Catalyst development for biochemical cascade reactions often follows a “whole-cell-approach” in which a single microbial cell is made to express all required enzyme activities. Although attractive in principle, the approach can encounter limitations when efficient overall flux necessitates precise balancing between activities. This study shows an effective integration of major design strategies from synthetic biology to a coherent development of plasmid vectors, enabling tunable two-enzyme co-expression in E. coli, for whole-cell-production of cellobiose. An efficient transformation of sucrose and glucose into cellobiose by a parallel (countercurrent) cascade of disaccharide phosphorylases requires the enzyme co-expression to cope with large differences in specific activity of cellobiose phosphorylase (14 U mg⁻¹) and sucrose phosphorylase (122 U mg⁻¹). Mono- and bicistronic co-expression strategies controlling transcription, transcription-translation coupling or plasmid replication are analyzed for effect on activity and stable producibility of the whole-cell-catalyst. A key role of bom (basis of mobility) for plasmid stability dependent on the ori is reported and the importance of RBS (ribosome binding site) strength is demonstrated. Whole cell catalysts show high specific rates (460 µmol cellobiose min⁻¹ g⁻¹ dry cells) and performance metrics (30 g L⁻¹; ∼82% yield; 3.8 g L⁻¹ h⁻¹ overall productivity) promising for cellobiose production.

1. Introduction

Enzyme cascade reactions represent an emerging paradigm in catalytic chemistry that adopts the fundamental “biological pathway” principle to organic synthesis. Limitations on applicability can arise, however, from a set of complex challenges that such cascades involve. Overcoming the challenges is essential in order to close a significant gap in number between the enzyme cascades demonstrated as proof of principle and the ones actually considered for production. Core problem for cascade reactions to implement is finding efficient ways to produce the required enzymes. Production of each enzyme individually is resource intensive and not generally viable. Therefore, catalyst development often proceeds according to a “whole cell approach” in which a single microbial cell is engineered to concurrently express the full set of enzymes required for pathway function.

A number of studies on enzyme co-expression show that obtaining the enzymes in high yield and suitable activity ratio represents a main bottleneck. The fundamental relationship between cascade flux efficiency dependent on individual activities and whole-cell reaction output has rarely been explored for process intensification. The net flux from substrate to product depends on the relative volumetric activities of the individual enzymes under prevailing conditions at steady state. The volumetric activity thus obtained from the enzyme specific activity (units mg⁻¹ protein) times the volumetric concentration of the functional protein. Modern synthetic biology offers a diverse set of tools, each applicable in principle to net flux optimization via tuning of the volumetric protein concentrations. However, the degree to which these tools can be effective, individually or in combination, is not generally known nor predictable. This restricts their targeted and time-efficient application to engineer cascade catalysis in whole cells for programmable performance. Enzymes differing widely in specific activity can be challenging for co-expression to balance their individual activities in whole cells. Existing technologies (e.g., pETDuet) are limited in scope to provide precise and independent expression control for each enzyme. Approaches using two plasmids of different copy number can expand the scope for adjusting the enzyme ratios, but are hardly amenable to industrial production due to the requirement for two or more antibiotic selection markers.
Here, we report on a synthetic biology-based engineering approach to two-enzyme co-expression in *Escherichia coli* for whole cell synthesis of cellobiose. The disaccharide cellobiose is an emerging food-and-feed ingredient whose bulk-scale production via a bottom-up biocatalytic synthesis has received considerable attention.[24–26] Cascade reaction of sucrose phosphorylase (SP) and cellobiose phosphorylase (CBP) in the presence of phosphate converts sucrose and glucose into cellobiose.[25,27–29] The overall reaction is a two-step glucosyl transfer via \(\alpha\)-glucose 1-phosphate (\(\alpha\)Glc 1-P), as shown in Figure 1. The SP from *Bifidobacterium adolescentis* (BaSP; 122 U mg\(^{-1}\)) has approximately tenfold higher specific activity than the CBP from *Cellulomonas uda* (CuCBP).[31] Co-expression of the two enzymes at an optimum activity ratio and in high-yield protein amounts is therefore required. Studies of cellobiose synthesis using isolated enzymes employed SP and CBP at activity ratios of 1.2/1[26,28] and 1/3.[27] Avoiding two-plasmid approaches for the above-mentioned reasons, we here focused on synthetic biology tools combinable on a single expression vector. Using mono- or bicistronic vector-based strategies, we explored genetic control elements of transcription, transcription-translation coupling or plasmid replication for effects on activity and stable producibility of the whole cell catalyst. We show that the bom site (basis of mobility), also named oriT (origin of transfer), has a key role for plasmid stability dependent on the origin of replication. We also demonstrate the importance of the RBS (ribosome binding site) strength for balanced bicistronic co-expression. We obtain whole cell catalysts with high specific rates (460 \(\mu\)mol cellobiose min\(^{-1}\) g\(^{-1}\) dry cells) and performance metrics (30 g L\(^{-1}\); \(\approx\)82% yield; 3.8 g L\(^{-1}\) h\(^{-1}\) overall productivity) that are promising for cellobiose production.

