TECHNISCHE UNIVERSITÄT MÜNCHEN

Integrative Research Center Campus Straubing für Biotechnologie und Nachhaltigkeit

Metabolic Engineering of *Paenibacillus polymyxa* DSM 365

Christoph Schilling

Vollständiger Abdruck der von der promotionsführenden Einrichtung Campus Straubing für Biotechnologie und Nachhaltigkeit der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Rubén Dario Costa Riquelme
Prüfer der Dissertation:	1. Prof. Dr. Volker Sieber
	2. Prof. Dr.-Ing. Jochen Schmid
	3. Prof. Dr. Bastian Blombach

Die Dissertation wurde am 17.02.2021 bei der Technischen Universität München eingereicht und von der promotionsführenden Einrichtung Campus Straubing für Biotechnologie und Nachhaltigkeit am 30.08.2021 angenommen.

Danksagung

Zu allererst möchte ich mich bei meinem Doktorvater Prof. Dr. Volker Sieber bedanken. Du hast mir nicht nur ein extrem vielschichtiges Promotionsthema gegeben, sondern mir auch stets die individuelle Freiheit gelassen, meine wissenschaftliche Neugier im ein oder anderen Nebenprojekt ausleben zu können. Stets konnte ich dabei auf deine fachliche Unterstützung zählen. Danke, dass du mich nach Straubing geholt hast und ich an deinem Lehrstuhl promovieren durfte.

Ein ganz besonderer Dank gilt auch Prof. Dr. Jochen Schmid. Selbst nach deiner Zeit hier in Straubing warst du mir während meiner gesamten Promotion mein Mentor und auch Freund. Danke für die zahlreichen Korrekturen, Fachgespräche und die Brainstorming Sessions über das nächste Graphical Abstract zusammen mit Anja. Auf das uns der Schleim niemals ausgeht.

F&E stehen gerade in Naturwissenschaften oftmals für Frust und Enttäuschung. Darum möchte ich mich ganz herzlich bei allen Mitgliedern des EPS-Teams bedanken, die mir stets mit Rat und Tat zur Seite gestanden sind. Danke Broder Rühmann für das gemeinsame Tüfteln an meinem Sudoku. Danke Moritz Gansbiller, dass du mich in die Welt der Rheologie eingeführt hast und danke Tristan Rath für den entscheidenden Hinweis zur Struktur von Paenan.

Auch möchte ich mich bei allen Kolleginnen und Kollegen des CBRs für die wahnsinnig gute und freundschaftliche Atmosphäre an unserem Lehrstuhl bedanken, sodass gerne ein Abend in einer Partie Risiko oder Siedler geendet hat. Ein Dank gilt auch Tobias Gmelch und Torben Hülsing, die mir gezeigt haben, dass man auch im Bayerischen Wald den ein oder anderen „Berg“ erklimmen kann.

Ein großer Dank gilt auch meinen Eltern Gertraud und Karl Schilling, sowie meinem Bruder Robert, für das wunderbare Leben, dass ihr mir ermöglicht habt.

Mein größter Dank geht an meinen Lebensmittelpunkt Julia Wünsche für die unglaubliche Unterstützung in sämtlichen Lebenslagen und die immense Hilfe beim Korrekturlesen dieser Arbeit. Ohne dich wäre das alles nicht möglich gewesen. Du bist und wirst immer meine „größte“ Entdeckung hier in Straubing sein.

Abstract

Paenibacillus polymyxa is a Gram-positiv, endospore forming soil bacteria belonging to the plant growth promoting rhizobacteria and has been been applied as a biofertilizer in agriculture and as a producer of numerous antibiotics. Furthermore, *P. polymyxa* is of biotechnological interest for the biosynthesis of highly enantiopure butanediol (2,3-BDL) and has been shown to be an avid producer of various exopolysaccharides (EPS).

In this work, *P. polymyxa* DSM 365 was metabolically engineered to increase the 2,3-BDL production by decoupling the proposed bottleneck enzyme butanediol dehydrogenase from its natural regulation system. Deletion of a lactate dehydrogenase resulted in an increased growth behaviour. Furthermore, undesirable side products such as spore formation and EPS that impede downstream processing production were eliminated. Thereby, 2,3-BDL synthesis was increased by 46 % achieving a final product titer of 43.8 g L^{-1} in batch fermentations, accounting for 86 % of the theoretical maximum.

In order to elucidate the structure of the heteroexopolysaccharide produced by *P. polymyxa* DSM 365, combinatorial knock-outs of glycosyltransferases and Wxy polymerases were performed. Through complementing analytical approaches including carbohydrate fingerprinting, partial hydrolysis, methylation analysis and NMR, it was discovered that the EPS in fact consists of three distinct polymers. Using a total of 32 mutant strains, it was possible to determine the monomer composition, sequence and linkage pattern of each polymer and attribute specific enzymes to the biosynthesis. Rheological properties of EPS compositions produced by knock-out strains were characterized in detail. Depending on the present polymers, unique physicochemical properties were observed ranging from a shear-thinning, highly viscous gel character to low viscosity Newtonian fluids. Thereby, this study provides the foundation for the rational design of tailor-made EPS, which might be applied in a wide range of applications.

Furthermore, a novel broad host tool was developed to enable the efficient, multiplex transcriptional perturbation of genes of interest. By transcriptional fusion of the activator of the superoxide regulon SoxS to a catalytically inactive variant of Cas12a, expression of target genes could be increased or repressed, depending on the positioning of the effector module. Thereby, it proved possible to rapidly emulate knock-out variants of the EPS cluster or control the flux of metabolites towards 2,3-BDL in *P. polymyxa*.

P. polymyxa is a potent alternative chassis organism for the production of numerous compounds of interest. CRISPR-Cas technology enabled the efficient metabolic engineering of this otherwise genetically hardly accessible organism towards a sustainable production of 2,3-BDL and polysaccharide based rheological modifiers.

Table of Content

List of Publications

Schilling C, Badri A, Sieber V, Koffas M & Schmid J (2020) Metabolic engineering for production of functional polysaccharides. *Current Opinion in Biotechnology* 66:44–51. https://doi.org/10.1016/j.copbio.2020.06.010

Schilling C, Ciccone R, Sieber V, Schmid J (2020) Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365. *Metabolic Engineering* 61:381–388. https://doi.org/10.1016/j.ymben.2020.07.009

Schilling C, Koffas MAG, Sieber V, Schmid J (2020) Novel Prokaryotic CRISPR-Cas12a-Based Tool for Programmable Transcriptional Activation and Repression. *ACS Synth Biol* 9:3353–3363.https://doi.org/10.1021/acssynbio.0c00424

Meliawati, M, Schilling, C & Schmid, J (2021) Recent advances of Cas12a applications in bacteria. *Appl Microbiol Biotechnol* 105**:**2981–2990. https://doi.org/10.1007/s00253-021-11243-9

Schilling C, Klau L, Rühmann B, Aachmann F, Schmid J & Sieber V (2022) Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365.*Carbohydrate Polymers* 278: https://doi.org/10.1016/j.carbpol.2021.118951

List of Abbreviations

2,3-BDL	2,3-butanediol
ABC	ATP-binding cassette
ABTS	2-azino-bis-(3-ethylbenzthiazoline)-6-sulfonic acid
Ac	acetate
acr	anti CRISPR protein
adh	alcohol dehydrogenase
ATP	adenosin triphosphate
bdh	butanediol dehydrogenase
bp	base pair
CARF	CRISPR associated Rossman fold
Cas	CRISPR associated protein
CPS	capsular polysaccharide
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
crRNA	CRISPR RNA
DA-64	N-(carboxymethylamino-carbonyl)-4.4'-bis-(dimethylamino)-diphenylamine sodium salt
ddH_2O	ultra-pure water
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EPS	exopolysaccharide
ESI	electrospray ionization
Fuc	L-fucose
G'	storage modulus
G"	loss modulus
GAG	glycosaminoglycan
Gal	D-galactose
GC	gas chromatography
Glc	D-glucose
GlcNAc	N-acetylglucoseamine
GOI	gene of interest
GPC	gel permeation chromatography
GRAS	generelly regarded as sage
GT	glycosyltransferase
GulA	L-guluronic acid
H2BC	heteronuclear two-bond correlation
HDR	homology directed repair
HEPN	higher eukaryotes and prokaryotes nucleotide-binding domain
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence

kb	kilobase
Kdo	3-deoxy-D-manno-octulosonic acid
KRAB	Krüppel-associated box
LB	lysogeny broth
ldh	lactate dehydrogenase
MALS	multi angle light scattering
Man	D-mannose
ManA	D-mannuronic acid
MES	2-(N-morpholino)-ethane-sulfonic acid
MOPS	3-(N-morpholino)-propane-sulfonic acid
NAD^+	nicotinamide adenine dinucleotide
neo	neomycin
NHEJ	non-homologous end joining
NMR	nuclear magnetic resonance
OPX	outer membrane export
PAM	protospacer adjacent motif
pfl	pyruvate formate lyase
PMAA	partially methylated alditol acetate
px	polymyxin
Pyr	pyruvate
RAST	rapid annotation using subsystems technology
Rha	L-rhamnose
RNA	ribonucleic acid
rpm	revolutions per minute
RT	room temperature
sgRNA	single guide RNA
SNFG	symbol nomenclature for glycans
SS	small subunit
TALEN	transcription activator like effector nuclease
TFA	trifluoroacetic acid
TRIS	trishydroxymethylaminomethane
trRNA	trans-activating RNA
UDP	uracil diphosphate
UHPLC	ultra-high performance liquid chromatography
UV	ultraviolet light
V	adenine, cytosin or guanine
v/v	volume per volume
w/v	weight per volume
w/w	weight per weight
x g	times g-force
η	dynamic viscosity

1. Introduction

1.1. *Paenibacillus spp.*

Paenibacillus spp. are characterized as rod-shaped, Gram-positive or Gram-variable endospore forming, facultative anaerobic bacteria. Currently, the diversity of the genus comprises over 239 characterized species that are found in highly variable environments ranging from the oceanic sediment to the Antarctica as well as desserts [1]. Though, *Paenibacillus* strains have been most commonly isolated from soil samples and are often associated with roots as part of the rhizosphere of various plants. Many species of *Paenibacillus* are important for plants, animals and humans.

Species of *Paenibacillus* were originally included in the genus *Bacillus*, which had been defined only loosely based on morphological characteristics and the ability to form endospores. With the rise of 16S rRNA gene sequencing in the early 1990s, phylogenetic classification redefined five new genera, including the novel genus *Paenibacillus* (*paene* lat. = almost) with the type strain *P. polymyxa* [2,3]. Since 1997, the genus is characterized based on its rod-shaped form with peritrichous flagella and the non-pigmented shaping on nutrient agar. In addition to the aforementioned morphological criteria for *Bacillus* spp., the genus shares more than 89.6 % 16S rRNA sequence similarity [4].

Paenibacillus spp. are known biofertilizer included in the group of plant growth promoting rhizobacteria [5]. Particular strains have demonstrated the ability to fixate nitrogen as well as to solubilize otherwise inaccessible phosphorus and thereby enhance plant growth. Genome analysis revealed over 20 root colonizing strains using a molybdenum dependent nitrogenase for the reduction of N_2 to bioavailable NH_3 [6]. In addition, *Paenibacillus* spp. are also known for positively contributing to plant growth through the production of phytohormones that are highly important regulators of gene expression and development of plants during their life cycle. The pathway for the production of indole-3 acetic acid, one of the most abundant auxins found in nature, is highly conserved among *Paenibacillus* spp. Genome analysis thereof revealed that all sequenced strains up to now encode the required key enzymes for its biosynthesis [7]. These multi-facetted features make *Paenibacillus* spp. an interesting agricultural biofertilizer. Moreover, it was shown that plant growth could be enhanced through seed treatment and soil inoculation by high doses of *P. polymyxa.* A positive effect in form of increased shoot fresh weight, plant height and ultimately yield was obtained for numerous model and crop plants including rice, cucumber, sesame, *Arabidopsis thaliana* and pumpkin [8].

Another important aspect in the evaluation of *Paenibacillus* spp. as plant growth promoting

rhizobacteria is its pronounced ability to produce an immense arsenal of antimicrobial biocontrol agents that efficiently suppress potential plant pathogens. In addition to chitinases that hydrolyze exoskeletons and gut intestines of pest insects, Cry proteins produced by *P. lentimorbus* are homologous to the insecticides produced by *Bacillus thuringiensis.* The structural similarity enables the formation of pores within the gut epithelia of insects, resulting in cell lysis and consequently the mortality of infected beetle larvae [9]. Further hydrolytic enzymes such as glucanases, cellulases, mannanases and amylases also contribute to the protection against pathogenic fungi [10]. The best described antimicrobial agents produced by *Paenibacillus* spp. are a class of cyclic cationic lipopeptides called polymyxins. These antibiotics consist of a diaminobutyryl-containing heptapeptide ring and a tripeptide side chain that binds to the lipid A compound of the lipopolysaccharide of Gram-negative bacteria and thereby permeabilize the outer, as well as the inner membrane [11]. Although several toxic side effects to the human central nervous system and kidneys are known, polymyxin B and E (colistin) remain important antibiotics of last resort against multidrug resistant Gram-negative pathogens. Furthermore, new semi-synthetic derivatives with improved pharmacokinetic properties have been developed [12]. Additionally, *Paenibacillus* spp. isolates have demonstrated the capability to produce fusaricidins, tridecaptins and lantibiotics such as paenilin and pediocin that further complement their antimicrobial arsenal against Gram-positive, Gram-negative and fungal competitors [13–17].

Contrary to the numerous positive effects as biofertilizers and strains of pharmaceutical interest, there are also pathogenic strains affecting mankind and animals. One of the best studied disease associated to the genus *Paenibacillus* spp. is the most destructive brood disease of honey bee hives named American Foulbrood [18]. The causative agent *P. larvae* afflicts bee larvae, where it rapidly proliferates in the midgut and breaches the epithelium ultimately resulting in the death of the host. In the attempt to clean the hive of deceased larvae, adult bees further spread the infectious spores, which in the worst case scenario induce the complete collapse of the colony when left untreated [19]. Additionally, *Paenibacillus* spp. are forming the predominant genus associated with the spoilage of dairy products. Endospores are highly resilient and resist extreme environmental conditions such as pasteurization and UV irradiation and therefore contribute to the limited shelf life time of milk products [20]. Due to their ability to grow at low temperatures, *Paenibacillus* spp. have demonstrated to represent more than 95 % of the bacterial population in pasteurized milk after 10 days of refrigerated storage making it an important agent of food spoilage [21].

In recent years, the non-pathogenic strain *P. polymyxa* became increasingly attractive for biotechnological industry as a production platform for bulk chemicals. On the one hand, *Paenibacillus* spp. are capable of synthesizing relevant amounts of 2,3-*R,R*-butanediol of

exceptionally high enantiomeric purity [22]. On the other hand, this genus is also known to produce a variety of exopolysaccharides that have demonstrated remarkable physicochemical properties for the application as rheological modifiers as well as antioxidant radical scavengers [23,24]. Due to previously established tools for the genetic modification of *P. polymyxa* DSM 365 [25], this thesis focuses on the metabolic engineering of this promising strain for the production of high-value products in the light of a growing interest for a sustainable bio economy.

1.2. Butanediol

2,3-butanediol (2,3-BDL) is considered as a platform chemical for a variety of applications. Currently, 2,3-BDL is mostly derived from petrochemical sources, however there is also a long history of research in its biotechnological production [26]. 2,3-BDL is mainly used as a precursor molecule for further conversion to other products with an annual world market capacity of 32 million tons in 2011, corresponding to a value of approximately $43 billion [27]. The wide range of applications of 2,3-BDL is highly attributed by its appearance in three stereoisomers – levo-BDL (*R,R*-BDL), dextro-BDL (*S,S*-BDL) and meso-BDL (*R,S*-BDL). In particular the optically active forms, levo-BDL and dextro-BDL are of high industrial interest and can be used for the synthesis of chiral specialty chemicals [28]. Dehydration result in the formation of methyl ethyl ketone, which is used as a fuel additive or solvent for resins and lacquers. Dehydrogenation of 2,3-BDL results in the buttery food aroma diacetyl [29]. During World War II, biotechnological production of 2,3-BDL was already of industrial interest for the production of 1,3-butadiene which was subsequently used in the organic synthesis of polyesters [30]. Due to a low freezing temperature at -60 °C, 2,3-BDL was proposed as an anti-freezing agent in the 1940s [31]. On the other hand, the boiling point ranges between 177 °C to 182 °C depending on the stereoisomer. This makes product recovery from the fermentation broth via distillation extremely cost intensive [32]. Therefore, downstream-processing represents a major bottleneck in the biotechnological production of 2,3-BDL and requires high product titers to ensure an economically feasible process.

With an ever-increasing awareness of the limited nature of fossil resources and the search for an environmentally friendly alternative, research in the biotechnological production of 2,3-BDL has increased over the past decades. A variety of different organisms including *Klebsiella* spp., *Serratia marcescens, Bacillus* spp. or *P. polymyxa* have demonstrated the capability to produce 2,3-BDL in remarkable amounts (Table 1). The highest product titer of more than 150 g L^{-1} with a productivity of 4.2 g L^{-1} has been described for *K. pneumoniae* in fed-batch fermentations [33]. Optimization of process parameters resulted in a product titer of 139.9 g L^{-1} for *S. marcescens* [34]. While the productivity and yield are worthy of attention, both organisms mainly produce the optically inactive meso-butanediol. Furthermore, the pathogenic nature of these strains makes them less desirable for the biotechnological production in an industrial scale process. In contrast, known production strains of the genus *Bacillus*, such as *B. amyloliquefaciens* and *B. licheniformis*, are generally recognized as safe (GRAS) and have also demonstrated high butanediol titers [35,36]. However, without any further genetic engineering the main product is composed of the less valuable, optically inactive meso-BDL. *P. polymyxa* on the other hand is a non-pathogenic Gram-positive organism that mainly produces *R,R*-butanediol with exceptionally high enantiomeric purity up to 98 % [37]. The enantioselectivity as well as

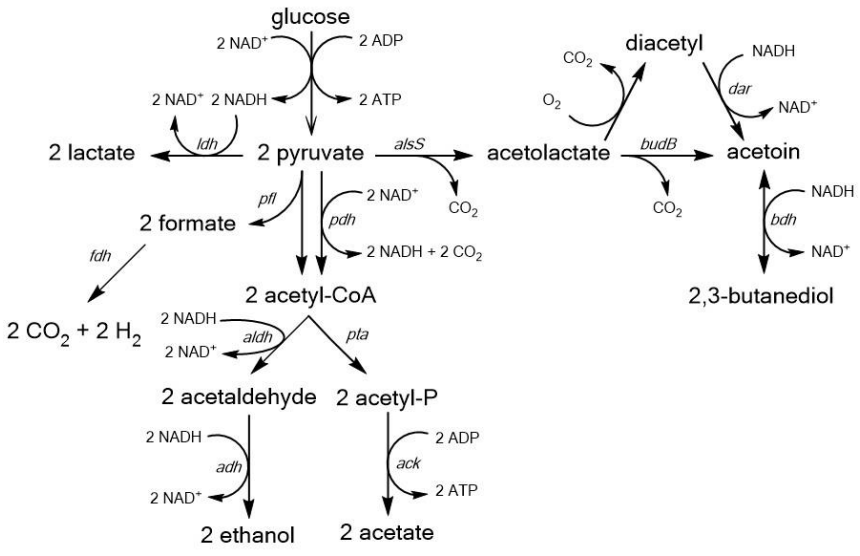
the non-pathogenic nature of this strain makes it a promising platform for large-scale production of 2,3-BDL. Existing studies to this topic report obtained product titer of 111 g L^{-1} by using a high nutritional media containing 60 g L^{-1} yeast extract [22]. By applying statistical approaches in order to optimize media composition and process parameters, it was also possible to produce up to 68.5 g L^{-1} without the industrially unfeasible use of enormous amounts of yeast extract [38].

Table 1: Overview of 2,3-BDL titer and productivity achieved by fermentation of different organism

Organism	Fermentation mode	2,3-BDL [g L^{-1}]	Productivity [g L^{-1} h^{-1}]	Reference
K. pneumoniae	Fed-Batch	150.0	4.21	[33]
K. oxytoca	Batch	127.9	1.78	[39]
K. oxytoca	Batch	95.5	1.17	[40]
S. marcescens	Fed-Batch	139.9	3.49	[34]
Enterobacter aerogenes	Fed-Batch	118.0	2.27	[41]
B. amyloliquefaciens	Fed-Batch	132.9	2.95	[42]
B. licheniformis	Fed-Batch	144.7	1.14	[35]
P. polymyxa DSM 365	Fed-Batch	111.0	2.05	[22]
P. polymyxa DSM 365	Fed-Batch	68.5	0.70	[38]
P. polymyxa DSM 365	Batch	51.1	1.70	[38]
P. polymyxa ATCC 12321	Fed-Batch	40.0	0.80	[43]
P. polymyxa ZJ-9	Fed-Batch	37.6	0.88	[44]

1.2.1. Biosynthesis of butanediol

The biosynthesis pathways of 2,3-butanediol in microorganisms is thoroughly studied and well understood (Figure 1). 2,3-BDL is produced by a mixed acid fermentation pathway under microaerobic conditions starting from the metabolic intermediate pyruvate. Using glucose as a carbon source, two mols of pyruvate are generated through glycolysis from one mol of glucose. Subsequently, two mol pyruvate are condensed to α-acetolactate, which is further decarboxylated to acetoin and finally reduced to butanediol while simultaneously regenerating one redox equivalent of NAD^+. In the absence of alternative electron acceptors at anaerobic or microaerobic conditions, the excessive NADH from glycolysis cannot easily be oxidized. Therefore, in order to maintain the NAD^+/NADH equilibrium, more undesired redox-neutral side products such as ethanol or lactate are formed [45]. Depending on oxygen availability, lactate dehydrogenase, pyruvate formate lyase, as well as the pyruvate dehydrogenase complex compete for the intermediate pyruvate and determine the end products [46]. Anaerobic conditions result in the production of lactate and ethanol, while an increase of dissolved oxygen favors acetoin and acetate production over 2,3-BDL biosynthesis [43]. A redox efficient route is represented by the oxygen sensitive enzyme pyruvate formate lyase (Pfl). Pfl catalyzes the conversion of pyruvate to acetyl-CoA without generating NADH. Subsequently, two mols of NAD^+ can be regenerated by the combined action of acetaldehyde dehydrogenase and alcohol dehydrogenase. Therefore, this pathway results in a net gain of one redox equivalent per mol of pyruvate that can compensate for the redox imbalance of the butanediol biosynthesis [47].



2.

Figure 1: Overview of the mixed acid pathway and butanediol biosynthesis under microaerobic conditions. Two mols of NADH are formed during glycolysis that need to be regenerated in order to maintain redox balance. Only one mol NADH is converted to NAD^+ in the 2,3-BDL pathway. Therefore, other redox neutral end products such as lactate or ethanol compete for the intermediate pyruvate.

However, the precise reason why microorganism produce 2,3-BDL in this redox inefficient pathway remains to be elucidated. A proposed explanation for this phenomenon is that the synthesis of 2,3-BDL provides a pyruvate drain to avoid elevated acidification of the organism [48]. Early research on butanediol production showed that the addition of acetic acid to the fermentation medium and a lower pH indeed increased 2,3-BDL biosynthesis in many organisms [49].

1.2.2. Metabolic engineering of butanediol producing organism

Biosynthesis of high amounts of butanediol is mostly directed by oxygen supply and the need to maintain redox balance under microaerobic conditions. With an increasing number of tools enabling the efficient genetic modification of widely used model organism such as *E. coli*, but also alternative host organisms, there have been plenty of publications aiming to optimize the biotechnological production of butanediol to enhance industrial feasibility. The most commonly used approaches of metabolic engineering to improve the production of 2,3-BDL can be categorized in three groups:

1) redirection of the carbon flux from undesirable side products towards butanediol

2) decoupling of genes of the butanediol pathway from their natural regulon

3) increasing the viability of cells and consequently biomass formation at conditions with low oxygen supply.

Jung et al. (2012) demonstrated that deletion of a lactate dehydrogenase (*ldh*) in *Enterobacter aerogenes* results in the attenuation of lactate as a side product while also increasing the viability of the cells, ultimately leading to 27.4 % higher product titers in fed-batch cultivation compared to the wildtype strain [41]. In *K. pneumoniae*, the combinatorial knock-out of *ldhA*, pyruvate formate lyase (*pfl*), as well the gene *wabG* encoding a glycosyltransferase reduced both the production of several by-products and the pathogenic character mediated by exopolysaccharide formation. Consequently, a yield of 0.46 g g^{-1} glucose, corresponding to 92.2 % of the theoretical maximum, was obtained [50]. In *K. oxytoca* yields were even further improved by deletion of multiple genes of the mixed-acid pathway. In the strain KMS005-73T Δ*adhE* Δ*ackA-pta* Δ*ldhA* lactate and formate formation were completely eliminated, while only trace amounts of acetate and ethanol were detected [51]. As a result, the main carbon flux was channeled into the butanediol biosynthesis pathway resulting in a product titer of 117.4 g L^{-1} and a yield of 0.49 g g^{-1}.

Gene expression levels of enzymes of the 2,3-BDL pathway play a crucial role, not only for the productivity but also for the product purity. Overexpression of butanediol dehydrogenase and glyceraldehyde-3-phosphate dehydrogenase in *B. amyloliquefaciens* significantly increased the molar yield and productivity, while also reducing the undesired pathway intermediate acetoin [42]. Similar effects were observed by decoupling the initial

step of the butanediol biosynthesis pathway. Already in 1988, Mallonee and Speckmann generated a *P. polymyxa* strain constitutively expressing the gene *alsS* by chemical mutagenesis [52]. Accordingly, 2,3-BDL production started during early log phase and a 4-fold elevated product titer compared to the wildtype strain was observed. Small amounts of meso-BDL are typically attributed to the spontaneous decaboxylation of α-acetolactate to diacetyl, which is further converted to acetoin by diacetyl reductase. However, thereby the chiral information is lost resulting in the less valuable optically inactive form of 2,3-BDL. Zhang et al (2018) demonstrated that the already high enantiomeric purity of 2,3-BDL produced by *P. polymyxa* could be even further increased by deletion of a diacetyl reductase [53].

Another interesting strategy represents the heterologous expression of microbial hemoglobin. The most common variant VHb of the bacterium *Vitreoscilla* sp. has demonstrated positive effects on the growth and viability of numerous host organisms that consequently led to improved production of the desired product [54]. For butanediol this effect could be shown in *K. oxytoca* as well as *E. aerogenes*, in which the episomal expression improved 2,3-BDL formation by approximately 30 % [55,56]. While the exact underlying mechanism is not fully understood, it is speculated that the hemoglobin like protein facilitates oxygen supply in microaerobic conditions and thereby promotes biomass formation.

Due to the inaccessibility of *P. polymyxa* to genetic engineering, most research using this organism was focused on the optimization of process parameters and media composition. However, recently a CRISPR-Cas9 based tool designed for *P. polymyxa* was developed, allowing more complex approaches for the rational metabolic engineering of this strain [25].

1.3. Microbial exopolysaccharides

Polysaccharides represent an extremely diverse class of carbohydrate polymers found in all domains of life [57]. Depending on the location of the polymers the biological function can be quite diverse. Intracellular polysaccharides are often used as energy storage (e.g. glycogen). Structural polysaccharides such as cellulose and lipopolysaccharides play an important role as integral parts of the cellular integrity of plants and microbes, while extracellular polysaccharides (EPS) cover numerous tasks depending on their physicochemical properties [58]. For example, microbial EPS serve as natural adhesives in biofilm formation and are crucial parts of the host-pathogen interaction [59] or act as barriers against abiotic stress factors such as extreme pH, antibiotics, temperature and desiccation. This list of physiological functions represents only a small overview of their natural tasks. Depending on their properties, these biopolymers have also shown industrial success in food, feed, cosmetic, pharmaceutical and technical applications or being used as rheological modifiers for example in oil recovery and construction science or as blood volume expanders in medicine [60]. The key reason for these multi-faceted fields of application lies in the enormous structural diversity of EPS found in nature. These biopolymers consist of charged and uncharged glycosidic repeating units that can be further decorated by organic and inorganic chemical moieties such as acetate, pyruvate, succinate or sulfate, resulting in a sheer endless number of potential structures [61].

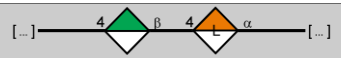
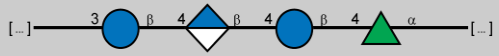
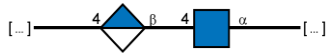
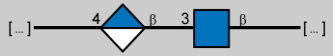
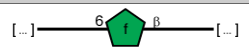
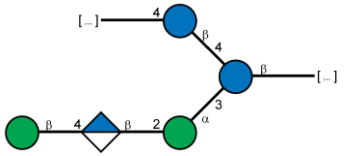
The global market is still dominated by plant and algae derived biopolymers such as starch, pectin, alginates and carrageenan [62]. However, bacterial exopolysaccharides like xanthan from *Xanthomonas campestris*, different sphingans from *Spingomonas* spp. [63] or hyaluronic acid produced by *Streptococcus* spp. and recombinant expression systems [64] have found commercial success as bulk products as well as in high-value niche applications. Contrary to their plant and algae derived counterparts, bacterial EPS can be produced by controlled large scale fermentation in bioreactors and are thereby independent of seasonal or geopolitical influences. In addition, due to the highly specific enzymes involved in the biosynthesis of these polymers, microbial EPS are organized in highly regular and uniform repetitive monomer sequences – the so-called repeating unit [65]. While it requires deep understanding of the underlying biosynthesis machinery, genetic engineering further allows rational tailoring of polysaccharide structures to modify their physicochemical properties towards specific applications [25,66]. An overview of commercially produced and applied bacterial EPS variants is given in Table 2.

Over the last 50 years many *Paenibacillus* spp. have been reported to produce variable exopolysaccharides with distinct physicochemical and rheological properties that could be of potential commercial interest for a variety of different applications [67,68]. Particularly isolates of *P. polymyxa* demonstrated to produce both, a levan-type polyfructan as well as

one or more heteropolysaccharides depending on the chosen cultivation conditions [69]. Levans produced by *P. polymyxa* vary in their branching pattern as well as their molecular weight, resulting in different rheological properties [70]. On the opposite, the heteroexopolysaccharides are reported to be more diverse. Most described monomer compositions comprise similar building blocks including glucose, mannose, galactose and glucuronic acid [67]. However, more complex EPS compositions additionally containing rhamnose, xylose or fucose were described [71–74]. Interestingly, despite the long history of research of the polysaccharides produced by *P. polymyxa*, no unambiguous structure of any heteroexopolysaccharide variant has been yet elucidated. Reasons for this lack of knowledge might be the insufficient separation of different polymer fractions as well as the high structural complexity of the polysaccharide, which has been demonstrated to contain multiple branching side chains [75].

The structural diversity and consequently different physicochemical properties of these polymers further resulted in a variety of proposed applications [76]. A heteroexopolysaccharide named paenan produced by *P. polymyxa* 2H2 demonstrated outstanding performance as a rheological modifier in surfactant systems and surpassed other commercial EPS variants like xanthan and gellan [24]. Further interesting (bio)activities include antioxidative properties [77,78], anti-tumor activity [79], as well as high water-binding capacity [80], which make the polysaccharides from *P. polymyxa* interesting for high value medical and cosmetic applications. In addition, the usage as a flocculant in bioremediation was proposed due to their bioabsorption ability of heavy metals [81,82]. Furthermore, it proved possible to artificially modify the structure of paenan by CRISPR-Cas9 mediated genome engineering resulting in an inherently different rheological behavior [25]. This strategy demonstrated for paenan allows to tailor the structure and hence physicochemical properties towards specific applications and could be transferred to other polysaccharides.

Table 2: Overview of bacterial exopolysaccharides and commercialized fields of application

EPS	Structure[a]	Organism	Field of application	Reference
Alginate		*Pseudomonas aeroginosa*	Food, medicine, cosmetics	[83]
Cellulose		*Escherichia coli, Komagotobacter* spp.	Food, medicine	[84]
Curdlan		*Alcaligenes faecalis var. myxogenes*	Food, medicine, cosmetics	[85]
Dextran		*Leuconostroc* spp., *Lactobacillus* spp.	Medicine	[86]
Gellan		*Sphingomonas elodea*	Construction chemistry, food, feed	[87]
Heparosan		*B. megaterium*[b]	Medicine	[88]
Hyaluronan		*Streptococcus* spp. *E. coli*[b], *B. subtillis*[b],	Medicine, cosmetics	[89]
Levan		*Bacillus* spp.	Food, medicine, cosmetics	[90]
Pullulan		*Aureobasidium pullulans*	Food, medicine, packaging	[91]
Xanthan		*Xanthomonas campestris*	Food, medicine, cosmetics, technical applications	[66]

a) Structures are depicted in the SNFG - format including their glycosidic linkage and conformation, but without decorations. green diamond: ManA; orange diamond: GulA; blue circle: Glc; blue diamond: GlcA; green triangle: Rha; blue rectangle: GlcNAc; green pentagon: fructose; green circle: Man b) recombinant strain

1.3.1. Biosynthesis of exopolysaccharides

EPS biosynthesis is a common feature among many organisms in all domains of life. While the structural diversity and consequently physicochemical properties of polysaccharides are enormous, the involved biosynthesis pathways are limited [65]. At present four different pathways involved in the assembly of EPS are characterized [92–95] 1) extracellular sucrase dependent pathway 2) Synthase dependent pathway 3) Wzx/Wzy pathway 4) ABC transporter dependent pathway.

Despite fundamental differences in the underlying enzyme machinery of these pathways apart from the sucrase dependent pathway, most key steps in the assembly of the polysaccharides remain quite similar [95]. The main building blocks of EPS consist of nucleotide-sugars. One or more glycosyltransferases (GTs) use the chemical binding energy of activated nucleotide precursors for the intracellular, stepwise assembly of either oligosaccharide repeating units or the entire polymer. Furthermore, the EPS chain is polymerized and secreted by various mechanisms towards the extracellular space. Many ABC-transporter dependent pathways also loosely link the secreted polysaccharide to the outer membrane of Gram-negative bacteria, which makes a clear discrimination between EPS and capsular polysaccharides (CPS) difficult [93].

Contrary to the other described pathways, sucrases are secreted to the extracellular space or bound to the outer membrane and polymerize the entire polysaccharide extracellularly. Instead of nucleotide-sugars, di- or trisaccharides such as sucrose, raffinose or nigerose are used as a substrate [65]. One sugar moiety is cleaved by the sucrase and transferred to a nascent polysaccharide chain [96]. Depending on the substrate and utilized sugar residue, homoglucans or homofructans such as dextran or levan are formed. Despite a seemingly simple system, different sucrases and also reaction conditions result in a variety of physiochemically diverse polysaccharides with variable molecular weight and branching pattern [90].

In the synthase dependent pathway, a multiprotein complex spanning the inner membrane simultaneously polymerizes and translocates the nascent polymer chain towards the periplasm. Further outer membrane polysaccharide exporters (OPX) realize the secretion of the polymer. Most homopolysaccharides such as curdlan or cellulose are produced by the synthase dependent pathway [93]. Interestingly, alginate, while strictly classified as a heteropolysaccharide, is first produced as a homopolymer solely consisting of mannuronic acid (M-blocks). Further epimerases in the periplasm or the extracellular space catalyze the conversion to guluronic acid (G-block) in a block like manner, while the AlgIJFX complex acetylates certain residues [97]. The monomer composition and decoration

ultimately determine the physicochemical properties of the obtained EPS. Synthase dependent EPS production is often directly regulated by the second messenger c-di-GMP. Both, the main catalytic subunit BcsA and secondary cytosolic subunit BcsE of bacterial cellulose synthase as well as Alg44 of the alginate synthase complex in *P. aeruginosa* contain c-di-GMP binding motives and are positively regulated by high concentrations of the signal molecule [98,99].

In the ABC-transporter dependent pathway the entire polysaccharide chain is assembled by one or more GTs at a lipid carrier at the inner membrane to form homo- or heteropolysaccharides [92]. Consequently, the name giving ABC-transporter catalyzes the ATP-dependent export of the EPS towards the periplasmic space. Additional proteins of the PCP and OPX family are required for the secretion. ABC transporter dependent pathways are often found to produce CPS. In the process, the final polysaccharide remains (loosely) attached to the cells and contributes to the pathogenic nature of many organisms, such as e.g. *Klebsiella* spp. [100].

Some of the most complex and diverse polysaccharides are produced via the Wzx/Wzy-pathway. Highly specific membrane-bound or cytosolic GTs assemble a nascent repeating unit at an undecaprenyl-phosphate lipid anchor. After translocation towards the periplasmic space by a Wzx flippase, multiple repeating units are polymerized by Wzy proteins before being exported by OPX proteins [94]. In addition to the GTs, the name-giving Wzx/Wzy proteins only accept a narrow substrate range of modified repeating units [101]. While it is possible to modify or truncate possible side chains, the complete removal of the branching residues is often detrimental to the assembly machinery [102,103]. Commercially successful EPS variants produced by the Wzx/Wzy pathway are for example xanthan and different sphingans.

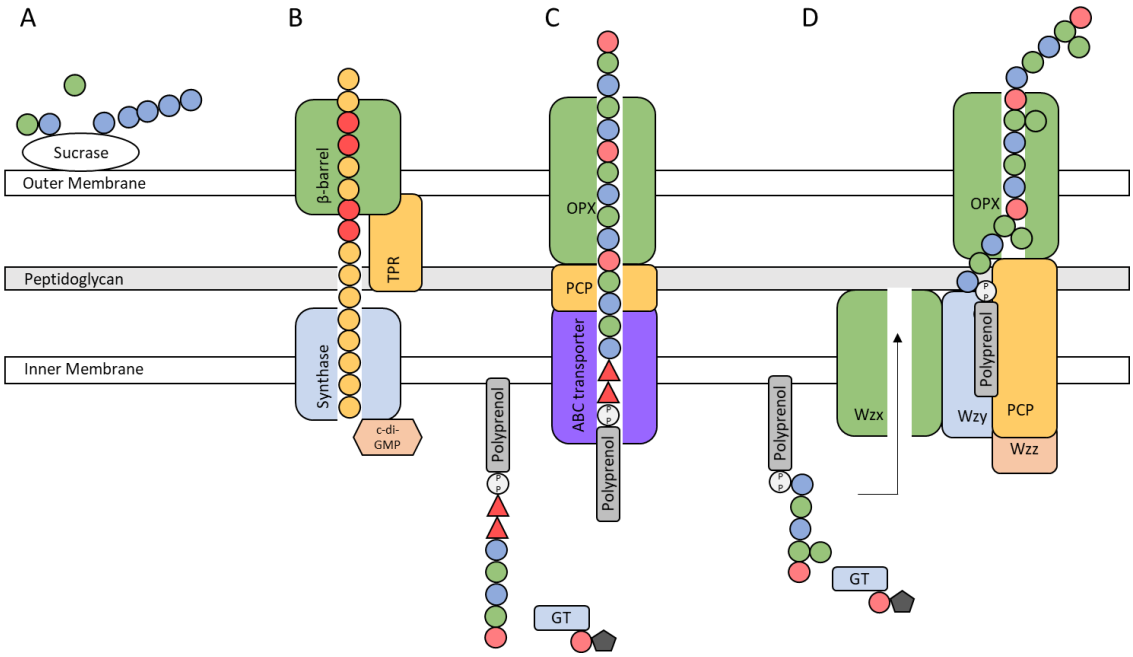


Figure 2: Schematic overview of the four major biosynthesis pathways for the production of EPS. A) Sucrases are bound to the outer membrane or extracellular and cleave di- or oligomeric carbohydrate substrates and use the energy for the polymerization of homopolysaccharides. B) In the synthase dependent pathway, a polymer is assembled and simultaneously exported by membrane-spanning multiprotein complex. Enzymes such as epimerases and sulfotransferases can further modify the nascent polysaccharide in the periplasmic space or the extracellular environment. Synthase dependent polysaccharides are often regulated by the secondary messenger c-di-GMP, which directly interacts with the synthase complex or a co-polymerase. C) In the ABC-transporter dependent pathway, intracellular GTs assemble the entire nascent polysaccharide at a lipid anchor via a Kdo-linker (red triangle) at the inner membrane, which is further secreted towards the cell environment by the concerted action of ABC-transporters and OPX-proteins spanning the inner and outer membrane. D) In the Wzx/Wzy pathway, oligomeric repeating units of a polysaccharide are assembled at a C55-lipid anchor at the inner membrane before being translocated towards the periplasmic space by a Wzx-flippase. Afterwards the Wzy protein polymerizes the polysaccharide before secretion. Adapted from Schmid (2018) [65]. Circle with pentagon: activated sugar nucleotide, TPR: tetratricopeptide repeat proteins; OPX: outer membrane polysaccharide export; PCP: polysaccharide co-polymerase

Interestingly, the genes encoding the functional elements of the EPS biosynthesis are typically found clustered in the genome of the host organism and comprise the genetic information of glycosyltransferases, polymerases, export proteins and precursor synthesis [104,105]. Due to their importance in other metabolic pathways, often multiple copies of genes involved in the synthesis of nucleotide precursors can be found in the genome. However, some organisms also encode multiple EPS clusters involved in the biosynthesis of different polysaccharides. For example, genome analysis of the plant pathogenic *Clavibacter michiganensis* LMG 26808 revealed four EPS clusters, which play an important role in the virulence of these bacteria [106]. Abiotic signals such as the ratio of carbon and nitrogen in the cell environment, phosphate concentration or oxygen availability are crucial in the transcriptional regulation of the corresponding gene clusters

and influence the EPS composition [69]. Bioinformatic tools allow the relatively easy identification of novel EPS clusters based on the genomic context and domain annotation of genes [107]. However, further biochemical analysis in the wet lab is required to assign specific functions to individual genes of a cluster. Particularly the annotation of structurally diverse GTs demonstrated to be challenging and often lacks unambiguous experimental evidence [108,109].

1.3.2 Engineering of functional polysaccharides

In order to improve production of bacterial polysaccharides with industrial interest, most approaches focus on process optimization to overcome limiting bottlenecks. Common strategies aim to improve mass transfer, which is impaired by highly viscous fermentation broth or optimizing the media composition to identify ideal conditions for EPS production [110]. In addition, metabolic engineering was applied to increase the biosynthesis of nucleotide precursors or remove undesired side-products affecting the quality of the polymer (e.g. pigments) [54,111]. Recombinant expression of the bio-machinery involved in EPS synthesis allowed for efficient production in microbial cell factories [88,112]. Particularly for glycosaminoglycans (GAGs), which naturally occur in higher animals, this strategy demonstrated a reliant way to produce these polysaccharides for high-value pharmaceutical and cosmetic applications [113]. Furthermore, genome engineering was used to knock-out GTs and enzymes involved in the chemical decoration of these biopolymers [25,66,114]. Consequently, for some well-studied biopolymers, artificial variants could be created with modified repeating units, molecular weight and altered decoration patterns resulting in a completely different rheological behavior. However, more knowledge about fundamental mechanisms of the underlying biosynthesis machinery is required for new polysaccharides. With the characterization of new polysaccharides and discovery of structure-functional relationship of these biopolymyers, it will be possible to expand the portfolio commercially feasible polysaccharides and rationally tailor EPS for distinct applications.

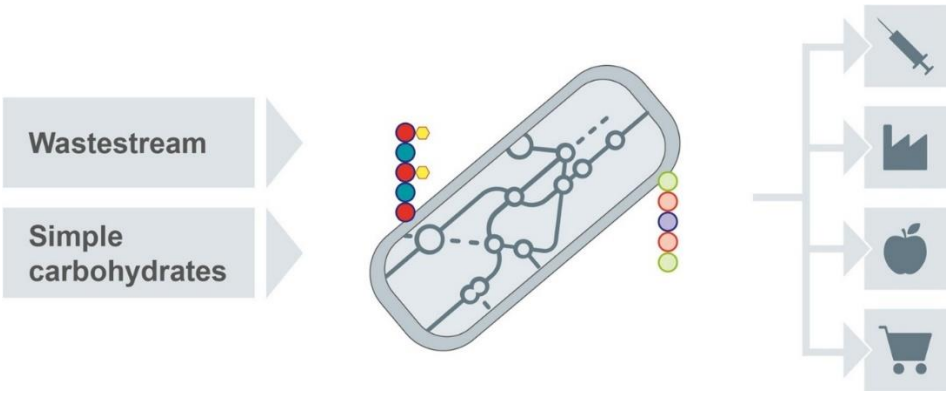
A concise review of the latest achievements in this field is presented in the following publication. Abinaya Badri and Christoph Schilling focused on literature research and writing of the initial draft. Abinaya Badri focused on glucosaminoglycans and metabolic engineering of appropriate production organisms. Christoph Schilling focused on bacterial heteroexopolysaccharides and homopolysaccharides, metabolic engineering strategies and different production pathways. Christoph Schilling prepared all tables and figures of the initial manuscript in close collaboration with Jochen Schmid. Volker Sieber, Mattheos Koffas and Jochen Schmid provided scientific advice. All authors contributed to content and language of the manuscript.

Metabolic engineering for production of functional polysaccharides

Christoph Schilling, Abinaya Badri, Volker Sieber, Mattheos Koffas & Jochen Schmid

Current opinion in Biotechnology

(2020)



DOI: 10.1016/j.copbio.2020.06.010

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Metabolic engineering for production of functional polysaccharides

Christoph Schilling[1,5], Abinaya Badri[2,5], Volker Sieber[1], Mattheos Koffas[2,4] and Jochen Schmid[3]

Functional carbohydrate polymers are of immense industrial interest for high value applications. Distinct biosynthetic pathways allow for metabolic engineering approaches for production in microbial cell factories. The most common strategies in recent years included the attenuation of central carbon metabolism, improved substrate utilization or enhanced intracellular sugar nucleotide precursor levels. Recombinant expression in more suitable surrogate host organisms has demonstrated remarkable results for the heterologous production of glycosaminoglycans. However, industrial application of pharmacological active functional polysaccharides is often limited by costly post-polymerization modifications and downstream processing. With increasing knowledge of bottleneck enzymes and fluxes, it will be possible to enable a sustainable microbial production of high value polysaccharides and tailor artificial polymers towards specific applications.

Addresses
[1] Chair of Chemistry of Biogenic Resources, Technical University of Munich, Campus for Biotechnology and Sustainability, 94315 Straubing, Germany
[2] Department of Chemical and Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[3] Norwegian University of Science and Technology, Department of Biotechnology and Food Science, Sem Sælands vei 6-8, 7491 Trondheim, Norway
[4] Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Corresponding authors: Koffas, Mattheos (koffam@rpi.edu), Schmid, Jochen (jochen.schmid@ntnu.no)
[5] Authors contributed equally.

Current Opinion in Biotechnology 2020, **66**:xx–yy

This review comes from a themed issue on **Tissue, cell and pathway engineering**

Edited by **Peng Xu**, **Li Tang** and **Haoran Zhang**

https://doi.org/10.1016/j.copbio.2020.06.010

Functional polysaccharides in today's world

Introduction and scope

Functional polysaccharides (FPs) are a class of compounds widely found in nature with diverse applications. Some well-known examples include alginates, xanthan, heparin, chondroitin sulfate and so on. Consequent to their diverse chemistries, the range of biological functions and applications of FPs are vast. They serve as natural adhesives in biofilm formation, host pathogen interaction or as barriers against environmental stress factors such as extreme pH and desiccation [1•]. Depending on their properties these biopolymers are vital in food, feed, cosmetic, medical and technical applications and a growing number of polysaccharides continue to be elucidated (Table 1) [2].

Exopolysaccharides (EPS) are high molecular weight carbohydrate polymers that are secreted by microorganisms into the surrounding medium during their life cycle. Glycosaminoglycans (GAGs) are a specific type of polysaccharides that are composed of hexuronic acids and hexosamines. Microbial EPS are used as bulk compounds in food, feed and technical applications, whereas most GAGs are utilized in high-value medical and cosmetic applications due to their attractive pharmacological properties. Though GAGs are found predominantly in higher eukaryotes, this review will focus on EPS and GAGs production by microbial hosts and will discuss new avenues which can be explored in that field by metabolic engineering and utilizing tools from ancillary studies.

EPS types and functions

EPS in general, consist of charged and uncharged glycosidic repeating units that can be modified by mainly acetate, pyruvate, succinate, sulfate and glycerol. They can be categorized by their monomer composition as homo-exopolysaccharides or hetero-exopolysaccharides or by the underlying synthesis pathways (Figure 1). While the various biosynthesis pathways of EPS are generally unraveled as previously reviewed [3], the precise function of individual genes has only been explored in some EPS like xanthan or alginates [4,5] which can be used for targeted functional modifications.

GAGs types and functions

GAGs play an important role as part of the extracellular matrix mainly in combination with proteins (proteoglycans) [6]. Their functional groups enable specific interactions with proteins, growth factors and hormones to realize essential physiological effects [7,8]. This functionality determines their high-value applications in the pharmaceutical industry, such as for heparins as

Table 1

Overview of commercially available polysaccharides, common sources and fields of application

Polymer	Natural source	Microbial cell factory	Composition	Application
Alginate	Brown seaweed	*Pseudomonas* spp., *Azotobacter* spp	ManA, GulA, Ac	Food, feed, medicine, cosmetics
Cellulose	Plants, *Acetobacteriacae*	*Komagataeibacter rhaeticus, E. coli*	Glc	Food, medicine
Chondroitin	Porcine trache, shark cartilage	*E. coli*	GalNAc, GlcA	Pharmaceuticals
Curdlan	*Agrobacterium* spp.	*Alcaligenes faecalis var. myxogenes IFO13140*	Glc	Food, feed, medicine, cosmetics
Dermatan	Bovine Mucose	n.a.	GalNAc, GlcA, IdoA	Pharmaceuticals
Diutan	*Sphingomonas* spp.	*Sphingomonas* sp. ATCC53159	Glc, Rha, GlcA, Ac	Construction chemistry
Gellan	*S. elodea*	*S. elodea* ATCC 31,461	Glc, Rha, GlcA, Ac, Gly	Construction chemistry, food, feed
Heparosan	Porcine intestine	*B. megaterium*	GlcA, GlcNAc, IdoA	Pharmaceuticals
Hyaluronan	Rooster comb	*E. coli*	GlcA, GlcNAc	Cosmetics
Levan	*Bacillus* spp., *Lactobacillus* spp.	*B. amyloliquefaciens*, *B. subtilis* (natto) CCT7712	Fru	Food, feed, medicine, cosmetics
Pullulan	*Aureobasidium pullulans*	*Aureobasidium pullulans*	Glc	Food, feed, packaging
Scleroglucan	*Sclerotium spp.*	*S. rolfsii*, *S. glucanicum*	Glc	Food, pharmaceuticals, oil drilling
Succinoglycan	*Agrobacterium* sp., *Pseudomonas* sp.	*Sinorhizobium meliloti*	Glc, Gal, Suc, Ac, Pyr	Cosmetics
Welan	*Sphingomonas* spp.	Sphingomonas sp. ATCC31555	Glc, Rha, GlcA, Man	Construction chemistry
Xanthan	*Xanthomonas* spp.	*X. campestris pv. campestris*	Glc, Man, GlcA, Ac, Pyr	Food, feed, oil drilling, cosmetics

Adapted from Schmid *et al.* [39]; ManA: D-mannuronic acid, GulA: L-guluronic acid; Ac: acetate; Glc: D-glucose; GalNAc: N-acetyl-D-galactosamine; GlcA: D-glucuronic acid; IdoA: L-Iduronic acid; Rha: L-rhamnose;; Gly: glycerate; GlcNAc: N-acetyl-D-glucosamine; Fru: D-fructose; Gal: galactose; Suc: succinate; Pyr: pyruvate; n.a.: not available.

anticoagulants. There are four major classes of GAGs - heparin/heparan sulfates (HS), chondroitin/chondroitin sulfates (CS), keratan sulfates (KS) and hyaluronic acids (HA). GAGs are mainly composed of uronic acid derivatives like glucuronic/iduronic acids and amine derivatives like *N*-acetylglucosamine/*N*-acetylgalactosamine. Residues can be sulfated at different positions giving rise to subtypes of GAGs (Figure 1). The biosynthetic routes of GAGs via the synthase-dependent pathway are well-studied, and the enzymes associated with these pathways are very well characterized [9•]. GAGs have several important and interesting pharmaceutical applications. Heparin is a potent anticoagulant and CS is widely used in the treatment of osteoarthritis. Moreover, several new pharmaceutical roles of GAGs including anti-inflammatory and anticancer effects, and nervous system recovery are being uncovered [10]. The main source of GAGs is still extraction from animal tissues, although animal-free alternatives including chemical, chemoenzymatic and bioengineering methods using heterologous hosts are being developed. Moreover, others have identified many GAG-producing marine sources for sustainable extraction (Table 2). The huge interest in microbial GAGs production is based on the limitations of extraction from animal resources and to circumvent problems such as contamination with other animal proteoglycans. Even though heterologous production by microbial cell factories is very promising, the economic feasibility is not realized yet, and further optimization of the production is still required (Figure 2).

Metabolic engineering for production of functional polysaccharides

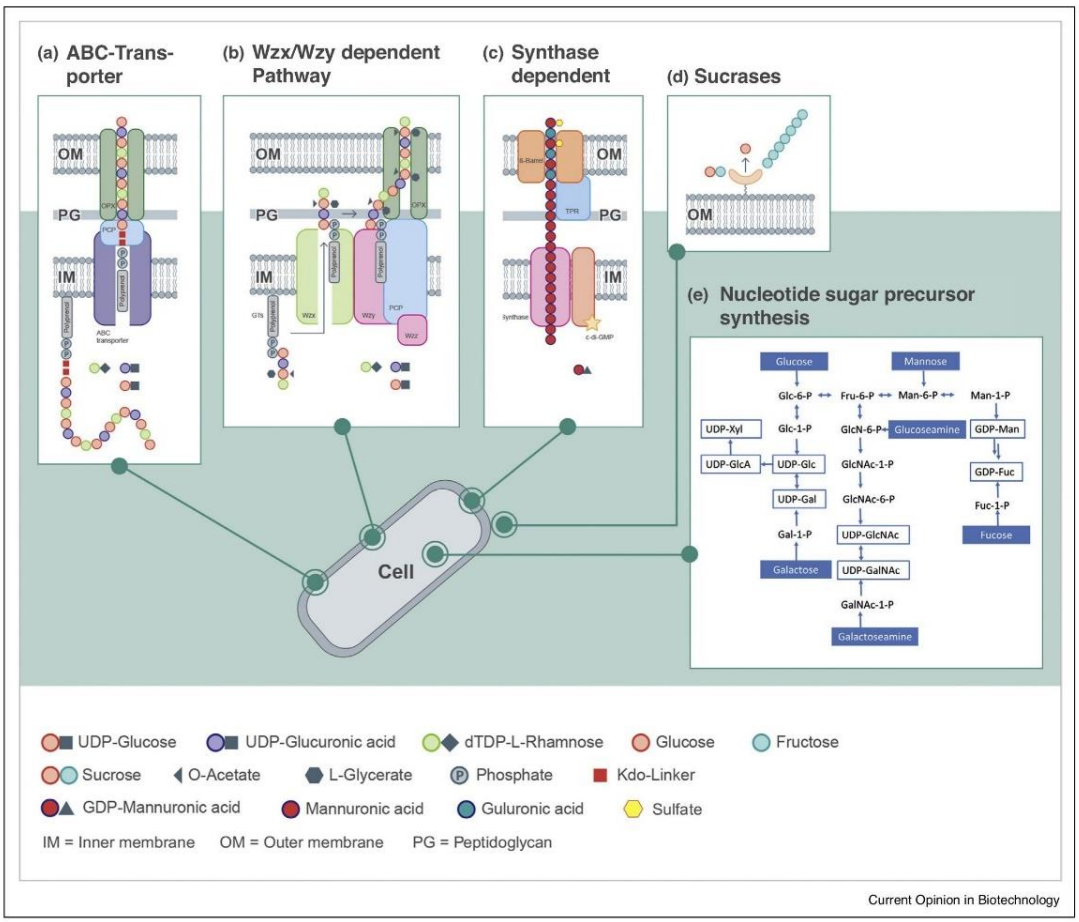
FPs for technical and food applications are applied as a bulk compound with prices in the range of 5–15 US$/kg. In contrast, commercialized pharmaceutical GAGs mainly rely on animal sources and thus, expensive extraction as well as purification steps to guarantee the necessary purity. Consequently, metabolic engineering approaches using microbial cell factories are highly attractive but need to cover the necessary amounts of FPs with distinct features. Predominantly applied engineering strategies are based on:

1) Improving nucleotide-precursor availability
2) Recombinant expression in suitable microbial cell factories
3) Optimizing substrate utilization
4) Attenuation of competing central carbon fluxes

In the next section we want to highlight proceedings of the recent years, in which these engineering strategies for the development of economically feasible production strains for bulk and high-value FPs were applied.

Engineering of the biosynthesis of homopolysaccharides such as levan or dextran relies on single enzymes that extracellularly polymerize di- or oligosaccharides as substrates. For *Bacillus amyloliquefaciens*, levan titers were

Figure 1



Overview of different biosynthesis pathways for microbial polysaccharide synthesis. **(a)** The ABC-transporter dependent pathway, glycosyltransferases assemble the polysaccharide anchored on a 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo), which is afterwards secreted by an ABC transporter. Subsequently, PCP and OPX proteins ensure the transport across the cell wall. **(b)** In the Wzx/Wzy pathway, multiple highly specific glycosyltransferases assemble a defined repeating unit of heteropolysaccharides on a polyprenol lipid anchor from nucleotide-activated sugar precursors. This repeating unit is translocated towards the periplasmic space by a Wzx flippase and polymerized by a Wzy protein before being exported. **(c)** A multiprotein synthase complex realizes the polymerization and simultaneous translocation of GAGs and homopolysaccharides such as curdlan, cellulose or alginate through the inner membrane. In the periplasmic space, epimerases modify the homopolymer in a block like manner in the case of alginate before the export through the outer membrane. For sulfated GAGs this will be highly specific epimerases and sulfotransferases. **(d)** Extracellular sucrases use the chemical binding energy of digosaccharide or oligosaccharide substrates to polymerize homopolysaccharides. **(e)** Intracellular EPS biosynthesis requires activated sugar nucleotide precursors for the polymerization of nascent polysaccharides. Sugar nucleotides are produced from simple saccharides of the central carbohydrate metabolism.

increased to 102 g/L in fed-batch fermentations by exchanging the natural promoter of the levansucrase SacB with the strong artificial promoter Pgrac. Additionally, the signal peptide was swapped to the YncM peptide from *Bacillus subtilis* to enhance secretion efficiency [11] and thus displaying the potential of this approach.

Further metabolic engineering approaches to increase the EPS titer included inhibition of lipopolysaccharide synthesis in *Xanthomonas campestris*. Consequently, intracellular nucleotide sugar pools were increased resulting in a 26 % increased EPS-yield [12]. In addition, deletion of *pigA*, involved in xanthomonadin synthesis, and

Table 2

Newly identified GAGs and their functions

GAG	Source	Function	Notes	Ref.
Fructosylated chondroitin	*E. coli* K4	Immune-stimulating effects	Wildtype K4 - partially fructosylated; Recombinant K4 - fully fructosylated	[40]
Crude Polysaccharides	Srikankan marine algae (*Chnoospora minima*)	Antioxidant	–	[41]
Fucosylated CS	Sea Cucumber (*Holothuria Sp.*)	Immuno-modulatory effects; hematopoietic function recovery	Activity is dependent on fucose sulfation pattern	[42]
	Sea Cucumber (*Apostichopus japonicus*)	Anti-hyperglycemic agent; Anti- insulin resistance	Effect via AMPK pathway activation	[43]
	Sea Cucumber (*Stichopus japonicus*)	Anti-obesity effect	Effect via regulation of gut microbiota	[44]
	Sea Cucumber (*Cucumaria djakonovi*)	High anti-inflammatory effect	Highly sulfated; higher anti-inflammatory effect than vertebrate CS	[45]
	Sea Cucumber (*Holothuria Mexicana*)	Growth factor binding; Anticoagulant	GlcA and GalNAc fucosylation; oligosaccharide fucosyl branches	[46]
AF-GAG	Edible Snail (*Achatina fulica*)	Not identified	Uniquely regular (IdoA(2S)-GlcNAc) disaccharides	[47]
CS	Skate (*Raja pulchra*)	Lipase-inhibitor; Adipocyte-inhibitor;	Oligosaccharides are easily absorbed *in vivo* → more effective	[48]
Insect GAG	Dung Beetle (*Catharsius molossus*); Bumble Bee (*Bombus ignitus*); Cicada (*Huechys sanguinea*)	Anticancer effect; Anti-aging effect	Major residues – *N*-acetylglucosamine and D-glucosaminic acid;	[49,50]
CS	Chicken Keel Bone	Antioxidant; Emulsifying agent	–	[51]
CS	*Homo sapiens*	Protease inhibition	CSA -urinary trypsin inhibitor	[52]

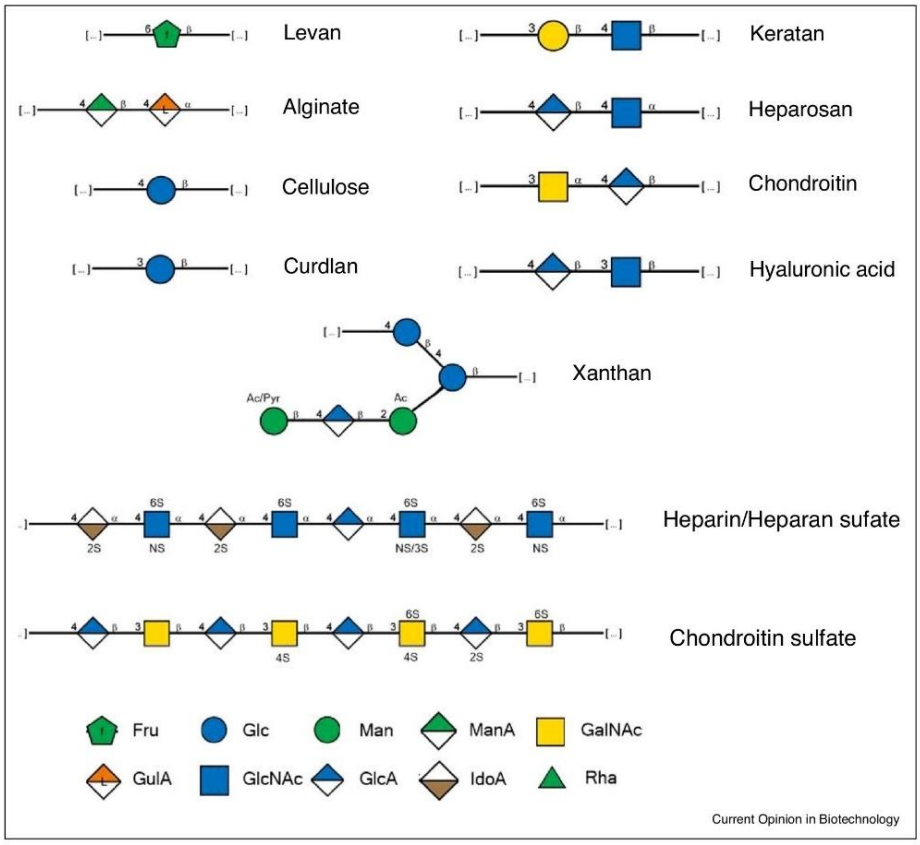
simultaneous heterologous expression of bacterial globin from *Vitreoscilla* sp. enhanced the metabolism under oxygen limiting conditions, while also improving the purity of the obtained EPS [13•].

Since native hyaluronic acid (HA) does not have any post polymerization modifications, it is the easiest GAG to produce recombinantly [14,15••,16]. Yu, Stephanopoulos *et al.* improved HA titers in an engineered *Corynebacterium glutamicum* strain by application of genome-scale metabolic modeling [15••]. They expressed HA pathway genes in a *C. glutamicum* strain with blocked pentose phosphate pathway and attenuated glycolysis. Additionally, they redirected the complete pyruvate flux to succinate and then attenuated the succinate pathway. By employing this strategy, they increased HA titers to 28.7 g/L from 1.3 g/L [15••].

Co-substrate utilization has also been explored as a strategy recently in *Escherichia coli*. By that Jang, Lee *et al.* improved HA titers 3.61-fold via utilization of glucose and galactose and circumventing glycolysis and pentose phosphate pathway [14]. Another study used a genome-scale model of *Lactococcus lactis* to identify inosine supplementation for improved HA production [17]. This substrate was also co-utilized with glucose and improved HA titers by 2.8-fold to 1.1 g/L. Other reports by this group describe addition of acetate to augment acetyl-coA production [18] and adopting a glucostat (to maintain the glucose level in fermentation medium) [19] to increase and maintain molecular weight throughout the fermentation time.

Metabolic engineering for HS and CS syntheses are lagging behind the strategies for HA. However, in the last three years, metabolic engineering approaches, such as glucose dehydrogenase overexpression or lactate dehydrogenase deletion, which have previously been demonstrated to improve HA production, have been adopted for heparosan and chondroitin production as well. Yu *et al.* demonstrated about 2 g/L chondroitin production from a new microbial host *C. glutamicum* upon adoption of the above strategies [20]. Another important requirement for heparosan or chondroitin synthesis is the availability of the two UDP-sugar precursor molecules that are condensed to form the repeating disaccharide unit. Balancing the availability of these UDP sugars is an established strategy in metabolic engineering of HA [21]. In an analogous study, Wu *et al.* report very high productivity of fructosylated chondroitin upon balancing UDP precursor pathways [22].

Figure 2



Overview of the structures of some functional polysaccharides. Polysaccharides are typically produced *in vivo* by microbial cell factories from simple carbohydrate substrates such as glucose. More complex glycosaminoglycans like heparin or chondroitin sulfate require further post-polymerization modification by highly specific epimerases and sulfotransferases to mediate specific functional roles. Particularly for heparan sulfate and chondroitin sulfate, residues are sulfated in different carbon positions giving rise to different subtypes.

Novel microbial hosts have also been engineered towards heparosan production [23,24••]. In a recent work, Koffas *et al.* heterologously expressed the heparosan synthase enzymes from *Pasteurella multocida* in *Bacillus megaterium* and demonstrated production of 2.74 g/L heparosan via a fed-batch fermentation [23]. In another study, this group also integrated heparosan synthase and glucose phosphate uridyltransferase in the genome of *Synechococcus* to demonstrate photoautotrophic synthesis of heparosan [24••]. Cimini, Restaino *et al.* have initiated one of the first studies of elucidating the molecular weight of the heparosan or chondroitin obtained from engineered *E. coli* [25]. They also showed higher molecular weights fractions from engineered strains overexpressing the positive gene regulator *rfaH* [26].

Limitations and perspective

Rational engineering strategies are not always evident, and bottlenecks can vary in different host organisms. Consequently, expression of individual pathway enzymes needs to be optimized individually for each approach [27••]. For recombinant strains, limited rare precursor nucleotide sugars offer promising targets to be engineered [5,22,25], however, depending on pathway bottlenecks this method cannot be universally applied for all FPs and production strains. UDP-glucose as used in

central metabolic pathways is typically supplied in sufficient amounts, while rare uronic acid precursor nucleotides are often limiting

While the general biosynthesis pathways are well-understood (Figure 1), the precise mechanism and mode of action of individual enzymes for new FPs are often hardly characterized. Particularly the function of hydrolases, that determine the molecular weight of FPs, are still only poorly understood, but are important for both pharmacological and technical applications. For example, dextran can only be applied as a clinical blood volume extender in a narrow range of molecular weight and can otherwise cause negative effects on coagulation [28]. An interesting example in which chain-length modifying enzymes were explored are alginates from *Pseudomonas* sp. or *Azotobacter* sp. Although not essential for the biosynthesis, overexpression of a putative alginate lyase has shown to reduce the molecular weight of alginate and increase the degree of *O*-acetylation, while deletion of the lyase increased the EPS yield [29].

Flippases that translocate nascent repeating units towards the periplasm can be often identified by bioinformatics means, however the mechanistic actions of different enzymes and their respective substrate spectrum remains to be elucidated. The most prominent EPS example of the Wzx/Wzy pathway is xanthan, from *X. campestris* (Figure 1). Both sidechain mannose residues of xanthan can be acetylated while the terminal residue can also be substituted with a pyruvate-ketal. Combinatorial deletion and overexpression of pyruvyltransferases and acetyltransferases allowed the production of polysaccharide variants with defined ratios of the respective decorations resulting in altered rheological behavior [30••,31]. Additionally, also truncated side chains were generated by targeted knock-out of the involved glycosyltransferases. However, due to the high specificity of flippases, modifications of the EPS structure often result in a significantly reduced yield of the polymer [32].

A major drawback in recombinant production of industrially relevant FPs of pharmacological interest is the necessity of distinct post-polymerization modifications that still remains to be addressed. Especially the enzymatic sulfatation as well as defined molecular weight adjustment represent the main challenges to realize economical competitive processes. The success of microbial metabolic engineering for the production of GAGs has relied on the presence of bacterial counterparts for GAG polymerases. However, there are no analogous bacterial GAG sulfotransferases and epimerases to carry out post polymerization modifications of sulfation and epimerization respectively [33]. These modifications contribute to the functional roles of GAGs in health and biology. This limitation in post polymerization processing enzymes has severely disrupted advances in direct production of functional GAGs through microbial metabolic engineering. However, we believe that advances in the expression, stability and solubility of glycosylated mammalian proteins in prokaryotic systems are a potential refuge to this limitation [34].

Highly sensitive analytical methods remain an important topic for all polymers with industrial applications and strong analytical techniques for differentiation, estimation and quality control of complex FP mixtures are required. Many FPs possess structural similarities but functional differences. Particularly for pharmacological applications nuances can influence their effectiveness. NMR fingerprinting and mass spectrometry are reliable methods which still must be optimized for complex FPs. Some new studies are focusing on improving the detection sensitivity by using standard HPLC equipment and fluorescence detectors [35]. Other studies employ methods like Fourier transform MS [36,37] and shotgun ion mobility MS [38]to resolve more complex oligosaccharide compositions.

With sufficient knowledge of enzymes involved in the biosynthesis of FPs, we envision a future in which it will be possible to tailor artificial polysaccharides towards specific applications. This approach allows harnessing optimized fermentation processes of commercially applied production strains to produce artificial FPs with new physicochemical properties, but will also need the development of novel chassis organisms.

Conclusion

Currently, for the bulk scale production of FPs, which are vastly applied in food and technical applications, mostly native production strains are used. Most studies focus on process optimization, which has to deal with highly viscous fermentation broths to increase product titers. However, we believe that the development of new genetic tools and better understanding of metabolic bottlenecks in these microorganisms will drastically increase metabolic engineering to enable industrially feasible production of FPs. For FPs of pharmacological interest, problems such as post-polymerization modifications must still be addressed.

New systems and synthetic biology approaches have demonstrated the capability to efficiently produce FPs in high titers in microbial cell factories as well as tailor them towards specific applications. Apart from highlighting recent studies, this review also presents a different point of view on the future of metabolic engineering in this field. With the development of enzymatic, analytical and regulatory tools, industry-scale production of many more natural and recombinant microbial FPs will be fast-approaching.

Conflict of interest statement

Nothing declared.

Acknowledgement

This work was supported by the U.S. National Science Foundation award CBET-1604547 awarded to MAGK.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. • Moradali MF, Rehm BHA: **Bacterial biopolymers: from pathogenesis to advanced materials**. *Nat Rev Microbiol* 2020 http://dx.doi.org/10.1038/s41579-019-0313-3.
This review summarizes in-depth the biosynthesis of bacterial biopolymers including polysaccharides and their potential application as biomaterials, while pointing out interesting targets for metabolic engineering as well as current challenges.

2. Toukach PV, Egorova KS: **Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts**. *Nucleic Acids Res* 2016, **44**:D1229-D1236.

3. Schmid J: **Recent insights in microbial exopolysaccharide biosynthesis and engineering strategies**. *Curr Opin Biotechnol* 2018, **53**:130-136.

4. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Pühler A, Ielpi L: ***Xanthomonas campestris* pv. campestrisgum mutants: effects on xanthan biosynthesis and plant virulence**. *J Bacteriol* 1998, **180**:1607-1617.

5. Maleki S, Mærk M, Hrudikova R, Valla S, Ertesvåg H: **New insights into *Pseudomonas fluorescens* alginate biosynthesis relevant for the establishment of an efficient production process for microbial alginates**. *New Biotechnol* 2017, **37**:2-8.

6. Ricard-Blum S: **Glycosaminoglycans: major biological players**. *Glycoconj J* 2017, **34**:275-276.

7. Ricard-Blum S: **Protein–glycosaminoglycan interaction networks: focus on heparan sulfate**. *Perspect Sci* 2017, **11**:62-69.

8. Kjellén L, Lindahl U: **Specificity of glycosaminoglycan–protein interactions**. *Curr Opin Struct Biol* 2018, **50**:101-108.

9. • Badri A, Williams A, Linhardt RJ, Koffas MA: **The road to animal-free glycosaminoglycan production: current efforts and bottlenecks**. *Curr Opin Biotechnol* 2018, **53**:85-92.
This review gives a comprehensive overview of all the different approaches for GAG production while pointing out major bottlenecks that lie in this field.

10. Lima M, Rudd T, Yates E: **New applications of heparin and other glycosaminoglycans**. *Molecules* 2017, **22**:749.

11. Gu Y, Zheng J, Feng J, Cao M, Gao W, Quan Y, Dang Y, Wang Y, Wang S, Song C: **Improvement of levan production in *Bacillus amyloliquefaciens* through metabolic optimization of regulatory elements**. *Appl Microbiol Biotechnol* 2017, **101**:4163-4174.

12. Steffens T, Vorhölter F-J, Giampà M, Hublik G, Pühler A, Niehaus K: **The influence of a modified lipopolysaccharide O-antigen on the biosynthesis of xanthan in *Xanthomonas campestris* pv. campestris B100**. *BMC Microbiol* 2016, **16**:93.

13. • Dai X, Gao G, Wu M, Wei W, Qu J, Li G, Ma T: **Construction and application of a *Xanthomonas campestris* CGMCC 15155 strain that produces white xanthan gum**. *Microbiol Open* 2019, **8**:e00631.
Dai *et al.* demonstrate the production of highly pure xanthan by deleting the xanthomonadin biosynthesis. The otherwise decreased yield was compensated by heterologous synthesis of bacterial hemoglobin Vgb.

14. Woo JE, Seong HJ, Lee SY, Jang Y-S: **Metabolic engineering of *Escherichia coli* for the production of hyaluronic acid from glucose and galactose**. *Front Bioeng Biotechnol* 2019, **7**:351.

15. •• Cheng F, Yu H, Stephanopoulos G: **Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid**. *Metab Eng* 2019, **55**:276-289.
The authors of this study used in silico model analysis and modified central carbon metabolism in *C. glutamicum* to produce 28.5 g/L.

16. Westbrook AW, Ren X, Oh J, Moo-Young M, Chou CP: **Metabolic engineering to enhance heterologous production of hyaluronic acid in *Bacillus subtilis***. *Metab Eng* 2018, **47**:401-413.

17. Badri A, Raman K, Jayaraman G: **Uncovering novel pathways for enhancing hyaluronan synthesis in recombinant *Lactococcus lactis*: genome-scale metabolic modeling and experimental validation**. *Processes* 2019, **7**:343.

