the nonconjugated soluble protein. As noted earlier, genetic fusion of long or structure antigenic sequences is unpredictable and frequently disrupts VLP assembly. Since chemical conjugation is performed post-assembly, it provides a means to overcome this limitation for larger antigens. Furthermore, it enables the presentation of antigens with posttranslational requirements such as glycosylation not possible in prokaryotic cells whereby antigens can be produced in suitable eukaryotic cells for subsequent conjugation to a VLP platform.

3.3 Biochemical Conjugation

Recently, a new conjugation strategy using the SpyTag-SpyCatcher system to display antigens on heterologous VLP has been reported. The SpyTag and SpyCatcher elements are optimized split components derived from the second immunoglobulin-like collagen adhesion domain (CnB2) from the fibronectin-binding protein FbaB of *Streptococcus pyogenes*. The folding of the CnB2 domain is known to be stabilized by an isopeptide bond between the C-terminal aspartate and a lysine in the protein, catalyzed by adjacent glutamine. Zakeri et al. (2012) split the C-terminal β strand of CnB2 which contains the aspartate (SpyTag) from the rest of the protein domain (SpyCatcher). Further modification and optimization (Zakeri et al. 2012; Li et al. 2014) resulted in accelerated binding just by mixing. The resulting amide bond is formed under physiological conditions, yet is resistant to extreme pH, temperature, and diverse buffer conditions and allowed the use of these split components as tag partners.

Thrane et al. (2016) tested several combinations of the tag partners on the surface of the A205 bacteriophage. They either inserted SpyCatcher to the N-terminus (SpyCatcher-VLP), SpyTag on the N-terminus (SpyTag-VLP), or both the N and C-terminus (2xSpyTag-VLP) of the A205 capsid protein. They paired these modified VLPs with their respective tag partners fused separately at either the N- or C-terminus of a variety of antigens. The antigens tested included a total of 12 antigens derived from malaria, tuberculosis, and cancer, the size of which ranged from 15 to 118 kDa. Their research showed that the coupling efficiency of a variety VLP-antigen combination using the SpyTag-SpyCatcher tag partners resulted in a 22–88% coupling efficiency, where the number of conjugated antigens decreased with increasing antigen size. Moreover, they concluded that SpyTag-VLP is preferable for small antigen display, while SpyCatcher-VLP is preferable for large antigen display. The variety of SpyCatcher-SpyTag antigen-tagged VLPs elicited high IgG titer and, in the case of self-antigens, were able to break B-cell tolerance. Similar results with malaria antigens have also been reported by Brune et al. (2016), and Janitzek et al. (2016) tested this strategy further by fusion of whole CSP with SpyCatcher and conjugating this large whole protein antigen into a SpyTag-VLP. Interestingly, while the SpyTag fused A205 VLP was produced in a bacterial host, the CSP-SpyCatcher antigen was produced in insect cells for the desired posttranslational modification important for antibody recognition. The development of the SpyCatcher-SpyTag tag partner strategy opens up new possibilities in displaying conformational antigens, especially large antigens, onto the surface of heterologous VLP platforms.

3.4 Capsid Protein Stoichiometry

Aside from modification at the capsid monomer protein level, bioengineering strategies to develop effective vaccines also encompass modification at the VLP assembly level. This includes, for example, presentation of multiple antigens or a combination of differential capsid monomers or subunits to form mosaic VLPs. Mosaic VLPs are VLPs composed of differently modified monomers or subunit proteins. Mosaic VLPs are developed for different purposes, such as to restore assembly lost due to steric hindrance or other factors, to broaden immune response, and to reduce surface hydrophobicity-related aggregation.

Tyler et al. (2014) used the co-expression strategy to assemble mosaic VLPs containing two different antigens for broader cross-protection against heterologous HPV strains. They did this by co-expressing two dual tandem repeats of either the PP7 or MS2 capsid monomer, where each of the dual tandem repeats contains different HPV L2-derived antigens. Their experiments showed that the co-expression of dual tandem repeats, each displaying different antigens, resulted in recombinant mosaic VLPs which elicited broader cross-protection against a wider spectrum of heterologous HPV strains in comparison to their single dual tandem repeat designs.

Co-expression of modified and unmodified capsid monomers or subunit have also been applied to prevent surface crowding and steric hindrance or to reduce surface hydrophobicity to prevent aggregation. Though not in an explicit vaccine context, steric hindrance that interfered with assembly was experienced by Brown et al. (2009) and Pokorski et al. (2011) when they inserted the Z protein or the epidermal growth factor (EGF), respectively, onto Qβ particles assembled in vivo. Co-expression of the modified and unmodified Qβ monomers restored the assembly competency of recombinant mosaic VLPs. The former was able to bind IgG with high capacity (Brown et al. 2009), and the latter was shown to specifically react with EGF receptor on mammalian cells (Pokorski et al. 2011).

Tekewe et al. (2017) co-expressed unmodified MPyV VP1 with modular MPyV VP1 displaying the rotavirus VP8* antigen on the surface in a bacterial host. Co-expression in the bacterial host resulted in mosaic pentameric capsomeres that could then be assembled into recombinant mosaic VLPs in vitro. Without the co-expression strategy, the pentameric subunit experienced aggregation due to the hydrophobic nature of the antigen, and mosaic subunit formation with unmodified VP1 reduced aggregation and restored the capacity for assembly. The resulting recombinant mosaic VLP induced a high level of VP8* antibodies in immunized mice. The ability of expression host cells to accommodate in vivo assembly is preferable in several ways such as shorter bioprocess steps and the inclusion of nucleic acid which may stimulate the innate immunity via TLR activation (Bessa et al. 2009). However, platforms that employ in vitro assembly provide the ability to control the composition of the mosaic VLPs as in the above example and, theoretically, the ability to modulate encapsulation of antigens inside VLPs.

3.5 Encapsulation

Bioengineering strategy to encapsulate antigens inside the cavity of heterologous VLP platform could be an option for antigens which do not need to be externally displayed to stimulate the adaptive immunity. This is especially preferable for antigens to stimulate the cellular immunity due to several reasons. First, T-cell epitopes are mostly hydrophobic, and their display on the surface of the VLPs is prone to aggregation. Second, T-cell epitope recognition does not depend on antigen conformation but rather on the ability to be taken up by APCs, especially dendritic cells, and be presented by either the MHC Class I and Class I pathways. Third, to ensure that the antigen is able to reach the intracellular compartments, encapsulation would be preferable to avoid extracellular antibody neutralization. Fourth, the ability to encapsulate larger protein antigen consisted of T-cell epitopes specific to a variety of HLAs does not only prevent hydrophobicity-related aggregation but also protect broadly against heterologous strains and a broader population coverage.

The MPyV platform has been shown to provide programmed encapsulation of heterologous proteins and even to co-encapsulation of multiple proteins during in vitro assembly (Dashti et al. 2018). Incorporation is driven by specific interactions

between the pentameric capsomeres and the C-terminus of the minor capsid protein, VP2 (Abbing et al. 2004; Boura et al. 2005; Pleckaityte et al. 2015), and examples where production and in vivo assembly have been carried out in enkaryotic cells show the potential of this platform as carriers of encapsulated antigens. Eriksson et al. (2011) fused the prostate-specific antigen (PSA) to the C-terminus of VP2 and demonstrated that while no PSA-specific antibodies where produced, dendritic cells loaded with PSA-MPyV VLPs strengthen protection against PSA-expressing tumors. Cellular immunity such as Tc and Th response was also high in loaded dendritic cell treatment, indicating efficient delivery of the antigen to the cytosol of these cells. Similar results were shown when HER2, a breast cancer antigen, was fused to MPyV VP2 instead of PSA (Tegerstedt et al. 2007).

Co-encapsidation of two different proteins, such as the M and M2 proteins of the respiratory syncytial virus (RSV) inside the P22 VLP, has also been reported (Schwarz et al. 2016). The P22 VLP is able to encapsulate heterologous proteins fused to the C-terminus of the scaffold protein that binds to the interior of the capsid during assembly. Schwarz et al. concatenated the M and M2 gene for encapsulation in this manner. Despite the quaternary structure of the fusion, VLPs of well-rounded and correct sizes bearing both the CP and M/M2-SP proteins were assembled in vivo when co-expressed in *E. coli* (Schwarz et al. 2016). Administration of the M/M2-P22 VLP alone resulted in reduced lung viral titer in RSV challenged mice and elevated Tc response. However, it was not known whether loading M/M2 P22 VLP into dendritic cells prior to immunization in mice would improve protection and Tc response.

4 Bioengineering Strategies for Bioprocess Optimization

As has been described throughout this review, the bioengineering of VLPs to incorporate heterologous antigens often poses challenges to be addressed for successful manufacturing. These challenges include improvement in overall bioprocess development; overcoming biosynthetic challenges such as soluble protein yield, aggregation, upstream, and downstream method selection and validation; and reducing overall production cost. A few example approaches to address these challenges are described in this section that demonstrate the effectiveness of designing process optimization experiments with increased throughput.

4.1 Upstream Bioprocess Optimization

The optimal result in protein expression at the upstream bioprocess depends on many interacting factors such as temperature, oxygen availability, initial culture concentration, culture acidity, additives, inducer concentration, induction time, and harvest time. Assessing these factors one by one or in combination in laboratory scale would

be resource, labor, and time-consuming. Therefore, a high-throughput method to assess different culture condition combinations at once at milliliter scale would be a versatile and informative option to optimize upstream bioprocess.

In producing vaccine candidates based in MPyV VP1, Lad Effio et al. (2016a) examined a triplicate of 16 culture conditions varied by pH, the presence of additives, 2 different inducer concentrations, and 2 different induction times in a miniaturized 48 well plates which were then incubated at 3 different temperatures and 2 different shaking speeds. The high-throughput multiparametric culture conditions screen aimed to discover the best culture condition for the expression of untagged MPyV VP1 protein in *E. coli*, adapted from GST-tagged MPyV VP1 upstream platform previously reported (Chuan et al. 2008). The results showed that the best culture condition for the untagged protein was different from the previously reported optimum condition for GST-tagged MPyV VP1. The selected condition was scalable in 2.5 L culture flasks, and the amount of purified VP1 obtained from this laboratory scale expression was similar to quantitated soluble VP1 derived from the miniaturized culture. This work showed that a high-throughput culture condition screen can be a robust approach for bacterially produced VLP subunits and would be useful in upstream bioprocess optimization of other microbially expressed industrial proteins.

4.2 Downstream Bioprocess Optimization by High-Throughput Screening

Along with upstream bioprocess optimization, high-throughput approaches to optimize downstream bioprocess in bioengineering heterologous antigens of VLP platforms have also been reported. Abidin et al. (2015) reported the use of a multiparametric high-throughput buffer screen to determine the most suitable condition to prevent aggregation of MPyV capsomeres displaying hydrophobic epitopes. The buffer screen tested 40 buffer conditions in a miniaturized 96-well format examining 6 different additives across 5 different pH. The results, based on spectrophotometry and light scattering analysis, showed that the addition of L-arginine (L-Arg) was the most promising additive to recover soluble capsomere during tag removal before in vitro assembly. Tekewe et al. (2015) also reported light scattering-based rapid screening method to select appropriate additives to stabilize MPyV capsomeres displaying rotaviral epitopes, though at lower throughput.

The development of VLP-based vaccines necessitates a robust VLP purification method to isolate VLPs assembled in vivo or in vitro. An ideal VLP purification method would have a good resolution to separate between aggregates, VLPs, and unassembled VLP subunits as well as capacity for rapid throughput for condition screening. The current methods to assess the quality of developed VLP purification methods involve TEM and light scattering analysis. Ladd Effio et al. (2016b) recently reported a novel high-throughput interlaced size exclusion-ultrahigh-

performance chromatography (SE-UHPLC) which allows rapid assessment of VLP purification conditions. The interlaced SE-UHPLC was designed to shorten analysis time by overlapping the hold-up phase of one sample application with the lag phase of the next sample application, reducing run times to approximately 3 minutes per sample. The analysis tool was validated with multiple commercially available VLP vaccines as well as *E. coli*-produced hepatitis B and MPyV VLPs. The particle size distribution and aggregate percentage analysis of the iSE-UHPLC were compared with dynamic light scattering (DLS) and TEM analysis, and it was shown that the iSE-HPLC was more sensitive in detecting aggregates and more accurate in calculating particle size. The iSE-HPLC analysis tool was then used to assess four VLP purification methods and conduct stability studies.

4.3 Platform Development to Reduce Production Cost

One of the important aspects of vaccine development, besides as a preventive and therapeutic measure, is the reduction in production cost and time. Production cost and production time are necessary aspects to ensure that the vaccines are affordable, widely distributable, and timely available.

A bacterial platform and process was developed for the production of the MPyV VLP platform as a commercially viable system using *E. coli* as the expression host (Chuan et al. 2008; Liew et al. 2010). Briefly, at laboratory scale, capsomeres are produced in bacteria (Chuan et al. 2008; Middelberg et al. 2011) and harvested as capsomeres by affinity chromatography and further purified by gel filtration after tag removal (Lipin et al. 2009). The capsomeres can be maintained in capsomere form or assembled into VLP in vitro using optimized buffer conditions (Chuan et al. 2010). The laboratory scale production has subsequently been translated into a large-scale process, involving fed-batch *E. coli* expression with a volumetric protein yield of 4.38 g/l which is 15-fold laboratory scale production (Liew et al. 2010). Capsomere purification can also be achieved by ion exchange chromatography (IEX) (Chuan et al. 2014), and assembly into VLP can be performed by either diafiltration or dilution methods (Liew et al. 2012a; Liew et al. 2012b). Optimized assembly buffer for the assembly via diafiltration increased yield by 42–56% in comparison to laboratory scale assembly by dialysis (Liew et al. 2012a). Analysis of this extensive process development estimated the cost of optimized large-scale production of capsomeres and VLPs at 500 L and 1500 L volume, respectively, to be at less than 1 cent (USD) per dose (50 μg per dose). Moreover, it was predicted that in a 10 k-L production scale, 320 million capsomere or VLP vaccine doses could be produced within 2.3 or 4.7 days, respectively (Chuan et al. 2014). In addition, an ammonium sulfate-based precipitation method to harvest MPyV capsomeres to skip laborious and costly chromatographic process has also been developed (Wibowo et al. 2015). The precipitation method was successfully used to purify mosaic capsomeres, and in vitro assembled VLPs elicited high titer antibody in immunized mice (Tekewe et al. 2017). Although this approach is not suitable for human vaccine production

due to unacceptable levels of host cell contaminants, it may stimulate use in veterinary settings and demonstrate the remarkable capacity of prokaryotes for cost-effective vaccine production.

5 Conclusions and Perspective

VLPs have a now long track record as successful subunit vaccines. Although those VLPs that are currently available commercially as vaccines are designed to raise immunity against the virus from which they are derived, there is a growing interest in the development of VLPs as carriers of heterologous antigens. Their inherent physical properties, leading to high immunogenicity, make them ideal platforms to develop as vaccines against not just infectious diseases but also nonpathogen-related cancers and other chronic diseases and conditions. They are amenable to precise modification by genetic or chemical means and are characterized by structural fidelity. As we have described herein, the use of prokaryotes, specifically *E. coli*, to produce VLPs provides an accessible and low-cost source of VLPs or VLP subunits for further development. One of the drawbacks of marketed VLPs vaccines has been relatively high cost, limiting their immediate use to developed countries. Bioengineering and bioprocessing improvements, including the use of prokaryotic production hosts such as *E. coli*, aim to reduce the cost and improve the efficacy of VLP vaccines. While challenges remain with regard to the manufacture of VLP vaccines incorporating heterologous antigens, a myriad of bioengineering approaches combined with recent developments in bioprocessing hold great promise.

References

Abbing A, Blaschke UK, Grein S, Kretschmar M, Stark CM, Thies MJ, Walter J, Weigand M, Woith DC, Hess J, Reiser CO (2004) Efficient intracellular delivery of a protein and a low molecular weight substance via recombinant polyomavirus-like particles. J Biol Chem 279(26): 27410–27421

Abidin RS, Lua LH, Middelberg AP, Sainsbury F (2015) Insert engineering and solubility screening improves recovery of virus-like particle subunits displaying hydrophobic epitopes. Protein Sci 24(11):1820–1828

Anggraeni MR, Connors NK, Wu Y, Chuan YP, Lua LH, Middelberg AP (2013) Sensitivity of immune response quality to influenza helix 190 antigen structure displayed on a modular virus-like particle. Vaccine 31(40):4428–4435

Babin C, Majeau N, Leclerc D (2013) Engineering of papaya mosaic virus (PapMV) nanoparticles with a CTL epitope derived from influenza NP. J Nanobiotechnol 11:10

Bachmann MF, Jennings GT (2010) Vaccine delivery: a matter of size, geometry, kinetics and molecular patterns. Nat Rev Immunol 10(11):787–796

Bachmann MF, Zinkernagel RM (1997) Neutralizing antiviral B cell responses. Annu Rev Immunol 15:235–270

Barrington R, Zhang M, Fischer M, Carroll MC (2001) The role of complement in inflammation and adaptive immunity. Immunol Rev 180:5–15

Bayer ME, Blumberg BS, Werner B (1968) Particles associated with Australia antigen in the sera of patients with leukaemia, Down's Syndrome and hepatitis. Nature 218(5146):1057–1059

Bessa J, Jegerlehner A, Hinton HJ, Pumpens P, Saudan P, Schneider P, Bachmann MF (2009) Alveolar macrophages and lung dendritic cells sense RNA and drive mucosal IgA responses. J Immunol 183(6):3788–3799

Blumberg BS, Alter HJ, Visnich S (1965) A "new" antigen in leukemia sera. JAMA 191:541–546

Boura E, Liebl D, Spisek R, Fric J, Marek M, Stokrova J, Holan V, Forstova J (2005) Polyomavirus EGFP-pseudocapsids: analysis of model particles for introduction of proteins and peptides into mammalian cells. FEBS Lett 579(29):6549–6558

Braun H, Boller K, Lower J, Bertling WM, Zimmer A (1999) Oligonucleotide and plasmid DNA packaging into polyoma VP1 virus-like particles expressed in *Escherichia coli*. Biotechnol Appl Biochem 29:31–43

Brown SD, Fiedler JD, Finn MG (2009) Assembly of hybrid bacteriophage Qbeta virus-like particles. Biochemistry 48(47):11155–11157

Brune KD, Leneghan DB, Brian IJ, Ishizuka AS, Bachmann MF, Draper SJ, Biswas S, Howarth M (2016) Plug-and-display: decoration of virus-like particles via isopeptide bonds for modular immunization. Sci Rep 6:19234

Chackerian B, Lowy DR, Schiller JT (1999) Induction of autoantibodies to mouse CCR5 with recombinant papillomavirus particles. Proc Natl Acad Sci U S A 96(5):2373–2378

Cheong WS, Reiseger J, Turner SJ, Boyd R, Netter HJ (2009) Chimeric virus-like particles for the delivery of an inserted conserved influenza A-specific CTL epitope. Antivir Res 81(2):113–122

Chuan YP, Fan YY, Lua LH, Middelberg AP (2010) Virus assembly occurs following a pH- or Ca2+−triggered switch in the thermodynamic attraction between structural protein capsomeres. J R Soc Interface 7(44):409–421

Chuan YP, Lua LH, Middelberg AP (2008) High-level expression of soluble viral structural protein in *Escherichia coli*. J Biotechnol 134(1–2):64–71

Chuan YP, Wibowo N, Lua LHL, Middelberg APJ (2014) The economics of virus-like particle and capsomere vaccines. Biochem Eng J 90:255

Dashti NH, Abidin RS, Sainsbury F (2018) Programmable in vitro coencapsidation of guest proteins for intracellular delivery by virus-like particles. ACS Nano 12(5):4615–4623

Eriksson M, Andreasson K, Weidmann J, Lundberg K, Tegerstedt K, Dalianis T, Ramqvist T (2011) Murine polyomavirus virus-like particles carrying full-length human PSA protect BALB/c mice from outgrowth of a PSA expressing tumor. PLoS One 6(8):e23828

Glasgow J, Tullman-Ercek D (2014) Production and applications of engineered viral capsids. Appl Microbiol Biotechnol 98(13):5847–5858

Hagensee ME, Yaegashi N, Galloway DA (1993) Self-assembly of human papillomavirus type 1 capsids by expression of the L1 protein alone or by coexpression of the L1 and L2 capsid proteins. J Virol 67(1):315–322

Janitzek CM, Matondo S, Thrane S, Nielsen MA, Kavishe R, Mwakalinga SB, Theander TG, Salanti A, Sander AF (2016) Bacterial superglue generates a full-length circumsporozoite protein virus-like particle vaccine capable of inducing high and durable antibody responses. Malar J 15(1):545

Jegerlehner A, Storni T, Lipowsky G, Schmid M, Pumpens P, Bachmann MF (2002) Regulation of IgG antibody responses by epitope density and CD21-mediated costimulation. Eur J Immunol 32(11):3305–3314

Kawano M, Morikawa K, Suda T, Ohno N, Matsushita S, Akatsuka T, Handa H, Matsui M (2014) Chimeric SV40 virus-like particles induce specific cytotoxicity and protective immunity against influenza a virus without the need of adjuvants. Virology 448:159–167

Khan F, Porter M, Schwenk R, DeBot M, Saudan P, Dutta S (2015) Head-to-head comparison of soluble vs. Qβ VLP circumsporozoite protein vaccines reveals selective enhancement of NANP repeat responses. PLoS One 10(11):e0142035

Kirnbauer R, Booy F, Cheng N, Lowy DR, Schiller JT (1992) Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc Natl Acad Sci U S A 89(24):12180–12184

Kushnir N, Streatfield SJ, Yusibov V (2012) Virus-like particles as a highly efficient vaccine platform: diversity of targets and production systems and advances in clinical development. Vaccine 31(1):58–83

Ladd Effio C, Baumann P, Weigel C, Vormittag P, Middelberg A, Hubbuch J (2016a) High-throughput process development of an alternative platform for the production of virus-like particles in *Escherichia coli*. J Biotechnol 219:7–19

Ladd Effio C, Oelmeier SA, Hubbuch J (2016b) High-throughput characterization of virus-like particles by interlaced size-exclusion chromatography. Vaccine 34(10):1259–1267

Lee KL, Twyman RM, Fiering S, Steinmetz NF (2016) Virus-based nanoparticles as platform technologies for modern vaccines. Wiley Interdiscip Rev Nanomed Nanobiotechnol 8(4): 554–578

Li L, Fierer JO, Rapoport TA, Howarth M (2014) Structural analysis and optimization of the covalent association between SpyCatcher and a peptide tag. J Mol Biol 426(2):309–317

Li M, Cripe TP, Estes PA, Lyon MK, Rose RC, Garcea RL (1997) Expression of the human papillomavirus type 11 L1 capsid protein in *Escherichia coli*: characterization of protein domains involved in DNA binding and capsid assembly. J Virol 71(4):2988–2995

Li W, Jing Z, Wang S, Li Q, Xing Y, Shi H, Li S, Hong Z (2021) P22 virus-like particles as an effective antigen delivery nanoplatform for cancer immunotherapy. Biomaterials 271:120726

Liew MW, Rajendran A, Middelberg AP (2010) Microbial production of virus-like particle vaccine protein at gram-per-litre levels. J Biotechnol 150(2):224–231

Liew MWO, Chuan YP, Middelberg APJ (2012a) Reactive diafiltration for assembly and formulation of virus-like particles. Biochem Eng J 68:120–128

Liew MWO, Chuan YP, Middelberg APJ (2012b) High-yield and scalable cell-free assembly of virus-like particles by dilution. Biochem Eng J 67:88–96

Lipin DI, Raj A, Lua LH, Middelberg AP (2009) Affinity purification of viral protein having heterogeneous quaternary structure: modeling the impact of soluble aggregates on chromatographic performance. J Chromatogr A 1216(30):5696–5708

Lua LH, Connors NK, Sainsbury F, Chuan YP, Wibowo N, Middelberg AP (2014) Bioengineering virus-like particles as vaccines. Biotechnol Bioeng 111(3):425–440

Marusic C, Rizza P, Lattanzi L, Mancini C, Spada M, Belardelli F, Benvenuto E, Capone I (2001) Chimeric plant virus particles as immunogens for inducing murine and human immune responses against human immunodeficiency virus type 1. J Virol 75(18):8434–8439

Maurer P, Jennings GT, Willers J, Rohner F, Lindman Y, Roubicek K, Renner WA, Muller P, Bachmann MF (2005) A therapeutic vaccine for nicotine dependence: preclinical efficacy, and phase I safety and immunogenicity. Eur J Immunol 35(7):2031–2040

Middelberg AP, Rivera-Hernandez T, Wibowo N, Lua LH, Fan Y, Magor G, Chang C, Chuan YP, Good MF, Batzloff MR (2011) A microbial platform for rapid and low-cost virus-like particle and capsomere vaccines. Vaccine 29(41):7154–7162

Mitic NS, Pavlovic MD, Jandrlic DR (2014) Epitope distribution in ordered and disordered protein regions - part a. T-cell epitope frequency, affinity and hydropathy. J Immunol Methods 406:83–103

Murphy K, Travers P, Walport M, Janeway C (2008) Janeway's Immunobiology, 7th edn. Garland Science, New York

Oussoren C, Zuidema J, Crommelin DJ, Storm G (1997) Lymphatic uptake and biodistribution of liposomes after subcutaneous injection. II. Influence of liposomal size, lipid composition and lipid dose. Biochim Biophys Acta 1328(2):261–272

Pastori C, Tudor D, Diomede L, Drillet AS, Jegerlehner A, Rohn TA, Bomsel M, Lopalco L (2012) Virus like particle based strategy to elicit HIV-protective antibodies to the alpha-helic regions of gp41. Virology 431(1–2):1–11

Pattinson DJ, Apte SH, Wibowo N, Chuan YP, Rivera-Hernandez T, Groves PL, Lua LH, Middelberg APJ, Doolan DL (2019) Chimeric murine polyomavirus virus-like particles induce plasmodium antigen-specific CD8+ T cell and antibody responses. Front Cell Infect Microbiol 9:215

Pleckaityte M, Bremer CM, Gedvilaite A, Kucinskaite-Kodze I, Glebe D, Zvirbliene A (2015) Construction of polyomavirus-derived pseudotype virus-like particles displaying a functionally active neutralizing antibody against hepatitis B virus surface antigen. BMC Biotechnol 15:85

Pokorski JK, Hovlid ML, Finn MG (2011) Cell targeting with hybrid Qbeta virus-like particles displaying epidermal growth factor. Chembiochem 12(16):2441–2447

Pokorski JK, Steinmetz NF (2011) The art of engineering viral nanoparticles. Mol Pharm 8(1): 29–43

Reddy ST, Rehor A, Schmoekel HG, Hubbell JA, Swartz MA (2006) In vivo targeting of dendritic cells in lymph nodes with poly(propylene sulfide) nanoparticles. J Control Release 112(1): 26–34

Rivera-Hernandez T, Hartas J, Wu Y, Chuan YP, Lua LH, Good M, Batzloff MR, Middelberg AP (2013) Self-adjuvanting modular virus-like particles for mucosal vaccination against group a streptococcus (GAS). Vaccine 31(15):1950–1955

Rodriguez-Limas WA, Sekar K, Tyo KE (2013) Virus-like particles: the future of microbial factories and cell-free systems as platforms for vaccine development. Curr Opin Biotechnol 24(6):1089–1093

Salunke DM, Caspar DL, Garcea RL (1986) Self-assembly of purified polyomavirus capsid protein VP1. Cell 46(6):895–904

Schiller J, Chackerian B (2014) Why HIV virions have low numbers of envelope spikes: implications for vaccine development. PLoS Pathog 10(8):e1004254

Schwarz B, Morabito KM, Ruckwardt TJ, Patterson DP, Avera J, Miettinen HM, Graham BS, Douglas T (2016) Virus like particles encapsidating respiratory syncytial virus M and M2 proteins induce robust T cell responses. ACS Biomaterials Sci Eng 2(12):2324–2332

Swartz MA (2001) The physiology of the lymphatic system. Adv Drug Deliv Rev 50(1–2):3–20

Tegerstedt K, Franzen A, Ramqvist T, Dalianis T (2007) Dendritic cells loaded with polyomavirus VP1/VP2Her2 virus-like particles efficiently prevent outgrowth of a Her2/neu expressing tumor. Cancer Immunol Immunother 56(9):1335–1344

Tekewe A, Connors NK, Sainsbury F, Wibowo N, Lua LH, Middelberg APJ (2015) A rapid and simple screening method to identify conditions for enhanced stability of modular vaccine candidates. Biochem Eng J 100:50–58

Tekewe A, Fan Y, Tan E, Middelberg AP, Lua LH (2017) Integrated molecular and bioprocess engineering for bacterially produced immunogenic modular virus-like particle vaccine displaying 18 kDa rotavirus antigen. Biotechnol Bioeng 114(2):397–406

Teunissen EA, de Raad M, Mastrobattista E (2013) Production and biomedical applications of virus-like particles derived from polyomaviruses. J Control Release 172(1):305–321

Thrane S, Janitzek CM, Matondo S, Resende M, Gustavsson T, de Jongh WA, Clemmensen S, Roeffen W, van de Vegte-Bolmer M, van Gemert GJ, Sauerwein R, Schiller JT, Nielsen MA, Theander TG, Salanti A, Sander AF (2016) Bacterial superglue enables easy development of efficient virus-like particle based vaccines. J Nanobiotechnol 14:30

Tumban E, Peabody J, Peabody DS, Chackerian B (2011) A pan-HPV vaccine based on bacteriophage PP7 VLPs displaying broadly cross-neutralizing epitopes from the HPV minor capsid protein, L2. PLoS One 6(8):e23310

Tumban E, Peabody J, Tyler M, Peabody DS, Chackerian B (2012) VLPs displaying a single L2 epitope induce broadly cross-neutralizing antibodies against human papillomavirus. PLoS One 7(11):e49751

Tyler M, Tumban E, Peabody DS, Chackerian B (2014) The use of hybrid virus-like particles to enhance the immunogenicity of a broadly protective HPV vaccine. Biotechnol Bioeng 111(12): 2398–2406

Wibowo N, Chuan YP, Lua LH, Middelberg AP (2012) Modular engineering of a microbially-produced viral capsomere vaccine. Chem Eng Sci 103:12–20

Wibowo N, Wu Y, Fan Y, Meers J, Lua LH, Middelberg AP (2015) Non-chromatographic preparation of a bacterially produced single-shot modular virus-like particle capsomere vaccine for avian influenza. Vaccine 33(44):5960–5965

Yan D, Wei YQ, Guo HC, Sun SQ (2015) The application of virus-like particles as vaccines and biological vehicles. Appl Microbiol Biotechnol 99(24):10415–10432

Yin Z, Comellas-Aragones M, Chowdhury S, Bentley P, Kaczanowska K, Benmohamed L, Gildersleeve JC, Finn MG, Huang X (2013) Boosting immunity to small tumor-associated carbohydrates with bacteriophage qbeta capsids. ACS Chem Biol 8(6):1253–1262

Yusibov V, Modelska A, Steplewski K, Agadjanyan M, Weiner D, Hooper DC, Koprowski H (1997) Antigens produced in plants by infection with chimeric plant viruses immunize against rabies virus and HIV-1. Proc Natl Acad Sci U S A 94(11):5784–5788

Zakeri B, Fierer JO, Celik E, Chittock EC, Schwarz-Linek U, Moy VT, Howarth M (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc Natl Acad Sci U S A 109(12):E690–E697

Zamora E, Handisurya A, Shafti-Keramat S, Borchelt D, Rudow G, Conant K, Cox C, Troncoso JC, Kirnbauer R (2006) Papillomavirus-like particles are an effective platform for amyloid-beta immunization in rabbits and transgenic mice. J Immunol 177(4):2662–2670

Zeltins A (2013) Construction and characterization of virus-like particles: a review. Mol Biotechnol 53(1):92–107

Functional Inclusion Bodies

Ricardo Baltà-Foix, Ramon Roca-Pinilla, Adria López-Cano, Laia Gifre-Renom, Anna Arís, and Elena Garcia-Fruitós

Contents

Abstract Inclusion bodies (IBs) are protein aggregates formed under recombinant protein production processes in microbial cell factories. Their characterization has shown that they are self-assembling and biologically active protein nanoparticles with promising properties for a wide range of applications, including biocatalysis, tissue engineering, and therapy. Besides, different protocols have also been developed to obtain soluble protein from IBs using non-denaturing conditions.