### 2. Results and Discussion

Phosphorylase cascade reaction for cellobiose synthesis was originally developed by Kitaoka and co-workers.[28] They, and later other authors,[25,27] used cell-free phosphorylase preparations and additionally employed xylose isomerase for in situ supply of glucose from fructose. We considered fructose as a valuable co-product of the sucrose conversion (https://www.carbafin.eu/) and therefore added glucose to the reaction. The biocatalytic process under development for industrial production of cellobiose[26] uses sucrose and glucose as substrates, as shown in Figure 1.

#### 2.1. Synthetic Biology Tools Enable a Variety of Activity Ratios

Activity balancing was targeted via down- and upregulation of the expression levels for BaSP (122 U mg\(^{-1}\))[30] and CuCBP (14 U mg\(^{-1}\)),[31] respectively, to control the produced protein amounts. We designed two strategies to that end, built upon mono- and bicistronic expression systems. Transcription and transcription-translation coupling were considered as relevant control points in both systems.

The pDUAL vector (Figure S1B, Supporting Information) implemented the monocistronic expression of each gene under the control of an individual promoter (Figure 2A) whose relative strength can be varied. Promoters are oriented in opposite directions, away from each other, thus allowing for transcriptions that are independent one from another. The strong T7\(\text{lacO}\) promoter[32,33] was used for CuCBP expression whereas the promoters for BaSP expression were varied. The weak lacZ promoter[33,34] as well as the relatively stronger tacI,[33,35] araBAD,[32,36] and T5 promoters[37] were used. P\(_{T7\text{lacO}}\)

\(\text{P}_{\text{lacZ}}\), \(\text{P}_{\text{lacI}}\), and \(\text{P}_{\text{T5}}\) are inducible by isopropyl \(\beta\)-D-1-thiogalactopyranoside (IPTG) and regulated by the LacI repressor[38] (Figure 2A). \(\text{P}_{\text{araBAD}}\) instead, must be induced by L-arabinose and is regulated by AraC.[39] These control elements were integrated into the vector backbone, in addition to LacI regulation (Figure S1B, Supporting Information). In total, four different pDUAL constructs were generated (Figure 2A).
Figure 2. Plasmid design strategies for enzyme co-expression. A) Monocistronic co-expression plasmid (pDUAL): Transcription of CuCBP is controlled by the strong $P_{\text{T7}}$ promoter and BaSP by either $P_{\text{lacz}}$, $P_{\text{tac}}$, $P_{\text{T5}}$, or $P_{\text{araBAD}}$. Each promoter is regulated by LacI, except for $P_{\text{araBAD}}$. pDUAL-$P_{\text{araBAD}}$, therefore, additionally bears the gene for the regulatory protein AraC. B) Bicistronic co-expression plasmid (pBICI): Transcription of both genes is controlled by the strong $P_{\text{T7}}$ promoter. CuCBP bears the strong T7-RBS, BaSP translation is controlled by either a weak, medium or strong RBS.

The bicistronic vector pBICI (Figure S1A, Supporting Information) facilitated the transcription of both genes from one operon under $T7_{\text{laci}}$ promoter control. The downstream order of gene placement, that is, BaSP after CuCBP, reflects the idea that in a polycistronic operon, open reading frames closer to the end of the operon are translated at a lower level than open reading frames in the vicinity of the operon start.\(^\text{[40]}\) To modulate the translation efficiency in a transcript-specific fashion, we used combinatorial fine-tuning of the RBS. Affinity of the ribosome to the RBS\(^\text{[41,42]}\) constitutes a promising element of differential expression control. Here, the strong native RBS of the phage $T7$\(^\text{[42]}\) was used for CuCBP expression whereas the RBS for BaSP expression varied from weak (http://parts.igem.org/Part:BBa_B0033), medium (http://parts.igem.org/Part:BBa_B0032) to strong RBS\(^\text{[44]}\). Hence, three bicistronic expression constructs, pBICI_weak, -medium and -strong, were designed (Figure 2B).

2.1.1. Strong RBS Enhances Expression of BaSP by One Order of Magnitude

Expression from pBICI_strong yielded BaSP in a specific activity of 10.8 U mg\(^{-1}\) (10% of total protein), which is about tenfold the specific activity received by expression from vectors harboring the weak and medium RBS. Under conditions of weak and medium RBS, therefore, the BaSP expression appears to have been limited by translation initiation. Co-expression of CuCBP and BaSP from pBICI_weak and pBICI_medium succeeded in offsetting the enzyme-specific differences in activity almost completely. However, since co-expression from pBICI_strong gave the highest activities of both enzymes, the activity ratio of BaSP and CuCBP becomes secondary for vector selection, and pBICI_strong is the best choice for cellulose production (see “Optimized cell-factory for future industrial cellulose production”).

The general concept of tuning enzyme activity ratios via molecular RBS engineering is strongly supported from these results. To enable more finely grained and better programmable adjustments of the activities, RBS sequences should be further explored for structure-function relationships in enzyme (co)-expression. Bacterial transcription and translation are closely linked molecular events, whereby translation initiation is the rate-limiting step of protein biosynthesis.\(^\text{[45]}\) Ribosome recruitment to the mRNA and translation initiation are highly regulated by the sequence of the RBS.\(^\text{[45]}\) Mathematical models have been used to engineer and optimize RBSs.\(^\text{[42,46–51]}\) The design of tailor-made RBSs for translation optimization of specific mRNAs considers the mRNA folding in the cell, which is determined by the sequence of the transcribed gene.\(^\text{[42,46]}\) We made use of the RBS calculator devised by Salis in 2011,\(^\text{[42]}\) whose thermodynamic model has been constantly updated.\(^\text{[49,50,52]}\) We used the reverse engineering mode of version 2.1 on the webpage http://salislab.net/software/.