18. Puvendran K, Jayaraman G: **Enhancement of acetyl-CoA by acetate co-utilization in recombinant *Lactococcus lactis* cultures enables the production of high molecular weight hyaluronic acid**. *Appl Microbiol Biotechnol* 2019, **103**:6989-7001.

19. Jeeva P, Shanmuga Doss S, Sundaram V, Jayaraman G: **Production of controlled molecular weight hyaluronic acid by glucostat strategy using recombinant *Lactococcus lactis* cultures**. *Appl Microbiol Biotechnol* 2019, **103**:4363-4375.

20. Cheng F, Luozhong S, Yu H, Guo Z: **Biosynthesis of chondroitin in engineered *Corynebacterium glutamicum***. *J Microbiol Biotechnol* 2019, **29**:392-400.

21. Badle SS, Jayaraman G, Ramachandran KB: **Ratio of intracellular precursors concentration and their flux influences hyaluronic acid molecular weight in *Streptococcus zooepidemicus* and recombinant *Lactococcus lactis***. *Bioresour Technol* 2014, **163**:222-227.

22. Zhang Q, Yao R, Chen X, Liu L, Xu S, Chen J, Wu J: **Enhancing fructosylated chondroitin production in *Escherichia coli* K4 by balancing the UDP-precursors**. *Metab Eng* 2018, **47**:314-322.

23. Williams A, Gedeon KS, Vaidyanathan D, Yu Y, Collins CH, Dordick JS, Linhardt RJ, Koffas MAG: **Metabolic engineering of *Bacillus megaterium* for heparosan biosynthesis using *Pasteurella multocida* heparosan synthase, PmHS2**. *Microb Cell Factories* 2019, **18**:132.

24. •• Sarnaik A, Abernathy MH, Han X, Ouyang Y, Xia K, Chen Y, Cress B, Zhang F, Lali A, Pandit R *et al.*: **Metabolic engineering of cyanobacteria for photoautotrophic production of heparosan, a pharmaceutical precursor of heparin**. *Algal Res* 2019, **37**:57-63.
This study demonstrated for photoautotrophic production of heparosan the first time using cyanobacteria. The authors brought out the GAG producing potential in this organism by genomic integration of two genes from the heparosan biosynthetic pathway.

25. Restaino OF, di Lauro I, Di Nuzzo R, De Rosa M, Schiraldi C: **New insight into chondroitin and heparosan-like capsular polysaccharide synthesis by profiling of the nucleotide sugar precursors**. *Biosci Rep* 2017, **37**:BSR20160548.

26. Restaino OF, D'ambrosio S, Cassese E, Ferraiuolo SB, Alfano A, Ventriglia R, Marrazzo A, Schiraldi C, Cimini D: **Molecular weight determination of heparosan- and chondroitin-like capsular polysaccharides: figuring out differences between wild -type and engineered Escherichia coli strains**. *Appl Microbiol Biotechnol* 2019, **103**:6771-6782.

27. •• Naseri G, Koffas MAG: **Application of combinatorial optimization strategies in synthetic biology**. *Nat Commun* 2020, **11**:2446.
An excellent review about optimization strategies to balance metabolic fluxes, identify pathway bottlenecks as well as generating and evaluating combinatorial libraries. Methodical and theoretical insights of state of the art techniques are given, which makes this article a highly recommended read for both, beginners and advanced scientists in the field of metabolic engineering and synthetic biology.

28. McCahon R, Hardman J: **Pharmacology of plasma expanders**. *Anaesth Intensive Care Med* 2007, **8**:79-81.

29. Wang Y, Moradali MF, Goudarztalejerdi A, Sims IM, Rehm BHA: **Biological function of a polysaccharide degrading enzyme in the periplasm**. *Sci Rep* 2016, **6**:31249.

30. •• Gansbiller M, Schmid J, Sieber V: **In-depth rheological characterization of genetically modified xanthan-variants**. *Carbohydr Polym* 2019, **213**:236-246.

This study demonstrates that the rheological properties of Xanthan can be modified by combinatorial expression and deletion of acetyl-transferases and pyruvyl-transferases. Genetic engineering resulted in uniform decoration patterns that enabled new possible applications based on the modified rheological behavior.

31. Wu M, Qu J, Shen Y, Dai X, Wei W, Shi Z, Li G, Ma T: **Gel properties of xanthan containing a single repeating unit with saturated pyruvate produced by an engineered *Xanthomonas campestris* CGMCC 15155**. *Food Hydrocolloids* 2019, **87**:747-757.

32. Rütering M, Cress BF, Schilling M, Rühmann B, Koffas MAG, Sieber V, Schmid J: **Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in Paenibacillus polymyxa**. *Synth Biol* 2017, **2**.

33. DeAngelis PL: **Glycosaminoglycan polysaccharide biosynthesis and production: today and tomorrow**. *Appl Microbiol Biotechnol* 2012, **94**:295-305.

34. Wayman JA, Glasscock C, Mansell TJ, DeLisa MP, Varner JD: **Improving designer glycan production in *Escherichia coli* through model-guided metabolic engineering**. *Metab Eng Commun* 2019, **9**:e00088.

35. Maciej-Hulme ML, Leprince ACN, Lavin A, Guimond SE, Turnbull JE, Pelletier J, Yates EA, Powell AK, Skidmore MA: **High sensitivity (zeptomole) detection of BODIPY heparan sulfate (HS) disaccharides by ion-paired RP-HPLC and LIF detection enables analysis of HS from mosquito midguts**. *Biochemistry* 2020.

36. Lin L, Yu Y, Zhang F, Xia K, Zhang X, Linhardt RJ: **Bottom-up and top-down profiling of pentosan polysulfate**. *Analyst* 2019, **144**:4781-4786.

37. Yan L, Li L, Li J, Yu Y, Liu X, Ye X, Linhardt RJ, Chen S: **Bottom-up analysis using liquid chromatography–Fourier transform mass spectrometry to characterize fucosylated chondroitin sulfates from sea cucumbers**. *Glycobiology* 2019, **29**:755-764.

38. Miller RL, Guimond SE, Schwörer R, Zubkova OV, Tyler PC, Xu Y, Liu J, Chopra P, Boons G-J, Grabarics M *et al.*: **Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides**. *Nat Commun* 2020, **11**:1481.

39. Schmid J, Sieber V, Rehm B: **Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies**. *Front Microbiol* 2015, **6**.

40. Xu S, Qiu M, Zhang Q, Wu J, Huimin X, Chen J: **Chain structure and immunomodulatory activity of a fructosylated chondroitin from an engineered *Escherichia coli* K4**. *Int J Biol Macromol* 2019, **133**:702-711.

41. Shanura Fernando IP, Asanka Sanjeewa KK, Samarakoon KW, Lee WW, Kim H-S, Ranasinghe P, Gunasekara UKDSS, Jeon Y-J: **Antioxidant and anti-inflammatory functionality of ten Sri Lankan seaweed extracts obtained by carbohydrase assisted extraction**. *Food Sci Biotechnol* 2018, **27**:1761-1769.

42. Niu Q, Li G, Li C, Li Q, Li J, Liu C, Pan L, Li S, Cai C, Hao J *et al.*: **Two different fucosylated chondroitin sulfates: structural elucidation, stimulating hematopoiesis and immune-enhancing effects**. *Carbohydr Polym* 2020, **230**:115698.

43. Chen Y, Wang Y, Yang S, Yu M, Jiang T, Lv Z: **Glycosaminoglycan from *Apostichopus japonicus* improves glucose metabolism in the liver of insulin resistant mice**. *Mar Drugs* 2019, **18**:1.

44. Zhu Z, Zhu B, Sun Y, Ai C, Wang L, Wen C, Yang J, Song S, Liu X: **Sulfated polysaccharide from sea cucumber and its depolymerized derivative prevent obesity in association with modification of gut microbiota in high-fat diet-fed mice**. *Mol Nutr Food Res* 2018, **62**:1800446.

45. Ustyuzhanina N, Bilan M, Panina E, Sanamyan N, Dmitrenok A, Tsvetkova E, Ushakova N, Shashkov A, Nifantiev N, Usov A: **Structure and anti-inflammatory activity of a new unusual fucosylated chondroitin sulfate from Cucumaria djakonovi**. *Mar Drugs* 2018, **16**:389.

46. Li Q, Cai C, Chang Y, Zhang F, Linhardt RJ, Xue C, Li G, Yu G: **A novel structural fucosylated chondroitin sulfate from Holothuria Mexicana and its effects on growth factors binding and anticoagulation**. *Carbohydr Polym* 2018, **181**:1160-1168.

47. Liu J, Zhou L, He Z, Gao N, Shang F, Xu J, Li Z, Yang Z, Wu M, Zhao J: **Structural analysis and biological activity of a highly regular glycosaminoglycan from Achatina fulica**. *Carbohydr Polym* 2018, **181**:433-441.

48. Li W, Kobayashi T, Moroi S, Kotake H, Ikoma T, Saeki H, Ura K, Takagi Y: **Anti-obesity effects of chondroitin sulfate oligosaccharides from the skate *Raja pulchra***. *Carbohydr Polym* 2019, **214**:303-310.

49. Ahn MY, Kim BJ, Kim HJ, Hwang JS, Jung Y-S, Park K-K: **Anti-aging effect and gene expression profiling of dung beetle glycosaminoglycan in aged rats**. *Biomater Res* 2017, **21**:5.

50. Ahn MY, Kim BJ, Kim HJ, Jin JM, Yoon HJ, Hwang JS, Park K-K: **Anti-cancer effect of dung beetle glycosaminoglycans on melanoma**. *BMC Cancer* 2019, **19**:9.

51. Rani A, Baruah R, Goyal A: **Physicochemical, antioxidant and biocompatible properties of chondroitin sulphate isolated from chicken keel bone for potential biomedical applications**. *Carbohydr Polym* 2017, **159**:11-19.

52. Teshigahara Y, Kakizaki I, Hirao W, Tanaka K, Takahashi R: **A chondroitin sulfate chain of urinary trypsin inhibitor enhances protease inhibitory activity of the core protein**. *J Appl Glycosci* 2020, **67**:63-66 http://dx.doi.org/10.5458/jag.jag.JAG-2019_0021.

1.4. CRISPR

Ever since the initial discovery of unusual microbial repeating sequences in 1987 [115,116] to the recognition as an adaptive bacterial immune system [117] and the application as a tool for the precise cleavage of nucleic acids [118,119], CRISPR (clustered regularly interspaced short palindromic repeats) has revolutionized modern biotechnology as we know it. Early publications by Mojica and colleagues [116, 118], who were aware of the relevance of their discovery, were rejected by prestigious journals such as Nature and Proceedings of the National Academics of Science due to the lack of sufficient novelty and importance. Despite those early setbacks, to date over 18,000 publications have been published using CRISPR derived tools in numerous applications leading to a Nobel price in 2020.

1.4.1. CRISPR/Cas classes and functionality

In nature, CRISPR-Cas (CRISPR associated protein) systems represent an adaptive immune system against phages and other mobile genetic elements such as transposons or plasmids [120]. Despite its late discovery, extensive research over the last two decades revealed that the anti-viral system is quite common among bacteria as well as archaea. It has been described in various forms in approximately 40 % of the sequenced bacterial genomes and nearly 80 % of sequenced archaea [121]. Even though evolution has led to a high diversity of different CRISPR-systems, the general mode of action remains similar (Figure 3). During an initial adaption and immunization phase, short nucleotide sequences of an invading species are cleaved and integrated into the name-giving CRISPR locus by nearly universal Cas1 and Cas2 proteins [122]. These 20 to 60 nt spacers act as a molecular fingerprint of the invading species. Expression of the CRISPR locus results either in a large single effector protein or a multiprotein complex that binds its cognate pre-crRNA. The pre-crRNA is then further processed by the effector module itself or by other non-Cas proteins to form a single spacer fused to structural parts of the CRISPR repeats [123]. In a next step, the matured crRNA guides the effector protein that also possesses nuclease activities to its homologous nucleic acid sequence. Depending on the CRISPR system, DNA, RNA or both can be targeted by the effector module [124]. Upon binding, the nuclease induces a double strand break, which ultimately destroys the invading nucleic acid species [118]. In order to avoid self-cleavage within the CRISPR-array, an additional sequence motif outside of the spacer is required. This so called protospacer adjacent motif (PAM) is a short, conserved sequence of 2-5 bp and varies between different CRISPR-Cas systems [118].

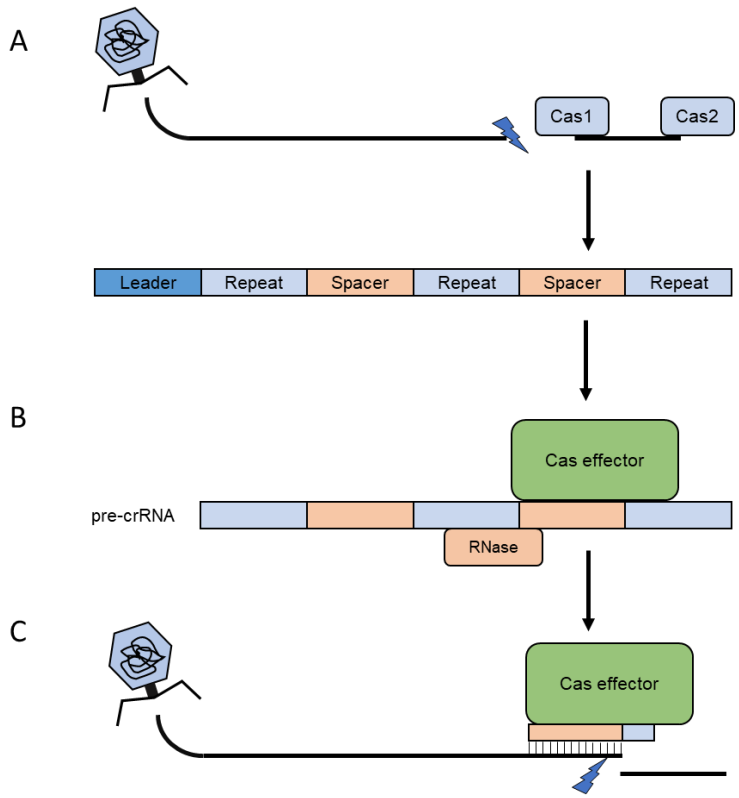


Figure 3: Overview of the CRISPR adaptive immune system. The acquisition of the immune response can be divided into three major phases: A) Adaption: small pieces of foreign DNA are cleaved and integrated as a new spacer into the CRISPR array by the universal Cas1 and Cas2 proteins. B) Processing: pre-crRNA is transcribed and recognized by its cognate Cas effector module. Depending on the type of the CRISPR system, this effector can be either a multiprotein complex or just a single protein. Typically, further processing of the pre-crRNA is required. C) Interference: in case of further invasion of pathogens, the matured crRNA guides the effector module to its homologous target site, which then cleaves the foreign nucleic acid in close proximity to a protospacer adjacent motif (PAM).

Due to its natural function as an adaptive immune system, there has been an evolutionary arms race between bacteria and phages. Therefore, it should not surprise that also bacteriophages and other mobile genetic elements have developed strategies to evade CRISPR-Cas systems. Such anti-CRISPR proteins (acr) bind to the Cas effector module and prevent interference with the target sequence of the invading phage [125]. This evolutionary pressure led to a remarkable diversity of CRISPR-systems that differ in their composition of required Cas proteins, target nucleotide sequences, PAMs and architecture of the genomic loci and is one of the reasons for its high diversity. One of the major criteria in classification is based on the presence of key signature genes and their phylogenetic relationship in order to established CRISPR-classes and subtypes [126]. Up to now, there are only two categories of CRISPR-systems classified. While both groups encode the name-giving CRISPR arrays, in which newly acquired spacers are integrated, class 1 CRISPR-Cas systems are comprised of a multiprotein effector module that binds to its cognate pre-crRNA. After maturation of the pre-crRNA, the crRNA-effector complex binds

its target nucleic acid sequence and induces cleavages. Contrary, in class 2 CRISPR-Cas systems the effector module consists of a single multidomain protein that combines all of the functions of the multiprotein complex of class 1 systems. Only type II CRISPR nucleases, such as the most commonly used Cas9, additionally require RNase III as a non-Cas protein to process the pre-crRNA [119]. The diversity of the characterized CRISPR-Cas systems has massively increased since the early discovery by Mojica et al. [116]. The current classification distinguishes between two major classes, six types and 33 subtypes and with ongoing research further sub-categorization is evident (Figure 4) [126].

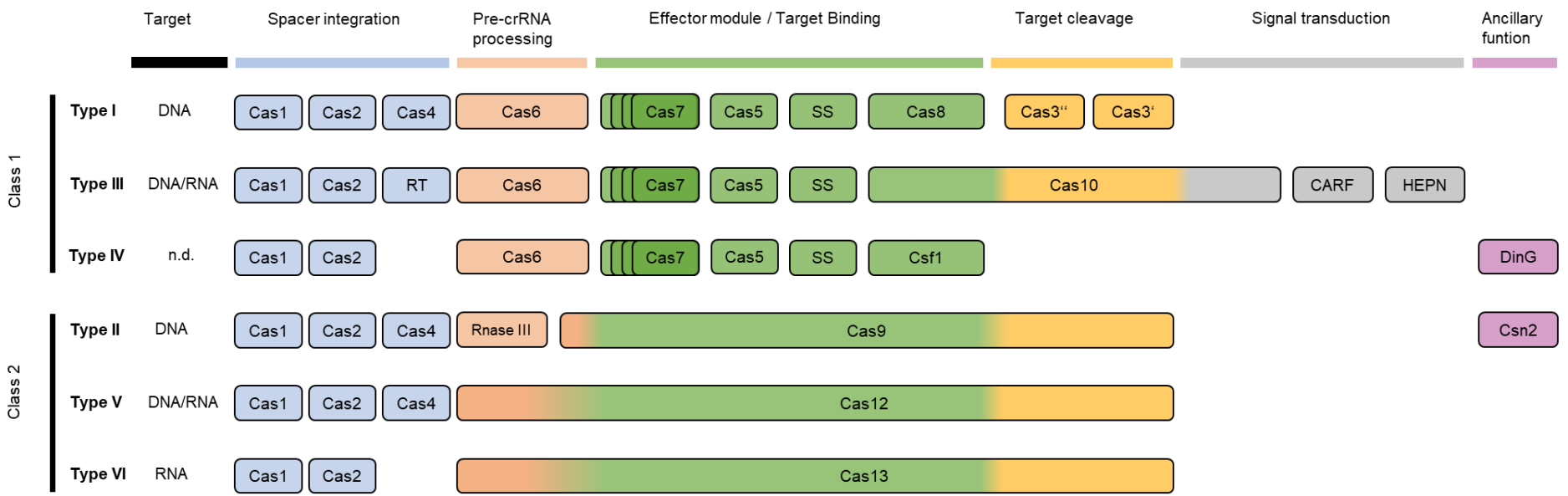


Figure 4: Overview of the current classification of CRISPR-Cas systems based on the presence of signature proteins adopted from Makarova et al. (2020) [126]. Class 1 systems comprise CRISPR-systems, in which a multiprotein complex binds the pre-crRNA forming the crRNA-effector complex. After processing of the crRNA, the effector complex is guided to the target sequence and cleaves the cognate DNA or RNA. Contrary, class 2 only require a single multidomain protein, that performs the same actions analogous to the multiprotein complex of class 1. While Cas1 and Cas2 proteins are important for spacer acquisition and are universally found within all types of CRISPR-Cas systems, it must be noted that still functional exceptions without both of these proteins have been discovered. Multicolored proteins indicate that these Cas proteins possess multiple functional domains in a single peptide chain and contribute to different stages of the immune response. DNA/RNA indicates that CRISPR-Cas systems found in the corresponding type have demonstrated to target either DNA or RNA or both. n.d.: not determined; RT: reverse transcriptase, SS: small subunit; CARF: CRISPR associated Rossman fold, HEPN: higher eukaryotes and prokaryotes nucleotide-binding domain.

Cas9 from *Streptococcus pyogenes* is the most commonly used CRISPR endonuclease for genome editing experiments. After expression of the pre-crRNA from the CRISPR-array, the RNA strand hybridizes with the so-called trans-activating (tracrRNA) and binds to the multifunctional Cas9 nuclease. Processing with bacterial RNase III results in the formation of mature crRNA that guides the riboprotein complex to its cognate target DNA and induces a double strand cleavage 3 bp upstream of the PAM [118,127]. Already in the beginning of the research on CRISPR-Cas systems, this process was facilitated by the fusion of artificial crRNAs to the tracrRNA forming a stable hairpin after expression [119]. Through the expression of this single guide RNA the required processing by RNase III, which is often problematic in heterologous hosts, could be circumvented. The blunt double strand cleavage of Cas9 is induced by the concerted actions of two nuclease domains – RuvC and HNH [128].

A different CRISPR system using Cas12a, formerly known as Cpf1, has been gaining increasing attention as a genome editing tool over the last years. This endonuclease originating from various organisms such as *Acidaminococcus* sp., *Francisella novicida* or *Lachnospiraceae* sp. requires neither a tracrRNA nor additional enzymes to process its pre-crRNA. Instead, Cas12a matures the pre-crRNA in its own ribonuclease catalytic domain resulting in an even simpler system than Cas9 [129]. Due to the self-processing ability of the spacers, these type V CRISPR-Cas systems have been particularly useful for multiplex genome editing approaches [130]. Leveraging this dual RNase/DNase function, the simultaneous perturbation of 25 individual targets was demonstrated in mammalian cell lines using a single transcript harboring both, the open reading frame of Cas12a and a CRISPR array [131]. Contrary to *Sp*Cas9, Cas12a only contains a single RuvC nuclease domain that induces a staggered double strand cleavage at the target site [132]. Additionally, the PAM motif is significantly different. Cas9 cleaves its target DNA 3 bp upstream of the PAM 5'-NGG-3', while Cas12a recognizes A-T rich sequences with the PAM 5'-TTTV-3' located upstream of spacer sequence [133]. While the required PAM is more restrictive for Cas12a endonucleases, it also enables genome editing in loci otherwise inaccessible for Cas9.

1.4.2. Genome Editing

The major breakthrough of the CRISPR-Cas technology came with its application as a highly efficient tool for genome editing. Using *Sp*Cas9 and the previously developed chimeric single guide RNA (sgRNA) containing the functions of matured crRNA and tracrRNA, a very minimalistic two component tool for the precise introduction of double strand cleavage was developed [134]. In order to direct the CRISPR nuclease to the desired locus, only a 20 bp fragment encoding the crRNA needs to be adapted to the sequence of the target site. The high efficiency of genome editing using CRISPR-Cas to target specific genes in the genome can be explained by the high activity of the Cas nuclease in addition to the mostly lethal effects of the double strand cleavage in the host

organism. To evade cell death, endogenous repair mechanisms are activated. Consequently, error prone, non-homologous end joining (NHEJ) events occur through the blunt ended binding of Ku and LigD, which ultimately leads to variable deletions or insertions during the repair process [135]. These 1-10 bp indels typically cause a frame shift within the open reading frame and thereby inactivate the gene of interest. Strategies abusing the natural NHEJ mechanism are often applied for eukaryotic hosts [136]. Alternatively, homology directed repair (HDR) can be used to implement *a priori* designed modifications in the genome by providing tailored donor repair templates with homologous flanks to the target site. This strategy has been used to delete complete genes or operons in various host organism and also allowed the integration of new genes [25,137].

In the beginning of CRISPR-Cas research, genome editing methods based on homologous recombination were already known. In 1979, HDR methods were successfully applied in *S. cerevisiae* harnessing the effective recombination mechanism of these yeast [138]. 15 years later in 1996, zinc finger nucleases were developed, which were also capable of double strand cleavage after the concerted binding of two endonucleases [139,140]. TALENs (transcription activator-like effector nuclease) derived from *Xanthomonas* sp. represent a very similar system, in which peptide sequences can be tailored to specific DNA sequences in order to initiate double strand breaks by using the identical nuclease domains as zinc fingers (FokI) [141,142]. However, due to the lethal character of the more efficient double strand breaks induced by Cas effectors, genome editing using the CRISPR technology does not require the integration of additional resistance markers in the repair template to select for the successful genome engineering. Furthermore, CRISPR-Cas9 systems can be easily programmed by changing the sequence of a short gRNA. In contrast TALENS require laborious protein engineering to modify the target binding site. This fact highly contributed to the success of CRISPR-Cas9 as it not only accelerated the pace of introducing new distinct modifications, but also circumvented possible negative side effects of recombinase mediated excision of marker cassettes, which are essential in other genome engineering methods [143]. The biggest limitation of CRISPR-Cas systems that restricts the number of possible target sites is the requirement for a PAM sequence (NGG for *Sp*Cas9) in close proximity to the cleavage point. However, there have been also substantial efforts to alter the PAM specificities of various CRISPR-Cas systems [144,145]. An extreme case is represented by xCas9, an engineered version of *Sp*Cas9, which recognizes a broad range of PAM sequences including NG, GAA and GAT without negatively effecting the DNA specificity or off-target activity [146].

Despite the generally high precision and efficiency of Cas9, the nuclease allows for single base mismatches within the core region of the gRNA [147]. Particularly in eukaryotic cell lines, off-targeting poses a severe problem, in which the cleavage of

highly similar sequences causes undesirable indel formations in undesired loci. A solution to this problem is provided by *Sp*Cas9-nickase. Here, one out of the two conserved nuclease functions is inactivated by a distinct point mutation (D10A in RuvC, H810A in HNH). While the mutant enzyme as a consequence can only cleave a single strand of the DNA helix, two paired gRNAs can be directed to a target site inducing a staggered double strand break [148]. As a result, potential off-targets are repaired by the highly efficient BER pathway [149] and only on-target binding of the nickase results in double strand cleavage leading to the indel formation.

1.4.3. CRISPRa & CRISPRi

A further step in the development of the CRISPR-Cas technology was the generation of catalytically deactivated variants of Cas nucleases (dCas) with mutated nuclease domains (D10A and H810A for *Sp*Cas9). Consequently, the effector module is still directed by its cognate gRNA to the target site but cannot cleave the DNA. Directing the inactive nuclease in the promotor region or the open reading frame of a target gene prevents the binding of the corresponding RNA-polymerase or stops the elongation of a nascent mRNA [150,151]. In eukaryotes, effects of this so-called CRISPR interference (CRISPRi) can be further intensified by fusing effector domains to dCas9. Examples are the Mxi1-deacetylase or Krüppel-associated box (KRAB) domain, which modify the chromatin structure of the target area leading to an even tighter repression of gene expression [152]. Thereby, CRISPRi allows the transcriptional perturbation of gene expression without completely deleting the target sequence.

In addition, gene expression can also be enhanced by transcriptional fusion of the dCas effector module to activator domains of transcription factors and locating the riboprotein complex upstream of the target promoter. The activator domain facilitates the recruitment of the corresponding RNA-polymerase and thereby increases the transcription and consequently the expression of the target gene. Particularly in eukaryotes, CRISPR-activation (CRISPRa) has demonstrated remarkable effects and is used in combination with CRISPRi as a tool for large-scale functional screenings in biomedical research [153]. Interestingly, to date only two publications established a functional CRISPRa system in bacteria that showed a distinct dynamic range of activation output [154,155]. In contrast to eukaryotic systems that mainly utilize chromatin restructuring effects of effector domains, prokaryotic systems need direct interaction with the corresponding RNA-polymerase to enhance recruitment. Therefore, the precise distance of the dCas9-activator complex to its promoter plays a crucial role. Computational analysis of previously established tools determined a distance of the effector module of 40-100 bp upstream of the target promoter [156].

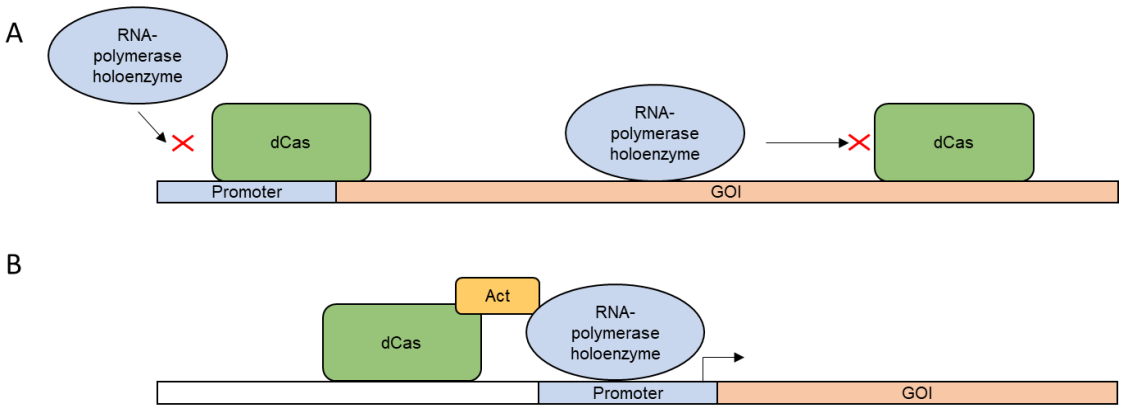


Figure 5: Overview of the mode of action of CRISPRi (A) and CRISPRa (B) systems A) Catalytically inactive Cas9 (dCas) either blocks the binding of the corresponding RNA-polymerase holoenzyme or the elongation of the transcription process to knock-down the expression of a gene of interest (GOI). B) CRISPRa enhances gene expression by the precise positioning of dCas9 fused to activator domains (Act) 40-100 bp upstream of the transcription start site to enhance the recruitment of the RNA-polymerase and consequently enhance the expression of a target gene.

Multiplexing approaches, targeting several loci in parallel, are of particular interest for metabolic engineering strategies using CRISPRa/i for flux balance optimization [157]. A major drawback of many, currently established systems is the use of *Sp*dCas9 as an effector module for those applications [158]. Although *Sp*Cas9 is historically the most well-studied CRISPR-endonuclease, it typically requires the expression of single gRNAs from distinct operons. Strategies to overcome this constraint include the expression of sgRNA transcripts from multiple plasmids or the co-expression of RNA processing enzymes such as RNase III [159] and Csy4 [160,161]. Additionally, flanking of consecutive gRNAs by ribozymes or tRNAs enabled the efficient processing of the mature gRNA [162,163] from a single transcript. However, these approaches are still limited in the number of gRNAs due to cytotoxic side effects [164]. An up-striving contender in multiplex CRISPRa/i experiments are Cas12a effector proteins. While the DNA cleavage is mediated by a single RuvC domain, these enzymes belonging to the CRISPR class 2 type V family also encode a RNase domain which processes transcribed spacers of its cognate CRISPR-array [130]. Thereby, both Cas12a and the CRISPR-array can be encoded in a single transcript allowing multiplex approaches without complex cloning efforts or putative negative side effects. While Cas12a from *Acidaminococcus* sp. has demonstrated remarkable CRISPRa results in mammalian cell lines [130,144], no comparable system has been established in bacterial hosts yet.

1.5. Aims and scope

The scope of this thesis comprises the metabolic engineering of *P. polymyxa* DSM 365 for the production of industrially relevant bulk chemicals.

On the one hand, *P. polymyxa* is an avid producer of 2*R*-3*R*-BDL with exceptionally high enantiomeric purity. While there is already detailed knowledge on the process optimization of *P. polymyxa* to increase productivity, more complex metabolic engineering approaches proved difficult due the genetic inaccessibility of this organism. Using a previously developed CRISPR-Cas9 toolkit, particularly designed for *P. polymyxa*, the focus was to transfer established engineering strategies from common 2,3-BDL producing organisms to *P. polymyxa*. Carbon fluxes of redox-neutral side products should be redirected towards butanediol by deletion of the respective pathways. Furthermore, bottlenecks within the 2,3-BDL biosynthesis pathway should be overcome by overexpression of the respective genes either by genomic integration or episomal expression from a compatible plasmid.

On the other hand, *P. polymyxa* has demonstrated the production of highly interesting exopolysaccharides that can be used in various applications. By combinatorial knock-outs of functional genes of the paenan biosynthesis cluster, modified EPS variants should be generated. The EPS variants must be characterized by a combination of analytical methods to establish the monomer composition, linkage pattern and sequence in order to elucidate the structure of the wildtype polysaccharide as well as the corresponding enzyme function of the knock-out strains. Furthermore, detailed rheological analysis is required to propose suitable applications for the different variants.

In order to accomplish these goals, this work aims to develop new genetic tools to further enhance the metabolic engineering of *P. polymyxa*. Ideally, these tools comprise CRISPR-Cas based technology using catalytically inactivated effector modules to knock-down multiple genes of interest simultaneously. Due to a lack of available prokaryotic CRISPRa plasmids, multiplex tools to induce gene expression should be developed. To enable transcriptional regulation of organisms other than *P. polymyxa* DSM 365, these new tools should be developed as broad-host range kits. If successful, this CRISPRa/i system can be applied for further metabolic engineering approaches of both, 2,3-BDL synthesis as well as EPS production and even further targets.

2. Materials and Methods

2.1. Materials

2.1.1. Equipment

Table 3: Equipment and devices used in this study

Device	Supplier	Model
1 L bioreactor	Eppendorf (Germany)	DASGIP SSL 1000
1 L bioreactor – gas analytics	Eppendorf (Germany)	DASGIP GA4
2 L bioreactor	Sartorius (Germany)	Biostat B plus
10 L bioreactor	Sartorius (Germany)	Biostat C plus 10 L
30 L bioreactor	Sartorius (Germany)	Biostat C plus 30 L
Autoclave	Thermo Scientific (Germany)	Varioklav 135S
Ball mill	Retsch (Germany)	Mixer Mill MM 400
Centrifuge	Thermo Scientific (Germany)	Sorvall Lynx 6000
Clean bench	Thermo Scientific (Germany)	MSC Advantage
Cross flow filtration system	Sartorius (Germany)	stedim biotech Sartojet
Filtration probe	TRACE analytics (Germany)	0.2 µm probe
Freezer -20 °C	Liebherr-Hausgeräte (Germany)	Fv913
Freezer -80 °C	Thermo Scientific (Germany)	Forma 906-86 C ULT
Gas chromatography	Thermo Scientific (Germany)	Trace GC 2000 with Ultra Trace
GC - autosampler	Thermo Scientific (Germany)	TriPlus Autosampler
GC - column	Restek Germany	Rxi-5 Sil MS
GC - mass spectrometer	Restek Germany	Ultra Trace DSQ II
Gel electrophoresis device	Bio-Rad (US)	Mini-Sub Cell GT
Heating block	VLM (Germany)	EC model series
Incubator	Binder (Germany)	KBF 240 E5.1/C
Magnetic stirrer	Thermo Scientific (Germany)	Variomag Telesystem
pH-meter	Mettler Toledo (Germany)	InLab Expert Pro
Pipettes	Brand (Germany)	Transferpette
Portable photometer	Amersham (UK)	Ultraspec 10
qPCR cycler	Bio-Rad (US)	CFX96

Device	Supplier	Model
Refrigerated Vapor trap	Thermo Scientific (Germany)	RVT 100
Rheometer	Anton Paar (Austria)	MCR3000
Shaker	Thermo Scientific (Germany)	MaxiQ 2000
Spectrophotometer	Thermo Scientific (Germany)	Multiskan
Speed mixer	Hauschild (Germany)	DAC 150 SP
Table top centrifuge	Thermo Scientific (Germany)	Hearaeus Fresco 21
Thermocycler	Bio-Rad (US)	MyCycler MJ Mini
UHPLC	Dionex (Germany)	Ultimate 3000RS
UHPLC - refractive index detector	Shodex (Japan)	RI 101
UHPLC- autosampler	Dionex (Germany)	WPS 3000RS
UHPLC - column	YMC (Japan) Macherey-Nagel (Germany) Phenomenex (USA)	Triart Diol (1.9 µm 100 x 2 mm) Gravity C18 (1.8 µm 100 x 2 mm) Rezex ROA-H^+ (300 mm x 7.8 mm)
UHPLC - degasser	Dionex (Germany)	SRD 3400
UHPLC - diode array detector	Dionex (Germany)	DAD 3000RS
UHPLC - high capacity iron trap	Bruker Daltonic (Germany)	HCT
UHPLC - pump module	Dionex (Germany)	HPG 3400RS
Ultrapure water system		
Vacuum drying oven	Binder (Germany)	VDL53
Vortexer	Scientific industries (US)	Vortex Genie 2
Water bath	Huber (Germany)	CC1

2.1.2. Chemicals and reagents

Table 4: Chemicals used in this study and corresponding vendors

Chemical	Supplier	Catalog number
2*R*,3*R*-butanediol	Acros Organics	217980050
32 % Ammonium solution	Carl Roth	P093.1
ABTS	Sigma-Aldrich	A1888
Acetic acid	Carl Roth	6755.1
Acetic anhydride	Carl Roth	CP28
acetoin	Sigma-Aldrich	A17951-50G
Acetonitrile	VWR	83.040.320
Acetyl chloride	TCI	A0082
Agar-Agar, Kobe I	Carl Roth	5210.4
Agarose	Biozym	840004
Ammonium acetate	Carl Roth	7869.1
Ammonium sulfate	VWR	A1032.5000
Antifoam B emulsion	Sigma-Aldrich	MKCG1647
Biotin	Carl Roth	3822
Boric acid	Sigma-Aldrich	B7091
Calcium chloride dihydrate	Carl Roth	5239.1
Chloroform	Carl Roth	7331.2
Cobalt(II) nitrate hexahydrate	Sigma-Aldrich	239267
Copper(II) sulphate pentahydrate	Sigma-Aldrich	61245
DA–64	Wako Chemicals	043-22351
dATP	Promega	U1205
dCTP	Promega	U1225
dGTP	Promega	U1215
Dimsyl	P. Lommes	M3RC1
Dipotassium phosphate	Carl Roth	P749.3
Dithiothreitol	Carl Roth	6908.2
dNTP Mix	VWR	733-1363
dTTP	Promega	U1235
Ethyl acetate	Carl Roth	AE69.1
Ethylenediamine	Fluka	3550
Formic acid	Sigma-Aldrich	06440-1L
Glucose monohydrate	Carl Roth	6887.5
Glycerol	Carl Roth	3783.2
Hydrochloric acid 37 %	Carl Roth	4625.2
Iron(II) sulphate heptahydrate	Sigma-Aldrich	31236
Isopropoanol	VWR	ACRO423830250
Lactate	Carl Roth	4071.2
Lithium granular	Sigma-Aldrich	499811
Lithium nitrate	Alfa Aesar	13405.3
Magnesium chloride hexahydrate	Carl Roth	2189.1
Magnesium sulfate heptahydrate	Merck	1.058.861.000
Manganese(II) chloride	VWR	1.05927-0100
MES	Carl Roth	4256.3
meso-butanediol	Sigma-Aldrich	361461-10G
Methanol	VWR	83.638.320
Methyliodine	VWR	25.596.154
Methyl-α-D-fucopyranoside	TCI	M1051

Chemical	Supplier	Catalog number
Methyl-α-D-galactopyranoside	TCI	M1047
Methyl-α-D-glucopyranoside	Alfa Aesar	A12484
Methyl-α-D-mannopyranoside	TCI	M0368
Monopotassium phosphate	Carl Roth	3904.3
MOPS	Carl Roth	6979.4
N-(3-dimethylaminopropyl)-N-ethylcarbodiimide hydrochloride	Sigma-Aldrich	E1769
NAD^+	Carl Roth	AE11.2
Neomycin sulphate	Carl Roth	8668
Phenol	VWR	20.599.231
Phenyl red	Alfa Aesar	B21710
Phosphoric acid	Carl Roth	6366.1
PMP	Sigma-Aldrich	M70800
Polymyxin B sulphate	Carl Roth	0235.1
Potassium phosphate dibasic	Carl Roth	P749.3
Potassium phosphate monobasic	Carl Roth	1.04873.1
Pullulan	Sigma-Aldrich	53168
Pyridine	Sigma-Aldrich	270970
Quick-Load Purple 2-Log Ladder	New England Biolabs	N0550L
Rubidium chloride	Alfa Aesar	12892.14
Serva DNA Stain Clear G	Serva	39804.02
Silver carbonate	Sigma-Aldrich	179647-5G
Silver(I) oxide	Alfa Aesar	11407
Sodium borohydride	Sigma-Aldrich	452882
Sodium chloride	Carl Roth	P029.3
Sodium hydrid, 60 % dispersion in oil	Alfa Aesar	H36490.18
Sodium hydroxide	Carl Roth	6771.2
Sodium molybdate dihydrate	Merck	331058
Thiamine pyrophosphate	Sigma-Aldrich	C8754
Toluene	Carl Roth	7346.1
Trifluoroacetic acid	Sigma-Aldrich	T6508
TRIS-HCl	Carl Roth	9090.3
Tryptone from casein	Carl Roth	8952.5
Vitamin solution RPMI 1640 (100x)	Sigma-Aldrich	R7256
Yeast extract	Carl Roth	2363.5
Zinc sulphate heptahydrate	Merck	1.088.831.000

2.1.3. Enzymes and consumables

Table 5: Enzymes, kits and special consumables used in this study

Enzyme	Supplier	Catalog number
Accuzyme	BioCat	BIO-25028-BL
Aurum Total RNA extraction Kit	BioRad	732-6820
BbsI	New England Biolabs	R0539L
BsaI	New England Biolabs	R0176L
DNase	VWR	A3778.0500
GeneJET Plasmid Miniprep Kit	Thermo Fischer Scientific	10242490
Glucose oxidase	VWR	G2133-50KU
Horseradish Peroxidase	Sigma-Aldrich	P6782
iScript cDNA Synthesis Kit	BioRad	170-8891
Lysozyme	Sigma-Aldrich	71110-4
Nucleospin Gel and PCR Clean-up Kit	Macherey-Nagel	740.609.250
Phusion high-fidelity polymerase	New England Biolabs	M0530L
Pyruvate oxidase	Sigma-Aldrich	P4591
Quick Dephosphorylation Kit	Thermo Fischer Scientific	M0508S
SpeI	New England Biolabs	R0133L
SsoAdvanced™ Universal SYBR® Green Supermix	BioRad	1725270
T4 ligase	New England Biolabs	M0202L
T4 Polynucleotide kinase	New England Biolabs	M0201S
T5 exonuclease	New England Biolabs	M0363S
Taq ligase	New England Biolabs	M0208S
Taq polymerase	New England Biolabs	M0267L
XbaI	New England Biolabs	R0145S

2.1.4. Plasmids and oligonucleotides

Plasmids used or constructed in this work are summarized in the table below (**Table 6**). A list of all oligonucleotides and gene fragments can be found in the supplemental materials for each manuscript or publication respectively (Chapter 7.1). All oligonucleotides were obtained from Eurofins Genomics Germany (Germany). Gene fragment synthesis was performed by ATG:biosynthetics (Germany).

Table 6: Plasmids used and constructed in this work

Plasmids	Description	Reference
pCasPP	*P. polymyxa* CRISPR-Cas9 genome editing plasmid; PsgsE-*cas9*; Pgapdh-off-target gRNA; *neo*; *oriT*; *repU*	[25]
pCasPPH_clu1	*clu1* targeting knock-out plasmid containing repair template	[25]
pCasPPH_ldh1	*ldh1* targeting knock-out plasmid containing repair template	This study
pCasPPH_spo0A	*spo0A* targeting knock-out plasmid containing repair template	This study
pCasPPH_pfl	*pfl* targeting knock-out plasmid containing repair template	This study
pCasPPH_spoIIE	*spoIIE* targeting knock-out plasmid containing repair template	This study
pCasPPH_sacB	*sacB* targeting knock-out plasmid containing repair template	This study
pHEiP_Ppbdh	constitutive expression of *Ppbdh* under control of Pspac	This study
pCasPPH_pepC	*pepC* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepD	*pepD* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepE	*pepE* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepF	*pepF* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepG	*pepG* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepH	*pepH* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepI	*pepI* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepJ	*pepJ* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepK	*pepK* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepL	*pepL* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepQ	*pepQ* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepT	*pepT* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepU	*pepU* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepV	*pepV* targeting knock-out plasmid containing repair template	This study
pCasPPH_epsO	*epsO* targeting knock-out plasmid containing repair template	This study
pCasPPH_fcl	*fcl* targeting knock-out plasmid containing repair template	This study

Plasmids	Description	Reference
pCRai	PsgsE-*Asdcas12a*; Pgapdh-off target gRNA; *neo*; *oriT*, *repU*	This study
pCRaiGFP	PsgsE-*Asdcas12a*; Pgapdh-off target gRNA;*neo*; *oriT*, *repU*; PsgsE-*sfGFP*	This study
pCRaiGmR_soxS	PsgsE-*Asdcas12a-soxS*; Pgapdh-off target gRNA; *neo*; *oriT*, *repU*; PsgsE_*mRFP*-US-PsgsE-*sfGFP*	This study
pCRaiGFP_GltC	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - off target gRNA	This study
pCRaiGFP_GltC_a1	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a1	This study
pCRaiGFP_GltC_a2	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a2	This study
pCRaiGFP_GltC_a3	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a3	This study
pCRaiGFP_GltC_a4	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a4	This study
pCRaiGFP_sig70	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - off target gRNA	This study
pCRaiGFP_sig70_a1	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a1	This study
pCRaiGFP_sig70_a2	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a2	This study
pCRaiGFP_sig70_a3	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a3	This study
pCRaiGFP_sig70_a4	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a4	This study
pCRai_GFP_rpoD_a1	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a1	This study
pCRai_GFP_rpoD_a2	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a2	This study
pCRai_GFP_rpoD_a3	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a3	This study
pCRai_GFP_rpoD_a4	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a4	This study
pCRai_GFP_rpoD	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - off target gRNA	This study
pCRai_GFP_soxS_a1	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a1	This study
pCRai_GFP_soxS_a2	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a2	This study
pCRai_GFP_soxS_a3	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a3	This study
pCRai_GFP_soxS_a4	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a4	This study
pCRai_GFP_soxS	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - off target gRNA	This study
pCRai_GFP_CRP_a1	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a1	This study
pCRai_GFP_CRP_a2	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a2	This study

Plasmids	Description	Reference
pCRai_GFP_CRP_a3	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a3	This study
pCRai_GFP_CRP_a4	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - sgRNA_a4	This study
pCRai_GFP_CRP	Test plasmid for CRISPRa targeting US region of PsgsE-*sfGFP* - off target gRNA	This study
pCRaiGFP_soxS_sfGFP_T1	Test plasmid for sfGFP repression sgRNA_T1	This study
pCRaiGFP_soxS_sfGFP_T2	Test plasmid for *sfGFP* repression sgRNA_T2	This study
pCRaiGFP_soxS_sfGFP_T3	Test plasmid for *sfGFP* repression sgRNA_T3	This study
pCRai_pepQ_T1	Plasmid for repression of *pepQ* sgRNA T1	This study
pCRai_pepCQ_T1	Plasmid for dual repression of *pepQ* and *pepC*	This study
pCRai_soxS_ldhmultibdh1	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 1)	This study
pCRai_soxS_ldhmultibdh2	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 2)	This study
pCRai_soxS_ldhmultibdh3	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 3)	This study
pCRaiGmR_soxS_mT2	PsgsE-*mRFP*-US-PsgsE_*sfGFP* expression plasmid with gRNA for *mRFP* repression (sgRNA T2)	This study
pCRaiGmR_soxS_dualT2	PsgsE-*mRFP*-US-PsgsE_*sfGFP* expression plasmid with gRNA for *mRFP* repression (sgRNA T2) and *GFP* activation (sgRNA a1)	This study
pCRaiGmR_soxS_GFPa1	PsgsE-*mRFP*-US-PsgsE_*sfGFP* expression plasmid with gRNA and *GFP* activation (sgRNA a1)	This study

2.1.5. Gibson isothermal assembly master mix

A reaction master mix by mixing 320 µL 5x isothermal reaction buffer (25 % w/v PEG-6000, 500 mM Tris-HCl pH 7.5, 50 mM $MgCl_2$, 50 mM DTT, 1 mM dATP, 1 mM dCTP, 1mM dGTP, 1mM dTTP, 5 mM NAD^+) with 0.64 µL T5 exonuclease (10 U μL^{-1}), 20 µL Phusion High Fidelity DNA Polymerase (2 U μL^{-1}), 160 µL Taq DNA ligase (40 U μL^{-1}) and 699 µL H_2O. Afterwards, 30 µL aliquots were stored at -20 °C and thawed on ice when required.

2.1.6. Software

Table 7: Software, tools and databases used in this study

Product	Supplier	Application
Basic Local Alignment Search Tool (BLAST)	National Center for Biotechnology Information (NCBI) (US)	DNA/Protein alignments
Carbohydrate Structure Database	see reference [165]	Database for carbohydrate structures and visualization
CAZy	see reference [109]	Database of carbohydrate active enzymes
Chemdraw Professional v.17.1	Perkin Elmer (US)	Visualization
Chromeleon v.6.80	Dionex (US)	HPLC UV/RI data analysis
CRISPR Guide RNA Design Tool	Benchling (US)	gRNA design
DataAnalysis v.4.0	Bruker Daltonics (Germany)	MS-Data analysis
GIMP v.2.10.18	GIMP	Visualization
Magellan v.6.6	Tecan (Switzerland)	Photometer analysis
MS Office 2016	Microsoft Corporation (US)	Text processing, data analysis, visualization
NUPACK	see reference [166]	In silico DNA analysis
OligoAnalyzer	Integrated DNA Technologies (IDT) (US)	In silico oligonucleotide analysis
PRISM GraphPad v.7	GraphPad Software	Visualization
PSS WinGPC UniChrom v.8.1	PSS Polymer Standards Service (Germany)	GPC Data analysis
QuantAnalysis v.2.0	Bruker Daltonics (Germany)	MS-Data analysis
Rheoplus v.3.61	Anton Paar (Austria)	Rheometry
SnapGene 2.3.2	GSL Biotech (US)	In silico DNA analysis and visualization
Xcalibur v.4.0	Thermo Scientific (US)	GC analysis
Zotero v.5	Corporation of Digital Scholarship (US)	Citation

2.1.7. Bacterial Strains

Table 8: Bacterial strains and genotypes of strains used in this study

Bacterial Strains	Genotype	Reference
E. coli S17-1	Conjugation strain; recA pro hsdR RP42 Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC 47055
P. polymyxa DSM 365	wild type	DSMZ
P. polymyxa DSM 365 *cas9*	DSM 365 espI::Psgse-*cas9*	This study
P. polymyxa DSM 365 *Δclu1*	DSM 365 *Δclu1*	[25]
P. polymyxa DSM 365 *ΔsacB*	DSM 365 *ΔsacB*	This study
P. polymyxa DSM 365 *Δclu1 ΔsacB*	DSM 365 *Δclu1 ΔsacB*	This study
P. polymyxa DSM 365 *Δldh1*	DSM 365 *Δldh1*	This study
P. polymyxa DSM 365 *Δpfl*	DSM 365 *Δpfl*	This study
P. polymyxa DSM 365 *Δspo0A*	DSM 365 *Δspo0A*	This study
P. polymyxa DSM 365 *ΔspoIIE*	DSM 365 *ΔspoIIE*	This study
P. polymyxa DSM 365 *Δldh1* pHEiP_Ppbdh	DSM 365 *Δldh1* pHEiP_Ppbdh	This study
P. polymyxa DSM 365 *ΔpepC*	DSM 365 *ΔpepC*	This study
P. polymyxa DSM 365 *ΔpepD*	DSM 365 *ΔpepD*	This study
P. polymyxa DSM 365 *ΔpepE*	DSM 365 *ΔpepE*	This study
P. polymyxa DSM 365 *ΔpepF*	DSM 365 *ΔpepF*	[25]
P. polymyxa DSM 365 *ΔpepG*	DSM 365 *ΔpepG*	This study
P. polymyxa DSM 365 *ΔpepH*	DSM 365 *ΔpepH*	This study
P. polymyxa DSM 365 *ΔpepI*	DSM 365 *ΔpepI*	This study
P. polymyxa DSM 365 *ΔpepJ*	DSM 365 *ΔpepJ*	[25]
P. polymyxa DSM 365 *ΔpepK*	DSM 365 *ΔpepK*	This study
P. polymyxa DSM 365 *ΔpepL*	DSM 365 *ΔpepL*	This study
P. polymyxa DSM 365 *ΔpepQ*	DSM 365 *ΔpepQ*	This study
P. polymyxa DSM 365 *ΔpepT*	DSM 365 *ΔpepT*	This study
P. polymyxa DSM 365 *ΔpepU*	DSM 365 *ΔpepU*	This study
P. polymyxa DSM 365 *ΔpepV*	DSM 365 *ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepCQ*	DSM 365 *ΔpepC ΔpepQ*	This study
P. polymyxa DSM 365 *ΔpepCQ ΔsacB*	DSM 365 *ΔpepC ΔpepQ ΔsacB*	This study
P. polymyxa DSM 365 *ΔpepCF*	DSM 365 *ΔpepC ΔpepF*	This study
P. polymyxa DSM 365 *ΔpepDJ*	DSM 365 *ΔpepD ΔpepJ*	This study
P. polymyxa DSM 365 *ΔpepEI*	*DSM 365 ΔpepE ΔpepI*	This study
P. polymyxa DSM 365 *ΔpepETUV*	DSM 365 *ΔpepE ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepFJ*	DSM 365 *ΔpepF ΔpepJ*	This study
P. polymyxa DSM 365 *ΔpepQF*	DSM 365 *ΔpepQ ΔpepF*	This study
P. polymyxa DSM 365 *ΔpepTL*	DSM 365 *ΔpepT ΔpepL*	This study
P. polymyxa DSM 365 *ΔpepUL*	DSM 365 *ΔpepU ΔpepL*	This study
P. polymyxa DSM 365 *ΔpepLJ*	DSM 365 *ΔpepJ ΔpepL*	This study
P. polymyxa DSM 365 *ΔpepVJ*	DSM 365 *ΔpepV ΔpepJ*	This study

Bacterial Strains	Genotype	Reference
P. polymyxa DSM 365 *ΔpepQTUV*	DSM 365 *ΔpepQ ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepTUV*	DSM 365 *ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepITUV*	DSM 365 *ΔpepI ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepKTUV*	DSM 365 *ΔpepK ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepLTUV*	DSM 365 *ΔpepL ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔepsO*	DSM 365 *ΔepsO*	This study
P. polymyxa DSM 365 *Δfcl*	DSM 365 *Δfcl*	This study

In addition to the strains listed above, each plasmid listed in **Table 6** was used for the transformation of *E. coli* S17-1 as well as for the conjugational transformation of *P. polymyxa* DSM 365.

2.2. Media and Buffer

2.2.1. Media preparation

All growth media were sterilized by autoclaving at 121 °C for 20 min. Glucose was sterilized separately to avoid Maillard reactions. Vitamin solution, trace element solution and antibiotics were filter-sterilized and added to the medium after autoclaving. 50 µg mL^{-1} neomycin and 20 µg mL^{-1} polymyxin B were added as selection markers if required.

2.2.2. General cultivation medium

General cloning procedures were performed using lysogeny broth (LB) medium containing 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 10 g L^{-1} sodium chloride. To prepare agar plates, 20 g L^{-1} agar-agar was added. For the conjugational transformation of *P. polymyxa* DSM 365 50 mM calcium chloride dihydrate were added to all media.

2.2.3. Butanediol fermentation media

Medium composition was adapted from a previous study by Okonkwo et al. (2017) [38]. Pre-culture medium containing 60 g L^{-1} glucose, 5 g L^{-1} yeast extract, 5 g L^{-1} tryptone, 0.2 g L^{-1} magnesium sulfate heptahydrate, 3.5 g L^{-1} potassium dihydrogen phosphate and 2.5 g L^{-1} dipotassium phosphate. Fermentation medium components were autoclaved separately and contained 120 g L^{-1} glucose, 7 g L^{-1} glycerol, 5 g L^{-1} yeast extract, 3.5 g L^{-1} tryptone, 0.2 g L^{-1} magnesium sulfate heptahydrate, 3.5 g L^{-1} potassium dihydrogen phosphate, 2.5 g L^{-1} dipotassium phosphate, 5 g L^{-1} ammonium acetate, 4 g L^{-1} ammonium sulfate. After autoclaving, 3 mL L^{-1} of sterile filtered trace elements per liter of medium were added.

2.2.4. EPS fermentation media

Preculture medium contained 10 g L^{-1} glucose, 0.05 g L^{-1} calcium chloride dihydrate, 5 g L^{-1} tryptone, 1.33 g L^{-1} magnesium sulfate heptahydrate, 1.67 g L^{-1} potassium dihydrogen phosphate, 2 mL L^{-1} RPMI 1640 vitamins solution (Merk, Germany) and 1 mL L^{-1} trace elements solution. 20 g L^{-1} MOPS (3-N-morpholino propanesulfonic acid) were used to buffer the medium to pH 7. Fermentation medium was prepared equal to the preculture medium with an increased glucose concentration of 30 g L^{-1} and exclusion of MOPS.

2.2.5. Trace element solution

Trace element solution contained 2.5 g L^{-1} iron sulfate heptahydrate, 2.1 g L^{-1} sodium tartrate dihydrate, 1.8 g L^{-1} manganese chloride dihydrate, 0.075 g L^{-1} cobalt chloride hexahydrate, 0.031 g L^{-1} copper sulfate pentahydrate, 0.258 g L^{-1} boric acid, 0.023 g L^{-1} sodium molybdate dihydrate and 0.021 g L^{-1} zinc chloride. Trace element solution was filter-sterilized and added to the indicated media after cooling down to room temperature.

2.3. Microbiological methods

2.3.1. Strain storage and cultivation

Strains were stored at -80 °C in cryo tubes by mixing 900 µL of an overnight culture of the bacterial strain in LB – medium with the corresponding antibiotic and 900 µL 60 % (w/w) glycerol solution. For reactivation of strains, cells were spread on the corresponding agar plate from the frozen cryo culture and incubated at 37 °C for *E. coli* and 30 °C for *P. polymyxa* DSM 365. Liquid cultures were inoculated with single colonies from the agar plates. Depending on the size of the required liquid cultures, cultivation conditions were adjusted. For small scale, 3 mL medium in 12 mL single-use plastic tubes were inoculated and incubated at 300 rpm. Baffled shaking flasks filled with 20 % medium were sealed with aluminum caps and incubated at 160 rpm.

2.3.2. Chemically competent *E. coli* cells

An overnight pre-culture of *E. coli* was used for the inoculation of 100 mL LB-medium in a baffled shaking flask to an initial OD_{600} of 0.1. Cells were grown until mid-log phase (OD_{600} 0.5-0.6) at 37 °C, 280 rpm. Cells were pelleted by centrifugation (3,500 g, 15 min, 4 °C) and afterwards reconstituted in 50 mL pre-cooled 50 mM $CaCl_2$ and incubated on ice for 60 min. Cells were centrifuged again (3,500 g, 15 min, 4 °C), resuspended in 10 mL 50 mM $CaCl_2$ solution containing 15 % (v/v) glycerol and incubated on ice for 60 min. 100 µL aliquots were transferred to sterile 1.5 mL reaction tubes, frozen with liquid nitrogen and stored at -80 °C.

2.3.3. Conjugation

P. polymyxa was transformed by conjugation using *E. coli* S17-1 harboring the target plasmids. Overnight cultures of donor and recipient strains were diluted 1:100 with

selective or non-selective LB media supplemented with 50 mM $CaCl_2$ and cultivated at 37 °C for 3 h, 280 rpm. 900 µL of the recipient culture was heat shocked at 42 °C for 15 min and mixed with 300 µL of the donor strain. Cells were centrifuged at 6,000 g for 2 min, resuspended in 400 µL LB media and dropped on non-selective LB agar plates. After 24 h of incubation at 30 °C, cells were scrapped off, resuspended in 500 µL LB-broth and 100 µL thereof were plated on selective LB-agar containing 50 µg mL^{-1} neomycin and 20 µg mL^{-1} polymyxin for counter selection. *P. polymyxa* conjugants were analyzed for successful transformation after 48 h incubation at 30 °C by cPCR. Confirmed knock out strains were plasmid cured by cultivation in LB broth at 37 °C without antibiotic selection pressure and subsequent replica plating on LB agar plates both with and without neomycin. Strains that did not grow on plates with selection marker were verified by sequencing of the target region and used for further experiments.

2.3.4. Sporulation assay

Five mL of LB medium were inoculated with 1 % preculture of the respective *P. polymyxa* strain and incubated at 37 °C, 300 rpm for 4 days. 500 µL of the culture broth were transferred to 1.5 mL reaction tubes and incubated at 85 °C in a water bath. Serial dilutions were afterwards plated on LB-agar plates and incubated at 30 °C for two days in order to determine the colony forming units (cfu).

2.4. Molecular biology

2.4.1. Oligonucleotide phosphorylation and annealing

1 µL of each oligonucleotide (100 µM) containing overhangs for Golden Gate assembly was mixed with 1 µL T4 ligase buffer, 1 µL polynucleotide kinase and 6 µL H_2O. The reaction mix was incubated at 37 °C for 30 min and afterwards heat inactivated at 95 °C for 5 min. Thereafter, the temperature was slowly decreased to 25 °C with a ramp of 0.5 °C every 6 seconds. Annealed oligonucleotides were diluted with 990 µL H_2O before utilization in Golden Gate assemblies.

2.4.2. Golden Gate assembly

For Golden Gate assembly, 0.75 µL BbsI, 0.25 µL T7-ligase, 1 µL 10x CutSmart Buffer, 1 µl 10 mM ATP, 1 µL 10 mM DTT, 100 ng plasmid backbone and 1 µL annealed gRNA oligos were mixed and filled with H_2O to a final volume of 10 µL. For assembly, the mixture was subjected to 20 cycles of heating to 37 °C for 5 min and afterwards incubating at 20 °C for 5 min. After heat inactivation for 20 min at 80 °C, the sample was cooled down to 12 °C. All of the reaction mix was used for the transformation of chemically competent *E. coli* cells without further purification.

2.4.3. Restriction cloning

For conventional restriction cloning, 3 µg of purified DNA fragments or plasmids were

digested using 1 µL restriction endonuclease and the appropriate buffer provided by the manufacturer for 30 min at 37 °C. For inserts, 1 µL Quick CIP were added to the reaction mix and incubated for additional 10 min at 37 °C. Digested products of the correct size were recovered after agarose gel electrophoresis (1 % (w/v) agarose, 1x TAE buffer, 120 V, 25 min) using the Nucleospin Gel and PCR Clean-up Kit according to the manufacturer's instructions. Digested plasmid and insert with complementary sticky overhangs were mixed in a 1:3 molar ratio and ligated overnight at 16 °C. After heat inactivation at 80 °C for 20 min, the reaction mix was used for the transformation of chemically competent *E. coli* cells.

2.4.4. Gibson assembly

Equimolar amount of each DNA fragment (minimum 50 ng) containing 20-30 bp overhangs were mixed with an equal volume of the 2 x Gibson master mix. For cloning of smaller fragments, a threefold excess of the insert was used. For assembly of complex constructs using four or more DNA fragments, overhangs were further analyzed using the NUPACK tool to avoid hairpin formation of single stranded DNA overhangs at 50 °C [166]. Isothermal cloning via Gibson assembly was performed at 50 °C for 60 min in a heating block or thermocycler. After cooling down to 4 °C, all of the reaction mix was used for the transformation of chemically competent *E. coli* cells without further purification.

2.4.5. CRISPR-Cas9 mediated genome editing

All gene knock-outs were performed as previously described by Rütering et al. (2017) [25]. In brief, gRNAs for the targeted genome regions were designed using Benchling CRISPR Design Tool. For each target, a minimum of two gRNAs were designed typically targeting distinct regions of the open reading frame of the target gene. Oligonucleotides were phosphorylated, annealed and cloned into pCasPP by Golden Gate assembly. Approximately 1 kB up- and downstream homology flanks for each targeted nucleotide sequence were amplified from genomic DNA of *P. polymyxa* DSM 365 using Phusion Polymerase according to the manufacturer's instructions and fused by overlap extension PCR via a 20 bp overlap. Homology flanks were cloned into pCasPP via Gibson assembly or molecular cloning after linearization by use of SpeI. For integration of new DNA After transformation of *E. coli* NEB Turbo, clones were analyzed for correct construct assembly by colony PCR (cPCR) and sequencing of the amplicons. Finally, correct constructs were transferred to chemical competent *E. coli* S17-1 cells for the following conjugational transformation of *P. polymyxa*.

2.5. Fermentation

2.5.1. Butanediol fermentation

2.5.1.1. Batch fermentation

Batch fermentations were conducted in 1 L DASGIP bioreactors (Eppendorf, Germany) with an initial volume of 550 mL. A single colony from a freshly streaked plate was used to inoculate 100 mL pre-culture medium by following incubation for 16 h at 30 °C, 160 rpm. 50 mL of this cultivation broth (diluted with pre-culture medium if necessary) were used to inoculate the bioreactor by an initial OD_{600} of 0.1. Fermentation was performed at 35 °C and constant aeration of 0.075 vvm. The stirrer was equipped with a 6-plate-rushton impeller placed 4 cm from the bottom of the shaft and constantly stirring at 300 rpm. The pH value was maintained at 6.0 and automatically adjusted with 2 M NaOH or 1.35 M H_3PO_4 as required. Foam control was performed using 1 % of antifoam B (Merck, Germany). For monitoring the process parameters, reactors were equipped with redox and pH probes. Oxygen transfer rate (OTR) and carbon dioxide production rate (CPR) were determined by online off-gas measurements by a DASGIP GA4 exhaust gas analyzer (Eppendorf, Germany). Under the assumption of pseudo-steady state and oxygen limited conditions, in which liquid gas concentrations are close to zero, the oxygen uptake rate (OUR) can be assumed as followed (Eq. 1):

$$OUR = OTR - \frac{tc_L}{td} \triangleq OTR \qquad (1)$$

Respiratory quotient (RQ) was calculated based on OUR and CPR (Eq. 2):

$$RQ = \frac{CPR}{OUR} \qquad (2)$$

2.5.1.2. Continous fermentation

Continuous fermentations were performed in 2 L reactors (Sartorius Biostat B plus, Sartorius AG, Germany) with a constant volume of 1.5 L. Aeration, agitation and temperature parameters were kept identical to the previously described batch fermentation. The pH value was maintained at 6.0 and automatically adjusted with 2 M NaOH or 1.35 M H_3PO_4 as required. For monitoring the process parameters, reactors were equipped with redox and pH probes. Foam control was performed using 1 % of antifoam B. Two 6-plate-rushton impellers were placed 3 cm and 6 cm from the bottom respectively. Continuous flow started 24 h after inoculation with a constant dilution of 0.042 h^{-1} and fermentation medium containing 46.2 g L^{-1} of glucose. Cell retention was guaranteed by use of a 0.2 µm filtration probe (TRACE analytics, Germany). Steady-state was assumed when product concentrations remained constant for at least 24 h.

2.5.2. EPS fermentation

Fermentative production of EPS was performed in bioreactors of variable scale. Small scale fermentation was performed in a 1 L benchtop DASGIP parallel bioreactor systems (Eppendorf, Germany) with a working volume of 500 mL equipped with a 6-blade Rushton impeller (2.5 cm from the bottom of the stirrer) over 28 h with a controlled pH of 6.8 and pO_2 saturation of 30 %. Aeration cascade was initially controlled by increasing stirring speed starting from 200 rpm to 600 rpm. Afterwards, gas flow was automatically increased starting from an initial rate of 2 L h^{-1} increasing to 10 L h^{-1}. Foam control was performed by the automatic addition of 1 % antifoam B. Medium scale fermentations were performed in 2 L Sartorius Biostat 2 reactor systems with a working volume of 1 L. The stirrer was equipped with two Rushton impellers (0.5 cm and 3.5 cm from the bottom respectively). Gas flow was adjusted to a 0.3 L min^{-1}- 1.5 L min^{-1} for the aeration cascade. Batch cultivations were started with an initial OD_{600} of 0.1 by inoculation with an appropriate volume of preculture. Cultivation were terminated after 26-30 h after depletion of glucose.

2.5.3. EPS downstream processing

After fermentation, biomass was separated by centrifugation (15,000 g, 20 °C, 20 min) followed by cross-flow filtration of the supernatant using 100 kDa filtration cassette (Hydrosart, Sartorius AG, Germany). Highly viscous EPS variants were diluted 1:10 with H_2O prior to centrifugation. Crossflow filtration was typically performed until the initial volume of the fermentation broth was reached or until the pressure of the cross-flow system was too high. Concentrated supernatant was afterwards slowly poured into two volumes of isopropanol. Precipitated EPS was collected with a spatula and dried for 24 h in a VDL53 vacuum oven at 40 °C (Binder, Germany). Dry weight of the obtained EPS was determined gravimetrically, before milling to a fine powder in a ball mill at 30 Hz for 1 min (Mixer Mill MM400, Retsch GmbH, Germany).

2.6. Analytical methods

2.6.1. Cell growth

Cell growth was determined by measuring optical density at 600 nm (OD_{600}) using a Ultraspec 10 spectrophotometer (Amersham Biosciences, UK). For cell dry weight (CDW) determination, 1 mL of fermentation broth was centrifuged at 24,000 g for 5 min. The remaining cell pellet was dried over night at 105 °C and the weight determined gravimetrically. Living cell count was determined by serial dilutions of the fermentation broth with 0.9 % NaCl and appropriate plating on LB-agar plates. Colony forming units (cfu) were determined after 48 h of incubation at 30 °C.

2.6.2. Photometric fluorescence assay

For fluorescence experiments, 3 mL of EPS medium supplemented with 50 µg mL^{-1} neomycin was inoculated with a single colony of the respective strains and grown overnight at 30 °C, 200 rpm. After 18 h, each strain was sub-cultured 1:100 in 3 mL selective EPS medium and grown for 24 h at 30 °C, 200 rpm. After 24 h, 100 µL were transferred to a 96 well microtiter plate and OD_{600}, GFP fluorescence (Ex. 488 nm Em. 515 nm) and mRFP fluorescence (Ex. 560 nm Em. 600 nm) measured in a Ultraspec 10 spectrophotometer (Amersham Biosciences, UK). Fluorescence values were normalized to OD_{600} in all experiments.

2.6.3. Quantitative RT-PCR

RNA extraction of positive samples of the GFP fluorescence assay as well as butanediol fermentation processes was performed using the Aurum Total RNA Mini Kit (BioRad, USA) according to the manufacturer's instructions. cDNA synthesis was conducted using iScript reverse transcriptase (BioRad, USA) according to the manufacturer's instructions using 1 µg total RNA template. qPCR reactions were performed in triplicates on a CFX-96 thermocycler using SsoAdvanced Universal SYBR Green Supermix (BioRad, USA) according to the manufacturer's instructions using 5 ng of cDNA as a template in 10 µL reaction volume. Negative controls without reverse transcriptase during cDNA synthesis were used in order to evaluate the absence of gDNA contaminations. Relative gene expression levels were calculated using the ΔΔCq method [167] using *gyrA* as a reference gene. After qPCR, a melting curve analysis was performed to confirm the presence of a single PCR product for each target. Designed primers were analyzed by the OligoAnalyzer Tool (IDT, USA) to avoid hairpin formation and self- and hetero dimer formation with free energy values more than 10 kcal mol^{-1}.

2.6.4. Carbohydrate fingerprint

Monomer composition of engineered EPS variants were analyzed by the 1-phenyl-3-methyl-5-pyrazolone-high throughput method (HT-PMP) [168]. In short, 0.1 % EPS solutions were hydrolyzed in a 96 well plate, sealed with a rubber mat and further covered by a custom-made metal device with 2 M TFA (90 min, 121 °C). Samples were neutralized with 3.2 % NH_4OH. 75 µL of PMP master mix (0.1 M methanolic PMP:0.4 % ammonium hydroxide 2:1) were added to 25 µL of neutralized hydrolysate and incubated at 70 °C for 100 min in a thermal cycler. 20 µL of derivatized samples were mixed with 130 µL of a 1:26 dilution of 0.5 M acetic acid and filtered with a 0.2 µm filter plate (1000 g, 2 min) followed by HPLC-UV-MS using an Ultimate 3000 RS HPLC system (Dionex, USA). Separation was performed on a reverse phase column (Gravity C18, 100 x 2 mm, 1.8 µm particle size, Macherey-Nagel, USA) set to 50 °C. Gradient elution was performed using a mobile phase A (5 mM ammonium acetate adjusted to pH 5.6 with 15 % acetonitrile) and mobile phase B (100 % acetonitrile) with a constant pump rate of

0.6 mL min^{-1}.

2.6.5. Partial hydrolysis of EPS samples

Analysis of partially hydrolyzed EPS samples was performed as described for carbohydrate fingerprinting (2.6.4). Differently, hydrolysis was performed at 90 °C for 105 min. After PMP derivatization, 25 µL 0.5 M acetic acid, 125 µL dH_2O and 500 µL acetonitrile were added to the reaction mix to quench the derivatization process and precipitate non-hydrolyzed polymer. Precipitate was collected by centrifugation (24,000 g, 5 min) and the obtained supernatant transferred to HPLC-vials. HPLC-UV-MS analysis was performed using a YMC-Triart Diol-HILIC column (100 x 2 mm, 1.9 µm particle size). Column temperature was set to 7 °C and a constant flow of 0.3 mL min^{-1} 5 mM ammonium-acetate (pH 5.6) with 15 % acetonitrile was applied.

2.6.6. Pyruvate assay

Pyruvate concentration was determined enzymatically of 1 g L^{-1} solutions of native and hydrolyzed EPS solutions. 100 µL of 1:10 diluted sample were mixed with 100 µL assay master mix (50 µM DA-64, 50 µM thiamine pyrophosphate, 100 µM $MgCl_2$ x 6 H_2O, 0.05 U mL^{-1} pyruvate oxidase, 0.2 U mL^{-1} horseradish peroxidase, 20 mM potassium phosphate buffer adjusted to pH 6.0) in 96 well microtiter plates and incubated at 37 °C, 700 rpm for 30 min. Quantification was performed by photometrical measurement of the color change at 727 nm and subtraction of the background signal at 540 nm. For the calibration curve, standards ranging from 0.5-100 µM pyruvate were used for each measurement.

2.6.7. Glucose assay

For the glucose assay, supernatant of fermentation broth or EPS solutions was diluted with H_2O to be within the calibration range of a glucose standard (2.5 µM – 500 µM). 50 µL sample or standard were mixed with 50 µL assay master mix (40 mM potassium phosphate buffer adjusted to pH 6.0, 1.5 mM ABTS, 0.4 U glucose oxidase, 0.02 U horseradish peroxidase) in a 96-well plate. The plate was sealed with a silicon cap mat and incubated on a microplate shaker (800 rpm, 30 °C, 30 min). Afterwards, glucose concentration was determined photometrically by measuring extinction at 418 nm and subtracting the background signal at 480 nm. Each sample was measured in technical triplicates.

2.6.8. Total carbohydrate analysis

Total carbohydrate content was determined with the phenol-sulphuric acid method. 20 µL of sample or standard solution (0.05 g L^{-1} – 5 g L^{-1}) were mixed with 180 µL assay mix (5 % (w/v) phenol solution in dH_2O and pre-cooled 96 % sulfuric acid in a 1:6 ratio) in a 96-well plate. The plate was covered with a plastic lid, incubated on a microplate shaker (800 rpm, RT, 5 min) and afterwards transferred to an incubation oven (80 °C, 35 min). Extinction was measured at 480 nm to calculate the total carbohydrate

concentration as glucose equivalents. Each sample was measured in technical triplicates.