R. Baltà-Foix · A. López-Cano · A. Arís · E. Garcia-Fruitós (✉)
Department of Ruminant Production, Institute of Agriculture and Food Research and Technology (IRTA), Barcelona, Spain
e-mail: anna.aris@irta.cat; elena.garcia@irta.cat

R. Roca-Pinilla
Translational Vectorology Research Unit, Children's Medical Research Institute, Faculty of Medicine and Health, The University of Sydney, Westmead, NSW, Australia

L. Gifre-Renom
Centre for Molecular and Vascular Biology, Department of Cardiovascular Sciences, KU Leuven, Leuven, Belgium

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_11

1 Introduction

Inclusion bodies (IBs) are protein aggregates formed during recombinant protein production in microbial cell factories (Fig. 1). For decades they have been considered a useless byproduct, and various strategies have been developed aimed to increase protein solubility thereby reducing the aggregation phenomenon. However, for over 15 years now, this vision has started to change. Different research groups have demonstrated that bacterial IBs are far from being merely inactive recombinant protein deposits (de Marco et al. 2019; Rinas et al. 2017). They have been characterized as self-assembling protein nanoparticles with a dual composition: (1) an amyloid-like scaffold and (2) folded or partially folded protein species with β-amyloid structure (Cano-Garrido et al. 2013; Cano-Garrido et al. 2016). The folded or partially folded protein conformers are biologically active and can be easily released from the scaffold under physiological conditions (Seras-Franzoso et al. 2016; Carratalá et al. 2021a). In addition, the formation of these protein aggregates is a general phenomenon (Villaverde et al. 2015), rather than being a specific trait of *Escherichia coli*, observed in different microbial cell factories, including *Lactococcus lactis* (Cano-Garrido et al. 2016) and *Pichia pastoris (*Rueda et al. 2016*;* Carratalá et al. 2020a*)*. Thus, IBs can be easily produced through a scalable process, and straightforward purification protocols have been optimized consisting in multiple centrifugation and washing steps (Seras-Franzoso et al. 2015). Altogether, IB features have radically changed the concept of protein aggregation and become an attractive protein-based biomaterial with a variety of applications, including biocatalysis, biomedical therapy, and tissue engineering (de Marco et al. 2019; Rinas et al. 2017; Gifre-Renom et al. 2020a; Hrabárová et al. 2015; Liovic et al. 2012; García-Fruitós et al. 2012). Moreover, the demonstration that IBs are nanoparticles containing proteins with their native structure has led to a modification of the protocols used to obtain soluble protein from IBs which proven that denaturing and resolubilization steps are not necessary (Gifre-Renom et al. 2018; Singhvi

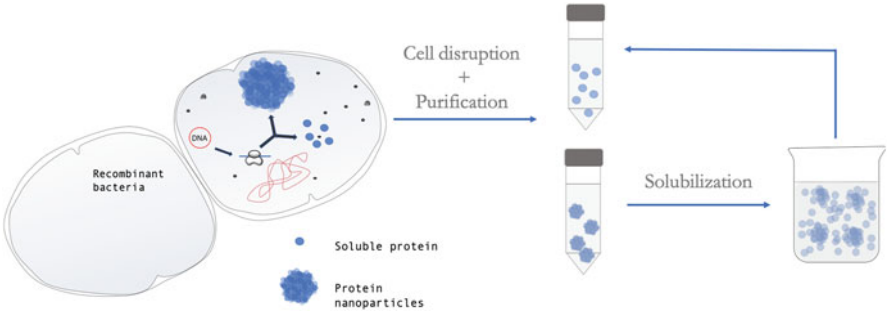


Fig. 1 Recombinant protein production and purification process. Recombinant bacteria produce the protein of interest in both soluble and aggregated (nanoparticle or inclusion bodies) forms. After that, cells are disrupted and soluble protein and protein nanoparticles are purified. If necessary, the soluble form can be obtained from inclusion bodies through a solubilization process

et al. 2020; Singh et al. 2015a; Ferrer-Miralles et al. 2018). Alternatively, mild or non-denaturing solubilization protocols have been developed by different research groups to solubilize proteins from IBs maintaining protein structure and activity (Fig. 1).

2 Protein Production and IB Formation

The bioactivity of releasable proteins forming IBs coupled with the high stability of these protein nanoparticles has made IBs an appealing alternative form to their soluble counterpart (Gifre-Renom et al. 2020a; Gifre-Renom et al. 2020b; Pesarrodona et al. 2019). Thus, contrary to what has been done for years, several research groups have focused on optimizing the formation and purification of IBs (Seras-Franzoso et al. 2015). Production time and temperature are the two cultivation parameters that can significantly influence the formation of aggregates (Vera et al. 2007; Garcia-Fruitós 2009). Increased production temperature has an impact not only on the size of IBs, i.e., larger size at higher growth temperatures (García-Fruitós et al. 2007), but also on the conformational quality of the aggregated proteins (Vera et al. 2007). Lower growth temperatures improve the conformational quality of both soluble and insoluble protein fractions (Vera et al. 2007; Jevsevar et al. 2005), meaning that IBs containing more active proteins can be easily obtained. The bacterial strain used to produce the recombinant protein of interest can also determine the properties of these proteins. Whereas in most cases IBs are spherical-like nanoparticles (Cano-Garrido et al. 2016; Garcia-Fruitós 2009; García-Fruitós et al. 2010), some *E. coli* strains produced tear-shaped aggregates with particular characteristics for tissue engineering applications (García-Fruitós et al. 2010). IBs of the same proteins produced in different bacterial systems (*E. coli* and *L. lactis*) also show differences in size and surface functional group density when used for micropatterned surface decoration (Martínez-Miguel et al. 2020).

Another strategy that has been used to optimize the formation of IBs is based on the use of protein tags based on specific aggregation-prone proteins, including the foot-and-mouth disease virus capsid protein (VP1) (García-Fruitós et al. 2005a), a variant of the human β-amyloid peptide (Morell et al. 2008), the maltose-binding protein mutant (Arié et al. 2006), poxB from *Paenibacillus polymyxa* (Park et al. 2012), and the cellulose-binding domain of *Clostridium cellulovorans (*Nahalka and Nidetzky 2007*)*. Also, shorter aggregation peptides (Carratalá et al. 2020a; Carratalá et al. 2021b; Wang et al. 2015a; Wu et al. 2011; Zhou et al. 2012; Jiang et al. 2019; Küsters et al. 2021), such as coiled-coil domains (Küsters et al. 2021; Jäger et al. 2019; Jäger et al. 2018; Gil-Garcia et al. 2020; Lamm et al. 2020) and leucine zippers (Choi et al. 2014; Roca-Pinilla et al. 2020a), have been successfully used to promote protein aggregation. Although all these tags have been described to promote aggregation, a screening for each specific protein would be needed to gain optimal results.

3 Structure, Composition, and Activity of IBs

The study and characterization of IBs have been mainly focused on their structure and composition. Regarding the structure, as already mentioned, there are two main parts conforming the aggregates: the β-sheet skeleton, which is a common structure in all IBs (García-Fruitós et al. 2011; de Groot et al. 2009; Castillo et al. 2011), and the fractions formed by the native or native-like recombinant proteins, which is protein-dependent (Rinas et al. 2017). The former provides the mechanical and chemical stability, while the latter is responsible for the specific activity of the IBs (Carratalá et al. 2020a; García-Fruitós et al. 2005a).

Despite all IBs present a common structural pattern, their composition varies. It depends on the recombinant protein that is being produced but also on the specific recombinant cell factory used. Thus, other molecules or impurities such as host cell proteins like chaperones, lipids, lipopolysaccharide (LPS), and/or nucleic acids, can be accumulated inside the IBs increasing the variability of their composition (Roca-Pinilla et al. 2020a; Rinas and Bailey 1992; Valax and Georgiou 1993). Therefore, the molecular complexity of IBs requires a huge variety of techniques for their characterization (Rinas et al. 2017). In this section, techniques extensively used to describe IBs physicochemical and biological features are provided in detail (Fig. 2).

Transmission electron microscopy (TEM) and some of its variants, such as Cryo-TEM, are techniques used for the visualization of intracellular structures, such as organelles or membranes, at high resolution. Furthermore, it is possible to detect the cell components, including IBs, at nanoscale. Various publications have shown that these microscopical techniques allow to identify the presence of the IBs and to determine their exact location inside the recombinant cell factories (Rinas et al. 2017; Cano-Garrido et al. 2016; Rueda et al. 2016; Zhou et al. 2012). Normally, IBs present a polar distribution inside the producer cells, showing an electrodense pattern under the microscope. Moreover, using purified IBs as a sample, it is possible to

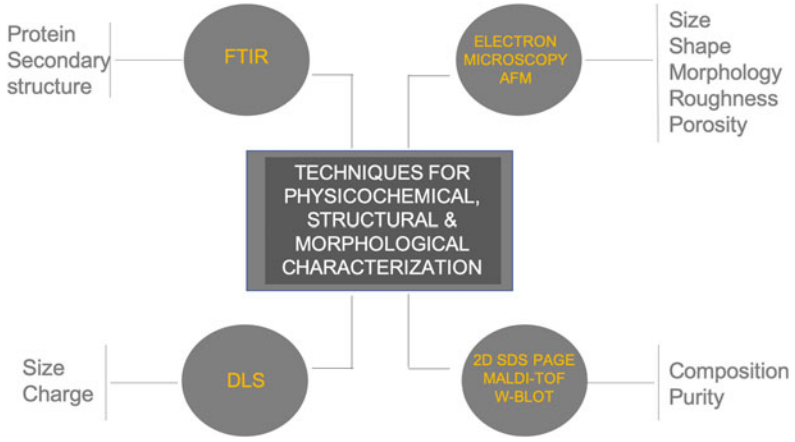


Fig. 2 Main techniques for IB characterization

determine their size and shape (Cano-Garrido et al. 2013; Garcia-Fruitós 2009; García-Fruitós et al. 2010; Zhou et al. 2012; Wang et al. 2008). The size of these protein aggregates usually ranges between 50 and 500 nm with a pseudospherical shape in most of the cases.

Scanning electron microscopy (SEM) is another technique widely used to determine physical parameters of IBs like shape, roughness, volume, and size (Cano-Garrido et al. 2016; Pesarrodona et al. 2019; Díez-Gil et al. 2010), which complements TEM analysis. SEM allows to work directly with IBs samples without any preparation prior to their visualization. For that, both SEM and TEM are usually used together in many IBs studies, obtaining a complete physical characterization of these protein aggregates (Carratalá et al. 2020a; Garcia-Fruitós 2009; Torrealba et al. 2016a). With 3D visualization, it has been possible to observe a porous structure in IBs, which is an interesting property not only for its application as biocatalysts but also as biomaterials to decorate surfaces and promote cell adhesion and proliferation (see the following sections).

Atomic force microscopy (AFM) is a microscopy technique with a wide applicability in IBs research (Garcia-Fruitós 2009; Díez-Gil et al. 2010; Sanagavarapu et al. 2019). This microscopy technique perfectly completes the information obtained by TEM and SEM, capable of characterizing morphology and stiffness of the samples directly under environmental conditions.

Dynamic light scattering (DLS) is used to measure the size distribution of nanoparticles or protein samples in solution. It allows researchers to determine the volume size distribution of soluble proteins in either unassembled or monomeric form (Unzueta et al. 2020), as well as the size and zeta potential of IBs (Carratalá et al. 2020a; Garcia-Fruitós 2009; Díez-Gil et al. 2010). The zeta potential provides information on the superficial charge of IBs, in which negatively charged IBs indicate the aggregation-prone nature of these protein-based aggregates (Garcia-Fruitós 2009; Díez-Gil et al. 2010). Furthermore, in some cases, DLS is used in the stability studies of IBs, by performing the measurements at different time points to see if the physical properties are maintained. DLS is very useful to obtain the particles' mean diameter when the sample is homogenous, and there is no important variance in the diameter of the particles. Moreover, DLS can differentiate populations of different particle sizes within the sample if the size of each population is homogenous (Martínez-Miguel et al. 2020). Thus, DLS offers many possibilities for the study of IBs, as it is relatively easy to measure under environmental conditions (i.e., buffers, solution, etc.).

Other techniques have allowed studying in detail the secondary structure of proteins forming these aggregates. Fourier transform infrared spectroscopy (FTIR) is the most widely used technique for this purpose, and different articles have proven that IBs contain correctly folded proteins but also an amyloid structure inside the IBs (Jevsevar et al. 2005; Natalello and Doglia 2015; Ami et al. 2006). FTIR is capable to detect and differentiate between β-sheet and α-helix structures, opening the possibility to define the structure and the organization of the proteins inside the aggregates (Garcia-Fruitós 2009). FTIR also allows researchers to establish comparisons between the soluble and aggregated forms of the same protein, elucidating

differences in the structure or the appearance of new patterns during the aggregation process (Cano-Garrido et al. 2016; Wang 2009). This fact is an important advantage compared to other techniques used for secondary structure analyses, such as circular dichroism (CD), that only permits the analysis of soluble protein samples. Specifically, the pattern described in IBs by FTIR comprises a part of β-sheet structure which corresponds to the IB scaffold and serves as the link to α-helix parts conforming to the native-folded protein inside the aggregates (Roca-Pinilla et al. 2020a; Ami et al. 2006). In addition, FTIR has been used for the in vivo detection of IBs in bacteria, following their rate of formation at different temperatures (Ami et al. 2005). The secondary structure evolves from α-helixes predomination at earlier stages of the aggregation process, corresponding to the native-folded protein, to β-sheet forms showing the growth of the IBs inside the cells.

To determine the specific composition of IBs, techniques such as 2D sodium dodecyl sulfate-polyacrylamide gel electrophoresis (2D SDS-PAGE) have been applied. As previously mentioned, recombinant protein coexists with other cell proteins. This has been studied in detail using 2D SDS-PAGE, alloweing researchers to take a step forward to elucidate the heterogeneity inside these aggregates (Jevsevar et al. 2005; Rinas and Bailey 1992; Rinas et al. 1993; Jürgen et al. 2010). Alternatively, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-ToF MS) can be used to identify cell proteins present in the aggregate (Gardner et al. 2019). Besides, western blot (WB) is widely used to specifically quantify the amount of recombinant protein present in IBs. Protocols for protein quantification in IBs have been well established, and this has allowed not only to determine IB recombinant protein yields but also to determine the soluble/aggregated protein ratio during recombinant protein production processes (Gifre-Renom et al. 2018; García-Fruitós et al. 2007).

However, IBs conformational heterogeneity is not only determined by different types of proteins; many molecules can be embedded inside these aggregates. In this context, recombinant protein production using *E. coli* as the cell factory has a big concern: the presence of LPS in the recombinant products. LPS, even at low amounts, can induce an endotoxic immune response in mammals, which greatly limits their biomedical applicability. Thus, its detection is crucial for knowing possible drawbacks or alterations to the final activity of the proteins that conform the aggregates (Rueda et al. 2014). Due to the possible obstacles that this could cause, the use of LPS-free systems or generally recognized as safe (GRAS) microorganisms, such as lactic acid bacteria, has increased significantly in the last few years (García-Fruitós 2012; Song et al. 2017). In addition to LPS impurities, carbohydrates or lipids could be present inside the IBs, being possible to determine the specific amounts in each case (Roca-Pinilla et al. 2020a).

Beyond the physicochemical, structural, and morphological characterization of IBs, different research groups have been working on the analysis of the activity of recombinant proteins in form of IBs. It has been widely demonstrated that IBs are formed (at least partially) by active proteins in a very important percentage. It is well-known that the composition of the IBs is directly related to the recombinant protein produced and that is why the assays to study the activity are

protein-dependent. Thus, different protocols have been optimized for the determination of IBs activity including enzymatic activity assays (Hrabárová et al. 2015; García-Fruitós et al. 2005a; Tokatlidis et al. 1991; García-Fruitós et al. 2005b; Worrall and Goss 1989; Gifre-Renom et al. 2020c), cytokine activity (Carratalá et al. 2020a), antimicrobial activity (Roca-Pinilla et al. 2020b), and also fluorescence emission (Garcia-Fruitós 2009). The presence of activity in these protein aggregates has opened a large number of possibilities in terms of applicability, as detailed in the following sections.

4 Stability of IBs

IBs are protein aggregates with a notable mechanical stability (Rinas et al. 2017). It has been widely described that their amyloid structure allows them to preserve their integrity and morphology upon mechanical, chemical, and enzymatic cell disruption and upon long-term storage under different conditions. Recent research has correlated these in vitro observations with the in vivo efficiency of this new biomaterial while comparing the effect of other types of nanoparticles. Matrix metalloproteinase-9 (MMP-9) protein, which has an important role in facilitating the migration of immune cells, has been used as a model protein in these studies (Gifre-Renom et al. 2020b). The MMP-9 IBs were compared with their soluble counterpart and MMP-9 encapsulated in polymeric-based micelles (PM) through ionic and covalent binding. The soluble MMP-9 and the MMP-9-ionic PM showed the highest activity values in vitro, whereas IBs showed the lowest activity values. However, the in vitro stability test in 50% bovine serum at room temperature proved that the IBs were the most stable format. Interestingly, the data were well correlated in vivo using an intra-dermal air-pouch model in mice. MMP-9 IBs appeared to be the biomaterial with the highest in vivo activity compared to the soluble MMP-9 form, that was associated with a low and a transitory peak of activity. These results demonstrated that, although the IBs are not always the most active format in vitro, their stability can switch this biomaterial in the most active form once administrated in vivo thanks to their slow-release properties and resilience to protein degradation.

5 Inclusion Bodies as Active Nanoparticles: Applications

Acknowledging IBs as functional protein nanoparticles (Jevsevar et al. 2005; García-Fruitós et al. 2005a; Peternel and Komel 2011) boosted a new perspective in research that aimed to exploit the different properties of these nanoparticles in a wide range of applications. As such, during the last decades, IBs have exhibited an important economic potential in industrial catalysis and an intrinsic biological interest in tissue engineering and in therapeutic research, including regenerative

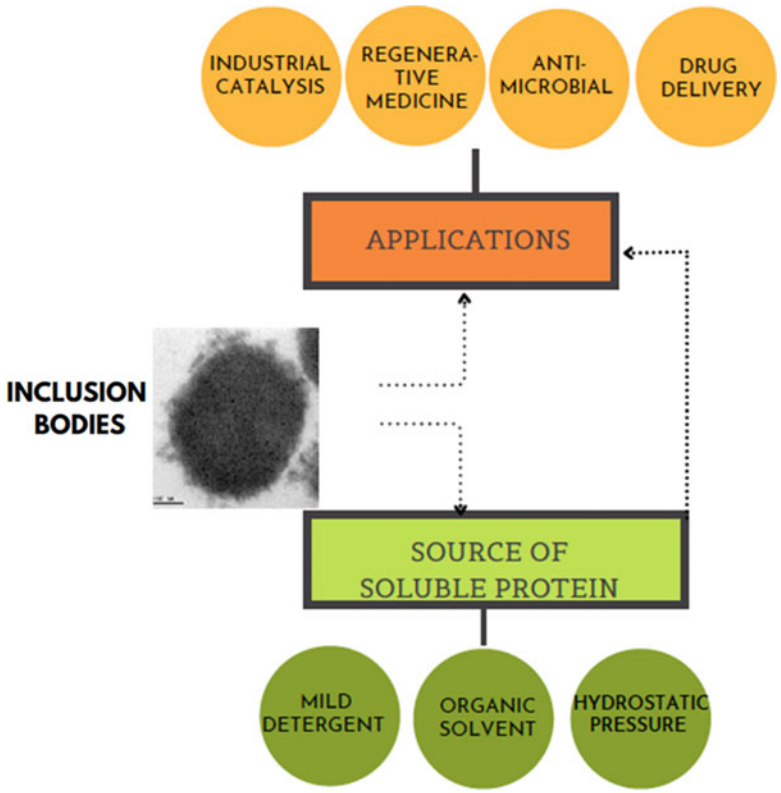


Fig. 3 Schematic representation of IBs applications

medicine, drug delivery in cancer and immunotherapy, and their antimicrobial use as alternatives to classic antibiotics (Fig. 3).

5.1 IBs in Biocatalysis

Pharmaceuticals, food-related products, biofuels, detergents, and other everyday goods, such as paper and textile, are being produced at high rates by engineered biocatalysts. The use of enzymes at industrial scale was enhanced after the DNA technology discovery, which provided the tools to obtain recombinant enzymes on demand. The main limitations were then the lifetime and stability of these enzymes in which, together with the high costs involved in protein purification processes, they made industrial scale-up expensive. The immobilization of the enzymes with either organic or inorganic carriers, or through carrier-free cross-linked enzyme aggregation (further reviewed methods in Wang et al. (2015b)), allowed the recyclability of these catalysts for several enzymatic reactions, importantly reducing the overall

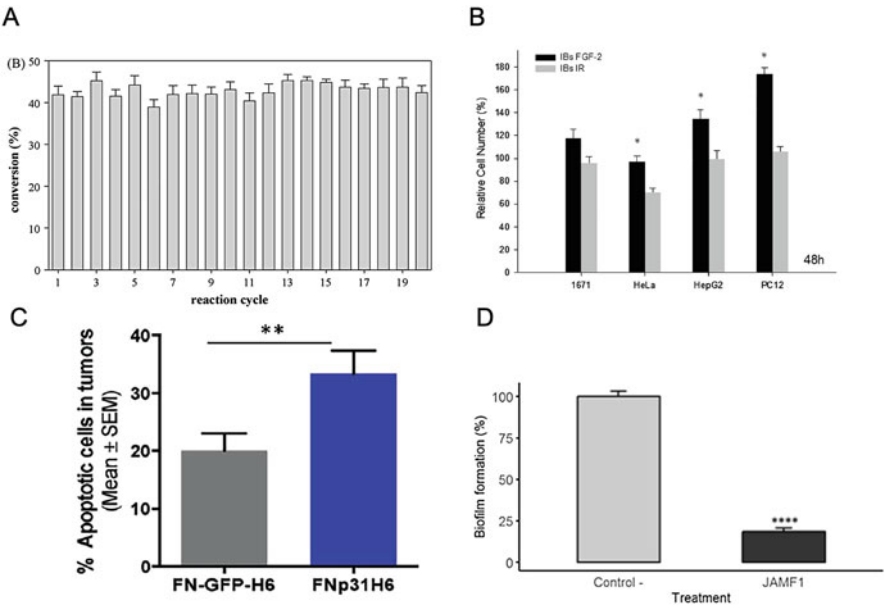


Fig. 4 Examples of IBs activity. (**a**) Yields of substrate to product conversion using cross-linked IBs. Reprinted from Nahalka et al., 2008, copyright (2008) with permission from Elsevier. (**b**) Cell growth of 1671, HeLa, HepG2, and PC12 cell cultures on FGF-2 IBs and nonfunctional IBs (IBs IR). Reprinted from Seras-Franzoso et al., 2014, copyright (2014) with permission from Elsevier. (**c**) Quantification of apoptotic nuclei in treated tumors with (FN-GFP-H6) and FN-p31-H6 IBs. Reprinted from Pesarrodona et al., 2019 (open access article distributed under the terms of the Creative Commons CC BY license). (**d**) Antibiofilm activity of JAMF1 IBs against carbapenem-resistant *Klebsiella pneumoniae* (KPC). Reprinted from Roca-Pinilla et al., 2020 (open access article licensed under a Creative Commons Attribution 4.0 International License)

costs. However, laborious and expensive chromatographic purification of the enzymes and covalent or non-covalent attachments to carriers are inherent in these approaches. In fact, these are still in the focus of biocatalysts optimization.

In this regard, IBs have demonstrated to be a highly convenient type of carrier-free immobilizing method for biocatalysts (Fig. 4a). The fact that these aggregates do not require any chromatographic purification steps makes them already appealing from their production process, which is important for the reduction in industrial scaling-up costs (Kloss et al. 2018). Hence, abundant research has been dedicated to the study of IBs as biocatalysts, as summarized in the book chapter by Hrabárová et al. (2015) and in the mini-review by Jäger et al. (2020). However, the benefits of using IBs go beyond purification-related processes. On the one hand, since IBs are spontaneously formed as aggregates (otherwise promoted by fusion of aggregation-inducing tags (Jäger et al. 2019)), no further immobilization procedures are required afterward, reducing not only any adsorption efficiency issue but also time and enzyme manipulation. Moreover, these new class of biocatalysts are reusable, maintaining their activity after multiple cycles of catalysis (Fig. 4a) (Rinas et al.

2017; Nahálka et al. 2008). On the other hand, IBs can be easily tailored to fit a desired size, resistance to solubilizing agents, porosity, purity, or specific activity by simply modulating the production conditions (e.g., temperature, pH, etc.) (de Marco et al. 2019) or the expression system. The porous nature of IBs, for instance, can benefit its adhesion in case of solid phase catalysis. Moreover, improved IB biocatalysts can be obtained by combining different proteins or by fusion of specific protein domains (Jäger et al. 2019). For example, the fusion of an hydroxynitrile lyase to a coiled-coil domain improved the enzyme resistance to acidic pH (Diener et al. 2016), or two-step cascade reactions could be facilitated by colocalizing different enzymes within the IBs, further improving the stability of the involved enzymes (Jäger et al. 2018).

Thus, so far, many enzymes have proven to form catalytically active IBs exploitable as industrial biocatalysts. Some examples are hydrolases (Park et al. 2012; Tokatlidis et al. 1991; Worrall and Goss 1989; Li et al. 2013; Dong et al. 2014), oxidoreductases (García-Fruitós et al. 2005a), oxidases (Hrabárová et al. 2015), lyases (Kloss et al. 2018), phosphatases (Huang et al. 2013), kinases (Nahálka et al. 2006), aldolases (Nahálka et al. 2008; Sans et al. 2012), phosphorylases (Nahálka 2008), and transferases (Korovashkina et al. 2012; Mestrom et al. 2020). Nevertheless, IBs do not guarantee per se the best (although still high) catalytic performance, as described by a recent work in which IBs and carrier-immobilization methods are compared for trehalose transferase (Mestrom et al. 2020); instead, this might be an enzyme-to-enzyme matter. The benefits of using IBs as biocatalysts are, however, not only drastically reducing the overall procedure costs and providing stability and tailorable properties to the enzyme but also conferring the possibility to design cascade-like bioreactions relatively easily.