Expression from pBICI_weak, -medium and -strong yielded CuCBP in specific activities of 1.0, 1.2, and 1.4 U mg\(^{-1}\) that correspond to protein amounts of 12%, 15%, and 17%, respectively, of the total protein in the E. coli cell extract. For cellulose production according to Figure 1, the relevant activity of CuCBP is the one in synthesis direction of reaction. For convenience, the activity assay is however performed in the phosphorylation direction. We determined that the synthesis exceeds the phosphorylation activity by a factor of 1.8 ($\pm$ 0.1; $N = 6$), consistent with literature.\(^\text{[31]}\)
The translation initiation rate for the start codon of the BaSP gene was predicted as 2.34 au, 2.08 au and 233.46 au for the weak, medium and strong RBS, respectively. Experimental data (Table 1) show a similar trend: There is almost no expression difference between pBICI_weak and pBICI_medium, whereas pBICI_strong shows a tenfold improvement. Nevertheless, the RBS calculator predicted an expression increase of two orders of magnitude for the strong RBS. This difference between the experimental data and the prediction can partly be explained by two facts: First, we assessed the soluble and active BaSP in our assay. The total translated amount of protein is slightly higher due to the formation of inclusion bodies (∼25% of soluble protein, see Figure S2A, Supporting Information), which are not taken into account by our assay. Second, the RBS calculator model v2.1 stated a warning about a potential prediction failure for our specific mRNAs of all pBICI constructs due to their putative slow folding in vivo. Borujeni and Salis[50] concluded that certain RNA structures may not have enough time to fold inside the cell, creating a non-equilibrium effect. In such cases, the thermodynamic modeling framework fails to deliver accurate predictions due to their proposed "ribosome drafting mechanism."[50]

2.1.2. Expression Instabilities and Transformation Difficulties

To obtain biological replicates of the results, a series of co-expression experiments was performed in which cells from the same glycerol stock were cultivated consecutively. Strikingly, already in proceeding from the first to the second replication, we observed a marked decrease by 52–70% in all activities, and throughout all pBICI constructs, which was accompanied by a corresponding decrease in enzyme expression (Table 1). Figure S2, Supporting Information, shows the corresponding SDS polyacrylamide gels of the soluble and insoluble protein fractions of two consecutive cultivations of all four co-expression plasmids. The second cultivation did not show an increase in inclusion body formation (Figure S2B, Supporting Information). The third round featured a complete loss of activity and enzyme expression. Subsequently, several freshly transformed colonies were tested, but none of them reached the initial expression again and lost all of the enzyme activity after three cultivations at latest.

In addition, we experienced severe difficulties in introducing the pDUAL vectors into E. coli hosts. All attempts to transform the pDUAL_P_plac, P_T5 and P_WorAD vectors into E. coli BL21(DE3)agg proved not fruitful. The same vectors could be amplified in E. coli Top10F, albeit with considerably lower efficiency. Of the four constructs, pDUAL_P_T5 harboring the strongest promoter for BaSP expression in addition to the very strong P_T7lacO-CuCBP expression cassette showed the lowest transformation efficiency. Metabolic burden associated with plasmid replication and protein production is arguably at the limit of E. coli BL21(DE3)aggT7 cell viability under the employed conditions.[51] In comparison, recombinant protein production is greatly reduced in E. coli Top10F which can explain the difference between the amplification and the expression strain. Lack of the T7 RNA polymerase gene renders the E. coli Top10F incapable of transcribing the CuCBP gene. Transcription of the BaSP gene, by contrast, would be possible from the pDUAL constructs, based on the protein machinery encoded in the E. coli Top10F genome.[54]

Table 1. Stability of enzyme production over multiple rounds of co-expression.

| Expression round | pBICI weak RBS | pBICI medium RBS | pBICI strong RBS | pDUAL P\textsubscript{lacZ} |
|------------------|----------------|-----------------|-----------------|-------------------|
| CuCBP [U mg\textsuperscript{-1}] | 0.97 ± 0.09 | 1.22 ± 0.08 | 1.35 ± 0.14 | 1.86 ± 0.19 |
| 1st              |                |                 |                 |                   |
| 2nd              | 0.40 ± 0.01    | 0.31 ± 0.01     | 0.48 ± 0.05     | 0.38 ± 0.02       |
| BaSP [U mg\textsuperscript{-1}] | 1.15 ± 0.09 | 1.16 ± 0.02 | 10.8 ± 1.5 | n.d. |
| 1st              |                |                 |                 |                   |
| 2nd              | 0.41 ± 0.04    | 0.36 ± 0.03     | 5.19 ± 0.13     | n.d. |

*U mg\textsuperscript{-1} soluble E. coli protein. Activities were measured in phosphorolysis direction and standard deviation was calculated from three technical replicates. n.d. not detectable.