2.6.9. Selective degradation of glycosyluronic acids

For glycosyluronic acid degradation 75 mg of polymer were dissolved in 10 mL anhydrous ethylenediamine. Lithium scrapes were added until a deep blue color was obtained and kept in this state for 1.5 h. The reaction was stopped by cooling down the solution in an ice bath and addition of 5 mL dH_2O. The solution was co-distilled with toluene until dryness. The polymer was reconstituted in 5 mL dH_2O and adjusted to a pH of 6 with 4 M TFA. Samples were then derivatized with PMP and analyzed as described above.

2.6.10. Methylation analysis of carbohydrate polymers

Linkage analysis was performed according to the Hakomori procedure [169]. Briefly, 1 mg polymer was lyophilized overnight and reconstituted in 250 µL anhydrous DMSO by stirring. The polyanion of the polymer was formed by the addition of 500 µL 2 M dimsyl sodium, stirring overnight. After freezing the sample, 250 µL methyl iodine were added and incubated for 1 h. Residual methyl iodine was blown off and the methylated polymer recovered by MeOH/$CHCl_3$ extraction. Polymer hydrolysis was performed using 2 M TFA (121 °C 2 h), reduced with 0.25 M NaH_4B (25 °C, 1 h) and acetylated (250 µL acetic anhydride, 250 µL pyridine, 100 °C, 20 min). Partially methylated alditol acetates were extracted with ethyl acetate and used for GC-MS analysis. For analysis of uronic acids, polymer samples were reduced prior to methylation analysis. 2 g L^{-1} polymer solutions were incubated in 156 mM *N*-ethyl-*N*-(3-dimethylaminopropyl)-carbodiimide hydrochloride at pH 4.75 for 1.5 h and reduced with 1 mM $NaHB_4$ for 1 h, followed by pH neutralization and dialysis of the samples. For GC analysis, a Rxi-5 Sil MS column (30 m, 0.25 mm ID, d_f 0.25 µm) column was used with a constant helium flow of 0.3 mL min^{-1}. Initial temperature was set to 120 °C and increased with a heat ramp of 7.5 °C min^{-1} to 180 °C, held constant for five minutes and afterwards further ramped up to 230 °C. Internal references to determine the retention times of partially methylated glycoside were obtained via Purdie-Irvine methylation of methyl-glycosides by removing samples of the incremental per-methylation every 3 h for a time periode of 12 h [170,171].

2.6.11. NMR analysis

The EPS was dissolved in ddH_2O and degraded using mechanical shearing by ten cycles on a Star Burst Mini (Wet Jet HJP-25001 CE, Sugino Machine Ltd., Uozu, Japan) and the water was afterwards removed by freeze-drying. To remove remaining protein material the EPS was redissolved in ddH_2O (~2% w/v solution), heated to 95 °C for 10 mins, centrifuged, and the supernant was transferred to a new tube and dried again by freeze-drying. From this sample, 10 mg of EPS was redissolved in 600 µL of deuterated water (99.9% D_2O, Sigma Aldrich) transferred to 5 mm NMR tube (Bruker,

LabScape Stream 103.5 mm length). All homo- and heteronuclear experiments were acquired at a temperature of 273.1 K on a Bruker AV-IIIHD 800 MHz spectrometer (Bruker BioSpin AG, Fälladen, Switzerland) equipped with a 5 mm cryogenic CP-TCI z-gradient probe. The following experiments were collected: 1D proton, 1H-13C heteronuclear single quantum coherence (HSQC) with multiplicity editing, 1H-13C heteronuclear two bond correlation (H2BC) spectroscopy, 1H-13C heteronuclear multiple bond coherence (HMBC) with suppression of one-bond correlations, and 1H-1H double quantum filtered correlation spectroscopy (DQF-COSY). The spectra were recorded, processed, and analyzed using the TopSpin 3.5 or 4.0.1 software (Bruker BioSpin AG).

2.6.12. HPLC analysis of organic acids and alcohols

For butanediol cultivations organic acids, alcohols and glucose were determined via a HPLC-UV-RID system (Dionex, USA) equipped with Rezex ROA-H^+ organic acid column (300 mm x 7.8 mm Phenomenex, USA). Column temperature was set to 70 °C and 2.5 mM H_2SO_4 was used as the mobile phase with a flow rate of 0.5 mL min^{-1}. Lactate and formate were measured via a photodiode array detector. Glucose and all other products were quantified using a refractive index detector.

2.6.13. Rheology

For rheological analysis, 1 % (w/w) solutions of each polymer were prepared in ddH_2O and 0.5% NaCl (85 mM) respectively. Conductivity of each solution was measured using an LF413T-ID electrode (Schott instruments, Germany) to determine residual salt concentrations of the fermentation broth (Table S2). Rheological measurements were conducted using a MCR 300 stress controlled rotational rheometer (Anton Paar, Austria) equipped with a CP 50-1 cone-plate measuring system (50 mm diameter, 1° cone angle, 50 µm cone truncation). All measurements, except temperature sweeps, were carried out at 20°C controlled by a TEK 150P temperature unit. After applying of the solution to the rheometer, all samples were incubated at 20 °C for 5 minutes before beginning the measurements. All experiments were conducted in technical triplicates.

2.6.13.1. Viscosity curves

Viscosity curves were measured using a logarithmically increased shear rate from 10^{-3} to 10^3 s^{-1} by measuring 3 data points per decade with decreasing measuring time of 100 - 5 s per data point.

2.6.13.2. Amplitude sweeps

Amplitude sweeps were measured using a logarithmically increasing shear stress amplitude from 10^{-1} to 10^3 Pa with a frequency of 1 Hz.

2.6.13.3. Frequency test

Frequency sweeps were carried out within the linear viscoelastic range (LVE) at a

logarithmically increasing frequency from 10^{-2} to 10 Hz.

2.6.13.4. Temperature sweeps

Temperature sweeps were performed within the LVE at a frequency of 1 Hz applying a temperature ramp from 20 to 75 °C with a heating rate of 4 °C min^{-1}. The edge of the cone-plate measuring system was covered with low viscosity paraffin oil (Carl Roth, Germany) to prevent evaporation.

2.6.13.5. Thixotropy test

Thixotropic behavior was evaluated by a three-stage oscillatory shear sequence. In the first stage, samples were subjected to shear stress within the LVE region followed by a high oscillatory shear of 10^3 Pa for 30 s. The structural recovery was then measured over 10 min within the LVE range.

3. Results

3.1. Engineering of the 2,3-butanediol pathway *of Paenibacillus polymxa* DSM 365

This publication deals with the metabolic engineering of the butanediol pathway in *P. polymyxa*. First, EPS production was eliminated by the combinatorial knock-out of the cluster encoding paenan synthesis (clu1), as well as *sacB* encoding a levansucrase. Furthermore, the formation of metabolically dormant endospores should be inhibited. Knock-out of the master transcription factor of the sporulation *spo0A* indeed resulted in the inability to form spores. However, both growth and 2,3-BDL production were severely reduced. In one of the model organisms for sporulation, *B. subtilis*, *spo0A* has demonstrated to directly or indirectly influence the expression of over 500 genes [172], which might also negatively affect 2,3-BDL synthesis in *P. polymyxa*. Consequently, *spoIIE*, encoding a serine phosphatase further downstream in the sporulation cascade, was deleted. While the endospore formation was inhibited, no negative influence on the growth behavior or the 2,3-BDL production was observed.

In a next step, the mixed-acid pathway was engineered to reduce the amount of side products and redirect the carbon flux towards 2,3-BDL. A pyruvate formate lyase was knocked-out to omit the production of formate. However, the strain showed severely reduced biomass formation as well as 2,3-BDL synthesis. In contrast, deletion of a lactate dehydrogenase resulted in a 59 % increased biomass formation and a 10 % improved 2,3-BDL titer. Analysis of continuous cultivation experiments revealed that at low product concentrations the productivity of the knock-out mutant was increased by 50 % indicating product inhibition in the batch fermentations. Decoupling of the proposed bottleneck enzyme *bdh* demonstrated highly increased production rates resulting in a volumetric productivity of 0.61 g L^{-1} h^{-1} and a final titer of 43.80 g L^{-1} after 72 h batch fermentation.

Design and execution of all experiments in this publication were performed by the first author in close collaboration with Jochen Schmid and Volker Sieber. Cloning and strain construction were performed by Christoph Schilling. Bioreactor fermentations were conducted by Rosario Ciccone and Christoph Schilling. Volker Sieber and Jochen Schmid provided technical and scientific advice. All authors contributed to content and language of the manuscript.

Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365

Christoph Schilling, Rosario Ciccone, Volker Sieber & Jochen Schmid

Metabolic Engineering

(2020)

DOI: 10.1016/j.ymben.2020.07.009

Metabolic Engineering 61 (2020) 381–388

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365

Christoph Schilling [a], Rosario Ciccone [a], Volker Sieber [a,b,c,d], Jochen Schmid [a,e,*]

[a] Chair of Chemistry of Biogenic Resources, Technical University of Munich, Campus for Biotechnology and Sustainability, Schulgasse 16, 94315, Straubing, Germany
[b] Fraunhofer IGB, Straubing Branch BioCat, Schulgasse 23, 94315, Straubing, Germany
[c] TUM Catalysis Research Center, Ernst-Otto-Fischer-Straße1, 85748, Garching, Germany
[d] The University of Queensland, School of Chemistry and Molecular Biosciences, 68 Copper Road, St. Lucia, 4072, Australia
[e] Department of Biotechnology and Food Science, Norwegian University of Science and Technology, 7034, Trondheim, Norway

ARTICLE INFO

Keywords:
Butanediol
Paenibacillus polymyxa
Lactate dehydrogenase
Mixed acid pathway
CRISPR-Cas9

ABSTRACT

Paenibacillus polymyxa is a Gram-positive, non-pathogenic soil bacterium that has been extensively investigated for the production of *R*,*R*-2,3-butanediol in exceptionally high enantiomeric purity. Rational metabolic engineering efforts to increase productivity and product titers were restricted due to limited genetic accessibility of the organism up to now. By use of CRISPR-Cas9 mediated genome editing, six metabolic mutant variants were generated and compared in batch fermentations for the first time. Downstream processing was facilitated by completely eliminating exopolysaccharide formation through the combined knockout of the *sacB* gene and the *clu1* region, encoding for the underlying enzymatic machinery of levan and paenan synthesis. Spore formation was inhibited by deletion of *spoIIE*, thereby disrupting the sporulation cascade of *P. polymyxa*. Optimization of the carbon flux towards 2,3-butanediol was achieved by deletion of the lactate dehydrogenase *ldh1* and decoupling of the butanediol dehydrogenase from its natural regulation via constitutive episomal expression. The improved strain showed 45 % increased productivity, reaching a final concentration of 43.8 g L^{-1} butanediol. A yield of 0.43 g g^{-1} glucose was achieved, accounting for 86 % of the theoretical maximum.

1. Introduction

2,3-Butanediol (2,3-BDL) is a high value platform chemical with a wide range of applications. In an industrial context, it is most commonly used as a precursor molecule for the conversion to other end products. It can be dehydrated to methyl ethyl ketone, which can be used as a fuel additive or solvent for paints and lacquers. Dehydrogenation yields diacetyl that is industrially utilized in high-value applications as buttery food aroma (Isagulyants and Belomestnykh, 1997). Furthermore, it can be converted to 1,3-butadiene which is applied as a bulk chemical for the production of polymers (Duan et al., 2016). A wide range of applications of 2,3-BDL is also based on its appearance in three stereoisomers – levo-BDL (*R*,*R*-BDL), dextro-BDL (*S*,*S*-BDL) and meso-BD (*R*,*S*-BDL). Particularly the optically active forms, levo-BDL and dextro-BDL are of high industrial interest and can be used for the synthesis of chiral specialty chemicals (Celińska and Grajek, 2009).

Seeking a bio-based industry independent of fossil based resources, there has been a long history of research in biotechnological production of 2,3-BDL (Freeman and Morrison, 1947; Fulmer et al., 1933). A variety of different organisms including *Klebsiella pneumoniae*, *Serratia marcescens*, *Bacillus* spp. or *Paenibacillus polymyxa* have been shown to produce 2,3-BDL in remarkable amounts. The highest product titer of 152 g L^{-1} with a productivity of 4.2 g L^{-1} h^{-1} has been described by *K. pneumoniae* in fed-batch fermentations (Ma et al., 2009). Remarkable titers of 139 g L^{-1} were also obtained with *S. marcescens* (Zhang et al., 2010). However, the pathogenic nature of *K. pneumoniae* and *S. marcescens*, which defines them a risk class 2 organism limits their utilization for industrial scale production processes. In contrast, *Bacillus* spp. such as *B. amyloliquefaciens* and *B. licheniformis* are generally regarded as safe (GRAS) and have also demonstrated high 2,3-BDL titers (Jurchescu et al., 2013; Yang et al., 2011). However, mainly the optical inactive meso-BDL isomer is produced by these organisms. Interestingly, high titers of 2,3-BDL up to 98 % purity can be produced by the Gram-positive, non-pathogenic, spore forming, soil bacterium *P. polymyxa* under microaerobic conditions and therefore represents a promising starting point for future optimization (Huo et al., 2010; Jeong

Abbreviations: 2,3-BDL, 2,3-butanediol; CDW, cell dry weight; wt, wildtype.

* Corresponding author. Chair of Chemistry of Biogenic Resources, Technical University of Munich, Campus for Biotechnology and Sustainability, Schulgasse 16, 94315, Straubing, Germany.

https://doi.org/10.1016/j.ymben.2020.07.009
Received 3 June 2020; Received in revised form 12 July 2020; Accepted 27 July 2020
Available online 7 August 2020

et al., 2019; Nakashimada et al., 1998).

The biosynthesis of 2,3-BDL follows a mixed acid fermentative pathway at microaerobic conditions (Fig. 1) (Ji et al., 2011a). In short, one mol of glucose is converted to pyruvate by the simultaneous generation of two mols of ATP and NADH via glycolysis. Consequently, two mol of pyruvate are converted to α-acetolactate, which is further decarboxylated to acetoin and finally dehydrogenated to butanediol while simultaneously one redox equivalent of NAD^+ is produced (Fig. 1). In absence of alternative electron acceptors during anaerobic and microaerobic conditions, the redox balance is maintained by the formation of additional side products (Clark, 1989). Depending on oxygen availability, lactate dehydrogenase, pyruvate formate lyase as well as the pyruvate dehydrogenase complex compete for the intermediate pyruvate. This finally results in the production of lactate, formate, ethanol or acidic acid (Alexeeva et al., 2003; Nakashimada et al., 1998).

The production of butanediol in *P. polymyxa* has been extensively studied towards improved process parameters and media composition (Häßler et al., 2012; Laube et al., 1984a, 1984b; Okonkwo et al., 2017b). While other organisms have been successfully engineered on a molecular level in order to increase the overall 2,3-BDL production, previous efforts on this in *P. polymyxa* were restricted due to its limited genetic accessibility. Based on our recently developed tool for CRISPR-Cas9 mediated genome editing in *P. polymyxa* DSM 365 (Rütering et al., 2017) we performed in this work the first intensified rational metabolic engineering of *Paenibacillus polymyxa* towards an increased 2,3-BDL biosynthesis. The focus of this study was to address aspects of downstream processing such as increased viscosity due to exopolysaccharide production or spore formation. *P. polymyxa* is known for the production of the exopolysaccharide (EPS) levan when growing on sucrose and a heteropolysaccharide named paenan when grown on monomeric carbohydrates such as glucose (Rütering et al., 2016). Both polymers can interfere with the downstream processing due to the increased viscosity of the fermentation broth and also can impair mass transfer during fermentation (Häßler et al., 2012). In addition, *P. polymyxa* forms metabolically dormant endospores as a stress response, which do not contribute to butanediol or EPS formation (Dworkin and Shah, 2010). Furthermore, spores represent a severe issue of bio-based industrial scale processes in which genetically modified organisms are used and any leakage from the manufacturing plant in form of product contamination must be prevented. Thus, sporulation should be eliminated without negatively impacting 2,3-BDL formation. Interestingly, the sporulation cascade in *P. polymyxa* differs from the well described cascade in *B. subtilis*. Particularly histidine kinases, which sense sporulation-specific signals and initiate a phosphorelay reaction are only poorly conserved or completely missing and require further investigation (Park et al., 2012).

Besides downstream impeding by-products such as EPS or spores, the mixed acid pathway itself should be engineered to redirect carbon flux from other side products towards 2,3-BDL in context of a rational metabolic engineering approach. 2,3-BDL constitutes a toxic end product in higher concentrations for the producing organism and biosynthesis is tightly regulated (Okonkwo et al., 2017a). Consequently, deregulation of the pathway might allow to overcome this bottleneck in *P. polymyxa*. Overexpression of genes involved in the butanediol pathway (Guo et al., 2014; Lu et al., 2010; Yang et al., 2013) or deletion of side products directly branching from the pathway, which affect redox equivalents and improve the carbon flux towards 2,3-BDL have previously proven to be valid strategies in other organism (Ji et al., 2011b; Yang et al., 2015; Zhang et al., 2018). In this study, we evaluated the effects of rational metabolic engineering strategies of the mixed-acid pathway on 2,3-BDL biosynthesis by a comparative experimental setup in *P. polymyxa* DSM 365 for the first time.

2. Materials and methods

2.1. Strain

P. polymyxa DSM 365 was acquired from the German Collection of Microorganisms and Cell Culture (DSMZ), Braunschweig, Germany. *Escherichia coli* NEB Turbo cells (New England Biolabs, USA) were used for any plasmid construction presented in this study. *E. coli* S17-1 (DSMZ strain DSM 9079) was used for transformation of *P. polymyxa* DSM 365 via conjugation. All strains were grown in LB media (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 10 g L^{-1} NaCl) and additionally supplemented with 50 μg ml^{-1} neomycin and 20 μg ml^{-1} polymyxin if required. All strains were stored in 30 % glycerol at −80 °C. Prior to liquid cultivation, strains were streaked on LB agar plates and incubated at 30 °C. All strains used or constructed in this study are listed in Table S1.

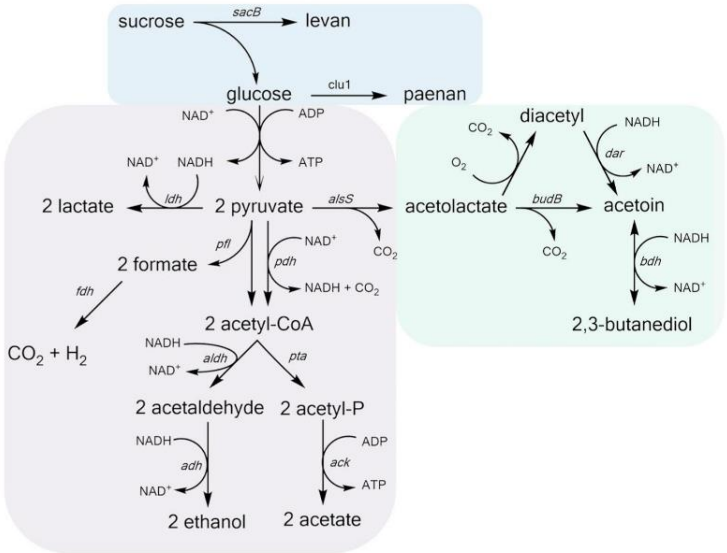


Fig. 1. Overview of 2,3-BDL biosynthesis pathway and byproducts in microaerobic conditions of *P. polymyxa* DSM 365 starting from sucrose as a substrate. Sucrose can be converted to the homopolysaccharide levan or to the heteroexopolysaccharide paenan. Butanediol is produced in a mixed acid fermentation pathway starting from pyruvate. Other side products such as lactate, ethanol, acetate or formate are competing over pyruvate as a substrate. Blue: polysaccharide synthesis, green: butanediol pathway, purple: mixed acid pathway, *sacB*: levansucrase, clu1: operon responsible for the production of paenan, *fdh*: formate dehydrogenase; *alsS*: acetolactate synthase; *budB*: acetolactate decarboxylase; *dar*: diacetyl reductase; *bdh*: butanediol dehydrogenase *aldh*: acetaldehyde dehydrogenase; *adh*: alcohol dehydrogenase; *pta*: phosphate acetyltransferase; *ack*: acetate kinase. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

2.2. CRISPR-Cas9 mediated genome editing

All gene knock-outs were performed as previously described (Rütering et al., 2017). In brief, gRNAs for the targeted genome regions were designed using Benchling CRISPR Design Tool. Oligonucleotides were phosphorylated, annealed and cloned into pCasPP by Golden Gate assembly. Approximately 1 kB up- and downstream homology flanks for each targeted nucleotide sequence were amplified from genomic DNA of *P. polymyxa* DSM 365, fused via overlap extension PCR and cloned into pCasPP after linearization by use of SpeI. After transformation of *E. coli* NEB Turbo, clones were analyzed for correct construct assembly by colony PCR (cPCR) and sequencing of the amplicons. Finally, correct constructs (Table S2) were transferred to chemical competent *E. coli* S17-1 cells for the following conjugational transfer to *P. polymyxa*. All primers used for the construction of the knockout plasmids are listed in Table S3. For each target, two gRNAs were tested and listed in the supplemental material if successful.

2.3. Conjugation based transformation of P. polymyxa DSM 365

P. polymyxa was transformed by conjugation using *E. coli* S17-1 harboring the various plasmids. Overnight cultures of donor and recipient strains were diluted 1:100 with selective or non-selective LB media respectively and cultivated at 37 °C for 3 h, 280 rpm. 900 µl of the recipient culture was heat shocked at 42 °C for 15 min and mixed with 300 µl of the donor strain. Cells were centrifuged at 6000 g for 2 min, resuspended in 400 µl LB media and dropped on non-selective LB agar plates. After 24 h of incubation at 30 °C, cells were scrapped off, resuspended in 500 µl LB-broth and 100 µl thereof were plated on selective LB-agar containing 50 µg ml^{-1} neomycin and 20 µg ml^{-1} polymyxin for counter selection. *P. polymyxa* conjugants were analyzed for successful transformation after 48 h incubation at 30 °C by cPCR. Confirmed knock out strains were plasmid cured by cultivation in LB broth at 37 °C without antibiotic selection pressure and subsequent replica plating on LB agar plates both with and without neomycin. Strains that did not grow on plates with selection marker were verified by sequencing of the target region and used for further experiments.

2.4. Fermentation media

All medium components were obtained from Carl Roth GmbH (Germany) if not indicated differently. Medium composition was adapted from a previous study (Okonkwo et al., 2017b). A single colony was used for inoculation of 50 ml pre-culture medium containing 60 g L^{-1} glucose, 5 g L^{-1} yeast extract, 5 g L^{-1} tryptone, 0.2 g L^{-1} $MgSO_4x7\ H_2O$ (Sigma Aldrich, USA), 3.5 g L^{-1} KH_2PO_4, 2.5 g L^{-1} K_2HPO_4. Fermentation medium components were autoclaved separately and contained 120 g L^{-1} glucose, 7 g L^{-1} glycerol, 5 g L^{-1} yeast extract, 3.5 g L^{-1} tryptone, 0.2 g L^{-1} $MgSO_4x7\ H_2O$, 3.5 g L^{-1} KH_2PO4, 2.5 g L^{-1} K_2HPO_4, 5 g L^{-1} ammonium acetate, 4 g L^{-1} $(NH_4)_2SO_4$ and 3 ml L^{-1} trace elements. Trace element solution contained 2.5 g L^{-1} iron sulfate heptahydrate, 2.1 g L^{-1} sodium tartrate dihydrate, 1.8 g L^{-1} manganese chloride dihydrate, 0.075 g L^{-1} cobalt chloride hexahydrate, 0.031 g L^{-1} copper sulfate pentahydrate, 0.258 g L^{-1} boric acid, 0.023 g L^{-1} sodium molybdate dihydrate and 0.021 g L^{-1} zinc chloride. Trace element solution was filter-sterilized and added to the media after cooling down to room temperature.

2.5. Batch fermentation

Batch fermentations were conducted in 1 L DASGIP bioreactors (Eppendorf, Germany) with an initial volume of 550 ml. A single colony from a freshly streaked plate was used to inoculate 100 ml pre-culture medium by following incubation for 16 h at 30 °C, 160 rpm. 50 ml of this cultivation broth (diluted with pre-culture medium if necessary) were used to inoculate the bioreactor by an initial OD_{600} of 0.1. Fermentation was performed at 35 °C and constant aeration of 0.075 vvm. The stirrer was equipped with a 6-plate-rushton impeller placed 4 cm from the bottom of the shaft and constantly stirring at 300 rpm. The pH value was maintained at 6.0 and automatically adjusted with 2 M NaOH or 1.35 M H_3PO_4 as required. Foam control was performed using 1 % of antifoam B (Merck, Germany). For monitoring the process parameters, reactors were equipped with redox and pH probes. Oxygen transfer rate (OTR) and carbon dioxide production rate (CPR) were determined by online off-gas measurements by a DASGIP GA4 exhaust gas analyzer (Eppendorf, Germany). Under the assumption of pseudo-steady state and oxygen limited conditions, in which liquid gas concentrations are close to zero, the oxygen uptake rate (OUR) can be assumed as followed (Eq. (1)):

$$OUR = OTR - \frac{ic_L}{td} \triangleq OTR \tag{1}$$

Respiratory quotient (RQ) was calculated based on OUR and CPR (Eq. (2)):

$$RQ = \frac{CPR}{OUR} \tag{2}$$

2.6. Continuous fermentation

Continuous fermentations were performed in 2 L reactors (Sartorius Biostat B plus, Sartorius AG, Germany) with a constant volume of 1.5 L. Aeration, agitation and temperature parameters were kept identical to the previously described batch fermentation. The pH value was maintained at 6.0 and automatically adjusted with 2 M NaOH or 1.35 M H_3PO_4 as required. For monitoring the process parameters, reactors were equipped with redox and pH probes. Foam control was performed using 1 % of antifoam B. Two 6-plate-rushton impellers were placed 3 cm and 6 cm from the bottom respectively. Continuous flow started 24 h after inoculation with a constant dilution of 0.042 h^{-1} and fermentation medium containing 46.2 g L^{-1} of glucose. Cell retention was guaranteed by use of a 0.2 µm filtration probe (TRACE analytics, Germany). Steady-state was assumed when product concentrations remained constant for at least 24 h.

2.7. Analytical methods

Cell growth was determined by measuring optical density at 600 nm (OD_{600}) using a Ultraspec 10 spectrophotometer (Amersham Biosciences, UK). For cell dry weight determination, 1 ml of fermentation broth was centrifuged at 24,000 g for 5 min. The resulting supernatant was filtered with 0.2 µm PTFE filters and utilized for HPLC-UV-RID analysis. The remaining cell pellet was dried over night at 105 °C. Living cell count was determined by serial dilutions of the fermentation broth with 0.9 % NaCl and appropriate plating on LB-agar plates. Colony forming units were determined after 48 h of incubation at 30 °C. Glucose and product concentrations were determined via a HPLC-UV-RID system (Dionex, USA) equipped with Rezex ROA-H^+ organic acid column (300 mm × 7.8 mm Phenomenex, USA). Column temperature was set to 70 °C and 2.5 mM H_2SO_4 was used as the mobile phase with a flow rate of 0.5 ml min^{-1}. All measured concentrations of 2,3-BDL in this publication represent solely the levo-stereoisomer of the alcohol if not explicitly noted differently. EPS concentrations were determined by centrifugation (8000 g, 15 min) of 100 ml aliquots of fermentation broth at the end of the cultivation process. Supernatant was slowly poured into 200 ml of isopropanol while stirring. Precipitated EPS was collected, dried overnight in a VDL53 vacuum oven at 40 °C (Binder, Germany) and following gravimetrical weight determination.

2.8. Sporulation assay

Five ml of LB medium were inoculated with 1 % preculture of the respective *P. polymyxa* strain and incubated at 37 °C, 300 rpm for 4 days. 500 µl of the culture broth were transferred to 1.5 ml reaction tubes and

incubated at 85 °C in a water bath. Serial dilutions were afterwards plated on LB-agar plates and incubated at 30 °C for two days in order to determine the colony forming units (cfu).

3. Results and discussion

3.1. Construction of knock out variants

Knockout constructs were generated to eliminate undesirable side-products of 2,3-BDL fermentations. EPS formation was targeted by deletion of a heteroexopolysaccharide cluster (*clu1*) combined with the levansucrase *sacB*. Sporulation was inhibited by deletions of key steps in the sporulation cascade, *spo0A* and *spoIIE*. In order to target metabolic pyruvate drains to redirect the carbon flux towards 2,3-BDL production pyruvate formate lyase (*pfl*) and a lactate dehydrogenase were deleted. Open reading frames of each targeted gene were completely removed by homologous repair assisted mechanisms mediated by CRISPR-Cas9 using 1 kb flanks up- and downstream of the target region, which were provided by the pCasPP plasmid. Deletion of each target was confirmed by cPCR and further verified by sequencing (Fig. S1). After plasmid curing, all verified knockout variants were evaluated and compared via 72 h batch fermentations for their yield in 2,3-BDL production as well as side product formation (Fig. 2). In addition, the process characteristics were analyzed (Table 1). The results of each generated metabolic mutant variant are discussed individually in the sections below.

3.2. Deletion of EPS formation

Recently, it has been shown that deletion of *sacB*, encoding for a levansucrase forming the fructose based carbohydrate polymer levan from sucrose, had a positive impact on 2,3-BDL synthesis (Okonkwo et al., 2020). However, *P. polymyxa* is known to produce an additional heteroexopolysaccharide, which might additionally interfere with the downstream processing of the targeted product 2,3-BDL (Rütering et al., 2016). In order to redirect the carbon flux towards butanediol formation, essential parts of the operon, responsible for the heteropolysaccharide biosynthesis (*clu1*) were deleted in addition to the *sacB* gene. As a result, no EPS formation was observed, and the cell growth was directly affected (Table 1). Both, OD_{600} and CDW increased over the course of the batch fermentation compared to the wildtype strain, but the final 2,3-BDL titer remained nearly unaffected (Fig. S3). However, lactate production was increased by 70 %, while the BDL-yield was slightly decreased to 0.37 g g^{-1} (Fig. S4). In previous studies (Häßler et al., 2012; Okonkwo et al., 2017b) EPS concentrations of 5–54 g L^{-1} were reported using similar cultivation conditions to our experimental set up. Compared to previous studies, the obtained EPS concentrationsin this study were generally very low with approximately 450 mg L^{1} for the wildtype strain by use of the DASGIP bioreactors. Utilization of glucose as a carbon source already prevented levan production such giving BDL titers of up to 54 g L^{-1} (Häßler et al., 2012). In combination with low aeration and low inoculum volume, this resulted in only small amounts of heteroexopolysaccharide formation by *P. polymyxa*. Nevertheless, by genetic engineering it was possible for the first time to completely eliminate EPS formation in *P. polymyxa* by use of the Δ*clu1* Δ*sacB* strain. However, it appeared that only minor amounts of carbon could be successfully redirected towards 2,3-BDL production (Fig. 2). Despite higher biomass and increased OD_{600} values after 72 h of batch cultivation, the final 2,3-BDL titer of 30.39 g L^{-1} remained comparable to the wildtype strain (Table 1). While the glucose consumption slightly increased, mostof the carbon source was redirected to the non-toxic end product lactate, rather than 2,3-BDL (Table S6).

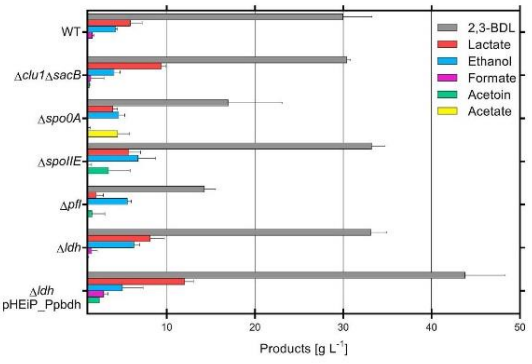


Fig. 2. Product profiles of *P. polymyxa* DSM 365 wt and engineered variants after 72 h of controlled batch cultivation in DASGIP bioreactors applying microaerobic conditions (0.075 vvm). Engineering of EPS formation (Δ*clu1* Δ*sacB*), sporulation (Δ*spo0A*, Δ*spoIIE*), or the mixed acid pathway (Δ*pfl*, Δ*ldh*) altered the final product titer of 2,3-BDL and side products. Plasmid based overexpression of butanediol dehydrogenase increased 2,3-BDL concentration by 46 %. Product concentrations are the result of three independent fermentation processes.

3.3. Deletion of spore formation

To prevent initiation of the sporulation cascade, *spo0A* encoding the corresponding sporulation stage 0 master transcription factor that initiates the sporulation cascade upon phosphorylation (Fujita et al., 2005), was deleted. As a result, no spores were detected within the sporulation assay and microscopic analysis for any sample taken during the batch fermentations (Table S5). However, the growth rate of 0.225 h^{-1} of the Δ*spo0A* mutant was significantly reduced compared to 0.311 h^{-1} of the wildtype (Table 1). For *P. polymyxa* wt and all other knockout variants, the stationary phase of the batch fermentations was reached within the first 24 h. The *spo0A* knockout variant continued to show exponential growth at later stages and only surpassed OD and CDW values of the wildtype strain towards the end of the process (Figure S3). In parallel, formation of 2,3-BDL was strongly affected. After 72 h of batch fermentation, only 16.97 g L^{-1} of 2,3-BDL were obtained (Fig. 2). The wildtype of P. polymyxa exclusively produced R,R-BDL under microaerobic conditions, whereas for the *spo0A* knockout variant elevated concentrations of 0.6 g L-*1* meso-BDL where obtained. As a result, enantiomeric purity of R,R-BD decreased to 96.5 %. I *B. subtilis,* Spo0A is known to directly or indirectly influence the expression of over 500 genes, which might explain the negative impact on BDL-biosynthesis (Molle et al., 2003). No consensus binding sites of Spo0A can be found in close proximity to any gene of the butanediol pathway, but Spo0A has been shown to control the expression of several other transcription factors. RAST (Rapid Annotation using Subsystem Technology) analysis (Overbeek et al., 2014) of the genome of *P. polymyxa* revealed only one single butanediol dehydrogenase, which belongs to the medium-chain dehydrogenase/reductase family known to form the R-alcohol (Yu et al., 2011). Meso-BDL is expected to be formed by spontaneous conversion of acetolactate to diacetyl and the following transformation to S-acetoin by diacetylreductase (Haukeli and Lie, 1978; Zhang et al., 2018). Therefore, we hypothesize that the master regulator Spo0A indirectly affects butanediol synthesis in *P. polymyxa* by downregulation of acetolactate decarboxylase. Consequently, less *R, R*-BDL is produced, while in contrast more S-acetoin is formed. Interestingly, butanediol dehydrogenase accepts both acetoin enantiomers as a substrate (Yu et al., 2011) and can thus further convert S-acetoin to meso-BDL resulting in a lower enantiomeric purity.

In addition to 2,3-BDL synthesis, the *spo0A* deletion had an impact on the acetate metabolism (Fig. S4). Initial acetate levels have been shown to shift mixed acid fermentation towards the synthesis of butanediol (Nakashimada et al., 2000; Yu and Saddler, 1985) and were therefore part of the media composition. While all other strains used in this study consumed initial acetate to the end of the growth phase, the *spo0A* deletion mutant did not deplete acetate at all. Instead, the concentration slightly

Table 1
Overview of process characteristics of *P. polymyxa* DSM 365 wt and metabolic mutant variants. Results were obtained by three independent 72 h batch fermentations at microaerobic conditions. The CDW was determined after 72 h of cultivation from a single bioreactor.

Strain	Y_{PS} [g g^{-1}]	Productivity [g L^{-1} h^{-1}]	CDW [g L^{-1}]	μ [h^{-1}]	2,3-BDLTiter [g L^{-1}]	EPS [mg L^{-1}]
wt	0.42 ± 0.01	0.42 ± 0.04	3.40 ± 0.12	0.311 ± 0.026	29.96 ± 2.69	443 ± 48
Δ*clu1* Δ*sacB*	0.37 ± 0.02	0.42 ± 0.00	4.60 ± 0.17	0.330 ± 0.012	30.39 ± 0.34	0 ± 0
Δ*spo0A*	0.23 ± 0.08	0.24 ± 0.07	4.80 ± 0.14	0.225 ± 0.019	16.97 ± 4.99	240 ± 14
Δ*spoIIE*	0.36 ± 0.02	0.46 ± 0.02	4.70 ± 0.26	0.337 ± 0.012	33.21 ± 1.25	413 ± 12
Δ*pfl*	0.25 ± 0.07	0.20 ± 0.01	2.73 ± 0.57	0.125 ± 0.014	14.24 ± 1.04	187 ± 17
Δ*ldh1*	0.39 ± 0.01	0.46 ± 0.02	5.40 ± 0.95	0.295 ± 0.015	33.11 ± 1.44	480 ± 22
Δ*ldh1* pHEiP_Ppbdh	0.43 ± 0.06	0.61 ± 0.05	4.93 ± 0.41	0.290 ± 0.026	43.80 ± 3.66	403 ± 12

increased. Studies conducted in *Clostridium cellulolyticum* and *C. beijerinckii* revealed that the shift from acetogenesis to solvent formation in these organism was controlled by Spo0A. (Ravagnani et al., 2000). While the precise mechanism of Spo0A influencing mixed acid fermentation in *P. polymyxa* has not been extensively investigated yet, we assume that the transcription factor putatively acts in a similar way as in *Clostridium* spp. Consequently, the knockout of *spo0A* partially redirected the carbon flux from butanediol towards acetate.

Due to the observed negative global effects of *spo0A* deletion on 2,3-BDL synthesis, a second gene downstream of the sporulation cascade, not known to interfere with other metabolic pathways was targeted. The *spoIIE* gene encodes a serine phosphatase involved in the early stage II sporulation phase (Bi et al., 2011; Scotcher and Bennett, 2005). As a result of the knockout, no spore formation could be observed in the batch fermentation and negative effects on the mixed-acid metabolism of *P. polymyxa* were eliminated. The yield slightly decreased to 0.36 g g^{-1}, however the final titer of 33.21 g L^{-1} was similar to the wildtype strain (Table 1). While deletion of *spo0A* had a negative impact on biomass formation, the *spoIIE* mutant did not show decreased cell growth (Fig. S3). Instead, a rather slightly increased biomass formation was observed, which might explain the 10 % increased 2,3-BDL concentration at the end of batch fermentation.

Deficiency in spore formation was proven by heat treatment of cultures grown for 4 days in LB broth (Table S4). While the wildtype strain survived 45 min of heat exposure at 85 °C, both sporulation knockout variants Δ*spo0A* and Δ*spoIIE* did not display any vegetative cell growth after the same time of heat treatment. Thus, spore free mutants were successfully established by gene deletion-based disruption of the sporulation cascade. Due to multiple negative side effects of the deletion of the master regulator Spo0A, only *P. polymyxa* Δ*spoIIE* proved to be suitable for butanediol production.

3.4. Elimination of formate formation

At anaerobic and microaerobic conditions, glucose is converted to pyruvate through the Embden-Meyerhof-Parnas pathway, which then branches and the three enzymes, lactate dehydrogenase, pyruvate formate lyase as well as acetoacetate synthase compete for the intermediate pyruvate. The expression level of the individual enzymes highly depends on the oxygen availability under microaerobic conditions (De Mas et al., 1988; Nakashimada et al., 1998). Deletion of *pfl* has previously been shown to be an efficient strategy to increase 2,3-BDL production by decreasing formate formation (Jung et al., 2014). Only minor levels of formate were detected for all investigated strains of this study. However, formate can be further converted to CO_2 and H_2 by formate dehydrogenase, thus removing the intermediate from the fermentation broth. Deletion of *pfl* indeed resulted in the complete elimination of formate (Fig. 2). Although, cell growth was severely reduced and OD_{600} as well as cell dry weight were decreased by 46 % and 19 % respectively compared to the wild type (Table 1). Consequently, the 2,3-BDL titer was also significantly attenuated to 14.21 g L^{-1}. Decreased 2,3-BDL concentrations were not only a result of lower biomass, but also the product yield was reduced significantly compared to other mutant variants as well as the wild type. At microaerobic conditions, both pyruvate formate lyase (Pfl) and the pyruvate dehydrogenase complex (Pdhc) compete for the intermediate pyruvate (Alexeeva et al., 2003). In comparison to previous studies for 2,3-BDL optimization, the applied aeration of 0.075 vvm was rather low explaining the shift towards *pfl* expression. Pyruvate formate lyase represents a redox valve, that can compensate for the redox imbalance of the butanediol pathway (Adlakha et al., 2015). Consequently, the decreased cell growth might be a result of intracellular redox imbalance as well as the limited availability of acetyl-CoA, which is also a key intermediate in the biosynthesis of fatty acids (Krivoruchko et al., 2015). Due to the low 2,3-BDL yield and volumetric productivity of the Δ*pfl* strain, no further characterization of this strain was conducted even though we demonstrated that the formate production can be completely omitted.

3.5. Knockout of a lactate dehydrogenase to redirect the carbon flux

The most prominent side product under the applied condition was lactate. In sum, four putative genes encoding lactate dehydrogenases were identified by RAST analysis of the genome of *P. polymyxa* DSM 365. Out of those genes, *ldh1* showed 60.19 % sequence identity to *ldh1* of *B. subtilis* and was therefore chosen as a target in order to remove or reduce lactate formation as a by-product (Hecker et al., 2009). However, during batch fermentations no attenuation of the lactate concentrations could be observed (Fig. S4). The final butanediol titer after 72 h of fermentation slightly increased by 10 % towards 33.11 g L^{-1}, but also lactate concentration was significantly elevated (Fig. 2). Based on our results, we concluded that the remaining three putative lactate dehydrogenases in *P. polymyxa* DSM 365 might compensate the deletion of *ldh1*. Interestingly, OD_{600} as well as the cell dry weight was significantly

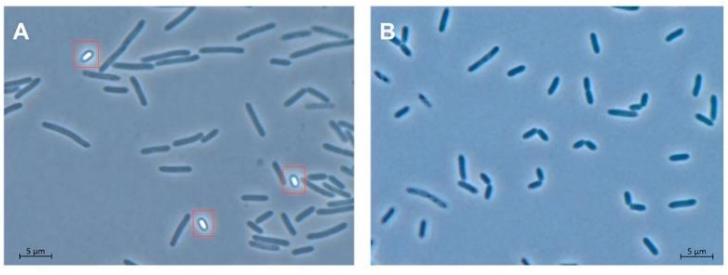


Fig. 3. Phase-contrast microscopic image (100× magnifications, Axio Observer Z1, Zeiss, Germany) of a 24 h *P. polymyxa* wt (A) and Δ*ldh1* (B) culture in pre-culture medium at 30 °C and 300 rpm. The wildtype strain showed matured rod-shaped cell morphology, while Δ*ldh1* formed averaged smaller cells that were observed in a permanent stage of cell division. Spores are marked by red rectangles. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

Table 2
Substrate consumption and product formation rates of *P. polymyxa* DSM 365 wt and Δ*ldh1* in continuous process mode with a dilution rate of 0.042 h^{-1} during steady-state conditions.

Volumetric rates [g L^{-1} h^{-1}]							
Strain	r_{Gluc}	r_{BDL}	$r_{Ethanol}$	$r_{Acetate}$	$r_{Acetoin}$	$r_{Lactate}$	OD_{600}
wt	−1.87 ± 0.03	0.60 ± 0.03	0.19 ± 0.01	−0.13 ± 0.02	0.05 ± 0.01	0.02 ± 0.00	6.9 ± 0.10
Δldh1	−1.99 ± 0.08	0.90 ± 0.03	0.31 ± 0.01	−0.13 ± 0.01	0.03 ± 0.00	0.17 ± 0.02	14.9 ± 0.10

increased compared to the wildtype strain (Table 1). Additionally, the early stage productivity of fermentation prior to product inhibition was significantly increased. Microscopic analysis of the fermentation broth (Fig. 3) revealed that the wildtype strain mainly formed matured cells with an average length of 5.4 μm (n = 19). In contrast, the Δ*ldh1* mutant showed most cells in a permanent stage of cell division, thus reducing the average length to 3.3 μm (n = 17). Additionally, no spore formation was observed over the whole course of the fermentation. Contrary, the wildtype strain did show spore formation, even in early stages of the fermentation course. The exact mechanism for this behavior remains to be investigated but similar effects have previously been described for *Enterobacter aerogenes, K. pneumoniae and B. subtilis* (Guo et al., 2014; Jung et al., 2012; Lu et al., 2010; Yang et al., 2015).

For further investigation of the effects of *ldh1* knockout on the productivity of *P. polymyxa* DSM 365, the knockout as well as the wildtype strain were cultivated in a continuous reactor system with cell retention to reduce effects of product inhibition on the organism. When product concentration remained constant for a minimum of 24 h, cells were considered in steady-state and production rates were calculated (Table 2).

Although the glucose consumption rate slightly increased from 1.87 g L^{-1} for the wildtype to 1.99 g L^{-1} for the Δ*ldh*1 variant, the butanediol production rate increased by 50 % from 0.60 to 0.90 g L^{-1} h^{-1}. Maximal OD_{600} values of Δ*ldh*1 were significantly higher with maximum values of 15 compared to 7 for the wildtype strain. Further, the cell dry mass increased by 25 % to 5.31 g L^{-1} in comparison to 4.24 g L^{-1} for the wildtype. Thus, the increase in optical density values could be mostly attributed to the different light scattering caused by changed cell morphology. Living cell count was increased 4.4-fold, thus confirming the presence of many more small cells rather than a few larger cells (Fig. 4).

3.6. Decoupling of butanediol dehydrogenase increased productivity

During the continuous process mode, a constant production rate of 0.90 g L^{-1} h^{-1} under low product concentrations was observed for the Δ*ldh1* strain. (Table 2). In batch fermentations, productivity in the early phase of fermentation reached similar values (Fig. S3). However, after an initial peak in production until 24 h, the production rates declined (Fig. S3). This is particularly based on the downregulation of the expression of the butanediol dehydrogenase (*bdh*) which is in correlation with increasing acetoin concentrations. In order to evaluate this observation, a plasmid encoding a PSP01-*bdh* expression cassette for the constitutive expression of bdh was designed and transferred to *P. polymyxa* Δ*ldh1* (Lee et al., 1980) (Fig. S2). Overexpression of *bdh* has previously been shown to positively impact 2,3-BDL production in *B. amyloliquefaciens* (Yang et al., 2013). Batch fermentation of, *P. polymyxa* Δ*ldh1* pHEiP_Ppbdh showed that the formation of side products such as ethanol or lactate were not affected and similar titers as for the Δ*ldh1* variant without additional expression of *bdh* were obtained (Fig. 2). Solely acetoin concentrations at the end of the cultivation were significantly increased from 0.96 g L^{-1} to 2.35 g L^{-1}. Furthermore, OD_{600} as well as CDW remained comparable to the parent strain indicating a neglectable metabolic burden by the constitutively plasmid-based expression (Table 1).

In the early process phase, no significant differences in the production rates of 2,3-BDL were observed. Volumetric productivity remained high until a product concentration of about 30 g L^{-1} was reached, resulting in a final space time yield (STY) of 0.61 g L^{-1} h^{-1} at 72 h after inoculation (Fig. 4 A). In contrast, the volumetric productivity of the wildtype strain without episomal expression of *Ppbdh* already declined after an initial spike surpassing 10 g L^{-1} of the product. Over the complete process a significantly reduced productivity of 0.42 g L^{-1} h^{-1} was obtained. Therefore, it is concluded that *bdh* expression was successfully decoupled from the native regulation system resulting in a 45 % increased volumetric productivity in *P. polymyxa* Δ*ldh1* pHEiP_Ppbdh. The respiratory quotient (RQ) is a parameter to easily monitor productivity in anaerobic and microaerobic process conditions (Zeng et al., 1994). After an initial increase of the RQ peak at 17 h after inoculation, values of the wildtype strain quickly dropped to a level of 2 and remained constant for the rest of the process. The initial spike has previously been associated with increased ethanol production proceeding butanediol biosynthesis in *B. licheniformis*,

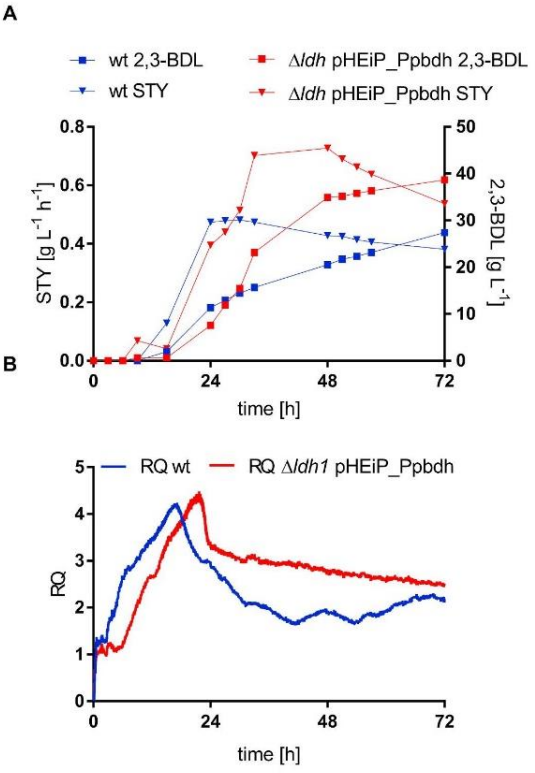


Fig. 4. A) Space time yield (STY) of 2,3-BDL over the course of a 72 h batch cultivation (n = 1). The wildtype productivity (blue) stagnated at 0.45 g L^{-1} h^{-1} after reaching a product titer of 10 g L^{-1}, while *ldh1* pHEiP_Ppbdh (red) with decoupled *bdh* expression showed significantly increased productivity, which only declined after surpassing 30 g L^{-1} 2,3-BDL. Triangles: STY; Squares: 2,3-BDL titer B) Increased productivity was reflected in the respiratory quotient (RQ). Decoupled expression of *bdh* led to increased RQ values over the course of the fermentation, while in the wildtype strain CO_2 production rates quickly dropped after reaching inhibitory 2,3-BDL concentrations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

which could be also observed in *P. polymyxa* (Heyman et al., 2020). In contrast to the wildtype, the strain with decoupled *bdh* expression reached higher product concentrations, which was also reflected in higher RQ values only slowly declining subsequent to the initial peak emergence (Fig. 5 B). For that reason, the monitoring of the RQ moreover confirmed higher metabolic activity of *P. polymyxa Δldh1* pHEiP_Ppbdh.

Shaking flask experiments in *B. licheniformis* showed the active conversion of 2,3-BDL to acetoin after the complete consumption of glucose by the reversed reaction of butanediol dehydrogenase (Heyman et al., 2020).*P. polymyxa Δldh1* pHEiP_Ppbdh also accumulated increased concentrations of 2.35 g L^{-1} acetoin compared to 0.86 g L^{-1} in the wildtype strain after 72 h of fermentation. However, in the same time frame glucose was never fully depleted in any fermentation process and also no decrease in 2,3-BDL was determined for *P. polymyxa Δldh1* pHEiP_Ppbdh. Even though the volumetric productivity was increased in our experiments, a 2,3-BDL mediated feedback inhibition was still observed at higher product concentrations indicating additional regulatory mechanisms limiting the production of butanediol.

Due to the redox imbalance of butanediol synthesis, other pathways are required to ensure NAD^+ equilibrium. When the NAD^+/NADH balance cannot be restored otherwise under microaerobic conditions, glucose is converted to redox-neutral end products such as ethanol and lactate. Pyruvate formate lyase enables conversion of pyruvate to acetyl-CoA without generating any redox equivalents. Subsequently, two mols of NAD^+ are regenerated by the combined action of aldehyde dehydrogenase and alcohol dehydrogenase. This route represents the most efficient pathway for redox regeneration in *P. polymyxa* which allows for the compensation of the redox imbalance of the butanediol biosynthesis (Adlakha et al., 2015). Interestingly, in *P. polymyxa Δldh1* pHEiP_Ppbdh the redox balance was not accomplished by expanding the carbon flux towards ethanol. (Table S6). However, pyruvate was used for the redox neutral production of lactate. With declining productivity of 2,3-BDL after approximately 48 h of cultivation, lactate formation started to result in a final titer of 12.0 g L^{-1} (Fig. 5, Table S4). Ethanol production is more efficient for the redox balance, but also represents additional solvent stress next to 2,3-BDL. Even though the conversion of pyruvate to lactate yields only one mol NAD^+ per mol substrate, solvent stress mitigation by production of lactate seemed to be favorable in pH-controlled bioreactor cultivations. In contrast, the more efficient ethanol formation was favored over lactate production during continuous fermentation mode of *P. polymyxa Δldh1*, in which both products ethanol and 2,3-BDL were constantly removed from the fermentation broth. Therefore, we conclude that not only oxygen supply but also alcohol concentration is highly important for the regulation of the mixed acid pathway in *P. polymyxa*.

Previous studies have demonstrated a toxicity threshold concentration of 50 g L^{-1} 2,3-BDL for *P. polymyxa* (Okonkwo et al., 2017a). Due to the toxicity of 2,3-BDL, an economically feasible process might not be easily achieved without further engineering strategies in *P. polymyxa*. Notably, other studies demonstrated significantly higher 2,3-BDL titers of up to 150 g L^{-1} for *Bacillus* strains such as *B. amyloliquefaciens* and *B. licheniformis* (Jurchescu et al., 2013; Yang et al., 2011). A better understanding of the underlying resistance mechanism of these strains is required to further engineer *P. polymyxa* towards similar product titers.

Even so, by overexpression of *bdh* in *P. polymyxa Δldh1* we successfully demonstrated the possibility of decoupling the last enzymatic step in butanediol formation. Most importantly we were able to increase the productivity of 2,3-BDL by 45 % compared to the wildtype strain.

4. Conclusion

P. polymyxa revealed to be an highly interesting GRAS platform organism with the advantage of producing exceptionally pure *R,R*–BDL. Based on our recently developed novel CRISPR-Cas9 mediated genome editing tools this study represents the first targeted approach of rational metabolic engineering of *P. polymyxa* DSM 365 for a simplified downstream processing of 2,3-BDL by complete elimination of EPS production

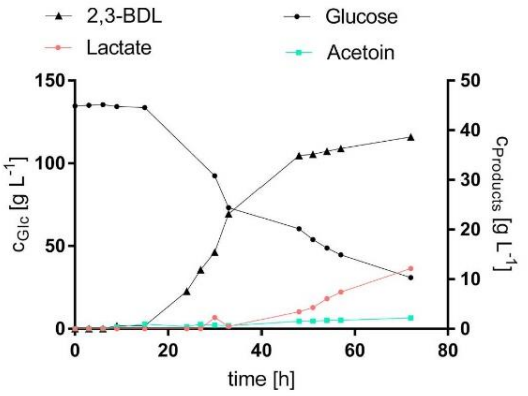


Fig. 5. Production of 2,3-BDL (black triangles) in *P. polymyxa* DSM 365 Δ*ldh* pHEiP Ppbdh showed increased productivity. After 48 h acetoin (green) started to accumulate while also lactate biosynthesis (red) was induced to ensure a redox neutral end product for glucose. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

as well as spore formation. Six mutant variants were generated to comparatively study the effects on the 2,3-BDL metabolism in *P. polymyxa* DSM 365. For the first time, EPS formation was eliminated by the combined knockout of the cluster responsible for heteroexopolysaccharide synthesis and the levansucrase SacB. Furthermore, we demonstrated that deletion of the sporulation stage 0 transcription factor Spo0A negatively impacted 2,3-BDL formation as well as the enantiomeric purity of the diol. However, sporulation was successfully inhibited by the knockout of *spoIIE* without affecting 2,3-BDL synthesis. In addition, by deletion of *ldh1*, which showed the highest similarity to the lactate dehydrogenase, proved to be responsible for lactate production in *B. subtilis*, we realized a completely different growth behavior of *P. polymyxa* in combination with reduction of spore formation. By that, we were able to obtain much higher cell densities, product titers as well as cell viability. This represents a highly promising approach towards optimized production strains, by massive engineering of the native growth behavior, what might be transferred to already established production strains, to further increase the efficiency.

While similar strategies to engineer the mixed acid pathway have been applied in *Klebsiella sp.* and *Enterobacter sp.*, we demonstrated that not all of them could be universally transferred to *P. polymyxa.* Novel insight of the metabolism of this organism at microaerobic conditions and the role of key enzymes such as Pfl were obtained, further contributing to future research on this interesting alternative host organism and closely related *Bacillus* species. Finally, rational modification of the mixed acid pathway and flux optimization towards butanediol by decoupling the expression of *bdh* resulted in a strain with 45 % increased productivity and 2,3-BDL titer.

Author statement

Christoph Schilling, Jochen Schmid: Conceptualization, Methodology, **Rosario Ciccone, Christoph Schilling, Jochen Schmid:** Experiments, Data Analysis, Writing- Original draft preparation. **Jochen Schmid:** Supervision.: **Rosario Ciccone, Christoph Schilling, Jochen Schmid, Volker Sieber:** Writing- Reviewing and Editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymben.2020.07.009.

References

Adlakha, N., Pfau, T., Ebenhöh, O., Yazdani, S.S., 2015. Insight into metabolic pathways of the potential biofuel producer, *Paenibacillus polymyxa* ICGEB2008. Biotechnol. Biofuels 8, 159. https://doi.org/10.1186/s13068-015-0338-4.

Alexeeva, S., Hellingwerf, K.J., Teixeira de Mattos, M.J., 2003. Requirement of ArcA for redox regulation in *Escherichia coli* under microaerobic but not anaerobic or aerobic conditions. J. Bacteriol. 185, 204–209. https://doi.org/10.1128/JB.185.1.204-209.2003.

Bi, C., Jones, S.W., Hess, D.R., Tracy, B.P., Papoutsakis, E.T., 2011. SpoIIE is necessary for asymmetric division, sporulation, and expression of F, E, and G but does not control solvent production in *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. 193, 5130–5137. https://doi.org/10.1128/JB.05474-11.

Celińska, E., Grajek, W., 2009. Biotechnological production of 2,3-butanediol—current state and prospects. Biotechnol. Adv. 27, 715–725. https://doi.org/10.1016/j.biotechadv.2009.05.002.

Clark, D.P., 1989. The fermentation pathways of *Escherichia coli*. FEMS Microbiol. Lett. 63, 223–234. https://doi.org/10.1111/j.1574-6968.1989.tb03398.x.

De Mas, C., Jansen, N.B., Tsao, G.T., 1988. Production of optically active 2,3 butanediol by *Bacillus polymyxa*. Biotechnol. Bioeng. 31, 366–377. https://doi.org/10.1002/bit.260310413.

Duan, H., Yamada, Y., Sato, S., 2016. Future prospect of the production of 1,3-butadiene from butanediols. Chem. Lett. 45, 1036–1047. https://doi.org/10.1246/cl.160595.

Dworkin, J., Shah, I.M., 2010. Exit from dormancy in microbial organisms. Nat. Rev. Microbiol. 8, 890–896. https://doi.org/10.1038/nrmicro2453.

Freeman, G.G., Morrison, K.I., 1947. Production of 2:3-butylene glycol by fermentation of molasses. J. Soc. Chem. Ind. 66, 216–221. https://doi.org/10.1002/jctb.5000660706.

Fujita, M., Gonzalez-Pastor, J.E., Losick, R., 2005. High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. J. Bacteriol. 187, 1357–1368. https://doi.org/10.1128/JB.187.4.1357-1368.2005.

Fulmer, E.I., Christensen, L.M., Kendall, A.R., 1933. Production of 2,3 butylene glycol by fermentation. Ind. Eng. Chem. 25, 798–800. https://doi.org/10.1021/ie50283a019.

Guo, X.-W., Zhang, Y.-H., Cao, C.-H., Shen, T., Wu, M.-Y., Chen, Y.-F., Zhang, C.-Y., Xiao, D.-G., 2014. Enhanced production of 2,3-butanediol by overexpressing acetolactate synthase and acetoin reductase in *Klebsiella pneumoniae*: enhanced Production of 2,3 Butanediol. Biotechnol. Appl. Biochem. 61, 707–715. https://doi.org/10.1002/bab.1217.

Häßler, T., Schieder, D., Pfaller, R., Faulstich, M., Sieber, V., 2012. Enhanced fed-batch fermentation of 2,3-butanediol by *Paenibacillus polymyxa* DSM 365. Bioresour. Technol. 124, 237–244. https://doi.org/10.1016/j.biortech.2012.08.047.

Haukeli, A.D., Lie, S., 1978. Conversion of α-acetolactate and removal of diacetyl A kinetic study. J. Inst. Brew. 84, 85–89. https://doi.org/10.1002/j.2050-0416.1978.tb03843.x.

Hecker, M., Reder, A., Fuchs, S., Pagels, M., Engelmann, S., 2009. Physiological proteomics and stress/starvation responses in *Bacillus subtilis* and *Staphylococcus aureus*. Res. Microbiol. 160, 245–258. https://doi.org/10.1016/j.resmic.2009.03.008.

Heyman, B., Tulke, H., Putri, S.P., Fukusaki, E., Büchs, J., 2020. Online monitoring of the respiratory quotient reveals metabolic phases during microaerobic 2,3-butanediol production with *Bacillus licheniformis*. Eng. Life Sci. 20, 133–144. https://doi.org/10.1002/elsc.201900121.

Huo, Z., Yang, X., Raza, W., Huang, Q., Xu, Y., Shen, Q., 2010. Investigation of factors influencing spore germination of *Paenibacillus polymyxa* ACCC10252 and SQR 21. Appl. Microbiol. Biotechnol. 87, 527–536. https://doi.org/10.1007/s00253-010-2520-8.

Isagulyants, G.V., Belomestnykh, I.P., 1997. Butadione synthesis by dehydrogenation and oxidative dehydrogenation of 2,3-butanediol. In: Studies in Surface Science and Catalysis. Elsevier, pp. 415–420. https://doi.org/10.1016/S0167-2991(97)80932-9.

Jeong, H., Choi, S.-K., Ryu, C.-M., Park, S.-H., 2019. Chronicle of a soil bacterium: *Paenibacillus polymyxa* E681 as a tiny guardian of plant and human health. Front. Microbiol. 10, 467. https://doi.org/10.3389/fmicb.2019.00467.

Ji, X.-J., Huang, H., Ouyang, P.-K., 2011a. Microbial 2,3-butanediol production: a state-of-the-art review. Biotechnol. Adv. 29, 351–364. https://doi.org/10.1016/j.biotechadv.2011.01.007.

Ji, X.-J., Nie, Z.-K., Huang, H., Ren, L.-J., Peng, C., Ouyang, P.-K., 2011b. Elimination of carbon catabolite repression in *Klebsiella oxytoca* for efficient 2,3-butanediol production from glucose–xylose mixtures. Appl. Microbiol. Biotechnol. 89, 1119–1125. https://doi.org/10.1007/s00253-010-2940-5.

Jung, M.-Y., Mazumdar, S., Shin, S.H., Yang, K.-S., Lee, J., Oh, M.-K., 2014. Improvement of 2,3-butanediol yield in *Klebsiella pneumoniae* by deletion of the pyruvate formate-lyase gene. Appl. Environ. Microbiol. 80, 6195–6203. https://doi.org/10.1128/AEM.02069-14.

Jung, M.-Y., Ng, C.Y., Song, H., Lee, J., Oh, M.-K., 2012. Deletion of lactate dehydrogenase in *Enterobacter aerogenes* to enhance 2,3 butanediol production. Appl. Microbiol. Biotechnol. 95, 461–469. https://doi.org/10.1007/s00253-012-3883-9.

Jurchescu, I. M., Hamann, J., Zhou, X., Ortmann, T., Kuenz, A., Prüße, U., Lang, S., 2013. Enhanced 2,3-butanediol production in fed-batch cultures of free and immobilized *Bacillus licheniformis* DSM 8785. Appl. Microbiol. Biotechnol. 97, 6715–6723. https://doi.org/10.1007/s00253-013-4981-z.

Krivoruchko, A., Zhang, Y., Siewers, V., Chen, Y., Nielsen, J., 2015. Microbial acetyl-CoA metabolism and metabolic engineering. Metab. Eng. 28, 28–42. https://doi.org/10.1016/j.ymben.2014.11.009.

Laube, V.M., Groleau, D., Martin, S.M., 1984a. The effect of yeast extract on the fermentation of glucose to 2,3-butanediol by *Bacillus polymyxa*. Biotechnol. Lett. 6, 535–540. https://doi.org/10.1007/BF00139998.

Laube, V.M., Groleau, D., Martin, S.M., 1984b. 2,3 Butanediol production from xylose and other hemicellulosic components by *Bacillus polymyxa*. Biotechnol. Lett. 6, 257–262. https://doi.org/10.1007/BF00140047.

Lu, Y., Zhao, H., Zhang, C., Lai, Q., Wu, X., Xing, X.-H., 2010. Alteration of hydrogen metabolism of ldh-deleted *Enterobacter aerogenes* by overexpression of NAD(+)-dependent formate dehydrogenase. Appl. Microbiol. Biotechnol. 86, 255–262. https://doi.org/10.1007/s00253-009-2274-3.

Ma, C., Wang, A., Qin, J., Li, L., Ai, X., Jiang, T., Tang, H., Xu, P., 2009. Enhanced 2,3-butanediol production by *Klebsiella pneumoniae* SDM. Appl. Microbiol. Biotechnol. 82, 49–57. https://doi.org/10.1007/s00253-008-1732-7.

Molle, V., Fujita, M., Jensen, S.T., Eichenberger, P., González-Pastor, J.E., Liu, J.S., Losick, R., 2003. The Spo0A regulon of *Bacillus subtilis*: the Spo0A regulon. Mol. Microbiol. 50, 1683–1701. https://doi.org/10.1046/j.1365-2958.2003.03818.x.

Nakashimada, Y., Kanai, K., Nishio, N., 1998. Optimization of dilution rate, pH and oxygen supply on optical purity of 2, 3-butanediol produced by *Paenibacillus polymyxa* in chemostat culture. Biotechnol. Lett. 20, 1133–1138. https://doi.org/10.1023/A:1005324403186.

Nakashimada, Y., Marwoto, B., Kashiwamura, T., Kakizono, T., Nishio, N., 2000. Enhanced 2,3-butanediol production by addition of acetic acid in *Paenibacillus polymyxa*. J. Biosci. Bioeng. 90, 661–664. https://doi.org/10.1016/S1389-1723(00)90013-6.

Okonkwo, C., Ujor, V., Cornish, K., Ezeji, T.C., 2020. Inactivation of levansucrase gene in *Paenibacillus polymyxa* DSM 365 diminishes exopolysaccharide biosynthesis during 2,3 butanediol fermentation. Appl. Environ. Microbiol. AEM. https://doi.org/10.1128/AEM.00196-20, 00196-20, acm;AEM.00196-20v1.

Okonkwo, C., Ujor, V., Ezeji, T.C., 2017a. Investigation of relationship between 2,3-butanediol toxicity and production during growth of *Paenibacillus polymyxa*. N. Biotech. 34, 23–31. https://doi.org/10.1016/j.nbt.2016.10.006.

Okonkwo, C., Ujor, V., Mishra, P., Ezeji, T., 2017b. Process development for enhanced 2,3-butanediol production by *Paenibacillus polymyxa* DSM 365. Fermentation 3, 18. https://doi.org/10.3390/fermentation3020018.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A.R., Xia, F., Stevens, R., 2014. The SEED and the rapid annotation of microbial genomes using subsystems Technology (RAST). Nucleic Acids Res. 42, D206–D214. https://doi.org/10.1093/nar/gkt1226.

Park, S. Y., Park, S. H., Choi, S. K., 2012. Characterization of sporulation histidine kinases of *Paenibacillus polymyxa*. Res. Microbiol. 163, 272–278. https://doi.org/10.1016/j.resmic.2012.02.003.

Ravagnani, A., Jennert, K.C.B., Steiner, E., Grunberg, R., Jefferies, J.R., Wilkinson, S.R., Young, D.I., Tidswell, E.C., Brown, D.P., Youngman, P., Morris, J.G., Young, M., 2000. Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. Mol. Microbiol. 37, 1172–1185. https://doi.org/10.1046/j.1365-2958.2000.02071.x.

Rütering, M., Cress, B.F., Schilling, M., Rühmann, B., Koffas, M.A.G., Sieber, V., Schmid, J., 2017. Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. Synth. Biol. 2 https://doi.org/10.1093/synbio/ysx007.

Rütering, M., Schmid, J., Rühmann, B., Schilling, M., Sieber, V., 2016. Controlled production of polysaccharides–exploiting nutrient supply for levan and heteropolysaccharide formation in *Paenibacillus* sp. Carbohydr. Polym. 148, 326–334. https://doi.org/10.1016/j.carbpol.2016.04.074.

Scotcher, M.C., Bennett, G.N., 2005. SpoIIE regulates sporulation but does not directly affect solventogenesis in *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. 187, 1930–1936. https://doi.org/10.1128/JB.187.6.1930-1936.2005.

Yang, T., Rao, Z., Hu, G., Zhang, X., Liu, M., Dai, Y., Xu, M., Xu, Z., Yang, S.-T., 2015. Metabolic engineering of *Bacillus subtilis* for redistributing the carbon flux to 2,3-butanediol by manipulating NADH levels. Biotechnol. Biofuels 8. https://doi.org/10.1186/s13068-015-0320-1.

Yang, T., Rao, Z., Zhang, X., Lin, Q., Xia, H., Xu, Z., Yang, S., 2011. Production of 2,3-butanediol from glucose by GRAS microorganism *Bacillus amyloliquefaciens*. J. Basic Microbiol. 51, 650–658. https://doi.org/10.1002/jobm.201100033.

Yang, T., Rao, Z., Zhang, X., Xu, M., Xu, Z., Yang, S.-T., 2013. Improved production of 2,3-butanediol in *Bacillus amyloliquefaciens* by over-expression of glyceraldehyde-3-phosphate dehydrogenase and 2,3-butanediol dehydrogenase. PloS One 8, e76149. https://doi.org/10.1371/journal.pone.0076149.

Yu, B., Sun, J., Bommareddy, R.R., Song, L., Zeng, A.-P., 2011. Novel (2 *R* ,3 *R*)-2,3-Butanediol dehydrogenase from potential industrial strain *Paenibacillus polymyxa* ATCC 12321. Appl. Environ. Microbiol. 77, 4230–4233. https://doi.org/10.1128/AEM.02998-10.

Yu, E.K.C., Saddler, J.N., 1985. Biomass conversion to butanediol by simultaneous saccharification and fermentation. Trends Biotechnol. 3, 100–104. https://doi.org/10.1016/0167-7799(85)90093-9.

Zeng, A. P., Byun, T. G., Posten, C., Deckwer, W. D., 1994. Use of respiratory quotient as a control parameter for optimum oxygen supply and scale up of 2,3 butanediol production under microaerobic conditions. Biotechnol. Bioeng. 44, 1107–1114. https://doi.org/10.1002/bit.260440912.

Zhang, L., Cao, C., Jiang, R., Xu, H., Xue, F., Huang, W., Ni, H., Gao, J., 2018. Production of R,R-2,3-butanediol of ultra-high optical purity from *Paenibacillus polymyxa* ZJ-9 using homologous recombination. Bioresour. Technol. 261, 272–278. https://doi.org/10.1016/j.biortech.2018.04.036.

Zhang, L., Yang, Y., Sun, J., Shen, Y., Wei, D., Zhu, J., Chu, J., 2010. Microbial production of 2,3 butanediol by a mutagenized strain of *Serratia marcescens* H30. Bioresour. Technol. 101 https://doi.org/10.1016/j.biortech.2009.10.052, 1961–1967.

3.2. CRISPR driven functional analysis of the biosynthesis of paenan

3.2.1. Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365

P. polymyxa is an avid producer of various carbohydrate polymers [71,77]. Using sucrose as the carbon source, a levan type polyfructan was produced by numerous strains [79,173]. Using glucose as a carbon source, *P. polymyxa* 2H2 has shown to secret a highly complex heteroexopolysaccharide termed paenan [69]. Multiple publications in the past have dealt with this polymer and structurally related EPS variants from other *P. polymyxa* strains [68,75,174]. However, due to the seemingly high complexity of this biopolymer, structural elucidation remained impossible or inconclusive so far.

In the following manuscript, a previously developed tool for CRISPR-Cas9 mediated genome engineering was used to rationally engineer single and combinatorial knock-out variants of the paenan cluster of *P. polymyxa* DSM 365 [25]. By this approach, a fucose containing polysaccharide was enriched and characterized in detail. A combinatorial approach using complementary analytical methods to investigate the underlying polymer structure was applied. In order to determine the monomer composition, the previously established HT-PMP method was used [175]. Furthermore, to identify the sequence of individual monomers, partial hydrolysis experiments as well as the specific degradation of uronic acid residues were performed. Linkage pattern analysis was conducted by methylation experiments and NMR. By combining results of all these analytical methods, it was discovered that the heteroexopolysaccharide produced by *P. polymyxa* DSM 365 contains a secondary fucose containing polysaccharide termed ‘paenan II’ with a distinct substructure.

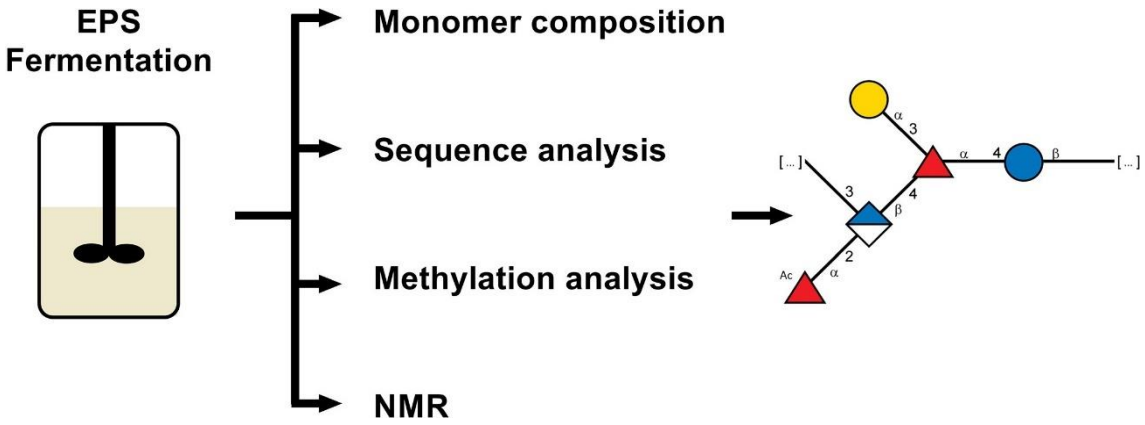
Design and planning of this study was performed by Christoph Schilling in collaboration with Jochen Schmid, Broder Rühmann and Volker Sieber. Strain construction, fermentation and carbohydrate analysis were performed by the author. Broder Rühmann gave scientific advice and support for analytical measurements. NMR analysis and evaluation were performed by Leesa J. Klau and Finn Aachmann. All authors contributed to content and language of the manuscripts and provided scientific or technical advice.