5.2 IBs in Therapy/Nanopills

Another remarkable potential of IBs is their use for drug delivery and regenerative medicine approaches in biomedicine and animal science. The fact that the nanoparticle form of proteins provides stability to the embedded proteins, along with their slow-release properties in physiological conditions and the capability of engineering the proteins for targeting purposes, motivated an extensive study of IBs to identify their therapeutic value as nanoscale aggregates with releasable protein (nanopills) (Vázquez et al. 2012). For example, IBs formed by biologically irrelevant proteins (i.e., green fluorescent protein (GFP)) were used to decorate surfaces for in vitro cell culture and demonstrated to increase human mesenchymal stem cells adherence and to promote their differentiation to osteoblasts (Seras-Franzoso et al. 2014). This bottom-up approach was further investigated with IBs composed of extracellularly and/or intracellularly acting biologically relevant proteins (e.g., fibroblast growth factor 2 and human chaperone Hsp70), with results suggesting the cellular internalization of the slowly released proteins from the immobilized IBs (Fig. 4b) (Seras-Franzoso et al. 2014; Seras-Franzoso et al. 2013a; Seras-Franzoso et al. 2013b).

These results are highly valuable for tissue engineering and regenerative medicine applications (Martínez-Miguel et al. 2020). This example has already been materialized in in vitro wound healing studies accomplished by protein hormone-releasing IBs (Stamm et al. 2018). In addition, IBs in suspension can also be internalized by mammalian cells with a high efficiency (Seras-Franzoso et al. 2016; Vázquez et al. 2012), suggesting their potential for drug delivery in vivo and inspiring tumor-targeting research, as will be discussed later on. In the context of the aforementioned applications of IBs, it is relevant to point out the absence of toxicity signs neither when injected (intraperitoneal or intratumorally) nor when administrated orally in mice and zebrafish (Vázquez et al. 2012; Torrealba et al. 2016b). Further, orally administered tumor necrosis factor (TNF)-α IBs protected zebrafish from a lethal infection through their immunostimulant action (Torrealba et al. 2016b). Other proteins forming IBs have been studied also for animal science applications. The low costs of IB production and their versatility make these nanomaterials especially appealing for veterinary applications.

5.3 IBs in Cancer

IBs have several properties that make them also attractive as peptide/protein drug delivery systems (DDS) for cancer (Vázquez et al. 2012). The IBs stability, penetrability, and slow-release properties, therefore, might be useful for the delivery of the proteins to tumoral tissues, if the proteins can target a specific biomarker that is only present or overexpressed in malignant cells. Taking advantage of this, researchers sought to target $CD184^+$ colorectal cancer cells, which overexpress the chemokine receptor CD184 (Unzueta et al. 2018). They designed two modular constructs based on the T22 and R9 proteins using GFP as a fusion partner. These two proteins, T22 and R9, are highly specific ligands of the CD184 receptor and, thus, are able to home in cells that overexpress it. In addition, T22 and R9, when fused to GFP, and in combination with a poly-His tag, can self-assemble in protein nanoparticles (pNPs) that can efficiently direct and increase pNPs cell penetration in tumoral cells. These pNPs can also aggregate forming IBs and provide an immobilized DDS that maintained the sustained release of these tumor-targeting pNPs both in vitro and in vivo. Remarkably, T22-GFP IBs provided high amounts of pNPs for up to 10 days after subcutaneous administration. This work only proved the appealing slow-release properties of IBs for cancer applications, with an efficient accumulation in the desired malignant tissue, yet it did not use any therapeutic proteins.

Moving a step further, Pesarrodona et al. designed two different modular proteins formed by the polypeptides p31 and Omomyc (Pesarrodona et al. 2019). The first protein (p31) is based on a fragment from p130cas which can promote apoptosis (Casanova et al. 2006). The second protein (Omomyc) is a dominant negative mutant of the Myc protein which is involved in the cell cycle and possesses antitumoral properties (Whitfield et al. 2017). These two proteins were each fused to a

tumor-homing peptide (FN) that binds to the CD44 receptor, which is a widely recognized tumoral marker and is related to tumor progression and metastasis. In this study, these fusion antitumoral constructs were produced recombinantly as IBs, which were able to selectively enter and kill $CD44^+$ cells in vitro in cell culture, and in vivo in a mice model of human breast cancer (Fig. 4c). Moreover, an important conclusion of both studies was that, in any case, there was no off-target accumulation and toxicity.

5.4 Antimicrobial IBs

Antibiotic-resistant bacteria (ARB) are one of the biggest threats to global health today. Many infectious diseases that could be successfully treated with antibiotic drugs have acquired resistance to most or nearly all of the available compounds, including a growing list of infections, such as pneumonia, tuberculosis, gonorrhea, and foodborne diseases (Laxminarayan et al. 2013). These infections are becoming harder, and sometimes impossible, to treat with the current generation of medicines. To make matter worse, no new class of antibiotics has been discovered since the 1980s, and few antibiotics are being developed to face the challenge of ARBs. So, there is an urgent need for the development of new antimicrobials to enable the treatment for those infections effectively.

Among the potential alternatives to conventional antibiotics, antimicrobial IBs might be a promising one, as they are highly stable and have slow-release properties likely able to maintain antimicrobial activity. Constant administration of antimicrobial compounds that have a short half-life is required, or otherwise, the use of concentrations that are under the minimum inhibitory concentration (MIC) will probably increase antimicrobial resistances (AMRs) (Gao et al. 2011). A slow-release profile seems to be vital to maintain constant antimicrobial levels for long periods of time, to obtain the optimal therapeutic benefits while minimizing AMRs. Another benefit might be that one long-lasting administration is more advantageous than multiple short-lived ones. Further, IBs offer all these properties without the need to manipulate them further (i.e., encapsulation processes or protein embedding in a matrix), because they are produced in a one-step process.

A study published in 2020 found out that the antimicrobial peptide (AMP) GWH1 had a therapeutic effect in a mouse mastitis model (Carratalá et al. 2020b). The GWH1 peptide was fused to the N-terminus of either of GFP or IFN-γ in a form of IBs and resulted in diminished bacterial loads by five- and sixfold, respectively. The study also found that the IB form per se can be antimicrobial, as GFP or IFN-γ IBs showed no antimicrobial activity and, in some cases, achieve a similar reduction in bacterial loads when compared to GWH1 fusion constructs (Carratalá et al. 2020b). Another study found that a multidomain AMP named JAMF1 displayed a clear antimicrobial effect against several ARB, such as *Klebsiella pneumoniae*, *E. coli*, *Enterococcus spp*., and *E. faecalis*, in both planktonic and biofilm state (Fig. 4d) (Roca-Pinilla et al. 2020b). The JAMF1 construct was based on the

combination of human α-defensin-5 (HD5), human XII-A secreted phospholipase A2 ($sPLA_2$), and a gelsolin-based bacterial-binding domain. Two aggregation-seeding domains based on leucine zippers were also added to promote IB formation during the recombinant production of JAMF1.

5.5 *IBs a Source of Soluble Protein*

Inclusion bodies have also been used for years as a source to obtain soluble proteins when these cannot be obtained directly from the cell cytoplasm or from the growth media if secreted. Although some heterologous proteins are produced mainly in the soluble form, most of them are in the aggregated form. In some cases, the aggregation levels are so high that the only option to obtain the soluble form is through the solubilization of IBs. One example of this is matrix metalloproteinase 9 (MMP-9); when using *L. lactis* as the cell factory, MMP-9 is only produced in an aggregated form (Gifre-Renom et al. 2018). Other proteins of interest are produced in a soluble form, but they are co-purified with host cell proteins. This occurs, for example, with mammary serum amyloid protein A3 (M-SAA3), and their solubilization from IBs is the only protocol that allows to have good purity levels (Gifre-Renom et al. 2018). Thus, for both aggregation-prone and difficult-to-produce/purify proteins, the solubilization of IBs is the only strategy for their purification.

The strategy that has been traditionally used to recover soluble proteins from IBs has been based on the following sequential steps: IB isolation and purification from the recombinant culture, IB solubilization with harsh detergents and chaotropic agents (e.g., urea or guanidine hydrochloride) at high concentration (i.e., 6–8 M), refolding of the solubilized protein to reach their native conformation, and purification by chromatography methods (Singh et al. 2015a). However, the use of these solubilization agents lead to a protein denaturation, and refolding steps are necessary to recover the original structure and function of the target proteins. Solubilization and refolding are critical steps limiting the yield of recovered protein, and have been subjected to an intense and expanding research (Burgess 2009; Singh et al. 2015b). Usually, proteins can be refolded upon removal of denaturing agents jointly providing favorable conditions to reach their native state (Singhvi et al. 2020). The underlying principle in the refolding process is that the proteins are able to switch back into their native conformation from a denatured condition. However, during this process, non-desirable aggregated intermediates can be formed. In the refolding process, quality and quantity of the folded protein rely on the buffer, protein concentration, and the method used (Singhvi et al. 2020). Alternative and simpler strategies considering the biological nature of bacterial IBs have been developed over the last few years. The new description of IBs as protein aggregates containing properly folded and biologically active recombinant proteins has permitted to advance significantly in the development of new protocols without the need of using denaturing agents to solubilize proteins from IBs (Ferrer-Miralles et al. 2018). For that, mild solubilization agents, such as *n*-lauroylsarcosine, dimethyl

sulfoxide (DMSO), and organic solvents such as *n*-propanol and isopropanol are used, which are strong enough to solubilize IBs without disturbing the native structure of the protein (Singhvi et al. 2020; Ferrer-Miralles et al. 2018; Peternel et al. 2008; Sarker et al. 2019). The use of *n*-lauroylsarcosine to recover soluble protein from IBs has allowed the development of a non-denaturing protocol for IBs that are produced from *E. coli* (Roca-Pinilla et al. 2020b; Peternel et al. 2008; Park et al. 2018; Francis et al. 2012) and *L. lactis* (Gifre-Renom et al. 2018). Target proteins are gradually solubilized in a single step, keeping their native structure. In addition, mild detergents enhance membrane protein recovery due to stabilization of their hydrophobic groups, playing both roles of solubilization and retaining its folded and active state (Francis et al. 2012). Once solubilized, the protein can be further purified by chromatography. An alternative to mild detergents are organic solvents, which have been proven to be plausible alternative to obtain high-quality protein from IBs. The use of *n*-propanol (Singh et al. 2012), for example, has been reported to be an efficient bioactive protein solubilization method. In addition, trifluoroethanol (TFE) has also demonstrated a remarkable potential to achieve functional protein from IBs (Upadhyay et al. 2016). Nevertheless, organic solvents can generate non-desirable chemical modifications, and in many cases they need to be used in combination with low concentrations of chaotropic agents (Singhvi et al. 2020).

Another strategy is based in the use of high hydrostatic pressure to disaggregate the insoluble aggregates and refold them back in their native structure (Ferrer-Miralles et al. 2018; St John et al. 2001). In addition, the use of heat (Cai et al. 2020), pH oscillations (Panda 2003), and freeze-thaw cycles (Qi et al. 2015) with the combination of low amounts of denaturant agents has been postulated as efficient methods to solubilize and recover bioactive protein from IBs (Ferrer-Miralles et al. 2018).

6 Conclusions

Inclusion bodies arise as a new protein-based biomaterial that is spontaneously formed in bacterial hosts during recombinant protein production. Several characterization procedures based on electron microscopy, molecular techniques, and physicochemical analyses can be used to finely determine their structure and composition. IBs have been demonstrated to be a promising biomaterial for biocatalysis, tissue regeneration, drug delivery, and antimicrobial therapy applications. It is likely that the IBs' stability along with their slow-release properties is the basis for its potential and successful use. Although the soluble version of a recombinant protein is needed in some cases, the target protein can be easily obtained from IBs under mild solubilization protocols and free of host cell proteins and other impurities. In short, IBs are valuable in the world of recombinant proteins and new protein-based biomaterials, which motivates to deeply study new microorganisms as IBs producers to tune those features that can be further optimized.

Acknowledgments The authors acknowledge the financial support granted to A.A and E.G.F from Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Spain (RTA2015-00064-C02-01) and Ministerio de Ciencia e Innovación (MICINN), Spain (grant PID2019-107298RB-C21). This work was also supported by Marató de TV3 foundation (201812-30-31-32-33). We are also indebted to CERCA Programme (Generalitat de Catalunya) and European Social Fund for supporting our research. R.B.F. and A.L.C received a predoctoral fellowship from INIA (FPI-INIA, MINECO) and Generalitat de Catalunya (FI-AGAUR), respectively.

References

Ami D, Natalello A, Gatti-Lafranconi P, Lotti M, Doglia SM (2005) Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy. FEBS Lett 579(16):3433–3436

Ami D, Natalello A, Taylor G, Tonon G, Maria DS (2006) Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy. Biochim Biophys Acta 1764(4): 793–799

Arié JP, Miot M, Sassoon N, Betton JM (2006) Formation of active inclusion bodies in the periplasm of *Escherichia coli*. Mol Microbiol 62(2):427–437

Burgess RR (2009) Refolding solubilized inclusion body proteins. Methods Enzymol 463:259–282

Cai H, Chen G, Yu H, Tang Y, Xiong S, Qi X (2020) One-step heating strategy for efficient solubilization of recombinant spider silk protein from inclusion bodies. BMC Biotechnol 20(1): 37

Cano-Garrido O, Rodríguez-Carmona E, Díez-Gil C, Vázquez E, Elizondo E, Cubarsi R et al (2013) Supramolecular organization of protein-releasing functional amyloids solved in bacterial inclusion bodies. Acta Biomater 9(4):6134–6142

Cano-Garrido O, Sanchez-Chardi A, Pares S, Giro I, Tatkiewicz WI, Ferrer-Miralles N et al (2016) Functional protein-based nanomaterial produced in microorganisms recognized as safe: a new platform for biotechnology. Acta Biomater 43:230–239

Carratalá JV, Cano-Garrido O, Sánchez J, Membrado C, Pérez E, Conchillo-Solé O et al (2020a) Aggregation-prone peptides modulate activity of bovine interferon gamma released from naturally occurring protein nanoparticles. New Biotechnol 57:11–19

Carratalá JV, Brouillette E, Serna N, Sánchez-Chardi A, Sánchez JM, Villaverde A et al (2020b) In vivo bactericidal efficacy of GWH1 antimicrobial peptide displayed on protein nanoparticles, a potential alternative to antibiotics. Pharmaceutics. 12(12):1217

Carratalá JV, Gifre-Renom L, Roca-Pinilla R, Villaverde A, Arís A, Garcia-Fruitós E et al (2021a) Selecting subpopulations of high-quality protein conformers among conformational mixtures of recombinant bovine MMP-9 solubilized from inclusion bodies. Int J Mol Sci 22(6):3020

Carratalá JV, Cisneros A, Hellman E, Villaverde A, Ferrer-Miralles N (2021b) Title: insoluble proteins catch heterologous soluble proteins into inclusion bodies by intermolecular interaction of aggregating peptides. Microb Cell Factories 20(1):30

Casanova I, Parreño M, Farré L, Guerrero S, Céspedes MV, Pavon MA et al (2006) Celecoxib induces anoikis in human colon carcinoma cells associated with the deregulation of focal adhesions and nuclear translocation of p130Cas. Int J Cancer 118(10):2381–2389

Castillo V, Graña-Montes R, Sabate R, Ventura S (2011) Prediction of the aggregation propensity of proteins from the primary sequence: aggregation properties of proteomes. Biotechnol J 6(6): 674–685

Choi SL, Lee SJ, Yeom SJ, Kim HJ, Rhee YH, Jung HC et al (2014) Controlled localization of functionally active proteins to inclusion bodies using leucine zippers. PLoS One 9(6):e97093

de Groot NS, Sabate R, Ventura S (2009) Amyloids in bacterial inclusion bodies. Trends Biochem Sci 34(8):408–416

de Marco A, Ferrer-Miralles N, Garcia-Fruitós E, Mitraki A, Peternel S, Rinas U et al (2019) Bacterial inclusion bodies are industrially exploitable amyloids. FEMS Microbiol Rev 43(1): 53–72

Diener M, Kopka B, Pohl M, Jaeger K-E, Krauss U (2016) Fusion of a coiled-coil domain facilitates the high-level production of catalytically active enzyme inclusion bodies. ChemCatChem 8(1): 142–152

Díez-Gil C, Krabbenborg S, García-Fruitós E, Vazquez E, Rodríguez-Carmona E, Ratera I et al (2010) The nanoscale properties of bacterial inclusion bodies and their effect on mammalian cell proliferation. Biomaterials 31(22):5805–5812

Dong Q, Yan X, Zheng M, Yang Z (2014) Characterization of an extremely thermostable but cold-adaptive β-galactosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* for use as a recombinant aggregation for batch lactose degradation at high temperature. J Biosci Bioeng 117(6):706–710

Ferrer-Miralles N, Saccardo P, Garcia-Fruitós E. Chapter 6. Protein purification from protein aggregates. In: Publishers NS, editor. Handbook on Protein Purification: Industry Challenges and Technological Developments. 2018

Francis VG, Majeed MA, Gummadi SN (2012) Recovery of functionally active recombinant human phospholipid scramblase 1 from inclusion bodies using N-lauroyl sarcosine. J Ind Microbiol Biotechnol 39(7):1041–1048

Gao P, Nie X, Zou M, Shi Y, Cheng G (2011) Recent advances in materials for extended-release antibiotic delivery system. J Antibiot 64(9):625–634

García-Fruitós E (2012) Lactic acid bacteria: a promising alternative for recombinant protein production. Microb Cell Factories 11:157

García-Fruitós E, González-Montalbán N, Morell M, Vera A, Ferraz RM, Arís A et al (2005a) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. Microb Cell Factories 4:27

García-Fruitós E, Carrió MM, Arís A, Villaverde A (2005b) Folding of a misfolding-prone beta-galactosidase in absence of DnaK. Biotechnol Bioeng 90(7):869–875

García-Fruitós E, Martínez-Alonso M, Gonzàlez-Montalbán N, Valli M, Mattanovich D, Villaverde A (2007) Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. J Mol Biol 374(1):195–205

García-Fruitós E, Sabate R, de Groot NS, Villaverde A, Ventura S (2011) Biological role of bacterial inclusion bodies: a model for amyloid aggregation. FEBS J 278(14):2419–2427

García-Fruitós E, Seras-Franzoso J, Vazquez E, Villaverde A (2010) Tunable geometry of bacterial inclusion bodies as substrate materials for tissue engineering. Nanotechnology 21(20):205101

García-Fruitós E, Vázquez E, Díez-Gil C, Corchero JL, Seras-Franzoso J, Ratera I et al (2012) Bacterial inclusion bodies: making gold from waste. Trends Biotechnol 30(2):65–70

Garcia-Fruitós E-C (2009) EscarlataDíez-Gil, CésarFerraz, Rosa MVázquez, EstherCorchero, José Luis Cano-Sarabia, Mary Ratera, Imma, Ventosa N, Veciana J. Villaverde A Surface cell growth engineering assisted by a novel bacterial nanomaterial Adv Mater 21:1–5

Gardner QA, Hassan N, Hafeez S, Arif M, Akhtar M (2019) Exploring the nature of inclusion bodies by MALDI mass spectrometry using recombinant proinsulin as a model protein. Int J Biol Macromol 139:647–653

Gifre-Renom L, Cano-Garrido O, Fàbregas F, Roca-Pinilla R, Seras-Franzoso J, Ferrer-Miralles N et al (2018) A new approach to obtain pure and active proteins from Lactococcus lactis protein aggregates. Sci Rep 8(1):13917

Gifre-Renom L, Ugarte-Berzal E, Martens E, Boon L, Cano-Garrido O, Martínez-Núñez E et al (2020a) Recombinant protein-based nanoparticles: elucidating their inflammatory effects in vivo and their potential as a new therapeutic format. Pharmaceutics. 12(5):450

Gifre-Renom L, Seras-Franzoso J, Rafael D, Andrade F, Cano-Garrido O, Martinez-Trucharte F et al (2020b) The biological potential hidden in inclusion bodies. Pharmaceutics 12(2):157

Gifre-Renom L, Carratalá JV, Parés S, Sánchez-García L, Ferrer-Miralles N, Villaverde A et al (2020c) Potential of MMP-9 based nanoparticles at optimizing the cow dry period: pulling apart the effects of MMP-9 and nanoparticles. Sci Rep 10(1):11299

Gil-Garcia M, Navarro S, Ventura S (2020) Coiled-coil inspired functional inclusion bodies. Microb Cell Factories 19(1):117

Hrabárová E, Achbergerová L, Nahálka J (2015) Insoluble protein applications: the use of bacterial inclusion bodies as biocatalysts. Methods Mol Biol 1258:411–422

Huang Z, Zhang C, Chen S, Ye F, Xing XH (2013) Active inclusion bodies of acid phosphatase PhoC: aggregation induced by GFP fusion and activities modulated by linker flexibility. Microb Cell Factories 12:25

Jäger VD, Kloss R, Grünberger A, Seide S, Hahn D, Karmainski T et al (2019) Tailoring the properties of (catalytically)-active inclusion bodies. Microb Cell Factories 18(1):33

Jäger VD, Lamm R, Kloß R, Kaganovitch E, Grünberger A, Pohl M et al (2018) A synthetic reaction cascade implemented by colocalization of two proteins within catalytically active inclusion bodies. ACS Synth Biol 7(9):2282–2295

Jäger VD, Lamm R, Küsters K, Ölçücü G, Oldiges M, Jaeger KE et al (2020) Catalytically-active inclusion bodies for biotechnology-general concepts, optimization, and application. Appl Microbiol Biotechnol 104(17):7313–7329

Jevsevar S, Gaberc-Porekar V, Fonda I, Podobnik B, Grdadolnik J, Menart V (2005) Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. Biotechnol Prog 21(2):632–639

Jiang L, Xiao W, Zhou X, Wang W, Fan J (2019) Comparative study of the insoluble and soluble Ulp1 protease constructs as carrier free and dependent protein immobilizates. J Biosci Bioeng 127(1):23–29

Jürgen B, Breitenstein A, Urlacher V, Büttner K, Lin H, Hecker M et al (2010) Quality control of inclusion bodies in *Escherichia coli*. Microb Cell Factories 9:41

Kloss R, Limberg MH, Mackfeld U, Hahn D, Grünberger A, Jäger VD et al (2018) Catalytically active inclusion bodies of L-lysine decarboxylase from *E. coli* for 1,5-diaminopentane production. Sci Rep 8(1):5856

Korovashkina AS, Rymko AN, Kvach SV, Zinchenko A (2012) Enzymatic synthesis of c-di-GMP using inclusion bodies of *Thermotoga maritima* full-length diguanylate cyclase. J Biotechnol 164(2):276–280

Küsters K, Pohl M, Krauss U, Ölçücü G, Albert S, Jaeger KE et al (2021) Construction and comprehensive characterization of an EcLDCc-CatIB set-varying linkers and aggregation inducing tags. Microb Cell Factories 20(1):49

Lamm R, Jäger VD, Heyman B, Berg C, Cürten C, Krauss U et al (2020) Detailed small-scale characterization and scale-up of active YFP inclusion body production with *Escherichia coli* induced by a tetrameric coiled coil domain. J Biosci Bioeng 129(6):730–740

Laxminarayan R, Duse A, Wattal C, Zaidi AK, Wertheim HF, Sumpradit N et al (2013) Antibiotic resistance-the need for global solutions. Lancet Infect Dis 13(12):1057–1098

Li S, Lin K, Pang H, Wu Y, Xu J (2013) Production, characterization, and application of an organic solvent-tolerant lipase present in active inclusion bodies. Appl Biochem Biotechnol 169(2): 612–623

Liovic M, Ozir M, Zavec AB, Peternel S, Komel R, Zupancic T (2012) Inclusion bodies as potential vehicles for recombinant protein delivery into epithelial cells. Microb Cell Factories 11:67

Martínez-Miguel M, Kyvik AR, Ernst LM, Martínez-Moreno A, Cano-Garrido O, Garcia-Fruitós E et al (2020) Stable anchoring of bacteria-based protein nanoparticles for surface enhanced cell guidance. J Mater Chem B 8(23):5080–5088

Mestrom L, Marsden SR, McMillan DGG, Schoevaart R, Hagedoorn P-L, Hanefeld U (2020) Comparison of enzymes immobilised on immobeads and inclusion bodies: a case study of a trehalose transferase. ChemCatChem 12(12):3249–3256

Morell M, Bravo R, Espargaró A, Sisquella X, Avilés FX, Fernàndez-Busquets X et al (2008) Inclusion bodies: specificity in their aggregation process and amyloid-like structure. Biochim Biophys Acta 1783(10):1815–1825

Nahálka J (2008) Physiological aggregation of maltodextrin phosphorylase from *Pyrococcus furiosus* and its application in a process of batch starch degradation to alpha-D-glucose-1-phosphate. J Ind Microbiol Biotechnol 35(4):219–223

Nahálka J, Gemeiner P, Bucko M, Wang PG (2006) Bioenergy beads: a tool for regeneration of ATP/NTP in biocatalytic synthesis. Artif Cells Blood Substit Immobil Biotechnol 34(5): 515–521

Nahalka J, Nidetzky B (2007) Fusion to a pull-down domain: a novel approach of producing *Trigonopsis variabilis* D-amino acid oxidase as insoluble enzyme aggregates. Biotechnol Bioeng 97(3):454–461

Nahálka J, Vikartovská A, Hrabárová E (2008) A crosslinked inclusion body process for sialic acid synthesis. J Biotechnol 134(1–2):146–153

Natalello A, Doglia SM (2015) Insoluble protein assemblies characterized by Fourier transform infrared spectroscopy. Methods Mol Biol 1258:347–369

Panda AK (2003) Bioprocessing of therapeutic proteins from the inclusion bodies of Escherichia coli. Adv Biochem Eng Biotechnol 85:43–93

Park AR, Jang SW, Kim JS, Park YG, Koo BS, Lee HC (2018) Efficient recovery of recombinant CRM197 expressed as inclusion bodies in *E. coli*. PLoS One 13(7):e0201060

Park SY, Park SH, Choi SK (2012) Active inclusion body formation using *Paenibacillus polymyxa* PoxB as a fusion partner in *Escherichia coli*. Anal Biochem 426(1):63–65

Pesarrodona M, Jauset T, Díaz-Riascos ZV, Sánchez-Chardi A, Beaulieu ME, Seras-Franzoso J et al (2019) Targeting antitumoral proteins to breast cancer by local administration of functional inclusion bodies. Adv Sci 6(18):1900849

Peternel S, Grdadolnik J, Gaberc-Porekar V, Komel R (2008) Engineering inclusion bodies for non denaturing extraction of functional proteins. Microb Cell Factories 7:34

Peternel S, Komel R (2011) Active protein aggregates produced in *Escherichia coli*. Int J Mol Sci 12(11):8275–8287

Qi X, Sun Y, Xiong S (2015) A single freeze-thawing cycle for highly efficient solubilization of inclusion body proteins and its refolding into bioactive form. Microb Cell Factories 14:24

Rinas U, Bailey JE (1992) Protein compositional analysis of inclusion bodies produced in recombinant *Escherichia coli*. Appl Microbiol Biotechnol 37(5):609–614

Rinas U, Boone TC, Bailey JE (1993) Characterization of inclusion bodies in recombinant *Escherichia coli* producing high levels of porcine somatotropin. J Biotechnol 28(2–3):313–320

Rinas U, Garcia-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A (2017) Bacterial inclusion bodies: discovering their better half. Trends Biochem Sci 42(9):726–737

Roca-Pinilla R, Fortuna S, Natalello A, Sánchez-Chardi A, Ami D, Arís A et al (2020a) Exploring the use of leucine zippers for the generation of a new class of inclusion bodies for pharma and biotechnological applications. Microb Cell Factories 19(1):175

Roca-Pinilla R, López-Cano A, Saubi C, Garcia-Fruitós E, Arís A (2020b) A new generation of recombinant polypeptides combines multiple protein domains for effective antimicrobial activity. Microb Cell Factories 19(1):122

Rueda F, Cano-Garrido O, Mamat U, Wilke K, Seras-Franzoso J, García-Fruitós E et al (2014) Production of functional inclusion bodies in endotoxin-free *Escherichia coli*. Appl Microbiol Biotechnol 98(22):9229–9238

Rueda F, Gasser B, Sanchez-Chardi A, Roldan M, Villegas S, Puxbaum V et al (2016) Functional inclusion bodies produced in the yeast *Pichia pastoris*. Microb Cell Factories 15:1–12

Sanagavarapu K, Nüske E, Nasir I, Meisl G, Immink JN, Sormanni P et al (2019) A method of predicting the in vitro fibril formation propensity of Aβ40 mutants based on their inclusion body levels in *E. coli*. Sci Rep 9(1):3680