The extremely low transformation efficiency of E. coli Top10F with the pDUAL_P\textsubscript{T5} construct, and the absence of viable transformants of E. coli BL21(DE3)aggT7, suggest that the E. coli strains approached, or even exceeded, their uppermost resource limits when plasmid replication and BaSP expression occurred at the same time. Interestingly, the pDUAL_P\textsubscript{lacZ} vector was taken up by the expression host, but no BaSP activity was detectable. This observation supports the notion that the expression of BaSP causes an unmanageable metabolic burden.[53] The specific CuCBP activity, instead, was determined as 1.86 U mg\textsuperscript{-1}, which corresponds to 23% of total soluble E. coli protein (Table 1). Nevertheless, similar to the bicistronic co-expression vectors, the activity was lost latest after the third round of expression from pDUAL_P\textsubscript{lacZ}.

In respect to the observed expression instabilities, earlier studies of Ellis et al.[52] showed, that the metabolic load increases with increase in both plasmid size (which is often the case for co-expression plasmids) and copy number. The energy required for plasmid maintenance is substantial and can result in lowering of the specific growth rate (Figure S4, Supporting Information) and the level of gene expression/plasmid.[53]

2.2. Redesigned Plasmids Improved Expression Stability and Growth Rate

To overcome the above-mentioned challenges, we redesigned the plasmid backbones. The newly constructed plasmids (pDUAL_2 and pBICI_2) carry a medium copy number origin of replication and, additionally, the so-called lacO site was introduced. The previously used pUC19 ori produces several hundred plasmid copies per cell,[55] whereas in its parental plasmid, with which it was replaced, the pBR322 ori reduces the copy number to just ≈30.[56] This significant difference results due to a G→A point mutation in the pUC19 ori that maps to the RNAII sequence and impedes RNAI:RNAII interaction.[57] This in turn lowers the plasmid replication frequency and results in a lower plasmid copy number.[57]
The second feature introduced, the bom site, is usually known as oriT and is required in cis for conjugal transfer between bacteria.\textsuperscript{58} OriT is the locus at which a strand-specific nick is introduced. To indeed transmit DNA from one cell to another, more than 20 tra genes need to be provided in trans (F-factor) for erection of the pilus.\textsuperscript{39} Commonly used laboratory E. coli strains, also those used in this study, do not contain F-factor in order to prevent spreading of antibiotic resistances (https://blog.addgene.org/plasmids-101-common-lab-e-coli-strains). The motivation for including the bom site was mainly its presence in commercially available plasmids, such as the pET-series\textsuperscript{60} as well as the pQE-series.\textsuperscript{61} As we exchanged the origin of replication with the one from pET21, we decided to keep the bom site.

Providentially, the lower copy number origin of replication was highly beneficial for expression stability of the pBICI\textsubscript{2} constructs (Figure 3A). The variation coefficient for expression from all three pBICI\textsubscript{2} constructs was within 15–25%, when considering at least three independent biological replicates. Similar improvements were noted for the pDUAL\textsubscript{2} constructs (Figure 3B). Replacement of the origin of replication abolished the transformation defect and resulted in a largely constant enzyme expression over at least three independent biological replicates. The variation coefficient for expression from pDUAL\textsubscript{2}P\textsubscript{araBAD} and -P\textsubscript{lac} was $\approx$30%. Expression from the remaining constructs showed a lower variation of $\approx$20%. Moreover, we achieved noteworthy percentages of recombinant protein in the cell-free extracts of cells carrying pDUAL\textsubscript{2}P\textsubscript{lac} and pDUAL\textsubscript{2}P\textsubscript{araBAD} (55% and 49%, respectively, Figure 3B). The corresponding SDS polyacrylamide gels of each representative biological replicate can be found in Figure S3, Supporting Information.

When cultivating cells in shake flasks, we noticed an improvement in growth rate attributable to the redesigned plasmids compared to the plasmids bearing the pUC19 ori. Therefore, we performed bioreactor cultivations (for details, see the Supporting Information) in which the growth of E. coli strains harboring pBICI\textsubscript{2} strong and pBICI\textsubscript{strong} vector could be compared under exactly comparable conditions. As shown in Figure S4, Supporting Information, cells harboring pBICI\textsubscript{2} strong instead of pBICI\textsubscript{strong} have a twofold shortened lag phase and a 1.6-fold enhanced maximum growth rate at a temperature of 25°C (0.85 h\textsuperscript{-1} compared to 0.52 h\textsuperscript{-1}). These results support the notion that salient characteristics of the combined production of CuCBP and BaSP are governed by the strong metabolic burden that the enzyme co-expression involves. The effect on lag-phase shortening when pBICI\textsubscript{2} strong instead of pBICI\textsubscript{strong} is used can be understood as an accelerated adaptation of the gene expression to the environmental conditions, which in turn would result from the decrease in the metabolic overload.\textsuperscript{53}

2.3. Uncoupling of the Effects of the bom-Site and the Origin of Replication

The possible contribution of the bom site (oriT) to plasmid stability immediately raised our interest. We considered that relevant insight would be of high fundamental importance and could also have significant practical impact. We chose to reconstruct the best performing co-expression plasmid, that is, pBICI equipped with the strong RBS, and used an approach targeted at uncoupling of the effects of the ori and the bom site. Two plasmids were designed to that end: pBICI\textsubscript{2} strong bom\textsuperscript{−} which harbors the medium copy origin of replication (pBR322) without the bom site, and pBICI\textsubscript{strong} bom\textsuperscript{−}, which carries the high copy origin of replication (pUC19) and the bom site. Thus, we ended up with four different constructs of pBICI\textsubscript{strong}, namely pUC19+bom: high origin without bom (pBICI\textsubscript{strong}), pUC19+bom: high origin with bom (designed pBICI\textsubscript{strong}), pBR322−bom: medium origin without bom (designed pBICI\textsubscript{2} strong), pBR322+bom: medium origin with bom (pBICI\textsubscript{2} strong).