Structural elucidation of the fucose containing exopolysaccharide produced by *Paenibacillus polymyxa* DSM 365

Christoph Schilling, Leesa J. Klau, Broder Rühmann, Finn Aachmann, Jochen Schmid & Volker Sieber

Carbohydrate Polymers

(2022)



DOI: 10.1016/j.carbpol.2021.118951

Carbohydrate Polymers 278 (2022) 118951

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365

Christoph Schilling [a], Leesa J. Klau [b], Finn L. Aachmann [b], Broder Rühmann [a], Jochen Schmid [a,b,c], Volker Sieber [a,d,e,*]

[a] Chair of Chemistry of Biogenic Resources, Technical University of Munich, Campus for Biotechnology and Sustainability, Schulgasse 16, 94315 Straubing, Germany
[b] Department of Biotechnology and Food Science, NTNU Norwegian University of Science and Technology, Sem Sælands vei 6/8, 7491 Trondheim, Norway
[c] Institute for Molecular Microbiology and Biotechnology, Westfälische Wilhelms-Universität Münster, Corrensstrasse 3, 48149 Münster, Germany
[d] School of Chemistry and Molecular Biosciences, The University of Queensland, 68 Copper Road, St. Lucia 4072, Australia
[e] TUM Catalysis Research Center, Ernst-Otto-Fischer-Straße1, 85748, Garching, Germany

ARTICLE INFO

Keywords:
Paenibacillus polymyxa
Paenan
CRISPR-Cas9
Exopolysaccharides
Fucose

ABSTRACT

Paenibacillus polymyxa is an avid producer of exopolysaccharides of industrial interest. However, due to the complexity of the polymer composition, structural elucidation of the polysaccharide remained unfeasible for a long time. By using a CRISPR-Cas9 mediated knock-out strategy, all single glycosyltransferases as well as the Wzy polymerases were individually deleted in the corresponding gene cluster for the first time. Thereby, it was observed that the main polymer fraction was completely suppressed (or deleted) and a pure minor fucose containing polysaccharide could be isolated, which was named paenan II. Applying this combinatorial approach, the monosaccharide composition, sequence and linkage pattern of this novel polymer was determined via HPLC-MS, GC–MS and NMR. Furthermore, we demonstrated that the knock-out of the glycosyltransferases PepQ, PepT, PepU and PepV as well as of the Wzy polymerase PepG led to the absence of paenan II, attributing those enzymes to the assembly of the repeating unit.

1. Introduction

Bacterial exopolysaccharides (EPS) represent a highly diverse class of carbohydrate polymers that are secreted to the extracellular environment (Sutherland, 1972). The natural function of these polysaccharides comprises the colonialization of surfaces as biofilms and barrier affects against abiotic environmental stress factors such as extreme pH, desiccation or antibiotics (Lehman & Long, 2013; Moradali & Rehm, 2020). EPS can also play an essential role in the host-pathogen interaction of certain microorganisms (Cress et al., 2014). Bacterial biopolymers consist of structurally diverse glycosidic repeating units, which can be decorated by pyruvate, acetate, succinate, glycerate or other chemical modifications. Due to this enormous structural variability and therefore physicochemical versatility, various EPS are of high interest for different industrial applications as e.g. rheological modifiers in food, cosmetics, and pharmaceuticals or as additives for oil recovery, as well as in construction chemistry (Jang et al., 2015; Paul et al., 1986; Plank, 2005; Schilling, Badri, et al., 2020).

The biosynthetic pathways for the production of bacterial exopolysaccharides can be divided into four groups. Homopolymers, that are composed of a single type of glycoside are typically produced in a cell membrane associated, synthase dependent pathway or by extracellular enzymes that use the energy of the glycosidic bond of di- or trimeric carbohydrates (Moradali & Rehm, 2020; Schmid, 2018). In the so-called ABC-pathway, linear heteroglycans are often intracellularly assembled and exported by the name-giving ATP-dependent transporters (Moreira et al., 2000). Complex branched bacterial heteroexopolysaccharides are commonly produced by the so called Wzx/Wzy pathway (Cuthbertson et al., 2009; Islam & Lam, 2014; Whitfield, 1995).

For this pathway, an initiating glycosyl-phosphotransferase transfers the first sugar residue to an undecaprenylphosphate (UndPP) lipid anchor at the inner cell membrane. Subsequently, highly specific

Abbreviations: GT, glycosyltransferase; PMP, 1-phenyl-3-methyl-5-pyrazolone; Hex, hexose; Gal, galactose; GlcA, glucuronic acid; Fuc, fucose; Glc, glucose; KO, knock-out.

* Corresponding author at: Technical University of Munich, Campus for Biotechnology and Sustainability, Schulgasse 16, 94315 Straubing, Germany.
E-mail addresses: christoph.schilling@tum.de (C. Schilling), leesa.klau@ntnu.no (L.J. Klau), finn.l.aachmann@ntnu.no (F.L. Aachmann), broder.ruehmann@tum.de (B. Rühmann), jochen.schmid@uni-muenster.de (J. Schmid), sieber@tum.de (V. Sieber).

https://doi.org/10.1016/j.carbpol.2021.118951
Received 14 September 2021; Received in revised form 11 November 2021; Accepted 26 November 2021
Available online 2 December 2021

glycosyltransferases (GT) elongate the nascent oligomer using the binding energy of activated sugar nucleotides. The fully assembled repeating unit is then translocated towards the periplasmic space by a Wzx flippase (Liu et al., 1996). Afterwards, polymerization is performed by a Wzy protein and consequently exported to the cellular environment (Woodward et al., 2010). A Wzz chain length regulator interacts with the Wzy polymerase to confer a distinct molecular weight of the biopolymer chain length (Morona & Nath, 2015; Nath & Morona, 2015). While the general biosynthesis pathway of heteroexopolysaccharides is well characterized, the specific function of individual glycosyltransferases involved in the stepwise assembly of the repeating unit is often poorly understood and the precise mechanism of the underlying enzymatic machinery is only described for few polymers (Becker, 1993; Ielpi et al., 1993; Katzen et al., 1998; Pollock et al., 1998). For one of the most prominent representatives, xanthan, it took a total of 13 years from the initial characterization of the polymer to the final elucidation of the structure and additional 23 years to determine the functional relationship of individual enzymes of the biosynthesis cluster (Becker et al., 1998; Hassler & Doherty, 1990; Jansson et al., 1975; Jeanes et al., 1961; Katzen et al., 1998). However, this knowledge is ultimately required for the rational genetic engineering of the host organism in order to tailor the physicochemical properties of any polymer towards specific applications (Gansbiller et al., 2019).

Paenibacillus polymyxa is a Gram-positive, non-pathogenic, spore forming, soil bacteria capable of nitrogen fixation and phosphate solubilization (Jeong et al., 2019). It is commercially used as a biofertilizer in agriculture (Mohd Din et al., 2019; Xie et al., 2014) and has demonstrated promising properties for the fermentative production of optically pure 2,3-butanediol (De Mas et al., 1988; Schilling, Ciccone, et al., 2020). Depending on the applied C/N ratio and carbon source, *P. polymyxa* revealed to be a potent producer of a levan-type polyfructan when utilizing sucrose as a substrate and a heteroexopolysaccharide named paenan when using glucose or glycerol (Rütering et al., 2016; Rütering et al., 2018). Despite intensive research focusing on the heteroexopolysaccharides of *P. polymyxa* using conventional methods a complete structural elucidation of the polymer remained elusive due to the high complexity (Liang et al., 2014; Madden et al., 1986; Raza et al., 2011). The complex structure of the polymer is also reflected by the underlying 35 kb biosynthesis cluster of *P. polymyxa* DSM 365, comprising 28 coding sequences with two initiating glycosyltransferases, eleven GTs, two Wzx-flippases and two Wzy polymerases (Rütering et al., 2017). Due to the presence of additional redundant copies of the flippase and polymerase proteins within this cluster, it has been hypothesized that the operon might encode for two distinct exopolysaccharides. While it is quite common among bacteria to produce different EPS depending on the environmental conditions, the corresponding biosynthesis machinery is usually encoded in separate gene clusters (Bentley et al., 2008; Pérez-Mendoza et al., 2015). By conventional precipitation with isopropanol a separation of different polymers could not be achieved for *P. polymyxa* when cultivated on heteropolysaccharides forming conditions. Gel permeation chromatography of the obtained polymer did not conclusively confirm the presence of a second minor polymer (Rütering et al., 2017). However, due to the high selectivity of GTs it is possible to delete individual genes and generate modified polysaccharide variants of distinct rheological properties (Katzen et al., 1998). Analysis of combinatorial variants can then be used to elucidate both, the structure as well as the putative enzyme function of the knocked-out genes. Clustered regularly interspaced short palindromic repeats (CRISPR) endonuclease mediated genome engineering allows the generation of combinatorial knock-outs by precisely deleting the open reading frame of targeted genes via homologous recombination (Fig. S1) in a relatively short time span (Cobb et al., 2015; Jiang et al., 2013; Rütering et al., 2017).

In this study, we describe the CRISPR-Cas9 driven structural elucidation of a highly complex polysaccharide mixture produced by *P. polymyxa* DSM 365. Via the combinatorial deletion of the underlying biosynthesis encoding genes, both the biopolymer structure, as well as the function of certain GTs was characterized. A detailed understanding of the biosynthesis pathways for exopolysaccharides is ultimately required to realize tailor-made biopolymers for the rational modification towards specific applications.

2. Materials and methods

2.1. Strains and media

P. polymyxa DSM 365 was acquired from the German Collection of Microorganisms and Cell Culture (DSMZ, Germany). *Escherichia coli* NEB Turbo cells (New England Biolabs, USA) were used for any plasmid construction presented in this study. *E. coli* S17–1 (DSMZ strain DSM 9079) was used for transformation of *P. polymyxa* DSM 365 via conjugation. Strains constructed in this study are listed in Table S1. All medium components were obtained from Carl Roth GmbH (Germany) if not indicated differently. For cloning procedures, strains were grown in LB media (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 10 g L^{-1} NaCl) and additionally supplemented with 50 μg mL^{-1} neomycin and 20 μg mL^{-1} polymyxin if required. All strains were stored in 30% glycerol at −80 °C. Prior to cultivation, strains were streaked on LB agar plates and grown at 30 °C for 24 h.

Fermentation medium contained 30 g L^{-1} glucose, 0.05 g L^{-1} $CaCl_2 \times 2\ H_2O$, 5 g L^{-1} tryptone, 1.33 g L^{-1} $MgSO_4 \times 7\ H_2O$, 1.67 g L^{-1} KH_2PO_4, 2 mL L^{-1} RPMI 1640 vitamins solution (Merck, Germany) and 1 mL L^{-1} trace elements solution (2.5 g L^{-1} $FeSO_4$, 2.1 g L^{-1} $C_4H_4O_6Na_2 \times 2\ H_2O$, 1.8 g L^{-1} $MnCl_2 \times 4\ H_2O$, 0.258 g L^{-1} H_3BO_3, 0.031 g L^{-1} $CuSO_4 \times 5\ H_2O$, 0.023 g L^{-1} $NaMoO_4 \times 2\ H_2O$, 0.075 g L^{-1} $CoCl_2 \times 7\ H_2O$, 0.021 g L^{-1} $ZnCl_2$). Preculture medium was prepared equal to the fermentation medium except of a reduced glucose concentration of 10 g L^{-1} and additional 20 g L^{-1} MOPS buffered to pH 7.

2.2. CRISPR-Cas9 mediated knock-out of glycosyltransferases

All knock-outs were performed as described previously (Rütering et al., 2017). Briefly, gRNAs for each target were cloned into the plasmid pCasPP via Golden Gate assembly using the restriction endonuclease *Bbs*I. Afterwards, 1 kb up- and downstream homology flanks of the gene of interest were ligated into a unique *Spe*I site, followed by transformation of chemically competent *E. coli* S17–1. *P. polymyxa* was transformed by conjugation using *E. coli* S17–1 harboring the various plasmids. Overnight cultures of donor and recipient strains were diluted 1:100 with selective or non-selective LB media respectively and cultivated at 37 °C for 3 h, 280 rpm. 900 μL of the recipient culture was heat shocked at 42 °C for 15 min and mixed with 300 μL of the donor strain. Cells were centrifuged at 6000 xg for 2 min, resuspended in 800 μL LB media and dropped on non-selective LB agar plates. After 24 h of incubation at 30 °C, cells were scrapped off, resuspended in 500 μL LB-broth and 100 μL thereof plated on selective LB-agar containing 50 μg mL^{-1} neomycin and 20 μg mL^{-1} polymyxin for counter selection. *P. polymyxa* conjugants were analyzed for successful transformation after 48 h incubation at 30 °C by colony PCR. In confirmed knock-out strains the corresponding plasmid was removed by cultivation in LB broth without antibiotic selection pressure and subsequent replica plating on LB agar plates both with and without neomycin. Strains that grew on plates without neomycin but not on agar with selection marker were considered as plasmid-free. Mutant strains were verified by sequencing of the target region and used for further experiments. All plasmids and oligonucleotides used to obtain knock-out strains are listed in the supplemental materials (Tables S2 and S3).

2.3. Fermentative EPS production

Fermentative production of EPS was performed in 1 L benchtop DASGIP parallel bioreactor systems (Eppendorf, Germany) with a

working volume of 500 mL equipped with a 6-blade Rushton impeller over 28 h with a controlled pH of 6.8 and pO_2 saturation of 30%. Batch cultivations were started with an initial OD_{600} of 0.1 by inoculation with an appropriate volume of preculture. After fermentation, biomass was separated by centrifugation (15,000 xg, 20 °C, 20 min) followed by cross-flow filtration of the supernatant using 100 kDa filtration cassette (Hydrosart, Sartorius AG, Germany). Highly viscous EPS variants were diluted 1:10 with ddH_2O prior to centrifugation. Concentrated supernatant was afterwards slowly poured into two volumes of isopropanol. Precipitated EPS was collected and dried overnight in a VDL53 vacuum oven at 40 °C (Binder, Germany). Dry weight of the obtained EPS was determined gravimetrically, before milling to a fine powder in a ball mill at 30 Hz for 1 min (Mixer Mill MM400, Retsch GmbH, Germany).

2.4. Carbohydrate fingerprinting

Monomer composition of engineered EPS variants was analyzed by the 1-phenyl-3-methyl-5-pyrazolone-high throughput method (HT-PMP) as previously described (Rühmann et al., 2014). HPLC quantification of acetate was performed as previously described (Gansbiller et al., 2019).

2.5. Enzymatic assays for glucose and pyruvate quantification

Glucose and pyruvate concentrations of obtained EPS samples were determined enzymatically before and after hydrolysis to deduct residual contaminations from the fermentation medium as previously described (Rühmann et al., 2016).

2.6. Partial hydrolysis of EPS samples

Analysis of partially hydrolyzed EPS samples was performed as described for the carbohydrate fingerprinting (Rühmann et al., 2016). Differently, hydrolysis conditions were modified to 90 °C for 105 min to obtain oligosaccharides. After PMP derivatization, 25 µL 0.5 M acetic acid, 125 µL ddH_2O and 500 µL acetonitrile were added to the reaction mix to quench the derivatization process and precipitate non-hydrolyzed polymer. The precipitate was collected by centrifugation (24,000 xg, 5 min) and the obtained supernatant transferred to HPLC-vials. HPLC-UV-MS analysis was performed using a YMC-Triart Diol-HILIC column (100 × 2 mm, 1.9 µm particle size). Column temperature was set to 7 °C and a constant flow of 0.3 mL min^{-1} of 5 mM ammonium-acetate at pH 5.6 with 15% acetonitrile was applied.

2.7. Degradation of glycosyluronic acids

For glycosyluronic acid degradation 75 mg polymer were dissolved in 10 mL anhydrous ethylenediamine. Lithium scrapes were added until a deep blue colour was obtained and kept in this state for 1.5 h as previously described (Lau et al., 1987). The reaction was stopped by cooling down the solution on ice and the addition of 5 mL dH_2O. The solution was co-distilled with toluene until dryness. The polymer was reconstituted in 5 mL dH_2O and adjusted to a pH of 8 with 4 M trifluoroacetic acid (TFA). Samples were then derivatized with PMP and analyzed by HPLC-MS as described in Section 2.6.

2.8. Linkage analysis

Linkage analysis was performed according to the Hakomori procedure (Hakomori, 1964). Briefly, 1 mg of the polymer was lyophilized overnight and reconstituted in 250 µL anhydrous DMSO by stirring on a magnetic stirrer overnight. The polyanion of the polymer was produced by the addition of 500 µL 2 M dimsyl sodium under stirring overnight at RT. After freezing the sample, 250 µL methyl iodide were added and incubated for 1 h at RT. Residual methyl iodide was blown off and the methylated polymer was recovered by MeOH/$CHCl_3$ extraction. Polymer hydrolysis was performed using 2 M TFA (121 °C, 2 h), reduced with 0.25 M NaH_4B (25 °C, 1 h) and acetylated (250 µL acetic anhydride, 250 µL pyridine, 100 °C, 20 min). Partially methylated alditol acetates were extracted with ethyl acetate and applied for GC–MS analysis. Analysis of uronic acids was performed by reducing the polymer samples prior to methylation analysis. For this, 2 g L^{-1} polymer solutions were incubated in 156 mM *N*-ethyl-*N*-(3-dimethylaminopropyl)-carbodiimide hydrochloride at pH 4.75 for 1.5 h and reduced with 1 mM $NaHB_4$ for 1 h, followed by pH neutralization and dialysis of the samples. A Rxi-5 Sil MS column (300 mm, 0.25 mm ID, d_f 0.25 µm) column was used with a constant helium flow of 0.3 mL min^{-1}. For the following GC–MS analysis the initial temperature was set to 120 °C which was increased via a heat ramp of 7.5 °C min^{-1} to 180 °C, following held constant for five minutes and afterwards ramping to 230 °C. Internal references for determination of the retention times of partially methylated glycoside were obtained via Purdie-Irvine methylation of methyl-glycosides by removing samples of the incremental per-methylation every 3 h for a time span of 12 h (Purdie & Irvine, 1903; Sassaki et al., 2005).

2.9. NMR analysis

For structural elucidation by NMR spectroscopy two samples were prepared. One sample was partially degraded with acid hydrolysis and the other was mechanically sheared. The acid hydrolyzed sample was prepared by first dissolving the EPS in ultrapure water (~0.1% *w/v*), acidifying with 1.0 M hydrochloric acid to pH 1.0, and heating (in 95 °C hot water bath) for 15 min. The solution was then cooled to room temperature using a cool-water bath, neutralized with sodium hydroxide, and freeze dried. Sample was resuspended in ultrapure water (~2% *w/v*) and dialyzed (Spectra/Por Biotech cellulose ester dialysis tubing, 100–500 Da MWCO) against ultrapure water and freeze-dried. To remove remaining protein material the resuspended sample (~1% w/v) was centrifuged (3720 xg, 24 °C, 10 min). The EPS was precipitated from the supernatant by adding four volumes of isopropanol (chilled to ~4 °C) and separated from supernatant by centrifugation (3720 xg, 24 °C, 10 min). The EPS pellet was washed with isopropanol, then resuspended in ultrapure water and freeze-dried.

The sheared sample was prepared by first dissolving EPS in ultrapure water (~ 0.5% w/v) and degraded using mechanical shearing by ten cycles with pressure of 200 MPa on a Star Burst Mini (Wet Jet HJP-25001 CE, Sugino Machine Ltd., Japan). The sample was heated (in 95 °C hot water bath for 15 min) following shearing to denature remaining protein material, centrifuged (3720 xg, 24 °C, 10 min) and then the supernatant freeze dried. The EPS was precipitated from the supernatant with four volumes chilled (~4 °C) isopropanol, separated from supernatant by centrifugation (3720 xg, 24 °C, 10 min), washed with isopropanol, then resuspended in ultrapure water and freeze-dried.

The acid hydrolyzed and sheared EPS samples (20.9 mg and 14.2 mg, respectively) were redissolved in 600 µL of D_2O (99.9%;D, Sigma Aldrich) and transferred to 5 mm NMR tube (LabScape Stream, Bruker BioSpin AG, Switzerland). All homo- and heteronuclear experiments were recorded at a temperature of 50 °C on a Bruker AV-IIIHD 800 MHz spectrometer (Bruker BioSpin AG, Switzerland) equipped with a 5 mm cryogenic CP-TCI z-gradient probe. The following experiments were collected: 1D proton with water suppression (noesygppr1d), $^{1}H—^{13}C$ HSQC (heteronuclear single quantum coherence) spectrum with multiplicity editing (hsqcedetgpsisp2.3), $^{1}H—^{13}C$ H2BC (heteronuclear two bond correlation) spectrum (h2bcetgpl3pr), $^{1}H—^{13}C$ HMBC (heteronuclear multiple bond coherence) spectrum with suppression of one-bond correlations (hmbcetgpl3nd), $^{1}H—^{1}H$ DQF-COSY (double quantum filtered correlation spectroscopy) spectrum (cosydfph), and $^{1}H—^{1}H$ NOESY (nuclear Overhauser effect spectroscopy) spectrum (noesyesgpph). Acquisition parameters are provided in supplemental methods. The spectra were recorded, processed and analyzed using the TopSpin 3.5 or 4.0.1 software (Bruker BioSpin AG, Switzerland). The ^{1}H chemical shift was internally referenced to the residual water signal

(4.50 ppm for experiments acquired at 50 °C) and the ^{13}C chemical shift was referenced indirectly to DSS, using a $^{13}C/^{1}H$ frequency ratio = 0.251449530 (Wishart et al., 1995).

3. Results and discussion

3.1. CRISPR-Cas9 mediated genome engineering of the paenan cluster

To elucidate the structure of the polysaccharides produced by *P. polymyxa* DSM 365 and to understand the function of genes encoded in the paenan cluster, each glycosyltransferase as well as both Wzy polymerases were individually knocked out by CRISPR-Cas9 mediated genome engineering (Fig. 1, Fig. S1). The obtained knock-out (KO) strains were used for the production of exopolysaccharides in 500 mL scale bioreactor cultivations. Batch fermentations were performed to the complete depletion of the substrate which was achieved after approximately 26 h. While EPS production of KO strains was mostly significantly reduced compared to the wildtype, growth rate and glucose consumption was just slightly affected in most cases, resulting in a reduced overall yield (Table S4). Interestingly, none of the single gene KO variants resulted in the complete deletion of EPS production, which would have been expected result if the gene deletion affects synthesis of the backbone of a single polymer (Katzen et al., 1998). This strongly suggested the presence of at least two polymers or the possibility of further glycosyltransferases encoded outside of the paenan cluster, complementing the function of those enzymes.

3.2. Carbohydrate fingerprinting of knock-out variants

The monomer composition of KO variants was determined by PMP-derivatization and subsequent HPLC-UV-MS analysis (Fig. 2, Fig. S2). A low fucose content was observed for the wildtype polymer. However, a dimer fragment with a *m/z* of 671 was detected for all variants containing the deoxyhexose. The cumulative masses of PMP binding to the reducing end of the dimer in addition to the molecular weight of a glucuronic acid and fucose suggested a GlcA-Fuc dimer. Analysis of the MS/MS fragmentation pattern revealed the initial mass loss of −176 Da corresponding to GlcA. Based on this observation, it was concluded that fucose is positioned at the reducing end of the dimer (Fig. S3). Furthermore, MS analysis of the wildtype EPS revealed two peaks with an identical *m/z* of 581 indicating two hexose-pyruvate ketals. Additionally, a dimer with a *m/z* of 687 consisting of a hexose and GlcA residue was observed. Due to the early retention time of this particular dimer, the hexose at the reducing end was identified as mannose (Fig. S3). The high stability of uronic acid linkages towards acid hydrolysis can be valuable for sequence analysis in terms of dimer formation. However, quantification is highly affected by both, the dimer formation as well as the susceptibility towards degradation after the release of free uronic acid residues (Rodén et al., 1972). Consequently, the glucosyluronic acid content presented in the EPS variants of *P. polymyxa* is underestimated in most cases.

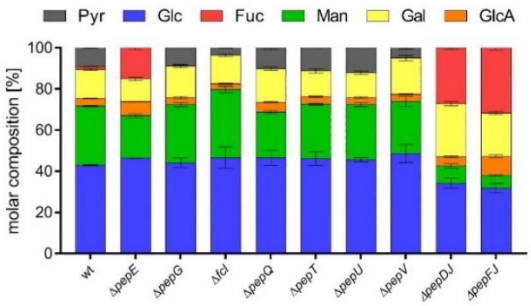


Fig. 2. Molar composition of carbohydrates and pyruvate in the EPS obtained from *P. polymyxa* DSM 365 (wt) as well as selected KO variants. Both genes encoding Wzy polymerases (*pepE* and *pepG*) as well as glycosyltransferases of clu2 (Δ*pepQ*-Δ*pepV*) were knocked-out. The gene *fcl* encodes a GDP-L-fucose synthase. Knock-outs of genes involved in biosynthesis of the fucose containing EPS (now named paenan II) resulted in a complete absence of the deoxyhexose in the monomer composition. Two combined KO variants (Δ*pepDJ* and Δ*pepFJ*) with highly enriched fucose content were designed, in which presumably the EPS paenan I cannot be produced anymore.

Interestingly, the knock-out of any single glycosyltransferase gene (*pepQ–pepV*) resulted in the complete absence of fucose. Therefore, it was assumed that these enzymes are involved in the assembly of the fucose containing polysaccharide. Notably, deletions of any other glycosyltransferases also resulted in significantly different fucose contents. Both Wzy polymerases PepE and PepG were deleted. For the Δ*pepG* strain the fucose and the corresponding dimer completely vanished. Contrary, deletion of *pepE* resulted in increased fucose content. This further supports the hypothesis of two different polymers, in which PepE is the responsible polymerase for the polymer henceforward termed paenan I, whereas PepG polymerizes a fucose containing EPS named paenan II. Notably, fucose containing polymers produced by isolates of *P. polymyxa* have previously been reported (Lee et al., 1997; Madden et al., 1986). However, the underlying polymer structure has to the best of our knowledge never been elucidated. Due to the complexity of the chemical structure of the *P. polymyxa* DSM 365 polymers, we continued to focus on the fucose containing polymer paenan II.

Two double knock-out variants, namely Δ*pepDJ* and Δ*pepFJ* with highly increased fucose content were generated (Fig. 2). Identical to

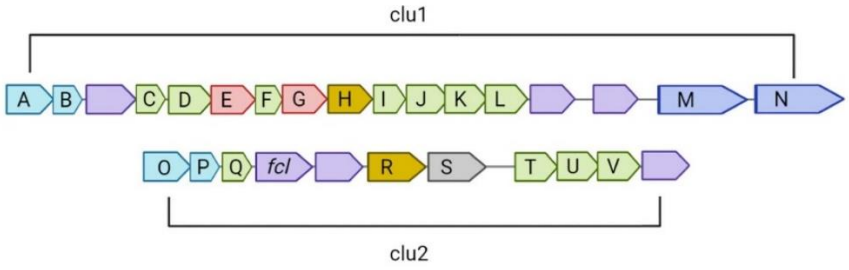


Fig. 1. CRISPR-Cas9 mediated knock-out study of the paenan cluster in *P. polymyxa* DSM 365. Putatively two transcripts encoding genes *pepA* – *pepV* (A - V) are expressed from the locus (clu1 and clu2) (Rütering et al., 2017). Each glycosyltransferase (green arrows) as well as every Wzy-polymerase (red arrows) were individually knocked-out and the EPS products of the resulting strains was characterized. Yellow: Wzx flippase; Purple: precursor synthesis; Light blue: transport/chain length control; Dark blue: Glycosyl hydrolase; Grey: uncharacterized. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ΔpepE, the enzymatically determined pyruvate concentration of the polymer was below the limit of detection. In addition, both the pyruvate ketal and the Man-GlcA dimer attributed to paenan I could not be detected in MS-analysis. Therefore, the results suggest that these variants mainly produce paenan II. Only small amounts of mannose were detected in the analyzed EPS, which might be the result of minor biomass contaminations in the obtained polymer precipitate. Based on the results of the HT-PMP analysis, a molar ratio of 1:1:1:1 Glc:Gal:Fuc:GlcA was determined for this polymer.

In addition to glycosyltransferases and the Wzy polymerases, the gene *fcl* encoding for a GDP-L-fucose synthase was deleted. Fcl has demonstrated to catalyze the second step of the nucleotide sugar synthesis by a subsequent epimerase and reduction reaction from GDP-4-dehydro-6-deoxy-D-mannose to GDP-L-fucose (Somers et al., 1998). In the Δ*fcl* strain, the nucleotide sugar was not available anymore resulting in the absence of paenan II. Therefore, we concluded that contrary to other activated sugar precursors of the paenan cluster, no functional secondary copies of the enzymes involved in the precursor synthesis of GDP-fucose exist in the genome of *P. polymyxa* DSM 365.

3.3. Partial hydrolysis confirmed the presence of two polymers

Partial hydrolysis experiments were performed to gain more information about the monomer sequence of the polymer mixture from *P. polymyxa* DSM 365. Therefore, the previously performed carbohydrate fingerprint HT-PMP method was adjusted by lowering the hydrolysis temperature to 90 °C resulting in incomplete polymer hydrolysis.

The HPLC-UV-MS analysis of the wildtype polymer revealed numerous oligosaccharides composed of hexoses with a terminally linked pyruvate ketal attributed to paenan I. Additionally, due to the different mass of fucose and the previously identified GlcA-Fuc dimer, minor oligosaccharide fractions specifically associated to paenan II were obtained. The biggest fragment detected this way displayed an *m/z* of 995, corresponding to two hexose monomers in addition to the previously determined GlcA-Fuc dimer (Fig. 3). The fragmentation pattern was analyzed for deducing the monomer sequence of the oligosaccharide. A mass shift of −162 Da to *m/z* 833 can be explained through the elimination of a hexose. Furthermore, a mixed spectrum with *m/z* 671 and *m/z* 657 attributed to the loss of GlcA (−176 Da) or the second hexose (−162 Da) was obtained. These findings indicate fucose as the branching residue. Sequence analysis for all further fragments was performed analogous to the previously described oligosaccharide.

Partial hydrolysis of EPS variants *ΔpepFJ* and *ΔpepDJ* with enriched fucose content revealed minor peaks attributed to bigger oligosaccharides with up to *m/z* 1317.5, including the GlcA-Fuc dimer. The fragmentation pattern of *m/z* 1317.5 putatively revealed two partial repeating units representing Hex-GlcA-Fuc-Hex-GlcA-Fuc-PMP. Again, the obtained mixed spectrum for the terminal Hex-GlcA residues confirmed our previous assumption of fucose as the polymer's branching point. Due to the repetitive nature of the oligosaccharide, a repeating unit as shown in Fig. 3 for paenan II is concluded. Further, smaller derivates of the 1317.5 Da fragment confirmed the proposed sequence (Table S5).

3.4. Lithium degradation of glycosyluronic acid residues

The putative sequence of the repeating unit of paenan II was further validated by specific degradation of the glycosyluronic acids using the lithium degradation method established by Lau and colleagues (Lau et al., 1987). Inhere, again the fucose enriched polysaccharide of *P. polymyxa ΔpepFJ*, was used to focus on the degradation products of paenan II. After derivatization of the obtained fragments with PMP, HPLC-MS analysis revealed mainly one peak with an *m/z* of 819 (Fig. 4). Indeed, the MS/MS fragmentation pattern confirmed the sequence Hex-Fuc-Hex-PMP. Only minor peaks with *m/z* 657 corresponding to Hex-

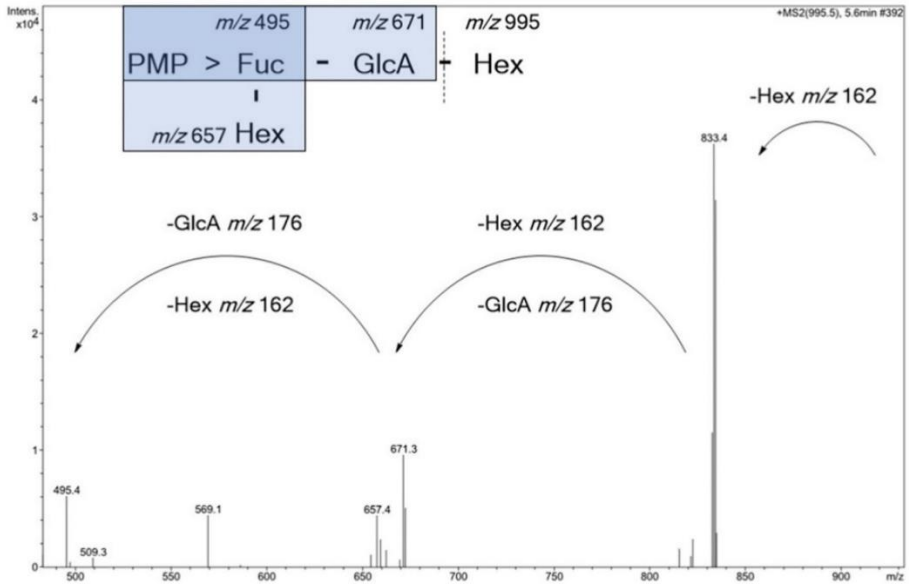


Fig. 3. Partial hydrolysis of EPS obtained from *P. polymyxa ΔpepFJ* revealed a major product with *m/z* 995. Tandem MS fragmentation indicated the initial degradation of a hexose (− 162 Da) followed by a mixed fragmentation spectrum. The difference of *m/z* 162 can be explained by a dispatch of hexose, while *m/z* 176 corresponds to GlcA. Sequential fragmentation of the remaining sugar resulted in a main peak at *m/z* 495 corresponding to Fuc-PMP. Hex: hexose; GlcA: glucuronic acid; Fuc: fucose.

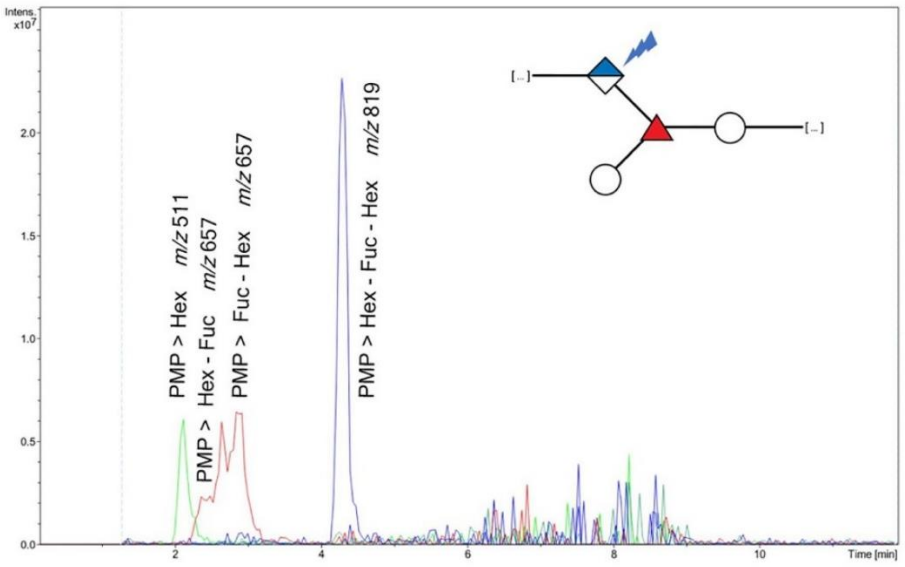


Fig. 4. Oligomers of EPS from *P. polymyxa* Δ*pepFJ* obtained after lithium degradation of uronic acids and PMP derivatization. A main fragment with *m/z* 819 was detected corresponding to Hex-Fuc-Hex-PMP. Small peaks with *m/z* 657 and *m/z* 511 show degradation products of the main peak. Corresponding MS spectra can be found in the supplemental materials (Fig. S3). red triangle: fucose; blue diamond: GlcA; empty circle: hexose. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fuc-PMP and Fuc-Hex-PMP were detected. However, the low peak area rather suggests a degradation product of the obtained trimer rather than an additional substructure. This suggests that the side chain does not consist of an irregular hexose and that the obtained double repeating units of the previous hydrolysis experiments (*m/z* 1317.5) indeed were a result of partial degradation.

While other studies have shown that the specificity of Wzx flippases allows for a certain substrate promiscuity with truncated or modified side chains, the identity of the proximal UndPP-linked sugar residue plays an essential role in the recognition of the repeating unit (Islam & Lam, 2013; Marolda, 2004). Furthermore, it was also demonstrated that Wzx flippases of *Salmonella enterica* serotypes show a high specificity for particular branching residues (Hong & Reeves, 2014; Yaoqin Hong et al., 2012). We hypothesize that a similar effect was observed for *P. polymyxa* DSM 365. The knock-out of each single glycosyltransferase of cluster 2 (PepQ – PepV) resulted in the disappearance of fucose and consequently the absence of paenan II. Therefore, we concluded that the identified branching side chain is essential for the recognition of the repeating unit by the corresponding Wzx flippase and consequently the biosynthesis.

3.5. Linkage pattern analysis of paenan II

The linkage pattern of paenan II was examined by GC–MS analysis of permethylated alditol acetates of EPS obtained from *P. polymyxa* Δ*pepFJ* (Table 1). The presence of 2-*O*-methylfucose confirmed the deoxyhexose as the branching point linked via C-3 and C-4. Additionally, 2,3,4,6-tetra-*O*-methylgalactose could be detected in equal amounts to the methylated fucose, indicating a terminal galactose residue as the branching hexose within the side chain of paenan II. A 1,4-linked glucose was identified as the second hexose of the polymer backbone. For the determination of the linkage type of the glucuronic acids, samples were subjected to chemical reduction with sodium borohydride prior to the methylation experiment. GC–MS analysis of carboxyl-reduced alditol acetates showed a new peak corresponding to a 4,6-methylated hexose. Therefore, we concluded that the glucuronic acid is linked via C-1, C-2 and C-3, corresponding to an additional branching point. Interestingly, 2,3,4-tri-*O*-methylfucose was detected in both the native, and the reduced polymer sample. However, overlapping impurities did not warrant proper integration of the peak. Consequently, also the mechanically sheared polymer sample used for NMR analysis was used for methylation. The putative removal of protein contaminations by heat precipitation allowed the detection of a second terminal sugar residue corresponding to fucose. Approximately the same peak integrals were obtained for the described sugars, confirming the equal molar ratio already proposed based on the carbohydrate fingerprints of the EPS from *P. polymyxa* DSM 365 Δ*pepFJ* and Δ*pepDJ* respectively. Considering the stable nature of GlcA-Fuc dimer, which was not quantified by the HT-PMP method due to the lack of appropriate standards, two fucosyl

Table 1
Methylation analysis of native and carboxyl reduced EPS samples of the double knock-out *P. polymyxa* DSM 365 Δ*pepFJ*.

		Area [%]			
PMAA	Linkage	RRT	A	B	C
2,3,4 Me-Fuc	1-Fuc*p*	0.81	*	*	23.2
2,3,4,6 Me-Gal	1-Gal*p*	1.05	25.5	13.7	27.3
2 Me-Fuc	1,3,4-Fuc*p*	1.07	34.7	20.3	17.3
2,3,6 Me-Glc	1,4-Glc*p*	1.25	39.8	40.6	32.1
4,6 Me Glc	1,2,3 Glc*p*	1.41	0	25.1	0.0

PMAA: partially methylated alditol acetate; 2,3,4,6 Me-Gal = 1,5-di-acetyl-2,3,4,6-tetra-*O*-methylgalactitol etc.; RRT: Relative retention time of the corresponding partially methylated alditol acetate relative to 2,3,4,6-tetra-*O*-methylglucose using a Rxi-5 Sil MS column; Percent area under curve in TIC of all integrated peaks. *: expected MS fragmentation observable, but overlapping impurities do not allow integration; A: native polymer sample; B: sample was reduced prior to Hakomori-methylation to enable detection of GlcA. C: sample was sheared prior to methylation. An extended overview of methylation products is found in the supplemental materials (Fig. S5).

residues per repeating unit seems reasonable as evidenced by the high fucose content of the carbohydrate fingerprints (Fig. 2). Interestingly, the terminal fucose residue could not be detected in the partial hydrolysis experiments. We hypothesize that due to the sensitive nature of the glycosidic linkage of the L-deoxyhexose, this terminal fucose is always cleaved off even using the applied milder hydrolysis conditions.

3.6. NMR analysis

Spectra recorded on the acid hydrolysed EPS from *ΔpepFJ* were well resolved and the core structure could be assigned by using a combination of HSQC, COSY, H2BC, HMBC, and NOESY spectra (Figs. S6–S10). In the anomeric region five signals were present (Fig. 5b). Scalar coupling constants ($^3J_{HH}$) measured in the ^{1}H spectrum (Fig. S11) and proton chemical shifts confirm that three are α-anomeric signals (δ_H 4.86, 5.21, and 5.65) and two are β-anomeric signals (δ_H 4.64 and 4.76). Two methyl signals (Fig. 5a: $\delta_H \sim 1.30$, $\delta_C \sim 18.0$) are evidence of 6-deoxy residues and two sets of methylene signals (D6 and E6 in Fig. 5) are consistent with unsubstituted hydroxymethylenes (position-6) of hexose residues. These signals are consistent with the detection of both the 6-deoxy sugar fucose, and the two hexoses glucose and galactose detected in the monosaccharide analyses.

Proton and carbon chemical shifts for each residue were assigned (Table 2) using HSQC to establish chemical shifts and COSY, H2BC, and HMBC correlations to establish connectivity between C—H spin pairs. The HMBC spectrum also confirmed the presence of a uronic acid

Table 2
^{1}H and ^{13}C chemical shifts[a] for acid hydrolysed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Correlations are annotated with bold letter that refers to residue and number indicate the carbon number for position within residue.

	Residue	Chemical shift (ppm)					
		C-1	C-2	C-3	C-4	C-5	C-6
		H-1	H-2	H-3	H-4	H-5	H-6
A	α-L-Fuc*p*-(1-	100.7	70.7	72.0	74.8	69.2	18.3
		5.65	3.78	3.87	3.72	4.57	1.17
B	2,3)-β-D-GlcAp-(1-	104.8	77.3	88.6	73.5	80.9	177.0
		4.64	3.77	3.96	3.74	3.67	–
C	3,4)-α-L-Fuc*p*-(1-	102.3	71.0	76.0	82.5	70.2	18.2
		4.86	3.91	4.01	4.12	4.37	1.30
D	4)-β-D-Glc*p*-(1-	104.5	76.6	77.2	79.4	78.0	62.7
		4.76	3.38	3.61	3.56	3.57	3.78, 3.94
E	α-D-Gal*p*-(1-	102.8	71.5	72.0	72.4	74.1	64.5
		5.21	3.73	3.88	3.95	4.11	3.68

[a] Chemical shift in ppm, EPS dissolved in D_2O, spectra recorded at 50 °C on 800 MHz spectrometer, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS based on ^{1}H/^{13}C frequency ratio = 0.251449530.

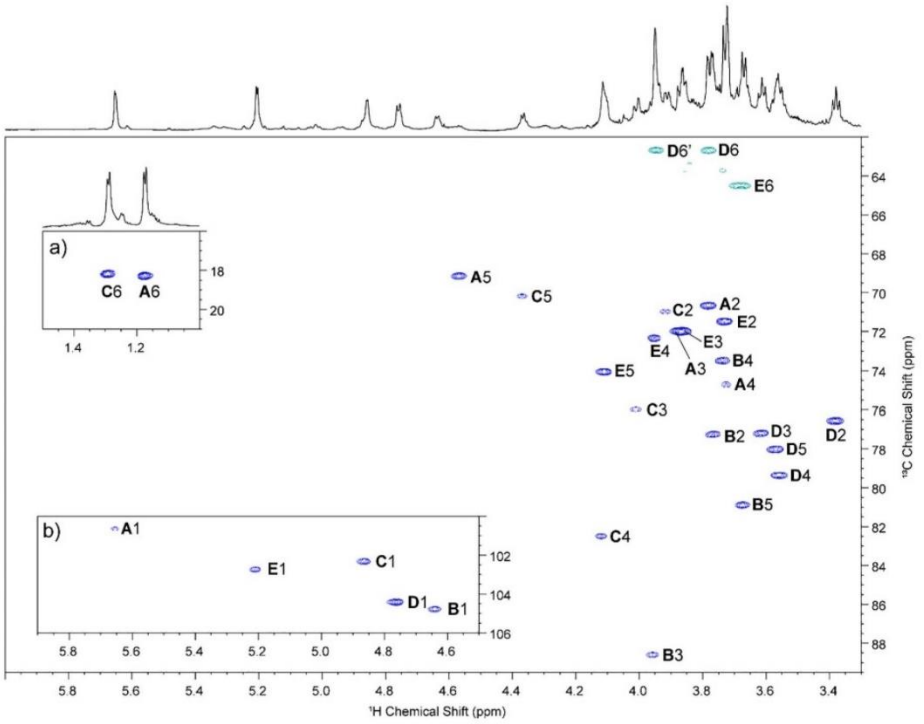


Fig. 5. ^{1}H—^{13}C HSQC spectrum with multiplicity editing of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Projection across top shows ^{1}H spectrum, fucose methyl groups are shown in inset a) and the anomeric signals are show in inset b). Correlations are annotated with bold letter that refers to residue and number that indicates the carbon number for position within residue. Blue correlations indicate C—H single bond correction for methyl (CH_3) and methine (CH) groups while cyan correlations indicate C—H single bond correction for methylene (CH_2) groups. Sample dissolved in D_2O (99.9%), spectrum recorded at 50 °C and 800 MHz. A = α-L-Fuc*p*-(1-, B = 2,3)-β-D-GlcA*p*-(1-, C = 3,4)-α-L-Fuc*p*-(1-, D = 4)-β-D-Glc*p*-(1-, E = α-D-Gal*p*-(1-. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(residue B) revealing a carboxylic acid at position-6 (δ_C 177.0). Using the results from carbohydrate fingerprinting and linkage analyses in combination with literature (Lundborg & Widmalm, 2011) the five residues were determined to be two α-L-Fuc*p* (residues A and C), β-D-GlcA*p* (residue B), β-D-Glc*p* (residue D), and α-D-Gal*p* (residue E).

Inter-residue HMBC correlations (Fig. S12) were used to establish the position of glycosidic linkages between residues and were further verified by NOESY correlations (Fig. S13). These positions are confirmed by their carbon chemical shifts, which are typically shifted to higher chemical shifts when involved in glycosidic linkages (Agrawal, 1992). Using HMBC and NOESY correlations the main backbone was established as a three-residue repeating unit comprising β-D-GlcA*p* (residue **B**) 1,4-linked to α-L-Fuc*p* (residue **C**) 1,4-linked to β-D-Glc*p* (residue **D**) and that is 1,3-linked to β-D-GlcAp (residue **B**). The remaining α-L-Fuc*p* (residue **A**) and α-D-Gal*p* (residue **E**) were both determined to be terminal residues. The α-L-Fuc*p* residue (**A**) is 1,2-linked to β-D-GlcA*p* (residue **B**) and the α-D-Gal*p* residue (**E**) is 1,3-linked to α-L-Fuc*p* (residue **C**).

HMBC correlations from H-1 of both fucose residues were weak and difficult to observe, but glycosidic linkages were confidently assigned using correlations from H-2 of β-D-GlcA*p* (residue **B**) and H-4 of Glc (residue **D**) into the anomeric carbon of their respective fucose residues (residue **A** and **C**, respectively). The positions were further supported by H—H NOESY correlations between protons involved in the glycosidic linkages. These NMR data provides further confirmation that fucose (residue **C**) is a branching point in the main chain with glycosidic linkages present at both C-3 and C-4.

Acid hydrolysis results in deacetylation, therefore, to investigate the nature of acetylation in the native polymer a second sample was prepared by mechanical shearing. Comparison of proton spectra (Fig. 6) and HSQC spectra (Fig. S14) between the sheared and acid hydrolysed EPS samples confirms the same five-residue repeating unit is present in the sheared sample as well as several new signals. Anomeric signals for the five characterized residues are present in the sheared sample, with the addition of two new anomeric signals (δ_H 5.73 and 5.72). Evidence of *O*-acetylation can be observed as two methyl singlets (approximately δ_H 2.15) and two new signals (approximately δ_H 5.10) that are the sites of *O*-acetylation. There are also several new methyl singlets (approximately δ_H 1.20–1.30) congruent with acetylation of fucose residues.

Using HSQC, H2BC, COSY and HMBC spectra (Figs. S15–S18) two partial residues (**A'** and **A"**) were assigned (Table 3). From these partial fragments it was determined that the residues are mono-acetylated and are located at either position-3 or − 4. It is presumed the acetylation occurs on the terminal fucose residue indicated by a decrease in the anomeric signal associated with residue A in the sheared sample compared to the acid hydrolysed sample (Fig. S19). This is also supported by the appearance of new C-5 methine and C-6 methyl signals (Fig. S14).

Integrals of relevant proton spectra were used to estimate the degree of acetylation (Fig. S19). The degree of acetylation was estimated at approximately 70% by comparing the integral of the anomeric signals of acetylated residues (**A'** and **A"** combined) with non-acetylated residue (**A**). As the anomeric signals for the acetylated residues (**A'** and **A"**) are overlapping the methyl signals of the acetyl groups was used to determine the proportion for each position. This showed that 4-O-Ac were more abundant accounting for approximately 2/3 of acetylated residues. Acetylation of paenan II was further cross-validated by HPLC analysis. Indeed, for the polymer obtained from *P. polymyxa* DSM 365 Δ*pepFJ* an acetate content of 7.2% was detected, indicating a complete acetylation of the terminal fucose residue.

From the combined results of carbohydrate fingerprinting, partial hydrolysis and methylation analysis as well as NMR the following structure as depicted in Fig. 7 can be derived for the repeating unit of paenan II.

4. Conclusion

P. polymyxa is a producer of a highly complex exopolysaccharide mixture. However, structural elucidation of the biopolymer based on conventional methods remained unfeasible so far due to the complexity of the polymer composition. Using systematic CRISPR-Cas9 mediated gene knock-outs, we achieved a deletion of the main polymer and were able to isolate a minor fucose containing polysaccharide named paenan II for the first time. By a combinatorial approach, the monosaccharide composition, sequence, and linkage pattern were resolved for paenan II. Furthermore, we observed that the knock-out of the glycosyltransferases PepQ, PepT, PepU and PepV resulted in the absence of paenan II. Therefore, these enzymes were attributed to the assembly of the polysaccharide indicating that truncated polysaccharides without a branching side chain cannot be produced by the organism (Fig. S20).

These results have paved a strategy to elucidate the structure of main Paenan I using combinatorial knock-out combined with classical carbohydrate analyses. This approach can establish a structure-function relationship between the chemical structure and the underlying biosynthesis machinery of the EPS.

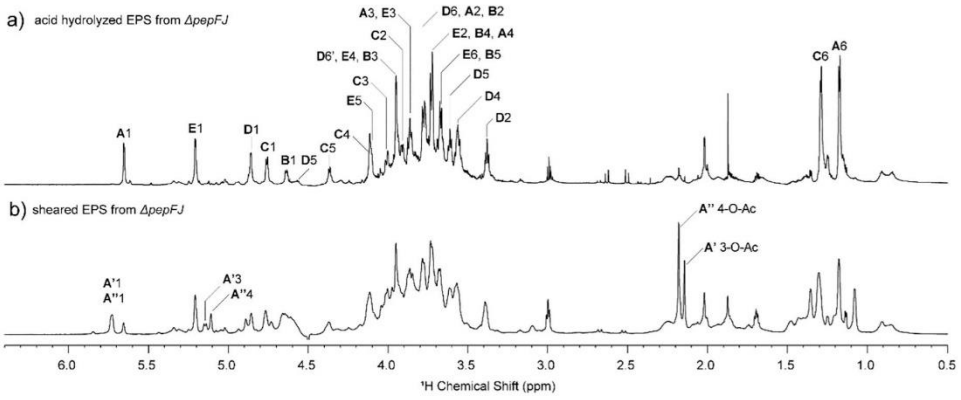


Fig. 6. ^{1}H spectra of a) acid degraded and b) mechanically sheared EPS produced by *P. polymyxa* DSM 365 Δ*pepFJ*. Spectra are annotated with residue assignments. Sample dissolved in D_2O (99.9%), spectrum acquired at 50 °C and 800 MHz. **A** = α-L-Fuc*p*-(1-, **A'** = α-L-Fuc*p*-3-O-Ac-(1-, **A"** = α-L-Fuc*p*-4-O-Ac-(1-, **B** = 2,3)-β-D-GlcA*p*-(1-, **C** = 3,4)-α-L-Fuc*p*-(1-, **D** = 4)-β-D-Glc*p*-(1-, **E** = α-D-Gal*p*-(1-.

Table 3
^{1}H and ^{13}C chemical shifts[a] for two partial residues in sheared EPS produced by *P. polymyxa* DSM 365 Δ*pepFJ*.

	Residue	Chemical shift (ppm)							
		C-1	C-2	C-3	C-4	C-5	C-6	CO	C-Me
		H-1	H-2	H-3	H-4	H-5	H-6		H-Me
A'	α-L-Fucp-3-O-Ac-(1-	100.6	68.5	75.5	n.a.	n.a.	n.a.	176.2	23.1
		5.73	4.00	5.14	n.a.	n.a.	n.a.		2.15
A''	α-L-Fucp-4-O-Ac-(1-	100.7	70.9	70.5	77.3	n.a.	n.a.	176.8	22.9
		5.72	3.83	4.04	5.11	n.a.	n.a.		2.18

[a] Chemical shift in ppm, EPS dissolved in D_2O (99.9%), spectra recorded at 50 °C on 800 MHz spectrometer, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS based on $^{1}H/^{13}C$ frequency ratio = 0.251449530. CO = carbon chemical shift of carbonyl of *O*-acetyl-group, C-Me and H-Me = carbon and proton chemical shifts (respectively) for the methyl signal of *O*-acetyl-group, n.a. = not assigned.

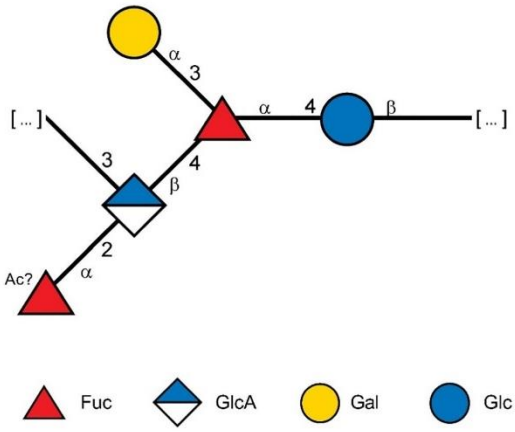


Fig. 7. Elucidated structure of the penta-saccharide repeating unit of paenan II derived from a combinatorial analysis of monomer composition, sequence and linkage pattern. Acetylation of the terminal fucose residue seems to be promiscuous at position C-3/4.

CRediT authorship contribution statement

CS: Conceptualization, Investigation, Data analysis, Writing – Original Draft; LJK: Methodology, Investigation, Writing – Original Draft; BR: Data analysis, Writing – Review & Editing; FLA: Data analysis, Funding acquisition, Writing – Review & Editing; JS: Conceptualization, Writing – Review & Editing, Funding acquisition, Supervision VS: Conceptualization, Writing – Review & Editing, Funding acquisition, Supervision.

Declaration of competing interest

CS, BR, JS and VS declare financial interest in the EPS of *P. polymyxa* DSM 365.

Acknowledgements

The authors would like to acknowledge the technical support from Petra Lommes and Julia Wünsche for methylation analysis experiments. This work was supported by the German Federal Ministry of Education and Research (BMBF) in frame of the project MaPolKo (number 03VP02560). LJK and FLA acknowledge the Research Council of Norway who has contributed though the grants 294946 (The Norwegian Seaweed Biorefinery Platform) and 226244 (Norwegian NMR platform-NNP).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.carbpol.2021.118951.

References

Agrawal, P. (1992). NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. *Phytochemistry, 31*(10), 3307–3330.
Becker, A. (1993). Analysis of the Rhizobium meliloti genes exo U, exo V, exo W, exo T, and exo I involved in exopolysaccharide biosynthesis and nodule invasion: Exo U and exo W probably encode glucosyltransferases. *6*(6), 735.
Becker, A., Katzen, F., Pühler, A., & Ielpi, L. (1998). Xanthan gum biosynthesis and application: A biochemical/genetic perspective. *Applied Microbiology and Biotechnology, 50*(2), 145–152.
Bentley, S. D., Corton, C., Brown, S. E., Barron, A., Clark, L., Doggett, J., Harris, B., et al. (2008). Genome of the actinomycete plant pathogen Clavibacter michiganensis subsp. sepedonicus suggests recent niche adaptation. *Journal of Bacteriology, 190*(6), 2150–2160.
Cobb, R. E., Wang, Y., & Zhao, H. (2015). High-efficiency multiplex genome editing of streptomyces species using an engineered CRISPR/Cas system. *ACS Synthetic Biology, 4*(6), 723–728.
Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J., & Koffas, M. A. G. (2014). Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews, 38*(4), 660–697.
Cuthbertson, L., Mainprize, I. L., Naismith, J. H., & Whitfield, C. (2009). Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiology and Molecular Biology Reviews, 73*(1), 155–177.
De Mas, C., Jansen, N. B., & Tsao, G. T. (1988). Production of optically active 2,3-butanediol by bacillus polymyxa. *Biotechnology and Bioengineering, 31*(4), 366–377.
Gansbiller, M., Schmid, J., & Sieber, V. (2019). In-depth rheological characterization of genetically modified xanthan-variants. *Carbohydrate Polymers, 213*, 236–246.
Hakomori, S. (1964). A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *Journal of Biochemistry, 55*, 205–208.
Hassler, R. A., & Doherty, D. H. (1990). Genetic engineering of polysaccharide structure: Production of variants of xanthan gum in Xanthomonas campestris. *Biotechnology Progress, 6*(3), 182–187.
Hong, Y., Cunneen, M. M., & Reeves, P. R. (2012). The Wzx translocases for Salmonella enterica O-antigen processing have unexpected serotype specificity: Wzx specificity in Salmonella enterica. *Molecular Microbiology, 84*(4), 620–630.
Hong, Y., & Reeves, P. R. (2014). Diversity of O-antigen repeat unit structures can account for the substantial sequence variation of Wzx translocases. *Journal of Bacteriology, 196*(9), 1713–1722.
Ielpi, L., Couso, R. O., & Dankert, M. A. (1993). Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in Xanthomonas campestris. *Journal of Bacteriology, 175*(9), 2490–2500.
Islam, S. T., & Lam, J. S. (2013). Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria: Role of Wzx in surface polysaccharide synthesis. *Environmental Microbiology, 15*(4), 1001–1015.
Islam, S. T., & Lam, J. S. (2014). Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Canadian Journal of Microbiology, 60*(11), 697–716.
Jang, H. Y., Zhang, K., Chon, B. H., & Choi, H. J. (2015). Enhanced oil recovery performance and viscosity characteristics of polysaccharide xanthan gum solution. *Journal of Industrial and Engineering Chemistry, 21*, 741–745.
Jansson, P., Kenne, L., & Lindberg, B. (1975). Structure of the extracellular polysaccharide from Xanthomonas campestris. *Carbohydrate Research, 45*(1), 275–282.
Jeanes, A., Pittsley, J. E., & Senti, F. R. (1961). Polysaccharide B-1459: A new hydrocolloid polyelectrolyte produced from glucose by bacterial fermentation. *Journal of Applied Polymer Science, 5*(17), 519–526.
Jeong, H., Choi, S.-K., Ryu, C.-M., & Park, S.-H. (2019). Chronicle of a soil bacterium: Paenibacillus polymyxa E681 as a tiny guardian of plant and human health. *Frontiers in Microbiology, 10*, 467.

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology, 31*(3), 233–239.
Katzen, F., Ferreiro, D. U., Oddo, C. G., Ielmini, M. V., Becker, A., Pühler, A., & Ielpi, L. (1998). Xanthomonas campestris pv. campestris gum mutants: Effects on xanthan biosynthesis and plant virulence. *Journal of Bacteriology, 180*(7), 1607–1617.
Lau, J. M., McNeil, M., Darvill, A. G., & Albersheim, P. (1987). Selective degradation of the glycosyluroic acid residues of complex carbohydrates by lithium dissolved in ethylenediamine. *Carbohydrate Research, 168*(2), 219–243.
Lee, I. Y., Seo, W. T., Kim, G. J., Kim, M. K., Ahn, S. G., Kwon, G. S., & Park, Y. H. (1997). Optimization of fermentation conditions for production of exopolysaccharide by Bacillus polymyxa. *Bioprocess Engineering, 16*(2), 71.
Lehman, A. P., & Long, S. R. (2013). Exopolysaccharides from Sinorhizobium meliloti can protect against H2O2 dependent damage. *Journal of Bacteriology, 195*(23), 5362–5369.
Liang, T.-W., Wu, C.-C., Cheng, W.-T., Chen, Y.-C., Wang, C.-I., Wang, I.-L., & Wang, S.-L. (2014). Exopolysaccharides and antimicrobial biosurfactants produced by Paenibacillus macerans TKU029. *Applied Biochemistry and Biotechnology, 172*(2), 933–950.
Liu, D., Cole, R. A., & Reeves, P. R. (1996). An O-antigen processing function for wzx (RfbX): A promising candidate for O-unit flippase. *Journal of Bacteriology, 178*(7), 2102–2107.
Lundborg, M., & Widmalm, G. (2011). Structural analysis of glycans by NMR chemical shift prediction. *Analytical Chemistry, 83*(5), 1514–1517.
Madden, J. K., Dea, I. C. M., & Steer, D. C. (1986). Structural and rheological properties of the extracellular polysaccharides from bacillus polymyxa. *Carbohydrate Polymers, 6*(1), 51–73.
Marolda, C. L. (2004). Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology, 150*(12), 4095–4105.
Mohd Din, A. R. J., Rosli, M. A., Mohamad Azam, Z., Othman, N. Z., & Sarmidi, M. R. (2019). Paenibacillus polymyxa role involved in phosphate solubilization and growth promotion of Zea mays under abiotic stress condition. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, 90*(1), 63–71.
Moradali, M. F., & Rehm, B. H. A. (2020). Bacterial biopolymers: From pathogenesis to advanced materials. *Nature Reviews Microbiology, 18*(4), 195–210. https://doi.org/10.1038/s41579-019-0313-3
Moreira, L. M., Becker, J. D., Pühler, A., & Becker, A. (2000). The Sinorhizobium meliloti ExpE1 protein secreted by a type I secretion system involving ExpD1 and ExpD2 is required for biosynthesis or secretion of the exopolysaccharide galactoglucan. *Microbiology, 146*(9), 2237–2248.
Morona, R., & Nath, P. (2015). Mutational analysis of the major periplasmic loops of Shigella flexneri Wzy: Identification of the residues affecting O antigen modal chain length control, and Wzz-dependent polymerization activity. *Microbiology, 161*(4), 774–785.
Nath, P., & Morona, R. (2015). Detection of Wzy/Wzz interaction in Shigella flexneri. *Microbiology, 161*(9), 1797–1805.
Paul, F., Morin, A., & Monsan, P. (1986). Microbial polysaccharides with actual potential industrial applications. *Biotechnology Advances, 4*(2), 245–259.
Pérez-Mendoza, D., Rodríguez-Carvajal, M.Á., Romero-Jiménez, L., Farias, G.d. A., Lloret, J., Gallegos, M. T., & Sanjuán, J. (2015). Novel mixed-linkage β-glucan activated by c di GMP in Sinorhizobium meliloti. *Proceedings of the National Academy of Sciences, 112*(7), E757–E765.
Plank, J. (2005). Applications of biopolymers in construction engineering. In A. Steinbüchel (Ed.), *Biopolymers online*. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.
Pollock, T. J., van Workum, W. A., Thorne, L., Mikolajczak, M. J., Yamazaki, M., Kijne, J. W., & Armentrout, R. W. (1998). Assignment of biochemical functions to glycosyl transferase genes which are essential for biosynthesis of exopolysaccharides in Sphingomonas strain S88 and Rhizobium leguminosarum. *Journal of Bacteriology, 180*(3), 586–593.
Purdie, T., & Irvine, J. C. (1903). C.—The alkylation of sugars. *Journal of the Chemical Society, Transactions, 83*, 1021–1037.
Raza, W., Makeen, K., Wang, Y., Xu, Y., & Qirong, S. (2011). Optimization, purification, characterization and antioxidant activity of an extracellular polysaccharide produced by Paenibacillus polymyxa SQR-21. *Bioresource Technology, 102*(10), 6095–6103.
Rodén, L., Baker, J. R., Anthony Cifonelli, J., & Mathews, M. B. (1972). Isolation and characterization of connective tissue polysaccharides. *Methods in Enzymology, 28*, 73–140.
Rühmann, B., Schmid, J., & Sieber, V. (2014). Fast carbohydrate analysis via liquid chromatography coupled with ultra violet and electrospray ionization ion trap detection in 96-well format. *Journal of Chromatography A, 1350*, 44–50.
Rühmann, B., Schmid, J., & Sieber, V. (2016). Automated modular high throughput exopolysaccharide screening platform coupled with highly sensitive carbohydrate fingerprint analysis. *Journal of Visualized Experiments, 110*, 53249.
Rütering, M., Cress, B. F., Schilling, M., Rühmann, B., Koffas, M. A. G., Sieber, V., & Schmid, J. (2017). Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in Paenibacillus polymyxa. *Synthetic Biology, 2*(1).
Rütering, M., Schmid, J., Gansbiller, M., Braun, A., Kleinen, J., Schilling, M., & Sieber, V. (2018). Rheological characterization of the exopolysaccharide Paenan in surfactant systems. *Carbohydrate Polymers, 181*, 719–726.
Rütering, M., Schmid, J., Rühmann, B., Schilling, M., & Sieber, V. (2016). Controlled production of polysaccharides–exploiting nutrient supply for Levan and heteropolysaccharide formation in Paenibacillus sp. *Carbohydrate Polymers, 148*, 326–334.
Sassaki, G. L., Gorin, P. A. J., Souza, L. M., Czelusniak, P. A., & Iacomini, M. (2005). Rapid synthesis of partially O methylated alditol acetate standards for GC–MS: Some relative activities of hydroxyl groups of methyl glycopyranosides on Purdie methylation. *Carbohydrate Research, 340*(4), 731–739.
Schilling, C., Badri, A., Sieber, V., & Schmid, J. (2020). Metabolic engineering for production of functional polysaccharides. *Current Opinion in Biotechnology, 66*, 44–51.
Schilling, C., Ciccone, R., Sieber, V., & Schmid, J. (2020). Engineering of the 2,3-butanediol pathway of Paenibacillus polymyxa DSM 365. *Metabolic Engineering, 61*, 381–388.
Schmid, J. (2018). Recent insights in microbial exopolysaccharide biosynthesis and engineering strategies. *Current Opinion in Biotechnology, 53*, 130–136.
Somers, W. S., Stahl, M. L., & Sullivan, F. X. (1998). GDP-fucose synthetase from Escherichia coli: Structure of a unique member of the short chain dehydrogenase/reductase family that catalyzes two distinct reactions at the same active site. *Structure, 6*(12), 1601–1612.
Sutherland, I. W. (1972). Bacterial exopolysaccharides. *Advances in Microbial Physiology, 8*, 143–213.
Whitfield, C. (1995). Biosynthesis of lipopolysaccharide O antigens. *Trends in Microbiology, 3*(5), 178–185.
Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. S., & Sykes, B. D. (1995). 13C and 15N chemical shift referencing in biomolecular NMR. *Journal of Biomolecular NMR, 6*, 135–140.
Woodward, R., Yi, W., Li, L., Zhao, G., Eguchi, H., Sridhar, P. R., Guo, H., et al. (2010). Guoln vitro bacterial polysaccharide biosynthesis: Defining the functions of Wzy and Wzz. *Nature Chemical Biology, 6*(6), 418–423.
Xie, J.-B., Du, Z., Bai, L., Tian, C., Zhang, Y., Xie, J.-Y., & Wang, T. (2014). Comparative genomic analysis of N2-fixing and non-N2-fixing Paenibacillus spp.: Organization, evolution and expression of the nitrogen fixation genes. In P. M. Richardson (Ed.), *10 (3). PLoS Genetics*.

3.2.2. CRISPR-Cas9 mediated structural elucidation of the heteroexopolysaccharides of *P. polymyxa* DSM 365

P. polymyxa is an avid producer of various carbohydrate polymers [71,77] and has been shown to secret a highly complex heteroexopolysaccharide termed paenan under specific cultivation conditions [69]. Using combinatorial knock-outs of glycosyltransferases, it was possible to enrich a fucose containing polymer termed ‘paenan II’ and characterized it in detail [176]. However, the structure of the main polymer remained to be elucidated.

Following up on our recent approach on the elucidation of polysaccharides produced by *P. polymyxa* DSM 365, additional knock-out combinations of glycosyltransferases encoded in the paenan cluster were performed. A total of 32 combinatorial knock-outs targeting the biosynthesis of the heteroexopolysaccharides were generated by CRISPR-Cas9 mediated genome engineering. These knock-out strains were used for the fermentative production of EPS and all polymer variants were characterized as previously described [176]. By combining results of analytical methods including monomer, sequence and linkage analysis, it was discovered that the heteroexopolysaccharide produced by *P. polymyxa* DSM 365 in fact consists of three distinct polymers, which were named paenan I, paenan II and paenan III respectively. Using the knock-out based approach, each GT of the paenan cluster could be attributed to its corresponding polymer. Furthermore, a secondary locus was detected in the genome encoding a pyruvyltransferase, a Wzx flippase as well as two additional glycosyltransferases, which are necessary for the biosynthesis of paenan I. By that approach two structures for the repeating unit of paenan I and III were postulated, which require further verification by NMR.

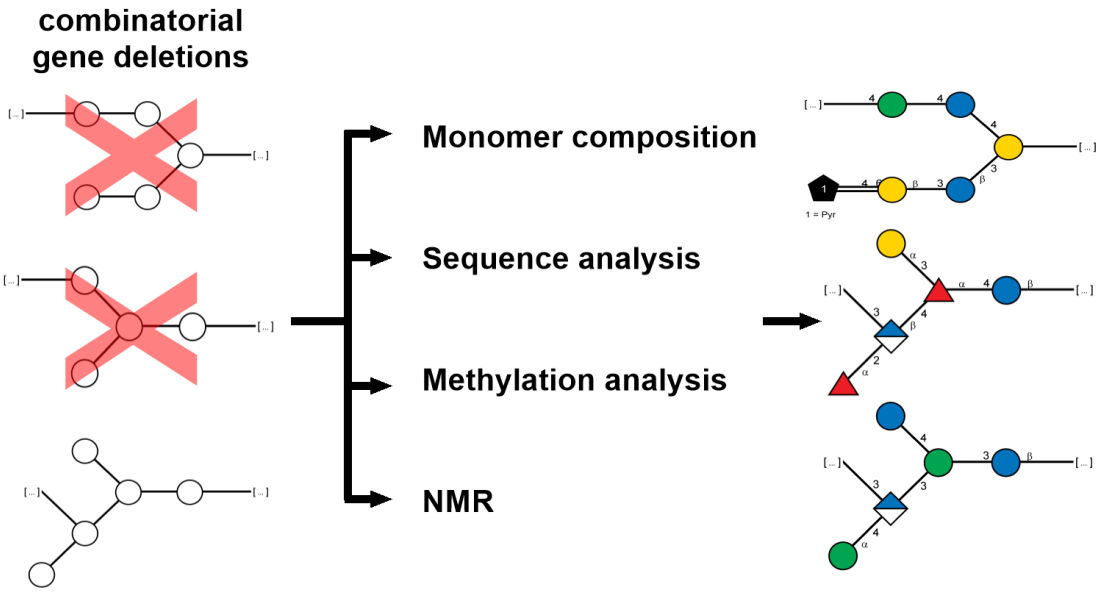
Design and planning of this study was performed by Christoph Schilling in collaboration with Jochen Schmid, Broder Rühmann and Volker Sieber. Strain construction, fermentation and carbohydrate analysis were performed by the author. Broder Rühmann gave scientific advice and support for analytical measurements. NMR analysis and evaluation were performed by Leesa J. Klau and Finn Aachmann. All authors contributed to content and language of the manuscripts and provided scientific or technical advice.

CRISPR-Cas9 driven structural elucidation of the heteroexopolysaccharides from *Paenibacillus polymyxa* DSM 365

Christoph Schilling, Leesa J. Klau, Broder Rühmann, Finn Aachmann, Jochen Schmid & Volker Sieber

Prepared manuscript

(2022)



Abstract

Paenibacillus polymyxa is a Gram-positive soil bacterium, which has been known for the production of a variety of exopolysaccharides. However, due to the complexity of the biopolymer, structural elucidation remained inconclusive so far. Combinatorial knock-outs of glycosyltransferases were generated in order to separate distinct polysaccharides produced by *P. polymyxa*. Using a complementary analytical approach consisting of carbohydrate fingerprints, sequence analysis, methylation analysis as well as NMR, the structure of the repeating units of two additional exopolysaccharides termed paenan I and paenan III were elucidated. Results for paenan I demonstrated a trisaccharide backbone consisting of 1→4-D-Glc, 1→4-D-Man and a 1,3,4-branching D-Glc residue with a side chain comprising of a terminal β-D-Gal$^{3,4\text{-Pyr}}$ and 1→3-D-Glc. For paenan III, results indicated a backbone consisting of 1→3-β-D-Glc, 1,3,4-linked D-Man and 1,3,4-linked GlcA. NMR analysis indicated monomeric D-Glc and α-D-Man side chains for the branching Man and GlcA residues respectively.

Introduction

Paenibacillus polymyxa (formerly classified as Bacillus polymyxa) is a Gram-positive, endospore forming, rod-shaped soil bacterium (Ash, Priest, & Collins, 1994; Shida, Takagi, Kadowaki, Nakamura, & Komagata, 1997), which is used as a biofertilizer in agriculture due to its abilities to fixate nitrogen and solubilize phosphate. In addition, it produces an arsenal of different antibiotic compounds that act as pest control against various insects and fungi (Jeong, Choi, Ryu, & Park, 2019; Wang, Shi, Li, Ibraham, & Sun, 2012; J.-B. Xie et al., 2014). Furthermore, *P. polymyxa* demonstrated promising results for the biotechnological production of *R,R*-2,3-butanediol in high enantiomeric purity (De Mas, Jansen, & Tsao, 1988) as well as a production organism for the synthesis of exopolysaccharides (EPS) (Grady, MacDonald, Liu, Richman, & Yuan, 2016). Naturally, these biopolymers are synthesized as barriers against abiotic stress factors (e.g. desiccation, extreme pH, osmotic stress etc.), but they can also play a crucial role in the host-pathogen interaction of certain pathogenic organisms (Cress et al., 2014; Moradali & Rehm, 2020). EPSs are also applied in food, feed, cosmetics, as well as technical applications such as packaging or oil recovery (Diab, Biliaderis, Gerasopoulos, & Sfakiotakis, 2001; Freitas, Alves, & Reis, 2011; Schilling, Badri, Sieber, Koffas, & Schmid, 2020). Most commonly, these biopolymers are applied as rheological modifiers, but depending on their physicochemical properties they are also applied as high value biomaterials in pharmaceutical and cosmetic products due to anti-inflammatory, low immunogenic or hygroscopic effects (Badri, Williams, Linhardt, & Koffas, 2018; Burdick & Prestwich, 2011; Miao, Wang, Zeng, Liu, & Chen, 2018).