Sans C, García-Fruitós E, Ferraz RM, González-Montalbán N, Rinas U, López-Santín J et al (2012) Inclusion bodies of fuculose-1-phosphate aldolase as stable and reusable biocatalysts. Biotechnol Prog 28(2):421–427

Sarker A, Rathore AS, Gupta RD (2019) Evaluation of scFv protein recovery from E. coli by in vitro refolding and mild solubilization process. Microb Cell Factories 18(1):5

Seras-Franzoso J, Peebo K, Luis Corchero J, Tsimbouri PM, Unzueta U, Rinas U et al (2013a) A nanostructured bacterial bioscaffold for the sustained bottom-up delivery of protein drugs. Nanomedicine 8(10):1587–1599

Seras-Franzoso J, Steurer C, Roldán M, Vendrell M, Vidaurre-Agut C, Tarruella A et al (2013b) Functionalization of 3D scaffolds with protein-releasing biomaterials for intracellular delivery. J Control Release 171(1):63–72

Seras-Franzoso J, Peternel S, Cano-Garrido O, Villaverde A, García-Fruitós E (2015) Bacterial inclusion body purification. Methods Mol Biol 1258:293–305

Seras-Franzoso J, Sanchez-Chardi A, Garcia-Fruitos E, Vazquez E, Villaverde A (2016) Cellular uptake and intracellular fate of protein releasing bacterial amyloids in mammalian cells. Soft Matter 12(14):3451–3460

Seras-Franzoso J, Tsimbouri PM, Burgess KV, Unzueta U, Garcia-Fruitos E, Vazquez E et al (2014) Topographically targeted osteogenesis of mesenchymal stem cells stimulated by inclusion bodies attached to polycaprolactone surfaces. Nanomedicine 9(2):207–220

Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK (2015a) Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. Microb Cell Factories 14:41

Singh A, Upadhyay V, Panda AK (2015b) Solubilization and refolding of inclusion body proteins. Methods Mol Biol 1258:283–291

Singh SM, Sharma A, Upadhyay AK, Singh A, Garg LC, Panda AK (2012) Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. Protein Expr Purif 81(1):75–82

Singhvi P, Saneja A, Srichandan S, Panda AK (2020) Bacterial inclusion bodies: a treasure trove of bioactive proteins. Trends Biotechnol 38(5):474–486

Song AA, In LLA, Lim SHE, Rahim RA (2017) A review on *Lactococcus lactis*: from food to factory. Microb Cell Factories 16(1):55

St John RJ, Carpenter JF, Balny C, Randolph TW (2001) High pressure refolding of recombinant human growth hormone from insoluble aggregates. Structural transformations, kinetic barriers, and energetics. J Biol Chem 276(50):46856–46863

Stamm A, Strauß S, Vogt P, Scheper T, Pepelanova I (2018) Positive in vitro wound healing effects of functional inclusion bodies of a lipoxygenase from the Mexican axolotl. Microb Cell Factories 17(1):57

Tokatlidis K, Dhurjati P, Millet J, Béguin P, Aubert JP (1991) High activity of inclusion bodies formed in *Escherichia coli* overproducing *clostridium thermocellum* endoglucanase D. FEBS Lett 282(1):205–208

Torrealba D, Seras-Franzoso J, Mamat U, Wilke K, Villaverde A, Roher N et al (2016a) Complex particulate biomaterials as immunostimulant-delivery platforms. PLoS One 11(10):e0164073

Torrealba D, Parra D, Seras-Franzoso J, Vallejos-Vidal E, Yero D, Gibert I et al (2016b) Nanostructured recombinant cytokines: a highly stable alternative to short-lived prophylactics. Biomaterials 107:102–114

Unzueta U, Cespedes MV, Sala R, Alamo P, Sánchez-Chardi A, Pesarrodona M et al (2018) Release of targeted protein nanoparticles from functional bacterial amyloids: a death star-like approach. J Control Release 279:29–39

Unzueta U, Roldán M, Pesarrodona M, Benitez R, Sánchez-Chardi A, Conchillo-Solé O et al (2020) Self-assembling as regular nanoparticles dramatically minimizes photobleaching of tumour-targeted GFP. Acta Biomater 103:272–280

Upadhyay V, Singh A, Jha D, Panda AK (2016) Recovery of bioactive protein from bacterial inclusion bodies using trifluoroethanol as solubilization agent. Microb Cell Factories 15:100

Valax P, Georgiou G (1993) Molecular characterization of beta-lactamase inclusion bodies produced in *Escherichia coli*. 1. Composition Biotechnol Prog 9(5):539–547

Vázquez E, Corchero JL, Burgueño JF, Seras-Franzoso J, Kosoy A, Bosser R et al (2012) Functional inclusion bodies produced in bacteria as naturally occurring nanopills for advanced cell therapies. Adv Mater 24(13):1742–1747

Vera A, González-Montalbán N, Arís A, Villaverde A (2007) The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol Bioeng 96(6):1101–1106

Villaverde A, Corchero JL, Seras-Franzoso J, Garcia-Fruitos E (2015) Functional protein aggregates: just the tip of the iceberg. Nanomedicine 10(18):2881–2891

Wang L (2009) Towards revealing the structure of bacterial inclusion bodies. Prion 3(3):139–145

Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R (2008) Bacterial inclusion bodies contain amyloid-like structure. PLoS Biol 6(8):e195

Wang X, Zhou B, Hu W, Zhao Q, Lin Z (2015a) Formation of active inclusion bodies induced by hydrophobic self-assembling peptide GFIL8. Microb Cell Factories 14:88

Wang M, Qi W, Su R, He Z (2015b) Advances in carrier-bound and carrier-free immobilized nanobiocatalysts. Chem Eng Sci 135:21–32

Whitfield JR, Beaulieu ME, Soucek L (2017) Strategies to inhibit myc and their clinical applicability. Front Cell Dev Biol 5:10

Worrall DM, Goss NH (1989) The formation of biologically active beta-galactosidase inclusion bodies in *Escherichia coli*. Aust J Biotechnol 3(1):28–32

Wu W, Xing L, Zhou B, Lin Z (2011) Active protein aggregates induced by terminally attached self-assembling peptide ELK16 in *Escherichia coli*. Microb Cell Factories 10:9

Zhou B, Xing L, Wu W, Zhang XE, Lin Z (2012) Small surfactant-like peptides can drive soluble proteins into active aggregates. Microb Cell Factories 11:10

Encapsulin Nanocompartments for Biomanufacturing Applications

Taylor N. Szyszka, Lachlan S. R. Adamson, and Yu Heng Lau

Contents

Abstract Enzyme scaffolding is an emerging technique for enhancing yield and efficiency of biomanufacturing processes for generating high-value products. Of the many scaffolds available for in vivo and cell-free applications, encapsulin protein nanocompartments have garnered recent attention due to their desirable properties as a scaffold, such as robust self-assembly, high thermal stability, and the inherent ability to package and display enzymatic cargo. In this chapter, we discuss basic and advanced methods for encapsulin engineering, as well as the many encapsulin-based

T. N. Szyszka · L. S. R. Adamson · Y. H. Lau (✉)
School of Chemistry, The University of Sydney, Camperdown, NSW, Australia

The University of Sydney Nano Institute, The University of Sydney, Camperdown, NSW, Australia
e-mail: taylor.szyszka@sydney.edu.au; yuheng.lau@sydney.edu.au

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*, Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_12

biomanufacturing systems that have been built. We subsequently discuss how these advances can be applied to increasingly complex encapsulin systems and look to the future of biomanufacturing in tunable protein nanocompartments.

1 Introduction to Encapsulins

Encapsulins are prokaryotic self-assembling nanocompartments that typically package other proteins within their interior. In 1994, the first encapsulin was identified in *Brevibacterium linens*, though it was originally called a high-molecular-weight aggregate, annotated as a bacteriocin (Valdés-Stauber and Scherer 1994). Concurrently, similar aggregates were isolated from other prokaryotic organisms such as *Mycobacterium leprae* and *Streptomyces* spp. (Triccas et al. 1996; Winter et al. 1995; Kawamoto et al. 2001; Kwak et al. 2001; Saito et al. 2003). The first encapsulin protein structures were published in the 2000s, from *Thermotoga maritima* and *Pyrococcus furiosus* (Akita et al. 2007; Sutter et al. 2008). These structures revealed that these encapsulin proteins formed large assemblies with other cargo proteins sequestered inside. Following these pioneering studies, the field of encapsulin research and engineering has progressed rapidly, leading to an increasing number of advanced applications in biomanufacturing.

1.1 Encapsulin Structure

Encapsulins are hollow icosahedral protein nanocompartments self-assembled from 60, 180, or 240 copies of a single protein monomer (Fig. 1). The monomers share a striking similarity to viral capsids, adopting a fold most reminiscent of the *Escherichia coli* (*E. coli*) phage HK97 capsid protein (Akita et al. 2007; Sutter et al. 2008; McHugh et al. 2014; Giessen et al. 2019). The HK97-like encapsulin fold consists of three conserved structural motifs: a peripheral P domain, an axial A-domain, and an elongated E loop (Akita et al. 2007; Sutter et al. 2008; McHugh et al. 2014; Giessen et al. 2019; Wikoff et al. 2000). Assembled encapsulins range between 24 and 42 nm in external diameter, with the composition of the assembled state classified according to their triangulation or T number (Caspar and Klug 1962), where $T = 1$ corresponds to a small 60-mer, $T = 3$ a larger 180-mer, and $T = 4$ the largest currently known 240-mer. Although the details of the encapsulin assembly process have not been fully elucidated, recent work suggests that the E loop may determine the T number, as a rotation in the E loop of *T. maritima* encapsulin ($T = 1$) is not present in its $T = 3$ counterparts (Akita et al. 2007; Sutter et al. 2008; McHugh et al. 2014; Nichols et al. 2017).

Encapsulins form robust and stable assemblies that can generally tolerate pH changes and resist digestion by non-specific proteases (Cassidy-Amstutz et al. 2016; Tamura et al. 2015). Some encapsulins from thermophilic organisms are highly

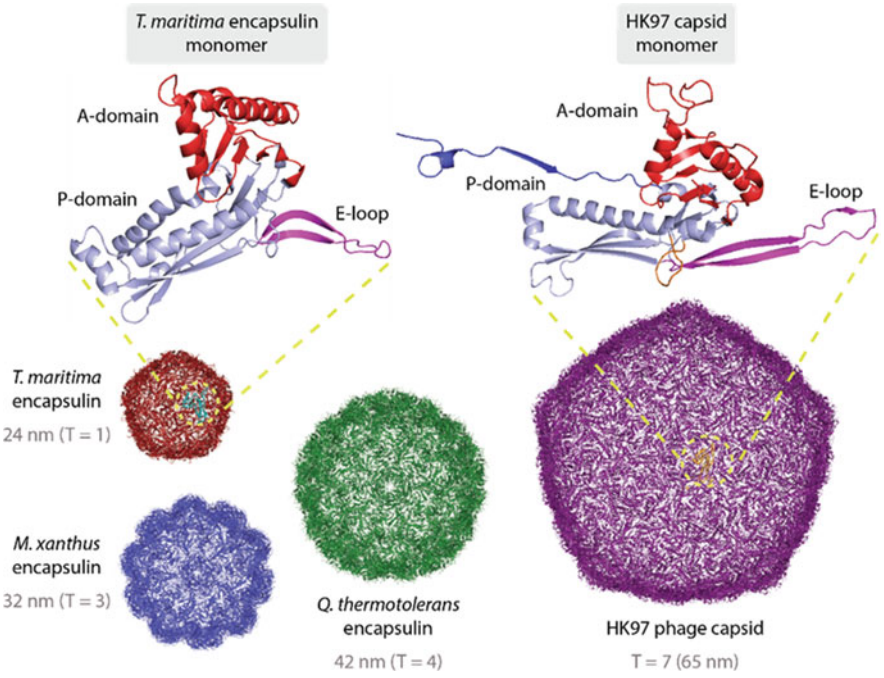


Fig. 1 Encapsulins are self-assembled protein nanocompartments composed of multiple copies of a single protein monomer. Representative Family 1 encapsulins of each triangulation number are shown: $T = 1$ from *Thermotoga maritima* (PDB: 3DKT (Sutter et al. 2008)), $T = 3$ from *Myxococcus xanthus* (PDB: 4PT2 (McHugh et al. 2014)), and $T = 4$ from *Quasibacillus thermotolerans* (PDB: 6NJ8 (Giessen et al. 2019)). Encapsulin monomer folds show high structural similarity to bacteriophage HK97 (PDB: 1OHG (Helgstrand et al. 2003)) with an axial domain, peripheral domain, and elongation loop conserved

thermostable, tolerating near-boiling temperatures (Tamura et al. 2015). It is purported that encapsulins can impart this stability on their cargo proteins, which provides a glimpse into their potential native functions (Nichols et al. 2017).

1.2 Encapsulin Function

The native function of encapsulins is primarily determined by the enzymatic function of the cargo proteins packaged. Cargo encapsulation is mediated by a short targeting peptide (Sutter et al. 2008), appended as a fusion to the terminus of the cargo protein, which interacts with the encapsulin shell protein. This targeting peptide system will be elaborated upon further in subsequent sections.

Though the full gamut of native functions remains to be fully elucidated, encapsulins have been proposed to play a role in iron storage and responding to oxidative stress, based on studies over the past two decades in native and

heterologous organisms (McHugh et al. 2014; Nichols et al. 2017; Giessen and Silver 2017). One of the most enlightening studies in the area was conducted by McHugh et al. in 2014, wherein the authors studied the $T = 3$ encapsulin from *M. xanthus*, which encapsulates three ferritin-like proteins (Flp) as cargo (McHugh et al. 2014). They found that, when strains were grown under starvation conditions, the encapsulin and its cargo were all upregulated. This finding suggests a potential role for encapsulin as an iron storage system or iron sequestration pathway. Growing cells under oxidative stress conditions in the absence of encapsulin resulted in a heightened sensitivity to overall oxidative stress.

There is a suggested link between the notable stability of encapsulins and their function (Nichols et al. 2017). Supporting this idea, work by Giessen and Silver involved expressing a *Streptomyces* encapsulin with its native hemerythrin cargo in *E. coli* and treating the cells with nitric oxide and hydrogen peroxide (Giessen and Silver 2017). Hemerythrin is suggested to play a role in assisting with oxidative and nitrosative (nitric oxide) stress, featuring a diiron centre that is susceptible to strong reducing agents. Encapsulation of the hemerythrin cargo was able to shield it from reducing agents. Taking this further, they found that hemerythrin-expressing cells displayed greater protection from H_2O_2 stress when co-expressed with the corresponding encapsulin.

Despite these studies, the core function of the encapsulin shell itself is still debated. Postulated functions include the aforementioned stability enhancements (Nichols et al. 2017; Giessen and Silver 2017), sequestration of unstable or toxic molecular species (Lau et al. 2018; Sigmund et al. 2018), and a size exclusion filter for gating different molecular substrates (Lončar et al. 2020; Adamson et al. 2022; Williams et al. 2018).

1.3 Encapsulin Genetics and Evolution

Bioinformatic analysis by Andreas and Giessen has revealed close to 6000 putative encapsulin systems across bacteria and archaea (Andreas and Giessen 2021). Encapsulins appear to encapsulate a host of cargo proteins, with common cargo including dye-decolourising peroxidases (DyPs) and Flps (Sutter et al. 2008; Giessen and Silver 2017). Encapsulin shell and cargo genes tend to be found in the same operon (with the cargo identifiable via their targeting peptide), but secondary cargo proteins are sometimes identified outside of this operon (Sutter et al. 2008; Giessen and Silver 2017). Giessen and Silver identified up to four co-encapsulated cargo proteins in individual encapsulin systems, including different classes of cargo (hemerythrins, ferredoxins, peroxidases, Flps) (Giessen and Silver 2017).

While some phylogenetic analyses have suggested the existence of two different encapsulin families (Nichols et al. 2021), Andreas and Giessen have most recently grouped encapsulins into four distinct families (Andreas and Giessen 2021), analysing sequence similarity, Pfam relationships, and genome-neighbourhood clusters. Family 1 is considered to be classical encapsulins, with over 2000 of these

systems identified with 6 different cargo types. Family 1 encapsulins are the best characterised of all encapsulins, as they are found in the conventional operon format with the primary cargo protein encoded directly upstream from the shell protein (Giessen and Silver 2017). The cargo proteins are most often DyPs or Flps. Additionally, Family 1-type encapsulins in anammox bacteria are often found with an additional *N*-terminal diheme cytochrome C fusion, and these encapsulins generally form $T = 3$ assemblies (Giessen and Silver 2017; Andreas and Giessen 2021).

Family 2 encapsulins are the most common type, with over 3500 identified in the study, encapsulating four different cargo types. These encapsulins are structurally distinct from other encapsulins, having an extended *N*-terminus containing a short helix, more akin to HK97 bacteriophages (Andreas and Giessen 2021; Nichols et al. 2021). A recent structure of a Family 2 encapsulin from *Synechococcus elongatus* shows the extended *N*-terminus pointing outside of the shell, whereas in Family 1, the *N*-terminus is usually pointed into the lumen (Nichols et al. 2021). Additionally, some Family 2 encapsulins have a cNMP-binding domain, purported to assist in regulating cargo (Andreas and Giessen 2021). Of particular note, some Family 2 encapsulins are purported to have two shell components encoded in their operon, potentially representing the first two-component encapsulin shell systems uncovered (Andreas and Giessen 2021). Family 2 encapsulins have been suggested to have a role in redox homeostasis, terpenoid biosynthesis, and the utilisation of xylose and sulphur (Andreas and Giessen 2021; Nichols et al. 2021).

Family 3 encapsulins are referred to as natural product encapsulins, as they are found within biosynthetic gene clusters (Andreas and Giessen 2021). Of note, one type of Family 3 encapsulin has a *C*-terminal extension that is a purported major facilitator superfamily (MFS) domain, containing up to five transmembrane helices. The function of these helices has not been determined, but Andreas and Giessen suggest that the helices may form a hydrophobic gated channel near the shell pores or may recruit lipids to potentially form a structure akin to enveloped viruses. Tailoring enzymes in the biosynthesis gene clusters seem to have terminal extensions which may be disordered or unstructured and may act at targeting motifs for encapsulation, suggesting a putative role for encapsulins as cellular nanoreaction vessels (Andreas and Giessen 2021).

Family 4 are referred to as the A-domain encapsulins, the most structurally divergent of the families. These encapsulins are smaller than traditional HK97-fold proteins, only forming the A-domain (Andreas and Giessen 2021; Kelley et al. 2007). Family 4 are more common in archaeal systems, where their operons consist of enzymes and the encapsulin shell downstream. Their bacterial counterparts are not arranged in an operon, which makes them difficult to identify. All Family 4 members were identified in thermophilic anaerobic organisms, suggesting their functions may be relevant to their specific, harsh environment (Andreas and Giessen 2021). For the biotechnological purposes of this chapter, we will be discussing Family 1 encapsulins exclusively.

Evolutionarily speaking, encapsulins present an interesting test case wherein the different encapsulin species share high sequence similarity, but this similarity is not shared with other HK97-like capsids, a feature which prevented their identification

as HK97-like until structural data became available (Sutter et al. 2008; Bamford et al. 2005). In 2005, Bamford et al. proposed that different viral lineages (differentiated by their fold) likely evolved from a single common ancestor predating the split of their host organisms (Bamford et al. 2005). It is likely that encapsulins fall within this family tree and share origins with modern viruses. Sutter et al. proposed that perhaps viruses evolved from an encapsulin-like ancestor by changing their preferred cargo from protein to nucleic acids (Sutter et al. 2008). Later work by Andreas and Giessen discussed how each of these encapsulin families belong to the same Pfam clan (Andreas and Giessen 2021). Generally, only families linked by a common evolutionary origin fall within the same clan, suggesting that all encapsulin families are evolutionarily related to the other clan members, which include all of the HK97-fold proteins, such as viruses and phages (Andreas and Giessen 2021). Andreas and Giessen carried out a phylogenetic analysis of this clan, narrowing their search to just the most well-conserved areas such as the A-domain and parts of the P-domain, given the poor sequence conservation among those with HK97 folds. Through this analysis, Families 1–3 were found to be more closely related to each other than the other HK97-fold proteins and thus may share a recent common ancestor (Andreas and Giessen 2021). Family 4 is less related to the other encapsulin families and more closely related to the other HK97 proteins. While families have their own characteristics and suggested evolutionary linkages, these studies support the notion that encapsulins may be evolved from viruses.

2 Engineering Basics for Encapsulins

As discussed above, encapsulins house a variety of native cargo proteins. In 2008, the structure of the *T. maritima* encapsulin by Sutter et al. provided the first glimpse at the mechanism of encapsulation (Sutter et al. 2008). This structure showed extra electron density on the internal face of the encapsulin monomer, which could be fit to a short peptide at the *C*-terminus of a cistronic ferritin-like protein, coined the *C*-terminal extension. Using a similar system in *B. linens*, the group co-expressed both the encapsulin monomer and the cistronic DyP protein both with and without the *C*-terminal extension. Electron microscopy data validated that encapsulated DyP was only present inside the encapsulin when the *C*-terminal extension was expressed. The use of both *C*- and *N*-terminal peptides to target for encapsulation (henceforth referred to as targeting peptides or TPs) has been widely reported (McHugh et al. 2014; Nichols et al. 2017; Cassidy-Amstutz et al. 2016; Tamura et al. 2015; Giessen and Silver 2017; Sigmund et al. 2018; Rurup et al. 2014; He et al. 2016; Rahmanpour and Bugg 2013). TPs are required for the native functions of encapsulins and have been used in engineering studies to target heterologous cargo to the interior of encapsulins. This process occurs spontaneously in vivo, though in vitro packaging has also been demonstrated. In this section, we discuss methods and examples of both in vivo and in vitro cargo encapsulation (Fig. 2).

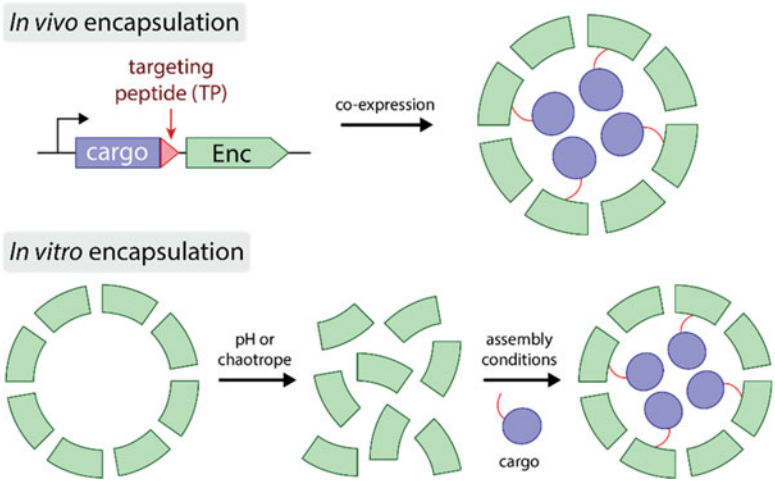


Fig. 2 Cargo encapsulation is possible in vivo or in vitro. Native encapsulin cargo is often encoded upstream of the encapsulin shell protein, with a TP on its terminus for in vivo encapsulation (Sutter et al. 2008; Giessen and Silver 2017). Encapsulins can be disassembled using harsh pH conditions or high concentrations of chaotrope and reassembled in the presence of the cargo protein for in vitro encapsulation (Cassidy-Amstutz et al. 2016; Rahmanpour and Bugg 2013)

2.1 Methods for In Vivo Encapsulation

In vivo encapsulation is a powerful tool for the customisation and spatial organisation of biosynthetic pathways in cellular factories.

In 2014, Rurup et al. demonstrated the encapsulation of heterologous cargo, a teal fluorescent protein, in the *B. linens* encapsulin with the help of a *C*-terminal "docking sequence" (i.e. the TP) (Rurup et al. 2014). Concurrently to this, Tamura et al. found that enhanced green fluorescent protein (eGFP) and firefly luciferase could both be packaged within the *Rhodococcus erythropolis* N771 encapsulin using the *C*-terminally fused TP derived from the native cargo protein (Tamura et al. 2015). The eGFP maintained its fluorescent properties, while the packaged luciferase was still active when incubated with substrate; however, the luminescence was lower than anticipated, likely owing to poor diffusion of the substrate through the pores of the encapsulin shell (Tamura et al. 2015).

The *C*-terminal TPs used in these studies were relatively long (37 residues in Rurup et al.), as a minimal tag length for targeting had yet to be established. A study by Cassidy-Amstutz et al. compared sequence alignments of the *C*-termini from many different protein cargoes, noticing that only the last ten amino acids were strongly conserved across the different proteins. Trialling different lengths of *C*-terminally fused TP (0, 5, 15, 30 residues) for the packaging of superfolder GFP (sfGFP) in the *T. maritima* encapsulin, equally efficient cargo loading was observed for the 15 and 30 residue TPs, leading to the assertion that 15 or more residues are sufficient for cargo loading (Cassidy-Amstutz et al. 2016).

While these earliest studies showed that single heterologous proteins could be packaged in encapsulins using the targeting peptide, multiple heterologous cargo proteins can be encapsulated, given that some encapsulins natively have multiple cargoes. In 2018, Lau et al. demonstrated targeting of multiple heterologous protein cargoes to the *M. xanthus* encapsulin when expressed in yeast (Lau et al. 2018). Using a split-Venus system (Kodama and Hu 2010), both halves of the split fluorescent protein were targeted to the encapsulin shell. Upon co-expression, co-encapsulation was confirmed upon observing higher fluorescence intensity when the encapsulin was expressed (Lau et al. 2018).

In all the above examples, cargo loading took place via a *C*-terminal TP, which is the most common route of encapsulation. In 2017, Giessen and Silver uncovered an encapsulin with an iron-mineralising encapsulin-associated Firmicute (IMEF) cargo protein with a standard *C*-terminal TP, plus an additional ferredoxin cargo in the same operon with an *N*-terminal TP (Giessen and Silver 2017). When both cargoes were co-expressed with the encapsulin shell, only the IMEF protein was detected. However, expression of only the ferredoxin and encapsulin shell led to detectable encapsulation. This result raises interesting questions regarding the efficiency of this *N*-terminal TP loading system and whether changing the fusion terminus can affect the stoichiometry of loaded cargo.

While use of the native TP sequence is the most popular route to achieving cargo encapsulation, alternative strategies may be possible. Akita et al. reported the structure of *Pyrococcus furiosus* encapsulin with a disordered *N*-terminus consisting of 109 residues, which is purported to be the cargo protein (Akita et al. 2007; Nichols et al. 2017). This shows it is possible to achieve encapsulation via the creation of a direct monomer-cargo fusion, rather than relying on the stochastic packaging which arises from a non-covalent interaction.

2.2 Methods for In Vitro Encapsulation

In vitro loading of encapsulins can be used for the encapsulation of metals, inorganic materials, therapeutics, and other such products, with applications such as "cell-free" nano-factories or targeted delivery vehicles.

An early study in 2013 by Rahmanpour and Bugg showed that the encapsulin from *Rhodococcus jostii* RHA1 could be disassembled in acetate buffer at pH 3 and reassembled using buffer at pH 7, although dynamic light scattering (DLS) suggested that the resulting reassembled nanocompartments may be slightly distended (Rahmanpour and Bugg 2013). The authors incubated the disassembled encapsulin sample with its native DyP cargo and brought the system back to the refolding pH before separating the mixture using size exclusion chromatography (SEC). The DyP cargo was found to co-elute with encapsulin, suggesting that the cargo was loaded in vitro following controlled disassembly and reassembly (Rahmanpour and Bugg 2013).

In 2016, Cassidy-Amstutz et al. conducted a similar study using the *T. maritima* encapsulin, achieving complete shell disassembly at pH 1 or 14 with minimal precipitation (Cassidy-Amstutz et al. 2016). Tracking protein structure by circular dichroism, the encapsulin protomer showed a high degree of helicity under neutral conditions, whereas at pH 13, the majority of the protomer was unfolded. Upon returning to neutral pH, an almost complete conversion back to the native spectrum was observed. In strongly acidic (pH 1) conditions, rather than unfolding, the protomers appeared to adopt a majority β-sheet alternative structure, reverting back to the native spectrum upon neutralisation. Reversible disassembly was attempted using urea and guanidine hydrochloride (GdnHCl) as chaotropes. While concentrations up to 12 M urea did not completely unfold the encapsulin, 7 M GdnHCl was successful in unfolding the protein. However, only about 60% of the encapsulin was able to refold following treatment as opposed to 88% or 95% recovery post pH-denaturation. For the in vitro loading of sfGFP appended with a *C*-terminal TP, minimal loading was observed in the pH denatured refolded encapsulin samples, while significant in vitro loading was observed for the encapsulins that were disassembled with GdnHCl (Cassidy-Amstutz et al. 2016). As an interesting note, a short 5 residue TP was found to have similar in vitro loading efficiency to that of 15 or 30 residue targeting peptides, despite the fact that the 5 residue TP was previously unable to load cargo in vivo (Cassidy-Amstutz et al. 2016).