To investigate plasmid and expression stability of each construct over several replication cycles, we designed a cell-passaging experiment. An aliquot of each cell culture (100 µL) was passed from the first uninduced main culture to the second preculture. Inoculation of the third preculture was done similarly. In total, three main cultures of each strain were tested regarding expression performance and stability, with a special focus on whether the bom site or the origin of replication showed any effects. Additionally, from each main culture and strain, the plasmids were isolated, transformed into E. coli Top10F\textsuperscript{+} to amplify the plasmids, isolated again, digested by HindIII and sequenced to exclude the possibility of any plasmid modifications.
Interestingly, the different plasmids bearing the pUC19 ori showed huge variation in expression performance and stability (Figure 4). When the bom site was present, expression from the high copy plasmid was even higher than from the plasmid with the lower copy number (∼1.4-fold). Moreover, the expression from the high copy number plasmid was highly stable, with a variation coefficient of only 5% considering all three cell passages. By contrast, the plasmid without the bom site gave almost no activity already in the first main culture. Plasmids bearing the lower origin of replication showed a highly stable expression with variations of only 4% and 17% for −bom and +bom, respectively (Figure 4). Nonetheless, a 1.7-fold higher CuCBP activity with pBR322 −bom in comparison to pBR322 +bom was observed. BaSP instead showed similar activities with both plasmids. Effect of the bom site varies apparently in strong dependence on the origin of replication.

2.3.1. Modification of High Copy Plasmid Causes Complete Expression Loss

Plasmids of each main culture were analyzed for modifications. All pBR322 ori plasmids (+bom and −bom) as well as the pUC19 +bom plasmid stayed unchanged during passaging, while in contrast the pUC19 −bom plasmid got modified (Figures S5 and S6, Supporting Information). Restriction analysis revealed a 4 kb smaller sized plasmid already in the first main culture (Figure S5, Supporting Information). This modification was further confirmed by sequence analysis (Figure S6, Supporting Information). The altered plasmid consisted of ampicillin resistance, LacI, some base pairs of the BaSP gene and the origin of replication. Thus, the gene of CuCBP and a major part of BaSP was missing (Figure S6A, Supporting Information).

Plasmid modifications are often the result of replication-transcription conflicts, as summarized by Merrikh et al.[62] The authors discuss that head-on encounters of the relevant enzyme machineries lead to over-wounding of the DNA template. The replisome and the transcription machinery simultaneously generate positive supercoiling in front, which in turn causes arrest of the replication fork.[62] One can expect, therefore, that the combination of a high-copy number origin of replication with a strong promoter can lead to the accumulation of such conflicts. This is most likely the reason, why 70% of the essential genes and even 100% of the tRNA and rRNA genes of E. coli are located on the leading strand of the chromosome, co-oriented...
with replication.\textsuperscript{[62]} Furthermore, it was shown before that prolonged stalling of the replication blocks the next round of replication. The replication fork collapses, and double-strand DNA breaks are generated.\textsuperscript{[62]} Such breaks are repaired by homologous recombination.\textsuperscript{[62]} Here, the action of recombinase RecA is central. First, RecA is recruited to randomly occurring ssDNA and forms a filament. The RecA-ssDNA filament binds to a dsDNA and samples for ssDNA-dsDNA homology. Once homology is encountered the actual strand exchange reaction results.\textsuperscript{[63]} Especially the abundance of plasmids in the cell and the more frequently occurring dsDNA breaks could reinforce such plasmid modifications caused by unnecessary homologous recombination events. For this very reason RecA is deleted in plasmid amplification strains.\textsuperscript{[54]} Expression hosts usually do express RecA, as its deletion causes less vital cells.\textsuperscript{[64]} Hence, we believe that RecA may have played a crucial role in the alteration of pBICI\_strong.

Discovery of the bom site involved in plasmid stability suggests further studies of the underlying mechanisms. It has been shown that bom sites (oriT) exhibit highly conserved and specific structural features.\textsuperscript{[65]} The DNA sequence around the nick site defines extensive secondary structures (e.g., hairpins) and acts as protein-DNA recognition site. The integration host factor (IHF), for instance, binds to these regions and stabilizes distinct DNA conformations that are required for several bacterial processes, including replication and transcription.\textsuperscript{[65,66]} Thus, biological context is suggested in which the bom site might be beneficial for expression and plasmid stability.

### 2.4. Optimized Cell-Factory for Future Industrial Cellobiose Production

The two best expressing vectors, pBICI\_2\_strong and pDUAL\_2\_PT5 were used in cellobiose synthesis. The least performing plasmid, pBICI\_2\_medium, was applied for comparison (Figure 5A–C, respectively). We distinguished cell-free extracts and freeze-thawed whole cells to investigate limitations on the reaction rate caused by mass transfer over the cell membrane. The initial space-time yields, calculated from the initial 7 to 9 h of reaction, were similar for whole cells and cell-free extracts (Figure 5D), excluding a significant effect of mass transfer. Initial space-time yield slightly higher (4–16%) for whole cells than cell-free extracts (Figure 5D) can arguably be explained by minor effects of the cell disruption on enzyme stability.