Different strains of *Paenibacillus* spp. have been reported to produce variable types of

EPS with distinct physicochemical and rheological properties that makes them interesting for a variety of different applications (Aguilera et al., 2001; Çolak, Olgun, Atar, & Yazıcıoğlu, 2013; Raza, Makeen, Wang, Xu, & Qirong, 2011). *P. polymyxa* demonstrated to synthesize a levan-type polyfructan if sucrose is used as a carbon source (Liu et al., 2010a). Contrarily, using glucose and a high C/N ratio results in the production of a heteroexopolysaccharide (Rütering, Schmid, Rühmann, Schilling, & Sieber, 2016). Despite multiple attempts over the last 50 years, structural elucidation of the polysaccharide of *P. polymyxa* has ultimately failed due to the high complexity of the biopolymer (Liang & Wang, 2015; Madden, Dea, & Steer, 1986; Ninomiya & Kizaki, 1969). Presumably, there is a large variety of structurally related polysaccharides produced by different *P. polymyxa* isolates. Previous studies typically reported monomer compositions consisting of glucose, galactose, mannose, fucose, glucuronic acids and pyruvate, while rarely also xylose has been described (Jung et al., 2007; Lee et al., 1997; Li et al., 2013; Liu et al., 2010b; Raza et al., 2011). However, a clear comparison of the obtained results is difficult due to the use of different strains, cultivation conditions and media components. For *P. polymyxa* DSM 365, the identification and preliminary annotation of the gene cluster responsible for the biosynthesis of its heteroexopolysaccharide was performed on the published genome (Rütering et al., 2017; N.-Z. Xie et al., 2015). The complexity of the produced EPS is mirrored in the underlying gene cluster, comprising 28 coding sequences encoding eleven glycosyltransferases, two polymerases and flippases as well as further genes attributed to export, chain-length control and nucleotide precursor synthesis (Rütering et al., 2017) (Figure 1). Typically, Wzx flippases and glycosyltransferases are highly specific towards their natural substrate and accept only minor differences in the composition of the repeating units (Galili, 1993; Islam & Lam, 2013; Riva, 2001). Consequently, bioinformatic analysis already indicated the ability of *P. polymyxa* DSM 365 to produce multiple heteroexopolysaccharides via the Wzx/Wzy pathway. While it is quite common among bacteria to produce variable types of EPS depending on environmental conditions, it is interesting that the genes corresponding to these putative heteroexopolysaccharides are encoded in a single cluster (Bentley et al., 2008; Rütering et al., 2016).

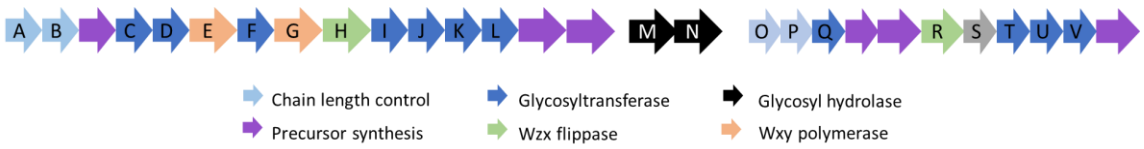


Figure 1: 22 *pep* genes (*pepA-V*) are encoded in the paenan cluster of *P. polymyxa* 365, which is responsible for the biosynthesis of heteroexopolysaccharides following the Wzx/Wzy pathway. Furthermore, additional copies of genes involved in nucleotide sugar precursor synthesis (Purple arrows) are encoded in the cluster. Adapted from (Rütering et al., 2017)

Using a previously developed tool for CRISPR-Cas9 mediated knock-outs in *P. polymyxa* our group recently confirmed the presence of multiple

heteroexopolysaccharides. Additionally, we were able to elucidate the structure of a minor fucose containing EPS termed paenan II (Rütering et al., 2017; Schilling et al., 2022). Furthermore, specific genes (*pepQ-V*) could be allocated to the synthesis of paenan II. Continuing these efforts, this study aims to generate combinatorial knock-outs of the paenan cluster to enable the structural elucidation of the main polymer paenan I and establish a structure – function relationship to the underlying gene cluster.

Materials and Methods

Strains and media

P. polymyxa DSM 365 was acquired from the German Collection of Microorganisms and Cell Culture (DSMZ), Braunschweig, Germany. Escherichia coli NEB Turbo cells (New England Biolabs, USA) were used for any plasmid construction presented in this study. E. coli S17-1 (DSMZ strain DSM 9079) was used for transformation of *P. polymyxa* DSM 365 via conjugation. All medium components were obtained from Carl Roth GmbH (Germany) if not indicated differently. For cloning procedures, strains were grown in LB media (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, 10 g L^{-1} NaCl) and additionally supplemented with 50 µg mL^{-1} neomycin and 20 µg mL^{-1} polymyxin if required. All strains were stored in 30 % glycerol at -80 °C. Prior to cultivation, strains were streaked on LB agar plates and grown at 30 °C.

Fermentation medium contained 30 g L^{-1} glucose, 0.05 g L^{-1} $CaCl_2$ x 2 H_2O, 5 g L^{-1} tryptone, 1.33 g L^{-1} $MgSO_4$ x 7 H_2O, 1.67 g L^{-1} KH_2PO_4, 2 mL L^{-1} RPMI 1640 vitamins solution (Merk, Germany) and 1 mL L^{-1} trace elements solution (2.5 g L-1 $FeSO_4$, 2.1 g L^{-1} $C_4H_4O_6Na_2$ x 2 H2O, 1.8 g L^{-1} $MnCl_2$ x 4 H_2O, 0.258 g L^{-1} H_3BO_3, 0.031 g L^{-1} $CuSO_4$ x 5 H_2O, 0.023 g L^{-1} $NaMoO_4$ x 2 H_2O, 0.075 g L^{-1} $CoCl_2$ x 7 H2O, 0.021 g L^{-1} $ZnCl_2$). Preculture medium was prepared equal to the fermentation medium apart from a reduced glucose concentration of 10 g L^{-1} and additional 20 g L^{-1} MOPS adjusted to pH 7.

CRISPR/Cas9 mediated knock-out of glycosyltransferases

All knock-outs were performed as described previously (Rütering et al., 2017). Briefly, gRNAs for each target were cloned into pCasPP via Golden Gate assembly using BbsI. Afterwards, 1 kb up- and downstream homology flanks of the gene of interest were ligated into a unique SpeI site, followed by transformation of chemically competent *E. coli* S17-1. *P. polymyxa* was transformed by conjugation using *E. coli* S17-1 harboring the various plasmids. Overnight cultures of donor and recipient strains were diluted 1:100 with selective or non-selective LB media respectively and cultivated at 37 °C for 3 h, 280 rpm. 900 µL of the recipient culture was heat shocked at 42 °C for 15 min and mixed with 300 µL of the donor strain. Cells were centrifuged at 6,000 g for 2 min, resuspended in 800 µL LB media and dropped on non-selective LB agar plates. After 24 h of incubation at 30 °C, cells were scrapped off, resuspended in 500 µL LB broth and 100 µL thereof plated on selective LB-agar containing 50 µg mL^{-1} neomycin and 20 µg mL^{-1} polymyxin

for counter selection. *P. polymyxa* conjugants were analyzed for successful transformation after 48 h incubation at 30 °C by cPCR. Confirmed knock-out strains were plasmid cured by cultivation in LB broth without antibiotic selection pressure and subsequent replica plating on LB agar plates both with and without neomycin. Strains that did not grow on plates with selection marker were verified by sequencing of the target region and used for further experiments.

Fermentative EPS production

Fermentative production of EPS was performed in 1 L benchtop DASGIP parallel bioreactor systems (Eppendorf, Germany) with a working volume of 500 mL equipped with a 6-blade Rushton impeller over 28 h with a controlled pH of 6.8 and pO2 saturation of 30 %. Batch cultivations were started with an initial OD600 of 0.1 by inoculation with an appropriate volume of preculture. After fermentation, biomass was separated by centrifugation (15,000 g, 20 °C, 20 min) followed by cross-flow filtration of the supernatant using a 100 kDa filtration cassette (Hydrosart, Sartorius AG, Germany). Highly viscous EPS variants were diluted 1:10 with ddH2O prior to centrifugation. Concentrated supernatant was afterwards poured slowly into two volumes of isopropanol. Precipitated EPS was collected and dried overnight in a VDL53 vacuum oven at 40 °C (Binder, Germany). Dry weight of the obtained EPS was determined gravimetrically, before milling to a fine powder in a ball mill at 30 Hz for 1 min (Mixer Mill MM400, Retsch GmbH, Germany)

Carbohydrate fingerprinting

Monomer composition of engineered EPS variants were analyzed by the 1-phenyl-3-methyl-5- pyrazolone high throughput method (HT-PMP) (Rühmann, Schmid, & Sieber, 2014). In short, 0.1 % EPS solutions were hydrolyzed in a 96 well plate, sealed with a rubber mat and further covered by a custom-made metal device with 2 M TFA (90 min, 121 °C). Samples were neutralized with 3.2 % NH4OH. 75 µL of PMP master mix (0.1 M methanolic PMP:0.4 % ammonium hydroxide 2:1) were added to 25 µL of neutralized hydrolysate and incubated at 70 °C for 100 min in a thermal cycler. 20 µL of derivatized samples were mixed with 130 µL of a 1:26 dilution of 0.5 M acetic acid and filtered with a 0.2 µm filter plate (1000 g, 2 min) followed by HPLC-UV-MS using an Ultimate 3000 RS HPLC system (Dionex, USA). Separation was performed on a reverse phase column (Gravity C18, 100 x 2 mm, 1.8 µm particle size, Macherey-Nagel, USA) set to 50 °C. Gradient elution was performed using a mobile phase A (5 mM ammonium acetate set to pH 5.6 with 15 % acetonitrile) and mobile phase B (100 % acetonitrile) with a constant pump rate of 0.6 mL min^{-1}.

Enzymatic assays for glucose and pyruvate quantification

Glucose and pyruvate concentrations of the obtained EPS samples were determined enzymatically before and after hydrolysis to deduct residual contaminations from the fermentation medium as previously described (Rühmann, Schmid, & Sieber, 2015).

Partial hydrolysis of EPS samples

Analysis of partially hydrolyzed EPS samples was performed as described for carbohydrate fingerprinting. Differently, hydrolysis was performed at 90 °C for 105 min. After PMP derivatization, 25 µL 0.5 M acetic acid, 125 µL dH_2O and 500 µL acetonitrile were added to the reaction mix in order to quench the derivatization process and precipitate non-hydrolyzed polymer. Precipitated polymer was collected by centrifugation (24,000 g, 5 min) and the obtained supernatant transferred to HPLC-vials. HPLC-UV-MS analysis was performed using a YMC-Triart Diol-HILIC column (100 x 2 mm, 1.9 µm particle size). Column temperature was set to 7 °C and a constant flow of 0.3 mL min^{-1} 5 mM of ammonium-acetate at pH 5.6 with 15 % acetonitrile was applied.

Degradation of glycosyluronic acids

For glycosyluronic acid degradation 75 mg of polymer were dissolved in 10 mL anhydrous ethylenediamine. Lithium scrapes were added until a deep blue color was obtained and kept in this state for 1.5 h. The reaction was stopped by cooling down the solution on ice and the addition of 5 mL dH_2O. The solution was co-distilled with toluene until dryness. The polymer was then reconstituted in 5 mL dH_2O and adjusted to a pH of 6 with 4 M TFA. Finally, samples were derivatized with PMP and analyzed as described above.

Linkage analysis

Linkage analysis was performed according to the Hakomori procedure (Hakomori, 1964). Briefly, 1 mg polymer was lyophilized overnight and reconstituted in 250 µL anhydrous DMSO by stirring on a magnetic stirrer overnight. If polymer flakes remained insoluble, samples were further treated by 10 min sonication. The polyanion of the polymer was formed by the addition of 500 µL 2 M dimsyl sodium under stirring overnight. After freezing the sample, 250 µL methyl iodide were added and incubated for 1 h at 121 °C. Residual methyl iodide was blown off and the methylated polymer was recovered by MeOH/$CHCl_3$ extraction. Polymer hydrolysis was performed using 2 M TFA (121 °C, 2 h), reduced with 0.25 M NaH_4B (25 °C, 1 h) and acetylated (250 µL acetic anhydride, 250 µL pyridine, 100 °C, 20 min). Partially methylated alditol acetates were extracted with ethyl acetate and used for GC-MS analysis. Analysis of uronic acids was performed by reducing polymer samples prior to methylation analysis. Therefore, 2 g L^{-1} polymer solutions were incubated in 156 mM *N*-ethyl-*N*-(3-dimethylaminopropyl)-carbodiimide hydrochloride at pH 4.75 for 1.5 h and reduced with 1 mM $NaHB_4$ for 1 h, followed by pH neutralization with 2 M TFA and dialysis of the samples. For GC analysis, a Rxi-5 Sil MS column (30 m, 0.25 mm ID, d_f 0.25 µm) column was used with a constant helium flow of 0.3 mL min^{-1}. The initial temperature was set to 120 °C and was increased via a heat ramp of 7.5 °C min^{-1} to 180 °C followed by a constant temperature for five minutes and afterwards further ramped up to 230°C. Internal references to determine the retention times of partially methylated glycoside were obtained via Purdie-Irvine methylation of

methyl-glycosides by removing samples of the incremental per-methylation every 3 h for a time period of 12 h (Purdie & Irvine, 1903; Sassaki, Gorin, Souza, Czelusniak, & Iacomini, 2005).

NMR analysis

The EPS was dissolved in ddH_2O and degraded using mechanical shearing by ten cycles on a Star Burst Mini (Wet Jet HJP-25001 CE, Sugino Machine Ltd., Uozu, Japan) and the water was afterwards removed by freeze-drying. To remove remaining protein material the EPS was redissolved in ddH_2O (~2% w/v solution), heated to 95 °C for 10 mins, centrifuged, and the supernant was transferred to a new tube and dried again by freeze-drying. From this sample, 10 mg of EPS was redissolved in 600 µL of deuterated water (99.9% D_2O, Sigma Aldrich) transferred to 5 mm NMR tube (Bruker, LabScape Stream 103.5 mm length). All homo- and heteronuclear experiments were acquired at a temperature of 273.1 K on a Bruker AV-IIIHD 800 MHz spectrometer (Bruker BioSpin AG, Fälladen, Switzerland) equipped with a 5 mm cryogenic CP-TCI z-gradient probe. The following experiments were collected: 1D proton, 1H-13C heteronuclear single quantum coherence (HSQC) with multiplicity editing, 1H-13C heteronuclear two bond correlation (H2BC) spectroscopy, 1H-13C heteronuclear multiple bond coherence (HMBC) with suppression of one-bond correlations, and 1H-1H double quantum filtered correlation spectroscopy (DQF-COSY). The spectra were recorded, processed, and analyzed using the TopSpin 3.5 or 4.0.1 software (Bruker BioSpin AG).

Results and Discussion

Carbohydrate fingerprints

As it was not possible, to properly separate the different polymers produced by *P. polymyxa* DSM 365 via gel permeation chromatography, combinatorial knock-outs of glycosyltransferases were performed. In a first step, carbohydrate fingerprints of different knock-out variants were obtained using the HT-PMP method (Figure 2). In addition to the monomer composition, partial sequence information was obtained by detecting dimers that withstood chemical hydrolysis by 2 M TFA (Table 1). The glycosyltransferases PepQ, PepT, PepU and PepV corresponding to the polymer paenan II were responsible for the formation of a GlcA-Fuc dimer. We recently demonstrated that the gene *epsO* corresponds to a pyruvyltransferase that forms the pyruvate ketal of paenan I. Deletion results in the absence of pyruvate in the polymer, without affecting the carbohydrate composition of the EPS. It was also possible to link the genes *pepC-F* to the formation of the pyruvate ketal (Table 1). For the glycosyltransferases PepI-L, knock-out of the underlying genes resulted in the absence of a GlcA-Hex dimer, which was present in all other single knock-out variants. Due to the

early retention time of the dimer, it was possible to identify the hexose as a mannose. MS/MS fragmentation revealed that the hexose was positioned at the reducing end of the dimer (Figure S 1).

Table 1: Key features of EPS variants obtained from knock-out strains of *P. polymyxa* DSM 365. GlcA dimers and a putative galactose pyruvate ketal endured chemical hydrolysis with 2 M TFA and were detected by MS-MS analysis. Knock-out of a single glycosyltransferase resulted in the absence of distinct key features.

	Gal$^{4,6\text{-Pyr}}$	**GlcA-Fuc**	**GlcA-Man**
Strain	***m/z* 581**	***m/z* 671**	***m/z* 687**
WT	✓	✓	✓
ΔpepC		✓	✓
ΔpepD		✓	✓
ΔpepF		✓	✓
ΔpepI	✓	✓	
ΔpepJ	✓	✓	
ΔpepK	✓	✓	
ΔpepL	✓	✓	
ΔpepQ	✓		✓
ΔpepT	✓		✓
ΔpepU	✓		✓
ΔpepV	✓		✓

Interestingly, knock-out of single glycosyltransferases only resulted in the absence of one of these key features, indicating the presence of multiple polymers or side chains. Therefore, we tried to combine knock-outs of paenan II (indicated by the missing GlcA-Fuc dimer) with deletions of genes responsible for other key features. When both initiating glycosyltransferases PepC and PepQ were deleted, no EPS was produced under the applied production conditions by *P. polymyxa* (Figure 2). Thus, we could experimentally confirm that no further initiating glycosyltransferases complementing the function of PepC or PepQ are present in the genome. Interestingly, the EPS of the double-knock out strain *P. polymyxa ΔpepQF* did not contain any traces of galactose. Furthermore, pyruvate could not be detected. Due to the susceptibility of uronic acids towards degradation during chemical hydrolysis with TFA, the amount of glucuronic acids in all variants is underestimated. Nevertheless, for the variant *ΔpepQF* a molar ratio of 2:2:1 Glc:Man:GlcA (1:0.7:0.1 without correction) could be determined.

Combinatorial knock-outs of the glycosyltransferases involved in the formation of the GlcA-Man dimer (PepI-L) with genes responsible for the biosynthesis of paenan II (*pepQ-V*) demonstrated similar monomer compositions. Both, GlcA as well as the GlcA-Man dimer could not be detected anymore. However, the pyruvate content significantly increased. As a result, a molar ratio of 2:1:1:1 Glc:Man:Gal:Pyr (1.5:1:1:1 without correction) was determined for these variants.

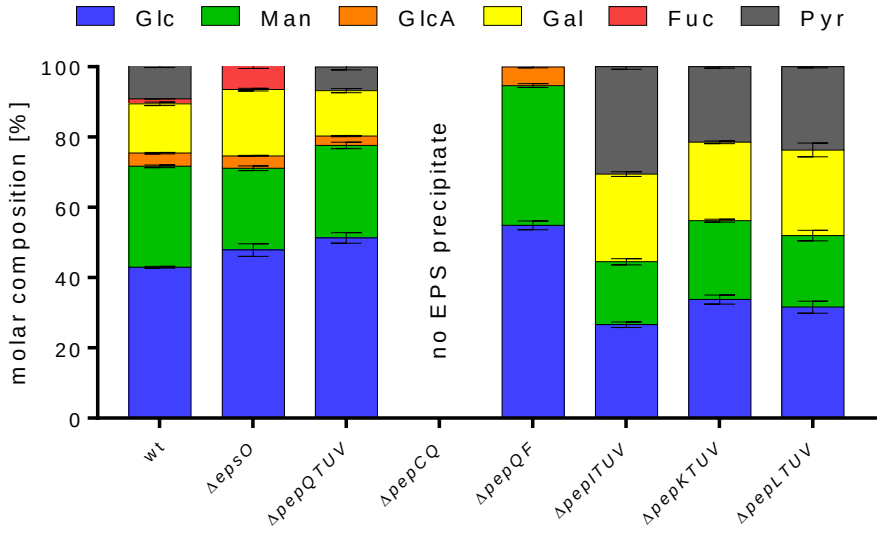


Figure 2: Carbohydrate fingerprints of EPS variants obtained from *P. polymyxa* DSM 365. Genes corresponding to glycosyltransferases were deleted by CRISPR-Cas9 mediated genome engineering resulting in different polymer compositions. Deletion of genes corresponding to a minor fucose containing polymer paenan II (*pepQ-V*) resulted in the absence of fucose. Carbohydrates were quantified by the HT-PMP method (Rühmann et al., 2014), while the pyruvate content was determined enzymatically.

Partial hydrolysis

In order to obtain initial information on the monomer sequence of the polysaccharide, partial hydrolysis experiments were performed at reduced temperature to achieve incomplete hydrolysis (Table 2, Table S4). For all variants in which pyruvate was previously determined, pyruvate linked to hexose oligomers of variable length were detected (Figure S2). The largest fragments obtained in considerable amounts consisted of pyruvate residue linked to five hexose monomers (*m/z* 1229). Due to the fragmentation pattern of oligosaccharides containing the pyruvate ketal, the pyruvic acid was positioned at the terminal end of a putative side chain consisting of galactose and and a second hexose (Table S4). Contrarily, for both knock-out variants *ΔpepD* and *ΔpepF*, which did not contain any pyruvate, the largest obtained oligomer was significantly smaller with *m/z* 835 corresponding to a hexose trimer. Interestingly, even though we could observe a GlcA-Man dimer using harsher hydrolysis conditions (121 °C, 120 min) in HT-PMP analysis, hardly any GlcA-Hex oligomers were detected in the EPS of any knock-out variant when hydrolysis temperature and incubation time were reduced. Only in the

wildtype polymer, minor peaks with MS-spectra corresponding to hexose oligomers containing an uronic acid were detected (Table 2). While free uronic acids monomers are prone towards degradation, the glycosidic linkage of uronic acids to its adjacent sugar at the reducing end is very stable (Rodén, Baker, Anthony Cifonelli, & Mathews, 1972). As a result, it might be possible that the polymer corresponding to the GlcA-Man dimer was not hydrolyzed properly to allow detection using the HILIC method. Particularly for the variant *ΔpepQF*, in which both, paenan II as well as the pyruvate ketal were deleted, only hexose monomers rather than any oligomers were detected. This might indicate that the larger hexose oligomers obtained from other variants might be attributed to a separate polymer and that the polymer corresponding to the GlcA-Man dimer is quite resilient towards chemical hydrolysis.

Table 2: Partial hydrolysis of EPS knock-out variants resulted in different polysaccharide fragments corresponding to previously detected GlcA-dimers and pyruvate ketals**.** Absence of key features resulted in hexose oligomers of variable length. A detailed overview of different variants and an exemplary fragmentation is given in the supplemental information (Table S4, Figure S2)

	PMP>Hex *m/z* 511	**PMP>Hex-Hex *m/z* 673**	**PMP>Hex-Hex-Hex *m/z* 835**	**PMP>Hex-Hex-Hex-Hex *m/z* 997**	**PMP>Hex-Hex-Gal$^{4,6\text{-Pyr}}$-Hex-Hex *m/z* 1229**	**PMP>Fuc-GlcA-Hex-Hex *m/z* 997**	**PMP>Hex-GlcA-Hex-Hex *m/z* 1011**
WT	✓	✓	✓	✓	✓	✓	✓
Δ*pepD*	✓	✓				✓	
Δ*pepF*	✓		✓			✓	
Δ*pepJ*	✓				✓	✓	
Δ*pepQ*	✓	✓	✓	✓	✓		*
Δ*pepQF*	✓						

Colored background: oligomers corresponding to key features detected in carbohydrate fingerprints. *: traces detectable

Lithium degradation of uronic acids

To follow up on the hypothesis that the heteroexopolysaccharide produced by *P. polymyxa* DSM 365 consists of three distinct polymers, lithium degradations of uronic acids was performed for two EPS variants. While the carbohydrate fingerprint for the

variant *ΔepsO* remained identical to the wildtype strain (Figure 2), viscosity of the EPS solution was severely reduced, facilitating lithium degradation of the uronic acid. Furthermore, the doube knock-out variant *ΔpepQF* was analyzed to determine the position of the uronic acid.

For *ΔepsO*, the obtained main fragment was related to the fucose containing polymer paenan II with *m/z* 819 corresponding to Hex-Fuc-Hex-PMP (Figure 4 A). However, additional hexose oligomers were detected with *m/z* 835 relating to a hexose trimer and *m/z* 673 corresponding to a hexose dimer. Even though partial hydrolysis of this polymer showed hexose oligomers up to pentamers, only hexose trimers were obtained by lithium degradation of GlcA. As a result, this indicates that the GlcA-Man indeed corresponds to yet another polymer. The results were further cross-verified by lithium degradation of the EPS obtained from *P. polymyxa ΔpepQF*. HPLC-UV analysis of the degradation product of this polymer also confirmed the presence of a hexose trimer as well as a dimer (Figure 4 B). Due to the knock-out of pepQ, no fragmentation products associated with paenan II were obtained. Owing to the clear fragmentation of the polymer, it was expected that also in the third heteroexopolysaccharide produced by *P. polymyxa* DSM 365 GlcA is positioned in the backbone of the polymer. Consequently, the repeating unit of this polymer, henceforward named paenan III, consists of GlcA and at least three hexoses with mannose bound to the reducing end of the uronic acid. During lithium degradation of uronic acids, sugar residues adjacent to the uronic acid are often also affected by the β-elimination event (Lau, McNeil, Darvill, & Albersheim, 1987). Consequently, it remains to be investigated, whether the obtained dimer is a distinct fragmentation product or if it corresponds to a degradation product of the hexose trimer. The obtained carbohydrate fingerprints of the *ΔpepQF* variant suggested a molar ratio of 2:2:1 Glc:Man:GlcA. Therefore, it seems reasonable that GlcA is positioned in the backbone as the branching glycosyl residue and that the missing hexose residue might belong to a putative side chain.

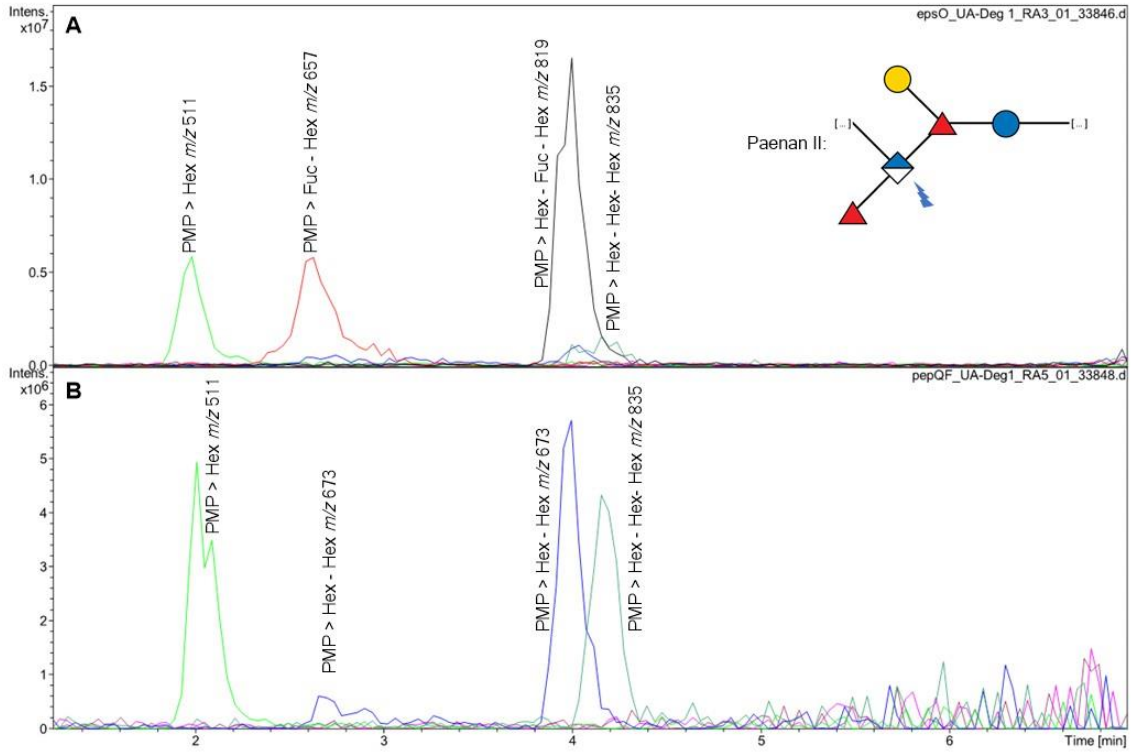


Figure 4: A) EPS oligomers obtained from *P. polymyxa ΔepsO* after lithium degradation of uronic acids and PMP derivatization. Fragments containing fucose correspond to a minor polymer termed paenan II. Additionally, hexose di- and trimers were detected in low amounts corresponding to another polymer named paenan III. B) EPS oligomers obtained from *P. polymyxa ΔpepQF* after lithium degradation of uronic acids and PMP derivatization. Carbohydrate fingerprinting indicated that this polymer does not contain the pyruvate ketal associated with paenan I and fucose associated with paenan II. Consequently, obtained hexose oligomers correspond to a third polymer paenan III. blue diamond: GlcA; red triangle: Fuc; blue circle: Glc; yellow circle: Gal. Linkage pattern analysis

Linkage pattern analysis

Linkage pattern analysis was performed by Hakomori methylation of various EPS variants (Table 3). For the wildtype strain, the EPS composition demonstrated a complex array of different linkage types, which was expected as the polysaccharide mixture consists of three distinct polymers. Using CRISPR-Cas9 mediated knock-out variants, the linkage pattern could be significantly simplified by deletion of individual polymers. Based on previous results, individual polymers were assigned by key features of their carbohydrate fingerprints (Table 1). The pyruvate ketal was attributed to paenan I, presence of fucose to paenan II and a GlcA-Man dimer to paenan III. Using this preliminary information, it was possible to construct combinatorial knock-out variants to specifically investigate individual polymers and assign glycosidic linkages to each paenan variant (Table 1).

Table 3: Linkage pattern analysis of partially methylated alditol acetates (PMAAs) of EPS obtained from *P. polymyxa* DSM 365 (wt) and different knock-out variants.

		Area [%]					
PMAA	**RT**	**wt**	***ΔpepFJ***	***ΔpepQTUV***	***ΔpepLTUV***	***ΔpepQF***	***ΔpepQF* red.**
Paenan		**I+II+III**	**II**	**I+II**	**I**	**III**	**III**
2,3,4,6-Me-Glc	1.00	12.0		5.9		21.4	14.0
2,3,4,6-Me-Man	1.01	20.5		11.9		27.5	25.3
2,3,4,6-Me-Gal	1.05	*	25.5				
2-Me-Fuc	1.07	*	34.7				
2,4,6-Me-Glc	1.22	21.8		24.9	15.1	27.0	12.3
2,3,6-Me-Man	1.23	8.5		10.2	20.6		
2,3,6-Me-Glc	1.25	10.5	39.8	10.4	25.6		16.4
2,4,6-Me-Man	1.26	3.2		4.1		12.9	12.6
2,6-Me-Man	1.38	5.1		10.1		11.1	12.4
2,6-Me-Gal	1.39	13.5		14.2	22.8		7.1
2,3-Me-Gal	1.49	4.7		8.5	15.9		

2,3,4,6-Me-Gal = 1,5-di-acetyl-2,3,4,6-tetra-O-methylgalactitol etc.; RT: Relative retention time of the corresponding PMAA relative to 2,3,4,6-Me-Glc using a Rxi-5 Sil MS column; Percent area under curve in TIC of all integrated peaks; red.: EPS sample was chemically reduced prior to methylation to enable detection of uronic acids;*: traces detectable; Paenan: the first row indicates which polysaccharides are putatively present in the respective knock-out variants based on the detection of key features in previous analysis. An exemplary methylation analysis and MS spectra are given in Figure S3.

We previously characterized paenan II in detail (Schilling et al., 2022). However, both the fucose content as well as the linkage pattern analysis suggested that the EPS produced by the wildtype strain was composed of only a small amount of paenan II. Consequently, deletion of the genes corresponding to this polymer (*pepQ-V*) did not significantly alter the linkage pattern (Table 1). However, combining the knock-out of paenan II with knock-outs attributed to the previously detected GlcA-Man dimer (*pepI-L*) significantly reduced obtained glycosidic linkage types. Both, terminal glucose and mannose were present in the wildtype strain but diminished in the variant Δ*pepLTUV*. Additionally, 1,3-linked mannose and a 1,3,4-linked mannose residue could not be detected anymore. While both terminal hexoses vanished in this variant, 2,6-Me-Gal and 2,3-Me-Gal were still detected, indicating branching sugar residues and a side chain. In the HT-PMP analysis, two peaks corresponding to a hexose with a pyruvate ketal were observed in each polymer containing pyruvate. Typically, pyruvate ketals are acid-labile and do not endure hydrolysis conditions used in this study (Hager, Sützl, Stefanović, Blaukopf, & Schäffer, 2019). However, galactose is the C4 epimer of glucose, resulting in a more resilient pyruvate ketal compared to other hexoses and allowed the detection in HT-PMP analysis. Consequently, the detected 2,3-Me-Gal corresponds to a terminal galactose residue linked at C4 and C6 to a pyruvate ketal. Similar residues have previously been identified in the EPS from *P. polymyxa* (Madden et al., 1986).

Furthermore, 1,3- and 1,4-linked glucose as well as 1,4-linked mannose were attributed to paenan I. Interestingly, all detected partially methylated alditol acetates (PMAAs) appeared in approximately equimolar ratios, confirming the results of HT-PMP analysis and suggesting a molar ratio of 2:1:2:1 Glc:Man:Gal:Pyr.

To identify the linkage pattern of paenan II, the previously investigated variant *ΔpepQF* was further investigated. As expected, this polymer variant showed both, the terminal glucose and mannose residues indicating two sidechains. In addition to 1,3-linked glucose, 1,3-linked and 1,3,4-linked mannose were detected. Both non-terminal mannose residues appeared in stoichiometrically lower amounts compared to the other PMAAs. However, the previously detected GlcA-Man dimer is resilient towards chemical hydrolysis and cannot be detected via GC-MS. Consequently, the corresponding mannose might be underrepresented. The appearance of two non-terminal mannose residues might further indicate an irregular branching sidechain. To identify the linkage pattern of the uronic acid in paenan III, the sample *ΔpepQF* was reduced prior to methylation to enable the detection (Janson, Kenne, Liedgren, Lindberg, & Lönngren, 1976). Compared to the non-reduced sample, two additional PMAAs were detected. 2,3,6-Me-Glc indicates 1,4-linked GlcA, while 2,6-Me-Glc corresponds to a branching 1,3,4-linked GlcA. As a result, these PMAAs might be attributed to a branching GlcA residue with an irregular sidechain. However, the ratio of terminal sugars to branching residues was approximately 2:1 indicating that the branching glycosyl residues were clearly underrepresented. Another explanation might be that the sonication procedure prior to Hakomori-methylation resulted in the cleavage of terminal sugar residues giving rise to 1,3-linked Man and 1,4-linked GlcA respectively.

Due to the presence of solely hexoses in paenan I, no clear structure of the repeating unit could be deduced from the combined results of sequence and linkage analysis. However, for the GlcA-Man containing paenan III a structure as depicted in Figure 5 can be assumed.

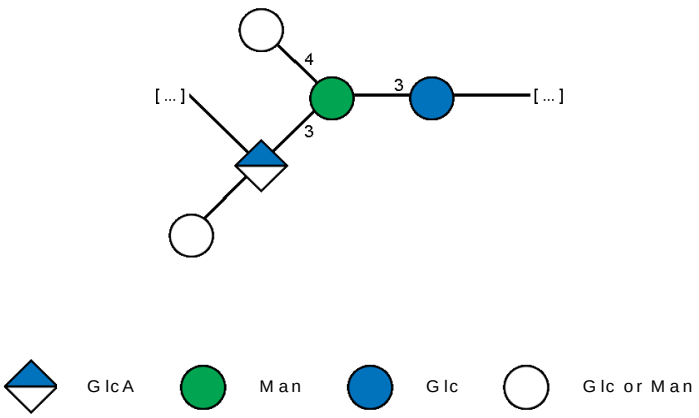


Figure 5: Putative structure of the repeating unit of paenan III derived from uronic acid degradation and methylation analysis. Terminal hexoses can be Glc or Man and might correspond

to irregular side chains. The GlcA residue is linked via C3 and C4 to a second side chain and to the backbone of the repeating unit.

NMR

Preliminary NMR data suggested that the terminal galactose of paenan I is linked β-1-3 to a glucose residue. For paenan III, the side chain hexose linked to a branching glucuronic acid residue could be attributed to an α-1-4-linked mannose. Furthermore, conformation of the initial glucose residue was determined as a β-1-3 linkage. HSQC spectra were able to confirm the preliminary structure given in Figure 5 and used to elucidate the repeating unit of paenan III. While NMR analysis of distinct variants of the heteroexopolysaccharide from *P. polymyxa* DSM 365 is currently still going on, a preliminary structure for paenan I and paenan III can be postulated (Figure 6).

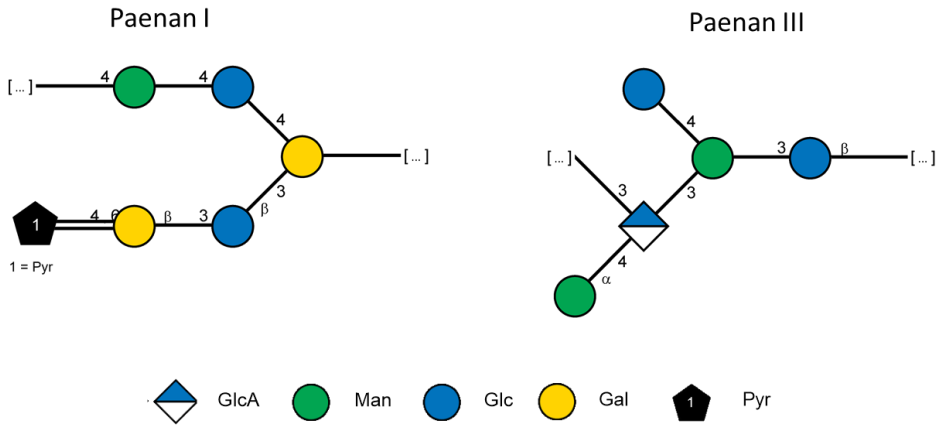


Figure 6: Preliminary structures of paenan I and paenan III were obtained by combinatorial carbohydrate analytics. For paenan I, sequence of the polymer backbone consisting of 1,4-linked Glc, 1,4-linked Man and a branching Gal residue remains to be elucidated. Structure of paenan III was elucidated for most part, however, further NMR analysis is required to identify the conformation of remaining sugars.

Conclusion

We recentely elucidated the structure of a minor fucose containing heteropolysaccharide produced by *P. polymyxa* DSM 365. Continuing our efforts to construct CRISPR-Cas9 mediated knock-outs of glycosyltransferases, it was possible to isolate individual polymers of the EPS composition obtained from *P. polymyxa*. Using a combinatorial analytical approach of complementary methods, structures of the repeating unit of paenan I and paenan III could be postulated. However, further NMR analysis is required to identify the sequence of the glycosyl residues of the corresponding backbone. An additional to the preliminary polysaccharide structures, distinct glycosyltransferases could be attributed to the biosynthesis of the individual polymers. By that, this study lays the foundation for further structure-function analysis of the EPS from *P. polymyxa* DSM 365.

Acknowledgement

The authors would like to acknowledge the technical support from Petra Lommes and Julia Wünsche for methylation analysis experiments. This work was supported by the German Federal Ministry of Education and Research (BMBF) in frame of the project MaPolKo (number 03VP02560).

References

Aguilera, M., Monteoliva-Sánchez, M., Suárez, A., Guerra, V., Lizama, C., Bennasar, A., & Ramos-Cormenzana, A. (2001). *Paenibacillus jamilae* sp. Nov., an exopolysaccharide-producing bacterium able to grow in olive-mill wastewater. *International Journal of Systematic and Evolutionary Microbiology*, *51*(Pt 5), 1687–1692.

Ash, C., Priest, F. G., & Collins, M. D. (1994). Molecular identification of rRNA group 3 Bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test: Proposal for the creation of a new genus *Paenibacillus*. *Antonie van Leeuwenhoek*, *64*(3–4), 253–260.

Badri, A., Williams, A., Linhardt, R. J., & Koffas, M. A. (2018). The road to animal-free glycosaminoglycan production: Current efforts and bottlenecks. *Current Opinion in Biotechnology*, *53*, 85–92.

Bentley, S. D., Corton, C., Brown, S. E., Barron, A., Clark, L., Doggett, J., Harris, B., et al. (2008). Genome of the Actinomycete Plant Pathogen Clavibacter michiganensis subsp. Sepedonicus Suggests Recent Niche Adaptation. *Journal of Bacteriology*, *190*(6), 2150–2160.

Burdick, J. A., & Prestwich, G. D. (2011). Hyaluronic Acid Hydrogels for Biomedical Applications. *Advanced Materials*, *23*(12), H41–H56.

Çolak, F., Olgun, A., Atar, N., & Yazıcıoğlu, D. (2013). Heavy metal resistances and biosorptive behaviors of *Paenibacillus polymyxa*: Batch and column studies. *Journal of Industrial and Engineering Chemistry*, *19*(3), 863–869.

Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J., & Koffas, M. A. G. (2014). Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews*, *38*(4), 660–697.

De Mas, C., Jansen, N. B., & Tsao, G. T. (1988). Production of optically active 2,3-butanediol by *Bacillus polymyxa*. *Biotechnology and Bioengineering*, *31*(4), 366–377.

Diab, T., Biliaderis, C. G., Gerasopoulos, D., & Sfakiotakis, E. (2001). Physicochemical properties and application of pullulan edible films and coatings in fruit preservation. *Journal of the Science of Food and Agriculture*, *81*(10), 988–1000.

Freitas, F., Alves, V. D., & Reis, M. A. M. (2011). Advances in bacterial exopolysaccharides: From production to biotechnological applications. *Trends in Biotechnology*, *29*(8), 388–398.

Galili, U. (1993). Interaction of the natural anti-Gal antibody with α-galactosyl epitopes: A major obstacle for xenotransplantation in humans. *Immunology Today*, *14*(10), 480–482.

Grady, E. N., MacDonald, J., Liu, L., Richman, A., & Yuan, Z.-C. (2016). Current knowledge and perspectives of *Paenibacillus*: A review. *Microbial Cell Factories*, *15*(1), 203.

Hager, Sützl, Stefanović, Blaukopf, & Schäffer. (2019). Pyruvate Substitutions on Glycoconjugates. *International Journal of Molecular Sciences*, *20*(19), 4929.

Hakomori, S. (1964). A Rapid Permethylation of Glycolipid, and Polysaccharide Catalyzed by Methylsulfinyl Carbanion in Dimethyl Sulfoxide. *Journal of Biochemistry*, *55*, 205–208.

Islam, S. T., & Lam, J. S. (2013). Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria: Role of Wzx in surface polysaccharide synthesis. *Environmental Microbiology*, *15*(4), 1001–1015.

Janson, P.-E., Kenne, L., Liedgren, H., Lindberg, B., & Lönngren, J. (1976). A Practical Guide to the Methylation Analysis of Carbohydrates. *Chemical Communications*, *8*.

Jeong, H., Choi, S.-K., Ryu, C.-M., & Park, S.-H. (2019). Chronicle of a Soil Bacterium: *Paenibacillus polymyx*a E681 as a Tiny Guardian of Plant and Human Health. *Frontiers in Microbiology*, *10*, 467.

Jung, H.-K., Hong, J.-H., Park, S.-C., Park, B.-K., Nam, D.-H., & Kim, S.-D. (2007). Production and physicochemical characterization of β-glucan produced by *Paenibacillus polymyxa* JB115. *Biotechnology and Bioprocess Engineering*, *12*(6), 713–719.

Lau, J. M., McNeil, M., Darvill, A. G., & Albersheim, P. (1987). Selective degradation of the glycosyluroic acid residues of complex carbohydrates by lithium dissolved in ethylenediamine. *Carbohydrate Research*, *168*(2), 219–243.

Lee, I. Y., Seo, W. T., Kim, G. J., Kim, M. K., Ahn, S. G., Kwon, G. S., & Park, Y. H. (1997). Optimization of fermentation conditions for production of exopolysaccharide by *Bacillus polymyxa*. *Bioprocess Engineering*, *16*(2), 71.

Li, O., Lu, C., Liu, A., Zhu, L., Wang, P.-M., Qian, C.-D., Jiang, X.-H., et al. (2013). Optimization and characterization of polysaccharide-based bioflocculant produced by Paenibacillus elgii B69 and its application in wastewater treatment. *Bioresource Technology*, *134*, 87–93.

Liang, T.-W., & Wang, S.-L. (2015). Recent advances in exopolysaccharides from *Paenibacillus* spp.: Production, isolation, structure, and bioactivities. *Marine Drugs*, *13*(4), 1847–1863.

Liu, J., Luo, J., Ye, H., Sun, Y., Lu, Z., & Zeng, X. (2010a). In vitro and in vivo antioxidant activity of exopolysaccharides from endophytic bacterium *Paenibacillus polymyxa* EJS-3. *Carbohydrate Polymers*, *82*(4), 1278–1283.

Liu, J., Luo, J., Ye, H., Sun, Y., Lu, Z., & Zeng, X. (2010b). Medium optimization and structural characterization of exopolysaccharides from endophytic bacterium *Paenibacillus polymyxa* EJS-3. *Carbohydrate Polymers*, *79*(1), 206–213.

Madden, J. K., Dea, I. C. M., & Steer, D. C. (1986). Structural and rheological properties of the extracellular polysaccharides from *Bacillus polymyxa*. *Carbohydrate Polymers*, *6*(1), 51–73.

Miao, T., Wang, J., Zeng, Y., Liu, G., & Chen, X. (2018). Polysaccharide-Based Controlled Release Systems for Therapeutics Delivery and Tissue Engineering: From Bench to Bedside. *Advanced Science*, *5*(4), 1700513.

Moradali, M. F., & Rehm, B. H. A. (2020). Bacterial biopolymers: From pathogenesis to advanced materials. *Nature Reviews Microbiology*.

Ninomiya, E., & Kizaki, T. (1969). Bacterial polysaccharide from *Bacillus polymyxa* No. 271. *Die Angewandte Makromolekulare Chemie*, *6*(1), 179–185.

Purdie, T., & Irvine, J. C. (1903). C.—The alkylation of sugars. *J. Chem. Soc., Trans.*, *83*(0), 1021–1037.

Raza, W., Makeen, K., Wang, Y., Xu, Y., & Qirong, S. (2011). Optimization, purification, characterization and antioxidant activity of an extracellular polysaccharide produced by *Paenibacillus polymyxa* SQR-21. *Bioresource Technology*, *102*(10), 6095–6103.

Riva, S. (2001). Biocatalytic modification of natural products. *Current Opinion in Chemical Biology*, *5*(2), 106–111.

Rodén, L., Baker, J. R., Anthony Cifonelli, J., & Mathews, M. B. (1972). [7] Isolation and characterization of connective tissue polysaccharides. *Methods in Enzymology* (Vol. 28, pp. 73–140).

Rühmann, B., Schmid, J., & Sieber, V. (2014). Fast carbohydrate analysis via liquid chromatography coupled with ultra violet and electrospray ionization ion trap detection in 96-well format. *Journal of Chromatography A*, *1350*, 44–50.

Rühmann, B., Schmid, J., & Sieber, V. (2015). High throughput exopolysaccharide screening platform: From strain cultivation to monosaccharide composition and carbohydrate fingerprinting in one day. *Carbohydrate Polymers*, *122*, 212–220.

Rütering, M., Cress, B. F., Schilling, M., Rühmann, B., Koffas, M. A. G., Sieber, V., & Schmid, J. (2017). Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in Paenibacillus polymyxa. *Synthetic Biology*, *2*(1)

Rütering, M., Schmid, J., Rühmann, B., Schilling, M., & Sieber, V. (2016). Controlled production of polysaccharides–exploiting nutrient supply for levan and heteropolysaccharide formation in *Paenibacillus sp. Carbohydrate Polymers*, *148*, 326–334.

Sassaki, G. L., Gorin, P. A. J., Souza, L. M., Czelusniak, P. A., & Iacomini, M. (2005). Rapid synthesis of partially O-methylated alditol acetate standards for GC–MS: Some relative activities of hydroxyl groups of methyl glycopyranosides on Purdie methylation. *Carbohydrate Research*, *340*(4), 731–739.

Schilling, C., Badri, A., Sieber, V., Koffas, M., & Schmid, J. (2020). Metabolic engineering for production of functional polysaccharides. *Current Opinion in Biotechnology*, *66*, 44–51.

Schilling, C., Rühmann, B., Klau, L., Aachmann, F., Schmid, J., & Sieber, V. (2022). Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365. *Carbohydrate Polymers, 278*

Shida, O., Takagi, H., Kadowaki, K., Nakamura, L. K., & Komagata, K. (1997). Transfer of *Bacillus alginolyticus, Bacillus chondroitinus, Bacillus curdlanolyticus, Bacillus glucanolyticus, Bacillus kobensis,* and *Bacillus thiaminolyticus* to the Genus *Paenibacillus* and Emended Description of the Genus *Paenibacillus*. *International Journal of Systematic Bacteriology*, *47*(2), 289–298.

Wang, Y., Shi, Y., Li, B., Ibraham, M., & Sun, G. (2012). Phosphate solubilization of *Paenibacillus polymyxa* and *Paenibacillus macerans* from mycorrhizal and non-mycorrhizal cucumber plants. *African Journal of Microbiology Research*, *6*(21).

Xie, J.-B., Du, Z., Bai, L., Tian, C., Zhang, Y., Xie, J.-Y., Wang, T., et al. (2014). Comparative Genomic Analysis of N2-Fixing and Non-N2-Fixing *Paenibacillus* spp.: Organization, Evolution and Expression of the Nitrogen Fixation Genes. (P. M. Richardson, Ed.)*PLoS Genetics*, *10*(3), e1004231.

Xie, N.-Z., Li, J.-X., Song, L.-F., Hou, J.-F., Guo, L., Du, Q.-S., Yu, B., et al. (2015). Genome sequence of type strain *Paenibacillus polymyxa* DSM 365, a highly efficient producer of optically active (R,R)-2,3-butanediol. *Journal of Biotechnology*, *195*, 72–73.

3.3. Rheological characterization of artificial paenan compositions produced by *Paenibacillus polymyxa* DSM 365

Polysaccharides comprise a highly diverse group of biopolymers that have been industrially applied as rheology modifiers and constitute a sustainable alternative to commonly used petrochemical thickeners and viscosifying agents in numerous applications.

This study is focused on the rheological characterization of the heteroexopolysaccharides produced by *P. polymyxa* DSM 365. Using CRISPR-Cas9 mediated knock-outs of glycosyltransferases, individual paenan variants or compositions were produced via fermentation and their rheological material properties analysed in detail. With the previously determined structure of paenan II and the postulated structures of paenan I and paenan III, specific interactions are proposed between individual biopolymers resulting in highly different rheological behavior. While the wildtype composition containing all three paenan polymers forms highly viscous, soft hydrogels, individual paenan variants represent nearly Newtonian fluids. Addition of monovalent cations to the solution further strengthened interactions resulting in a more rigid gel character. The pronounced intermolecular network is hypothesized to result of the interaction of a pyruvyl residue of paenan I and the glucuronic acid found in the backbone of paenan III. Absence of the fucose containing paenan II results in a reduced temperature stability, suggesting further non-covalent intermolecular interactions. Using the structure-function relationship of different EPS compositions proposed in this study, paenan and different variants thereof might be applied in a variety of applications including thickening agents or high-value biomedical surface coatings and facilitate rational design of product formulations.

Design and planning of this study were performed by Christoph Schilling in collaboration with Moritz Gansbiller, Volker Sieber and Jochen Schmid. Strain construction, EPS production by fermentation and carbohydrate analysis were performed by Christoph Schilling. Rheological measurements were performed by Christoph Schilling. Rheological evaluation and modeling were performed by Moritz Gansbiller. Potential applications were evaluated by Christoph Schilling. All authors contributed to content and language of the manuscripts and provided scientific or technical advice.

Rheological characterization of artificial paenan compositions produced by *Paenibacillus polymyxa* DSM 365

Christoph Schilling‡, Moritz Gansbiller‡, Broder Rühmann, Volker Sieber & Jochen Schmid

Prepared manuscript

(2022)

‡: authors contributed equally

Abstract

Microbial exopolysaccharides (EPS) constitute a sustainable alternative to petroleum-based rheological modifiers. Recent studies revealed that the heteroexopolysaccharide produced by *Paenibacillus polymyxa* is composed of three distinct biopolymers termed paenan I, II and III. Using CRISPR-Cas9 mediated knock-out variants of glycosyltransferases, defined polysaccharide compositions were produced and rheologically characterized in detail. The high viscosity and gel-like character of the wildtype polymer is proposed to originate from the non-covalent interaction between a pyruvate residue of paenan I and the glucuronic acid found in the backbone of paenan III. Paenan II conveys thermostable properties to the EPS mixture. Opposed to the wildtype polymer mixture, knock-out variants demonstrated significantly altered rheological behavior. Using the structure-function relationship determined in this study, tailor-made paenan variants might be utilized in a wide range of applications including thickening agents, coatings or high value biomedical materials.

Introduction

Carbohydrate polymers represent a highly diverse class of functional polysaccharides found in all domains of life (Sutherland, 1972). Intracellular polysaccharides may function as energy storage (e.g. glycogen) or play a central role as integral parts of cellular structures (e.g. lipopolysaccharides, starch, cellulose etc.). Contrary, extracellular polysaccharides (EPS) cover numerous tasks as natural adhesives, barriers against abiotic stress factors such as desiccation and antibiotics or as protection against extreme environmental conditions (Danese, Pratt, & Kolter, 2000; Ophir & Gutnick, 1994; Poli, Anzelmo, & Nicolaus, 2010). Owing to their complex and diverse structures, a sheer endless variety of physicochemical properties can be derived from different EPS variants and compositions (Moradali & Rehm, 2020). Consequently, various polysaccharides have found commercial success as bulk products in food, feed and technical applications as well as in high-value niches in cosmetics and pharmaceutical industry (Freitas, Alves, & Reis, 2011; Jang, Zhang, Chon, & Choi, 2015; Kaur, Bera, Panesar, Kumar, & Kennedy, 2014; Kumar, Rao, & Han, 2018).

The global market of biopolymers is still dominated by plant and algae derived polysaccharides such as starch, alginates or carrageenans (Williams & Phillips, 2016). However, bacterial EPS variants from *Xanthomonas campestris*, different sphingans or hyaluronic acid produced by *Streptococcus* spp. have found commercial success in bulk products as well as in high-value niche applications (Bajaj, Survase, Saudagar, & Singhal, 2007; Becker, Katzen, Pühler, & Ielpi, 1998; Burdick & Prestwich, 2011). The main advantage of bacterial polysaccharides over their plant and algae derived counterparts is the ability to produce these polymers by large scale fermentation

processes in bioreactors under controlled conditions, making them independent of seasonal or geopolitical influences. Furthermore, ongoing research in the development of new genetic tools allowed the targeted modification of polymer structures to generate EPS variants with new physicochemical properties or enhance production via metabolic engineering (Hassler & Doherty, 1990; Schilling, Badri, Sieber, Koffas, & Schmid, 2020; Wu et al., 2019).

Paenibacillus polymyxa is a Gram-positive, spore-forming soil bacterium that has attracted increasing interest as a production platform for an arsenal of different antibiotics (Grady, MacDonald, Liu, Richman, & Yuan, 2016), 2,3-butanediol of extremely high enantiomeric purity (De Mas, Jansen, & Tsao, 1988; Schilling, Ciccone, Sieber, & Schmid, 2020) as well as for an heteroexopolysaccharide termed paenan with outstanding material properties in surfactant systems (Rütering et al., 2018). There have been multiple publications focusing on the rheology of heteroexopolysaccharides produced by *P. polymyxa* (Kahng, Lim, Yun, & Seo, 2001; Madden, Dea, & Steer, 1986; Ninomiya & Kizaki, 1969; Raza, Makeen, Wang, Xu, & Qirong, 2011). Even though monomer analysis typically identified glucose, mannose, galactose and sometimes fucose and pyruvate to be part of the EPS, the available rheological data can be hardly compared as these studies differ with respect to the strains as experimental setups for polymer production and rheological characterization. Even more important, recent advances in the elucidation of the genetics of Paenibacilli and the underlying polymer structure revealed that the heteroexopolysaccharide “paenan” is in fact composed of three polysaccharides with distinct substructures, which are simultaneously produced by *P. polymyxa* (Rütering et al., 2018, Schilling et al., 2021a, 2021b). The biosynthesis of the heteroexopolysaccharides from *P. polymyxa* DSM 365 following the Wzx/Wzy pathway is encoded in a 35 kb gene cluster, encoding eleven glycosyltransferases, two Wzx flippases and two Wzy polymerases as well as multiple genes associated to regulatory functions and precursor synthesis (Rütering et al., 2017). A secondary cluster encoding the pyruvyltransferase EpsO, another putative flippase and two additional glycosyltransferases correspond to the biosynthesis of paenan I. Using a combinatorial approach in performing CRISPR-Cas9 mediated gene deletions of the paenan cluster in *P. polymyxa*, it was possible to isolate the individual polymers and elucidate their structures (Fig. 1). Furthermore, by this approach, distinct glycosyltransferases were attributed to the biosynthesis of individual polymers establishing a structure-function relationship (Schillinget al., 2022a, 2022b).

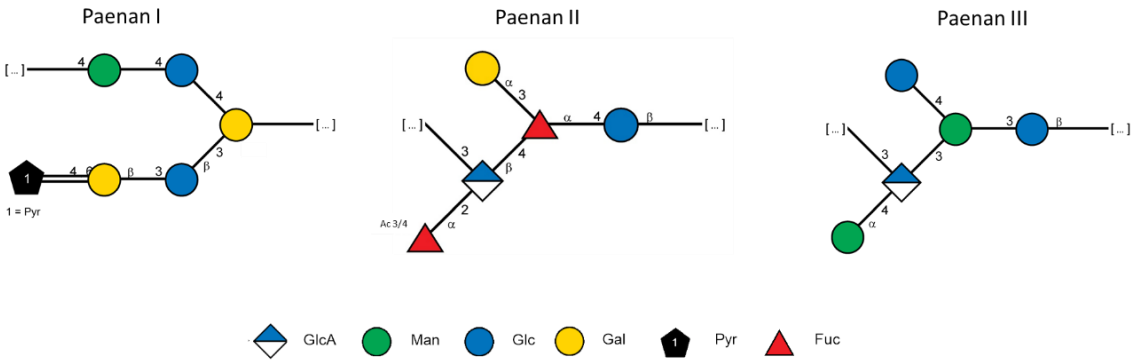


Fig. 1: Overview of the three heteroexopolysaccharides produced by *P. polymyxa* DSM 365. The repeating units of paenan I and paenan III are postulated structures based on current carbohydrate analysis and require further verification by NMR. Figure adapted from (Schilling et al., 2022a, 2022b).

For the first time, this study focuses on the in-depth rheological characterization of defined paenan compositions. Leveraging CRISPR-Cas9 mediated combinatorial knock-out strains, defined polysaccharide compositions were produced in their natural ratio. The rheological behavior of each polymer is characterized individually and in combination with other paenan variants and will elude how the EPS variants interact with each other, resulting in highly different physicochemical properties.

Materials and Methods

Strains and media

P. polymyxa DSM 365 was acquired from the German Collection of Microorganisms and Cell Culture (DSMZ, Germany). Combinatorial knock-out variants were obtained as previously described (Schilling, Rühmann, et al., 2022a). Strains used in this study are listed in Table S1. All medium components were obtained from Carl Roth GmbH (Germany) if not indicated differently. All strains were stored in 30 % glycerol at -80 °C. Prior to cultivation, strains were streaked on LB agar plates and grown at 30 °C for 24 h.

Fermentation medium contained 30 g L^{-1} glucose, 0.05 g L^{-1} $CaCl_2$ x 2 H_2O, 5 g L^{-1} tryptone, 1.33 g L^{-1} $MgSO_4$ x 7 H_2O, 1.67 g L^{-1} KH_2PO_4, 2 mL L^{-1} RPMI 1640 vitamins solution (Merck, Germany) and 1 mL L^{-1} trace elements solution (2.5 g L^{-1} $FeSO_4$, 2.1 g L^{-1} $C_4H_4O_6Na_2$ x 2 H_2O, 1.8 g L^{-1} $MnCl_2$ x 4 H_2O, 0.258 g L^{-1} H_3BO_3, 0.031 g L^{-1} $CuSO_4$ x 5 H_2O, 0.023 g L^{-1} $NaMoO_4$ x 2 H_2O, 0.075 g L^{-1} $CoCl_2$ x 7 H_2O, 0.021 g L^{-1} $ZnCl_2$). Preculture medium was prepared equal to the fermentation medium except of a reduced glucose concentration of 10 g L^{-1} and additional 20 g L^{-1} MOPS buffered to pH 7.

Fermentative EPS production.

Fermentative production of EPS was performed in 2 L Sartorius Biostat B plus bioreactor systems (Sartorius, Germany) with a working volume of 1 L equipped with two 6-blade Rushton impellers over 28 h at a controlled pH of 6.8 and pO_2 saturation of 30 %. Batch

cultivations were started with an initial OD_{600} of 0.1 by inoculation with an appropriate volume of preculture. After fermentation, biomass was separated by centrifugation (15,000 x g, 20 °C, 20 min) followed by cross-flow filtration of the supernatant using 100 kDa filtration cassette (Hydrosart, Sartorius AG, Germany). Highly viscous EPS variants were diluted 1:10 with ddH_2O prior to centrifugation. Concentrated supernatant was afterwards slowly poured into two volumes of isopropanol. Precipitated EPS was then collected and dried overnight in a VDL53 vacuum oven at 40 °C (Binder, Germany). Dry weight of the obtained EPS was determined gravimetrically, before milling to a fine powder in a ball mill at 30 Hz for 1 min (Mixer Mill MM400, Retsch GmbH, Germany).

Carbohydrate fingerprinting

Monomer compositions of all EPS variants were analyzed by the 1-phenyl-3-methyl-5-pyrazolone-high throughput method (HT-PMP) (Rühmann, Schmid, & Sieber, 2014). In short, 0.1 % EPS solutions were hydrolyzed in a 96-well plate, sealed with a silicone mat and further covered by a custom-made metal device with 2 M trifluoroacetic acid (90 min, 121 °C). Samples were neutralized with 3.2 % NH_4OH. 75 µL of PMP master mix (0.1 M methanolic PMP:0.4 % ammonium hydroxide 2:1) were added to 25 µL of neutralized hydrolysate and incubated at 70 °C for 100 min in a thermal cycler. 20 µL of derivatized samples were mixed with 25 µL 0.5 M acetic acid and 125 µL ddH_2O and filtered with a 0.2 µm filter plate (1,000 x g, 2 min) followed by HPLC-UV-MS using a Ultimate 3000 RS HPLC system (Dionex, USA). Separation was performed on a reverse phase column (Gravity C18, 100 x 2 mm, 1.8 µm particle size, Macherey-Nagel, USA) set to 50 °C. Gradient elution was performed using a mobile phase A (5 mM ammonium acetate adjusted to pH 5.6 with 15 % acetonitrile) and mobile phase B (100 % acetonitrile) with a constant pump rate of 0.6 mL min^{-1}.

Enzymatic assays for glucose and pyruvate quantification

Glucose and pyruvate concentrations of obtained EPS samples were determined enzymatically before and after hydrolysis to deduct residual contaminations from the fermentation medium as previously described (Rühmann, Schmid, & Sieber, 2016).

Molecular weight

The molecular weight of polymer variants was determined via size exclusion chromatography using an Agilent 1260 Infinity system (Agilent Technologies, Germany) equipped with a refractive index detector (SECcurity GPC1260) and a SECcurity SLD7000 seven-angle static light scattering detector (PSS Polymer Standards Service, Germany). For this, 0.5 g L^{-1} of each variant was reconstituted in 0.1 M $LiNO_3$ and 100 µL sample were injected to the system in 30 min intervals and analyzed using a TSKgel SuperMP(PW)-H guard column and two consecutive TSKGel SuperMultipore PW-H columns (6.0 mm ID x 15 cm, TOSOH Bioscience, Germany) kept at 50 °C. As an eluent 0.1 M $LiNO_3$ was used at a constant flow rate of 0.3 mL min^{-1}. Absolut molecular weight was determined via light scatter and polymer concentration and further cross-validated

using a 12 point-pullulan standard (384 Da – 2.35 MDa) and a 4.5 MDa xanthan reference.

Rheological analysis

For rheological analysis, 1 % (w/w) solutions of each polymer were prepared in ddH_2O and 0.5% NaCl (85 mM) respectively. Conductivity of each solution was measured using an LF413T-ID electrode (Schott instruments, Germany) to determine residual salt concentrations of the fermentation broth (Table S2). Rheological measurements were conducted using a MCR 300 stress controlled rotational rheometer (Anton Paar, Austria) equipped with a CP 50-1 cone-plate measuring system (50 mm diameter, 1° cone angle, 50 µm cone truncation). All measurements, except temperature sweeps, were carried out at 20°C controlled by a TEK 150P temperature unit. After applying of the solution to the rheometer, all samples were incubated at 20 °C for 5 minutes before beginning the measurements. All experiments were performed in technical triplicates.

Viscosity curves

Viscosity curves were measured using a logarithmically increased shear rate from 10^{-3} to 10^3 s^{-1} by measuring 3 data points per decade with decreasing measuring time of 100 – 5 s per data point.

Amplitude sweeps

Amplitude sweeps were measured using a logarithmically increasing shear stress amplitude from 10^{-1} to 10^3 Pa with a frequency of 1 Hz.

Frequency sweeps

Frequency sweeps were carried out within the linear viscoelastic range (LVE) at a logarithmically increasing frequency from 10^{-2} to 10 Hz.

Temperature sweeps

Temperature sweeps were performed within the LVE at a frequency of 1 Hz applying a temperature ramp from 20 to 75 °C with a heating rate of 4 °C min^{-1}. The edge of the cone-plate measuring system was covered with low viscosity paraffin oil (Carl Roth, Germany) to prevent evaporation.

Thixotropy test

Thixotropic behavior was evaluated by a three-stage oscillatory shear sequence. In the first stage, samples were subjected to shear stress within the LVE region followed by a high oscillatory shear of 10^3 Pa for 30 s. The structural recovery was then measured over 10 min within the LVE.

Results and discussion

Polymer characterization

In a first step, the EPS of *P. polymyxa* DSM 365 and combinatorial knock-out mutants were produced in bioreactor scale. During downstream processing, the fermentation supernatant was desalted via cross-flow filtration to enable uniform characterization of the rheological behavior. Due to the anionic character of all paenan variants, a residual conductivity between 0.7-1.0 mS cm^{-1} was determined for all 1 % polymer solutions (Table S2). In order to verify the presence of individual paenan polymers, the carbohydrate fingerprint was determined for each variant (Fig. 2). For paenan I, a monomer ratio of 2:1:2:1 (Glc:Man:Gal:Pyr) was expected, while for paenan II an equimolar ratio of 1:1:1:2 (Glc:Gal:GlcA:Fuc) and for paenan III, 2:2:1 (Glc:Man:GlcA) were anticipated. Due to the high susceptibility of uronic acids towards degradation during chemical hydrolysis, GlcA was severely underestimated for all polysaccharide compositions containing the uronic acid (Rodén, Baker, Cifonelli, & Mathews, 1972). In addition to the monomer composition, previously identified key dimers detected via MS/MS analysis were used to assign combinatorial knock-out variants to each paenan variant (Schilling et al., 2022b). By assigning the presence of pyruvate ketals to paenan I, a GlcA-Fuc dimer to paenan II and a GlcA-Man dimer to paenan III, the polymer composition of each knock-out variant was determined, while still retaining the natural polysaccharide distribution in those variants. Henceforward, each EPS variant is named by the present paenan variants, rather than the gene deletions leading to the respective phenotype.

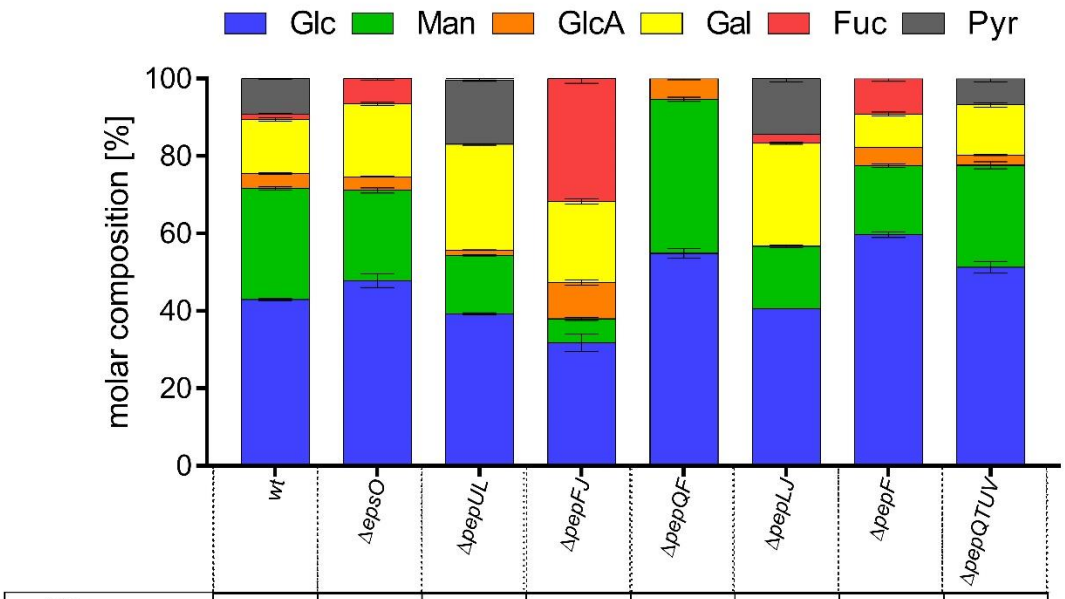


Fig. 2: Carbohydrate fingerprints of EPS obtained from *P. polymyxa* DSM 365 and mutant variants were analyzed via the HT-PMP method. Based on the obtained monomer ratios and the detection of specific dimers in MS analysis, polymer compositions were assigned to the presence of distinct paenan variants (I-III). * The *ΔepsO* variant demonstrated a similar monomer composition as the wildtype polymer, however, neither pyruvate nor the respective ketal were detected anymore.

The molecular weight of each variant was determined by size-exclusion chromatography (Table 1). Despite the presence of three distinct polymers in the wildtype EPS, no clear separation of individual paenan variants was possible (Figure S1). Analysis of individual paenan variants revealed a similar molecular weight distribution for paenan I and paenan III. Only paenan II seems to be significantly smaller with a size of $5.5 \cdot 10^5$ Da. Considering the low proportion of paenan II in the wildtype polymer, this might explain why previous attempts to analyse the heteroexopolysaccharide of *P. polymyxa* DSM 365 were not able to distinguish between multiple paenan variants (Madden et al., 1986; Rütering et al., 2017). Interestingly, while the depyruvylated polymer also showed a minor peak at approximately $3.0 \cdot 10^6$ Da, the main molar mass was detected at a considerable larger size of $8.8 \cdot 10^6$ Da compared to all other paenan variants. In comparison to the production of xanthan, for which side chains are irregularly decorated by acetyl- or pyruvyl residues, all repeating units of paenan I seem to be modified with a pyruvate ketal (Callet, Milas, & Rinaudo, 1987; Sandford et al., 1977). Consequently, the loss of this decoration in the Δ*epsO* knock-out variant might affect chain length control in *P. polymyxa,* resulting in an increased molecular weight and different rheological behavior. Alternatively, pyruvylation can also affect the hydrodynamic radius of the polymer and consequently influence SEC-

MALS analysis (Baumgartner, Pavli, & Kristl, 2008; Smolka & Belmonte, 2006).

Table 1: Calculated Mw of paenan variants obtained by GPC analysis using 0.5 % EPS solutions in 0.1 M LiNO3.

Paenan variant	Molecular weight
I & II & III	$2.6 \cdot 10^6$ Da
I & II & III depyruvylated	$8.8 \cdot 10^6$ Da
I	$5.8 \cdot 10^6$ Da
II	$5.5 \cdot 10^5$ Da
III	$3.9 \cdot 10^6$ Da
I & II	$2.8 \cdot 10^6$ Da
I & III	$2.8 \cdot 10^6$ Da
II & III	$2.7 \cdot 10^6$ Da
Xanthan reference	$4.5 \cdot 10^6$ Da

Rheological characteristics

Flow behavior of paenan-variants

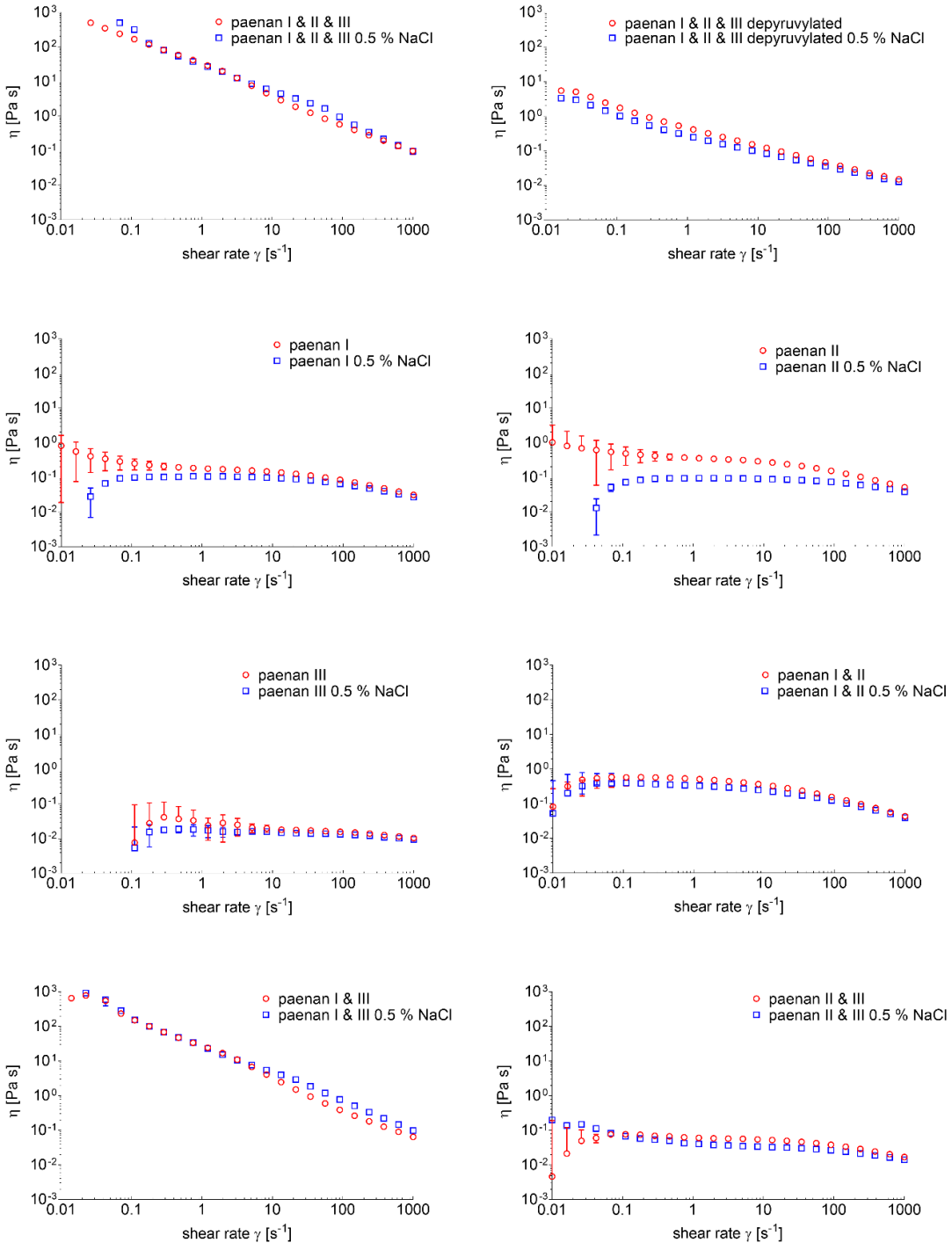


Fig. 3: Viscosity curves of paenan variants measured at logarithmically increasing shear rates from 0.01 to 1000 s^{-1} with a cone-and-plate geometry at 20 °C with (blue squares □) and without (red circles ○) the addition of 0.5 % NaCl. All measurements were carried out in triplicates, error bars show the standard deviation.