3 Examples of Biomanufacturing Using Encapsulins

The canonical role of encapsulins in biomanufacturing is for the construction of nanoreactors, providing a protective shell to surround enzymes that are capable of generating a desired product. In this section, we will provide examples of encapsulins being used to create simple nanoreactor systems, discussing the effects of encapsulation upon enzyme activity (Fig. 3).

While not discussed further in this chapter, we note that encapsulin-based biopharmaceuticals have also been developed, including vaccines (Choi et al. 2016; Lagoutte et al. 2018; Jones and Giessen 2021), imaging probes (Putri et al. 2016; Moon et al. 2016), and drug-delivery vehicles (Moon et al. 2014a). These products are made out of encapsulin proteins themselves, rather than the encapsulin serving as catalytic factories for biomanufacturing.

3.1 Examples of In Vitro Nanoreactors

In 2020, Lončar et al. built a variety of nanoreactors using the *Mycolicibacterium hassiacum* encapsulin via peptide-directed encapsulation (Lončar et al. 2020). First, they encapsulated an orthogonal DyP peroxidase from *Saccharomonospora viridis*

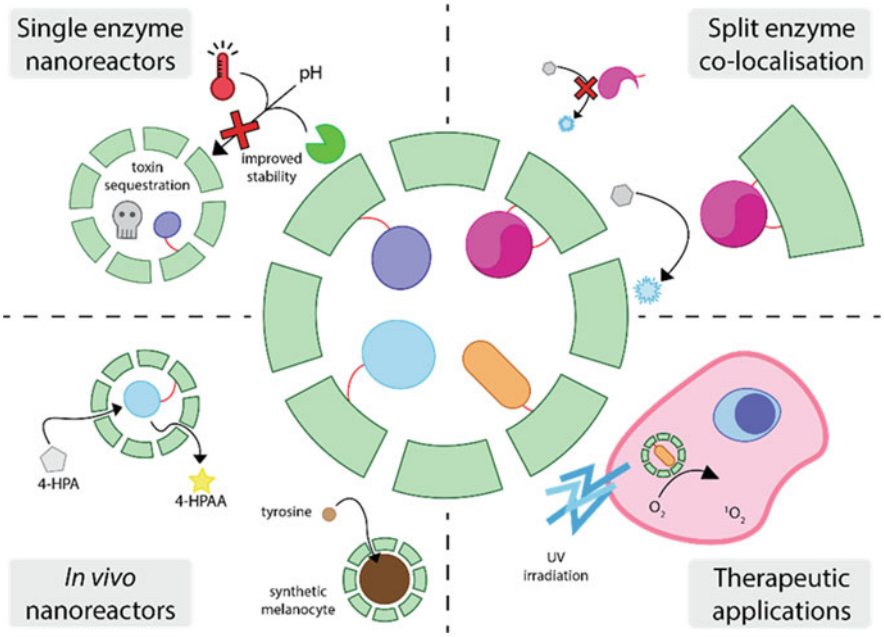


Fig. 3 Simple encapsulin nanoreactors consist of a single encapsulated enzyme. Top left: encapsulin shells confer stability and protection to enzymes during catalysis (Nichols et al. 2017; Giessen and Silver 2017). Bottom left: simple nanoreactors can be used for a variety of purposes including the generation of *in vivo* cell factories to recreate melanosomes (Sigmund et al. 2018) for imaging or biomanufacturing therapeutic precursors (Lau et al. 2018). Top right: split-enzyme systems allow for greater control over the sequestration of catalysis, as components are more likely to co-localise within the shell than in the cytosol (Sigmund et al. 2018). Bottom right: nanoreactors can be used to generate ROS for targeted cancer therapy (Diaz et al. 2021)

DSM43017 via appending the *C*-terminal TP from its native DyP cargo. The enzyme was successfully encapsulated, and upon assaying enzyme activity, the encapsulated version had a slightly lower k_{cat} and slightly higher K_m. Interestingly, when heated to 40 °C and assayed again, the free enzyme was no longer active after 30 min, whereas the activity of the encapsulated enzyme increased over the first few hours, with substantial loss in activity only after 25 h. The authors suggest that the stabilising effect of the encapsulin may be related to cargo crowding in the cage which might prevent aggregation. Subsequent encapsulation of a carbohydrate oxidase flavoenzyme also resulted in increased enzyme thermostability at 50 °C. Furthermore, the authors prepared an enzyme cascade consisting of a free eugenol oxidase and encapsulated peroxidase which, in tandem, converted eugenol to lignin-like materials. There were differences in the abundance of certain linkage types in the lignin products when using encapsulated and non-encapsulated peroxidase. This result suggests a potential selectivity shift due to enzyme encapsulation (Lončar et al. 2020). Thus, preferential production of certain product geometries through

encapsulation has great potential in tailoring the manufacturing of molecules where selectivity is required.

Similar to the split-Venus system discussed in Sect. 2.1 (Lau et al. 2018), Sigmund et al. encapsulated split luciferase inside *M. xanthus* encapsulin. In studies conducted on cell lysates, substantially more luminescent product was generated in the presence of the encapsulin shell, suggesting that the observed enzymatic activity of a split system may be enhanced upon encapsulation (Sigmund et al. 2018). Such split systems provide an additional element of control over nanoreactors, ensuring that catalysis predominantly occurs within the shell.

To use encapsulin nanoreactors for in vitro applications, it is often necessary to extract these systems from their expression host. Due to their large size, one of the most common initial steps in encapsulin purification is ultracentrifugation with sucrose or cesium chloride (Akita et al. 2007; Sutter et al. 2008; McHugh et al. 2014; Cassidy-Amstutz et al. 2016; He et al. 2016). In 2014, Rurup et al. (2015) published a protocol for extraction and purification of encapsulin nanocompartments featuring a sucrose cushion step followed by a sucrose gradient. Upon extracting the band of appropriate density, the mixture was further purified by size exclusion chromatography.

Other encapsulin purification methods achieve coarse size separation by inducing precipitation of large molecules through polyethylene glycol (PEG) or ammonium sulphate precipitation of crude mixtures (Giessen and Silver 2017; Lau et al. 2018; Williams et al. 2018; Lee et al. 2020). Following precipitation, the sample is resolubilised and applied to separation techniques such as size exclusion or ion-exchange chromatography.

Surface-exposed regions of some encapsulins have also been engineered to include polyhistidine tags (Tamura et al. 2015; Diaz et al. 2021; Lee et al. 2020; Sonotaki et al. 2017), enabling immobilised metal affinity chromatography (IMAC) as an initial purification step, followed by subsequent chromatographic techniques.

As some encapsulins are native to thermophilic organisms, such as *Thermotoga maritima*, impurities can be removed by heat denaturation. Work by Moon et al. in 2014 demonstrated that His tags could be inserted into exterior loops which could not only be used in IMAC but also increased the thermostability of the shell such that lysates could be heated to high temperatures, to precipitate less stable proteins. The remaining mixture could then be refined in downstream processes. While this approach may not be suitable for encapsulins carrying cargo, this method was used (Lee et al. 2020) to purify encapsulins fused to lumen-facing antimicrobial peptides (discussed in greater detail in Sect. 4.6).

3.2 In Vivo Applications of Nanoreactors

Encapsulins are also suitable for use in whole cell biocatalysis, avoiding the extra downstream processing steps required in the in vitro context**.** Early work on in vivo encapsulation of non-native cargo used bacterial hosts (Rurup et al. 2014). However,

depending on the criteria of the nanoreactor application (i.e., enzymes used, pathways involved, intermediates, or products produced), different hosts may be more suitable. Below, we describe examples of novel encapsulin nanoreactor systems produced in a range of cellular hosts.

In 2018, two parallel in vivo studies demonstrated the use of these prokaryotic nanocompartments as orthogonal nanoreactors in eukaryotic cells. Sigmund et al. used the *M. xanthus* encapsulin to build melanosomes in mammalian cells (Sigmund et al. 2018), using the enzyme tyrosinase which converts tyrosine into melanin, a gene reporter in optoacoustic tomography (Sigmund et al. 2018; Stritzker et al. 2013; Jathoul et al. 2015). Importantly, melanin is toxic to cells unless it is sequestered within melanosome compartments. After fusing tyrosinase to *M. xanthus* endogenous cargo protein D (McHugh et al. 2014) to encapsulate the enzyme, the resulting synthetic melanosomes were capable of melanin production without the toxicity often observed with tyrosinase expression, leading to more viable cells overall (Sigmund et al. 2018).

Within the context of producing pharmaceutically relevant molecular precursors, Lau et al. showed that *M. xanthus* encapsulin assembly and cargo packaging were robust processes in the yeast *Saccharomyces cerevisiae* (Lau et al. 2018). The yeast enzyme Aro10p was chosen as cargo for its ability to catalyse the conversion of 4-hydroxyphenylpyruvate (4-HPP) to 4-hydroxyphenylacetaldehyde (4-HPAA), a key intermediate for the production of benzylisoquinoline alkaloids (opioids). The in vivo synthesis of 4-HPAA in yeast is of great therapeutic value, but 4-HPAA is toxic to yeast and subject to endogenous degradation, thus justifying a nanoreactor setup. Upon yeast co-expression of encapsulin and Aro10p-TP with a targeting peptide appended to its *C*-terminus, enzymatic turnover to produce 4-HPAA was observed, providing a proof of concept for engineering nanoreactors as orthogonal synthetic organelles which can generate therapeutic precursors (Lau et al. 2018).

Encapsulin nanoreactors have also been explored as a system for facilitating photodynamic therapy. In a study by Diaz and coworkers, *T. maritima* encapsulin was loaded with a mini-singlet oxygen generator (miniSOG) via a *C*-terminal TP (Diaz et al. 2021). MiniSOG is a photosensitiser which converts molecular oxygen to reactive oxygen species (ROS) via blue light stimulation of its flavin mononucleotide (FMN) chromophore. The authors first validated the self-assembly of the encapsulin and the successful loading of cargo, finding that while regular assemblies were formed, mini-SOG fluorescence was partially attenuated when encapsulated. This attenuation was attributed to structural alterations near the chromophore site as a result of encapsulation. Measurements of ROS production upon irradiation with blue light for 10 min showed that empty encapsulin generated small amounts of ROS (likely due to the endogenous flavin of *T. maritima* encapsulin (Diaz et al. 2021; Künzle et al. 2018)), while free miniSOG produced 2.2-fold more ROS, and the encapsulated miniSOG system produced 4.3-fold more ROS than encapsulin alone, representing a 1.9-fold increase for encapsulated enzyme versus free enzyme. The increase was attributed to the combination of ROS generation from encapsulin and miniSOG (Diaz et al. 2021). ROS from the encapsulated miniSOG system led to structural damage to the shell, with larger and less-ordered assemblies observed by

DLS and transmission electron microscopy. The authors subsequently explored the effectiveness of ROS generation in encapsulins as a possible photodynamic therapy (PDT) for cancer. MiniSOG had previously been examined as a possible PDT candidate by Mironova et al. and was found to have cytotoxic activity (Mironova et al. 2013). Encapsulated and free miniSOG were incubated with A549 human lung adenocarcinoma cells, with the encapsulated miniSOG endocytosed over time (greatest uptake at 8 and 12 h incubation), while the free miniSOG never entered cells. Further, both empty and miniSOG encapsulins did not harm cells in the absence of irradiation. Upon irradiation, only the cells treated with encapsulated miniSOG showed reduced viability by approximately 34%. While this proof of concept ROS system has limited application in medicine in its current form, considering that blue light cannot deeply penetrate tissue, the creation of a light-responsive encapsulin could find utility in other applications where ROS generation is useful.

4 Advanced Engineering and Examples in Biomanufacturing

The previous examples of engineering and biomanufacturing with encapsulins have involved a single type of encapsulated enzyme. In this section, we will detail advanced engineering considerations beyond single enzymes and discuss the ways in which these methods have been used in practice (Fig. 4).

4.1 Molecular Display on the Encapsulin Surface

There are several methods used to display molecules on the surface of encapsulin shells, including chemical conjugation, genetic fusion, and post-translational protein-mediated covalent attachment, as reviewed by Szyszka et al. (2021). Chemical conjugation is a popular method for the attachment of dyes, fluorophores, or other small molecules (Lagoutte et al. 2018; Moon et al. 2014a; Bae et al. 2018; Klem et al. 2018), while genetic fusion is a popular method for the attachment of small peptides or heterologous proteins (Lagoutte et al. 2018; Moon et al. 2014a; Jenkins and Lutz 2021; Moon et al. 2014b).

Recent studies have identified a flexible surface-exposed loop on the *T. maritima* encapsulin which is amenable to the inclusion of peptides (Lagoutte et al. 2018; Moon et al. 2014a; Bae et al. 2018; Moon et al. 2014b). Landmark work by Moon and colleagues took advantage of both chemical conjugation and the fusion of a small peptide by inserting a cell-binding peptide into this surface loop and conjugating a prodrug aldoxorubicin (AlDox) to a surface-exposed cysteine. The AlDox drug was connected to the shell via a hydrazone linker which cleaves in low-pH environments such as a tumour cell. When cells were treated with these dual-display

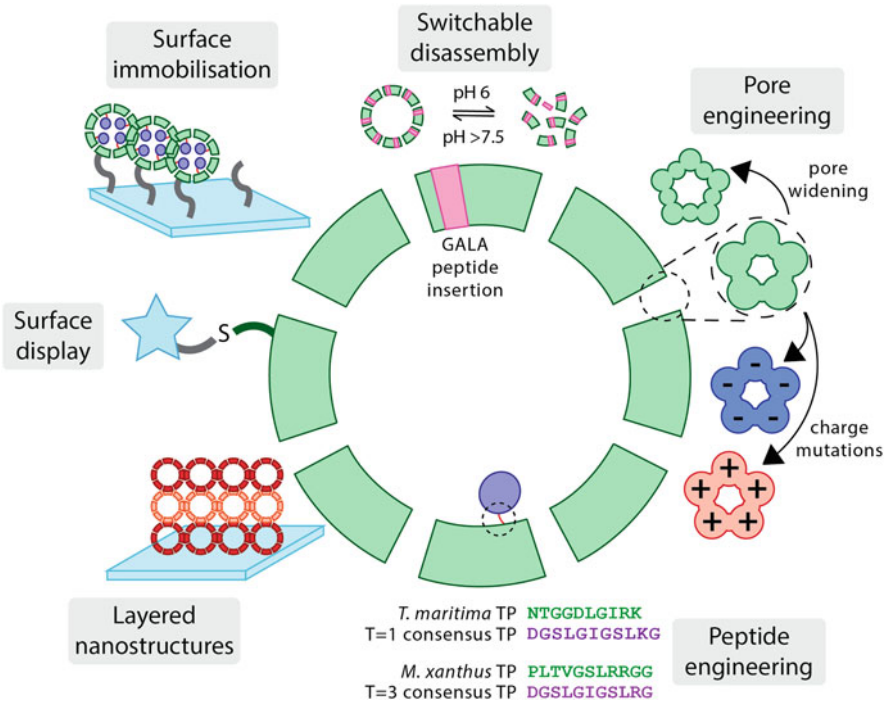


Fig. 4 Advances in encapsulin engineering have allowed for more sophisticated nanoreactor designs. Modulating the surface of encapsulins allows for the display of various moieties for catalysis (Jenkins and Lutz 2021), targeting (Moon et al. 2014a; Bae et al. 2018), surface immobilisation (Choi et al. 2021; Putri et al. 2017), and creating hierarchical nanostructures (Choi et al. 2021). Insertion of a GALA peptide in to the encapsulin monomer leads to controlled disassembly in mildly acidic conditions (Giessen et al. 2021). Pore engineering studies have demonstrated that encapsulin pores are amenable to mutation and these alterations affect molecular flux (Adamson et al. 2022; Williams et al. 2018). By learning the details of the TP-monomer interaction, protein engineers can affect the stoichiometry of cargo loading through engineered TPs (Altenburg et al. 2021)

encapsulins, the authors noted a similar level of cytotoxicity between the encapsulin drug and the free drug (Moon et al. 2014a).

Another popular method of display is post-translational protein-mediated covalent attachment using the SpyTag/SpyCatcher system. SpyTag and SpyCatcher were developed by the Howarth lab in 2012, in which two split fragments of the fibronectin-binding protein FbaB structure form an irreversible isopeptide bond, thus linking proteins which are fused to the split fragments (Zakeri et al. 2012). In one recent encapsulin example, Bae et al. genetically fused the SpyTag into the exposed loops of the *T. maritima* encapsulin and used affibody-SpyCatcher fusions to dual-functionalise the shells for targeting two different cell types (Bae et al. 2018).

Many applications of encapsulin display studies have been for the targeting of encapsulins to particular cell targets or for the display of cargo. One alternative

intriguing application of molecular display is to enhance the stability and protective ability of the shell itself. While encapsulins are incredibly robust and highly stable, they are not completely immune to disruption. Klem and co-authors sought to enhance the resistance of encapsulin to proteolysis by chemically conjugating cysteine knottin proteins to the encapsulin exterior (Klem et al. 2018). Knottins are a type of knot-shaped cysteine-stabilised miniprotein which are highly thermostable and resistant to proteases. Trypsin inhibitor EETI-II knottin was modified to make the *N*-terminus more flexible and surface-exposed. A heterobifunctional chemical linker was then used to chemically conjugate the knottin *N*-terminus to the exposed cysteine on the *T. maritima* encapsulin surface. After treatment with trypsin, they found that bare encapsulins showed a resistance to proteolysis, but this was enhanced by the presence of the knottins, though the inhibition was not to the extent of unbound knottins (Klem et al. 2018). While this system requires further optimisation, it presents an interesting idea of using surface modifications to improve the already impressive stability of encapsulins.

4.2 Engineering Encapsulin Pores

Encapsulins have numerous pores at their vertices as a result of their icosahedral symmetry. These pores govern which chemical species can freely diffuse in and out of the lumenal space and have been shown to be amenable to mutation (Adamson et al. 2022; Williams et al. 2018). Controlling the flux of different chemical species in and out of the encapsulin shell would provide a powerful dimension of control over the reactions catalysed by encapsulated cargo. Williams et al. created mutants of the *T. maritima* encapsulin with various deletions in the five-fold symmetric pore (Williams et al. 2018), finding that some assemblies still formed despite these drastic mutations and that widening of the pores increased the flux of Tb^{3+} ions into the shell in a lanthanide luminescence assay (Williams et al. 2018). Building upon this study, Adamson and coworkers created an extensive library of fivefold pore mutant encapsulins, altering both the pore size and charge through the choice of amino acids lining the pore (Adamson et al. 2022). In molecular dynamics simulations, both size and charge alterations affected the flux of chloride ions through the pores, but the flux of Tb^{3+} ions was largely unaffected by the charge and size differences across the mutants, a result confirmed experimentally by replicating the Tb^{3+} assay from Adamson et al. (2022) and Williams et al. (2018). Simulations suggest this unexpected effect may result from the highly charged ions non-specifically associating with negative patches on the encapsulin particle surface (Adamson et al. 2022). The results from these two studies suggest that both the size and chemistry of the pores are important for governing molecular flux but that the molecular details and experimental measurement of this process are complex. Future work in this area is required to ascertain the details of porosity, but the ability to control the flux of molecules in and out of shells would reveal an additional level of control over

nanoreactor designs and allow for more sophisticated biomanufacturing both in vivo and in vitro.

In a recent example utilising both molecular display and pore engineering, Jenkins and Lutz designed dual-enzyme metabolons using the *T. maritima* encapsulin (Jenkins and Lutz 2021), encapsulating the *Sphingomonas paucimobilis* demethylase LigM inside the shell via a *C*-terminal TP. Interestingly, SpyCatcher was also genetically fused to the *C*-terminus of the encapsulin protein such that each assembly displayed 60 copies of the fused SpyCatcher, notable given its ~12 kDa size. The fact that the addition of a full-length protein fusion to encapsulin did not disrupt assembly potentially paves the way for the display of other larger proteins. The displayed SpyCatcher was then ligated with SpyTag fused to dihydrofolate reductase (DHFR). The rationale behind this design was that in the external environment, DHFR would produce tetrahydrofolate (THF) from dihydrofolate (DHF), which would then enter the encapsulin shell via the pores to assist LigM in the demethylation of vanillate to form protocatechuate (PCA) as well as *N*5-methyltetrahydrofolate from the methylation of THF. Despite both enzymes being active, less LigM product was produced using this metabolon than the authors had expected. As the authors postulated that the pores of the encapsulin were preventing the diffusion of THF into the lumen, they used an aforementioned mutant with a significantly widened fivefold pore and attained expected levels of product. This example demonstrates the effects of enzyme encapsulation, enzyme display (both via SpyTag/SpyCatcher ligation and genetic fusion), and pore engineering to create a multi enzyme metabolon.

4.3 Immobilisation of Encapsulin onto a Surface

Surface immobilisation of nanoreactors can enable the creation of flow reactors, recyclable catalysts, and sensors (Putri et al. 2017; Sassolas et al. 2012; Sheldon et al. 2021; Sheldon and van Pelt 2013). Putri et al. constructed an immobilised system by tethering *B. linens* encapsulins containing DyP peroxidase onto glass. The glass surface was activated using pentafluorophenyl silicate, creating a surface capable of reacting with exposed lysine residues present on the surface of the *B. linens* encapsulin shell (Putri et al. 2017; Liu et al. 2017). Based on scanning electron microscopy images, they estimated that the glass surface was approximately 35% occupied with encapsulins. Testing the encapsulin/DyP system as a nanoreactor in concert with glucose oxidase for the H_2O_2-mediated oxidation of dye substrate 2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS), the k_{cat} of encapsulated DyP was measured to be 115.4 s^{-1}, which is within the same order of magnitude as the unencapsulated literature (Salvachua et al. 2013) k_{cat} of 224 s^{-1}. This study was a proof of concept that encapsulins containing a cargo enzyme can retain their activity when immobilised.

A similar study in 2021 used the SpyTag/SpyCatcher system to display glutathione-*S*-transferase (GST) on the *T. maritima* shell surface (Choi et al. 2021).

Encapsulins were loaded with NanoLuc luciferase and displayed GST to immobilise the NanoLuc-filled encapsulin on plates coated with glutathione (GSH). Luminescence from the turnover of substrate furimazine was observed, and the plates were found to be reusable for up to six wash and reaction cycles, demonstrating the feasibility of a regeneratable immobilised enzyme system.

4.4 Controlled Disassembly and Reassembly Using GALA Peptide

While cargo proteins can easily be packaged during in vivo co-expression, in vitro packaging is necessary when encapsulating synthetic non-protein cargo. While disassembly of native encapsulins is possible under harsh conditions (Cassidy-Amstutz et al. 2016), this may be suboptimal when packaging sensitive cargo that requires near-physiological conditions. In a study by Giessen et al. (2021), the group engineered the $T = 4$ *Q. thermotolerans* encapsulin to enable reversible assembly and disassembly under mild conditions. Engineering efforts were focused on the E loop of the encapsulin monomer, as it is not implicated in cargo loading but is vital for compartment assembly. The authors chose to use a GALA peptide, a pH-sensitive motif which switches from a coil to a more rigid helix under mildly acidic conditions, inserting the peptide in the E loop flanked by flexible triglycine linkers. Under neutral conditions, the coil conformation was expected to allow for the E loop to maintain its regular intersubunit contacts, while switching to the helix upon acidification would induce strain, leading to loss of intersubunit contacts and thus disassembly of the shell. Indeed, the engineered cage was able to form assemblies under neutral conditions similar to the wild-type encapsulin, while disassembling almost completely at pH 6.0. Measuring hydrodynamic radius by DLS, the encapsulin was observed to shift to a large size upon reassembly, which may be a result of disorder at the insert site, or potentially indicative of slight aggregation or the presence of some deformed assemblies. Functional disassembly was demonstrated by encapsulating one part of split GFP via the regular TP-mediated in vivo cargo encapsulation process and then mixing the other split part in either assembly or disassembly pH conditions. The observed fluorescence was much higher under disassembly conditions, suggesting that the encapsulin can safely house cargo until it is triggered for disassembly. In a separate experiment, encapsulated glucose dehydrogenase was found to retain its enzymatic activity upon encapsulin disassembly, although it is worth noting the kinetics of the free unencapsulated enzyme were much faster than the encapsulated enzyme or that of the exposed enzyme post-disassembly. This effect is likely due to documented slow down effects of tethered or encapsulated enzymes, or pore diffusion acting as a barrier (Giessen et al. 2021; Lee et al. 2011; Jordan et al. 2016). Loading of protein cargo was also demonstrated in vitro, by mixing disassembled encapsulin with TP-tagged mNeon then reassembling the shell, as indicated by a high-molecular-

weight fluorescent band in native PAGE. This work opens a new dimension of engineering possibilities based on controlled release of cargo and in vitro packaging.

4.5 Engineering Targeting Peptide Interactions

Modulation of the encapsulin TP system was reported by Altenburg et al., with the goal of achieving tighter control over the stoichiometry of cargo loading (Altenburg et al. 2021). The authors used Rosetta FlexPepDock (Raveh et al. 2011; London et al. 2011) to model the native targeting peptides for both *T. maritima* and *M. xanthus* encapsulins, consensus targeting peptides for $T = 1$ and $T = 3$ encapsulin systems, and a GGS negative control peptide interacting with encapsulin shells. Overall, the native targeting peptides resulted in the highest scoring interactions, with lower scores for the consensus peptides, and the lowest scores for the GGS peptide. While the $T = 1$ native and consensus targeting peptide scores had similar scores, the $T = 3$ native and consensus were quite different, as the consensus $T = 3$ TP scored similarly to the GGS peptide (Altenburg et al. 2021).

In general, the $T = 1$ targeting peptide interaction is governed by a hydrophobic pocket in the shell interacting with a key LGI motif in the targeting peptides, where the middle glycine provides the needed flexibility for this interaction to take place. *C*-terminal lysine residues in the native and consensus peptides interact with a negative patch on shell. Additionally, the native peptide has an internal salt bridge between an aspartic acid and arginine which was noted in the crystal structure (Sutter et al. 2008). The loss of this interaction and unfavourable conformation due to an additional leucine are suggested to weaken the binding of the consensus peptide (Altenburg et al. 2021). In the native $T = 3$ interaction, modelling suggests that the interaction is largely mediated by a VxxxxR motif, wherein the valine interacts with a hydrophobic patch and the arginine interacts with a negative patch via hydrogen bond formation. Despite a conserved GSL motif (Sutter et al. 2008) separating the valine and arginine residues, the $T = 3$ consensus peptide does not have optimal spacing between the key hydrophobic residue and the penultimate arginine to allow for the same interaction. Instead, the primary interaction with the shell is via the arginine, resulting in a low score comparable to the GGS peptide. The modelling results were validated in experiments where the different peptides were used to load a fluorescent cargo protein. The native and consensus $T = 1$ TPs showed similar levels of loading and the native $T = 3$ targeting peptide showed substantially more loading than the consensus and GGS peptides. Interestingly, the loading percentage was higher overall in the $T = 3$ system which authors attributed to the quicker in vivo assembly of the $T = 1$ shell than the $T = 3$ shell in their expression system (Altenburg et al. 2021).

Using this information, novel TPs were designed for the $T = 3$ system through a screen of point mutants, creating 220 novel sequences where each position was changed to one of the 20 canonical amino acids. In FlexPepDock, mutating the last three residues (R9, G10, G11) in the sequence had the largest detrimental effects to

binding. Mutating the conserved G5, S6, and L7 also lowered the scores, in part due to a hydrogen bond from S6 to the binding pocket. However, mutating residues 2–4 largely improved binding. The negative mutations were experimentally validated through three mutant peptides, S6A, R9A, and G10W, resulting in the expected attenuation of loading efficiency (Altenburg et al. 2021). While this study characterised the TPs of only two types of encapsulins, it demonstrates the possibility of controlled cargo loading as an additional parameter in nanoreactor design. By controlling the efficiency and stoichiometry of cargo loading (particularly in multiple cargo systems), more sophisticated designs for metabolic control may be possible.

4.6 Other Advanced Engineering Examples

A study by Choi et al. showcases the combination of multiple techniques for encapsulin engineering, using split inteins rather than TPs to encapsulate cargo molecules (Choi et al. 2021). Split inteins are split components of intervening proteins, which carry out post-translational splicing to covalently link proteins attached as fusions to each component (Shah and Muir 2014). NanoLuc luciferase was fused to an *N*-terminal split intein and encapsulated into *T. maritima* encapsulin bearing on its lumen-facing *N*-terminus the corresponding *C*-terminal split intein. The encapsulated NanoLuc showed similar activity to the unencapsulated version.

The authors subsequently created a layered encapsulin nanostructure by leveraging the rapamycin-induced heterodimerisation of FKB12 and FRB (Choi et al. 2021; Banaszynski et al. 2005). Either FKB12 or FRB was displayed on the outside of NanoLuc-filled encapsulins using SpyTag/SpyCatcher ligation. By alternating the addition of the two shell types with rapamycin onto a surface, they were able to observe the formation of the layered nanomaterial using a quartz crystal microbalance. The encapsulated enzyme retained activity, but as the turnover rate of NanoLuc for furimazine is extremely fast (Hall et al. 2012), it was difficult to compare the enzymatic activity of the nanomaterial to a single layer, as the top layer of the material could potentially turn over all available substrate faster than the ability of the substrate to permeate through the layers. The authors noted that using high concentrations of furimazine resulted in brighter intensities in the multilayer assemblies relative to single layer.