Table 2 summarizes the most relevant process parameters of the whole-cell conversions. In order to identify limitations caused by less than optimal activity ratios, the used catalysts were normalized to 1 U mL\(^{-1}\) CuCBP activity in the final reaction mixture. The highest initial catalyst activity and final space-time-yield were obtained with pBICI\_2\_strong, followed by the second-best expressing construct pDUAL\_P\_T5. The total turnover number was slightly increased with the pDUAL\_P\_T5 cells, which is due to the fact that CuCBP expression was approximately 1.2-fold higher with pDUAL\_P\_T5 than pBICI\_2\_strong. Hence, a proportionally lower cell concentration (0.39 g CDW (cell dry weight) L\(^{-1}\)) was applied to the conversion when the pDUAL\_P\_T5 cells were used.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Whole-cell conversions of sucrose (100 mm), glucose (100 mm), and phosphate (10 mm) to cellobiose. Cell catalysts carried either A) pBICI\_2\_strong, B) pDUAL\_2\_P\_T5 or C) pBICI\_2\_medium. D) Comparison of initial space-time yield (first 7 h) of whole-cell and cell-free extract conversions.}
\end{figure}
as compared to 0.46 gCDW L\(^{-1}\) when pBICL_2_strong was used. However, the two whole-cell systems showed very similar productivities, which is explainable on account of BaSP activity being limiting in the pDUAL_P\(_{5}\) reaction. Additionally, we consider that BaSP is slightly inhibited by glucose (≈20% activity loss at 100 mM glucose). BaSP activity in excess, as achieved with pBICL_2_strong, is further beneficial for the initial conversion rates (460 µmol cellobiose min\(^{-1}\) g\(^{-1}\) dry cells). Performance metrics of the synthesis (30 g product L\(^{-1}\); ≈82% yield; 3.8 g L\(^{-1}\) h\(^{-1}\) overall productivity) show that enzyme co-expression from pBICL_2_strong is highly promising for cellobiose production. The enzyme ratios established from the current whole cell approach appear to be more suitable for cellobiose production at high productivity than the SP-CBP enzyme ratios (1.2:1\(^{[26,28]}\); 1:3\(^{[27]}\)) used in earlier in vitro studies that employed isolated enzymes.

### 3. Conclusion

Combination of promoter and RBS engineering is here presented as a highly efficient strategy from synthetic biology to achieve tunable two-enzyme co-expression from a single plasmid vector. Activity ratios for BaSP and CuCBP were adjusted in a more than tenfold range (0.64–8.7), thus expanding the scope available to expression engineering for flux optimization. The activity ratio for the most active whole-cell catalyst had BaSP in substantial excess over CuCBP, which is a non-trivial finding of high practical significance. Earlier studies of cellobiose synthesis by isolated phosphorylases had used balanced activity ratios\(^{[26,28]}\) or CBP activity in excess.\(^{[27]}\) The performance metrics of the whole cell catalyst seem to be quite promising for cellobiose production. Due to the rational and semi-quantitative principle of promoter and RBS strength (strong/medium/weak) used in plasmid construction, we believe that the results of this study have excellent replication potential when applied to other enzyme cascades in different fields of bio-catalysis. Besides the well-known oxidoreductase systems that require regeneration of nicotinamide coenzymes and/or flavin cofactors\(^{[1,5,9,10,20]}\), two-enzyme cascades of glycoside phosphorylases similar to the one used here\(^{[67–69]}\), or cascades of two nucleoside phosphorylases\(^{[21,22,70]}\) are gaining increased importance. Two-enzyme cascades involving sugar nucleotide-dependent glycosyltransferases are emerging systems for synthesis.\(^{[71,72]}\) The co-expression approach developed herein can facilitate whole cell catalysis with each of these two-enzyme systems, but it is certainly not limited to them. It should therefore be of interest broadly across bio-catalysis. In the course of analyzing the expression stability of the plasmid constructs developed herein, we discovered an important role of the bom site in preventing expression loss dependent on the origin of replication. The evidence on the bom site can be quite relevant to those in the field who must (re)design co-expression vectors, in order to install features lacking in the ready-to-use commercial plasmids.

### 4. Experimental Section

**Plasmid Construction:** Custom vectors for the co-expression were designed using DNA parts amplified from templates available in-house. Constructs were designed using SnapGene software (GSL Biotech LLC, Chicago, IL, USA). Synthetic DNA fragments and primers were purchased from Integrated DNA Technologies, Inc. (Caroline, IA, USA). Parts were assembled by overlap extension PCR\(^{[73]}\) and restriction cloning. E. coli Top10F\(^{\circ}\) (Thermo Fisher Scientific Inc., Waltham, MA, USA) was used for plasmid amplification. Plasmids were sequence-verified by restriction analysis and sequencing (Microsynth Austria GmbH, Vienna, Austria). PCR products were isolated with GeneJET Gel Extraction Kit and plasmids were extracted using GeneJET Plasmid Miniprep Kit (both Thermo Fisher Scientific Inc.). Phusion Polymerase for PCR, restriction enzymes and T4 DNA ligase were from Thermo Fisher Scientific Inc.