Investigations of the flow behavior showed a general shear thinning behavior of all paenan variants (Fig. 3). The highest viscosities and highest shear thinning behavior was

observed for paenan wt (I & II & III) with the addition of NaCl, followed by paenan wt without monovalent cations. All individual variants of paenan I-III did not show significantly high viscosities and almost Newtonian flow behavior in the intermediated shear rate regions. Only the paenan wt as well as paenan I & III showed strictly power-law shear thinning behavior and high viscosities, which increased in the presence of 0.5 % NaCl. Upon depyruvylation of paenan wt, the viscosity drastically decreased and the addition of NaCl now had the adverse effect, comparable to the depyruvylation of xanthan (Gansbiller, Schmid, & Sieber, 2019). For xanthan, the effect is due to lower cation mediated intermolecular interactions between individual polymer molecules. The data of the different paenan polymers suggest a similar effect between the paenan I and paenan III molecules. The combination of paenan I & III showed almost the same properties as the combination of paenan I & II & III, underlining the interaction between paenan I and III as the main origin of the high viscosity of the polymers. This interaction may be explained by the cation mediated interaction of the terminal pyruvate residue of paenan I with the negative charge of the glucuronic acid in the backbone of paenan III. Interestingly, both combinations of paenan I & II and paenan II & III did not result in the behavior of the wildtype. Consequently, the glucuronic acid in the backbone of paenan II did not seem to interact with the pyruvyl residue of paenan I as seen by the low viscosity of this combination. This is particularly interesting, as the single side chain of paenan II would suggest higher accessibility of the charge in paenan II compared to the two side-chains of paenan III, linked to the glucuronic acid and the adjacent mannose (Fig. 1). There are two explanations for the interactions of the individual molecules. First, the formation of individual homohelices of paenan I & II & III molecules respectively and a subsequent interaction of these helices. Second, the formation of heterohelices of paenan I & II & III. Investigation of the individual polymers provides evidence for the latter. The formation of homohelices would also suggest strong interactions between the ones of paenan I alone, if, as suggested the pyruvylated side chain was protruding outwards, like in xanthan. On the other hand, an inwards-protruding side chain, if paenan I was forming homohelices, would result in a similar structure as it is described for diutan gum (Xu, Gong, Dong, & Li, 2015) and might also be the reason for decreased viscosity. As opposed to diutan gum, the main chain does not carry a negative charge due to the lack of an uronic acid in the backbone of paenan I. In all cases, differences in the helical arrangements of paenan II and paenan III could explain the strong interactions of paenan I & III but not between paenan I & II as well as between paenan II & III, giving rise to further investigations of the secondary and tertiary polymer structures.

A detailed investigation of the individual polymer variants and the combination of paenan II & III showed multiple shear thinning regions, indicated by up to three different K and n values of the power law fits of the individual sections (Table 2). This phenomenon was evident for paenan I and paenan II individually as well as for combinations of paenan I & II and paenan I & III. However, for paenan III, paenan I & III or the wildtype composition,

only a single shear thinning region was observed. In case of paenan I & II, the Newtonian region is more pronounced in the presence of NaCl, which could be the reason of the slight onset of a Newtonian region in paenan wt and paenan I & III in the presence of NaCl. Consequently, this effect might be attributed to paenan I and paenan II respectively, which was previously discussed to be caused by shear-banding (Gansbiller, Schmid, & Sieber, 2020; Rütering et al., 2018).

Table 2: Model parameters (K, n) of the power law fits of the different paenan combinations with and without the addition of 0.5 % NaCl. If multiple sections were fitted individually, K and n values of the individual section are shown in ascending order of the corresponding shear rate-region.

Paenan variant	Solution	K	n
wt (I & II & III)	aq.	29.02	0.152
	0.5 % NaCl	34.38	0.171
wt (I & II & III) depyruvylated	aq.	0.47	0.485
	0.5 % NaCl	0.29	0.534
I & II	aq.	0.01	0.830
	0.5 % NaCl	0.01	0.850
I & III	aq.	24.80	0.114
	0.5 % NaCl	27.41	0.203
II & III	aq.	0.06 /0.16	0.924/0.679
	0.5 % NaCl	0.03/0.05/0.10	0.677/0.890/0.718
I	aq.	0.05/0.18/0.52	0.390/0.898/0.597
	0.5 % NaCl	0.11/0.255	1*/0.685
II	aq.	0.20/0.37/1.24	0.649/0.895/0.542
	0.5 % NaCl	0.10/0.26	0.978/0.73
III	aq.	0.03	0.86
	0.5 % NaCl	0.02	0.917

*indicating Newtonian flow behavior

Viscoelastic properties

The basic viscoelastic properties determined by the amplitude sweep (Figure S2-3) are shown in Table 3. Paenan wt (I & II & III) displayed a soft and elastic gel-like behavior

with a yield point of 9.1 Pa (32 % strain) and a flow point of 51.1 Pa (550 % strain) with a damping factor of 0.3 within the LVE. The addition of 0.5 % NaCl resulted in an increased yield and flow point of 32.3 Pa and 90.8 Pa respectively, corresponding to a strain of 82 % and 590 %. Damping factor of 0.1 and the distinct G'' peak following the LVE indicates a stronger but more brittle gel character. Frequency sweeps (Figure S4) also revealed a viscoelastic fluid like behavior of G' and G'' for paenan I & II & III with predominant elastic behavior throughout the investigated frequency range, which shifted more towards a gel-like behavior upon the addition of NaCl with lower frequency dependency of both G' and G''. Both, with and without addition of NaCl, no crossover point was evident at low frequencies, indicating long term stability of the network.

This gel character is most likely caused by the cation mediated interactions between the pyruvyl residues of paenan I and the -COO^- group of the glucuronic acid of paenan III. Further proof for this was provided by the depyruvylation of paenan I & II & III, which resulted in a complete loss of the viscoelastic properties. Furthermore, all individual paenan polymers as well as the mixtures of paenan I & II and paenan II & III showed predominant fluid properties, with Maxwell-like behavior (Figure S4-5). Interestingly, except for paenan II in the presence of NaCl, no Maxwell-fluid-typical crossover point at higher frequencies could be observed and data suggested an onsetting decrease of G' at higher frequencies, resulting in fluid behavior at both low and high frequencies. This was especially evident for paenan I and paenan III.

The high viscosity and the pronounced intermolecular network resulting in a gel-like character make this polymer variant an interesting compound as a rheological thickening agent. Similar to other microbial polysaccharides, potential applications as rheology modifiers in food and beverages but also technical applications such as oil drilling seem promising (Schmid, Sieber, & Rehm, 2015). Compared to those polysaccharides, viscosifying effects are highly increased, suggesting that lower EPS concentrations are required to obtain similar results. Furthermore, the structurally related polysaccharide from *P. polymyxa* 2H2 has recently demonstrated excellent compatibility with commonly used surfactants such as lauryl sulfate or cocamidopropyl betaine, which are typically used cosmetics and daily care products (Rütering et al., 2018). Consequently, we propose the utilization of the wildtype EPS composition of *P. polymyxa* DSM 365, containing paenan I & II & III as a sustainable thickening agent for variable applications that could replace commercially available petroleum-based acrylic compounds (Kim, Song, Lee, & Park, 2003; Tafuro, Costantini, Baratto, Busata, & Semenzato, 2019).

Table 3: Viscoelastic properties of the paenan polymer variants. n.d.: not determined if measurement was not possible

Paenan variant	Solution	G' [Pa] (LVE)	tanδ (LVE)	Yield point [Pa]	Flow point [Pa]
wt (I & II & III)	aq.	28.5	0.3	9.1	51.1
	0.5 % NaCl	45.5	0.1	32.3	90.8
wt (I & II & III) depyruvylated	aq.	0.71	1.0	n.d.	n.d.
	0.5 % NaCl	n.d.	1.0	n.d.	n.d.
I	aq.	0.15	5.1	n.d.	n.d.
	0.5 % NaCl	0.10	6.7	n.d.	n.d.
II	aq.	0.60	3.0	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.	n.d.
III	aq.	n.d.	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.	n.d.
I & II	aq.	0.78	2.6	n.d.	n.d.
	0.5 % NaCl	0.40	3.6	n.d.	n.d.
I & III	aq.	13.9	0.25	11	40
	0.5 % NaCl	35.8	0.1	18	79
II & III	aq.	n.d.	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.	n.d.

The paenan I & III mixture showed gel-like properties very similar to the ones of paenan I & II & III, indicating that the interaction is mainly between paenan I and paenan III. However, compared to paenan I & II & III, paenan I & III showed lower gel-strength with yield points at 11 Pa and 18 Pa with and without the presence of 0.5 % NaCl, and a less pronounced G'' peak at the end of the LVE region in the presence of NaCl. This indicates weaker interactions of these polymers. Investigations of the amplitude sweeps of the individual polymers showed that paenan I and II both show viscoelastic fluid like behavior, while paenan III only shows strictly liquid behavior. Without the formation of gel-like networks, the addition of NaCl resulted in a decrease of both G' and G'', while paenan I exhibited higher salt stability compared to paenan II. This also becomes evident in the mixture of paenan I & II, where the effect of NaCl is more comparable to paenan I than paenan II. These effects point toward an interaction between paenan I and II, which

might be responsible for the increased gel strength in paenan I & II & III compared to paenan I & III. As both paenan II and paenan III have a glucuronic acid in the backbone, the strong interaction of paenan I & III suggests a better accessibility of the glucuronic acid of paenan III compared to the one of paenan II. Contrarily, interactions between paenan II and paenan III might lead to a different structural arrangement of these polymers, resulting in better accessibility of the glucuronic acid in paenan II and therefore to increased interactions between paenan I & II in the wt polymer mixture.

In contrast to the native polysaccharide composition containing paenan I & II & III, deletion of individual polymers resulted in significantly altered viscoelastic properties. While the combination of paenan I & III still demonstrated a pronounced intermolecular network resulting in a gel like character, individual biopolymers demonstrated fluid like behavior, which still form films when dried. Consequently, significantly different applications to the wildtype EPS composition arise. On the one hand, bulk application for the formation of edible films and packaging materials similar to pullulan seem practical (Diab, Biliaderis, Gerasopoulos, & Sfakiotakis, 2001). On the other hand, high value biomedical applications as coating materials in pharmaceutical controlled drug release systems should be further investigated (Miao, Wang, Zeng, Liu, & Chen, 2018). For other charged polysaccharides such as hyaluronic acid and alginates, the chemical modification of the functional groups improved targeting of specific cell types allowing effective drug delivery systems (Bhattacharya et al., 2017; Pawar & Edgar, 2012). Furthermore, polysaccharides produced by other strains of *P. polymyxa* have demonstrated anti-oxidant activities, which might further enhance pharmacological applications (Liu et al., 2010; Raza et al., 2011).

Temperature stability and thixotropic properties

Temperature sweeps showed high temperature dependency of the viscoelastic properties of paenan I & II & III both with and without the addition of NaCl (Figure S6-7). The viscoelastic properties of paenan I & III showed even higher temperature dependency, which could be reduced by the addition of NaCl. While the native polysaccharide composition containing all three paenan variants retained a weak gel character up to 75 °C in the presence of NaCl, deletion of paenan II resulted in a loss of temperature stability. This suggests a stabilizing effect of paenan II regarding to temperature stability of the polymer network, which is also evident by the high temperature stability of paenan II compared to paenan I & III. For paenan II, an increase of G’ and G’’ could be observed during the heating ramp, which was further pronounced by the addition of NaCl. This effect is similar to the one seen in the undecorated xanthan variant previously described (Gansbiller et al., 2019) and might also be explained by structural rearrangements of the single polymer. However, these effects did not occur in any other combination of paenan II with other paenan polymers, suggesting a different arrangement of the individual polymers in mixture, which underlines the previously

described differences between paenan I & II & III and paenan I & III in regards to their viscoelastic properties.

Table 4: Temperature stability of different paenan variants. n.d.: not determined if measurement was not possible within the LVE range

Paenan Variant	Solution	G' at 20°C	G' at 75°C	relative Gelstrength at 75°C [%]	tanδ at 75°C
wt (I & II & III)	aq.	29.43	1.15	3.91	1.49
	0.5 % NaCl	51.97	5.01	9.65	0.83
wt (I & II & III) depyruvylated	aq.	0.40	0.00	0.15	n.d.
	0.5 % NaCl	0.18	0.00	0.13	n.d.
I	aq.	0.17	0.01	6.19	22.29
	0.5 % NaCl	0.07	0.00	0.00	n.d.
II	aq.	0.269	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.	n.d.
III	aq.	0.139	n.d.	n.d.	n.d.
	0.5 % NaCl	0.139	n.d.	n.d.	n.d.
I & II	aq.	0.31	0.01	4.20	28.02
	0.5 % NaCl	0.08	0.00	0.00	n.d.
I & III	aq.	12.20	0.00	0.00	n.d.
	0.5 % NaCl	35.03	0.07	0.20	35.90
II & III	aq.	0.78	0.20	26.05	1.25
	0.5 % NaCl	0.78	0.12	15.71	1.23

Furthermore, thixotropic properties were determined by a three stage oscillatory shear stress test (Table S3). While structural recovery was observed for all combinatorial variants of paenan, only 86.8 % of the initial gel strength was measured after three minutes with a non-destructive shear stress for the wildtype EPS mixture. This further underlines a pronounced intermolecular network, which takes more time to recover and coordinate non-covalent interactions between individual polymers. Similar effects of a retarded structural recovery were observed for polysaccharide composition containing paenan I & III, consolidating the hypothesis that the gel-like character mainly originates from the interaction between the pyruvate of paenan I and the glucuronic acid residue of

paenan III. Contrarily, for all other knock-out variants, immediate structural recovery resulting in the initial gel strength was observed. Consequently, distinct variants might be applicable as binders that convey thixotropic behavior typically used for varnishes and coatings with divergent rheological profiles.

Conclusion

In this study, we have characterized the rheological behavior of the heteroexopolysaccharides produced by *P. polymyxa* DSM 365, using CRISPR-Cas9 mediated knock-outs of glycosyltransferases. Viscoelastic properties of individual paenan variants and combinations thereof were analyzed in detail. While the wildtype EPS composition demonstrated high viscosity and a gel-like behavior, knock-out variants showed significantly altered physicochemical properties depending on the present individual polysaccharides. Consequently, distinct polysaccharide compositions might be utilized for a wide range of applications such as thickening agents or coating materials. We propose specific intra- but also intermolecular network formations that seem to be highly affected by the presence of distinct pyruvyl- and glucuronic acid residues. Additional analytical approaches such as atomic force microscopy might be necessary to further investigate the precise interaction of individual biopolymers and putative secondary and tertiary structures.

Acknowledgements

The authors would like to acknowledge the technical support from Tristan Rath for SEC-MALS analysis. This work was supported by the German Federal Ministry of Education and Research (BMBF) in frame of the project MaPolKo (number 03VP02560).

References

Bajaj, I., Survase, S., Saudagar, P., & Singhal, R. (2007). Gellan Gum: Fermentative Production, Downstream Processing and Applications. Food Technology and Biotechnology, 45.

Baumgartner, S., Pavli, M., & Kristl, J. (2008). Effect of calcium ions on the gelling and drug release characteristics of xanthan matrix tablets. European Journal of Pharmaceutics and Biopharmaceutics, 69(2), 698–707.

Becker, A., Katzen, F., Pühler, A., & Ielpi, L. (1998). Xanthan gum biosynthesis and application: A biochemical /genetic perspective. Applied Microbiology and Biotechnology, 50(2), 145–152.

Bhattacharya, D. S., Svechkarev, D., Souchek, J. J., Hill, T. K., Taylor, M. A., Natarajan, A., & Mohs, A. M. (2017). Impact of structurally modifying hyaluronic acid on CD44 interaction. J. Mater. Chem. B, 5(41), 8183–8192.

Burdick, J. A., & Prestwich, G. D. (2011). Hyaluronic Acid Hydrogels for Biomedical Applications. Advanced Materials, 23(12), H41–H56.

Callet, F., Milas, M., & Rinaudo, M. (1987). Influence of acetyl and pyruvate contents on rheological properties of xanthan in dilute solution. International Journal of Biological

Macromolecules, 9(5), 291–293.
Danese, P. N., Pratt, L. A., & Kolter, R. (2000). Exopolysaccharide Production Is Required for Development of *Escherichia coli* K-12 Biofilm Architecture. Journal of Bacteriology, 182(12), 3593–3596.
De Mas, C., Jansen, N. B., & Tsao, G. T. (1988). Production of optically active 2,3-butanediol by *Bacillus polymyxa*. Biotechnology and Bioengineering, 31(4), 366–377.
Diab, T., Biliaderis, C. G., Gerasopoulos, D., & Sfakiotakis, E. (2001). Physicochemical properties and application of pullulan edible films and coatings in fruit preservation. Journal of the Science of Food and Agriculture, 81(10), 988–1000.
Freitas, F., Alves, V. D., & Reis, M. A. M. (2011). Advances in bacterial exopolysaccharides: From production to biotechnological applications. Trends in Biotechnology, 29(8), 388–398.
Gansbiller, M., Schmid, J., & Sieber, V. (2019). In-depth rheological characterization of genetically modified xanthan-variants. Carbohydrate Polymers, 213, 236–246.
Gansbiller, M., Schmid, J., & Sieber, V. (2020). Rheology of sphingans in EPS–surfactant systems. Carbohydrate Polymers, 248, 116778.
Grady, E. N., MacDonald, J., Liu, L., Richman, A., & Yuan, Z.-C. (2016). Current knowledge and perspectives of *Paenibacillus*: A review. Microbial Cell Factories, 15(1), 203.
Hassler, R. A., & Doherty, D. H. (1990). Genetic engineering of polysaccharide structure: Production of variants of xanthan gum in *Xanthomonas campestris*. Biotechnology Progress, 6(3), 182–187.
Jang, H. Y., Zhang, K., Chon, B. H., & Choi, H. J. (2015). Enhanced oil recovery performance and viscosity characteristics of polysaccharide xanthan gum solution. Journal of Industrial and Engineering Chemistry, 21, 741–745.
Kahng, G.-G., Lim, S.-H., Yun, H.-D., & Seo, W.-T. (2001). Production of extracellular polysaccharide, EPS WN9, from *Paenibacillus* sp. WN9 KCTC 8951P and its usefulness as a cement mortar admixture. Biotechnology and Bioprocess Engineering, 6(2), 112–116.
Kaur, V., Bera, M. B., Panesar, P. S., Kumar, H., & Kennedy, J. F. (2014). Welan gum: Microbial production, characterization, and applications. International Journal of Biological Macromolecules, 65, 454–461.
Kim, J.-Y., Song, J.-Y., Lee, E.-J., & Park, S.-K. (2003). Rheological properties and microstructures of Carbopol gel network system. Colloid & Polymer Science, 281(7), 614–623.
Kumar, A., Rao, K. M., & Han, S. S. (2018). Application of xanthan gum as polysaccharide in tissue engineering: A review. Carbohydrate Polymers, 180, 128–144.
Liu, J., Luo, J., Ye, H., Sun, Y., Lu, Z., & Zeng, X. (2010). In vitro and in vivo antioxidant activity of exopolysaccharides from endophytic bacterium *Paenibacillus polymyxa* EJS-3. Carbohydrate Polymers, 82(4), 1278–1283.
Madden, J. K., Dea, I. C. M., & Steer, D. C. (1986). Structural and rheological properties of the extracellular polysaccharides from *Bacillus polymyxa*. Carbohydrate Polymers, 6(1), 51–73.
Miao, T., Wang, J., Zeng, Y., Liu, G., & Chen, X. (2018). Polysaccharide-Based Controlled Release Systems for Therapeutics Delivery and Tissue Engineering: From Bench to Bedside. Advanced Science, 5(4), 1700513.
Moradali, M. F., & Rehm, B. H. A. (2020). Bacterial biopolymers: From pathogenesis to advanced materials. Nature Reviews Microbiology. Retrieved February 18, 2020, from http://www.nature.com/articles/s41579-019-0313-3
Ninomiya, E., & Kizaki, T. (1969). Bacterial polysaccharide from *Bacillus polymyxa* No. 271. Die Angewandte Makromolekulare Chemie, 6(1), 179–185.

Ophir, T., & Gutnick, D. L. (1994). A Role for Exopolysaccharides in the Protection of Microorganisms from Desiccation. Applied and Environmental Microbiology, 60(2), 740.

Pawar, S. N., & Edgar, K. J. (2012). Alginate derivatization: A review of chemistry, properties and applications. Biomaterials, 33(11), 3279–3305.

Poli, A., Anzelmo, G., & Nicolaus, B. (2010). Bacterial Exopolysaccharides from Extreme Marine Habitats: Production, Characterization and Biological Activities. Marine Drugs, 8(6), 1779–1802.

Raza, W., Makeen, K., Wang, Y., Xu, Y., & Qirong, S. (2011). Optimization, purification, characterization and antioxidant activity of an extracellular polysaccharide produced by *Paenibacillus polymyxa* SQR-21. Bioresource Technology, 102(10), 6095–6103.

Rodén, L., Baker, J. R., Cifonelli, J., & Mathews, M. B. (1972). [7] Isolation and characterization of connective tissue polysaccharides. Methods in Enzymology (Vol. 28, pp. 73–140). Elsevier. Retrieved April 3, 2020, from https://linkinghub.elsevier.com/retrieve/pii/0076687972280090

Rühmann, B., Schmid, J., & Sieber, V. (2014). Fast carbohydrate analysis via liquid chromatography coupled with ultra violet and electrospray ionization ion trap detection in 96-well format. Journal of Chromatography A, 1350, 44–50.

Rühmann, B., Schmid, J., & Sieber, V. (2016). Automated Modular High Throughput Exopolysaccharide Screening Platform Coupled with Highly Sensitive Carbohydrate Fingerprint Analysis. Journal of Visualized Experiments, (110), 53249.

Rütering, M., Cress, B. F., Schilling, M., Rühmann, B., Koffas, M. A. G., Sieber, V., & Schmid, J. (2017). Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. Synthetic Biology, 2(1). Retrieved July 1, 2019, from http://academic.oup.com/synbio/article/doi/10.1093/synbio/ysx007/4772606

Rütering, M., Schmid, J., Gansbiller, M., Braun, A., Kleinen, J., Schilling, M., & Sieber, V. (2018). Rheological characterization of the exopolysaccharide Paenan in surfactant systems. Carbohydrate Polymers, 181, 719–726.

Sandford, P. A., Pittsley, J. E., Knutson, C. A., Watson, P. R., Cadmus, M. C., & Jeanes, A. (1977). Variation in Xanthomonas campestris NRRL B-1459: Characterization of Xanthan Products of Differing Pyruvic Acid Content. In Paul A. Sandford & A. Laskin (Eds.), Extracellular Microbial Polysaccharides, ACS Symposium Series (Vol. 45, pp. 192–210). WASHINGTON, D. C.: AMERICAN CHEMICAL SOCIETY. Retrieved October 13, 2020, from https://pubs.acs.org/doi/abs/10.1021/bk-1977-0045.ch015

Schilling, C., Badri, A., Sieber, V., Koffas, M., & Schmid, J. (2020). Metabolic engineering for production of functional polysaccharides. Current Opinion in Biotechnology, 66, 44–51.

Schilling, C., Ciccone, R., Sieber, V., & Schmid, J. (2020). Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365. Metabolic Engineering, 61, 381–388.

Schilling, C., Klau, L.J., Rühmann, B., Aachmann, F.L., Schmid, J., & Sieber, V. (2022a). Structural elucidation of the fucose containing exopolysaccharide from *Paenibacillus polymyxa* DSM 365. Carbohydrate Polymers, 278

Schilling, C., Klau, L.J., Rühmann, B., Aachmann, F.L., Schmid, J., & Sieber, V. (2022b). CRISPR-Cas9 driven structural elucidation of the heteroexopolysaccharides from *Paenibacillus polymyxa* DSM 365. Prepared Manuscript.

Schmid, J., Sieber, V., & Rehm, B. (2015). Bacterial exopolysaccharides: Biosynthesis pathways and engineering strategies. Frontiers in Microbiology, 6. Retrieved July 1, 2019, from http://www.frontiersin.org/Microbiotechnology%2c_Ecotoxicology_and_Bioremediation/10.3389/fmicb.2015.00496/abstract

Smolka, L. B., & Belmonte, A. (2006). Charge screening effects on filament dynamics in

xanthan gum solutions. Journal of Non-Newtonian Fluid Mechanics, 137(1–3), 103–109.

Sutherland, I. W. (1972). Bacterial Exopolysaccharides. Advances in Microbial Physiology (Vol. 8, pp. 143–213). Elsevier. Retrieved October 12, 2020, from https://linkinghub.elsevier.com/retrieve/pii/S0065291108601903

Tafuro, G., Costantini, A., Baratto, G., Busata, L., & Semenzato, A. (2019). Rheological and Textural Characterization of Acrylic Polymer Water Dispersions for Cosmetic Use. Industrial & Engineering Chemistry Research, 58(51), 23549–23558.

Williams, P. A., & Phillips, G. O. (Eds.). (2016). Gums and stabilisers for the food industry 18: Hydrocolloid functionality for affordable and sustainable global food solutions. Special publication / Royal Society of Chemistry. Presented at the Gums and Stabilisers for the Food Industry Conference, Cambridge, UK: Royal Society of Chemistry.

Wu, M., Qu, J., Tian, X., Zhao, X., Shen, Y., Shi, Z., Chen, P., et al. (2019). Tailor-made polysaccharides containing uniformly distributed repeating units based on the xanthan gum skeleton. International Journal of Biological Macromolecules, 131, 646–653.

Xu, L., Gong, H., Dong, M., & Li, Y. (2015). Rheological properties and thickening mechanism of aqueous diutan gum solution: Effects of temperature and salts. Carbohydrate Polymers, 132, 620–629.

3.4. A broad-host range tool for CRISPRa/i driven flux control

Using CRISPR-Cas9 for genome editing has revolutionized modern biotechnology and accelerated rational strain development and metabolic engineering efforts tremendously. The use of catalytically inactive Cas proteins to repress gene expression by blocking transcription has demonstrated to be an efficient tool to precisely control metabolic fluxes [177]. Particularly in eukaryotes, dCas effectors are often linked to activator domains to enhance expression of targeted genes of interest [178]. For bacteria, first tools to enable prokaryotic CRISPRa were previously developed [154,155]. However, all of these systems used Cas9 as an effector module. Due to the required processing of its corresponding gRNA, Cas9 demonstrated many disadvantages for multiplexing approaches required for the efficient flux control of complex metabolic pathways. Therefore, in this study a novel Cas12a based broad-host range tool was designed for the simultaneous activation and repression of multiple target genes in *P. polymyxa*.

In a first step, five different activator domains from different protein families were linked to an engineered version of dCas12a and tested for their efficacy to activate the gene expression of a reporter gene. The activator of the superoxide regulon SoxS from *E. coli* showed the least stringency in the distance of the gRNA to the corresponding promoter. Leveraging the ability of Cas12a to process its own CRISPR array, multiplexing was demonstrated by simultaneously activating the expression of *gfp* and repressing *mrfp*. By targeting glycosyltransferases of the paenan cluster in the genome of *P. polymyxa*, it proofed possible to mimic EPS compositions of more laborious and time-consuming knock-out strains. Furthermore, a CRISPR array encoding eleven gRNAs targeting the mixed acid pathway and butanediol biosynthesis was used to enhance 2,3-BDL production. While transcriptional activation and repression under microaerobic conditions proofed to be more difficult, metabolic fluxes could be partially redirected from lactate to 2,3-BDL.

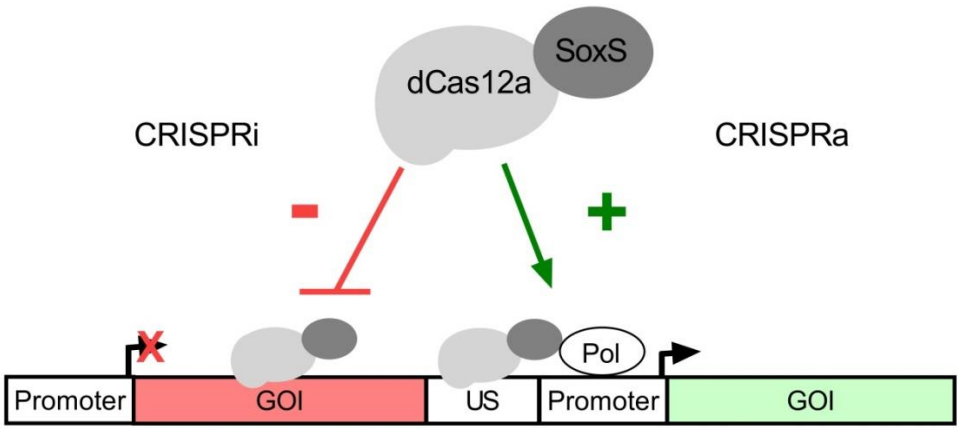
The author planned and conducted all experiments in this study. Mattheos Koffas, Volker Sieber and Jochen Schmid provided technical and scientific advice for design of experiments and execution. All authors contributed to content and language of the manuscript.

A novel prokaryotic CRISPR-Cas12a based tool for programmable transcriptional activation and repression

Christoph Schilling, Mattheos A.G. Koffas, Volker Sieber & Jochen Schmid

ACS Synthetic Biology

(2020)



DOI: 10.1021/acssynbio.0c00424

ACS SyntheticBiology

pubs.acs.org/synthbio Research Article

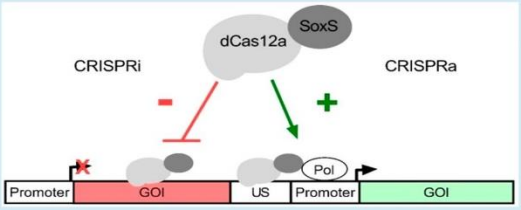
Novel Prokaryotic CRISPR-Cas12a-Based Tool for Programmable Transcriptional Activation and Repression

Christoph Schilling, Mattheos A. G. Koffas, Volker Sieber, and Jochen Schmid*

Cite This: *ACS Synth. Biol.* 2020, 9, 3353−3363 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Transcriptional perturbation using inactivated CRISPR-nucleases (dCas) is a common method in eukaryotic organisms. While rare examples of dCas9-based tools for prokaryotes have been described, multiplexing approaches are limited due to the used effector nuclease. For the first time, a dCas12a derived tool for the targeted activation and repression of genes was developed. Therefore, a previously described SoxS activator domain was linked to dCas12a to enable the programmable activation of gene expression. A proof of principle of transcriptional regulation was demonstrated on the basis of fluorescence reporter assays using the alternative host organism *Paenibacillus polymyxa* as well as *Escherichia coli*. Single target and multiplex CRISPR interference targeting the exopolysaccharide biosynthesis of *P. polymyxa* was shown to emulate polymer compositions of gene knockouts. The simultaneous expression of 11 gRNAs targeting multiple lactate dehydrogenases and a butanediol dehydrogenase resulted in decreased lactate formation, as well as an increased butanediol production in microaerobic fermentation processes. Even though Cas12a is more restricted in terms of its genomic target sequences compared to Cas9, its ability to efficiently process its own guide RNAs *in vivo* makes it a promising tool to orchestrate sophisticated genetic reprogramming of bacterial cells or to screen for engineering targets in the genome. The developed tool will accelerate metabolic engineering efforts in the alternative host organism *P. polymyxa* and might be also applied for other bacterial cell factories.



KEYWORDS: *Paenibacillus polymyxa, CRISPRa, CRISPRi, Cas12a, multiplex gene regulation*

Seeking a biobased and sustainable economy, bacterial cell factories have been used for the production of a variety of high-value products such as amino acids[1] and biofuels[2] or the biosynthesis of complex pharmaceutical compounds like artemisinic acid.[3] The development of robust production strains for industrial scale production typically requires a deep understanding of the underlying metabolic networks enabling sophisticated engineering technologies to optimize fluxes toward the product of interest and eliminating unwanted side products.[4] Within the past decade, the development of new technologies such as CRISPR-Cas9-mediated genome editing resulted in a dramatic increase in the complexity and scope of metabolic engineering approaches.[5−8]

CRISPR interference (CRISPRi), using catalytically inactive CRISPR nucleases (dCas) for transcriptional repression, has been successfully demonstrated in bacteria.[9−11] In eukaryotic organisms, CRISPRi has been further expanded by the direct fusion of Cas9 and effector domains to remodel the chromatin structure of target genes, resulting in an even tighter repression of expression.[12−14] On the contrary, CRISPR-mediated activation (CRISPRa) of gene expression was achieved by linking the deactivated CRISPR-nuclease from *Streptococcus pyogenes* dCas9 to transcriptional activation domains *via* translational fusion or recruitment *via* peptide epitopes or additional RNA scaffolds.[15−17] While CRISPRa has been extensively used and further developed for eukaryotic organisms to activate the transcription of target genes,[18] the number of synthetic tools for prokaryotes is still limited. Recently, new CRISPR-Cas9-based systems were developed for bacteria using effector domains such as RpoZ,[9] RpoA,[19,20] bacteriophage derived transcriptional activators like AsiA,[21,22] or the more effective AraC family transcription factor SoxS[23] that facilitate the recruitment of the RNA polymerase holoenzyme. In order to overcome narrow target site requirements, more flexible CRISPRa toolkits using σ^{54}-dependent promoters were established.[24] While σ^{70}-activators bind in close proximity to their cognate promoter and recruit RNA-polymerase, bacterial enhancer binding proteins (bEBPs) corresponding to σ^{54}-dependent promoters target long distance upstream activating sequences (UAS) and initiate a transcription similar to eukaryotic promoters.[24] Consequently, CRISPR-dCas-guided bEBPs were directed to UAS regions in

Received: August 9, 2020
Published: November 25, 2020

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

ACS Synthetic Biology pubs.acs.org/synthbio Research Article

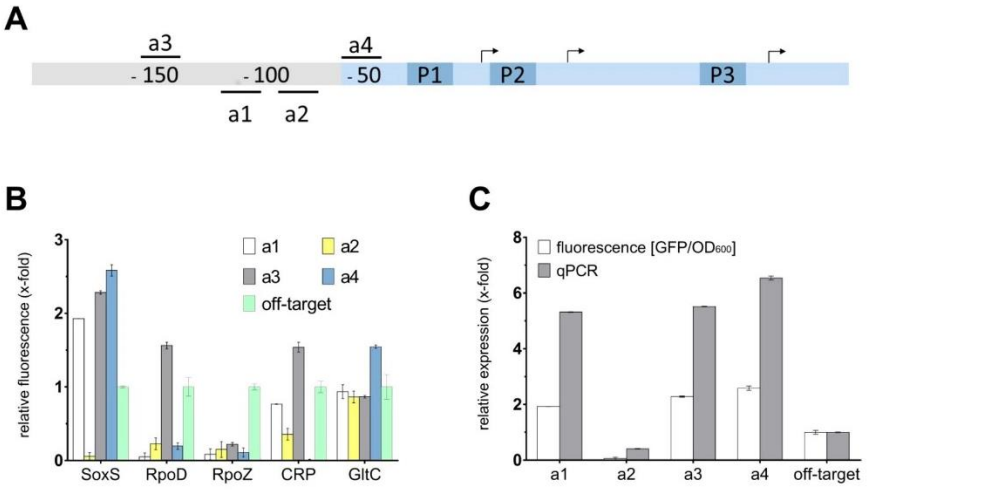


Figure 1. Establishment of a CRISPRa system using dCas12a linked to activator domains. (A) Schematic overview of spacers used for CRISPRa (a1−a4) upstream of the *sgsE* promoter (blue). dCas12a was fused to different activator domains (SoxS, RpoD, RpoZ, CRP, GltC) and positioned upstream of the *sgsE*-promoter with multiple spacers on the template and nontemplate strands. The promoter consists of three core promoter binding sites (P1−P3), of which the heat-inducible P1 site corresponds to the strongest expression.[15] CRISPRa experiments aim to activate expression from P1 already at low temperatures. Arrows indicate the TSS of the corresponding core promoter sites. (B) GFP expression using different activator domains and spacer sequences (a1−a4) relative to a corresponding off-target gRNA. SoxS showed up to 2.5-fold GFP fluorescence with three gRNAs, while RpoD, CRP, and GltC demonstrated elevated fluorescence for only one spacer, respectively. (C) Expression levels determined by qPCR (relative expression) showed up to 6.5-fold increased transcription levels of *gfp* for soxS variants compared to those of off-target gRNAs.

order to enable a more flexible, long distance regulation of target promoters, resulting in a remarkable dynamic output range.[24]

While Cas9 derived from *Streptococcus pyogenes* is the most well-studied RNA-guided endonuclease and was used in a multitude of studies, it has demonstrated several downsides in simultaneously targeting multiple loci. Although, multiplex genome editing can be realized, a uniform expression of multiple gRNAs proved to be challenging. Strategies to overcome this constraint include the expression of sgRNA transcripts from multiple plasmids, the coexpression of RNA processing enzymes such as RNase III[25] and Csy4,[26,27] or the flanking of consecutive gRNAs by ribozymes or tRNAs that enable efficient processing of the mature gRNA[28,29] from a single transcript. However, all these strategies are limited in the number of multiplex targets due to cytotoxic effects. Contrary, multiplex genome editing approaches using Cas12a nuclease orthologs (also known as Cpf1) from *Francisella novicida*, *Acidaminococcus* sp., or *Lachnospiraceae* sp. require only the expression of a single crRNA array.[30,31] Unlike Cas9, Cas12a additionally possesses RNase activities to process the precursor crRNA array and form the gRNAs necessary to direct the CRISPR nuclease to the target DNA.[31] By leveraging this dual RNase/DNase function, the simultaneous perturbation of 25 individual targets was demonstrated in mammalian cell lines using a single transcript harboring both the open reading frame of Cas12a and a CRISPR array.[32] In contrast to Cas9, the protospacer adjacent motif (PAM) is located upstream of the cleavage site and consists of a sequence with a very low GC content. For all commonly used Cas12a nucleases from *Francisella novicida*, *Acidaminococcus* sp., or *Lachnospiraceae* sp., the most efficient PAM required for cleavage was determined as TTTV.[33,34] While this particular PAM is more restrictive compared to NGG of *Sp*Cas9, protein engineering efforts to loosen the stringency of CRISPR nucleases to enable genome editing in otherwise inaccessible loci were successful.[35,36]

Due to these advantages for multiplex genome perturbation studies, dCas12a has been extensively used for the tunable transcriptional regulation of gene expression *via* CRISPRi and CRISPRa in eukaryotic cells.[30,32,36] Despite rapid advances in CRISPR-based technologies, to the best of our knowledge, only CRISPR interference studies have previously been reported for prokaryotic cells using dCas12a,[37] while publications demonstrating the targeted gene activation *via* CRISPRa are still missing for this promising approach. Therefore, the aim of this study was to establish a functional dCas12a-based multiplex gene modulation system capable of CRISPRa and CRISPRi using a broad host range plasmid.

Paenibacillus polymyxa is a Gram-positive, spore forming, nonpathogenic, soil bacterium[38] of biotechnological interest for its ability to produce enantiopure *R,R*-2,3-butanediol (2,3-BDL) and exopolysaccharides (EPS) with interesting material properties (Figure S1).[39,40] *P. polymyxa* DSM 365 putatively produces two distinct heteroexopolysaccharides that we termed Paenan I and Paenan II, which are reflected in two functionally complete EPS clusters encoding all genes required for the Wzx/Wzy biosynthesis pathway.[41,42] Knockouts of distinct glycosyltransferases within the clusters resulted in EPS variants with altered rheological behavior.[41] *P. polymyxa* is also applied in the production of 2,3-BDL *via* the mixed acid fermentation pathway in microaerobic conditions. Depending on oxygen availability, the production of side products such as lactate, formate, and ethanol is required to maintain the redox balance.[43] We recently showed that targeted engineering of the side pathways competing for pyruvate significantly increased

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

the productivity of 2,3-BDL biosynthesis.[44] However, the knockout of a specific lactate dehydrogenase did not result in decreased titers of the byproduct lactate due to the action of redundant homologues. In this study, lactate should be eliminated *via* the concerted knockdown of all four different lactate dehydrogenases found in the genome of *P. polymyxa*. Additionally, the carbon flux should be directed toward 2,3-BDL by inducing the expression of butanediol dehydrogenase (*bdh*) in parallel *via* a newly developed CRISPRa/i system.

RESULTS AND DISCUSSION

Identification of Functional Transcriptional Activator Domains. In a first step, the transcription activator domains of different regulatory protein families were tested in order to identify a suitable candidate for CRISPRa (Table S5). Each domain was linked by translational fusion to the C-terminal end of dCas12a through a 10 amino acid flexible linker peptide (-GSEASGSGRA-). As endogenous transcription activators from *P. polymyxa*, the cAMP receptor protein (CRP) RNA polymerase subunits σ^{70} (RpoD) and ω (RpoZ), as well as the regulator of the glutamate synthase operon (GltC) were evaluated. SoxS, an activator of the superoxide stress genes from *E. coli*, was chosen as an additional heterologous regulator. Out of these, only RpoZ and SoxS have previously been reported as suitable candidates using dCas9-based CRISPRa systems.[9,23] The plasmid pCRai_sfGFP was constructed by isothermal assembly on the basis of the previously established Cas9 genome editing plasmid pCasPP.[41] pCRai_sfGFP encodes dCas12a linked to the different transcriptional activators, the corresponding gRNA expression cassette, as well as *sfgfp* under the control of constitutive promoters.

The sgsE-promoter from *Geobacillus stearothermophilus* used for sfGFP expression is a temperature sensitive promoter containing three core promoter sites.[45] At low temperatures of 28 °C, the front most core promoter (P3) is active, resulting in a weak basal expression, while elevated temperatures of 37−45 °C lead to a highly increased expression from RNA-polymerase binding sites further upstream (P1 and P2), as shown in *B. subtilis*.[45] Furthermore, this well-characterized promoter has demonstrated a robust but weak expression in *P. polymyxa* for all cultivation conditions used in this study.[41,45] Therefore, the sgsE-promoter was chosen to test the CRISPRa activities of the different activator domains in order to induce strong expression levels even at a low temperature (Figure 1A). Eukaryotic CRISPRa systems allow for a relatively broad range, in which gRNAs mediate the binding of the CRISPR effector module to efficiently activate or repress the expression of target genes. Contrarily, bacterial CRISPRa systems have been demonstrated to be highly sensitive to the correct distance of the activated dCas12a target sequence to the promoter.[23] Bacterial CRISPRa systems act by facilitating the recruitment of the RNA-polymerase to the promoter, while eukaryotic systems typically cause chromatin rearrangements to interfere with the expression of target genes.[13] For bacterial dCas9-based systems, an optimal distance was determined in the range from 60 to 100 bp upstream of the transcriptional start site (TSS).[23] The optimal distance might vary depending on the different activator domains. Consequently, we tested four different spacers (target sequence of the gRNA) allowing the activated effector module to bind to the template and nontemplate strands in the range 40−120 bp upstream of the TSS to induce expression from the strong RNA-polymerase binding site P1 (Figure 1). The PAM sites for all spacers showed the same motif (TTTC). In order to test whether the observed effects actually arise from the binding of the dCas12a-activator complex to the respective target sites, additional constructs expressing off-target gRNAs not binding to the plasmid or genome of *P. polymyxa* DSM 365 were constructed.

Out of all tested activator constructs, *Ec*SoxS demonstrated the best performance in *P. polymyxa* (Figure 1B). Three out of the four tested gRNAs significantly increased the expression of sfGFP and showed an increased fluorescence signal during photometric evaluation. The qPCR experiments showed up to 6.5-fold increased transcription levels using gRNA_a4, but also, gRNA_a1 and gRNA_a3 bound to dCas12a-soxS displayed increased transcription and fluorescence signals (Figure 1C). Surprisingly, the highest fluorescence signal was achieved using the spacer a4, which was positioned 50 bp upstream of the TSS of the heat-inducible promoter site P1 (Figure 1A). While the close proximity of the binding site of this gRNA lies outside of the ideal distance determined for a dCas9-soxS construct,[23] the distance to the second RNA-polymerase binding site (P2) of 85 bp might result in transcription from the secondary heat-inducible promoter.

Additionally, dCas12a fused to other activator domains such as RpoD, GltC, and CRP, respectively, also demonstrated increased fluorescence results for individual spacer sequences (Figure 1B). However, some combinations also led to a decreased fluorescence signal of GFP in *P. polymyxa*, indicating that the effector module blocks the binding of the RNA-polymerase to the promoter and therefore effectively represses transcription of the gene of interest. For GltC and CRP, the use of gRNAs a3 and a4, which are located in close proximity to each other (10 bp), affects the GFP expression, changed from 2-fold increased GFP signal to transcriptional repression. All of the investigated activators act by direct interaction with the RNA-polymerase.[46−48] Contrary to other activators, there is experimental evidence suggesting that SoxS already forms a binary prerecruitment complex with the C-terminal domain of the α-subunit and scans DNA for cognate SoxS binding sites.[49,50] Therefore, we hypothesize that a similar prerecruitment is formed with the dCas12a-SoxS fusion protein, which allows for more flexibility in the correct distancing of the gRNA to the promoter binding site. Due to the consistent performance of the dCas12a-soxS constructs, this particular activator domain was used for all further experiments.

Interestingly, while all experiments using the fluorescence reporter system were performed in *P. polymyxa*, observed effects were almost identical in *E. coli* S17-1 that was used for the conjugational transformation of *P. polymyxa* DSM 365 (Figure S3). Consequently, we demonstrated a broad host range use of the constructed pCRai_soxS plasmid in both Gram-positive and Gram-negative bacteria. In case of the fluorescence reporter assays, in which all functional parts were encoded on a single plasmid, it was possible to accelerate the screening of potential guide RNAs by using *E. coli* S17-1 as a prescreening platform prior to the more time-consuming conjugational transformation of *P. polymyxa* DSM 365.

Our results exemplified that the stringency of gRNA positioning with the SoxS domain is lower compared to with other activators. Empirical testing of multiple spacers is still required to enable improved activation of target promoters. However, it might be possible to establish a design rule set to enable *a priori* construction of optimized spacers with more experimental data using different promoters.

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

Establishment of CRISPRi & Multiplexing CRISPRa/i. In a next step, we evaluated whether the use of the dCas12a-soxS activator constructs is also possible for CRISPRi by repositioning the activated dCas12a-gRNA complex within the open reading frame of sfGFP. Thereby, the effector module acts as a road block for the RNA-polymerase and inhibits the elongation of the nascent transcript. Three different spacer sequences were tested (Figure 2A). While expression levels of

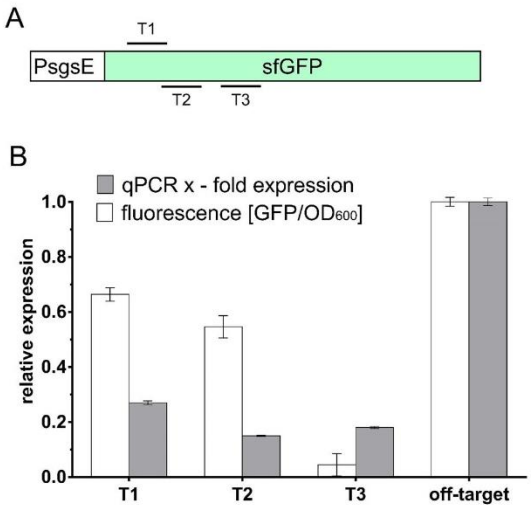


Figure 2. Establishment of CRISPRi in *P. polymyxa* using pCRai_soxS. (A) Schematic overview of gRNA binding sites within the ORF of sfGFP. (B) Three spacers within the ORF of *sfgfp* (T1–T3) were tested, respectively, by fluorescence experiments and qPCR expression analysis relative to an off-target gRNA. While the transcriptional expression levels were reduced by 75%–80% for all gRNAs, the measured fluorescence levels of GFP fluctuated more between different spacers. The activated complex using gRNA T3 demonstrated the best repression, resulting in a highly reduced fluorescence signal as well as a reduced transcription of *sfgfp*. The reporter expression was determined photometrically and by qPCR experiments in biological triplicates.

sfgfp were significantly decreased by approximately 80% for all constructs compared to those of an off-target construct, the actual fluorescence of GFP remained at higher levels using gRNAs T1 and T2 (Figure 2A,B). Even though, fluorescence signals were not fully eliminated, a severe decrease of up to 95% using gRNA T3 was observed.

In order to test the capability of our construct to simultaneously repress and activate different target genes, a constitutively expressed mRFP reporter cassette was cloned in pCRai_soxS in addition to the sfGFP reporter. For CRISPRa, the previously used gRNA_a1 was chosen to induce the expression of sfGFP. For CRISPRi, a new spacer was designed binding within the ORF of *mrfp* (Figure 3A). All the strains of *P. polymyxa* were compared to a strain harboring an off-target CRISPR array. Due to weak fluorescence signals after 24 h of inoculation, only transcriptional expression levels were determined *via* qPCR at this point of time, but a photometric evaluation of the reporters was performed after 48 h. When expressed individually, CRISPRi resulted in a reduction of the mRFP fluorescence signal by 74%, while CRISPRa increased

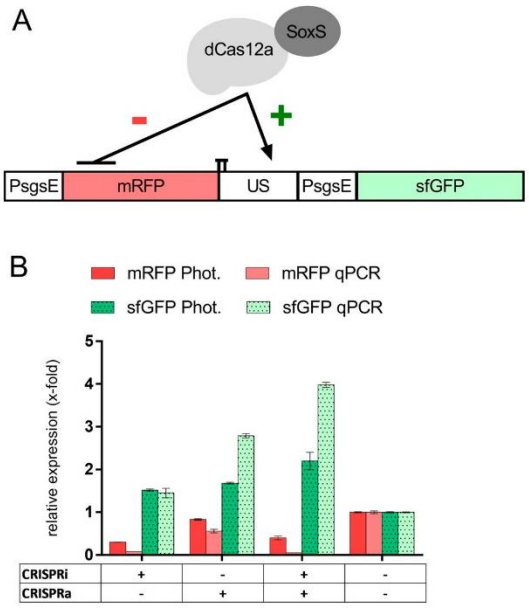


Figure 3. Simultaneous repression (CRISPRi, mRFP) and activation (CRISPRa, sfGFP) of fluorescence reporters. (A) Schematic display of spacer sequences within the ORF of mRFP and in the upstream region (US) of PsgsE controlling the expression of sfGFP. (B) Multiplex transcriptional perturbation was tested in *P. polymyxa* harboring a plasmid for the constitutive expression of GFP and mRFP fluorescence reporters. Single CRISPR arrays were designed targeting the ORF of mRFP or the upstream region of PsgsE controlling *sfgfp* expression (gRNA_a1). The expression of gRNAs resulted in the repression of mRFP or induction of sfGFP, respectively. When both activated gRNAs were expressed simultaneously, the obtained fluorescence results were similar to the expression of individual gRNAs alone. All results are depicted relative to an off-target CRISPR array encoding a spacer not present in the strain. The reporter expression was determined photometrically (after 48 h cultivation) and by qPCR experiments (after 24 h cultivation) in biological triplicates.

sfGFP expression by 68% (Figure 3B). The simultaneous expression of both activated gRNAs from a single CRISPR array decreased the mRFP fluorescence by 60% while increasing the fluorescence signal of sfGFP by 120%. Therefore, we demonstrated the efficient control of the expression of multiple genes using a single CRISPR array. Depending on the positioning of the spacer, it proved possible to activate or repress multiple target genes in parallel. Interestingly, the repression of mRFP alone seemed to positively influence the expression of sfGFP. We hypothesize that this synergistic effect resulted out of a lower competition of the PsgsE promoter corresponding to *sfgfp* and the PsgsE-*mrfp* expression cassette and a reduced metabolic burden.

Multiplex CRISPRi to Modify Exopolysaccharide Composition of *P. polymyxa* DSM 365. The most important advantage of Cas12a over the more commonly used Cas9 is the ability of the effector nuclease to process its own crRNA, allowing for the simultaneous targeting of multiple loci through a single CRISPR array.[32]

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353–3363

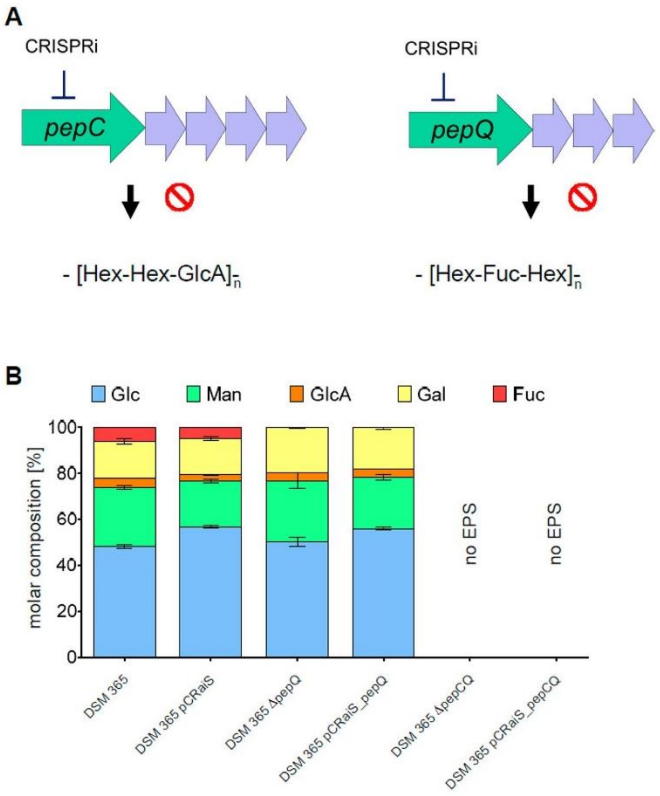


Figure 4. Carbohydrate fingerprint of the heteroexopolysaccharide of *P. polymyxa* DSM 365 and engineered variants. The transformation of *P. polymyxa* with a plasmid encoding the SoxS activator domain and off-target gRNA (pCRaiS) did not alter the EPS composition significantly. The expression of a gRNA targeting the ORF of the initiating glycosyltransferase *pepQ* (pCRaiS_pepQ) resulted in the loss of fucose within the EPS composition that was also observed in the knockout strain Δ*pepQ*. Targeting of the ORFs of both GTis (pCRaiS_pepCQ) did not yield any EPS resembling the same phenotype as the double knockout Δ*pepCQ*, Δ: gene deletion by CRISPR-Cas9-mediated genome engineering. DSM 365, *P. polymyxa* DSM 365; pCRaiS, pCRai_soxS; Hex, hexose; GlcA, glucuronic acid; Fuc, fucose; Glc, glucose; Man, mannose; Gal, galactose.

P. polymyxa DSM 365 is an avid producer of different exopolysaccharides. Depending on the process conditions, variable polymer mixtures are produced.[39] However, it has also been shown that the engineering of a polysaccharide structure with modified physicochemical properties is feasible.[41] The underlying gene cluster contains the two so-called initiating glycosyltransferases (GTi) PepC and PepQ, which are putatively responsible for the initiation of the biosynthesis of two distinct polysaccharides (Figure 4A). Contrary to the first polymer (Paenan I), the second polymer (Paenan II) initiated by PepQ contains the deoxyhexose fucose (unpublished data). To evaluate the effects of CRISPRi constructs on genomic targets, the GTi PepQ was targeted alone or in combination with PepC. Strains harboring the plasmids pCRaiS (pCRai_soxS encoding an off-target spacer), pCRaiS_pepQ (targeting the ORF of *pepQ*), and pCRaiS_pepCQ (targeting the ORFs of *pepC* and *pepQ*) were constructed and used in EPS batch fermentations. The carbohydrate fingerprints of the obtained EPS were performed and compared to those of the respective knockout strains (Figure 4B).

The strains harboring the pCRaiS plasmid expressing an off-target CRISPR array did not show an altered EPS composition. When the second GTi *pepQ* was targeted, fucose diminished, indicating the absence of Paenan II in the EPS mixture. In a next step, both initiating GTs were targeted simultaneously. With both GTis downregulated, no EPS at all was produced. In order to evaluate whether the observed effects actually resulted from interference with the respective target genes, knockout strains of the respective target genes with a previously established CRISPR-Cas9 genome editing system were constructed.[41] Our experiments confirmed that the monomer composition was comparable to the respective CRISPRi variants (Figure 4B). In the previous fluorescence reporter assays, a reduced signal of mRFP and GFP could still be detected in the CRISPRi approaches. Contrarily, for the glycosyltransferase targets, no significant differences in the EPS composition were observed between the CRISPRi constructs and the knockout strains, indicating a more efficient repression of the target genes in comparison to the fluorescence assays.

In conclusion, this approach showed that the developed pCRai_soxS tool can be used for the parallel screening of multiple interesting knockout targets in the genome. In addition, the efficiency of this tool proved to be comparable

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

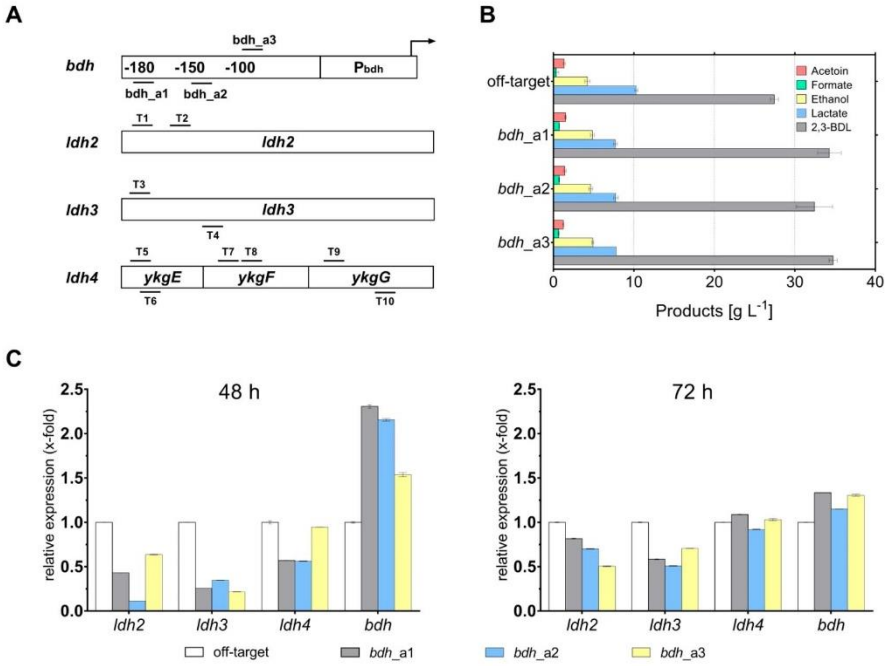


Figure 5. Multiplex CRISPRi and CRISPRa to engineer the mixed acid pathway of *P. polymyxa* and increase 2,3-BDL production. (A) Schematic overview of gRNA binding sites. Three different gRNAs binding upstream of the P_{bdh} promoter were tested individually (bdh_a1−3). Simultaneously, all constructs also targeted three putative lactate dehydrogenase genes with a total of nine spacers (T1−T9) to knock down the respective genes. (B) Product titer obtained after 72 h cultivation at microaerobic conditions. Compared to in an off-target construct, the lactate production was reduced by ∼20% in strains expressing the target gRNAs. 2,3-BDL production was increased and reached a maximum of 34.7 g L^{-1} in the construct using bdh_a3 to target *bdh* expression. The depicted values represent the mean of biological duplicates. (C) Expression of target genes was analyzed *via* qPCR after 48 and 72 h of cultivation. After 48 h, the transcription levels of lactate dehydrogenases (*ldh*2−4) were significantly reduced compared to those of a strain expressing off-target gRNAs. Furthermore, also, the expression of a butanediol dehydrogenase was increased. However, after 72 h of cultivation, the effects of CRISPRi and CRISPRa were severely reduced.

to more laborious and time-consuming genome-editing approaches.

Multiplex CRISPRa/i to Screen for Metabolic Engineering Targets of the Butanediol Biosynthesis Pathway in *P. polymyxa* DSM 365. In a previous study, we engineered the mixed acid pathway of *P. polymyxa* DSM 365 to increase the production of 2,3-BDL and remove undesirable side-products. Interestingly, the knockout of a lactate dehydrogenase (*ldh1*) resulted in an adapted growth behavior, increased biomass formation, and consequently enhanced 2,3-BDL formation.[44] However, due to the presence of additional homologues of *ldh* within the genome, lactate formation could not be completely eliminated. Therefore, in order to demonstrate the capabilities of our dCas12a-based CRISPR-tool, all additional copies of *ldh* homologues were targeted in parallel. By leveraging the CRISPR array processing abilities of Cas12a, each gene was targeted with two gRNAs at the same time. As the decoupling of the 2,3-BDL biosynthesis from its natural regulon showed positive effects, the expression of the butanediol dehydrogenase should also be transcriptionally activated in *P. polymyxa* DSM 365 Δ*ldh1*. Therefore, the corresponding promoter was predicted using the Softberry CNNPromoter_b tool.[51] Three distinct gRNAs binding 106−180 bp upstream of the putative TSS were tested separately (bdh_a1−a3). Thereby, three strains carrying plasmid constructs, each targeting 11 genomic sites in parallel were designed and evaluated in batch fermentations using microaerobic conditions (Figure 5A). Due to the more stringent PAM requirements of Cas12a (TTTV), spacers more distant to the TSS had to be used for the *bdh* promoter compared to those of previous fluorescence experiments.

The 2,3-BDL fermentations were conducted for 72 h. Despite the fact that each ORF encoding the different lactate dehydrogenases was targeted with two gRNAs, the lactate production in all strains was reduced only by ∼20% compared to that of the strain harboring the off-target spacer (Figure 5B). However, 2,3-BDL titers were increased from 27.5 to 34.7 g L^{-1} for the *P. polymyxa* expressing the activated complex bdh_a3, corresponding to a 26% increased product titer. Additionally, also, the strains encoding bdh_a1 and bdh_a2 showed 25% and 18% increased 2,3-BDL titers, respectively. All the other end products of the mixed acid pathway that were not targeted by any gRNA remained similar in all variants (Figure 5 and Figure S4). Furthermore, the 2,3-BDL yields were increased by approximately 20% for all strains encoding target gRNAs, indicating a redirection of the carbon flux from lactate to 2,3-BDL (Table S4).

While the general principle of our developed CRISPRa/i tool could be successfully demonstrated, the effects of both transcriptional repression and activation were not as pronounced as those observed in the fluorescence and EPS experiments. An expression analysis *via* qPCR revealed a

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

transcriptional perturbation of all *ldh* homologues ranging from 50% to 80% after 48 h of cultivation (Figure 5C). Furthermore, two gRNAs binding 106 and 146 bp, respectively, upstream of the TSS of the *bdh* promoter caused more than a 2-fold increased expression on the transcriptional level. However, after 72 h, the effects on the transcriptional perturbation were significantly decreased. We hypothesize that the observed decreased effects by our dCas12a tool are a combined result of a rather low expression of the CRISPR tool at microaerobic conditions and the long cultivation times used for the 2,3-BDL fermentations. Compared to other 2,3-BDL producing organisms such as *Serratia marcescens* or *Klebsiella pneumoniae*, *P. polymyxa* DSM 365 seems to be more sensitive toward high concentrations of 2,3-BDL with a toxic threshold at approximately 50 g L^{-1}.[52] As a result, using the applied microaerobic cultivation conditions, an increased 2,3-BDL production is typically accompanied by elevated levels of lactate, which is used as a nontoxic, redox neutral electron acceptor.[44] Consequently, the decreasing effects of transcriptional perturbation might be caused by genetic instabilities or by competing endogenous regulation mechanisms. Furthermore, the used spacers in a more distal region to the *bdh* promoter might be suboptimal for dCas12a-soxS. Due to the restrictive PAM site of Cas12a of *Acidaminococcus* sp., optimal distancing from the TSS might not always be possible and may impede genome wide screenings. However, engineered variants of Cas12a have shown expanded binding motifs and enabled the targeting of otherwise inaccessible PAMs.[36]

The dynamic range of transcriptional perturbation described in this study was rather low. Depending on the applied conditions and targets, a 2−6.5-fold increased transcription was achieved by CRISPRa. While the expression of some genes could be reduced by 80% using CRISPRi, other targets were only poorly affected. Currently, each spacer is required to be tested separately to determine the efficiency of the system. The efficiencies of different PAMs have to be investigated more closely in the context of catalytically inactive Cas nucleases. Protein engineering approaches to modify dCas12a or the activator domain have been previously demonstrated to be valid strategies to optimize and improve the binding efficiency of the effector module and dynamic output range for CRISPRa.[23,36]

■ CONCLUSION

While CRISPRi has been continuously demonstrated in bacteria, CRISPRa technology is lacking behind their eukaryotic counterparts. Currently available systems are still limited in their number of targets that can be modified in parallel due to the use of dCas9. In this study, we showed the first successful utilization of dCas12a for the simultaneous activation and repression of multiple genes in the alternative host organism *P. polymyxa* DSM 365.

While spacers particularly for CRISPRa still need to be optimized individually for each target promoter, the utilization of SoxS as an activator domain enables more flexibility in the correct distancing to the target promoter compared to other tested activator domains. In this study, we demonstrated an efficient broad host range tool for the parallel transcriptional modulation of expression patterns in bacteria that can be applied for both metabolic engineering efforts and screening of potential targets for further studies. With ongoing studies using dCas12a-SoxS-based tools for CRISPRa in bacterial hosts it will be possible to establish more precise design rule sets for the efficient positioning of the effector module for CRISPRi and CRISPRa to facilitate and accelerate the use of dCas12a-based transcriptional perturbation tools.

Even though Cas12a is more restricted in terms of its PAM sequence compared to that of Cas9, its ability to efficiently process its own gRNAs makes it a promising tool to orchestrate sophisticated genetic reprogramming of bacterial cells or to screen for engineering targets in the genome.

In conclusion, this work demonstrated both the simultaneous activation and repression of multiple targets in the genome of *P. polymyxa* using a single CRISPR array and represents therefore an important extension of current Cas9-based tools. We demonstrated that the developed tool is functional in common bacterial cell factories such as *E. coli* as well as in the Gram-positive alternative host organism *P. polymyxa*. The usage of multiplex transcriptional perturbation will facilitate the combinatorial regulation of complex pathways.

■ MATERIALS AND METHODS

Strains and Media. *P. polymyxa* DSM 365 was acquired from the German Collection of Microorganisms and Cell Culture (DSMZ). *E. coli* NEB Turbo cells (New England Biolabs) were used for any plasmid construction presented in this study. *E. coli* S17-1 (DSMZ strain DSM 9079) was utilized for the transformation of *P. polymyxa* DSM 365 *via* conjugation. All medium components were obtained from Carl Roth GmbH if not indicated differently. For cloning procedures, the strains were grown in LB media (5 g L^{-1} yeast extract, 10 g L^{-1} tryptone, and 10 g L^{-1} NaCl) and additionally supplemented with 50 μg mL^{-1} neomycin and 20 μg mL^{-1} polymyxin if required. All the strains were stored in 30% glycerol at −80 °C. Prior to cultivation, the strains were streaked on LB agar plates and grown at 30 °C. All the strains used or constructed in this study are listed in Table S1.

For 2,3-BDL fermentations, a single colony was used for the inoculation of 50 mL of preculture medium containing 60 g L^{-1} glucose, 5 g L^{-1} yeast extract, 5 g L^{-1} tryptone, 0.2 g L^{-1} $MgSO_4$ heptahydrate (Sigma-Aldrich), 3.5 g L^{-1} KH_2PO_4, and 2.5 g L^{-1} K_2HPO_4. The fermentation medium components were autoclaved separately and contained 120 g L^{-1} glucose, 5 g L^{-1} yeast extract, 3.5 g L^{-1} tryptone, 0.2 g L^{-1} $MgSO_4 \cdot 7H_2O$, 3.5 g L^{-1} KH_2PO_4, 2.5 g L^{-1} K_2HPO_4, 5 g L^{-1} ammonium acetate, 4 g L^{-1} $(NH_4)_2SO_4$, and 3 mL L^{-1} trace element solution. The trace element solution contained 2.5 g L^{-1} iron sulfate heptahydrate, 2.1 g L^{-1} sodium tartrate dihydrate, 1.8 g L^{-1} manganese chloride dihydrate, 0.075 g L^{-1} cobalt chloride hexahydrate, 0.031 g L^{-1} copper sulfate pentahydrate, 0.258 g L^{-1} boric acid, 0.023 g L^{-1} sodium molybdate dihydrate, and 0.021 g L^{-1} zinc chloride. The trace element solution was filter-sterilized and added to the media after cooling down to room temperature.

For EPS production, MM1 P100 medium[39] was used as described before, containing 30 g L^{-1} glucose and 5 g L^{-1} peptone. The corresponding preculture medium contained a reduced amount of 10 g L^{-1} glucose and was buffered to pH 6.8 with 20 g L^{-1} MOPS.

Plasmid Construction. The gene encoding for an engineered (E174R, N282A, S542R, K548R)[36] catalytically inactivate (D908A) variant of AsCas12a was codon optimized for *Bacillus* sp. and synthesized by ATG:biosynthetics. The basic plasmid pCRai (Figure S2) was assembled by isothermal Gibson Assembly[53] from three PCR-amplified fragments

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

consisting of a pUB110 derived backbone including oriT for conjugational transfer,[41] a lacZ replacement cassette for BbsI-based cloning of target gRNAs and the codon optimized *enAsdcas12a* cassette. Activator domains were PCR-amplified from extracted gDNA of *P. polymyxa* DSM 365 and *E. coli* NEB Turbo, respectively, and cloned into pCRai by Golden Gate Assembly using BsaI. The cloning of the gRNA sequences was conducted as previously described.[41] The PsgsE-sfGFP reporter was cloned *via* isothermal Gibson Assembly using a unique SpeI site of pCRai. The dual reporter plasmid was constructed by cloning a PsgsE-mRFP and a PsgsE-sfGFP reporter cassette in tandem by Golden Gate Assembly after linearization of pCRai_soxS with SpeI/*Sal*I. All the oligonucleotides used for the construction of plasmids are listed in Tables S2 and S3. The gRNA targeting sequences (spacers) and DNA sequences of the used activator domains are listed in Tables S4 and S5.

Conjugation-Based Transformation of *P. polymyxa* DSM 365. *P. polymyxa* was transformed by conjugation using *E. coli* S17-1 harboring the various plasmids. Overnight cultures of donor and recipient strains were diluted 1:100 with selective or nonselective LB media, respectively, and cultivated at 37 °C for 3 h, 280 rpm. The recipient culture (900 μL) was heat shocked at 42 °C for 15 min and mixed with 300 μL of the donor strain culture. Cells were centrifuged at 6000*g* for 2 min, resuspended in 800 μL of LB media, and dropped on nonselective LB agar plates. After 24 h of incubation at 30 °C, the cells were scrapped off and resuspended in 500 μL of LB broth, and 100 μL thereof was plated on selective LB agar containing 50 μg mL^{-1} neomycin and 20 μg mL^{-1} polymyxin for counter selection. *P. polymyxa* conjugants were analyzed for successful transformation after 48 h of incubation at 30 °C by cPCR. The confirmed knockout strains were plasmid cured by cultivation in LB broth without antibiotic selection pressure and subsequent replica plating on LB agar plates both with and without neomycin. Strains that did not grow on plates with selection marker were verified by sequencing of the target region and used for further experiments.

Photometric Assay. For sfGFP fluorescence experiments, 3 mL of EPS medium supplemented with 50 μg mL^{-1} neomycin was inoculated with a single colony of the respective strains and grown overnight at 30 °C, 200 rpm. After 18 h, each strain was subcultured 1:100 in 3 mL of selective MM1 P100 medium and grown for 24 h at 30 °C, 200 rpm. After 24 h, 100 μL was transferred to a 96-well microtiter plate, and the OD_{600}, GFP fluorescence (Ex. 488 nm Em. 515 nm), and mRFP fluorescence (Ex. 560 nm Em. 600 nm) were measured in an Ultraspec 10 spectrophotometer (Amersham Biosciences). The fluorescence values were normalized to OD_{600} in all experiments. In parallel, 1 mL of each culture was pelleted by centrifugation and used for qPCR experiments. The relative expression signals were calculated on the basis of an off-target construct expressing gRNAs not binding to the genome of *P. polymyxa* DSM 365 or any plasmid used in this study (Table S4).

Quantitative RT-PCR. RNA extraction of positive samples of the GFP fluorescence assay as well as the butanediol fermentation processes was performed using an Aurum Total RNA Mini Kit (BioRad) according to the manufacturer's instructions. The synthesis of cDNA was conducted using iScript reverse transcriptase (BioRad) using 1 μg of total RNA template. The qPCR reactions were performed in triplicate on a CFX-96 thermocycler using SsoAdvanced Universal SYBR Green Supermix (BioRad) using 5 ng of cDNA as a template in a 10 μL reaction volume. Negative controls without reverse transcriptase during cDNA synthesis were used in order to evaluate the absence of gDNA contaminations. The relative gene expression levels were calculated on the basis of the ΔΔCq method,[54] and *gyrA* was used as a reference gene. After qPCR, a melting curve analysis was performed to confirm the presence of a single PCR product for each target. The designed primers were analyzed by an OligoAnalyzer Tool (IDT) to avoid hairpin formation and self- and heterodimer formation with free energy values more than 10 kcal mol^{-1}. The oligonucleotides used for qPCR experiments are listed in Table S2.

CRISPR-Cas9-Mediated Genome Editing. All gene knockouts were performed as previously described by Rütering et al.[41] In brief, gRNAs for the targeted genome regions were designed using a Benchling CRISPR Design Tool. For each target, a minimum of two gRNAs were designed, typically targeting distinct regions of the open reading frame. Oligonucleotides were phosphorylated, annealed, and cloned into pCasPP by a Golden Gate assembly. Approximately, 1 kB up- and downstream homology flanks for each targeted nucleotide sequence were amplified from genomic DNA of *P. polymyxa* DSM 365 using Phusion Polymerase according to the manufacturer's instructions and fused by an overlap extension PCR *via* a 20 bp overlap. Homology flanks were cloned into pCasPP through Gibson Assembly or molecular cloning after linearization by the use of SpeI. After the transformation of *E. coli* NEB Turbo, clones were analyzed for a correct construct assembly by colony PCR (cPCR) and sequencing of the amplicons. Finally, the correct constructs were transferred to chemical competent *E. coli* S17-1 cells for the following conjugational transformation of *P. polymyxa*.

EPS Batch Fermentation. EPS fermentations were conducted in a 1 L DASGIP parallel bioreactor system with a working volume of 500 mL. A single colony from a freshly streaked plate was used to inoculate 100 mL of MM1 P100 preculture medium by following incubation for 16 h at 30 °C, 160 rpm. The bioreactors were inoculated to give an initial OD of 0.1. Fermentation was performed at 30 °C, and the stirrer speed (200−600 rpm) and gassing (6−10 L h^{-1}) with pressurized air through a L-sparger were controlled to maintain a 30% DO saturation. The stirrer was equipped with a 6-plate-rushton impeller placed 2.5 cm from the bottom of the shaft. The pH value was maintained at 6.8 and automatically adjusted with 2 M NaOH or 1.35 M H_3PO_4 as required. Foam control was performed using 1% of antifoam B (Merck). For monitoring of the process parameters, the reactors were equipped with redox and dissolved oxygen probes.