Another interesting study in 2020 showed that encapsulins could assist the expression of antimicrobial peptides (AMPs), which are not easily expressed in bacterial hosts due to their bactericidal activity (Lee et al. 2020). Lee et al. created fusion constructs connecting the AMP to the *T. maritima* encapsulin *N*-terminus via a TEV site-containing linker, resulting in fused AMP inside the encapsulin shell. The encapsulin was also engineered to feature TEV sites throughout the shell. The group found that the fusions could be expressed and that a fraction of the peptides were released upon TEV protease treatment. While the system requires additional

optimisation, this work represents an innovative use of encapsulation to produce challenging biopharmaceuticals.

An additional cargo loading feature is the specific enrichment of expressed cargo proteins in an encapsulin shell by fusing a degradation tag (Lau et al. 2018; Sigmund et al. 2018). Sigmund et al. investigated this method by appending a targeting peptide to the *C*-terminus of a fluorescent protein and a degradation tag (Maynard-Smith et al. 2007) to its *N*-terminus (Sigmund et al. 2018). Upon expression and cargo loading in mammalian cells, degradation of the fluorescent cargo was blocked upon its sequestration within the *M. xanthus* encapsulin. The same phenomenon was also demonstrated within yeast cells in a concurrent study by Lau et al. (2018). As there is often unencapsulated cargo after expression which can present issues during subsequent purification steps, this use of degradation tags provides a strategy for streamlining purification.

5 Conclusions and Future Opportunities

Encapsulin protein cages have become popular engineering targets for biomanufacturing applications in the past decade. While the native functions of encapsulin nanocompartments still remain to be fully elucidated, we have gained sufficient understand about how these systems are built to enable repurposing for alternate in vitro and in vivo applications, especially as nanoreactors for carrying out non-endogenous chemical transformations.

Of the many systems available in microorganisms to exploit for biomanufacturing, encapsulins possess several unique advantages. Encapsulins are simple, assembling from only a single protein monomer type, with robust in vivo self-assembly characteristics that include cargo packaging (Sutter et al. 2008; Tamura et al. 2015). They are remarkably stable to high temperatures, harsh pH, and proteolysis (Nichols et al. 2017; Cassidy-Amstutz et al. 2016; Tamura et al. 2015), further enhancing their suitability for nanoreactor applications.

Despite their structural stability, encapsulins are not impervious to perturbations that can disrupt the assembly. The work of Diaz et al. on the effects of ROS (Diaz et al. 2021) and work by both Cassidy-Amstutz et al., and Jones et al. on in vitro packaging shows that engineering can lead to abnormal assemblies or aggregation (Cassidy-Amstutz et al. 2016; Giessen et al. 2021). Further enhancement of encapsulin stability, including precise control over the stability of disassembled forms, would be of incredible value for biomanufacturing. An encapsulin shell with extreme stability could allow for longer enzymatic turnover within the shell or prolong the sequestration of potentially toxic products or intermediates.

We have learnt a substantial amount about the encapsulation of molecular cargo. This is most often achieved with the *C*-terminal TP appended to an exogenous protein or enzyme of interest. However, control of packing stoichiometry and the realistic upper occupancy limit is still unknown, though it is expected that steric limits exist within the shells, particularly for large or bulky cargo. Shell occupancy

limits will become more important in multienzyme systems where precise control over cargo loading is required. Currently, it is challenging to tune the expression levels of encapsulins and cargo molecules. Tuning the TP-monomer interaction by modulating key interacting residues may help control the stoichiometry of loading (Altenburg et al. 2021), though more studies are required to test this concept in multi-cargo systems. In addition, Choi et al.'s innovative use of inteins to encapsulate cargo provides an orthogonal route to encapsulate cargo (Choi et al. 2021), providing a potential route towards overcoming the native limit of a single bound cargo per encapsulin monomer.

Encapsulins have also proven to be highly amenable to mutation or structural alteration, which can potentially be exploited to increase control over substrate flux. Williams et al. and Adamson et al. both showed that encapsulins are highly tolerant to mutations at their pores, forming correct assemblies despite large deletions or charge alterations at these sites (Adamson et al. 2022; Williams et al. 2018). Controlling the permeability of encapsulin shells would allow for more complicated chemistry within encapsulin shells, as the flux of molecules into and out of the compartments could be regulated. These pore engineering approaches may allow encapsulins to emulate complicated biological systems such as carboxysomes (Borden and Savage 2021), while using simpler protein building blocks.

Another major area of interest in encapsulin engineering is the controlled assembly and disassembly of the compartments. While in vivo assembly is incredibly useful for making cellular nanoreactors (Lau et al. 2018; Sigmund et al. 2018), in vitro loading would allow for the encapsulation of inorganic or chemically synthesised small molecules for innumerable purposes. Switchable disassembly in non-denaturing conditions (Giessen et al. 2021) would enable engineers to control the release of cargo in targeted scenarios. While switchable encapsulin systems exist, further refinement may enable multiple cycles of assembly without loss of material to aggregation or mis-assembly pathways.

While there has been considerable focus on systems that exist or could exist in vivo, encapsulins are also well suited to fully in vitro scenarios. Cargo-loaded encapsulins can be immobilised onto surfaces, with enzymatic cargo retaining activity after multiple catalytic cycles (Choi et al. 2021; Putri et al. 2017), suggesting that encapsulins could be used to create catalytic materials for plate- or flow-based biological manufacturing. These in vitro applications are particularly valuable in cases where the metabolic complexity of cellular factories result in unwanted complications, such as the effects of metabolic burden, crosstalk with endogenous pathways, and the challenges associated with achieving precise and stable expression levels.

In summary, encapsulins have great potential as nanoreactors for biomanufacturing purposes, due to their robust structural properties and engineerability. In almost 30 years since their discovery, synthetic biologists have now brought encapsulins tantalisingly close to industrial applications. While there still are considerable practical challenges to overcome, there appears to be no shortage of innovation and interesting engineering strategies still to come.

Acknowledgements YHL acknowledges funding from the Australian Research Council (DE190100624) and the Westpac Scholars Trust (WRF2020). LSRA acknowledges funding from a CSIRO SynBio Future Science Platform top-up scholarship.

References

Adamson LSR, Tasneem N, Andreas MP, Close W, Jenner EN, Szyszka TN, Young R, Cheah LC, Norman A, Mac Dermott-Opeskin HI, O'Mara ML, Sainsbury F, Giessen TW, Lau YH (2022) Pore structure controls stability and molecular flux in engineered protein cages. Sci Adv 8(5): eabl7346. https://doi.org/10.1126/sciadv.abl7346

Akita F et al (2007) The crystal structure of a virus-like particle from the hyperthermophilic archaeon *Pyrococcus furiosus* provides insight into the evolution of viruses. J Mol Biol 368: 1469–1483

Altenburg WJ, Rollins N, Silver PA, Giessen TW (2021) Exploring targeting peptide-shell interactions in encapsulin nanocompartments. Sci Rep 11:1–9

Andreas MP, Giessen TW (2021) Large-scale computational discovery and analysis of virus-derived microbial nanocompartments. Nat Commun 12:4748

Bae Y et al (2018) Engineering tunable dual functional protein cage nanoparticles using bacterial superglue. Biomacromolecules 19:2896–2904

Bamford DH, Grimes JM, Stuart DI (2005) What does structure tell us about virus evolution? Curr Opin Struct Biol 15:655–663

Banaszynski LA, Liu CW, Wandless TJ (2005) Characterization of the FKBP. Rapamycin.FRB ternary complex. J Am Chem Soc 127:4715–4721

Borden JS, Savage DF (2021) New discoveries expand possibilities for carboxysome engineering. Curr Opin Microbiol 61:58–66

Caspar DLD, Klug A (1962) Physical principles in the construction of regular viruses. In: *Cold Spring Harbor symposia on quantitative biology*, vol 27. Cold Spring Harbor Laboratory Press, pp 1–24

Cassidy-Amstutz C et al (2016) Identification of a minimal peptide tag for in vivo and in vitro loading of encapsulin. Biochemistry 55:3461–3468

Choi B et al (2016) Effective delivery of antigen–encapsulin nanoparticle fusions to dendritic cells leads to antigen-specific cytotoxic T cell activation and tumor rejection. ACS Nano 10:7339–7350

Choi H et al (2021) Load and display: engineering encapsulin as a modular nanoplatform for protein-cargo encapsulation and protein-ligand decoration using split intein and SpyTag/SpyCatcher. Biomacromolecules 22:3028–3039

Diaz D, Vidal X, Sunna A, Care A (2021) Bioengineering a light-responsive encapsulin nanoreactor: a potential tool for in vitro photodynamic therapy. ACS Appl Mater Interfaces 13:7977–7986

Giessen TW, Jones JA, Cristie-David AS, Andreas MP (2021) Triggered reversible disassembly of an engineered protein nanocage. Angew Chemie Int Ed. https://doi.org/10.1002/anie.202110318

Giessen TW, Silver PA (2017) Widespread distribution of encapsulin nanocompartments reveals functional diversity. Nat Microbiol 2:1–11

Giessen TW et al (2019) Large protein organelles form a new iron sequestration system with high storage capacity. elife 8:e46070

Hall MP et al (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. ACS Chem Biol 7:1848–1857

He D et al (2016) Structural characterization of encapsulated ferritin provides insight into iron storage in bacterial nanocompartments. elife 5:e18972

Helgstrand C et al (2003) The refined structure of a protein catenane: the HK97 bacteriophage capsid at 3.44 Å resolution. J Mol Biol 334:885–899
Jathoul AP et al (2015) Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter. Nat Photonics 9:239–246
Jenkins MC, Lutz S (2021) Encapsulin nanocontainers as versatile scaffolds for the development of artificial metabolons. ACS Synth Biol 10:857–869
Jones JA, Giessen TW (2021) Advances in encapsulin nanocompartment biology and engineering. Biotechnol Bioeng 118:491–505
Jordan PC et al (2016) Self-assembling biomolecular catalysts for hydrogen production. Nat Chem 8:179–185
Kawamoto S et al (2001) Molecular and functional analyses of the gene (eshA) encoding the 52-kilodalton protein of *Streptomyces coelicolor* A3 (2) required for antibiotic production. J Bacteriol 183:6009–6016
Kelley L-LC et al (2007) Structure of the hypothetical protein PF0899 from *Pyrococcus furiosus* at 1.85 Å resolution. Acta Crystallogr Sect F 63:549
Klem R, De Ruiter MV, Cornelissen JJLM (2018) Protecting encapsulin nanoparticles with cysteine-knot miniproteins. Mol Pharm 15:2991–2996
Kodama Y, Hu C-D (2010) An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. BioTechniques 49:793–805
Künzle M, Mangler J, Lach M, Beck T (2018) Peptide-directed encapsulation of inorganic nanoparticles into protein containers. Nanoscale 10:22917–22926
Kwak J, McCue LA, Trczianka K, Kendrick KE (2001) Identification and characterization of a developmentally regulated protein, EshA, required for sporogenic hyphal branches in *Streptomyces griseus*. J Bacteriol 183:3004–3015
Lagoutte P et al (2018) Simultaneous surface display and cargo loading of encapsulin nanocompartments and their use for rational vaccine design. Vaccine 36:3622–3628
Lau YH, Giessen TW, Altenburg WJ, Silver PA (2018) Prokaryotic nanocompartments form synthetic organelles in a eukaryote. Nat Commun 9:1–7
Lee S-Y, Lee J-H, Chang J-H, Lee J-H (2011) Inorganic nanomaterial-based biocatalysts. BMB Rep 44:77–86
Lee T, Carpenter TS, D'haeseleer P, Savage DF, Yung MC (2020) Encapsulin carrier proteins for enhanced expression of antimicrobial peptides. Biotechnol Bioeng 117:603–613
Liu A, Yang L, Traulsen C-H, Cornelissen J (2017) Immobilization of catalytic virus-like particles in a flow reactor. Chem Commun 53:7632–7634
Lončar N, Rozeboom HJ, Franken LE, Stuart MCA, Fraaije MW (2020) Structure of a robust bacterial protein cage and its application as a versatile biocatalytic platform through enzyme encapsulation. Biochem Biophys Res Commun 529:548–553
London N, Raveh B, Cohen E, Fathi G, Schueler-Furman O (2011) Rosetta FlexPepDock web server—high resolution modeling of peptide–protein interactions. Nucleic Acids Res 39:W249–W253
Maynard-Smith LA, Chen L, Banaszynski LA, Ooi AGL, Wandless TJ (2007) A directed approach for engineering conditional protein stability using biologically silent small molecules. J Biol Chem 282:24866–24872
McHugh CA et al (2014) A virus capsid-like nanocompartment that stores iron and protects bacteria from oxidative stress. EMBO J 33:1896–1911
Mironova KE et al (2013) Genetically encoded immunophotosensitizer 4D5scFv-miniSOG is a highly selective agent for targeted photokilling of tumor cells in vitro. Theranostics 3:831
Moon H, Bae Y, Kim H, Kang S (2016) Plug-and-playable fluorescent cell imaging modular toolkits using the bacterial superglue, SpyTag/SpyCatcher. *Chem Commun*52:14051–14054
Moon H, Lee J, Min J, Kang S (2014a) Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform. Biomacromolecules 15:3794–3801
Moon H et al (2014b) Genetically engineering encapsulin protein cage nanoparticle as a SCC-7 cell targeting optical nanoprobe. Biomater Res 18:1–7

Nichols RJ, Cassidy-Amstutz C, Chaijarasphong T, Savage DF (2017) Encapsulins: molecular biology of the shell. Crit Rev Biochem Mol Biol 52:583–594

Nichols RJ et al (2021) Discovery and characterization of a novel family of prokaryotic nanocompartments involved in sulfur metabolism. elife 10:e59288

Putri RM, Fredy JW, Cornelissen JJLM, Koay MST, Katsonis N (2016) Labelling bacterial nanocages with photo-switchable fluorophores. ChemPhysChem 17:1815–1818

Putri RM et al (2017) Structural characterization of native and modified encapsulins as nanoplatforms for in vitro catalysis and cellular uptake. ACS Nano 11:12796–12804

Rahmanpour R, Bugg TDH (2013) Assembly in vitro of *Rhodococcus jostii* RHA 1 encapsulin and peroxidase DypB to form a nanocompartment. FEBS J 280:2097–2104

Raveh B, London N, Zimmerman L, Schueler-Furman O (2011) Rosetta FlexPepDock ab-initio: simultaneous folding, docking and refinement of peptides onto their receptors. PLoS One 6: e18934

Rurup WF, Cornelissen JJLM, Koay MST (2015) Recombinant expression and purification of "virus-like" bacterial encapsulin protein cages. In: *Protein Cages*. Springer, pp 61–67

Rurup WF, Snijder J, Koay MST, Heck AJR, Cornelissen JJLM (2014) Self-sorting of foreign proteins in a bacterial nanocompartment. J Am Chem Soc 136:3828–3832

Saito N, Matsubara K, Watanabe M, Kato F, Ochi K (2003) Genetic and biochemical characterization of EshA, a protein that forms large multimers and affects developmental processes in *Streptomyces griseus*. J Biol Chem 278:5902–5911

Salvachua D, Prieto A, Martinez AT, Martinez MJ (2013) Characterization of a novel dye-decolorizing peroxidase (DyP)-type enzyme from Irpex lacteus and its application in enzymatic hydrolysis of wheat straw. Appl Environ Microbiol 79:4316–4324

Sassolas A, Blum LJ, Leca-Bouvier BD (2012) Immobilization strategies to develop enzymatic biosensors. Biotechnol Adv 30:489–511

Shah NH, Muir TW (2014) Inteins: nature's gift to protein chemists. Chem Sci 5:446–461

Sheldon RA, Basso A, Brady D (2021) New frontiers in enzyme immobilisation: robust biocatalysts for a circular bio-based economy. Chem Soc Rev *50*(10):5850–5862

Sheldon RA, van Pelt S (2013) Enzyme immobilisation in biocatalysis: why, what and how. Chem Soc Rev 42:6223–6235

Sigmund F et al (2018) Bacterial encapsulins as orthogonal compartments for mammalian cell engineering. Nat Commun 9:1–14

Sonotaki S et al (2017) Successful PEGylation of hollow encapsulin nanoparticles from *Rhodococcus erythropolis* N771 without affecting their disassembly and reassembly properties. Biomater Sci 5:1082–1089

Stritzker J et al (2013) Vaccinia virus-mediated melanin production allows MR and optoacoustic deep tissue imaging and laser-induced thermotherapy of cancer. Proc Natl Acad Sci 110:3316–3320

Sutter M et al (2008) Structural basis of enzyme encapsulation into a bacterial nanocompartment. Nat Struct Mol Biol 15:939–947

Szyszka TN, Jenner EN, Tasneem N, Lau YH (2021) Molecular display on protein nanocompartments: design strategies and systems applications. Chem Syst Chem 4:2

Tamura A et al (2015) Packaging guest proteins into the encapsulin nanocompartment from *Rhodococcus erythropolis* N771. Biotechnol Bioeng 112:13–20

Triccas JA et al (1996) A 35-kilodalton protein is a major target of the human immune response to *mycobacterium leprae*. Infect Immun 64:5171–5177

Valdés-Stauber N, Scherer S (1994) Isolation and characterization of Linocin M18, a bacteriocin produced by *Brevibacterium linens*. Appl Environ Microbiol 60:3809–3814

Wikoff WR et al (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. Science 289:2129–2133

Williams EM, Jung SM, Coffman JL, Lutz S (2018) Pore engineering for enhanced mass transport in encapsulin nanocompartments. ACS Synth Biol 7:2514–2517

Winter N et al (1995) Characterization of the gene encoding the immunodominant 35 kDa protein of *mycobacterium leprae*. Mol Microbiol 16:865–876

Zakeri B et al (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc Natl Acad Sci 109:E690–E697

Lumazine Synthase Nanocompartments

Lukasz Koziej, Agnieszka Gawin, and Yusuke Azuma

Contents

Abstract Many naturally existing proteins possess a unique self-assembling property that enables them to form nanoscale cage-like structures capable of accommodating specific guest molecules in the lumenal space. The examples include coat proteins of viruses, ferritin, chaperones, and shells of bacterial nano-/microcompartments. Efforts have been directed towards the characterization and redesign of such native proteinaceous containers for the development of advanced biomaterials serving as reaction chambers, drug delivery systems, and modular building blocks for nanotechnology. The increasing interest in protein cages requires effective methods for their production, an area in which microorganisms have become a pliable and reliable workhorse. In this chapter, we describe strategies employed for expression, purification, quality control, and reengineering of lumazine synthase, a cage-forming enzyme distinguished by structural plasticity and robustness. Although initially extracted from the source organisms, lumazine synthases can now be produced recombinantly in a high-yield and cost-effective manner. This key advancement allowed not only detailed characterization of the cage assembly but also genetic, post-translational, and chemical modifications yielding useful nanocompartments.

L. Koziej · A. Gawin · Y. Azuma (✉)
Małopolska Centre of Biotechnology, Jagiellonian University, Krakow, Poland
e-mail: yusuke.azuma@uj.edu.pl

B. H. A. Rehm, D. Wibowo (eds.), *Microbial Production of High-Value Products*,
Microbiology Monographs 37, https://doi.org/10.1007/978-3-031-06600-9_13

1 Introduction

Protein cages are hollow three-dimensional compartments that are formed by multiple copies of self-assembling proteins. Their size ranges from a few nanometers to the submicrometer scale (Aumiller et al. 2018). Spherical structures are the most common, but other shapes such as ovoids and extended rods are also observed. Protein cages are ubiquitous across all domains of life from viruses and microorganisms to higher eukaryotes. They are responsible for a wide variety of tasks including mineral storage (Arosio et al. 2017), proteome homeostasis (Coux et al. 1996; Hayer-Hartl et al. 2016), nucleic acid protection and delivery (Buzón et al. 2020), or metabolic reaction sequestration (Heinhorst and Cannon 2020). The broad morphologies and functionalities demonstrated by naturally occurring protein cages have inspired researchers to utilize them as delivery/display vehicles (Choi et al. 2018; Nguyen and Tolia 2021), nanochambers for chemical reactions (Chakraborti et al. 2020), or building blocks to construct higher-order materials (Majsterkiewicz et al. 2020). With such potential applications in medicine and nanotechnology, protein cages have become lucrative bioproducts.

Protein cages have remarkable characteristics as platforms for biotechnology development. They spontaneously assemble from constituent building blocks into monodisperse cage-like structures with well-defined sizes and shapes. Their proteinaceous nature renders them biocompatible and biodegradable, particularly preferable in the context of biomedical applications (Morozova et al. 2020). Most importantly, these structures, particularly ones formed by one or a few distinct proteins, can be easily synthesized in a recombinant form. Practical expression systems such as *Escherichia coli* enable safe and scalable bioproduction as well as further modifications using genetic methods. These features facilitate the reengineering of protein cages to provide tailored nanocompartments with non-natural functions.

Lumazine synthase (LS), 5-amino-6-($_D$-ribitylamino)uracil butanedione transferase (EC 2.5.1.78), is an enzyme that catalyzes the penultimate reaction step of the riboflavin (vitamin B2) biosynthesis (Fig. 1a). While fungi, archaea, and some eubacteria have either pentameric or decameric LS, the variants derived from several organisms such as *Bacillus subtilis* and *Aquifex aeolicus* are known to form dodecahedral assemblies composed of 60 identical monomers (Fig. 1b, c, d). These cages encapsulate the next enzyme in the metabolic pathway, homotrimeric riboflavin synthase (RS), to enhance the overall reaction rate (Kis and Bacher 1995; Azuma et al. 2017). Initially, the enzymes from several pathogenic organisms were investigated as a potential antimicrobial target since animals do not have the riboflavin biosynthesis pathway (Klinke et al. 2005; Morgunova et al. 2005). Lately, cage-forming LS variants have been reengineered to repurpose them for a wide variety of applications in medicine and biotechnology (Azuma et al. 2018b). The development of effective production methods has laid the groundwork for such characterization and redesign. In this chapter, we overview the biosynthesis of LS using microbial factories, particularly focusing on recent advances in heterologous expression systems to yield complex nanocompartments with customized functions.

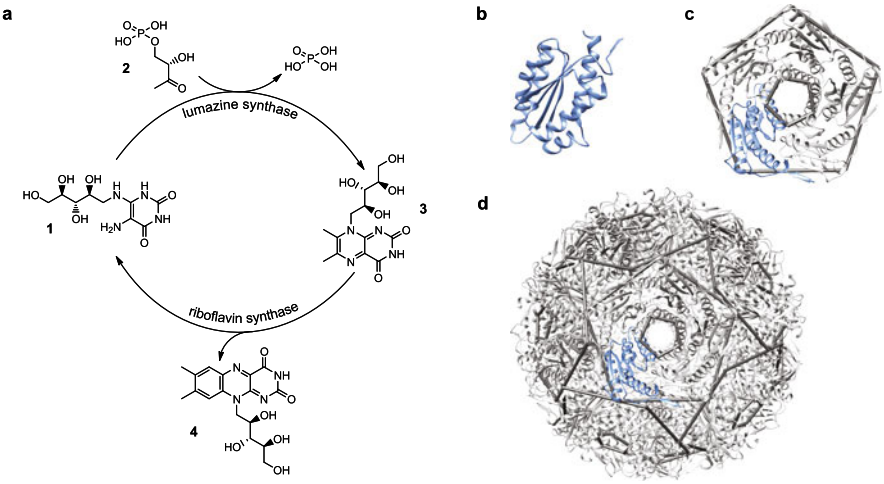


Fig. 1 Catalytic activity and structure of lumazine synthase. (**a**) The last two reaction steps of the riboflavin biosynthesis with 5-amino-6-(D-ribitylamino)uracil (1), L-3,4-dihydroxy-2-butanone 4-phosphate (2), 6,7-dimethyl-8-ribityllumazine (3), and riboflavin (4) depicted. (**b-d**) Ribbon diagrams of (**b**) monomer, (**c**) pentamer, and (**d**) dodecahedral cage formed by *Aquifex aeolicus* lumazine synthase (PDB ID, 5MPP). Pentameric units are indicated as pentagons. The structures are not to scale

2 Production of Lumazine Synthase Orthologs

Prospects for utilization of microbially produced riboflavin as a dietary supplement ignited interest in its biosynthetic enzymes (Bresler et al. 1972; Liu et al. 2020). *Bacillus subtilis* is often chosen for industrial production of high-value organic compounds due to its convenient characteristics, such as adaptation of cellular metabolism to low-cost growth conditions (Su et al. 2020). However, a tightly regulated genetic repression system limits the riboflavin production efficiency of wild-type *B. subtilis* (Abbas and Sibirny 2011). To remove the regulatory elements, the cells were exposed to a chemical mutagen methylnitronitroso-guanidine, and the clones overproducing fluorescent riboflavin were selected under UV light from agar plates (Bresler et al. 1972; Bacher et al. 1973). Among prospective candidates, a derepressed mutant strain H 94 demonstrated the highest concentration of intracellular LS and RS biosynthetic enzymes (Bacher and Mailaender 1978).

The ground-breaking production of the LS/RS complex commenced using flavinogenic *B. subtilis* strain H 94 (Bacher et al. 1980). A large scale microbial fermentation (500-litres batch) using Spizizen's minimal medium (Spizizen 1958) provided a sufficient amount of biomaterial for the extraction and enrichment of the protein products. After cell lysis by sonication, the enzymes were isolated using heat treatment (40 °C), ammonium sulphate fractionation, sedimentation by ultracentrifugation, and two rounds of sucrose gradient centrifugations. The enzymatic activity

assay and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) confirmed the high purity of the microbially produced LS/RS complex with a yield of 5 mg/500-liter batch.

Improvement in the protein extraction and purification procedure increased the production yield of LS and RS enzymes from *B. subtilis* (BsRS/BsLS) (Schott et al. 1990). The sonication process was substituted with more effective enzymatic lysis demanding supplementation of lysozyme, DNase, RNase, and protamine sulphate. Furthermore, ion-exchange chromatography (IEC) and size exclusion chromatography (SEC) techniques replaced ultracentrifugation. The final product was obtained by preparative batch crystallization. Although the quality was comparable with the previous attempts (Bacher et al. 1980), the new procedure increased the yield to 135 mg/500-liter batch (Schott et al. 1990).

Success in microbial production and isolation of the LS/RS complex paved the way for its structural characterization using X-ray crystallography (Ritsert et al. 1995). To elucidate the reaction mechanism, BsRS/BsLS was co-crystallized with substrate analogue inhibitors. Subsequent X-ray diffraction experiments led to the first high-resolution (2.4 Å) structure of LS, a cage-like hollow assembly composed of 12 pentameric subunits. The structure of co-purified RS was not determined, probably due to the crystallographic disorder of the guest enzyme within the host cages. Such cage-formation is, however, not a common feature of LS enzymes as a concurrent study indicated that the functional ortholog extracted from *Saccharomyces cerevisiae* (ScLS) lacks a higher-order structure and remains as a single pentamer unit (Garcia-Ramirez et al. 1995).

Despite successful production of BsLS and ScLS, extraction of enzymes from natural sources is often troublesome. Common issues include low expression levels of endogenous proteins and difficulty in culturing organisms of interest in the laboratory. To overcome these challenges, the LS genes from several organisms have been subcloned into inducible plasmids and successfully expressed in *E. coli* cells. Such heterologous expression system enables safe and scalable production of LSs with distant origins, including eukaryotic (Schneider et al. 2008), pathogenic (Klinke et al. 2005), or thermophilic (Zhang et al. 2001).

The first heterologous expression of LS was conducted with the ScLS variant using *E. coli* strain XL1-Blue (Mörtl et al. 1996). In this study, the cells were transformed with the *B. subtilis* shuttle vector p602 which harbors a constitutively expressed *lac* repressor gene. Furthermore, the plasmid encodes the ScLS (*RIB4*) controlled by *lac* operator and bacteriophage T5-derived promoter (P_{T5}/*lacO*) (Peschke et al. 1985), which guides inherent *E. coli* RNA polymerase (RNAP) (Gentz and Bujard 1985). *RIB4* gene expression was induced with isopropyl β-$_D$-1-thiogalactopyranoside (IPTG), and the cells were processed by enzymatic lysis. Subsequent protein purification using hydrophobic interaction chromatography (HIC) and hydroxyapatite chromatography yielded 50 mg of the enzyme from 1 liter of lysogeny broth (LB) culture (Bacher et al. 1997). A high-resolution (1.9 Å) structure of the recombinant ScLS was determined by X-ray crystallography confirming that the variant exists as a pentamer (Meining et al. 2000).