Table S1, Supporting Information, lists all parts used for construction of co-expression plasmids. Additionally, the complete sequences of all co-expression plasmids were uploaded to the Addgene Plasmid Database (plasmid-ID 155163 to 155178). Ribosome binding sites (RBS) were chosen that had been classified as weak and medium by virtue of their relative translation efficiencies of 0.35% and 33.94% in relation to the strong RBS published by Elowitz and Leibler\(^{[44]}\), whose translation efficiency was set to 100%. The RBS classification is based on experimental data generated by Team Warsaw in the course of their 2010 iGEM project (http://2010.igem.org/Team: Warsaw/Stage1/RBSMeas).

E. coli Top10F\(^{\circ}\) was used for plasmid amplification and E. coli BL21(DE3)agp for enzyme expression. E. coli strain BL21(DE3)agp\(^{47}\), which carries a deletion of the glucose-1-phosphatase (agp) gene to prevent degradation of aGlc1-P (Figure 1) in whole-cell conversions, was obtained from Tom Desmet (Chent University). Cells were transformed using electroporation\(^{[74]}\) and regenerated in 1 mL SOC medium\(^{[75]}\) for 1 h at 37 °C. Positive transformants were selected on lysogeny broth (LB)-agar plates containing 100 mg L\(^{-1}\) ampicillin.

**Cell Preparation:** E. coli strains were grown at 37 °C in LB-medium (5 g L\(^{-1}\) NaCl, 5 g L\(^{-1}\) yeast extract, 10 g L\(^{-1}\) peptone from casein) in baffled shake flasks containing 100 mg L\(^{-1}\) ampicillin. The main culture (250 mL in 1 L flasks) was inoculated with cells from an overnight grown preculture (50 mL in 300 mL flasks). Biomass was grown to OD\(_{600}\) of 0.8–1.0 and expression was induced with 1 mM IPTG. Incubation was overnight at 25 °C and 110 rpm in incubation shaker CERTOMAT BS-1 (Sartorius,

| Plasmid          | Concentration [gCDW L\(^{-1}\)] | 5 h | 24 h | Cell catalyst | Initial activity [U gCDW\(^{-1}\)] | Total turnover number [g CDW\(^{-1}\)] | Space-time yield [g L\(^{-1}\) h\(^{-1}\)] | Product [g L\(^{-1}\)] |
|------------------|----------------------------------|-----|-----|---------------|-----------------------------------|------------------------------------------|-------------------------------------------|---------------------|
| pBICL_2_strong   | 0.46                             | 64  | 82  | BaSP          | 461 ± 48                          | 62                                       | 3.8 ± 28                                  | 28                  |
| pDUAL_2_P\(_{5}\) | 0.39                             | 39  | 80  | CuCBP         | 329 ± 28                          | 69                                       | 1.7 ± 27                                  | 27                  |
| pBICL_2_medium   | 1.14                             | 23  | 76  |               | 64 ± 3                            | 23                                       | 1.1 ± 26                                  | 26                  |

"Gram cell dry weight per liter reaction mixture. Cell catalyst initial activity (U gCDW\(^{-1}\)) was calculated from four initial time points. The time point (h) from which the highest cellobiose yield (g L\(^{-1}\)) was reached, was used for calculation of space-time yield."
Göttingen, Germany). OD$_{600}$ was measured spectrophotometrically (DU 800 UV/Vis Spectrophotometer, Beckman Coulter, Brea, CA, USA). Cell dry weight (CDW) was determined by filtering cell culture (15–20 mL) over pre-weighed Whatman Nuclepore TrackEtched membrane (diameter 50 mm, pore size 0.4 µm, polycarbonate, Sigma-Aldrich/Merck, Darmstadt, Germany). Washed cells (10–15 mL water) were dried overnight at 70 °C and weighed. Centrifuged cells (20 min, 4 °C, 5000 rpm, Ultracentrifuge Sorvall RC-5B Superspeed) were resuspended in 50 mM MES (2-(N-morpholino)ethanesulfonic acid) buffer, pH 7.0 (6:1 cell wet weight, v:v), aliquoted (±10–15 mL portions) and stored at −80 °C until further use.

To prepare the cell-free extract, an aliquot of the cell suspension was thawed and ultra-sonicated (Branson Ultrasound Microtip Probe 1/8"). The mixture consisted of 100 mM sucrose, 100 mM glucose, 10 mM phosphate, 1 U mL$^{-1}$ CuCBP (synthesis direction, 1.8-fold phosphorolysis activity),[31] and 50 mM MES buffer. 50 mL reactions were performed in 100 mL borosilicate glass bottles equipped with Rotilabo magnetic sticks (25 × 8 mm, Carl Roth GmbH, Karlsruhe, Germany). OD$_{600}$ was measured spectrophotometrically at 340 nm (DU 800 UV/Vis Spectrophotometer, Beckman Coulter, Brea, CA, USA). Calibration was done with standards containing main components.