After the end of the process, the fermentation broth was diluted 1:10 with dH_2O and the biomass was separated by centrifugation (15 000*g*, 20 °C, 20 min) followed by cross-flow filtration of the supernatant using a 100 kDa filtration cassette (Hydrosart, Sartorius AG). EPS was precipitated by slowly pouring the concentrated fermentation supernatant into two volumes of isopropanol. EPS was collected and dried overnight in a VDL53 vacuum oven at 40 °C (Binder). The dry weight of the obtained EPS was determined gravimetrically prior to being milled to a fine powder in a ball mill at 30 Hz for 1 min (Mixer Mill MM400, Retsch GmbH).

Carbohydrate Fingerprinting. EPS monosaccharide composition was analyzed using the 1-phenyl-3-methyl-5-

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

pyrazolone high-throughput method (HT-PMP) as previously described using 1 g L^{-1} reconstituted EPS solutions.[55] In brief, 0.1% EPS solutions were hydrolyzed in a 96-well plate, sealed with a rubber mat, and further covered by a custom-made metal device with 2 M trifluoroacetic acid (90 min, 121 °C). The samples were neutralized with 3.2% NH_4OH. A PMP master mix (125 mg of PMP, 7 mL of MeOH, 3,06 mL of dH_2O, 437.5 μL of 3.2% NH_4OH) (75 μL) was mixed with 25 μL of neutralized hydrolysate and incubated at 70 °C for 100 min in a PCR cycler. Twenty microliters of derivatized samples was mixed with 130 μL of a 1:26 dilution of 0.5 M acetic acid, and the mixture was filtered with a 0.2 μm filter plate (1000g, 2 min), followed by HPLC−UV-MS analysis as previously described.[55]

Butanediol Batch Fermentation. Batch fermentations were conducted in 1 L DASGIP bioreactors (Eppendorf) with an initial volume of 550 mL. A single colony from a freshly streaked plate was used to inoculate 100 mL of preculture medium by following incubation for 16 h at 30 °C, 160 rpm. This cultivation broth (50 mL, diluted with preculture medium if required) was used to inoculate the bioreactor by an initial OD_{600} of 0.1. Fermentation was performed at 35 °C and a constant aeration of 0.075 vvm. The stirrer was equipped with a 6-plate-rushton impeller placed 4 cm from the bottom of the shaft and constantly stirred at 300 rpm. The pH value was maintained at 6.0 and automatically adjusted with the addition of 2 M NaOH or 1.35 M H_3PO_4 as required. The foam control was performed using 1% of antifoam B (Merck). In order to monitor the process parameters, the reactors were equipped with redox and pH probes. Glucose and product concentrations were determined *via* a HPLC-UV−RID system (Dionex) equipped with Rezex ROA-H^+ organic acid column (300 mm × 7.8 mm Phenomenex). The column temperature was set to 70 °C, and 2.5 mM H_2SO_4 was used as the mobile phase with a flow rate of 0.5 mL min^{-1}. All measured concentrations of 2,3-BDL in this publication represent solely the levo-stereoisomer of the alcohol if not explicitly noted differently.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00424.

Figures of metabolic pathway scheme, plasmid map, relative fluorescence levels, and fermentation profiles and tables of bacterial strains, plasmids, oligonucleotides and primers, spacer sequences and protospacer adjacent motifs, and nucleotide sequences of activator domains used and final product titers and 2,3-BDL yield of 2,3-BDL fermentations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jochen Schmid − *Chair of Chemistry of Biogenic Resources, Technical University of Munich, 94315 Straubing, Germany; Institute for Molecular Microbiology and Biotechnology, University of Münster, 48149 Münster, Germany;* orcid.org/0000-0003-2557-5532; Email: jochen.schmid@uni-muenster.de

Authors

Christoph Schilling − *Chair of Chemistry of Biogenic Resources, Technical University of Munich, 94315 Straubing, Germany;* orcid.org/0000-0002-7203-5611

Mattheos A. G. Koffas − *Department of Chemical and Biological Engineering and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York 12180, United States;* orcid.org/0000-0002-1405-0565

Volker Sieber − *Chair of Chemistry of Biogenic Resources, Technical University of Munich, 94315 Straubing, Germany; Fraunhofer IGB, Straubing Branch BioCat, 94315 Straubing, Germany; TUM Catalysis Research Center, 85748 Garching, Germany; School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia 4072, Australia;* orcid.org/0000-0001-5458-9330

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00424

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the German Federal Ministry of Education and Research (BMBF) in frame of the project MaPolKo (number 03VP02560). M.A.G.K. would like to acknowledge funding from the U.S. National Science Foundation, award MCB-1817631.

■ ABBREVIATIONS

CRISPR, clustered regularly interspaced short palindromic repeats; CRISPRi, CRISPR interference; CRISPRa, CRISPR activation; gRNA, guide RNA; GTi, initiating glycosyltransferase; 2,3-BDL, 2,3-butanediol; EPS, exopolysaccharide; *bdh*, butanediol dehydrogenase; *ldh*, lactate dehydrogenase

■ REFERENCES

(1) Wendisch, V. F., Bott, M., and Eikmanns, B. J. (2006) Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for biotechnological production of organic acids and amino acids. *Curr. Opin. Microbiol. 9*, 268−274.

(2) Jin, Y.-S., and Cate, J. H. (2017) Metabolic engineering of yeast for lignocellulosic biofuel production. *Curr. Opin. Chem. Biol. 41*, 99−106.

(3) Ro, D.-K., Paradise, E. M., Ouellet, M., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T. S., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., and Keasling, J. D. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature 440*, 940−943.

(4) Nielsen, J., and Keasling, J. D. (2016) Engineering Cellular Metabolism. *Cell 164*, 1185−1197.

(5) Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L. A. (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol. 31*, 233.

(6) Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S. J., Kristensen, M., Pedersen, L. E., Jensen, M. K., and Keasling, J. D. (2015) Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng. 28*, 213−222.

(7) Li, Y., Lin, Z., Huang, C., Zhang, Y., Wang, Z., Tang, Y., Chen, T., and Zhao, X. (2015) Metabolic engineering of *Escherichia coli* using CRISPR−Cas9 meditated genome editing. *Metab. Eng. 31*, 13−21.

(8) Roointan, A., and Morowvat, M. H. (2016) Road to the future of systems biotechnology: CRISPR-Cas-mediated metabolic engineering

for recombinant protein production. *Biotechnol. Genet. Eng. Rev. 32*, 74−91.
(9) Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L. A. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res. 41*, 7429−7437.
(10) Depardieu, F., and Bikard, D. (2020) Gene silencing with CRISPRi in bacteria and optimization of dCas9 expression levels. *Methods 172*, 61−75.
(11) Hogan, A. M., Rahman, A. S. M. Z., Lightly, T. J., and Cardona, S. T. (2019) A Broad-Host-Range CRISPRi Toolkit for Silencing Gene Expression in *Burkholderia*. *ACS Synth. Biol. 8*, 2372−2384.
(12) Gilbert, L. A., Horlbeck, M. A., Adamson, B., Villalta, J. E., Chen, Y., Whitehead, E. H., Guimaraes, C., Panning, B., Ploegh, H. L., Bassik, M. C., Qi, L. S., Kampmann, M., and Weissman, J. S. (2014) Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell 159*, 647−661.
(13) Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., and Qi, L. S. (2013) CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell 154*, 442−451.
(14) Yeo, N. C., Chavez, A., Lance-Byrne, A., Chan, Y., Menn, D., Milanova, D., Kuo, C.-C., Guo, X., Sharma, S., Tung, A., Cecchi, R. J., Tuttle, M., Pradhan, S., Lim, E. T., Davidsohn, N., Ebrahimkhani, M. R., Collins, J. J., Lewis, N. E., Kiani, S., and Church, G. M. (2018) An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods 15*, 611−616.
(15) Zalatan, J. G., Lee, M. E., Almeida, R., Gilbert, L. A., Whitehead, E. H., La Russa, M., Tsai, J. C., Weissman, J. S., Dueber, J. E., Qi, L. S., and Lim, W. A. (2015) Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell 160*, 339−350.
(16) Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W., Guilak, F., Crawford, G. E., Reddy, T. E., and Gersbach, C. A. (2013) RNA-guided gene activation by CRISPR-Cas9−based transcription factors. *Nat. Methods 10*, 973−976.
(17) Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., and Vale, R. D. (2014) A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell 159*, 635−646.
(18) Dominguez, A. A., Lim, W. A., and Qi, L. S. (2016) Beyond editing: repurposing CRISPR−Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol. 17*, 5−15.
(19) Dove, S. L., and Hochschild, A. (2004) A Bacterial Two-Hybrid System Based on Transcription Activation, *Protein-Protein Interactions*, pp 231−246, Humana Press, Totowa, NJ.
(20) Dove, S. L., Joung, J. K., and Hochschild, A. (1997) Activation of prokaryotic transcription through arbitrary protein−protein contacts. *Nature 386*, 627−630.
(21) Gregory, B. D., Deighan, P., and Hochschild, A. (2005) An Artificial Activator that Contacts a Normally Occluded Surface of the RNA Polymerase Holoenzyme. *J. Mol. Biol. 353*, 497−506.
(22) Ho, H., Fang, J. R., Cheung, J., and Wang, H. H. (2020) Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol. 16*, No. e9427.
(23) Dong, C., Fontana, J., Patel, A., Carothers, J. M., and Zalatan, J. G. (2018) Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat. Commun. 9*, 2489.
(24) Liu, Y., Wan, X., and Wang, B. (2019) Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun. 10*, 3693.
(25) Yan, Q., Xu, K., Xing, J., Zhang, T., Wang, X., Wei, Z., Ren, C., Liu, Z., Shao, S., and Zhang, Z. (2016) Multiplex CRISPR/Cas9-based genome engineering enhanced by Drosha-mediated sgRNA-shRNA structure. *Sci. Rep. 6*, 38970.
(26) Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., and Doudna, J. A. (2010) Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease. *Science 329*, 1355−1358.
(27) Nissim, L., Perli, S. D., Fridkin, A., Perez-Pinera, P., and Lu, T. K. (2014) Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol. Cell 54*, 698−710.
(28) Gao, Y., and Zhao, Y. (2014) Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. *J. Integr. Plant Biol. 56*, 343−349.
(29) Xie, K., Minkenberg, B., and Yang, Y. (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc. Natl. Acad. Sci. U. S. A. 112*, 3570−3575.
(30) Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E. M., Winblad, N., Choudhury, S. R., Abudayyeh, O. O., Gootenberg, J. S., Wu, W. Y., Scott, D. A., Severinov, K., van der Oost, J., and Zhang, F. (2017) Multiplex gene editing by CRISPR−Cpf1 using a single crRNA array. *Nat. Biotechnol. 35*, 31.
(31) Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., and Charpentier, E. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature 532*, 517−521.
(32) Campa, C. C., Weisbach, N. R., Santinha, A. J., Incarnato, D., and Platt, R. J. (2019) Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts. *Nat. Methods 16*, 887−893.
(33) Kim, H. K., Song, M., Lee, J., Menon, A. V., Jung, S., Kang, Y.-M., Choi, J. W., Woo, E., Koh, H. C., Nam, J.-W., and Kim, H. (2017) In vivo high-throughput profiling of CRISPR−Cpf1 activity. *Nat. Methods 14*, 153−159.
(34) Miao, C., Zhao, H., Qian, L., and Lou, C. (2019) Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells. *Synth. Syst. Biotechnol. 4*, 1−9.
(35) Hu, J. H., Miller, S. M., Geurts, M. H., Tang, W., Chen, L., Sun, N., Zeina, C. M., Gao, X., Rees, H. A., Lin, Z., and Liu, D. R. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature 556*, 57−63.
(36) Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., Welch, M. M., Horng, J. E., Malagon-Lopez, J., Scarfò, I., Maus, M. V., Pinello, L., Aryee, M. J., and Joung, J. K. (2019) Engineered CRISPR−Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol. 37*, 276−282.
(37) Zhang, X., Wang, J., Cheng, Q., Zheng, X., Zhao, G., and Wang, J. (2017) Multiplex gene regulation by CRISPR-ddCpf1. *Cell Discovery 3*, 17018.
(38) Jeong, H., Choi, S.-K., Ryu, C.-M., and Park, S.-H. (2019) Chronicle of a Soil Bacterium: *Paenibacillus polymyx*a E681 as a Tiny Guardian of Plant and Human Health. *Front. Microbiol. 10*, 467.
(39) Rütering, M., Schmid, J., Rühmann, B., Schilling, M., and Sieber, V. (2016) Controlled production of polysaccharides−exploiting nutrient supply for levan and heteropolysaccharide formation in Paenibacillus sp. *Carbohydr. Polym. 148*, 326−334.
(40) Rütering, M., Schmid, J., Gansbiller, M., Braun, A., Kleinen, J., Schilling, M., and Sieber, V. (2018) Rheological characterization of the exopolysaccharide Paenan in surfactant systems. *Carbohydr. Polym. 181*, 719−726.
(41) Rütering, M., Cress, B. F., Schilling, M., Rühmann, B., Koffas, M. A. G., Sieber, V., and Schmid, J. (2017) Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. *Synth. Biol. 2*, 007.
(42) Schmid, J. (2018) Recent insights in microbial exopolysaccharide biosynthesis and engineering strategies. *Curr. Opin. Biotechnol. 53*, 130−136.
(43) Ji, X.-J., Huang, H., and Ouyang, P.-K. (2011) Microbial 2,3-butanediol production: A state-of-the-art review. *Biotechnol. Adv. 29*, 351−364.

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

(44) Schilling, C., Ciccone, R., Sieber, V., and Schmid, J. (2020) Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365. *Metab. Eng. 61*, 381.
(45) Novotny, R., Berger, H., Schinko, T., Messner, P., Schäffer, C., and Strauss, J. (2008) A temperature-sensitive expression system based on the *Geobacillus stearothermophilus* NRS 2004/3a sgsE surface-layer gene promoter. *Biotechnol. Appl. Biochem. 49*, 35.
(46) Niu, W., Kim, Y., Tau, G., Heyduk, T., and Ebright, R. H. (1996) Transcription Activation at Class II CAP-Dependent Promoters: Two Interactions between CAP and RNA Polymerase. *Cell 87*, 1123−1134.
(47) Picossi, S., Belitsky, B. R., and Sonenshein, A. L. (2007) Molecular Mechanism of the Regulation of *Bacillus subtilis* gltAB Expression by GltC. *J. Mol. Biol. 365*, 1298−1313.
(48) Dove, S. L., and Hochschild, A. (1998) Conversion of the omega subunit of *Escherichia coli* RNA polymerase into a transcriptional activator or an activation target. *Genes Dev. 12*, 745−754.
(49) Zafar, M. A., Shah, I. M., and Wolf, R. E. (2010) Protein−Protein Interactions Between σ70 Region 4 of RNA Polymerase and Escherichia coli SoxS, a Transcription Activator That Functions by the Prerecruitment Mechanism: Evidence for "Off-DNA" and "On-DNA" Interactions. *J. Mol. Biol. 401*, 13−32.
(50) Shah, I. M., and Wolf, R. E. (2004) Novel Protein−Protein Interaction Between *Escherichia coli* SoxS and the DNA Binding Determinant of the RNA Polymerase α Subunit: SoxS Functions as a Co-sigma Factor and Redeploys RNA Polymerase from UP-element-containing Promoters to SoxS-dependent Promoters during Oxidative Stress. *J. Mol. Biol. 343*, 513−532.
(51) Umarov, R. Kh., and Solovyev, V. V. (2017) Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks. *PLoS One 12*, No. e0171410.
(52) Okonkwo, C., Ujor, V., and Ezeji, T. C. (2017) Investigation of relationship between 2,3-butanediol toxicity and production during growth of *Paenibacillus polymyxa*. *New Biotechnol. 34*, 23−31.
(53) Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods 6*, 343−345.
(54) Livak, K. J., and Schmittgen, T. D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods 25*, 402−408.
(55) Rühmann, B., Schmid, J., and Sieber, V. (2014) Fast carbohydrate analysis via liquid chromatography coupled with ultra violet and electrospray ionization ion trap detection in 96-well format. *J. Chromatogr. A 1350*, 44−50.

https://dx.doi.org/10.1021/acssynbio.0c00424
ACS Synth. Biol. 2020, 9, 3353−3363

4. Discussion

4.1. *Paenibacillus* sp. as a production organism for 2,3-BDL

The biosynthesis of of highly pure enantiomeric 2,3-BDL by *P. polymyxa*, renders this organism as a highly promising, non-pathogenic production strain [37]. Until now, the majority of efforts were focusing on medium and process optimization [37,38,43,179]. Using a previously developed tool for CRISPR-Cas9 mediated genome engineering, for the first time, metabolic engineering strategies targeting the mixed acid pathway and 2,3-BDL synthesis in *P. polymyxa* as well as further undesired by-products were applied.

Using combinatorial knock-outs of *sacB* and major parts of the paenan cluster resulted in the complete elimination of EPS during microaerobic fermentation. Due to relatively low aeration rates applied in this study, polysaccharide production was limited compared to previously reported titers [22,38]. As a result, despite a slightly increased cell dry weight, 2,3-BDL productivity remained at a similar level compared to the wildtype strain. However, advantages might be more prominent when higher aeration rates, favoring EPS production, are used during the cultivation process. Additionally, sporulation was inhibited by the deletion of key enzymes of the sporulation cascade. While deletion of *spo0A* demonstrated several downsides for 2,3-BDL production such as severely decreased productivity and enantiomeric purity, knock-out of *spoIIE* did not show any negative effects. Spo0A acts as a transcription factor and has been reported to directly or indirectly regulate the expression of more than 500 genes in *B. subtilis* [172]. Even though no binding site corresponding to Spo0A was detected in close proximity to any gene of the 2,3-BDL pathway, other regulatory proteins might be influenced by the master regulator of sporulation.

Furthermore, undesired by-products should be eliminated in order to redirect the carbon flux towards 2,3-BDL. The mixed acid pathway was engineered by deletion of *pfl* and a lactate dehydrogenase homolog (*ldh1*). Previous studies demonstrated positive effects on the production of 2,3-BDL, particularly in *Klebsiella* sp. [41,50,180]. However, using the cultivation conditions applied in this study, despite elimination of formate as a by-product, the knock-out of *pfl* highly impaired growth and consequently 2,3-BDL synthesis. Using anaerobic conditions, pyruvate formate lyase allows the synthesis of acetyl-CoA without generating additional redox equivalents. In the following pathway towards ethanol, two equivalents of NAD^+ are regenerated, allowing to counter the redox imbalance of the butanediol pathway [46]. Consequently, the knock-out might have been detrimental for *P. polymyxa* using a low aeration of 0.08 vvm. Contrarily, deletion of *ldh1* did not decrease lactate production. Three addtitional *ldh* homologs were identified in the genome. Due to the remaining expression of these genes, lactate was still produced.

A first attempt to delete the additional genes did not result in successful knock-outs and might require the testing of additional spacer sequences or extended flanking homology arms. Interestingly, the *Δldh1* showed a significantly changed growth behavior. While the wildtype strain typically started to form endospores within the first 24 h of cultivation, the knock-out strain did not produce any spores throughout the fermentation process. Furthermore, cells were significantly smaller and seemed to remain in permanent state of cell division, resulting in ~ 25 % increased biomass formation in a continuous fermentation process. The precise mechanism for this altered growth behavior remains to be investigated. Interestingly, similar effects were previously reported for other bacteria when *ldh1* homologs were knocked-out [41,181–183]. In most of these publications, increased biomass formation was explained by a reduced lactate production resulting in decreased acidification of the cultivation media. However, in the case of *P. polymyxa*, the pH value was controlled by the fermentation parameters of the bioreactor and therefore remained at a constant level. Finally, transformation of *P. polymyxa Δldh* with a plasmid constitutively expressing a butanediol dehydrogenase resulted in a highly increased productivity of 0.61 g L^{-1} h^{-1} and a titer of 43.80 g L^{-1} after 72 h cultivation. By decoupling of the final step of the 2,3-BDL pathway, initial product inhibition already observed at low 2,3-BDL concentrations could be circumvented resulting in a yield of 0.43 g g_{Glc}^{-1}, which corresponds to 86 % of the theoretical maximum.

Techno-economic assessment of current microbial fermentation processes revealed the possibility to establish an industrial production of 2,3-BDL. However, a minimal selling price of ~ 1 \$ kg^{-1}, which would be required for a bulk chemical to compete with fossil resources, is currently not conceivable [32]. Major cost factors include the high expenditure for fermentation medium, low productivity and an expensive downstream processing in order to isolate and purify the product [184,185]. For *P. polymyxa*, optimization of the media composition is still required. This should primarily focus on the substitution of expensive complex media compounds such as yeast extract and tryptone with cheaper alternatives or waste stream products. Batch cultivations reported in this study suffered from long lag phases of up to 24 h in the beginning of the fermentation. Consequently, altered inoculation strategies and increased aeration to compensate for the redox imbalance of the 2,3-BDL pathway might be suitable approaches for process optimization using *P. polymyxa* to increase productivity and product titers [22,38].

A major problem of *P. polymyxa* DSM 365 as a 2,3-BDL producing organism is the relatively low toxicity threshold of approximately 50 g L^{-1} [186]. While *P. polymyxa* still has some advantages over pathogenic *Klebsiella* or *Serratia* strains, which mainly produce the optically inactive meso-butanediol, other GRAS bacteria such as *B. licheniformis* or *B. amyloliquefaciens* have demonstrated to be competitive production platforms for 2,3-BDL [35,42]. In this study, we showed that metabolic engineering principles cannot be universally transferred between different organisms; however, it

might be possible to adapt strategies that increased 2,3-BDL production in *P. polymyxa* for closely related *Bacillus* strains. Due to the broad host range character of the used genetic tools in this study, both CRISPR systems for the generation of knock-outs or CRISPRa/i-mediated flux control could theoretically be transferred to related organisms. Furthermore, additional research is required to explore the reasons for the high solvent tolerance of other *Bacillus* spp. to allow a rational strain development of *P. polymyxa* DSM 365. Alternatively, adaptive laboratory evolution might be a suitable strategy to increase the toxicity threshold and consequently final product titers [187].

Considering these points, further work is required to make the microbial production of 2,3-BDL economically feasible. Apart from previously shown optimization of the fermentation process and downstream processing, metabolic engineering proved to be a valuable strategy to increase 2,3-BDL titers obtained by *P. polymyxa* DSM 365. The insights gained in this study might lay the foundation for further improvements of metabolic fluxes in *Paenibacillus* sp. as well as other organisms.

4.2. Quo vadis paenan? – *P. polymyxa* as an EPS producing organism

Paenibacillus spp. have been known to produce a variety of polysaccharides [67]. Among these, different isolates of *P. polymyxa* have shown to produce structurally related heteroexopolysaccharides, which often consist of similar carbohydrate building blocks of Glc, Man, Gal and GlcA [68,77,188]. Even though previous studies suggested partial structures or substructures of these EPSs, structural elucidation of the repeating unit ultimately failed due to the high complexity of the polymer [75,174]. Using a combinatorial knock-out approach by deleting essential genes of the EPS biosynthesis machinery of *P. polymyxa* DSM 365, three distinct polymers were identified (Figure 6). Furthermore, individual glycosyltransferases as well as Wzy polymerases could be specifically attributed to each EPS. Complementary carbohydrate analytical methods, e.g. the HT-PMP method, partial hydrolysis, specific degradation of uronic acid residues, linkage pattern analysis via different methylation methods and NMR analysis, were required to uncover the composition, sequence and linkage patterns of each polymer. Although each of these methods provides significant data on their own, a combinatorial approach was necessary to acquire a holistic, cross-validated picture of each repeating unit and elucidate their polymer structure.

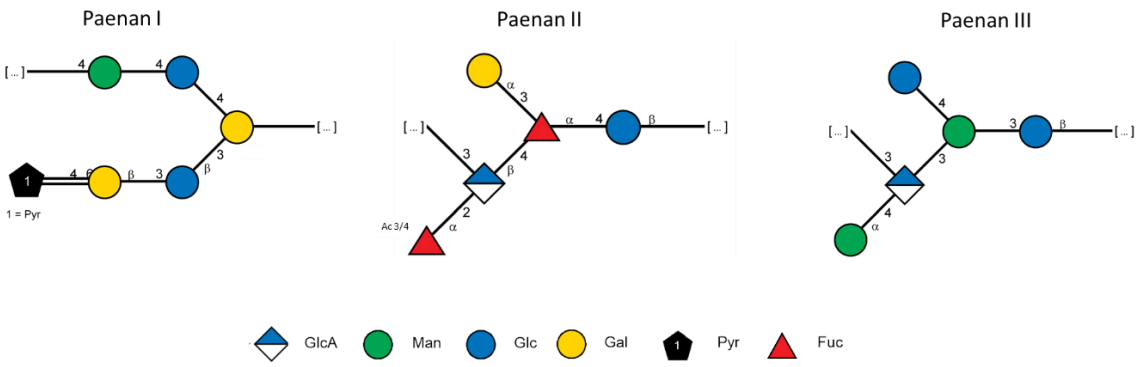


Figure 6: Postulated polymer structures of the three heteroexopolysaccharides produced by *P. polymyxa* DSM 365. Individual glycosyltransferases of the paenan cluster (Pep) were assigned to the assembly of specific polymers. Initiating glycosyltransferases (PepC and PepQ) seem to share the same function by transferring a glucose residue to a lipid anchor; however, deletions resulted in different EPS compositions. For paenan I, two additional glycosyltransferases putatively located in a cluster next to the corresponding pyruvyltransferase EpsO remain to be identified. Figure from [189].

Using CRISPR-Cas9 mediated gene deletions of GTs, it proved possible to attribute distinct glycosyltransferases to the assembly of specific repeating units (Figure 6). However, most gene deletions resulted in the complete loss of the ability to produce the respective paenan variant. Interestingly, our current research of the biosynthesis of

paenan polymers produced by *P. polymyxa* DSM 365 indicates that both initiating glycosyltransferases PepC and PepQ encode glucosyltransferses. Nevertheless, individual knock-outs result in distinct EPS compositions without the secondary initiating GT being able to complement the missing function. It is currently speculated that the deletion of the open reading frame of *pepC* and *pepQ* also effects the expression of additional GTs downstream in the operon resulting in the absence of specific paenan variants. Further investigation of the transcriptome of the Δ*pepC* and Δ*pepQ* strain is required in order to validate this hypothesis.

It has been demonstrated that Wzx flippases are highly sensitive towards the structure of their natural repeating unit [94]. While minor changes in side chains are often accepted by the flippase, deletion of the GT corresponding to the initial branching glycosylresidue and consequently loss of the side chain are often detrimental to the function of flippases [94]. Interestingly, deletion of *pepT*, *pepU* and *pepV* result in the absence of the fucose containing polysaccharide paenan II. However, the product yield of the variant *ΔpepV* is significantly lower in comparison to the other glycosyltransferase knock-outs. This observation might be explained by the putative function of PepV to link the terminal galactose residue in paenan II. Subsequently, its knock-out should result in the assembly of repeating units without a side chain, which are not recognized by the highly specific corresponding Wzx flippase and are thus not exported. To this end, the backbone of the repeating unit of paenan II would block undecaprenylphosphate lipid anchors for the biosynthesis of paenan I and paenan III, which could be an explanation for the reduced EPS yield of the *ΔpepV* variant. In order to fully understand the precise function of individual glycosyltransferases within the paenan cluster, a reasonable approach would include the heterologous expression of GTs with known enzymatic functions to complement missing enzyme activities in knock-out variants. If the assembly of certain paenan variants can be restored in a sufficient manner, this strategy might uncover specific activities of the GTs of *P. polymyxa* DSM 365.

For any application of bacterial EPS, a major criterion is its rheological properties. Consequently, the wildtype EPS composition as well as each paenan variant and mixture were rheologically characterized in detail. While the wildtype strain produces a highly viscous, shear-thinning polymer solution with a pronounced gel character, individual paenan variants demonstrated extremely low viscosity and high damping factors. Furthermore, the presence of paenan II in the polymer compositions is crucial for the temperature stability of the polymer gel. It is hypothesized that this indicates a strong intermolecular network between individual paenan variants that is responsible for the physicochemical properties of the wildtype polymer. In comparison to other commercially used bacterial polysaccharides (e.g. xanthan) paenan shows significantly higher viscosity at low shear rates, but demonstrates a less rigid gel character than gellan [190,191]. These properties make the wildtype polymer an interesting candidate as a

viscosifier in personal care products and cosmetics that could replace synthetic, acrylic co-polymers such as Carbopol® or Aristoflex® [192,193]. The low viscosity and the loss of the gel character of engineered paenan variants make them promising candidates for film formation and surface coatings. Similar polysaccharides such as pullulan from *Aureobasidium pullulans* are used in edible films or packaging materials [194–196]. For pharmaceutical applications, polysaccharide-based coatings are increasingly used in various controlled drug release systems due to anti-inflammatory effects, low immunogenic response and their biodegradability [23,197]. Therefore, the EPS produced by *P. polymyxa* DSM 365 and engineered variants thereof might be a green alternative to fossil-based polymers in a broad range of applications.

Recently, a structurally related EPS from *P. polymyxa* 2H2 demonstrated to be compatible with most commonly used anionic and cationic surfactants included in toiletry (e.g. sodium lauryl sulfate, cetrimonium chloride) [24]. With the rheological data of different paenan variants presented in this study, it is possible to tailor cosmetic formulations towards specific rheological requirements. To rationally predict interactions with other compounds of interest, further research regarding the secondary, tertiary and quaternary structure of paenan variants and their intermolecular network might be required to understand the positioning of side chains and exposure of functional groups. For xanthan, sophisticated NMR methods and atom force microscopy demonstrated that the polymer forms a double stranded helical structure [198–200]. Depending on the degree of acetylation and pyruvylation of the side chain, the EPS takes different conformations resulting in different rheological properties [66,201]. Similar investigations might be necessary for paenan to fully understand its rheological behavior.

Currently, a major bottleneck for any commercial application of the EPS from *P. polymyxa* lies in the operation costs of production and obtained product yields. While it was not the aim of this study to optimize EPS production of this organism, it is an important biotechnological aspect that should be considered. Without any process and media optimization, the wildtype strain produced 5 g L^{-1} EPS corresponding to a yield of 0.135 g g_{Glc}^{-1}. However, knock-out variants produced significantly less EPS ranging from 2-3 g L^{-1} and proportionally reduced yields. In comparison, for the small segment of microbial polymers that currently persist on the market, established fermentation processes resulted in remarkably higher product titers. For xanthan, industrial fed-batch cultivations result in 30-60 g L^{-1} EPS with a yield of 0.82 g g^{-1} glucose [60,202,203]. For the fermentative production of gellan using *Sphingomonas paucimobilis*, EPS titers of up to 35.7 g L^{-1} were demonstrated [204]. A major limitation for the production of EPS seems to be the high viscosity of the fermentation broth, which ultimately prevents mass and oxygen transfer within the bioreactor. In comparison, for polysaccharides with a low viscosity such as dextran or levan, production titers of 150 g L^{-1} by microbial fermentation have been reported [205]. Consequently, process and media optimization might be

suitable approaches to increase EPS titers obtained by *P. polymyxa* and knock-out variants. Due to the limited number of genetic tools for alternative host organisms, metabolic engineering efforts of EPS producing organisms have been scarcely reported. Using the tools developed for *P. polymyxa* [25,206], flux optimization via CRISPRa/i, overexpression of glycosyltransferases or knock-outs of competing by-products might be used to significantly enhance EPS production.

Most commercially available bacterial polysaccharides are regarded as bulk products with prices ranging from 3-30 \$ kg^{-1} [60,207]. Depending on the final application, some polymers also fill high-value niches and are utilized in cosmetic and pharmaceutical products. Despite low product titers of less than 10 g L^{-1}, microbial production of hyaluronic acid is still economically profitable due to high prices of up to 100 \$ mg^{-1} for pure product [113]. Depending on the final application, downstream processing can be modified for different paenan variants depending on the required purity for the final application and thereby reducing the operation costs [208]. The profit margin can be further expanded by a precise adjustment of the physicochemical properties of engineered paenan variants to their application. A major advantage of the EPS produced by *P. polymyxa* DSM 365 might be that process optimization and metabolic engineering strategies can putatively be transferred between different paenan variants.

4.3. CRISPRa/i mediated metabolic flux optimization

CRISPR-Cas mediated genome engineering advanced synthetic biology and metabolic engineering in an unprecedented way by enabling fast and precise genome editing in numerous organisms [134,209]. While CRISPRi is commonly used for bacteria to modulate the expression of target genes, CRISPRa is lacking behind. Transcriptional perturbation has been applied for genome screenings in eukaryotes to discover new biological mechanisms or to identify and leverage the mode of action of drugs, toxins or pathogens [210–212]. The few systems that have been developed for bacteria so far typically suffer from a limited range of promoters that can be targeted or a narrow range in which transcriptional activation can be achieved [150,154,155]. Therefore, this work also focused on the development of a novel prokaryotic tool for both, CRISPRi and CRISPRa, using dCas12a to facilitate processing of the CRISPR-array and enable multiplexing of numerous genes.

Using the regulator of the superoxide regulon of *E. coli* as an activator domain, the here developed tool was successfully used to activate or repress the expression of fluorescence reporter proteins in both, *E. coli* S17-1 as well as *P. polymyxa* DSM 365. Furthermore, initial glycosyltransferases of the paenan cluster were repressed, effectively producing the same EPS composition as corresponding knock-out variants. For the butanediol biosynthesis pathway, the tool was used to simultaneously knock-down the expression of three lactate dehydrogenases, while activating the expression of the endogenous butanediol dehydrogenase. However, only mediocre effects could be observed resulting in a 20 % reduced lactate concentration and a 26 % increased butanediol titer compared to an off-target construct. A major problem might lie in endogenous regulation mechanisms of *P. polymyxa* DSM 365. Already in initial knock-out mediated engineering strategies it was demonstrated that *P. polymyxa* increases redox-neutral lactate production in response to elevated 2,3-BDL levels [213]. Consequently, low effects of transcriptional perturbation might be a result of undesired cross-talk with endogenous regulatory mechanisms of the mixed acid pathway. Nevertheless, even though a proof of principle could be demonstrated on episomal and genomic targets, the dynamic output range remained low compared to previous systems targeting δ^{54}- dependent promoters [154]. Interestingly, efficacy of repression varied significantly between different spacer sequences. Despite utilizing same PAM sequences, some spacers reduced GFP signal by 90 %, where other gRNAs target sites in close proximity only reduced the fluorescence output by 50 %. With our current knowledge, it is not possible to design efficient spacer sequences *a priori* and empirical testing is required for each target. Chosen PAM sequences (TTTV) did not seem to affect the efficiency of gRNA in this study. Nevertheless, it might be interesting to characterize the seed sequences of each spacer (5-10 bases next to the PAM). Typically, for knock-out experiments, seed sequences are of lower importance. Even mismatches within the

seed sequence have demonstrated to result in efficient cleavage at the target sites [214]. However, for transcriptional perturbation using catalytically dead Cas nucleases, seed sequences can highly affect the binding efficiency of the activated Cas module and consequently the robustness of activation or repression [133,215]. It might be possible to intentionally exploit suboptimal spacers or mismatches within the seed region to tune the binding affinity and consequently the dynamic range in order to modulate the expression of target genes of a distinct pathway. For this approach, large datasets of spacer sequences have to be characterized in detail to generate rule sets on spacer design with predictable outcome.

Furthermore, particularly for CRISPRi, it might be interesting to investigate repression efficiency of the developed pCRai system, when the corresponding promoter region is targeted directly. Transcriptional repression might be improved by using the activated Cas module as a repressor that prevents binding of the RNA-polymerase rather than a roadblock within the ORF that blocks elongation of the transcript. Additionally, protein engineering might be applied to mutate both, Cas12a to increase binding efficiency to its cognate target site and the SoxS activator domain to improve the recruitment of RNA-polymerase. Similar strategies have previously demonstrated to improve the efficacy of CRISPR mediated transcriptional perturbation by increasing the dynamic output range [144,155].

There is still room for improvement of the developed pCRai tool for transcriptional perturbation in this study. However, proof of concept was established in *P. polymyxa* by modulating the expression of fluorescence reporters, glycosyltransferases and butanediol synthesis. Due to the design of the plasmid including commonly recognized origins of replication for both Gram-negative and Gram-positive bacteria as well as an origin of transfer enabling conjugation, the system might be suitable for a range of different organisms for which molecular tools and methods have not been developed yet. Consequently, the CRISPRa/i tool could be used to enable complex metabolic engineering efforts in otherwise genetically inaccessible host organism.

5. Conclusion

Over the last decades, the awareness of the limited nature of resources of our planet has been steadily rising and the demand for sustainable alternatives to fossil-based compounds is rapidly increasing. This study focused on the metabolic engineering of the alternative host *P. polymyxa* DSM 365 to establish a chassis organism for the production of 2,3-BDL and EPS with unique physicochemical properties.

Using a previously developed tool for CRISPR-Cas9 mediated genome engineering in *P. polymyxa*, the butanediol biosynthesis and the mixed-acid pathway were engineered to increase the productivity and yield. Deletion of a lactate dehydrogenase surprisingly resulted in an improved growth behaviour of the organism. By decoupling and overexpression of the last enzymatic step of the butanediol pathway 2,3-BDL titers were increased to 43.8 g L^{-1} after 72 h cultivation. A yield of 0.43 g g_{Glc}^{-1} was obtained, corresponding to 86 % of the theoretical maximum. Furthermore, undesired side products that disturb downstream processing such as EPS and spore formation were eliminated without negatively affecting 2,3-BDL synthesis. Even though final product titers could be increased by 45 %, the current engineered strain is not competitive to other biotechnological producers such as pathogenic *Klebsiella* spp., *Serratia* spp. or non-pathogenic *Bacillus* spp. The major bottleneck lies in the low toxicity threshold of *P. polymyxa* towards the final product. While this problem might be overcome by laboratory evolution of the optimized strain, we also propose the transfer of the applied engineering strategies to closely related *Bacillus* spp., which have demonstrated significantly increased 2,3-BDL tolerance.

Furthermore, *P. polymyxa* DSM 365 was also used as a production platform for heteroexopolysaccharides with unique physicochemical properties. Using combinatorial knock-out variants of glycosyltransferases, it was elucidated that the EPS of *P. polymyxa* DSM 365 in fact consists of three distinct biopolymers named Paenan I-III. Due to the complexity of polymers and EPS mixtures, complementary analytical methods were required to identify the carbohydrate composition, sequence and linkage type of each polysaccharide. Using this structural knowledge, it might be possible to rationally predict interactions and compatibility with typical compounds found in cosmetics, food and technical applications as well as other commodity items in which polysaccharides are applied as rheological modifiers. Individual paenan variants and EPS mixtures produced by knock-out strains of *P. polymyxa* were rheologically characterized in detail. While the wildtype strain produced a highly viscous, gel-forming polymer with a pronounced intramolecular network, individual paenan variants demonstrated significantly reduced viscoelastic properties. Particularly the interaction of pyruvate and GlcA found in paenan III seems to be crucial for the gel character of the produced EPS. As a result, different paenan variants and combination thereof can be applied in a wide range of

applications. Precise knowledge of the function of individual enzymes involved in the biosynthesis of paenan might enable tailoring of artificial EPS compositions with distinct material properties that allow the utilization in completely new fields of application. The portfolio of EPS variants produced by *P. polymyxa* DSM 365 could be further expanded by the expression of heterologous GTs to further alter the polymer structures.

The established tool for CRISPR mediated transcriptional perturbation provides a powerful method for genome scale, multiplex metabolic engineering efforts that allow the simultaneous activation and repression of genes of interest. Even though improvements to increase the dynamic range of the tool and empirical testing of optimized target sites are still required, a proof of principle could be established in both, *P. polymyxa* and *E. coli*. On the example of the EPS composition of *P. polymyxa* DSM 365, it was demonstrated that time-consuming knock-out variants can be efficiently emulated by transcriptional repression of the target genes. However, on the example of 2,3-BDL synthesis, it was also shown that endogenous regulatory systems might also influence the efficacy of the tool. Consequently, possible target genes have to be chosen carefully and more empirical testing is required to establish a robust design rule set. Owing to the broad-host range character of the plasmid, the developed CRISPRa/i tool might be used to optimize metabolic fluxes in numerous organisms other that *Paenibacillus* sp..

In this work, *P. polymyxa* was used as a platform organism for the biosynthesis of levo-butanediol and exopolysaccharide based rheological modifiers. Metabolic engineering allowed to improve 2,3-BDL synthesis and enabled the structural elucidation of biopolymers produced by *P. polymyxa* DSM 365. Furthermore, a broad host range tool for transcriptional activation and repression in bacteria was developed. The target products discussed in this work are only a small part of the wide range of potentially interesting biotechnological products and applications of *P. polymyxa*. Further research focusing on the vast variety of antibiotics, *Paenibacillus* sp. as an agricultural biofertilizer or pathogenic strains such as *P. larvae* seem obvious. Desirably, the findings of this study will contribute to make new biotechnological processes using *Paenibacillus* sp. but also other organisms economically feasible and push society further towards a sustainable future independent of fossil-based resources.

6. References

1. Grady EN, MacDonald J, Liu L, Richman A, Yuan Z-C: **Current knowledge and perspectives of *Paenibacillus*: a review**. *Microb Cell Factories* 2016, **15**:203.
2. Ash C, Priest FG, Collins MD: **Molecular identification of rRNA group 3 Bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test: Proposal for the creation of a new genus *Paenibacillus***. *Antonie Van Leeuwenhoek* 1994, **64**:253–260.
3. Ash C, Farrow JAE, Wallbanks S, Collins MD: **Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequences**. *Lett Appl Microbiol* 2008, **13**:202–206.
4. Shida O, Takagi H, Kadowaki K, Nakamura LK, Komagata K: **Transfer of *Bacillus alginolyticus, Bacillus chondroitinus, Bacillus curdlanolyticus, Bacillus glucanolyticus, Bacillus kobensis,* and *Bacillus thiaminolyticus* to the Genus *Paenibacillus* and Emended Description of the Genus *Paenibacillus***. *Int J Syst Bacteriol* 1997, **47**:289–298.
5. Timmusk S, Wagner EG: **The plant-growth-promoting rhizobacterium *Paenibacillus polymyxa induces* changes in*Arabidopsis thaliana* gene expression: a possible connection between biotic and abiotic stress responses**. *Mol Plant-Microbe Interact MPMI* 1999, **12**:951–959.
6. Xie J-B, Du Z, Bai L, Tian C, Zhang Y, Xie J-Y, Wang T, Liu X, Chen X, Cheng Q, et al.: **Comparative Genomic Analysis of N2-Fixing and Non-N2-Fixing *Paenibacillus* spp.: Organization, Evolution and Expression of the Nitrogen Fixation Genes**. *PLoS Genet* 2014, **10**:e1004231.
7. Xie J, Shi H, Du Z, Wang T, Liu X, Chen S: **Comparative genomic and functional analysis reveal conservation of plant growth promoting traits in *Paenibacillus polymyxa* and its closely related species**. *Sci Rep* 2016, **6**:21329.
8. Jeong H, Choi S-K, Ryu C-M, Park S-H: **Chronicle of a Soil Bacterium: *Paenibacillus polymyx*a E681 as a Tiny Guardian of Plant and Human Health**. *Front Microbiol* 2019, **10**:467.
9. Yokoyama T, Tanaka M, Hasegawa M: **Novel cry gene from *Paenibacillus lentimorbus* strain Semadara inhibits ingestion and promotes insecticidal activity in *Anomala cuprea* larvae**. *J Invertebr Pathol* 2004, **85**:25–32.
10. Bach E, Seger GD dos S, Fernandes G de C, Lisboa BB, Passaglia LMP: **Evaluation of biological control and rhizosphere competence of plant growth promoting bacteria**. *Appl Soil Ecol* 2016, **99**:141–149.
11. Velkov T, Thompson PE, Nation RL, Li J: **Structure−Activity Relationships of Polymyxin Antibiotics**. *J Med Chem* 2010, **53**:1898–1916.
12. Velkov T, Roberts KD, Nation RL, Wang J, Thompson PE, Li J: **Teaching 'Old' Polymyxins New Tricks: New-Generation Lipopeptides Targeting Gram-Negative 'Superbugs.'** *ACS Chem Biol* 2014, **9**:1172–1177.

13. Cochrane SA, Findlay B, Bakhtiary A, Acedo JZ, Rodriguez-Lopez EM, Mercier P, Vederas JC: **Antimicrobial lipopeptide tridecaptin A $_1$ selectively binds to Gram-negative lipid II**. *Proc Natl Acad Sci* 2016, **113**:11561–11566.

14. Cochrane SA, Vederas JC: **Lipopeptides from *Bacillus* and *Paenibacillus* spp.: A Gold Mine of Antibiotic Candidates: BACILLUS AND PAENIBACILLUS LIPOPEPTIDES**. *Med Res Rev* 2016, **36**:4–31.

15. He Z, Kisla D, Zhang L, Yuan C, Green-Church KB, Yousef AE: **Isolation and Identification of a *Paenibacillus polymyxa* Strain That Coproduces a Novel Lantibiotic and Polymyxin**. *Appl Environ Microbiol* 2007, **73**:168–178.

16. Park J-E, Kim H-R, Park S-Y, Choi S-K, Park S-H: **Identification of the biosynthesis gene cluster for the novel lantibiotic paenilan from *Paenibacillus polymyxa* E681 and characterization of its product**. *J Appl Microbiol* 2017, **123**:1133–1147.

17. Kajimura Y, Kaneda M: **Fusaricidin A, a New Depsipeptide Antibiotic Produced by *Bacillus polymyxa* KT-8 Taxonomy, Fermentation, Isolation, Structure Elucidation and Biological Activity.** *J Antibiot (Tokyo)* 1996, **49**:129–135.

18. Garcia-Gonzalez E, Genersch E: **Honey bee larval peritrophic matrix degradation during infection with *Paenibacillus larvae* , the aetiological agent of American foulbrood of honey bees, is a key step in pathogenesis: Peritrophic matrix degradation through *P. larvae***. *Environ Microbiol* 2013, doi:10.1111/1462-2920.12167.

19. Djukic M, Brzuszkiewicz E, Fünfhaus A, Voss J, Gollnow K, Poppinga L, Liesegang H, Garcia-Gonzalez E, Genersch E, Daniel R: **How to Kill the Honey Bee Larva: Genomic Potential and Virulence Mechanisms of *Paenibacillus* larvae**. *PLoS ONE* 2014, **9**:e90914.

20. Gopal N, Hill C, Ross PR, Beresford TP, Fenelon MA, Cotter PD: **The Prevalence and Control of *Bacillus* and Related Spore-Forming Bacteria in the Dairy Industry**. *Front Microbiol* 2015, **6**.

21. Ranieri ML, Boor KJ: **Short communication: Bacterial ecology of high-temperature, short-time pasteurized milk processed in the United States**. *J Dairy Sci* 2009, **92**:4833–4840.

22. Häßler T, Schieder D, Pfaller R, Faulstich M, Sieber V: **Enhanced fed-batch fermentation of 2,3-butanediol by *Paenibacillus polymyxa* DSM 365**. *Bioresour Technol* 2012, **124**:237–244.

23. Liu J, Luo J, Ye H, Sun Y, Lu Z, Zeng X: **In vitro and in vivo antioxidant activity of exopolysaccharides from endophytic bacterium *Paenibacillus polymyxa* EJS-3**. *Carbohydr Polym* 2010, **82**:1278–1283.

24. Rütering M, Schmid J, Gansbiller M, Braun A, Kleinen J, Schilling M, Sieber V: **Rheological characterization of the exopolysaccharide Paenan in surfactant systems**. *Carbohydr Polym* 2018, **181**:719–726.

25. Rütering M, Cress BF, Schilling M, Rühmann B, Koffas MAG, Sieber V, Schmid J: **Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa***. *Synth Biol* 2017, **2**.

26. Fulmer EI, Christensen LM, Kendali AR: **Production of 2,3-Butylene Glycol by Fermentation**. *Ind Eng Chem* 1933, **25**:798–800.

27. Köpke M, Mihalcea C, Liew F, Tizard JH, Ali MS, Conolly JJ, Al-Sinawi B, Simpson SD: **2,3-Butanediol Production by Acetogenic Bacteria, an Alternative Route to Chemical Synthesis, Using Industrial Waste Gas**. *Appl Environ Microbiol* 2011, **77**:5467–5475.

28. Celińska E, Grajek W: **Biotechnological production of 2,3-butanediol—Current state and prospects**. *Biotechnol Adv* 2009, **27**:715–725.

29. Isagulyants GV, Belomestnykh IP: **Butadione synthesis by dehydrogenation and oxidative dehydrogenation of 2,3-butanediol**. In *Studies in Surface Science and Catalysis*. . Elsevier; 1997:415–420.

30. Voloch M, Jansen NB, Ladisch MR, Tsao GT, Narayan R, Rodwell VW: **2,3 Butanediol**. In *Moo-Young M. (ed), Comprehensive biotechnology: the principles, applications and regulations of biotechnology in industry, agriculture and medicine*. . Pergamon Press; 1985:933–947.

31. Clendenning KA, Wright DE: **PRODUCTION AND PROPERTIES OF 2,3-BUTANEDIOL: XII. ANTIFREEZE PROPERTIES OF TERNARY AQUEOUS SOLUTIONS CONTAINING *LEVO* -2,3-BUTANEDIOL AS A MAJOR COMPONENT**. *Can J Res* 1946, **24f**:287–299.

32. Koutinas AA, Yepez B, Kopsahelis N, Freire DMG, de Castro AM, Papanikolaou S, Kookos IK: **Techno-economic evaluation of a complete bioprocess for 2,3-butanediol production from renewable resources**. *Bioresour Technol* 2016, **204**:55–64.

33. Ma C, Wang A, Qin J, Li L, Ai X, Jiang T, Tang H, Xu P: **Enhanced 2,3-butanediol production by *Klebsiella pneumoniae* SDM**. *Appl Microbiol Biotechnol* 2009, **82**:49–57.

34. Zhang L, Yang Y, Sun J, Shen Y, Wei D, Zhu J, Chu J: **Microbial production of 2,3-butanediol by a mutagenized strain of *Serratia marcescens* H30**. *Bioresour Technol* 2010, **101**:1961–1967.

35. Jurchescu I-M, Hamann J, Zhou X, Ortmann T, Kuenz A, Prüße U, Lang S: **Enhanced 2,3-butanediol production in fed-batch cultures of free and immobilized *Bacillus licheniformis* DSM 8785**. *Appl Microbiol Biotechnol* 2013, **97**:6715–6723.

36. Yang T, Rao Z, Zhang X, Lin Q, Xia H, Xu Z, Yang S: **Production of 2,3-butanediol from glucose by GRAS microorganism *Bacillus amyloliquefaciens***. *J Basic Microbiol* 2011, **51**:650–658.

37. Nakashimada Y, Kanai K, Nishio N: **Optimization of dilution rate, pH and oxygen supply on optical purity of 2, 3-butanediol produced by *Paenibacillus polymyxa* in chemostat culture**. *Biotechnol Lett* 1998, **20**:1133–1138.

38. Okonkwo C, Ujor V, Mishra P, Ezeji T: **Process Development for Enhanced 2,3-Butanediol Production by *Paenibacillus polymyxa* DSM 365**. *Fermentation* 2017, **3**:18.

39. Nie Z-K, Ji X-J, Huang H, Du J, Li Z-Y, Qu L, Zhang Q, Ouyang P-K: **An Effective and Simplified Fed-Batch Strategy for Improved 2,3-Butanediol Production by *Klebsiella oxytoca***. *Appl Biochem Biotechnol* 2011, **163**:946–953.

40. Ji X-J, Huang H, Du J, Zhu J-G, Ren L-J, Hu N, Li S: **Enhanced 2,3-butanediol production by *Klebsiella oxytoca* using a two-stage agitation speed control strategy**. *Bioresour Technol* 2009, **100**:3410–3414.

41. Jung M-Y, Ng CY, Song H, Lee J, Oh M-K: **Deletion of lactate dehydrogenase in *Enterobacter aerogenes* to enhance 2,3-butanediol production**. *Appl Microbiol Biotechnol* 2012, **95**:461–469.

42. Yang T, Rao Z, Zhang X, Xu M, Xu Z, Yang S-T: **Improved Production of 2,3-Butanediol in *Bacillus amyloliquefaciens* by Over-Expression of Glyceraldehyde-3-Phosphate Dehydrogenase and 2,3-butanediol Dehydrogenase**. *PLoS ONE* 2013, **8**:e76149.

43. De Mas C, Jansen NB, Tsao GT: **Production of optically active 2,3-butanediol by *Bacillus polymyxa***. *Biotechnol Bioeng* 1988, **31**:366–377.

44. Gao J, Xu H, Li Q, Feng X, Li S: **Optimization of medium for one-step fermentation of inulin extract from Jerusalem artichoke tubers using *Paenibacillus polymyxa* ZJ-9 to produce R,R-2,3-butanediol**. *Bioresour Technol* 2010, **101**:7076–7082.

45. Clark DP: **The fermentation pathways of *Escherichia coli***. *FEMS Microbiol Lett* 1989, **63**:223–234.

46. Alexeeva S, Hellingwerf KJ, Teixeira de Mattos MJ: **Requirement of ArcA for Redox Regulation in*Escherichia coli*under Microaerobic but Not Anaerobic or Aerobic Conditions**. *J Bacteriol* 2003, **185**:204–209.

47. Adlakha N, Pfau T, Ebenhöh O, Yazdani SS: **Insight into metabolic pathways of the potential biofuel producer, *Paenibacillus polymyxa* ICGEB2008**. *Biotechnol Biofuels* 2015, **8**:159.

48. Van Houdt R, Aertsen A, Michiels CW: **Quorum-sensing-dependent switch to butanediol fermentation prevents lethal medium acidification in *Aeromonas hydrophila* AH-1N**. *Res Microbiol* 2007, **158**:379–385.

49. Yu EKC, Saddler JN: **Biomass conversion to butanediol by simultaneous saccharification and fermentation**. *Trends Biotechnol* 1985, **3**:100–104.

50. Jung M-Y, Mazumdar S, Shin SH, Yang K-S, Lee J, Oh M-K: **Improvement of 2,3-Butanediol Yield in *Klebsiella pneumoniae* by Deletion of the Pyruvate Formate-Lyase Gene**. *Appl Environ Microbiol* 2014, **80**:6195–6203.

51. Jantama K, Polyiam P, Khunnonkwao P, Chan S, Sangproo M, Khor K, Jantama SS, Kanchanatawee S: **Efficient reduction of the formation of by-products and improvement of production yield of 2,3-butanediol by a combined deletion of alcohol dehydrogenase, acetate kinase-phosphotransacetylase,**

and lactate dehydrogenase genes in metabolically engineered *Klebsiella oxytoca* in mineral salts medium. *Metab Eng* 2015, **30**:16–26.

52. Mallonee DH, Speckman RA: **Development of a Mutant Strain of *Bacillus polymyxa* Showing Enhanced Production of 2,3-Butanediol**. *Appl Environ Microbiol* 1988, **54**:168–171.
53. Zhang L, Cao C, Jiang R, Xu H, Xue F, Huang W, Ni H, Gao J: **Production of R,R-2,3-butanediol of ultra-high optical purity from *Paenibacillus polymyxa* ZJ-9 using homologous recombination**. *Bioresour Technol* 2018, **261**:272–278.
54. Dai X, Gao G, Wu M, Wei W, Qu J, Li G, Ma T: **Construction and application of a *Xanthomonas campestris* CGMCC 15155 strain that produces white xanthan gum**. *MicrobiologyOpen* 2019, **8**:e00631.
55. Jeong J, Park K, Chung M, Won J-I: **Influence of *Vitreoscilla* hemoglobin gene expression on 2,3-butanediol production in *Klebsiella oxytoca***. *Biotechnol Bioprocess Eng* 2015, **20**:10–17.
56. Geckil H, Barak Z, Chipman DM, Erenler SO, Webster DA, Stark BC: **Enhanced production of acetoin and butanediol in recombinant Enterobacter aerogenes carrying Vitreoscilla hemoglobin gene**. *Bioprocess Biosyst Eng* 2004, **26**:325–330.
57. Moradali MF, Rehm BHA: **Bacterial biopolymers: from pathogenesis to advanced materials**. *Nat Rev Microbiol* 2020, doi:10.1038/s41579-019-0313-3.
58. Schmid J, Sieber V: **The Bacterial Glycome: From Monomers to Complex Carbohydrate Polymers**. In *Reference Module in Life Sciences*. . Elsevier; 2019:B978012809633890774X.
59. Cress BF, Englaender JA, He W, Kasper D, Linhardt RJ, Koffas MAG: **Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules**. *FEMS Microbiol Rev* 2014, **38**:660–697.
60. Freitas F, Alves VD, Reis MAM: **Advances in bacterial exopolysaccharides: from production to biotechnological applications**. *Trends Biotechnol* 2011, **29**:388–398.
61. *Essentials of Carbohydrate Chemistry*. Springer New York; 1998.
62. Williams PA, Phillips GO (Eds): *Gums and stabilisers for the food industry 18: hydrocolloid functionality for affordable and sustainable global food solutions*. Royal Society of Chemistry; 2016.
63. Schmid J, Sperl N, Sieber V: **A comparison of genes involved in sphingan biosynthesis brought up to date**. *Appl Microbiol Biotechnol* 2014, **98**:7719–7733.
64. Sze JH, Brownlie JC, Love CA: **Biotechnological production of hyaluronic acid: a mini review**. *3 Biotech* 2016, **6**:67.
65. Schmid J: **Recent insights in microbial exopolysaccharide biosynthesis and engineering strategies**. *Curr Opin Biotechnol* 2018, **53**:130–136.
66. Gansbiller M, Schmid J, Sieber V: **In-depth rheological characterization of genetically modified xanthan-variants**. *Carbohydr Polym* 2019, **213**:236–246.

67. Liang T-W, Wang S-L: **Recent advances in exopolysaccharides from *Paenibacillus* spp.: production, isolation, structure, and bioactivities**. *Mar Drugs* 2015, **13**:1847–1863.

68. Ninomiya E, Kizaki T: **Bacterial polysaccharide from *Bacillus polymyxa* No. 271**. *Angew Makromol Chem* 1969, **6**:179–185.

69. Rütering M, Schmid J, Rühmann B, Schilling M, Sieber V: **Controlled production of polysaccharides–exploiting nutrient supply for levan and heteropolysaccharide formation in *Paenibacillus sp.*** *Carbohydr Polym* 2016, **148**:326–334.

70. Hundschell C, Braun A, Wefers D, Vogel R, Jakob F: **Size-Dependent Variability in Flow and Viscoelastic Behavior of Levan Produced by *Gluconobacter albidus* TMW 2.1191**. *Foods* 2020, **9**:192.

71. Morillo JA, Guerra Del Águila V, Aguilera M, Ramos-Cormenzana A, Monteoliva-Sánchez M: **Production and characterization of the exopolysaccharide produced by *Paenibacillus jamilae* grown on olive mill-waste waters**. *World J Microbiol Biotechnol* 2007, **23**:1705.

72. Aguilera M, Monteoliva-Sánchez M, Suárez A, Guerra V, Lizama C, Bennasar A, Ramos-Cormenzana A: ***Paenibacillus jamilae* sp. nov., an exopolysaccharide-producing bacterium able to grow in olive-mill wastewater**. *Int J Syst Evol Microbiol* 2001, **51**:1687–1692.

73. Jung H-K, Hong J-H, Park S-C, Park B-K, Nam D-H, Kim S-D: **Production and physicochemical characterization of β-glucan produced by *Paenibacillus polymyxa* JB115**. *Biotechnol Bioprocess Eng* 2007, **12**:713–719.

74. Kahng G-G, Lim S-H, Yun H-D, Seo W-T: **Production of extracellular polysaccharide, EPS WN9, from *Paenibacillus* sp. WN9 KCTC 8951P and its usefulness as a cement mortar admixture**. *Biotechnol Bioprocess Eng* 2001, **6**:112–116.

75. Madden JK, Dea ICM, Steer DC: **Structural and rheological properties of the extracellular polysaccharides from *Bacillus polymyxa***. *Carbohydr Polym* 1986, **6**:51–73.

76. Yegorenkova IV, Tregubova KV, Matora LY, Burygin GL, Ignatov VV: **Biofilm formation by *Paenibacillus polymyxa* strains differing in the production and rheological properties of their exopolysaccharides**. *Curr Microbiol* 2011, **62**:1554–1559.

77. Raza W, Makeen K, Wang Y, Xu Y, Qirong S: **Optimization, purification, characterization and antioxidant activity of an extracellular polysaccharide produced by *Paenibacillus polymyxa* SQR-21**. *Bioresour Technol* 2011, **102**:6095–6103.

78. Wang C-L, Huang T-H, Liang T-W, Fang C-Y, Wang S-L: **Production and characterization of exopolysaccharides and antioxidant from *Paenibacillus* sp. TKU023**. *New Biotechnol* 2011, **28**:559–565.

79. Liu J, Luo J, Ye H, Zeng X: **Preparation, antioxidant and antitumor activities in vitro of different derivatives of levan from endophytic bacterium**

***Paenibacillus polymyxa* EJS-3**. *Food Chem Toxicol Int J Publ Br Ind Biol Res Assoc* 2012, **50**:767–772.

80. Liang T-W, Wu C-C, Cheng W-T, Chen Y-C, Wang C-L, Wang I-L, Wang S-L: **Exopolysaccharides and antimicrobial biosurfactants produced by *Paenibacillus macerans* TKU029**. *Appl Biochem Biotechnol* 2014, **172**:933–950.

81. Çolak F, Olgun A, Atar N, Yazıcıoğlu D: **Heavy metal resistances and biosorptive behaviors of *Paenibacillus polymyxa*: Batch and column studies**. *J Ind Eng Chem* 2013, **19**:863–869.

82. Li O, Lu C, Liu A, Zhu L, Wang P-M, Qian C-D, Jiang X-H, Wu X-C: **Optimization and characterization of polysaccharide-based bioflocculant produced by *Paenibacillus elgii* B69 and its application in wastewater treatment**. *Bioresour Technol* 2013, **134**:87–93.

83. Rehm BHA, Valla S: **Bacterial alginates: biosynthesis and applications**. *Appl Microbiol Biotechnol* 1997, **48**:281–288.

84. Buldum G, Bismarck A, Mantalaris A: **Recombinant biosynthesis of bacterial cellulose in genetically modified *Escherichia coli***. *Bioprocess Biosyst Eng* 2018, **41**:265–279.

85. Zhan X-B, Lin C-C, Zhang H-T: **Recent advances in curdlan biosynthesis, biotechnological production, and applications**. *Appl Microbiol Biotechnol* 2012, **93**:525–531.

86. Naessens M, Cerdobbel A, Soetaert W, Vandamme EJ: **Leuconostoc dextransucrase and dextran: production, properties and applications**. *J Chem Technol Biotechnol* 2005, **80**:845–860.

87. Bajaj I, Survase S, Saudagar P, Singhal R: **Gellan Gum: Fermentative Production, Downstream Processing and Applications**. *Food Technol Biotechnol* 2007, **45**.

88. Williams A, Gedeon KS, Vaidyanathan D, Yu Y, Collins CH, Dordick JS, Linhardt RJ, Koffas MAG: **Metabolic engineering of *Bacillus megaterium* for heparosan biosynthesis using *Pasteurella multocida* heparosan synthase, PmHS2**. *Microb Cell Factories* 2019, **18**:132.

89. Jin P, Kang Z, Yuan P, Du G, Chen J: **Production of specific-molecular-weight hyaluronan by metabolically engineered *Bacillus subtilis* 168**. *Metab Eng* 2016, **35**:21–30.

90. Srikanth R, Reddy CHSSS, Siddartha G, Ramaiah MJ, Uppuluri KB: **Review on production, characterization and applications of microbial levan**. *Carbohydr Polym* 2015, **120**:102–114.

91. Farris S, Unalan IU, Introzzi L, Fuentes-Alventosa JM, Cozzolino CA: **Pullulan-based films and coatings for food packaging: Present applications, emerging opportunities, and future challenges**. *J Appl Polym Sci* 2014, **131**:n/a-n/a.

92. Willis LM, Whitfield C: **Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways**. *Carbohydr Res* 2013, **378**:35–44.

93. Whitney JC, Howell PL: **Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria**. *Trends Microbiol* 2013, **21**:63–72.

94. Islam ST, Lam JS: **Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway**. *Can J Microbiol* 2014, **60**:697–716.

95. Schmid J, Sieber V, Rehm B: **Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies**. *Front Microbiol* 2015, **6**.

96. Robyt JF, Kimble BK, Walseth TF: **The mechanism of dextransucrase action**. *Arch Biochem Biophys* 1974, **165**:634–640.

97. Chanasit W, Gonzaga ZJC, Rehm BHA: **Analysis of the alginate O-acetylation machinery in *Pseudomonas aeruginosa***. *Appl Microbiol Biotechnol* 2020, **104**:2179–2191.

98. Krasteva PV, Bernal-Bayard J, Travier L, Martin FA, Kaminski P-A, Karimova G, Fronzes R, Ghigo J-M: **Insights into the structure and assembly of a bacterial cellulose secretion system**. *Nat Commun* 2017, **8**.

99. Whitney JC, Whitfield GB, Marmont LS, Yip P, Neculai AM, Lobsanov YD, Robinson H, Ohman DE, Howell PL: **Dimeric c-di-GMP Is Required for Post-translational Regulation of Alginate Production in *Pseudomonas aeruginosa***. *J Biol Chem* 2015, **290**:12451–12462.

100. Dorman MJ, Feltwell T, Goulding DA, Parkhill J, Short FL: **The Capsule Regulatory Network of *Klebsiella pneumoniae* Defined by density-TraDISort**. *mBio* 2018, **9**:e01863-18, /mbio/9/6/mBio.01863-18.atom.

101. Hong Y, Reeves PR: **Diversity of O-Antigen Repeat Unit Structures Can Account for the Substantial Sequence Variation of Wzx Translocases**. *J Bacteriol* 2014, **196**:1713–1722.

102. Hong Y, Cunneen MM, Reeves PR: **The Wzx translocases for Salmonella enterica O-antigen processing have unexpected serotype specificity: Wzx specificity in *Salmonella enterica***. *Mol Microbiol* 2012, **84**:620–630.

103. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Pühler A, Ielpi L: ***Xanthomonas campestris pv. campestrisgum* Mutants: Effects on Xanthan Biosynthesis and Plant Virulence**. *J Bacteriol* 1998, **180**:1607–1617.

104. Jolly L, Stingele F: **Molecular organization and functionality of exopolysaccharide gene clusters in lactic acid bacteria**. *Int Dairy J* 2001, **11**:733–745.

105. Long S, Reed JW, Himawan J, Walker GC: **Genetic analysis of a cluster of genes required for synthesis of the calcofluor-binding exopolysaccharide of *Rhizobium meliloti*.** *J Bacteriol* 1988, **170**:4239–4248.

106. Załuga J, Stragier P, Baeyen S, Haegeman A, Van Vaerenbergh J, Maes M, De Vos P: **Comparative genome analysis of pathogenic and non-pathogenic *Clavibacter* strains reveals adaptations to their lifestyle**. *BMC Genomics* 2014, **15**:392.

107. Terrapon N, Lombard V, Gilbert HJ, Henrissat B: **Automatic prediction of polysaccharide utilization loci in Bacteroidetes species**. *Bioinformatics* 2015, **31**:647–655.

108. Becker A: **Challenges and perspectives in combinatorial assembly of novel exopolysaccharide biosynthesis pathways**. *Front Microbiol* 2015, **6**.

109. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B: **The carbohydrate-active enzymes database (CAZy) in 2013**. *Nucleic Acids Res* 2014, **42**:D490–D495.

110. Freitas F, Torres CAV, Reis MAM: **Engineering aspects of microbial exopolysaccharide production**. *Bioresour Technol* 2017, **245**:1674–1683.

111. Maleki S, Mærk M, Valla S, Ertesvåg H: **Mutational Analyses of Glucose Dehydrogenase and Glucose-6-Phosphate Dehydrogenase Genes in *Pseudomonas fluorescens* Reveal Their Effects on Growth and Alginate Production**. *Appl Environ Microbiol* 2015, **81**:3349–3356.

112. Cheng F, Yu H, Stephanopoulos G: **Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid**. *Metab Eng* 2019, **55**:276–289.

113. Badri A, Williams A, Linhardt RJ, Koffas MA: **The road to animal-free glycosaminoglycan production: current efforts and bottlenecks**. *Curr Opin Biotechnol* 2018, **53**:85–92.

114. Wang Y, Moradali MF, Goudarztalejerdi A, Sims IM, Rehm BHA: **Biological function of a polysaccharide degrading enzyme in the periplasm**. *Sci Rep* 2016, **6**:31249.

115. Mojica FJM, Ferrer C, Juez G, Rodríguez-Valera F: **Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning**. *Mol Microbiol* 1995, **17**:85–93.

116. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A: **Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product**. *J Bacteriol* 1987, **169**:5429–5433.

117. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Soria E: **Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements**. *J Mol Evol* 2005, **60**:174–182.

118. Gasiunas G, Barrangou R, Horvath P, Siksnys V: **Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria**. *Proc Natl Acad Sci* 2012, **109**:E2579–E2586.

119. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity**. *Science* 2012, **337**:816–821.

120. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, Koonin EV: **A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action**. *Biol Direct* 2006, **1**:7.

121. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, Rocha EPC, Vergnaud G, Gautheret D, Pourcel C: **CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins**. *Nucleic Acids Res* 2018, **46**:W246–W251.

122. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P: **CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes**. *Science* 2007, **315**:1709–1712.

123. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, Dickman MJ, Makarova KS, Koonin EV, van der Oost J: **Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes**. *Science* 2008, **321**:960–964.

124. Makarova KS, Wolf YI, Koonin EV: **Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?** *CRISPR J* 2018, **1**:325–336.

125. Pawluk A, Staals RHJ, Taylor C, Watson BNJ, Saha S, Fineran PC, Maxwell KL, Davidson AR: **Inactivation of CRISPR-Cas systems by anti-CRISPR proteins in diverse bacterial species**. *Nat Microbiol* 2016, **1**:16085.

126. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, Charpentier E, Cheng D, Haft DH, Horvath P, et al.: **Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants**. *Nat Rev Microbiol* 2020, **18**:67–83.

127. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S, et al.: **Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation**. *Science* 2014, **343**:1247997–1247997.

128. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O: **Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA**. *Cell* 2014, **156**:935–949.

129. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, et al.: **Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System**. *Cell* 2015, **163**:759–771.

130. Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, DeGennaro EM, Winblad N, Choudhury SR, Abudayyeh OO, Gootenberg JS, et al.: **Multiplex gene editing by CRISPR–Cpf1 using a single crRNA array**. *Nat Biotechnol* 2016, **35**:31.

131. Campa CC, Weisbach NR, Santinha AJ, Incarnato D, Platt RJ: **Multiplexed genome engineering by Cas12a and CRISPR arrays encoded on single transcripts**. *Nat Methods* 2019, **16**:887–893.

132. Dong D, Ren K, Qiu X, Zheng J, Guo M, Guan X, Liu H, Li N, Zhang B, Yang D, et al.: **The crystal structure of Cpf1 in complex with CRISPR RNA**. *Nature* 2016, **532**:522–526.

133. Miao C, Zhao H, Qian L, Lou C: **Systematically investigating the key features of the DNase deactivated Cpf1 for tunable transcription regulation in prokaryotic cells**. *Synth Syst Biotechnol* 2019, **4**:1–9.

134. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA: **RNA-guided editing of bacterial genomes using CRISPR-Cas systems**. *Nat Biotechnol* 2013, **31**:233.

135. Shuman S, Glickman MS: **Bacterial DNA repair by non-homologous end joining**. *Nat Rev Microbiol* 2007, **5**:852–861.

136. Doudna JA, Charpentier E: **The new frontier of genome engineering with CRISPR-Cas9**. *Science* 2014, **346**:1258096.

137. So Y, Park S-Y, Park E-H, Park S-H, Kim E-J, Pan J-G, Choi S-K: **A Highly Efficient CRISPR-Cas9-Mediated Large Genomic Deletion in *Bacillus subtilis***. *Front Microbiol* 2017, **8**:1167.

138. Scherer S, Davis RW: **Replacement of chromosome segments with altered DNA sequences constructed in vitro.** *Proc Natl Acad Sci* 1979, **76**:4951–4955.

139. Kim YG, Cha J, Chandrasegaran S: **Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.** *Proc Natl Acad Sci* 1996, **93**:1156–1160.

140. Bibikova M, Golic M, Golic KG, Carroll D: **Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases**. *Genetics* 2002, **161**:1169–1175.

141. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U: **Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors**. *Science* 2009, **326**:1509–1512.

142. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF: **Targeting DNA Double-Strand Breaks with TAL Effector Nucleases**. *Genetics* 2010, **186**:757–761.

143. Turan S, Galla M, Ernst E, Qiao J, Voelkel C, Schiedlmeier B, Zehe C, Bode J: **Recombinase-Mediated Cassette Exchange (RMCE): Traditional Concepts and Current Challenges**. *J Mol Biol* 2011, **407**:193–221.

144. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, Clement K, Welch MM, Horng JE, Malagon-Lopez J, Scarfò I, et al.: **Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing**. *Nat Biotechnol* 2019, **37**:276–282.

145. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, Gonzales APW, Li Z, Peterson RT, Yeh J-RJ, et al.: **Engineered CRISPR-Cas9 nucleases with altered PAM specificities**. *Nature* 2015, **523**:481–485.

146. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, Sun N, Zeina CM, Gao X, Rees HA, Lin Z, et al.: **Evolved Cas9 variants with broad PAM compatibility and high DNA specificity**. *Nature* 2018, **556**:57–63.

147. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD: **High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells**. *Nat Biotechnol* 2013, **31**:822–826.

148. Ran FA, Hsu PD, Lin C-Y, Gootenberg JS, Konermann S, Trevino AE, Scott DA, Inoue A, Matoba S, Zhang Y, et al.: **Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity**. *Cell* 2013, **154**:1380–1389.

149. Dianov GL, Hübscher U: **Mammalian Base Excision Repair: the Forgotten Archangel**. *Nucleic Acids Res* 2013, **41**:3483–3490.

150. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA: **Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system**. *Nucleic Acids Res* 2013, **41**:7429–7437.

151. Depardieu F, Bikard D: **Gene silencing with CRISPRi in bacteria and optimization of dCas9 expression levels**. *Methods* 2020, **172**:61–75.

152. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, et al.: **Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation**. *Cell* 2014, **159**:647–661.

153. Fellmann C, Gowen BG, Lin P-C, Doudna JA, Corn JE: **Cornerstones of CRISPR–Cas in drug discovery and therapy**. *Nat Rev Drug Discov* 2017, **16**:89–100.

154. Liu Y, Wan X, Wang B: **Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria**. *Nat Commun* 2019, **10**.

155. Dong C, Fontana J, Patel A, Carothers JM, Zalatan JG: **Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria**. *Nat Commun* 2018, **9**.

156. Hanna RE, Doench JG: **Design and analysis of CRISPR–Cas experiments**. *Nat Biotechnol* 2020, doi:10.1038/s41587-020-0490-7.