Heterologous expression in non-pathogenic *E. coli* cells allows safe and easy laboratory production of the LSs from infectious microbes. X-ray crystallography studies revealed that these variants exhibit diverse quaternary arrangements: while the LS from *Mycobacterium tuberculosis* forms a single pentamer (Morgunova et al. 2005), the orthologs from *Bacillus anthracis* (Morgunova et al. 2010) and *Salmonella typhimurium* (Kumar et al. 2011) were found to assemble into dodecahedral cages, like the BsLS variant. *Brucella abortus* has two LS isoforms (Bonomi et al. 2010). RibH1 exists as a single pentamer, and the RibH2 paralogue is composed of two stacked pentamers giving a decameric structure (Klinke et al. 2005, 2007). Owing to the highly ordered homoassembly that can be produced in a scalable manner, the *Brucella* LS isoforms have been used for vaccine development to protect against brucellosis disease, or as a customizable platform allowing multivalent presentation of antigens aimed to prevent other pathologies (Wei et al. 2018).

There is an urgent requirement for new therapeutics amid the emergence of antibiotic-resistant microbes. Since biosynthesis of riboflavin does not take place in the human metabolic pathways, this window of opportunity set in motion structure-based exploration of LS orthologs from pathogenic bacteria (Long et al. 2010). Following co-crystallization with rationally designed compounds that mimic natural ligands, X-ray crystallography demonstrated direct inhibition of LSs from infectious organisms causing tuberculosis (Morgunova et al. 2005), anthrax (Morgunova et al. 2010), salmonellosis (Kumar et al. 2011), or brucellosis (Klinke et al. 2005). These structure-based pursuits were aided by the high-throughput screening, virtual screening, and molecular dynamics (Kundu et al. 2019). The search for the "magic bullet" drug targeting the riboflavin biosynthesis pathway is ongoing.

3 Production of *Aquifex aeolicus* Lumazine Synthase

Aquifex aeolicus LS (AaLS) presents industrially attractive features (Guiral et al. 2012; Sharma et al. 2019). Favorable characteristics include catalytic activity over a wide temperature range reaching at least 90 °C (Haase et al. 2003) and structural stability of the enzyme extending to 120 °C (Zhang et al. 2001). Furthermore, due to tolerance to modification and morphological plasticity, the AaLS has emerged as a customizable nanoscaffold with diverse applications (Azuma et al. 2018b). Here we describe the founding protein production developments that spearheaded the nanotechnological enhancements of AaLS cages.

Production of AaLS using the natural source is difficult to conduct in the laboratory. High-temperature (85 °C) and low-oxygen (microaerophilic) conditions are required for optimal growth of *A. aeolicus* (Guiral et al. 2012). Consequently, laboratory cultivation demands a specialized setting that delivers a pressured gaseous mixture at 80 °C but is limited to biomass yields of about 0.8 optical density units (Uzarraga et al. 2011). A potentially low inherent abundance in the natural source could further limit the yield of extracted AaLS proteins. The manipulation of

A. aeolicus DNA became feasible when the complete genome sequence was revealed (Deckert et al. 1998), a milestone achievement, being the first for a eubacterial hyperthermophile.

The complete genomic sequence of *A. aeolicus* enabled heterologous transfer of AaLS production into *E. coli* expression systems (Deckert et al. 1998). The AA-*ribH* gene encoding AaLS was optimized to match the *E. coli* codon usage, and it was cloned into pNCO113 vector under the control of P_{T5}/*lacO* (Zhang et al. 2001). The isolation of AaLS commenced in *E. coli* M15 cells harboring an additional plasmid, pREP4, which conferred constitutively expressed *lac* repressor to ensure tight transcriptional control of AA-*ribH* until derepressive induction with IPTG. The purification procedure consisted of cell lysis by ultrasonication, heat-induced denaturation of inherent *E. coli* proteins (85 °C), size exclusion chromatography (SEC), and ultracentrifugation. These noteworthy efforts yielded pure AaLS protein in a recombinant form.

As an alternative to heat-assisted isolation, AaLS can be produced as a His-tagged protein and isolated using immobilized metal affinity chromatography (IMAC) (Azuma et al. 2018d). High-level protein expression was achieved using the bacteriophage T7 derived promoter and *E. coli* BL21-(DE3) strain and its derivatives that carry the chromosomal T7 RNAP gene under control of the lacUV5 promoter (P_{lacUV5}). Typically, pMG plasmid possessing the gene for AaLS fused at the C-terminus to hexahistidine tag (6×His) is used for protein production (Fig. 2a). Purification using nickel(II)-loaded nitrilotriacetic acid (Ni-NTA) affinity chromatography gives nearly 100% pure AaLS protein, as judged by SDS-PAGE analysis (Fig. 2c). Further removal of aggregated populations using SEC (Fig. 2b) yields a very regular, spherical cage-like structure with a ~16-nm diameter. In addition to SEC, negative-stain transmission electron microscopy (TEM) and dynamic light scattering (DLS) are routine methods for quality control analysis of AaLS assemblies (Fig. 2d, e), while atomic force microscopy (Heinze et al. 2016) and small-angle X-ray scattering (Zhang et al. 2006) have also been used. Such recombinant production combined with purification using Ni-NTA affinity and SEC typically yields ~20 mg of monodisperse AaLS assemblies from 1 liter of LB culture.

Scalable production using *E. coli* cells provided AaLS protein in a quantity sufficient for crystallization (5.6 mg/mL), which led to a high-resolution (1.6 Å) structure determination of the cage assembly composed of 12 identical pentameric rings (Zhang et al. 2001). The atomic detail shows an ion-pair network formed between the constituent monomers and pentamers, which is not observed in the similar LS cage derived from *B. subtilis* (Ritsert et al. 1995). These amino acid contacts likely contribute to the extreme thermostability of AaLS (Zhang et al. 2001). Differential scanning calorimetry (DSC) showed the melting temperature of AaLS cage at 120 °C, significantly higher than 93 °C for BsLS. The well-defined 3D structure of AaLS, as well as reliable recombinant production in *E. coli*, laid a solid basis for subsequent functional and morphological diversification of this robust nanocage.

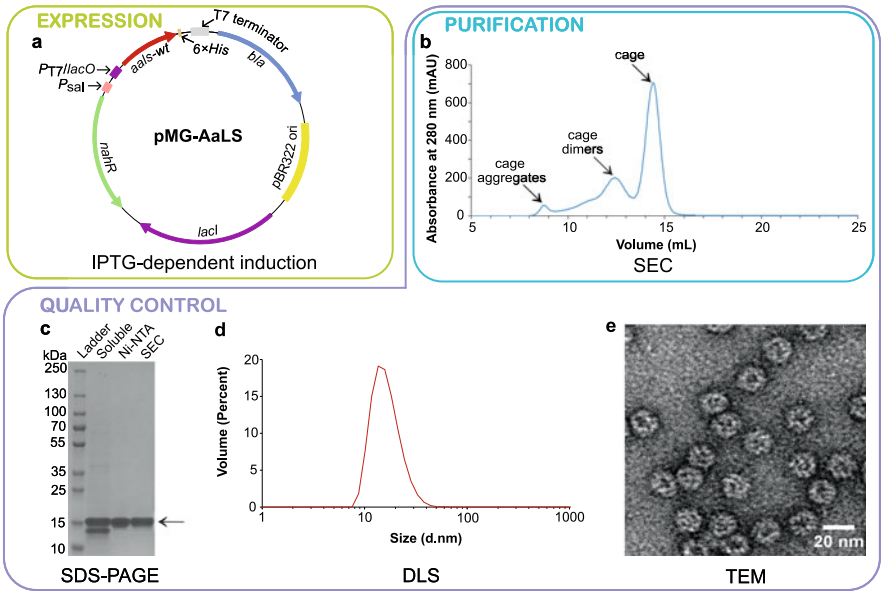


Fig. 2 Production and characterization of wild-type AaLS (AaLS-wt). (**a**) A bacterial expression vector, pMG-AaLS, used for production of His-tagged AaLS-wt. All the abbreviations on the plasmid map are summarized in Table 1. (**b**) Size exclusion chromatogram (SEC) of an AaLS-wt sample purified by Ni-NTA affinity chromatography. Column Superose 6 Increase 10/300; eluent, 50 mM sodium phosphate buffer (pH 8.0) containing 200 mM NaCl and 5 mM ethylenediaminetetraacetic acid (EDTA); flow rate, 0.5 mL/min; room temperature. (**c**) SDS-PAGE analysis depicting the 17.8 kDa AaLS-wt present in purification fractions obtained after cell lysis (soluble), Ni-NTA, and SEC. The bands are visualized by Coomassie staining. (**d**, **e**) Characterization of the SEC-isolated AaLS-wt cages using dynamic light scattering (DLS) (**d**) and negative-stain transmission electron microscopy (TEM) (**e**), suggesting monodisperse cage-assembly with a diameter of ~16 nm. For TEM, the sample was stained with 2% uranyl acetate

4 Production of AaLS-Based Nanocompartments

With the ambition to produce new nanomaterials for technological advancement, the AaLS enzyme has been subjected to extensive protein redesign efforts (Azuma et al. 2018b). Genetic, post-translational, and chemical editing tools have been brought into play to provide customizable packaging and exterior modification with functional moieties. Such modularity of the AaLS scaffold has become evident with a growing number of successfully added capabilities. Although AaLS assembly could give the impression of being structurally rigid and resistant to change, amino acid sequence substitutions and rearrangements, reinforced by directed evolution approaches, have allowed diversification of the morphologies. These achievements broadened the applicability of this enzyme, distinguishing it as one of the most successfully remodelled cage-forming proteins.

Table 1 The major (co-)expression systems for production of AaLS-based nanocompartments

AaLS-based nanocompartment	Gene of interest	Promoter/ operator system(s)	Regulator	Inducer(s)	Ori & selection marker
AaLS-neg with encapsulated GFP-R10 (Seebeck et al. 2006)	*aals-neg*[a] (6×His-tagged)	P_{T7}/*lacO* | P_{sal}[b]	LacI | NahR	IPTG | salicylate[c]	pBR322 | *bla*[d]
	gfp (R10-tagged)	P_{T7}/*lacO* | P_{sal}	LacI | NahR	IPTG | salicylate	p15A | *cat*
AaLS-neg with encapsulated HIVp (Wörsdörfer et al. 2011)	*aals-neg* (6×His-tagged)	P_{T7}/*lacO* | P_{sal}	LacI | NahR	IPTG | salicylate	pBR322 | *bla*
	hiv-p (R10-tagged)	P_{tet}/*tetO*	TetR	tetracycline	p15A | *cat*
AaLS patchwork cage with encapsulated GFP (Azuma et al. 2018d)	*aals-*(*wt*/*neg*/13) (6×His-tagged)	P_{T7}/*lacO* | P_{sal}	LacI | NahR	IPTG | salicylate	pBR322 | *bla*
	cp-aals(119)-*gfp*	P_{tet}/*tetO*	TetR	tetracycline	p15A | *cat*
AaLS-13 with monosaccharide units introduced onto the exterior surface (Tytgat et al. 2019)	*aals-13-GI*/*GIII* (6×His-tagged)	P_{T7}/*lacO*	LacI	IPTG | salicylate	pBR322 | *bla*
	apngt	P_{lacUV5}	-	IPTG	p15A | *cat*

[a]All the genes described in italic encode the proteins as follows; *aals-*(*wt*/*neg*/13), AaLS-wt, AaLS-neg, or AaLS-13, respectively; *gfp*, green fluorescent protein (GFP); *hiv-p*, the protease from human immunodeficiency virus (HIVp); *cp-aals(119)-gfp*, cpAaLS(119) fused to GFP; *aals-13-GI/GIII*, AaLS-13 variants with a monovalent (NAT) or trivalent unit (NATANATANAS) of the glycosylation site, respectively; *apngt*, the glycosyltransferase from *A. pleuropneumoniae* (ApNGT); *bla*, β-lactamase; *cat*, chloramphenicol acetyltransferase

[b]Promoter/operator systems coupled with the regulatory gene are shown as follows; P_{T7}/*lacO,* T7 promoter combined with lactose operator and the LacI repressor (*lacI*); P_{sal}, the salicylate promoter regulated by the NahR transcriptional activator (*nahR*); P_{tet}/*tetO*, the tetracycline promoter combined with tetracycline operator and the TetR repressor (*tetR*); P_{lacUV5}, the lacUV5 promoter

[c]The inducers used in the study are underlined

[d]pBR322 and p15A origins of replication

A general system to encapsulate non-native guest molecules inside AaLS cages was established using a simple electrostatic interaction (Seebeck et al. 2006). Four amino acid residues projecting into the AaLS cage interior (R83, T86, T120, and Q123) were replaced by site-directed mutagenesis into glutamates, yielding a variant called AaLS-neg. The resulting protein cage possessing net negative charge on the interior surface has been shown to package a green fluorescent protein (GFP) equipped at the C-terminus with a deca-arginine tag (R10). An analogous but reverse approach was also employed to design a AaLS-pos construct containing T86R, D90N, T120R, and E122R mutations (Lilavivat et al. 2012). Expression of the

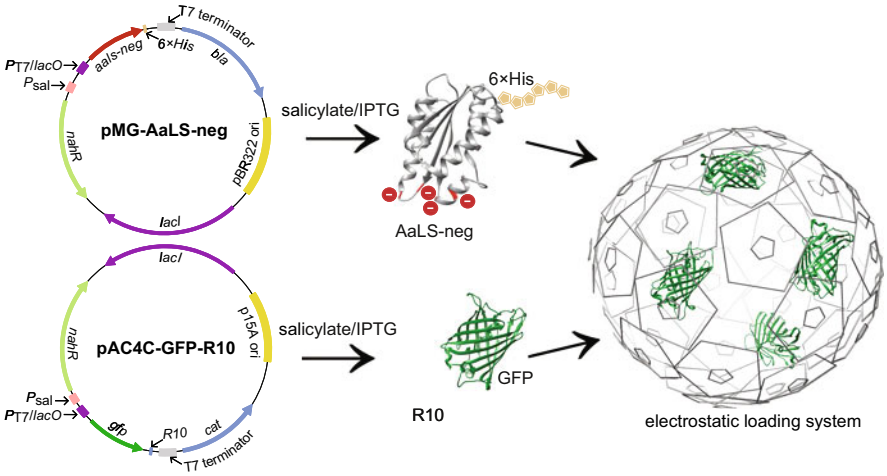


Fig. 3 Charge complementary encapsulation strategy. Both AaLS-neg cage possessing a negatively charged lumen and GFP fused to a deca-arginine R10 tag are co-expressed by induction with salicylate and IPTG. The structures and pentagon diagram of AaLS-neg (PDB ID, 5MQ3) are not to scale. All the abbreviations on the plasmid maps are summarized in Table 1

AaLS-pos variant in *E. coli* cells resulted in length-specific (200–350 bases) packaging of cellular RNA by these protein cages.

A protein co-expression system, as well as an affinity chromatography technique, served as bases to establish protein encapsulation into the redesigned AaLS cages (Seebeck et al. 2006). To test the packaging utility of the above-mentioned charge complementary system, *E. coli* BL21(DE3) cells were co-transformed with the expression vectors, pMG-AaLS-neg and pAC4C-GFP-R10 (Fig. 3 and Table 1). The plasmids had compatible origins of replication, pBR322 and p15A, respectively. AaLS-neg was appended with a 6×His tag at the C-terminus. Salicylate and IPTG induction ensured transcription of the cage and cargo from the dual P_{sal}/P_{T7} promoters included in both plasmids. In an analogous way to the wild-type AaLS variant (AaLS-wt), the His-tagged AaLS-neg protein was isolated using Ni-NTA affinity chromatography. Although AaLS variants exhibit increased resistance to thermal denaturation, heat-assisted extraction could be problematic for the encapsulated guests. To preserve the electrostatic interaction with the cage, salt-alternating methods like ammonium sulphate precipitation or IEC are also not optimal (Zhang et al. 2006; Chen and Woycechowsky 2012). The purification strategy using IMAC allowed one-step isolation of the protein cages while maintaining guest folding as well as host-guest interaction. Following SEC polishing step, the production strategy gave ~20 mg of the AaLS-neg cages containing cargo molecules from 1 liter of LB culture.

Although the same promoter/induction system was used to produce both protein cage and cargo in the proof-of-principle study, later the tetracycline-inducible system was employed for regulating the gene expression level of guest protein

individually. The benefits of tight expression control were demonstrated by the encapsulation of HIV protease into the AaLS-neg cages (Wörsdörfer et al. 2011) (Fig. 4a and Table 1). Cytoplasmic production of the protease is known to inhibit the growth of *E. coli* due to toxicity. The enzyme was genetically fused at the C-terminus to a deca-arginine tag (HIVp-R10) to grant electrostatic complementarity with the negatively charged interior of the AaLS-neg cage. To preserve the viability

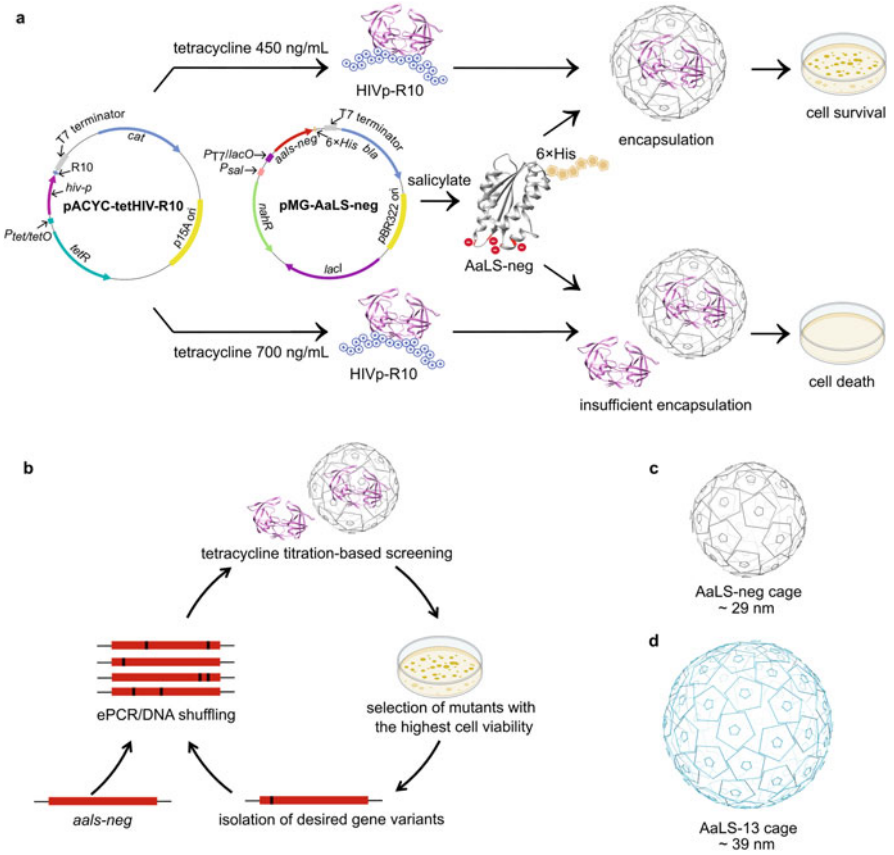


Fig. 4 Encapsulation of cytotoxic HIV protease (HIVp) into the AaLS-neg cages. (**a**) Co-expression of AaLS-neg and HIVp fused to a deca-arginine R10 tag (HIVp-R10) was induced by salicylate and tetracycline, respectively. Upon induction with a relatively low contraction of tetracycline (450 ng/mL), HIVp-R10 is encapsulated by AaLS-neg cages allowing cell growth despite the toxic effect of the protease. At higher tetracycline concentrations (700 ng/mL), cargo encapsulation was presumably insufficient and led to cell death. All abbreviations on the plasmid maps are summarized in Table 1. (**b**) Scheme for directed evolution of AaLS-neg towards efficient encapsulation of HIVp-R10. (**c**, **d**) Diagrams of the AaLS-neg cage (**c**) and the assembly formed by the evolutionally optimized AaLS-13 (**d**). The structure of HIV protease (PDB ID, 1DMP) is not to scale with monomer structure, or pentagon diagrams of AaLS-neg (PDB ID, 5MQ3) and AaLS-13 (PDB ID, 5MQ7)

of the cells, the cargo construct was placed within the context of the tetracycline promoter-operon system (P_{tet}/*tetO*). This gene regulation strategy permits tight control of transcription in absence of tetracycline inducer and gradually adjustable expression following induction (Neuenschwander et al. 2007). *E. coli* XL1-Blue was co-transformed with compatible pMG-AaLS-neg and pACYC-tetHIV-R10. Gene expression was induced with salicylate and tetracycline, recruiting inherent *E. coli* RNAP to P_{sal} or P_{tet} promoter, respectively. Whereas the protease encapsulation by AaLS-neg showed a rescue effect for the *E. coli* host at a relatively low tetracycline concentration (450 ng/mL), cell growth was impeded at a higher concentration (700 ng/mL).

By taking advantage of the genetic programmability and easy protein production associated with the use of a microbial host, the initially designed AaLS-neg was further optimized using directed evolution for more effective guest packaging (Wörsdörfer et al. 2011). The tunable co-expression system of toxic cargo and host cage within *E. coli* cells afforded an ideal selection pressure for such an evolutionary approach. The DNA coding for AaLS-neg was mutagenized by error-prone polymerase chain reaction (ePCR) and DNA shuffling, and a library of mutants was co-expressed with the HIVp-R10 under the growth-inhibiting titer of tetracycline (Fig. 4b). Ultimately, a selective replication scheme yielded a winner variant, called AaLS-13, containing additional seven mutations (D28G, R52C, T112S, V115D, A118D, R127C, and K131E) with increased net negative charge compared to AaLS-neg. The protein cage composed of this variant showed improved packaging efficiency for the HIVp-R10 protein in cells, demonstrating the utility of evolutionary optimization to customize protein cage functions.

The AaLS-13 cages were found to encapsulate a variety of positively charged cargo molecules not only in cells but also in test tubes. Such guests include fluorescent proteins (Wörsdörfer et al. 2012; Zschoche and Hilvert 2015), enzymes (Azuma et al. 2016, 2018a; Frey et al. 2016a, b), metal nanoparticles (Sasaki et al. 2020), and even other protein cages (Beck et al. 2015; Sasaki and Hilvert 2016). In case of positively supercharged variant of GFP, GFP(+36), and its fusion proteins, the guests are rapidly and quantitatively internalized in the empty AaLS-13 cages by simply mixing them in aqueous solution (Zschoche and Hilvert 2015; Azuma et al. 2016). The detailed protocols for production of the AaLS-13 cages and packaging with GFP(+36)-fusion enzymes were described elsewhere (Azuma and Hilvert 2018).

Structure determination by cryo-electron microscopy (cryoEM) single particle reconstruction revealed why the AaLS-13 cage demonstrates efficient cargo loading (Sasaki et al. 2017). While AaLS-wt forms a 16-nm assembly composed of 12 pentamers (Fig. 1d), the negative supercharged AaLS-neg and the evolutionally optimized AaLS-13 variants were found to assemble into the 29-nm and 39-nm spherical cages built from 36 and 72 pentameric subunits, respectively (Fig. 4c, d). The tiling of the expanded AaLS shell became skewed, leading to the formation of 4-nm keyhole-shaped pores, in contrast to the tightly closed AaLS-wt cage. These large openings may give a path for 2.4-nm-wide GFP(+36) to diffuse into the lumen of the preassembled cages. The volume expansion (3.4×) and additional net

negative charges gained through evolutionary optimization could explain the increased packaging density of protein cargo by the AaLS-13 cages, over the parent variant AaLS-neg.

The morphology of AaLS cage was further diversified by topological rearrangement of the secondary structure (Azuma et al. 2018d; Terasaka et al. 2018). Circular permutation is a protein engineering strategy used to change the connectivity of the amino acid sequence while maintaining the 3D shape (Yu and Lutz 2011). In the AaLS-wt protein, the N- and C-termini of each monomer are exposed on the cage exterior. These native termini were genetically connected via short peptidic linkers, and the new N- and C-termini were established within cavity-facing loops. Consequently, two circularly permuted variants were designed, cpAaLS(84) and cpAaLS(119), where the numbers in brackets denote the amino acid sequence position of the new C-termini (Fig. 5a, b). While the cpAaLS(84) variant forms 16-nm native-like structure (Terasaka et al. 2018), the cpAaLS(119) was obtained as a mixture of unassembled capsomers, 24-nm and 28-nm cages, and 24-nm wide rod-shaped structures of variable length (Azuma et al. 2018d). Whereas

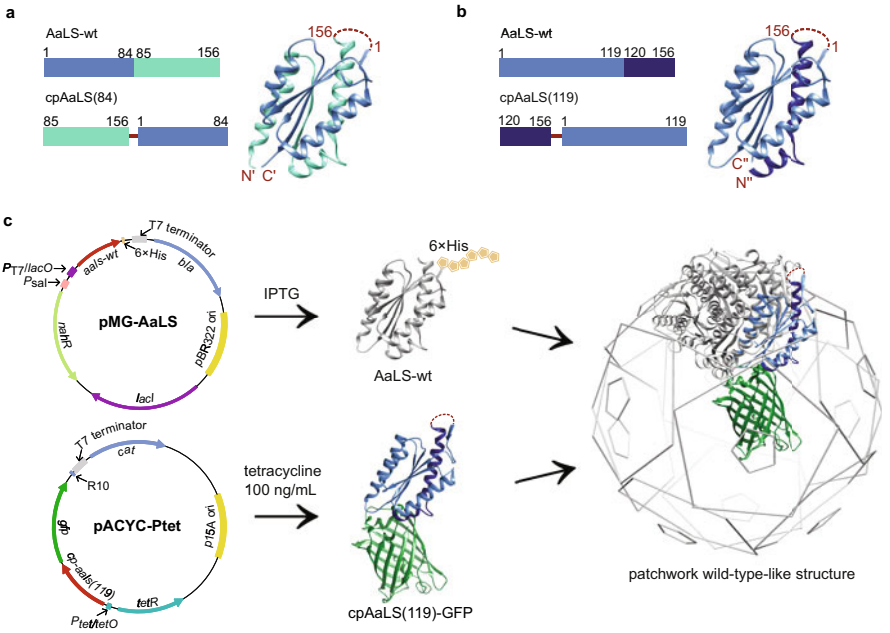


Fig. 5 Circular permutation of AaLS. (**a**, **b**) Design of cpAaLS(84) (**a**) and cpAaLS(119) (**b**), where the numbers in brackets indicate the position of the newly generated C-terminus in the wild-type sequence. The peptidic linkers connecting native N- and C-termini are marked in red. The single quote (‘) and double quote (“) indicate the new termini of cpAaLS(84) and cpAaLS(119), respectively. (**c**) Co-expression of AaLS-wt and cpAaLS(119) fused to GFP at the new C-terminus leading to a formation of patchworked cages encapsulating the guest protein in the lumen. The ribbon representation of AaLS-wt (PDB ID, 5MPP) and GFP fusion (PDB ID, 4KW4) were based on X-ray structures. All the abbreviations on the plasmid maps are summarized in Table 1

it is still unknown how the engineered variant adopts such morphological plasticity, this circular permutation together with protein supercharging strategy demonstrated the potential of AaLS to provide a wide variety of hollow nanostructures that can be easily constructed using the practical microbial system.

The circularly permutated AaLS variants, which possess cavity-facing N- and C-termini, enabled regulated, fusion-based encapsulation of protein cargo (Azuma et al. 2018d). In such an approach, a concern was that bulky guests would sterically hinder the cage assembly process. To address this issue, a patchwork formation strategy has been devised based on the above-mentioned protein co-expression using two compatible plasmids comprising tetracycline and IPTG induction systems. Namely, the gene encoding cpAaLS(119), fused at the C-terminus to a model protein GFP (minor cage component), was subcloned into the low-copy pACYC-Ptet plasmid, harboring p15A replicon and P_{tet}/*tetO* allowing tight regulation of the gene expression at a low level. The DNA sequence coding for C-terminally His-tagged AaLS-wt, AaLS-neg, or AaLS-13 (major cage component) was placed in a high-copy pMG vector, having pBR322 origin of replication and IPTG-inducible P_{T7}/*lacO* (Fig. 5c and Table 1). The *E. coli* BL21-gold(DE3) cells were co-transformed with these plasmids to allow co-expression of both components, while nickel affinity chromatography targeted the major cage component, which had the His-tag. Fluorescence of the SEC fractions indicated co-localization of GFP with the corresponding cages, and negative-staining TEM confirmed that all assemblies incorporated the guest in the lumen. The size of the patchworked cages corresponded exclusively to the major components, suggesting that most abundant AaLS variant dictates the overall morphology of the final assembly. This genetic fusion strategy combined with the protein co-expression system provided a facile microbial production of proteinaceous compartments of which guest density as well as morphology is under control.