HPLC Analysis of Reaction Compounds

Purified crude biomass was used for the analysis of reaction compounds. A master mix containing all reaction components, except the catalyst, was distributed to the bottles. Reactions were started by adding the catalysts. BaSP activities were determined in phosphorolysis direction of reaction (Figure 1) by a continuous coupled enzyme assay.[76] The aGlC1-P liberated on enzyme activity was converted by phosphoglucomutase from rabbit muscle (3 U mL$^{-1}$, Sigma-Aldrich/Merck, Darmstadt, Germany) and NAD$^+$-dependent β-glucose-6-phosphate dehydrogenase from Leuconostoc mesenteroides (3.4 U mL$^{-1}$, Sigma-Aldrich/Merck, Darmstadt, Germany) to NADH, which was then monitored spectrophotometrically at 340 nm (DU 800 UV/Vis Spectrophotometer, see the Supporting Information for details). Protein concentration was calculated from a calibration curve in the range of 0.1 to 1.0 g L$^{-1}$ protein (bovine serum albumin, Sigma-Aldrich/Merck, Darmstadt, Germany). Overexpression of the enzyme in total soluble E. coli protein was calculated from the specific phosphorolysis activities of cell extract and isolated enzyme. The specific activities used for BaSP$^{[10]}$ and CuCBP$^{[31]}$ were 122 and 8 U mg$^{-1}$ (or 14 U mg$^{-1}$ in synthesis direction), respectively. The literature values were confirmed from this study.

**Cellobiose Synthesis in Whole Cell and Cell-Free Systems:** The reaction mixture consisted of 100 mM sucrose, 100 mM glucose, 10 mM phosphate, 1 U mL$^{-1}$ CuCBP (synthesis direction, 1.8-fold phosphorolysis activity),[31] and 50 mM MES buffer. 50 mL reactions were performed in 100 mL borosilicate glass bottles equipped with Rotilabo magnetic sticks (25 × 8 mm, Carl Roth GmbH, Karlsruhe, Germany). OD$_{600}$ was measured at 30 °C and 300 rpm for 32 h on an Variomag Multi-Magnetic Stirrer (Thermo Fisher Scientific Inc. Walthman, MA, USA).

Activities of thawed cell-free extracts were measured shortly prior to the reaction start. Cell-free extracts and thawed cell suspensions (semi-permeabilized cells) were treated similarly. Therefore, it was assumed that semi-permeabilized cells show similar activities as corresponding cell-free extracts (prepared from the same cultivation batch), except for possible activity losses during cell-free extract preparations. A master mix containing all reaction components, except the catalyst, was distributed to the bottles. Reactions were started by adding the catalysts. BaSP activities were calculated based on cell-free extract activities. At certain times, 1 mL of the reaction mixture was withdrawn and heated to 99 °C for 5–10 min to stop the reaction (ThermoMixer C, E-5048, Eppendorf, Hamburg, Germany), precipitated proteins were centrifuged for 10 min at 15 000 rpm (Centrifuge Eppendorf 5424 R, Eppendorf, Hamburg, Germany) and the supernatant was stored at −20 °C until HPLC measurements were performed.

**HPLC Analysis of Reaction Compounds:** Quantification of main components (sucrose, glucose, fructose, cellobiose) in the reaction mixtures was performed by HPLC using a Merck Hitachi L-7100 system (Merck, Darmstadt, Germany) equipped with an autosampler and RI-detector. As YMC-Pack Polyamide II (5×5 µm) or 12 nm column (250 mm × 4.6 mm) was used and additionally a guard column (20 mm × 4.0 mm) was installed (YMC Co., Ltd., Shimogyo-ku, Kyoto, Japan). Elution was performed isocratically with an acetonitrile-water mixture (75:25, v:v) at a flow rate of 1 mL min$^{-1}$. Measurements were performed at room temperature, the injection volume per sample was set to 20 µL, and the running time was 35 min. Refractive index detected peaks were analyzed using the software Chromleon Chromatography Data System (Thermo Fischer Scientific Inc., Waltham, MA, USA). Calibration was done with standards containing main components.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 761030 (CARBAFIN). H.D. acknowledges support from Erasmus+ for student exchange from Slovak University of Technology in Bratislava to Graz University of Technology.

Conflict of Interest

The authors declare no conflict of interest.

Data Accessibility Statement

Data obtained in the current study are available from the DOI 10.5281/zenodo.3706423.

Keywords

cellobiose synthesis, co-expression, phosphorylases, ribosome binding sites, synthetic biology, whole-cell biocconversion

Received: March 17, 2020
Revised: June 16, 2020
Published online: August 7, 2020

[1] J. H. Schrittwieser, S. Velikogne, M. Hall, W. Kroutil, Chem. Rev. 2018, 118, 270.
[2] J. Muschiol, C. Peters, N. Oberleitner, M. D. Mihovilovic, U. T. Bornscheuer, F. Rudroff, Chem. Commun. 2015, 51, 5798.
[3] S. P. France, L. J. Hepworth, N. J. Turner, S. L. Flitsch, ACS Catal. 2017, 7, 710.
[4] S. Wu, Z. Li, ChemCatChem 2018, 10, 2164.
[5] F. Rudroff, Curr. Opin. Chem. Biol. 2019, 49, 84.
[6] B. Lin, Y. Tao, Microb. Cell Fact. 2017, 16, 106.
[7] J. C. C. R. de Carvão, Microb. Biotechnol. 2017, 10, 250.
[8] M. Liu, H. Yu, Biochem. Eng. J. 2012, 68, 1.
[9] K. B. Otte, J. Kittelberger, M. Kirtz, B. M. Nestl, B. Hauer, ChemCatChem 2014, 6, 1003.
[10] A. Kohl, V. Srinivasasamurthy, D. Böttcher, J. Kabisch, U. T. Bornscheuer, Enzyme Microb. Technol. 2018, 108, 53.