In addition to allowing effective protein encapsulation, the circularly permuted AaLS variants have been engineered and evolved into virus-like capsids capable of nucleic acid packaging. The efforts towards the conversion of the bacterial enzyme into nucleocapsids commenced with a genetic fusion of arginine-rich peptides with varying lengths (Rx) to the C-terminus of cpAaLS(119), which acquired spontaneous RNA-binding property inside *E. coli* cells (Azuma et al. 2018c). Co-production of this circularly permuted variant with AaLS-wt resulted in the formation of patchwork structures, where the expression level of cpAaLS(119)-Rx was modulated using the previously mentioned tetracycline system. Such tuning of peptide occupancy, with the additional control in the number of arginines, established a size-selective RNA encapsulation system. In separate developments, the N-terminus of cpAaLS(84) was fused to λN+ peptide (λcpAaLS), which specifically binds to the 19-nucleotide-long RNA called BoxB (Terasaka et al. 2018). The strategy followed by evolutionary optimization yielded a highly selective and protective variant able to pack its own mRNA (Tetter et al. 2021). Such a capsid that allows one-step loading of custom RNA molecules is potentially useful as a delivery vehicle for gene manipulation technology.

In a similar manner to the guest packaging systems, a simple genetic method has been used for exterior decoration of AaLS cages with peptides to yield customized delivery/display vehicles for medical applications (Min et al. 2014). Owing to biodegradability and amenability to both genetic and chemical modifications, protein nanoparticles promise an industrially attractive alternative to metal- or lipid-based carriers (Torres-Pérez et al. 2020). In these appealing systems, conjugation with peptides that recognize cell surface markers potentially presents a powerful approach for directing drugs to specific tissues (Worm et al. 2020). As such, a tumor-targeting peptide, RGD4C, was genetically introduced to the AaLS-wt variant possessing R108C mutation (Min et al. 2014). This site, found on the exterior region (between residues 70 and 71) of the assembly, became available for chemical conjugation. The gene encoding this variant was subcloned into an IPTG-inducible pET30 plasmid and the protein was produced using *E. coli* BL21(DE3) strain, followed by modification via thiol-maleimide chemistry with an anticancer drug aldoxorubicin. The protein retained the ability to form the native-like cage structure upon genetic and chemical engineering and was successfully delivered to the target cancer cells in vitro. In an analogous fashion, another cancer-targeting peptide SP94 (Min et al. 2014), ovalbumin fragments (Ra et al. 2014), and antibody-binding domain (Kim et al. 2016) were successfully displayed on the cage assemblies by genetic fusion to the N- or C-terminus of the AaLS-wt protein. These studies indicated the AaLS cage is a robust and potentially general scaffold capable of arranging various polypeptides in a high density and regular manner.

The spectrum of AaLS cage external surface decorations has been broadened by the utilization of the SpyTag/SpyCatcher, a "bacterial superglue" from *Streptococcus pyogenes* (Zakeri et al. 2012). The short SpyTag peptide (13 amino acids) and cognate partner SpyCatcher protein (12.3 kDa) spontaneously form a covalent isopeptide bond upon reconstitution. To utilize AaLS as a vaccine platform to accomplish multivalent decoration with immunizing antigens (Jardine et al. 2013; Ladenstein and Morgunova 2020; Kang et al. 2021), the SpyTag/SpyCatcher system has been successfully implemented without perturbation of the cage structure (Okba et al. 2020; Zhang et al. 2020). A significant advantage of this approach is that the bioconjugation targets can be expressed independently by specialized hosts. Of note is an effort in which SpyTag-fused AaLS-wt was produced in *E. coli* and the cage was reconstituted with SpyCatcher-fused glycoprotein (35 kDa), which was produced in insect cells (Wichgers Schreur et al. 2021). Following purification, the individual components were coupled in a test tube forming the desired complex. With such diverse possibilities, the modular SpyTag/SpyCatcher system is potentially a universal platform for surface decoration of proteins cages.

A combination of the antibody-binding domain (ABD) display using genetic fusion with the charge complementary guest packaging system leads to a construction of an ideal nanocage for molecular delivery (Levasseur et al. 2021). A well-characterized ABD derived from Staphylococcal protein A, which binds to the Fc portion of IgG (Nilsson et al. 1987), was genetically fused to the C-terminus of AaLS-13. The resulting protein cage can accommodate positively charged guests in the lumen and can display IgG antibodies on the exterior. GFP(+36) and an IgG

against human epidermal growth receptor 2 (HER2), which is highly expressed by human breast and ovarian cancer cell lines, were employed as a model guest and targeting moiety to test such a system. The AaLS-13 protein equipped with the ABD was produced using *E. coli* BL21(DE3) cells via overexpression from the pMG vector possessing an IPTG-inducible P_{T7} system. Following Ni-NTA affinity and SEC purification, isolated cages were mixed with GFP(+36) and the anti-HER2 IgG. As expected, the resulting protein cage complex led to cell-specific uptake and delivery of the encapsulated cargo. However, the study also highlighted that in rich serum, nonspecific coating by other IgGs can mask the efficacy of the targeting antibody, which in turn limits the practical applications of the system. Meanwhile, such camouflaging effect can potentially increase the circulatory half-life of protein cages in bodily fluids, being a useful characteristic for drug carriers.

Besides genetic fusion to peptides and proteins, the AaLS cage is amenable to further decoration with non-proteinaceous substances. This can be achieved by co-expression with post-translational modification enzymes in *E. coli* host (Tytgat et al. 2019). Monosaccharide units were introduced to the surface of the AaLS-13 cage by the asparagine (N)-glucosyltransferase derived from *Actinobacillus pleuropneumoniae* (ApNGT), an enzyme catalyzing the transfer of glucose to asparagine in the N-x-S/T consensus sequence (Schwarz et al. 2011). A monovalent (NAT) or trivalent unit (NATANATANAS) of the target site was genetically fused to the C-terminus of AaLS-13, yielding AaLS-13-GI and AaLS-13-GIII variants, respectively. These variants were further equipped with a 6×His-tag at the C-termini and co-produced with ApNGT in *E. coli* BL21-gold(DE3) cells using compatible pMG and pACYC plasmids (Fig. 6 and Table 1). Mass spectrometry analysis of the resulting assemblies indicated that glycosylation efficiency was nearly 100% for AaLS-13-GI, while on average two hexose units were introduced to the AaLS-13-GIII monomer. Upon the modifications, both variants retain the same morphology as the parent AaLS-13 cage, meaning 360 or 720 hexose units per cage were incorporated by using this simple enzyme co-expression system. Glycoengineering in *E. coli* has been gaining pace as a facile and inexpensive alternative to cellular systems from higher organisms or cell-free setups (Keys et al. 2017; Mueller et al. 2018; Du et al. 2019; Tytgat et al. 2019). The establishment of a microbial cell factory for generating homogenously glycosylated nanocages would extend their applicability in biomedicine.

5 Conclusions and Prospects

Improvements in microbial production methods have facilitated functional and structural investigations of LS orthologs as well as their customization to yield useful nanoparticles. The variant from *A. aeolicus* presents particularly attractive characteristics for technology development, including self-assembly into a monodisperse cage-like structure and amenability to mutagenesis and fusion to other proteins. Recombinant expression using *E. coli* cells enables simple and scalable

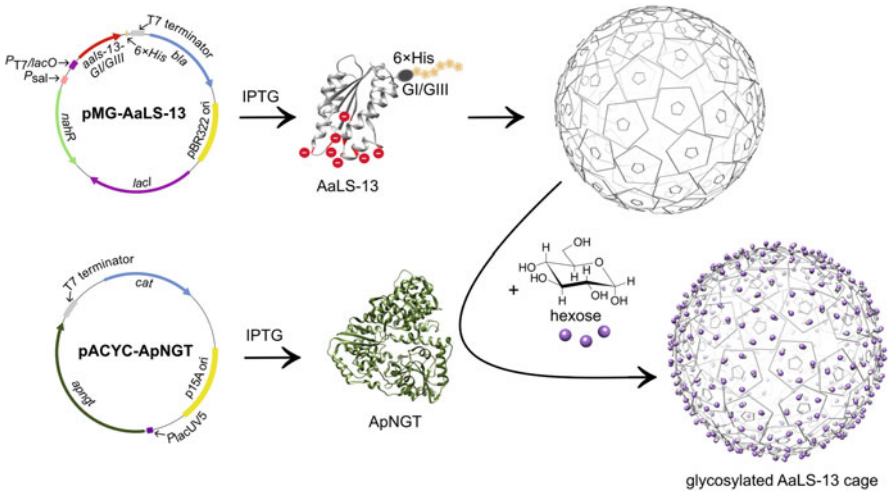


Fig. 6 Post-translational attachment of monosaccharides to the AaLS-13 cage. AaLS-13 with a monovalent (GI) or trivalent (GIII) glycosylation unit introduced at the C-terminus was co-expressed with *Actinobacillus pleuropneumoniae* (N)-glucosyltransferase (ApNGT, PDB ID, 3Q3E) that catalyzes the transfer of hexose onto the cage protein. The structures and pentagon diagram of AaLS-neg (PDB ID, 5MQ7) are not to scale. All the abbreviations on the plasmid maps are summarized in Table 1

production of the enzyme and permits modification by means of genetic manipulation. Additionally, advanced protein co-production systems allow guest packaging and post-translational modification. This range of strategies renders AaLS suitable for new applications beyond the native role of the biosynthetic enzyme.

As demonstrated by LS/RS complexes or bacterial nano/microcompartments, enzyme (co-)localization in a semi-isolated volume is a practical means of optimizing metabolic fluxes in cells (Polka et al. 2016). The protein shells limit diffusive escape of reaction intermediates to prevent undesired side reactions while promoting enzyme cascade. Owing to their efficient protein encapsulation properties, AaLS cages can be filled with a variety of foreign enzymes to mimic the naturally existing bacterial compartments (Azuma and Hilvert 2018). However, the current systems face major challenges in the control of guest stoichiometry and molecular permeability across the shells. These aspects have been suggested to be crucial to exert the expected functions in natural biological compartments, yet detailed molecular mechanisms are still not fully understood (Kis and Bacher 1995; Kerfeld and Melnicki 2016; Wang et al. 2019). Knowledge about the assembly and substrate channelling needs to be expanded further to allow the construction of smart nanoreactors that can endow new metabolic pathways upon cellular factories for future biosynthesis of high-value products (Yocum et al. 2021).

By using genetic fusion and bioconjugation methods, the external surface of the AaLS cage can be decorated with a variety of polypeptides in a high-density fashion. Such a multivalent display nanoplatform is ideal for vaccine development.

Additionally, combined with cargo packaging systems, the protein cages can be harnessed as a delivery carrier for therapeutic or imaging purposes (Ladenstein and Morgunova 2020). However, many concerns remain including encephalitic shock caused by unwanted immunogenicity towards the bacterial protein. Additionally, proteolytic degradation and unspecific cellular uptake could result in a limited circulatory half-life in bodily fluids. Coating of the protein scaffold with masking compounds such as long sugar chains (Kang et al. 2015), polyethylene glycol (PEG) (Suk et al. 2016), or human-derived proteins (Levasseur et al. 2021) is a potential solution to overcome these issues.

For prospective applications in catalysis and medicine, the construction of nanostructures with a defined morphology is currently a research topic of high interest (Stupka and Heddle 2020). The AaLS-based assembly system is an appealing platform in this regard due to the morphological plasticity observed upon reengineering. However, despite the recent cryoEM studies showing the AaLS cage structures (Sasaki et al. 2017; Tetter et al. 2021), how the variants derived from the common parent (Zhang et al. 2001) come together to form a variety of spherical and tubular assemblies is still largely unknown. Understanding the theory underlying the multimeric assembly will provide a blueprint to predict and to construct novel nanoarchitectures in a rational manner.

Acknowledgments We want to thank Dr. Jonathan G. Heddle and Dr. Dmitry Ghilarov (Małopolska Centre of Biotechnology, Jagiellonian University) for critical reading and comments on the manuscript. We are also grateful to Dr. Olga Woźnicka and Dr. Elżbieta Pyza (Institute of Zoology and Biomedical Research, Jagiellonian University) for their help with TEM imaging. This work was supported by National Science Centre (NCN Poland) grants Sonata 14 (2018/31/D/NZ1/01102) and Opus 18 (2019/35/B/NZ1/02044), as well as EMBO Installation Grant.

References

Abbas CA, Sibirny AA (2011) Genetic control of biosynthesis and transport of riboflavin and flavin nucleotides and construction of robust biotechnological producers. Microbiol Mol Biol Rev 75:321–360

Arosio P, Elia L, Poli M (2017) Ferritin, cellular iron storage and regulation. IUBMB Life 69:414–422

Aumiller WM, Uchida M, Douglas T (2018) Protein cage assembly across multiple length scales. Chem Soc Rev 47:3433–3469

Azuma Y, Bader DLV, Hilvert D (2018a) Substrate sorting by a supercharged nanoreactor. J Am Chem Soc 140:860–863

Azuma Y, Edwardson TGW, Hilvert D (2018b) Tailoring lumazine synthase assemblies for bionanotechnology. Chem Soc Rev 47:3543–3557

Azuma Y, Edwardson TGW, Terasaka N, Hilvert D (2018c) Modular protein cages for size-selective RNA packaging in vivo. J Am Chem Soc 140:566–569

Azuma Y, Herger M, Hilvert D (2018d) Diversification of protein cage structure using circularly permuted subunits. J Am Chem Soc 140:558–561

Azuma Y, Hilvert D (2018) Enzyme encapsulation in an engineered lumazine synthase protein cage. In: Protein Scaffolds. Humana Press, New York, NY, Methods in Molecular Biology, pp 39–55

Azuma Y, Zschoche R, Hilvert D (2017) The C-terminal peptide of *Aquifex aeolicus* riboflavin synthase directs encapsulation of native and foreign guests by a cage-forming lumazine synthase. J Biol Chem 292:10321–10327

Azuma Y, Zschoche R, Tinzl M, Hilvert D (2016) Quantitative packaging of active enzymes into a protein cage. Angew Chem Int Ed 55:1531–1534

Bacher A, Baur R, Eggers U et al (1980) Riboflavin synthases of Bacillus subtilis. Purification and properties. J Biol Chem 255:632–637

Bacher A, Eberhardt S, Fischer M et al (1997) Biosynthesis of riboflavin: Lumazine synthase and riboflavin synthase. Methods Enzymol 280:389–399

Bacher A, Eggers U, Lingens F (1973) Genetic control of riboflavin synthetase in *Bacillus subtilis*. Arch Mikrobiol 89:73–77

Bacher A, Mailaender B (1978) Biosynthesis of riboflavin in *Bacillus subtilis*: function and genetic control of the riboflavin synthase complex. J Bacteriol 134:476–482

Beck T, Tetter S, Künzle M, Hilvert D (2015) Construction of Matryoshka-type structures from supercharged protein nanocages. Angew Chem Int Ed 54:937–940

Bonomi HR, Marchesini MIS, Klinke S et al (2010) An atypical riboflavin pathway is essential for *Brucella abortus* virulence. PLoS One 5:e9435

Bresler SE, Glazunov EA, Perumov DA (1972) Study of the operon of riboflavin biosynthesis in *Bacillus subtilis*. Communication IV. Regulation of the synthesis of riboflavin synthetase. Investigation of riboflavin transport through the cell membrane. SovGenet 8:214–222

Buzón P, Maity S, Roos WH (2020) Physical virology: from virus self-assembly to particle mechanics. Wiley Interdiscip Rev Nanomed Nanobiotechnol 12:e1613

Chakraborti S, Lin TY, Glatt S, Heddle JG (2020) Enzyme encapsulation by protein cages. RSC Adv 10:13293–13301

Chen HN, Woycechowsky KJ (2012) Conversion of a dodecahedral protein capsid into pentamers via minimal point mutations. Biochemistry 51:4704–4712

Choi B, Kim H, Choi H, Kang S (2018) Protein cage nanoparticles as delivery nanoplatforms. In: Biomimetic Medical Materials. Advances in Experimental Medicine and Biology, Springer, Singapore, pp 27–43

Coux O, Tanaka K, Goldberg AL (1996) Structure and functions of the 20S and 26S proteasomes. Annu Rev Biochem 65:801–847

Deckert G, Warren PV, Gaasterland T et al (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 392:353–358

Du T, Buenbrazo N, Kell L et al (2019) A bacterial expression platform for production of therapeutic proteins containing human-like O-linked glycans. Cell Chem Biol 26:203–212.e5

Frey R, Hayashi T, Hilvert D (2016a) Enzyme-mediated polymerization inside engineered protein cages. Chem Commun 52:10423–10426

Frey R, Mantri S, Rocca M, Hilvert D (2016b) Bottom-up construction of a primordial carboxysome mimic. J Am Chem Soc 138:10072–10075

Garcia-Ramirez JJ, Santos MA, Revuelta JL (1995) The *Saccharomyces cerevisiae* RIB4 gene codes for 6,7-dimethyl-8-ribityllumazine synthase involved in riboflavin biosynthesis. Molecular characterization of the gene and purification of the encoded protein. J Biol Chem 270:23801–23807

Gentz R, Bujard H (1985) Promoters recognized by *Escherichia coli* RNA polymerase selected by function: highly efficient promoters from bacteriophage T5. J Bacteriol 164:70–77

Guiral M, Prunetti L, Aussignargues C et al (2012) The hyperthermophilic bacterium *Aquifex aeolicus*. From respiratory pathways to extremely resistant enzymes and biotechnological applications. In: Poole RK (ed) Advances in Bacterial Respiratory Physiology. Advances in Microbial Physiology. Academic Press, pp 125–194

Haase I, Fischer M, Bacher A, Schramek N (2003) Temperature-dependent presteady state kinetics of lumazine synthase from the hyperthermophilic eubacterium *Aquifex aeolicus*. J Biol Chem 278:37909–37915

Hayer-Hartl M, Bracher A, Hartl FU (2016) The GroEL-GroES chaperonin machine: a nano-cage for protein folding. Trends Biochem Sci 41:62–76

Heinhorst S, Cannon GC (2020) Bacterial microcompartments. In: Bacterial Organelles and Organelle-like Inclusions. Microbiology Monographs. Springer, Cham, pp 125–147

Heinze K, Sasaki E, King NP et al (2016) Protein nanocontainers from nonviral origin: testing the mechanics of artificial and natural protein cages by AFM. J Phys Chem B 120:5945–5952

Jardine J, Julien JP, Menis S et al (2013) Rational HIV immunogen design to target specific germline B cell receptors. Science 340:711–716

Kang B, Opatz T, Landfester K, Wurm FR (2015) Carbohydrate nanocarriers in biomedical applications: functionalization and construction. Chem Soc Rev 44:8301–8325

Kang YF, Zhang X, Yu XH et al (2021) Immunization with a self-assembled nanoparticle vaccine elicits potent neutralizing antibody responses against EBV infection. Nano Lett 21:2476–2486

Kerfeld CA, Melnicki MR (2016) Assembly, function and evolution of cyanobacterial carboxysomes. Curr Opin Plant Biol 31:66–75

Keys TG, Wetter M, Hang I et al (2017) A biosynthetic route for polysialylating proteins in *Escherichia coli*. Metab Eng 44:293–301

Kim H, Kang YJ, Min J et al (2016) Development of an antibody-binding modular nanoplatform for antibody-guided targeted cell imaging and delivery. RSC Adv 6:19208–19213

Kis K, Bacher A (1995) Substrate channeling in the lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*. J Biol Chem 270:16788–16795

Klinke S, Zylberman V, Bonomi HR et al (2007) Structural and kinetic properties of lumazine synthase isoenzymes in the order *Rhizobiales*. J Mol Biol 373:664–680

Klinke S, Zylberman V, Vega DR et al (2005) Crystallographic studies on decameric *Brucella* spp. lumazine synthase: a novel quaternary arrangement evolved for a new function? J Mol Biol 353: 124–137

Kumar P, Singh M, Karthikeyan S (2011) Crystal structure analysis of icosahedral lumazine synthase from *Salmonella typhimurium*, an antibacterial drug target. Acta Crystallogr Sect D Biol Crystallogr 67:131–139

Kundu B, Sarkar D, Ray N, Talukdar A (2019) Understanding the riboflavin biosynthesis pathway for the development of antimicrobial agents. Med Res Rev 39:1338–1371

Ladenstein R, Morgunova E (2020) Second career of a biosynthetic enzyme: Lumazine synthase as a virus-like nanoparticle in vaccine development. Biotechnol Reports 27:e00494

Levasseur MD, Mantri S, Hayashi T et al (2021) Cell-specific delivery using an engineered protein nanocage. ACS Chem Biol 16:838–843

Lilavivat S, Sardar D, Jana S et al (2012) In vivo encapsulation of nucleic acids using an engineered nonviral protein capsid. J Am Chem Soc 134:13152–13155

Liu S, Hu W, Wang Z, Chen T (2020) Production of riboflavin and related cofactors by biotechnological processes. Microb Cell Factories 19:31

Long Q, Ji L, Wang H, Xie J (2010) Riboflavin biosynthetic and regulatory factors as potential novel anti-infective drug targets: perspective. Chem Biol Drug Des 75:339–347

Majsterkiewicz K, Azuma Y, Heddle JG (2020) Connectability of protein cages Nanoscale Adv 2: 2255–2264

Meining W, Mörtl S, Fischer M et al (2000) The atomic structure of pentameric lumazine synthase from *Saccharomyces cerevisiae* at 1.85 Å resolution reveals the binding mode of a phosphonate intermediate analogue. J Mol Biol 299:181–197

Min J, Kim S, Lee J, Kang S (2014) Lumazine synthase protein cage nanoparticles as modular delivery platforms for targeted drug delivery. RSC Adv 4:48596–48600

Morgunova E, Illarionov B, Saller S et al (2010) Structural study and thermodynamic characterization of inhibitor binding to lumazine synthase from *Bacillus anthracis*. Acta Crystallogr Sect D Biol Crystallogr 66:1001–1011

Morgunova E, Meining W, Illarionov B et al (2005) Crystal structure of lumazine synthase from *Mycobacterium tuberculosis* as a target for rational drug design: binding mode of a new class of purinetrione inhibitors. Biochemistry 44:2746–2758

Morozova OV, Sokolova AI, Pavlova ER et al (2020) Protein nanoparticles: cellular uptake, intracellular distribution, biodegradation and induction of cytokine gene expression. Nanomedicine Nanotechnology, Biol Med 30:102293

Mörtl S, Fischer M, Richter G et al (1996) Biosynthesis of riboflavin: Lumazine synthase of *Escherichia coli*. J Biol Chem 271:33201–33207

Mueller P, Gauttam R, Raab N et al (2018) High level in vivo mucin-type glycosylation in *Escherichia coli*. Microb Cell Factories 17:168

Neuenschwander M, Butz M, Heintz C et al (2007) A simple selection strategy for evolving highly efficient enzymes. Nat Biotechnol 25:1145–1147

Nguyen B, Tolia NH (2021) Protein-based antigen presentation platforms for nanoparticle vaccines. NPJ Vaccines 6:70

Nilsson B, Moks T, Jansson B et al (1987) A synthetic IgG-binding domain based on staphylococcal protein a. Protein Eng Des Sel 1:107–113

Okba NMA, Widjaja I, van Dieren B et al (2020) Particulate multivalent presentation of the receptor binding domain induces protective immune responses against MERS-CoV. Emerg Microbes Infect 9:1080–1091

Peschke U, Beuck V, Bujard H et al (1985) Efficient utilization of *Escherichia coli* transcriptional signals in *Bacillus subtilis*. J Mol Biol 186:547–555

Polka JK, Hays SG, Silver PA (2016) Building spatial synthetic biology with compartments, scaffolds, and communities. Cold Spring Harb Perspect Biol 8:a024018

Ra J-S, Shin H-H, Kang S, Do Y (2014) Lumazine synthase protein cage nanoparticles as antigen delivery nanoplatforms for dendritic cell-based vaccine development. Clin Exp Vaccine Res 3: 227

Ritsert K, Huber R, Turk D et al (1995) Studies on the lumazine synthase/riboflavin synthase complex of *Bacillus subtilis*: crystal structure analysis of reconstituted, icosahedral β-subunit capsids with bound substrate analogue inhibitor at 2.4 Å resolution. J Mol Biol 253:151–167

Sasaki E, Böhringer D, Van De Waterbeemd M et al (2017) Structure and assembly of scalable porous protein cages. Nat Commun 8:14663

Sasaki E, Dragoman RM, Mantri S et al (2020) Self-assembly of proteinaceous shells around positively charged gold nanomaterials enhances colloidal stability in high-ionic-strength buffers. Chembiochem 21:74–79

Sasaki E, Hilvert D (2016) Self-assembly of proteinaceous multishell structures mediated by a supercharged protein. J Phys Chem B 120:6089–6095

Schneider G, Sandalova T, Persson K et al (2008) Crystal structure analysis of a pentameric fungal and an icosahedral plant lumazine synthase reveals the structural basis for differences in assembly. Protein Sci 8:2355–2365

Schott K, Ladenstein R, Konig A, Bacher A (1990) The lumazine synthase-riboflavin synthase complex of *Bacillus subtilis*. Crystallization of reconstituted icosahedral β-subunit capsids. J Biol Chem 265:12686–12689

Schwarz F, Fan YY, Schubert M, Aebi M (2011) Cytoplasmic N-glycosyltransferase of *Actinobacillus pleuropneumoniae* is an inverting enzyme and recognizes the NX(S/T) consensus sequence. J Biol Chem 286:35267–35274

Seebeck FP, Woycechowsky KJ, Zhuang W et al (2006) A simple tagging system for protein encapsulation. J Am Chem Soc 128:4516–4517

Sharma S, Vaid S, Bhat B et al (2019) Thermostable enzymes for industrial biotechnology. In: Biomass. Biofuels, Biochemicals. Advances in Enzyme Technology. Elsevier, pp 469–495

Spizizen J (1958) Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc Natl Acad Sci U S A 44:1072–1078

Stupka I, Heddle JG (2020) Artificial protein cages – inspiration, construction, and observation. Curr Opin Struct Biol 64:66–73

Su Y, Liu C, Fang H, Zhang D (2020) *Bacillus subtilis*: a universal cell factory for industry, agriculture, biomaterials and medicine. Microb Cell Factories 19:173
Suk JS, Xu Q, Kim N et al (2016) PEGylation as a strategy for improving nanoparticle-based drug and gene delivery. Adv Drug Deliv Rev 99:28–51
Terasaka N, Azuma Y, Hilvert D (2018) Laboratory evolution of virus-like nucleocapsids from nonviral protein cages. Proc Natl Acad Sci U S A 115:5432–5437
Tetter S, Terasaka N, Steinauer A et al (2021) Evolution of a virus-like architecture and packaging mechanism in a repurposed bacterial protein. Science 372:1220–1224
Torres-Pérez SA, Torres-Pérez CE, Pedraza-Escalona M et al (2020) Glycosylated nanoparticles for cancer-targeted drug delivery. Front Oncol 10:605037
Tytgat HLP, Lin CW, Levasseur MD et al (2019) Cytoplasmic glycoengineering enables biosynthesis of nanoscale glycoprotein assemblies. Nat Commun 10:5403
Uzarraga R, Auria R, Davidson S et al (2011) New cultural approaches for microaerophilic hyperthermophiles. Curr Microbiol 62:346–350
Wang H, Yan X, Aigner H et al (2019) Rubisco condensate formation by CcmM in β-carboxysome biogenesis. Nature 566:131–135
Wei Y, Kumar P, Wahome N et al (2018) Biomedical applications of lumazine synthase. J Pharm Sci 107:2283–2296
Wichgers Schreur PJ, Tacken M, Gutjahr B et al (2021) Vaccine efficacy of self-assembled multimeric protein scaffold particles displaying the glycoprotein Gn head domain of rift valley fever virus. Vaccine 9:301
Worm DJ, Els-Heindl S, Beck-Sickinger AG (2020) Targeting of peptide-binding receptors on cancer cells with peptide-drug conjugates. Pept Sci 112:e24171
Wörsdörfer B, Pianowski Z, Hilvert D (2012) Efficient in vitro encapsulation of protein cargo by an engineered protein container. J Am Chem Soc 134:909–911
Wörsdörfer B, Woycechowsky KJ, Hilvert D (2011) Directed evolution of a protein container. Science 331:589–592
Yocum HC, Pham A, Da Silva NA (2021) Successful enzyme colocalization strategies in yeast for increased synthesis of non-native products. Front Bioeng Biotechnol 9:606795
Yu Y, Lutz S (2011) Circular permutation: a different way to engineer enzyme structure and function. Trends Biotechnol 29:18–25
Zakeri B, Fierer JO, Celik E et al (2012) Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proc Natl Acad Sci U S A 109:E690–E697
Zhang B, Chao C, Tsybovsky Y et al (2020) A platform incorporating trimeric antigens into self-assembling nanoparticles reveals SARS-CoV-2-spike nanoparticles to elicit substantially higher neutralizing responses than spike alone. Sci Rep 10:18149
Zhang X, Konarev PV, Petoukhov MV et al (2006) Multiple assembly states of lumazine synthase: a model relating catalytic function and molecular assembly. J Mol Biol 362:753–770
Zhang X, Meining W, Fischer M et al (2001) X-ray structure analysis and crystallographic refinement of lumazine synthase from the hyperthermophile *Aquifex aeolicus* at 1.6 Å resolution: determinants of thermostability revealed from structural comparisons. J Mol Biol 306: 1099–1114
Zschoche R, Hilvert D (2015) Diffusion-limited cargo loading of an engineered protein container. J Am Chem Soc 137:16121–16132