157. McCarty NS, Graham AE, Studená L, Ledesma-Amaro R: **Multiplexed CRISPR technologies for gene editing and transcriptional regulation**. *Nat Commun* 2020, **11**:1281.

158. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al.: **Multiplex Genome Engineering Using CRISPR/Cas Systems**. *Science* 2013, **339**:819–823.

159. Yan Q, Xu K, Xing J, Zhang T, Wang X, Wei Z, Ren C, Liu Z, Shao S, Zhang Z: **Multiplex CRISPR/Cas9-based genome engineering enhanced by Drosha-mediated sgRNA-shRNA structure**. *Sci Rep* 2016, **6**.

160. Haurwitz RE, Jinek M, Wiedenheft B, Zhou K, Doudna JA: **Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease**. *Science* 2010, **329**:1355–1358.

161. Nissim L, Perli SD, Fridkin A, Perez-Pinera P, Lu TK: **Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells**. *Mol Cell* 2014, **54**:698–710.

162. Gao Y, Zhao Y: **Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing**. *J Integr Plant Biol* 2014, **56**:343–349.

163. Xie K, Minkenberg B, Yang Y: **Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system**. *Proc Natl Acad Sci* 2015, **112**:3570–3575.

164. Zhang S, Voigt CA: **Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design**. *Nucleic Acids Res* 2018, doi:10.1093/nar/gky884.

165. Toukach PV, Egorova KS: **Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts**. *Nucleic Acids Res* 2016, **44**:D1229–D1236.

166. Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA: **NUPACK: Analysis and design of nucleic acid systems**. *J Comput Chem* 2011, **32**:170–173.

167. Livak KJ, Schmittgen TD: **Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method**. *Methods* 2001, **25**:402–408.

168. Rühmann B, Schmid J, Sieber V: **Fast carbohydrate analysis via liquid chromatography coupled with ultra violet and electrospray ionization ion trap detection in 96-well format**. *J Chromatogr A* 2014, **1350**:44–50.

169. Hakomori S: **A Rapid Permethylation of Glycolipid, and Polysaccharide Catalyzed by Methylsulfinyl Carbanion in Dimethyl Sulfoxide**. *J Biochem (Tokyo)* 1964, **55**:205–208.

170. Purdie T, Irvine JC: **C.—The alkylation of sugars**. *J Chem Soc Trans* 1903, **83**:1021–1037.

171. Sassaki GL, Gorin PAJ, Souza LM, Czelusniak PA, Iacomini M: **Rapid synthesis of partially O-methylated alditol acetate standards for GC–MS: some relative activities of hydroxyl groups of methyl glycopyranosides on Purdie methylation**. *Carbohydr Res* 2005, **340**:731–739.

172. Molle V, Fujita M, Jensen ST, Eichenberger P, González-Pastor JE, Liu JS, Losick R: **The Spo0A regulon of *Bacillus subtilis*: The Spo0A regulon**. *Mol Microbiol* 2003, **50**:1683–1701.

173. Okonkwo C, Ujor V, Cornish K, Ezeji TC: **Inactivation of levansucrase gene in *Paenibacillus polymyxa* DSM 365 diminishes exopolysaccharide biosynthesis during 2,3-butanediol fermentation**. *Appl Environ Microbiol* 2020, doi:10.1128/AEM.00196-20.

174. Fukui H, Tanaka M, Misaki A: **Structure of a Physiologically Active Polysaccharide Produced by *Bacillus polymyxa* S-4**. *Agric Biol Chem* 1985, **49**:2343–2349.

175. Rühmann B, Schmid J, Sieber V: **High throughput exopolysaccharide screening platform: From strain cultivation to monosaccharide composition and carbohydrate fingerprinting in one day**. *Carbohydr Polym* 2015, **122**:212–220.

176. Schilling C, Klau L, Rühmann B, Aachmann F, Schmid J, Sieber V: **Structural elucidation of the fucose containing polysaccharide of Paenibacillus polymyxa DSM 365**. *Carbohydr Polym* 2022, **278**.

177. Cress BF, Leitz QD, Kim DC, Amore TD, Suzuki JY, Linhardt RJ, Koffas MAG: **CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production**. *Microb Cell Factories* 2017, **16**:10.

178. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, Polstein LR, Thakore PI, Glass KA, Ousterout DG, Leong KW, et al.: **RNA-guided gene activation by CRISPR-Cas9–based transcription factors**. *Nat Methods* 2013, **10**:973–976.

179. Nakashimada Y, Marwoto B, Kashiwamura T, Kakizono T, Nishio N: **Enhanced 2,3-butanediol production by addition of acetic acid in *Paenibacillus polymyxa***. *J Biosci Bioeng* 2000, **90**:661–664.

180. Kim D-K, Rathnasingh C, Song H, Lee HJ, Seung D, Chang YK: **Metabolic engineering of a novel *Klebsiella oxytoca* strain for enhanced 2,3-butanediol production**. *J Biosci Bioeng* 2013, **116**:186–192.

181. Guo X-W, Zhang Y-H, Cao C-H, Shen T, Wu M-Y, Chen Y-F, Zhang C-Y, Xiao D-G: **Enhanced production of 2,3-butanediol by overexpressing acetolactate synthase and acetoin reductase in *Klebsiella pneumoniae*: Enhanced Production of 2,3-Butanediol**. *Biotechnol Appl Biochem* 2014, **61**:707–715.

182. Yang T, Rao Z, Hu G, Zhang X, Liu M, Dai Y, Xu M, Xu Z, Yang S-T: **Metabolic engineering of *Bacillus subtilis* for redistributing the carbon flux to 2,3-butanediol by manipulating NADH levels**. *Biotechnol Biofuels* 2015, **8**.

183. Lu Y, Zhao H, Zhang C, Lai Q, Wu X, Xing X-H: **Alteration of hydrogen metabolism of ldh-deleted *Enterobacter aerogenes* by overexpression of NAD(+)-dependent formate dehydrogenase**. *Appl Microbiol Biotechnol* 2010, **86**:255–262.

184. Hakizimana O, Matabaro E, Lee BH: **The current strategies and parameters for the enhanced microbial production of 2,3-butanediol**. *Biotechnol Rep* 2020, **25**:e00397.

185. Xiu Z-L, Zeng A-P: **Present state and perspective of downstream processing of biologically produced 1,3-propanediol and 2,3-butanediol**. *Appl Microbiol Biotechnol* 2008, **78**:917–926.

186. Okonkwo C, Ujor, V., Ezeji TC: **Investigation of relationship between 2,3-butanediol toxicity and production during growth of *Paenibacillus polymyxa***. *New Biotechnol* 2017, **34**:23–31.

187. Dragosits M, Mattanovich D: **Adaptive laboratory evolution – principles and applications for biotechnology**. *Microb Cell Factories* 2013, **12**:64.

188. Tang J, Qi S, Li Z, An Q, Xie M, Yang B, Wang Y: **Production, purification and application of polysaccharide-based bioflocculant by *Paenibacillus mucilaginosus***. *Carbohydr Polym* 2014, **113**:463–470.

189. Schilling C, Gansbiller M, Rühmann, Broder, Sieber V, Schmid J: **Rheological characterization of artificial paenan compositions produced by *Paenibacillus polymyxa* DSM 365**. *Prep Manuscr* 2022,

190. Gansbiller M, Schmid J, Sieber V: **Rheology of sphingans in EPS–surfactant systems**. *Carbohydr Polym* 2020, **248**:116778.

191. Whitcomb PJ, Macosko CW: **Rheology of Xanthan Gum**. *J Rheol* 1978, **22**:493–505.

192. Kim J-Y, Song J-Y, Lee E-J, Park S-K: **Rheological properties and microstructures of Carbopol gel network system**. *Colloid Polym Sci* 2003, **281**:614–623.

193. Tafuro G, Costantini A, Baratto G, Busata L, Semenzato A: **Rheological and Textural Characterization of Acrylic Polymer Water Dispersions for Cosmetic Use**. *Ind Eng Chem Res* 2019, **58**:23549–23558.

194. Diab T, Biliaderis CG, Gerasopoulos D, Sfakiotakis E: **Physicochemical properties and application of pullulan edible films and coatings in fruit preservation**. *J Sci Food Agric* 2001, **81**:988–1000.

195. Morsy MK, Khalaf HH, Sharoba AM, El-Tanahi HH, Cutter CN: **Incorporation of Essential Oils and Nanoparticles in Pullulan Films to Control Foodborne Pathogens on Meat and Poultry Products: Antimicrobial pullulan films for muscle foods**. *J Food Sci* 2014, **79**:M675–M684.

196. Niu B, Shao P, Chen H, Sun P: **Structural and physiochemical characterization of novel hydrophobic packaging films based on pullulan derivatives for fruits preservation**. *Carbohydr Polym* 2019, **208**:276–284.

197. Miao T, Wang J, Zeng Y, Liu G, Chen X: **Polysaccharide-Based Controlled Release Systems for Therapeutics Delivery and Tissue Engineering: From Bench to Bedside**. *Adv Sci* 2018, **5**:1700513.

198. Moffat J, Morris VJ, Al-Assaf S, Gunning AP: **Visualisation of xanthan conformation by atomic force microscopy**. *Carbohydr Polym* 2016, **148**:380–389.

199. Milas M, Rinaudo M: **Conformational investigation on the bacterial polysaccharide xanthan**. *Carbohydr Res* 1979, **76**:189–196.

200. Rinaudo M, Milas M, Lambert F, Vincendon M: **Proton and carbon-13 NMR investigation of xanthan gum**. *Macromolecules* 1983, **16**:816–819.

201. Teckentrup J, Al-Hammood O, Steffens T, Bednarz H, Walhorn V, Niehaus K, Anselmetti D: **Comparative analysis of different xanthan samples by atomic force microscopy**. *J Biotechnol* 2017, **257**:2–8.

202. Amanullah A, Satti S, Nienow AW: **Enhancing Xanthan Fermentations by Different Modes of Glucose Feeding**. *Biotechnol Prog* 1998, **14**:265–269.

203. Matyjaszewski K, Möller M (Eds): *Polymer science: a comprehensive reference*. Elsevier; 2012.

204. Nampoothiri KM, Singhania RR, Sabarinath C, Pandey A: **Fermentative production of gellan using *Sphingomonas paucimobilis***. *Process Biochem* 2003, **38**:1513–1519.

205. Karthikeyan RS, Karthikeyan RS, Rakshit SK, Baradarajan A: **Optimization of batch fermentation conditions for dextran production**. *Bioprocess Eng* 1996, **15**:247–251.

206. Schilling C, Koffas MAG, Sieber V, Schmid J: **Novel Prokaryotic CRISPR-Cas12a-Based Tool for Programmable Transcriptional Activation and Repression**. *ACS Synth Biol* 2020, **9**:3353–3363.

207. Yang S-T (Ed): *Bioprocessing for value-added products from renewable resources: new technologies and applications*. Elsevier; 2007.

208. Misailidis N, Demetri Petrides: **Xanthan Gum Production via Fermentation - Process Modeling and Techno-Economic Assessment (TEA) using SuperPro Designer**. 2020, doi:10.13140/RG.2.2.36384.74245.

209. Adli M: **The CRISPR tool kit for genome editing and beyond**. *Nat Commun* 2018, **9**:1911.

210. Shi J, Wang E, Milazzo JP, Wang Z, Kinney JB, Vakoc CR: **Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains**. *Nat Biotechnol* 2015, **33**:661–667.

211. Manguso RT, Pope HW, Zimmer MD, Brown FD, Yates KB, Miller BC, Collins NB, Bi K, LaFleur MW, Juneja VR, et al.: **In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target**. *Nature* 2017, **547**:413–418.

212. Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, Chen Y, Villalta JE, Gilbert LA, Horlbeck MA, Hein MY, et al.: **A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response**. *Cell* 2016, **167**:1867-1882.e21.

213. Schilling C, Ciccone R, Sieber V, Schmid J: **Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365**. *Metab Eng* 2020, **61**:381–388.

214. Swarts DC, van der Oost J, Jinek M: **Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a**. *Mol Cell* 2017, **66**:221-233.e4.

215. Vigouroux A, Bikard D: **CRISPR Tools To Control Gene Expression in Bacteria**. *Microbiol Mol Biol Rev* 2020, **84**:e00077-19, /mmbr/84/2/MMBR.00077-19.atom.

7. Appendix

7.1. Supplemental information of manuscripts

7.1.1. Metabolic Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365

Table S1: Bacterial strains used in this study

Strain	**Genotype**	**Reference**
E. coli NEB Turbo	F´ *proA*+ B+ *lacIq* Δ *lacZ* M15/ fhuA2 Δ(lac-*proAB*) glnV gal R(zgb210::Tn10)TetS *endA1* thi-1 Δ(hsdS-mcrB)5	New England Biolabs C2984I
E. coli S17-1	Conjugation strain; recA pro hsdR RP42Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC 47055
P. polymyxa DSM365	wild type	DSMZ 365
P. polymyxa DSM365 *Δclu1*	DSM365 *Δclu1*	Rütering et al (2017)
P. polymyxa DSM365 *Δclu1 ΔsacB*	DSM365 *Δclu1 ΔsacB*	This study
P. polymyxa DSM365 *Δldh1*	DSM365 *Δldh1*	This study
P. polymyxa DSM365 *Δpfl*	DSM365 *Δpfl*	This study
P. polymyxa DSM365 *Δspo0A*	DSM365 *Δspo0A*	This study
P. polymyxa DSM365 *ΔspoIIE*	DSM365 *ΔspoIIE*	This study
P. polymyxa DSM365 *Δldh1* pHEiP_Ppbdh	DSM365 *Δldh1* pHEiP_Ppbdh	This study

Table S2: Plasmids used in this study

Plasmid	**Description**	**Reference**
pCasPP	*P. polymyxa* CRISPR-Cas9 genome editing plasmid	(Rütering et al., 2017)
pCasPPH_clu1	clu1 targeting knock out plasmid containing repair template	(Rütering et al., 2017)
pCasPPH_ldh1	*ldh1* targeting knock out plasmid containing repair template	This study
pCasPPH_spo0A	*spo0A* targeting knock out plasmid containing repair template	This study
pCasPPH_pfl	*pfl* targeting knock out plasmid containing repair template	This study
pCasPPH_spoIIE	*spoIIE* targeting knock out plasmid containing repair template	This study
pCasPPH_sacB	*sacB* targeting knock out plasmid containing repair template	This study
pHEiP_Ppbdh	constitutive expression of *Ppbdh* under control of Pspac	This study

Table S3: Primers used for plasmid construction and knock out verification. Overhangs used for Golden Gate assembly are depicted in lower case. Restriction sites for cloning of homology flanks are underlined.

Primer	Sequence 5'→3'
ldh_sgRNA1_fw	acgcCGTCGTCCAGTACAGACAGG
ldh_sgRNA1_rev	aaacCCTGTCTGTACTGGACGACG
ldh_sgRNA2_fw	acgcCGAGTTGGAACTAAGTCCAG
ldh_sgRNA_2_rev	aaacCTGGACTTAGTTCCAACTCG
ldh_Harms_US_fw	ATATTCTAGAGCGCTACCTGTAACACTGGG
ldh_Harms_US_rev_OE	CGAGGAGGCAACATACCTCATTGAGGATAAGGAGTTTTTGCAATCC
ldh_Harms_DS_fw_OE	CTCCTTATCCTCAATGAGGTATGTTGCCTCCTCGTTCCAGCGAA
ldh_Harms_DS_rev	CTATTCTAGAGGCGCTTGCTTCAATCTATACTCA
ldh_KO_proof_fw	GGAGAACTGCTACGACTCGGAT
ldh_KO_proof_rev	ACCAAACGGCTGTGCTCTGC
pfl_sgRNA1_fw	acgcACTGACTATTACTTCTAACG

pfl_sgRNA1_rev	aaacCGTTAGAAGTAATAGTCAGT
pfl_sgRNA2_fw	acgcAGTCACCGATAATACGACCA
pfl_sgRNA_2_rev	aaacTGGTCGTATTATCGGTGACT
pfl_Harms_US_fw_gib	CTCAAGGTCGCCCGCAACCGGCGCATCAAGCCCGCCGaGTGGCAATTATTGACCCAGTGTTTG
pfl_Harms_US_rev_OE	GCGTCATTGTAATACCAAATTGTTTGGTCTTCCGGACTTG
pfl_Harms_DS_fw_OE	GACCAAACAATTTGGGTATTACAATGACGCACGCATAGCG
pfl_Harms_DS_rev_gib	GAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCACGGTGAGCATGGTCCATTTCTTGAAATTCA
pfl_KO_proof_fw	CGCAGTAATTACGGACAAAAACCT
pfl_KO_proof_rev	AACGCTGGAATCAGCAACATTGG
spoIIE_sg1_fw	acgcGGGCGTATCACAAATCATGG
spoIIE_sg1_rev	aaacCCATGATTTGTGATACGCCC
spoIIE_sg2_fw	acgcTAACAGGCTGCCTAACACAA
spoIIE_sg2_rev	aaacTTGTGTTAGGCAGCCTGTTA
spoIIE_US_fw	AATTAACGTCTAGAGTTAGAGGGCAAAGTGACAGGC
spoIIE_US_rev_OE	ACAGTCGGCCGCGTCAGGCACCCCATTCAT
spoIIE_DS_fw	TGCCTGACGCGGCCGACTGTTTACTAGAAAGGG
spoIIE_DS_rev	AATTAAGGTCTAGAAGCATCCCTTCTCCTTTGCC
spoIIE_KO_proof_fw	TCACCACAGCGGGATTTTTTTGC
spoIIE_KO_proof_rev	TTCCACTTGCCCTTGATGACCG
BbsI_Ppbdh_fw	AATTGAAGACAAACGCATGCAAGCATTGAGATGGCATG
BBsI_Ppbdh_rev	ATTGAAGACATAAACTTAAGCTTGCGGAGATACCAGGA
CLU1_USHR_XBAI_F	GGCCTCTAGAGCAATACAAAATCTGCGGTAACAATCG
CLU1_USHR_R	CCGAGCAGACCTTCGTAATGACTTTGCGG-GAGGGAAATGTATACTTTCGGATACGG
CLU1_DSHR_F	CCGTATCCGAAAGTATACATTTCCCTCCCGCAAAGTCATTAC-GAAGGTCTGCTCGG
CLU1_DSHR_XBAI_R	GGCCTCTAGACCCTCTCTCCTCATATTTTTAGACCGG
CLU1_KOPROOF_F	CCATTCCGGTAAATGCTTATGAATGAGG
CLU1_KOPROOF_R	GCTACGACCGCTAATAGTCCAGATAAAACG
CLU1_KO_GBLOCK	GAGACATCTTTGAAGACAAACGCAATTGGCTCTGTGAACGACAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGGCTGCTCCTTCGGTCGGACGTGCGTCTACGGGCACCTTACCGCAGCCGTCGGCTGTGCGACACGGACGGATCGGGCGAACTGGCCGATGCTGGGAGAAGCGCGCTGCTGTACGGCGCGCACCGGGTGCGGAGCCCCTCGGCGAGCGGTGTGAAACTTCTGTGAATGGCCTGTTCGGTTGCTTTTTTTATACGGCTGCCAGATAAGGCTTGCAGCATCTGGGCGGCTACCGCTATGATCGGGGCGTTCCTGCAATTCTTAGTGCGAGTATCTGAAAGGGGATACGCGTATACCCCGAAATACCCGGGTTTAAGTCTTCTTTCACGTGGC

sacB_usHR_XbaI.FOR	TGCTCTAGATGTCTACATTGTTGTATGGACTGGG
sacB_dsHR_XbaI.REV	CCATTCTAGAGCCACTCAGGTTCACTGTGATCAGG
sacB_usHR.FOR	TGTCTACATTGTTGTATGGACTGGG
sacB_dsHR.REV	CACTCAGGTTCACTGTGATCAGG
sacB_KOproof.FOR	CCTCATTGATTGGACTGACCC
sacB_KOproof.REV	CCAGGCCAGAGCATAACGG
sg1_sacB.FOR	acgcACTACCTTGCTGTTAGGAGG
sg1_sacB.REV	aaacCCTCCTAACAGCAAGGTAGT
sg2_sacB.FOR	acgcGCACAGTAGCCAATTACAAA
sg2_sacB.REV	aaacTTTGTAATTGGCTACTGTGC
Spo0A_US_fw	AATTTCTAGAGAGGGGATTTCAAGGGGAAGTGACACCCCAA
Spo0A_US_revOE	GTCACTTCCCCTTGAAATCCCCTCCTCAAAAGTCTCC
Spo0A_DS_fwOE	AGGAAGTACAGGGTCAGAATGATTGATGAAGCAAG
Spo0A_DS_rev	AATTTCTAGAACCCCATGCCTGTCCTCTTC
Spo0A_sg1_fw	acgcATTGTTGGCTGATGACAACC
Spo0A_sg1_rev	aaacGGTTGTCATCAGCCAACAAT
Spo0A_KO_proof_fw	CACACATGAACGGCAAACCAC
Spo0A_KO_proof_rev	GTATGTCCTTAAAGCCCGCGTCAC

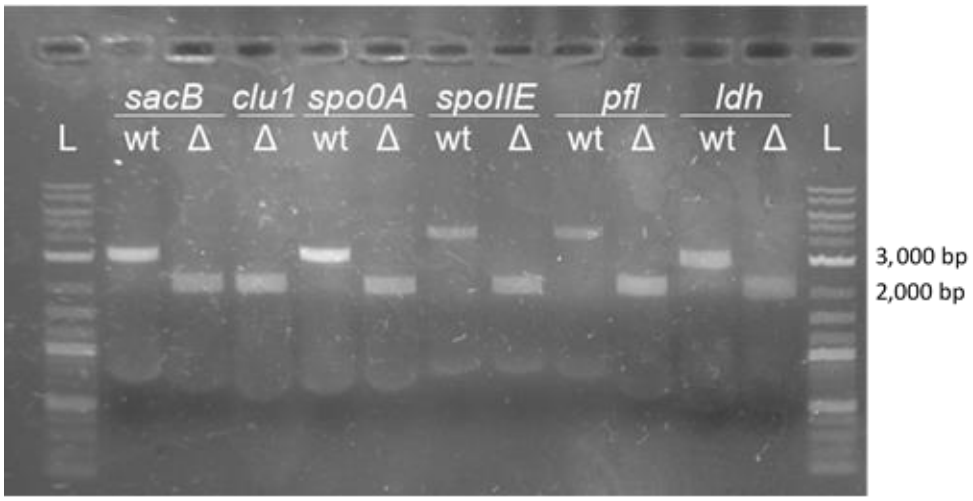


Figure S1: CRISPR/Cas9 mediated knockout variants of *P. polymyxa*. The 1 % agarose gel shows PCR products of wildtype (wt) and knockout mutants (Δ) for each gene of interest. For each knockout repair templates were provided by fusion of 1 kb up- and downstream regions of the target gene resulting in approximately 2 kb amplicons for successful knockouts. For cPCRs, primers (KO_proof in Table S3) located on the genome, outside of the flanking repair template were used for both amplicons. clu1 represents the 18 kb biosynthesis cluster required for paenan synthesis and did not yield a PCR product for the wt strain for the chosen amplification conditions. L: NEB 1 kb Plus DNA ladder; wt: wildtype amplicon; Δ: KO amplicon

Table S4: Sequences of targeted genes in *P. polymyxa* DSM 365

Gene	Sequence 5'-> 3'
sacB	TTGAAGTTTAACAAATGGTTCAGTAAAGCAGCCACTGCAACAGTAGCA ACTACCTTGCTGTTAGGAGGCGGCGTTCAGGCTTTTGCCGAGGAAAA TGATGCATCTGATCTCAAAACAGACAATGCATTTACGCAAATTACGCG TAATGACATGCTGAATATTTTCAAACAGCAAGGCAAAGAGAAATATGA GGTTCCGTCCTTCGATGCGTCTACAATTAAGAATATTCCTTCTGCTATT GGTCGCGACAGCTCAGGCAAGCTGACTGATTTTGATGTCTGGGACTC TTGGCCGTTACAAAATGCCGATGGCACAGTAGCCAATTACAAAGGGT ATAATATCGTTTTTGGATTGGCAGGCGATCCTAAAAGAGCAGAGGATA CATTTATATATCTGTTCTATCAAAAAGCGGGTAATACCTCGCTGAGCG GCTGGAAAAATGCCGGAAGAGTATTTAAAGACAATGATAAATTACTTG CCAACGATCCGATTCTAAAAAATCAGTCAGAAGAATGGTCCGGTTCTG CTACGTTAACCTCTGATGGACAAGTACGTTTATTCTATACCAGCAGAC AGCCCTTTGAGCCAAGTAACCAACTATACGGCAAACAAACCTTGAGCA CGGCTCAAATTAATGTCTCTCAGCCTGATGACAAGACACTGAAAGTAG ATGGAGTCGAGGATTTGAAATCGATCTATGACGGTGGAGATGGGAAA ACCTATCAAAATGTACAGCAAAGCGTCGGAATAGATATGGACAACCAT ACGTTTAGAGACCCTCACTATGTCGAGGATCAAGGCCACAAATATATT ATATTTGAAGCCAACACAGGAACAGAAACCGGCTACCAAGGCGAAGA TTCGCTTCAAAACCCAGCTTATTATGGCGGCAATAAGAAATTCTTCAC GGAAGAGCAACAAAATTTGTTGCAAAGTCCCAAGAAAAAGGGAGCTG AGCTGGCTAACGGCGCACTGGGCATCGTTGAATTGAATGATGATTATA CGTTGAAAAATGTAATGCCTCCTTTGATTGCTTCCAATTTGGTAACAGA TGAAATTGAACGGGCAAATGTGTTCAAAATGAACGGTTTGTGGTATTT ATTTACGAGTACGAGAGGCTCCAAGGTGACTGTAGATGCGATCGGTG ACGATGATATTTACATGCTGGGTTATGTTTCTACTTCCTTGACAGGGC CTTATAAACCCTTGAATGGAACCGGTCTCGTTCTTCATCAGGATCTGG ATCGTGATGATATTACCTGGACGTATGCCCATTTTGCTATCCCGCAAG GCAAAGGCAATAACGTAGTCGTTTCGAGTTATATGACCAATCGCGGG CTTTTCCCGGATCA
clu1	GCAATACAAAATCTGCGGTAACAATCGGATACATAGGGATTATGCAAA TCCTGAACTGTAGACAGGAGGCCGAATCATGAAAATTGTCATTGTCGA TGATCATCCTTTAGTACGGAAAGGACTCGCTGCGGTTATATCCATGCA GCCCAATGTTCAGTTTGCGGGAGAGGCCCAGAATCAACAGGAAGCAC TTGTGGTGATTGAGGAAACGGATCCGGATCTGGTGCTGCTGGATTTG AAACTAGCTGATGAATCAGGGCTGGATATTATCAAGATAGCGCGCGG ACGTGGTTTCAGAAGTAAATTTATTTTGCTGACCTCTTCCGCAACACG TGAAGATTTTCTGAAAGCAGAGGAAGCTTCCGTGGACGGTTATGTGCT TAAGGAAGCGTTGCCTGAGGAACTAATTTACGCGATTCATCTTGTCAA CAAAGGACGCAAATATTATGATCCGGCTATTCTCGAAAACAAGCTGCG TGAAGATGCGGGCAATCTGACGGACGACCTGACGCCAAAGGAAAAG GAAGTGCTGGTCGAGTTGGGACAAGGAGCTTGCAATCGTGAAATTGC ATCCCGATTATTTATCAGTGAGTTTACAGTCAAAAAGCATGTCAGTCA

Gene	Sequence 5'-> 3'
	GATCCTGGCCAAGCTGCGGCTGGCGGATCGCACGCAGGCAGCGTTG TATGCCAATGCCATTGGTCTGGCGAAGTACGAGCCAACAAACATGGT ATGATGCGCTGACTCTATAAGTTGCATTGCACTTTCATATGCGAGTGG AATGAAAAGCCTCAAAATCACCAATTTTTGTGATTTTGAGGCTTTTTTC GATTTCTTTTTCCATAATCAACAATGTATGGTACAAAAGTATACCATGC CGTATCCGAAAGTATACATTTCCCTCCCCCTGTAGTGAGGGAGAAAGT GCCATTCGGTTCCATGTCGTTCACAGAGCCAATTGTTACAATAAAGTC ATACTTAAATATGAAGAGAAATTGTATAGGAGGGAGGATTTTTGTGGA GAAAACCATTTTGGATTATCTAGTCCTGATTAAAAGAAGATTGTGGCT GATTGCGATTTTTGTCATTCTCTCATGTGTAACAACATACTTTGTGAGC GACAATGTCGTGAAGCCAGTCTACTCCGCTTCAAGCCAACTGCTGGT TAATCATATTTCCGCCAAGCAGGGGAGCAACAATCTGAACGACGTCAA CACAAGTCTTAATTTGGTGGAAAGCTACAAGCAAATTCTCACTTCACC TAAAATTATGGACGCTGTTGTAGCGACTCATCCCGAGTTCGGTTTGAC ACAGCAGAAGTTGGCAGAAAAGCTGCAGATTAAATCCTCGGATAAAA GTCAGGTTATCGGTTTGACGTTTGAGGCAGGGAGCTATCCACAGGCC GCAGCGGTTGTAAACGCGGTTACCCAGAAATTCGTACAGACGGTTCC GGGATTGATGGAATTGAATAATGTAAAAATTTTAAATGAAGCGGATCC TCAGGCCAATCCTGACCCGATTAATGGGAACCCGCTCATGAATATTGC GATTAGTTTTCTCGTATCGCTGATGATTGCATTAGGCACGATTGTGTT CCTGGAAAGCATTAACGGCACGCTGCGCTCGGAAAAAGAGGCAGAC GCAGACATCGGTATTCCGGTAATTGCATCGATTCCAGTCATTCGCAAG AAGGACATCGATGGAGCAGCCGGAAATTCCAAAGAACGGGTAGGGG AGCGTAAATATGCTGCGATTGAATGACATGTTGATTACAGAAAGTAAT CCATCATCGTATATCTCGGAATCATTTCGATCCCTCCGTACATACATTC GGCAGCAGGTGATTGGACAAAAGGACCGTGGGACGGTTCTATTGTTG ACCTCGCCGGAAAGCGGAGCGGGTAAAACAACACTGCTCGCCAACAT CGGAGTTTCTTTTGCACAAGAAGGCAAGAAGGTGGCTCTGGTGGATT GCAATCTGCATAAGCCAGCGCTGCACGAGATATTCGGAATGGAGAAT ACCGGAGGTCTATCGGCTTATTTGCGCGGTGGGGCTTCTTCCAAGTC TATCGTTAGGCATGGTGGCCTGGAGCTATCGGTTGTTCCCGGTGGAG ACACCCTGTACAACGCAGCGGATCTGCTCGGTAGCGAACGTATGGCT GAGCTGCTGGAAGAGCTCAAACGTGAATATGACATCATCTTACTCGAT TCAGCTCCGGCGCTGAATTATACGGATGCCCGCCTGATTGCAGGCTT GACAGATGGTGTGATTCTCGTTGCTAAACATGGTCGTACCAAACGGG AAGATCTGCGCAAAACAAAGCAACTCATGGAGCAGGCCAGTGCGACG CTGGTTGGAATTGTGATGAATCAGGTGAAGTAAGCAACACCAAAATTC GATGAAGCAAGGAGAATGAGTGCGATGACGAAAAGAGTGAAAAAGGC CATTATTCCAGCAGCAGGGTTAGGTACGCGCTTCCTCCCTGCCACCA AAGCGATGCCTAAGGAAATGCTTCCAATTATCAACAAGCCTACGATTC AATATATCGTGGAAGAAGCGATTGCTTCTGGTATTGAGGACATTATTA TCGTTACGGGTAAAGGCAAACGTGCCATTGAGGATCACTTTGACAAC GCATTTGAACTGGAGTCAAAACTGCTCGAGGACGGTAAGCTGAAGCT TCTGGAGGAAGTACAGCGTTCTTCGGGAGTGGAAATTCACTACATTC GCCAAAAAGAACCCAAAGGTCTTGGACATGCGGTATGGTGCGCAAGA CGCTTTATCGGGGACGAGCCGTTTGGTGTGCTGCTGGGCGATGATAT CGTAACAGGTCAGAAGCCTTGCCTGCGCCAGCTGATGGATCAGTATG AGGAAACACAAAATTCGGTTATCGGTGTTCAGCGGGTACCGCAGGAA TTTACGAACCGGTACGGCATCATTGAACCGGATCAGCAGGACGGACG CTTGTACCGAGTGAATAATTTTATAGAAAAACCGGCTCCAGGCACAGC

Gene	Sequence 5'-> 3'
	GCCATCCGATTTGGCTATTATGGGTCGCTACGTGTTCTCACCGAAAAT TTTCAAATATCTTGATCTTCAGGAAAAAGGAGCTGGCGGCGAAATTCA GTTGACGGATGCGATCCAAAAGCTGAATCAGAGTGAGCGTGTGTACG CCTATGATTTTGAAGGCACACGCTATGACGTAGGTGAGCGTCTTGGAT ATATTTTAACCACACTTGAATTTGCACTTGCCAGCGATGATTTGAAGTA TCCGGTTATGGAGGCTATGAATGAATGGTTGAAAAAGACCGAGAAAG CAACCACGGCAGGTTGAGAGGGGGTACGTCAGCATGAGCATGGAAA ATTTACCGGAGGACGGCGAGGCTGTCATGTCAGGTAAAGCGTTTTAC ATGCCATACGGAAATACACAAATTCAGGATAAAACGTCCTACCTGGTC ACAAAACGCCTCCTTGATATGCTGCTGTCGTTCGTGGGCTTGATCGTC TTGCTTCCGCTTTTTGTAGTGGTAGGCGTATTGATTAAGCTCGAAGAC CCAAAGGGAAGCGTGTTTTTCAAGCAAACTCGGGTGGGCAAGAATGA GAAGTTATTCGATATGTATAAATTCCGTTCAATGGTGTCCAATGCTGA GGAATTGAAAAAGGACTTGATGGAGCTGAATGAAGTGAGCGGCGCCA TGTTCAAAATTAAAAACGATCCTCGGATCACGAAAATTGGTAGTTTTCT GCGCAAGACGAGTATTGATGAGATTCCGCAGCTGTGGAACGTGTTGG TCGGCGACATGACTTTAGTCGGTCCGCGTCCGCCGTTGCCCAGCGA GGTCGAGCAATATTCCGATTATGATAAGCAACGCCTGATTGTAACACC GGGATGTACGGGTTACTGGCAGGTGAGCGCACGTAACAGTGTAGGC TTCGAAGAAATGGTGCAAATGGATTTGAAGTATATTCATGTACGCAAT ACGTGGCTCGATTTGAAAATCATCATGAAAACAGGCGTTAAAATGCTG TTCTCCAAGGATGCTTATTGATGGATGCATTCAAGCGATGTCCCGCAG TATCCGTCGTTATTTGCACCTATAACCGGGCGGATCTGCTGGAAAAAA CGTTAATGTCCCTGCTGGAGCTGGAAGACCTTGCACTGGCTGAGATC ATTGTAGTCGATAATCGGTCCACCGATCATACAGCGGCGACGATCAA AAGGTTCACTGTCGCACATGGTCAGAATATCCATCTCCGATATCATTA TGAGCGAGAGCAAGGACTGTCTGCAGCTCGTAATGCAGGCATTACGT TATCAAGAGCGGAAGTCATTGCTTTTCTCGATGATGATGCAATTCCGT GTGTTACATGGTTGCAAACGATTATGTCGGCATTCGGGAACAATCCTG AGCTGACGGCTATGGGTGGGAAAATTGACCCTATGTTTGAGACGGAA CGACCGGGTTGGCTGACAGGCCCGCTTGAACTGCCTTACACGATTAT CGATCTGGGCAAAGCCGTTCGTGAATACCCGTCGGGACTGAATCCAT TTGGAGCGAATATGGCGATGCGCAAAACGGCTTTTGCCGCCGTTATG TTCCCGCTTCATCTAGGCAGAAAGGGAGATATCCTTCTATCTGGTGAA GAATCGTGGGTATTCGAGCAAATCCGTAAAAATGGTGGAACGATTATG TATCATCCACACATGTCTGTGGAACATTTTGTGCCGGCTTCACGGTTA ACGAAAGAATGGATCATGAACCGCTACTATTGTCAGGGCATGTCTAAC GCTGCTCAGGCCGTAGGTCTGCGCGGCAACGCGATGCTGATGGCGA AGACGGCAGCCAAAGTGTTATACATCACGGCGGATTCCCTGCTGGCA CGTAGTGAGGGCAGGAAGCTGTTGAACCGGTGCAGGCTTGAAAGTAT CCGTGGGACGCTGGATACACTGCGCAACCGGAAAAGCGAGTCAGCT GCGGGGTGAAAGATGGAACATGAGTAACTATCCTTGGACTACTAGAA TAAACACCATACAGCATGTTTTTATTTTTCTACTGGGCGCTCTTGCTGT AGGAATTGCGGCTACCTATCAGCCGGTGGTCAGTATTGCGGCGGTAT GTTTACTGCTCTTGCTGGCTGTTTCTATCCATCATCCGGAGCGCATCA GCTATGCTGTTTTGCTAAGTACAGCAGTCTCAGTCGATTCTCTGTATC AGGGCGGCGTGTTCGGTATTGAGATTTTGTCATTGTACAAATTGGGCA TTTTGGCTCTGCTGGTGCCGTGCATGCTCGTGTATGGCATACGTCTCA AGTTCAGCTATCCGGTCTGGGCATTGGTGATCATGCTGGGGATCACA TTTGGATTTTCAGCATGGATGCCTTTGTTAACGACCTCGATTGCCGTC

Gene	Sequence 5'-> 3'
	AAGGCGTTCGTCGGTTTATCACTCCCTTTCTTTTTTTTACTGATCCAGT GGAAAAAGGATGTGGCAGAAAAGCATATCCGACTGATCAGTCTGTTG CCAATTATTAGTGTGGTCATTGGGCTGGGTTTGCAGGTTCTAGGGCT GCATTCGTTTACGGATGTGGAATTTACGGGCGCAGTCAGAGTACAAG GAGCGAATATACCGCCGCATCTGGCAATGCTCTCCTTTCTCGGAATTG CCGTCTCTCTAATTGAGGTCAAGCGCAAACCCAAGCAAGCCGCTTTTT ATTACACAGTGTTGGCACTGAACTTTTTGATTCTGATCGCCACCGGTA CACGTGGACCGATTCTGGCATTGTTATTAATGTTTGCTGTGTATCTGTT TGATATTGCCCGTCAGTACCTCAAGGGGAAAGTCAATTATTTGCTACC GCTGGCCGGATCGTTTTTGGTTGCTCTGGGGGCGGTGGCCCTGCAAT GGAACAATTTAAAAAAACGCTCCTTCGAACGGGAGACCGATACAGGC ATTGACCTGTCGGGACGCTCGGAGGCATGGGAATATTTTTTGAACAG AGTACATGATTACCCTTGGTCAGGTAGAGGTCTTGGCGCGGTAACGG TTGCGAACGACGGGACGTTATTCAATGGTTTCGTTGTACCGCATAATG AGTATATTCGCTTTTATTTTGACACGGGATATATCGGTTGCGGGCTATT AATGCTTTCATTGCTGATCGTGTTCGCTCTCATTTACCGGTCGCTTGC CAAGTCTATCAAACCTTATTTTGCAAGCTTAATTGCAGGCTTTCTGATC TACTCGTTTTCAGATAATACGTTGTCGACGGTTCAGATGATTATTCCAT TTTGCTGGTATTTAAACGCGTTATACCAGACATCTACTCAAACCGATTT TCGCAAAGAGAAGTGATACGATGAGCAGAGTGAACATGTTCGATGTC AACTTTGATAATTACGATTTTAATGATCTACTCGAATTTATAGATACTTC TATTGAACATCAGCGTCATTCCTACATTTTGACCTGTAATGTCGATCAT GTGATCAAGCTTCGCAAGGATGATGAATTCCGCAAGGTGTATTCAGAT GCCGGAGCTGTAGTGGCCGATGGGATGCCTATCATATGGGCGTCCAA GTTACTGAAAAAGCCCCTCAAACAAAAAGTGTCTGGGTCAGATTTGTT CACTCGGTTGGGGGAGGCTTTTGAGGATCGAGGCTATCGTTTATTTTT TCTGGGTGCAGCTGACGGTATTCCTGAGAAGGCTATGCTCAATCTCC AGGAATCTTATCCGAACATGAACGTGGTGGGCTGTTACTCTCCTTCTT ATGGATTTGAGAAAAACGAAGAGGAGAACAGACACATTGTCCAACTG CTTCAAGAGGCTCGCCCGGACATCGTTTTTGTAGGTGTAGGCGCACC TAAACAGGAAAAATGGATTTACAGATATTATGAGACTTACCGTGCTCC GATTTCGATTGGGGTGGGAGCGACCTTTGATTTCCTGTCGGGAAATG TGAAAAGAGCACCGGAGCTGATGCAGAAGACGGGATTTGAATGGTTC TGGAGATTGTCTCAGGAGCCTAAACGATTGTGGAAAAGATATTTGATT GAGGATTCCCAATTTCTCGTTCTGCTGTTTAAGGAGATGTTCAAGCGG AAGCGGATGAAGGGAGGTCAGGGGTGAGAAACATGAGTCGGTTGGC AGATCAGCAGATTTCCGTTTGGCTAGGAAATCGGCGTGGAATCGGCA TGATCGGCATGGGGTTGCTGGCTTGTATGCTGCCACTCGCCATCGGA TTTGCCAGCGCCAAGATGAATCCCACCATGAGCCAGCAGGGAGCCAT ATTGCTGGCGCTCGTTTTTCCGGCCTTCCTGCTGGCGATCTTGCAATC CCGGCTACTGATTCCTTACACGCTGGCAGTGTGGGCAGTAGGCCCG GAAATCCGCCGTATTGCGGATTGGTTGGAGGGAACGTACCACTCGGT TTCCTTGTTGAGCGTGGCACCGCTGCTGGTAAGCAGTATGTTGATTAT TCCTGTGCTGCGGGGTATTCATCTGGCGGAAAAGCCATTAACCCGAA TTGCTGTATTTTTCGGCATAGAATTGGCCTACGGAAGTGTCGTTGGGT TATTCAAAAATGGTATTGTTTTTGCTTATGATTTAGCCAATTATGTGGTT CCTTTGATCCTATTGCCATATCTTGCGATTAAGCCCATGAAAGCGAAG GAACTGGACCGACTGCTGTATTCATATGCCAACATTGCTGTTCTTGTG GCGATTTACGGCATTATTCAGTATTTGACAGTTCCTCCTTGGGATGCA TTTTGGATGAATCATGTAGAAATGAATTCCATTGGTGTTCCAGAACCG

Gene	Sequence 5'-> 3'
	CTTCAAATCAGAGTGTTTTCTTCCATGAACTCGCCCGGTCCGTGCGCT ATTTTCCTGGCGATGGCACTTGTACCCATGCTAATGGAAAAACGGTGG CGTGGCACTTTGGGATGGATCGGAATTCTGCTGACGGTCGTTTGTCT TTTGATCACGCTGGTACGTTCAGCCTGGCTAATTGCATTCGTCATGTT GTTGGCTTATATTCTCAGTTCATCTTCCAAGGGCAAGTGGAAGACCTT GGTTCAACTAGCTGTCGTCGGTCTTCTTCTCTACATTATTGTTCCAAAA CTTCCAGGGGCAGAGGGTTTGGTGGCTCGTATGCAGACGCTGACGG ATATTCAACAGGATCATTCGTACAACGAGCGTTTGGATTTACTGCATA CGATGCTTCCGGCGATTACTAGCAATCCAGTCGGGCAGGGGATCGG GAGTGTAGGAATCGGAACCAAGCTTGATAACGGCGGCGATCTCGGTG AACTTGGGATTATGGATAATGGGTATATTGCCATTTTCCTCACCTTCG GTATTTTCGGAGCATTCTTTTTCTTCGGTGGTCTGTTCGTTATCATCAA GCGCCTTCTCGTCCGTATTGCCGAACGAGATTCCAGTCAGCCTTATAT CAGACTGGCGTTGGCTACGTGGGCCGGAGCTGTAGCCAGTCTCATAT CGGATAATGGTTTCCCGGGTATGCGCGGCTATCTGATCTGGATGATG ATCGGAATAGGACTATGGGCAAAAGATGTCATCGCGGAAAGAAGGTA AAGGCTGTGCAAAACTTACAAGCATTATCCGCACCTGTCCGGACACTA TGGTCCACAGTGATTCGGTTTGCCAAAAGCAAGGATAACAGCTCTGC AGCAGTAAAAACGATGATATTCAGTATGCTGATTCTCGTCGTGAACAT GCTGACTGGTGTGCTAACGGCACGATTCCTGGGTCCTACTGGACGCG GGGAACAGACAGCGATGGTGAACTGGTCTCAATTTCTCGCCTTCTGC ATGAGCTTCGGTGTCCCTTCTGCATTGATTTACAATGCGAAAAGGAAA CCGGAGGAAACGGGCAAGCTGTATGGTTTGGCCTTGTTGCTGGCTAC CATGTTCGGGGGCTTGGCGACACTCGTCGGAGTATTTCTAATTCCTTA TTGGCTGCGCTCTTTTTCTTCATCGGTCATCCTTTTCGCGCAGTGTTC CATGATGATGTGTCCGCTCATAGCGATTTCACAGATCAATAATGCATT GCTTCAAGTTCGCTCGGAATATAAGCAGTACAATTTATTCCGGTATCT GGTACCGCTAAGTACGTTGCTGGGTCTGGGGATTTTGATCCTAACCG GAAATATGAATCCGTACACCTCGGCATTGGCTTATTTACTACCAGGCT TGCCGATTTATATCGGTGTCACGATACGGATGATTCGGCTGTACAAAC CCAAGGTGAAGAACAGCTGGACGCAATTTAAAAATCTCTTTACCTACG GTATGGGTTCATACGGAAATGATCTCATGGGGCAGGTTTCCACTTATA TTGATCAGATTTTGATTGCGGGTCTGCTCAGACCCGCTGATCTCGGTT TATATGCAGTAGCGGTCAGCTTGGCCAGGATGGTGAATGTATTTTCTA CCTCGATCATTGTCGTGCTGTTTCCCAAAGCCTCTGGTCTGAATAAGG AAGAGGCGGTGGACATAACATTTCGGGCATTTCGAGTGACATCAACA GCAACCTTTCTGGCTGCGATGATGCTGATGCTGATTGCTCCGTTCGTA TTTACATTGTTATATGGTCAAGAGTTTAAACAAGCGCTTACGGTATTCC GGTTCCTTGTGCTTGAGGTTGCTATCAGTGGAGGGACCATGGTGCTG GCACAACAATTTATGGCGTTGGGTAAGCCCAAACTGGTTACCATTCTA CAGGGTGTCGGATTGGCGCTCGTCATTCCATTGCTGTCCATACTCGT GCCGAGATATGGATTGACGGGTGCAGGGATAGCAATGCTGTCTTCGG GGATTCTACGGTTCATTTTCATTTTATGTAATGTTAAATTCGTTCTAAAA ATGAAGATTCCAAGACTGCTTATTTCCAAACAAGATTTTCAATGGCTCC GATCAACGATGTCACACTACATCCGGAAAAAGCCAGTTAACGGTTAGA TAAAGAAGGAGGCACCCGCTAATGAACCACATGTCTAGCGTTCCCGC GGACAACCGGATATCCGTAGTTATTATCGCTCAGGATGATGGCACTC GAATCACAGCCGCCATTAAATCCTGCAAGCCCTTTGCGGACGATATC GTCGTCATTGACGGAGGGAGCAAGGACAATACGATTCAGGTAGCTGA ATCCTTGGGCTGTCGGGTGTTTTCTAACCCGTGGCCCGGTTATGCCA

Gene	Sequence 5'-> 3'
	AGCAGCGCATGTTTGGTGTTGAAAAAGCGGAATTTGATTGGATATTCC TGATTGATACGGACGAAGTAGTGGATGCAGAGTTGCTTGGAGATTTGT TGAAAGTAAAGGAAACGCTCCACGATCCGGCCAAAGCATATTCCGTAT TTCGGATTGGCGATTTTCTCGGCAAATGGATGAATAAAGGCGAATATC TGGTTCGTTTGTACAATCGTCGTATTTATGGCATTACGAATAGTCTTGT TCATGAAATGCCGGATGTAGCCTCTTCCCAGATTGTCAATCTGGAGG GTGTGTTATGGCACTACGGCTTCCGCAGCATCAGTGATCATATGACC CGTTTTAATAAATATACAGATCTGGAAGCAGAGACGGCATTTAACAAA GGAAGATCCTTTCGGATGACCAGACTGTTGTGGCGGCCTCCTGCCCG CTTTGTGCAGAAGTATTTTGTCCAAGGCTTATACCGCAAGGGATTGGC GGGTTTTGCCGTAGCCGTTTTCTGGGTGATGTATGAGTTTCTCGTCTG CTTCAAGCATTATGAGCTAACCAAGCGGCGAAAAAAGATAGAAATCGT GGATGACGGACATACCGAGCAGAAAGGAGAAACCAGCTATGTCGTAC AATAACGGATTTCGTATTGCGGCAACGGGGTTGAGCTGGCCGTCTTT GCAGCCAGGAGGCCTCAACACATATTTCCAATCGATCTGTGAGCAGC TTACGTTAGAGCGCAATACGTTGGATGCTCTAATCTGTAGTGACGAGC AGCCGCAAGCGCCCGACCGTATTCGGATTCATACGATTGGCAGCAAG CAGCAATCCATCTGGAAGCGCCGGGAGCTAATGCAAAAATATGCAGC AGAGTTGTTCGACAAGCAACCGATGGATATATTGTATTCCCATTTTGC TCCCTATAGTGTCGGGCCAGCGCTGGAGGCGAAAAAAAGAGGAATCC CGGTAGTTACCACCTTTCACGGACCGTGGACGGAGGAAATGAAGATT GAAGGACAGGGTATCAAGCATTTTCTAAAAACGACGTTGGCTAAATCC ATTGAAATGAAGGCCTACGGTCTGTCGGATAAATTTATAGTGCTAAGC GAGACTTTTCGGGATATTTTGCATGAGCACTATAAGGTTCCACTCAGC AAAATTCACATTATTCCCGGAGCGGCGAATGTAGAAAGGTTCCACCCA GCAGAGGATCGGGGAGCAGTACGAGAGCGCCTGAATTTGCCGCAAA ACGCGACGATTGTGCTAACCGTTCGTCGTCTTGTTAATCGAATGGGAC TGCTACAGCTACTCGAAGCATGGCGACGAGTAACCGAGCGTCATCCT GATCATCTGCTGCTGATTGGAGGGAAGGGTCCCTTGATGGAGGAATT GGCTTCCAAGGTAGCGGAATACAATCTTCATAACCATGTAAGGCTGCT CGGTTATGTGTCGGATGAGGAGCTTCCACTGTACCATCAAGCATCGA ATTTGTTCGTCGTACCTACACAGGCGCTGGAAGGTTTCGGTCTGATCA CGGTGGAGGCAATGGCTTCGGGACTGCCAGTGCTGGCCACTCCTGT GGGCGGTAGCAAGGAAATTTTAAGAGGATTCCGTCCCGAGCTATTGT TCCAGGGAACGGACAGTGAAGCCATTGCTGAAGGGTTGCTGCGTGTG CTGGATCATCGTGAACTTCTACCGAATGCAAGGGAATGCCGGGACCA TGTCCTGAGCAGATATACATGGGGACATGTAGCGGAGCAGGTGGAAG AGGTGTTCCGACAGGTCCTTGACAGAAAGGCGGCGGCGAAATGATGA GAGTAGCTTATATTGACCATACGGCAAGATGGAGCGGCGGCGAAGTC GCTTTGTATAACATACTGACGAATATCGGAGAACATATCGACCCGCTC GTTATTTTGGCAGAGGAGGGTGATTTGGCAGATCGGCTTCGGCAGCG TGATATTGATGTGCGCATCGTTCCATTGGACGATTCGATTCGCAATCG GGGACGGAACGCCGTGAATCTAGGTGCGCCTGCAGCAGCTTTTCGTT TGCTGGCATACGGTAGAAAACTTGCGCCCTTGCTGCGTGAGGAAAAG GTCGTGTGCGTTCACACGAACTCGCTCAAGTCTGCACTGTATGGTAC AGTAGCTGCCAAATCGGCGAAGCTGCCCTTGATCTGGCATATCCGTG ACCACATCGGTCCGCCGTATCTGAAGCCGATTGTAGCCAAAGGGATT CGGCTGATGTCCCGCTTTTTGCCCAACGGAGTGATTGCCAATTCCAA GTCCACGCTGAGTGCATTGGAGTTACCCCCGGACAAAAAGACGCTCG TCGTTTATTCCGCTTTTGCCAAAGCAATTACAGCACGTGACACGGCTG

Gene	Sequence 5'-> 3'
	CGCATTCACGTGGTGATGACTCGTTTAACGTCGTATTGGTCGGCAGAT TGGCGGAATGGAAGGGACAACATATTTTACTGGAAGCGGCACGTTCC TTTTTACCAGATCAGCGCGTGAAATTCTGGCTGGCTGGAGATGCATTG TTCGGAGAAGAGGAATACAAGCAACGATTAGAGTCCACGATGCGCGA ATACGGGTTGGCTAATGTCAATCTGTTGGGGCACGTCGATGACATTCA AGGCCTGATGCAGCGTTGTGATTTACTGATTCATACCTCTATTACACC AGAACCATTCGGTCAGGTCATTATTGAAGGCATGGCTGCTGGTCTCC CAGTGATTGCTTCCAATGAGGGTGGGCCCAAGGAAACGGTGGTACCC CATGAAACCGGGCTACTTATTGAACCGGGTGACCCGGCTAAGCTGGA AGAGGCCATCCGCTGGATGTTGGAGCATCCACAGGAACGTCAGCAGA TGGGTGAGAGAGGGATGGAACGGGTCAAGAAGCATTTTGTCATCGAG AACACAGTTAAGGATATAGTTCATTACTATAAGGGTTTGTTGGCAGGA GTCTGACCGACATTCGCGTTAAAGGATGATTTAGATGAAAATAGCGAT AGCACACGATTACTTAATCCAAATGGGCGGAGCGGAAAGGGTAGTAG AGGTATTCCACGATATGTATCCGGAGGCTCCTATCTTCACAACGGTGT TCAGCGACGACCGCCTGTCGCGTAACTTGAAAGATGCGGATATTAGA GCCACTTGGCTACAAAAAGTCCCAGGAGTGAAGGCCAACTTCAAAGG AGTGCTGCCGCTATATCCCATCGCTATCCGTGATTTTGATTTTAGGGA TTTCGATATTGTTTTAAGCTCCAGCAGTGCATTTATGAAAAGCATCCAG GTACCCAAACATACGTTTCACATTTGTTATTGTCACACACCGATGCGG TTTGCCTGGGATTACGATACATATATGGCCCGGCAGTCCAAATCCAAC CTGCTCAAAAACATGCTAAAGCTGTACATGAACCGGCTGAAAACATGG GATGCCAAAACTTCGCGCAACGTAGACCAGTTTGTCGCCAACTCTTC GGTTGTTAAAAGAAGAATTTTGCACTATTACCAAAGGGATTCGGATGT CATTTTCCCACCTATTAATACATCCCGTTTTAAACGCGCTACAAACATC GGTGACTATTACCTGATCGTATCCCGATTGGTGTCCTACAAGCGCATC GATCTCGCGATCGAAGCATTCAAACGTAACGGTCTCAAGCTGCGCAT TGTGGGTGAAGGACCAGATCGAAAACGACTCGAAGGTATGGCTGCCC CGAACATTGAGTTCCTCGGCAGGCTTGAGGACGAAGAGGTTAACAAG CTGATGGCAGAATGCCGAGCGCTCGTTTTTCCGGGCGAAGAGGACTT CGGTATTACGCCTTTGGAGGCGAATGCGGCCGGACGGCCTGTTATCG CTTTTCAGGGTGGGGGAGCGCTGGATACGATTGTTCCTCATGTGAAC GGCGTATTTTTCCGCAAGCATCAAGTAGAGGATGTGCTAGAGGCCGT CGCAAAGGTCGAGCAGCATGCATGGAATGTGGACGATATCATTACCC ATGCCCGCAAATTTGATGAAGATAACTTCAAGGATCAACTGAAAAAGT ATGTAGAACAAGCCTATGTGAACTTTCTAAAAGGAGGATGAACGGAAT GAAGCTGGCTGTAATTGGTACTGGATACGTTGGGCTCGTCTCGGGTG TATGTTTTGCGCAAAAAGGTAATGAGGTGATTTGTGTCGATCTGGAGC AGTACAAAATCGACATGTTGAACCGGGCAGAATCTCCGATTTATGAAC CGGGTATTGAGGAATTGATCGCGCTGAATCTGGAAGCAGGTCGATTG GAGTTTACTTCGGATTTGGCTGATGCGGTACGCCGGTCGGATATCGT CATTTTGGCGGTAGGTACGCCTTCTCTGGCTAACGGCGAGGCGAACT TGTCCTACATTGAACAGGCGGCTGCTGATGTTGGGAAAGCCATGAAC GGATATAAGATTATTATGACCAAGAGTACTGTACCTGTCGGAACGAAC GAAAAGATTAAGGACGTATTGGCTCGTCACACGAGTTTGTCTTTCGAT ATCGTATCTGCGCCGGAGTTTCTGAGAGAAGGCTCAGCCATCAATGA TACGCTTCATCCGGATCGCATCATTATTGGTTTGGACAATACTGGATT GCGTGAAACAATGGTCACTCTGCATCAGGTGTTCACGGATAAAATCTA TGTGACGGATATCCGTAGTGCTGAAATGATCAAATACGCGTCAAATGC GTTCCTGGCAACTAAAATTTCGTTTATTAACGAAATCGCTAATATTTGC

Gene	Sequence 5'-> 3'
	GAAAAGGTTGGCGCCGATGTAACCTGTGTGGCTGATGGCATGGGTAT GGACAAACGGATCGGCTCCTCTTTCCTACAAGCAGGGATCGGGTATG GTGGCTCTTGCTTTCCAAAAGACACAAATGCGCTTATCCAAATTGCGG GTAATGTAGACTATGAATTTAAGTTGTTGAAATCGGTGGTCGAGGTGA ATACCGACCAGCGCTTTATGATTGTATCCAAGCTGCGTGAATCGCTGG GTCAACTGAACGGTGTGTCGATTGGCATTTGGGGCCTCGCCTTCAAG CCAAATACAGACGACATCCGCGAAGCTCCAGCACTGGAAATCGTAGA AACACTCATTCAGGCAGGCGCAATTGTCAAACTGTATGATCCCATTGC GATGGATAAATTCAAGGAGCACGTGGATCACCCGAATATTGTGTGGT GCTCATCCCCACAGCAAGCCGCCGAAGGTAGCGACGCTGTGTGCCT GTTGACCGATTGGGAAGAGTTCAAAAAGGTCGATCTCGTACAACTGG CCTCCCTTCTGCGCAAACCAGTTCTAATTGACGGTCGTAACGTATTCG CCGAGGAACAGATTCAAGGCTCCGGGCTGGAATACTATTCCGTGGGT CGTCCGCGGATGAGTGGGTGGAATCGGGATAAAGTGGAAGTTTAAAA AGTGACTGTGGCGGGGGTGGAGACGCTACACGTGCAGATTGTTCTTC CGATTGCTGTTACCCCCGGATTTTCTTGATTGTATAAAACATTCAAAGG TTAAAATCCGGGGATAAAGGCAAACGCTGACGCTTCTTCAGAACAATT CTGCCCGCTCCGCTGCTTTTCGCCGTGGAGTCACAGTTTTTAAACTCC CACCTTATGGGGGGAGTAGAAGGGCATCGGAAGGAGCGGCACGGTT GAAGCCGCTCCTTCCGATCAAGTACGCCATTGCGGTTGCCGAAGGCA ATCTGGGCAAGGCGTTACGAAACCTCGATTACTTGAAGCTTGCAAGC ACAGCGGGCAAGCTCTGAGTTCTCGAACATTTTACAGGCTGACCGGA CTTAATAATGTCCGGTCGTGCCTTTTGCACCACAAGGTATGATCGGGT ATGATGCATTTTTGAGTAGCGCAGGGATGGATTTGAATCTGGAGAAGC GGAGCGGTCGCCTTTGACTTGGGATTCTTACATATAAATGTCTTACAC AATCCAAAGAATCCCACGGCAACAGCGATCGTAAGATCAAATCCAGC CCGTAGCGGGCTACTCAACTTTGCCCCATCCGATCATCCAACCAATAT AGAGAGGAAGTGGAGAGTTATGAAGCTCGTACTTTTGTCTGGTGGTT CCGGTAAAAGATTATGGCCTTTGTCCAACGATTCCCGTTCCAAACAAT TTTTGAAGGTACTGGAGAGCCCGGAGGGGATTTCGGAATCCATGGTA CAGCGGGTATGGAGACAATTGGGAGACAACGGATTATCAGAATCCTC GTTTATTGCGACGGGGAGAGCACAGGTCGAAATGATCCAGAGTCAGG TCGGGCCGAATACGCGCATCATTGTTGAGCCGGAACGCAGAGACACC TTTCCTGCGATTGCGCTGACAGCGACTTATCTATATTCCATTGCTGGT GTTTCCCCTGATGAAATTATAGCTATATTGCCTGTAGATCCTTATGTTG AGGATGCTTTCTTCGCTTCCGTAGCCCAGCTGGAACATACACTCCAG GAGAGTGAAGGCAAGCTTGCTTTAATCGGGGTTGTCCCTTCCTATCCT TCAGAGAAATATGGTTACATTATTCCGAAAAACGATGCCACCCAAGGT AGCACTGGCTACCGCGAAGTGAGCCATTTTCAAGAGAAACCCGACCG GGAGCAGGCGGAGCGTCTGATTGAACGCAATGCTCTATGGAACTGCG GGGTTTTTGCATTCAAATTAAGTTATTTGCTGGATATCTTGGCATCCAA GGGTTTGCCGCTGAATTATGAAGAAATGCAAAAGCAATACGCATCTCT GGAAAAAATCAGCTTTGACTATGAAGTAGTCGAGAAGGAAAAGGATAT CGTAGTTCTTCCGTATGATGGTTTTTGGAAGGATTTGGGCACCTGGAA TACGCTCACAGAAGAAATGTCTAATCAACAAGTGGGCAGAGGTGTAG TAACGGAGGATTGCGTCAATACGAGTCTTATTAACGAGTTAGACATTC CTGTGGCGATCATTGGAACGAATGATCTGATTGTAGCTGCTAGCCCG GATGGTATTCTTGTGACCAATAAGGCTGAAAGTCCTCGCATTAAGGAA GTGCTCAAGGCTCATGATCAGCGTCCAATGTATGAAGAACGTCGTTG GGGGCAGTATCGGGTTGTGGATTATGTCAAATATGACGAGGGTAACG

Gene	Sequence 5'-> 3'
	AGGTATTGACCAAACGAATCCGCGTACGGAAAGGCAAAAATATCAGC TATCAGCTTCATTTCAAACGGAGTGAAATTTGGACCATTATTAGTGGT GAGGCGGAAATTATTTTGAATGAGAAGCTGCACAAAGTAAAAGCTGGT GATGTAGTTCGCATTCCTGAAGGAACGAAGCATAGCATTTTGGCTGTG ACCGATGTGGATTTTATTGAGGTGCAGACTGGTTCTGAATTGGTTGAA GAGGATATTGTACGCTTTTGTATGGATTGGAATGAAATTTCCCGTCAC CAATTTATCTCATAATTGTTTTCGTAAAAACAAAAAAAAGACGAATTCAT GAAAATTCATAGCCCGAACTGTCGAAATTGGGCGAAAACCTGTAAATT CATGAACATGACCTATTAAATTTTGGTAAAAAAAAGTCGATACTTAGGG TAAGAAGGAATCCAAACTCCTTCCAACGATAAACGCAGCCCTGTCGAA AGGCAGGTACGCAAAGCTAAAGGGCCTTCTACAAGGATGGCAGCCTG GCTACCGAAAGGAAGAGCGAAAGTGAGATTCTATTTTATTCCTCTGTA TCGACTGATACGTCATGTACTCCCCGTTGTTATGCTTTTTTCATTGTTG CCCGTCGGTTTTGGTGATTCTCTGTCTGTGGCGAATGCCGCTTCATTA CAACCTGTTAATTCGGACGCTTCTACTCAGGCCCGTGAACTATACAAT TATTTACTCAACATCTCAGGGAAAAAAATAGTCACCGGTCAGCATGAT TATTTGGAAAGTCCGGATGAATTAAGTAACAAGGTCCAAAAAATCAGT CAGGCATATGTTGGTCTGCACGGCTATGAGATGGGTGCCATCTCTGG TCAATCCGACACTACAGAAGCAGCACAACGTAAAAATGTCGTCGATAG TGCTATCCGTTGGAGCAGATCCGGTGGTATTGTGACGATGACTTTTCA TGAAGCGCTACCGGGTAGGCCACTGACTTGGGCCAATGTACAGGCG AAAGTAAGTCAGGCAGAGTTCAACAAATACGTGACCCCAGGTACTAC GCAGTATAACCTCCTCATTGCCGATCTGGATAAGGTCGCTGTATCATT AAAGCAGCTTCGTGATGCAGGGGTTCCGGTTTTGTGGAGACCTTACC ACGAAATGAATGGTGACTGGTTCTGGTGGGGAAACAAAACTAATTTTA ATCAGCTATGGAATATCATGTACGACCGTTTTACAAACACACATCAATT GAACAATCTGCTGTGGGTATGGAGTCCGAATGCGCCCAATTCGTATA CCGTCCCGTATACCCCGAAATACCCGGGGGATGACAAAGTGGACATT TTGGCAGTAGACATTTACAACAATGATTTTAAGCAAAGTCATTACGAAG GTCTGCTCGGGCTGGCACGGAATAAGCCGATTGCTATTGGGGAGCAC GGTGAGATGCCTAGTTCTGAGGTGTTGCAGGCTCAGCCTAAATGGGT ATACTCCATGACTTGGGGCAAGATGCTGACCGAGAATAACACAACAAA CCAAATACAGGATTATATGAATGATTCGAAGAGTTTGACACGGGATGA TGTGAAAAATGGGCTGGCAGCAATCGAACAACCTGGGACAGATGTAC CGACCTTACCGGACGAGGGACAAAGTGAGCCTATCCCGGCTCCGACT CCTCCTTCCGTCCCTGATGTGGTATATGTGAACGGGCTGCAGGGCGA ATATTTTAACAATATGAATTTGAATGGAAGTCCTGCTTTGATCCGTACA GATAGCAAGCTGGACTACAACTGGCGCGCCATTGCAGCTGATCCAAA GGTGAATGCCGATCAGTTCTCGGTTCGCTGGACGGGGAAAATCAAAC CACAATATAGCGAGACCTATACGTTCACAACGATATCCGATGATGGTA TTCGCGTATGGGTAGACGGCAAGTTGGTGATTGACAGCTGGTTCAAA CAAAGCTGGACGGAGCGTAAAGGCAGTATTACTTTGGAAGCAGGCAA AATGGTAGATTTAAAAGTGGAATACTATGATGAAAAAGGCGACGCGAT GGCACGTCTAATGTGGGAAAGTCAGCATGAAGCGAAAGCGGTGATTC CCGGAAACGCTTTGTTTCTTCCTTGATTGATACTCTTTCCCGGTCTAAA AATATGAGGAGAGAGGGT
spo0A	TTGCAAAAAATTGAGGTATTGTTGGCTGATGACAACCGGGAATTTACG AATTTGCTTGCTGAATATATTTCCGATCAGGAGGATATGGAAGTTACA

Gene	Sequence 5'-> 3'
	GGAATCGCCTATAACGGCGAAGAAGTGCTCCAACGCATCGCAGAATC CCGCAACATACCTGATGTACTTATTTTAGATATTATTATGCCTCATCTG GATGGTCTCGGTGTATTGGAGCGTTTAAGAGAAATGAACCTGACTCCA CAGCCGAAAATCATTATGCTGACTGCATTCGGTCAAGAAAATATTACA CAGAGGGCCGTACAGCTCGGGGCATCTTATTATATTTTGAAGCCGTTT GACATGGAAGTGCTGGCTAACCGTGTTCGCCAATTAGTGGGGCCGCA GTTAGTTAGCAGCAGTCCGGTGACCATTTCATCCATGCGGTCCAACG TTGTACCTATGGGTAAGACGAAAAACCTGGATGCCAGTATCACGGCC ATCATACATGAAATTGGTGTACCAGCTCACATCAAGGGCTATCAATAT TTACGCGAAGCCATTACGATGGTGTACAATAATATCGAAATTTTGGGA GCCATCACCAAAACATTATATCCGGCAATCGCCGAAAAGTTCAAAACG ACGGCATCTCGCGTAGAACGCGCCATTCGTCATGCTATCGAGGTAGC ATGGACCCGTGGAAATATCGACAGCATTTCACACCTGTTCGGGTACA CCATTAATATCAGTAAGTCCAAACCAACCAATTCAGAGTTCATAGCGA TGGTAGCAGATAAGCTGAGAATTGAGAACAAGGTGTCCTGA
spoIIE	ATGATGGAGAAATGGAACGTGGTACATTTTCCGGGGTTAAAGTCTAGA AAGTCCAAATCCAAAGAGCGTACGGAAGGTTGGTCTTCCAGACTCAG AAAACGGCTGGATTCACATAAAGTGGTGCAATGGGTAACAGCGCAAA GATGGATGCTGCTATTGTCGGCTATGGGAGTTCTACTGGGACGCGCC ACTATTTTGGACGAGCTTTCTCCTTTTGCCCTCGCTTATTTTGCGGTCA TTACGTTTCTTCGGCGTGACCTGATGCTGCCTGTAGCTGTATCCGTTG TGTTAGGCAGCCTGTTAGCTCCTTACCCATCGTTATTTATGATTGGTG GTGAGCTACTGGTATTTTATTTCCTGTATCGCGGGTTACGAACCTTCG ATCGTCTGGAACTATCCTATGCCCCGCTTATGGTTTTCTTGTCTGTGTT TATGGTCAATTTATTTAACGCCGTGATGGCTCCAACTTTCACATGGTAT GAACTTATCATTGTGGGAACGGATGCGGTTCTCAGCTTTATGCTTACG CTGGTATTCACACAGGCCATACCGTTGTTTACGTACCGCAAGAAGACA AGCCAGCTCAAAAATGAGGAAATTTTATGTCTTATTATTTTATTAGCCT CTGTGATGACGGGAGCAGTAGGATGGACTGTTCATTCATTGTCTGTC GATCATGTACTGTCACGATATTTGGTCTTGGTGTTTGCTTTGGTGGGA GGTGCTTCGCTGGGTGCATCGGTCGGTGTAATTACCGGTTTAATTTTA AGCTTGGCCAATATGTCTGCAATTGTGCAGATGAGTTTACTCGCTTTT GCAGGTTTGCTGGCAGGGATGCTGCGTGAAGGGAAGAAAGGTGCTG TTGCGCTTGGGATGCTGCTAGGCTCCTCCATTCTGACGATCTATTTGG ATGGGACCGGAGACGTGCTGACGTCAACCTGGGAGAGCTGTGCGGC GATTCTATTATTTTTAATGACCCCAAAAAGTTTCTTTAAAGCGGTTTCTA CCTATGTTCCCGGTACGCAAGACCATGCTAAATCTCAGCATGAATATG CCAAAAGGGTCCGCGATCTCACCGCAGAGCGAGTAACGAAGTTTTCC CGTGTATTCAGTCAGCTATCCCGCAGCTTTCACCAGATATCTGCAAGT GAGGGCGCGAAGTCGGATGGTGAAATCGAACACTTTATGAATGCGGT GGCGGAAGGGACATGTGCCTCCTGCTTCAAGTGTGAGCAGTGCTGG GACGGCAAGTTCATGCAAACCCATCAGTATATGACGGAAATGATGAG CGCCATTGAGGATAACCCGGATTTGGAGCCGGAACAGATACCCCCGC AATGGAGTAAGGCATGTGCTAAAACGGGCGCTGTCCTGCAAGTGATG AAGCAGCAGTATAGCCTCTACCAACATAATATGCAGTGGAAGCGCCA AGTATACGATAGTCGTCAGCTTGTTGCTGACCAGCTATCGGGCGTATC ACAAATCATGGAGGATTTAGCGCGAGAAATTCAGCGGGAAGGGCAGA CCATGCATCGGCAAGAGGAGCAAATACGGGAGGCTCTTGATAGACTG

Gene	Sequence 5'-> 3'
	GGCTTGTCCATCCAATCTATTGAGATTATTAATCTGGATGCAGGTCAG GTAGAAATTGAGATTGTCCATGCCTATACAAGGGGATTTGATGAATGT CGGAAAATCATTGCTCCACTGCTGTCGGACATACTGGACGAGCATATA GCTGTGATGCGCGAAACGGAGGCAGATCGCCGCCAAGGCTTGGCTA CAGTCATGTTCGGCTCTGCTAAAACGTATGAGATTGCTACAGGGGTA GCTAGTGCGGCCAAAGGAGGTGATTTTTTCTCAGGAGACAGCTACAG CATGATGGAGCTAGGGAACGGGACTTTTGCAGTCTCTTTAAGCGATG GTATGGGGAACGGAGAGCGCGCGCAGCAAGAGAGCAGTGCAGCGCT TAACATTTTGGAGGAATTACTGCAATCGGGTATGGACGAAAAATTGGC GATCAAGTCGGTTAACTCTGTCCTGATGCTGCGATCACCTGAGGAAAT GTATGCTACGGTCGATATGGCGCTGATTGATCAGTATACGGCCCAGA CAACCTTTATGAAAATTGGGTCTACGCCAAGCTTTATTAAAAGAGGAG AGGAGGTCATTCCCGTATCTGCCAGCAATTTGCCGATTGGCATTATAC AGGATATTGAAATTGATTTGGTATCCCTGCAGCTCCAGCCGGGTGATA TTCTCATTATGATTACAGATGGTATATACGACTCACCGGGTTATGCGG TTAATAAGGAAATATGGATGAAACGCATGATTCAGGAAATTGAAAGCG AAGATCCACAGCAGTTGGCCGACTGTTTACTAGAAAGGGTTATCCGAT ACCAACAAAATCAAATCTATGATGATATGACTGTAGTCGTTAGTAAAAT AGATCACTTCCGCCCAGAATGGGCAACTCTCCATGTACCTGGTATGA GGTGGTTGGAACGTCCACGGACGGTGAGCTAA
pfl	ATGTCGGTGATTGAGAGAGAATTACAAGAGCAACAATCTGGCTGGAG AGGTTTTAAAAAAGGCAAATGGACTAAAGAAGTCAACGTCAACGATTT TATTGAAACCAATATCTTACCTTATGTTGGCAATGAAGAATTTTTGGTT GGTCCTACGGCCAACACAACTGCATTGTGGGATATTGTATCCGATCTT ACCAAAAAAGAGTTGGCAAACGGCGGCGTGCTTGATGTAGACGTTAA TACGCCATCCACTATTATTTCACATAAACCGGGTTATCTAGACCGGGA TAAAGAGCAAATTGTCGGTGTACAAACAGATGCTCCATTCAAACGTTC TTTACAGCCGTTTGGCGGAATTCGAATGATGATCGACGCTTGCAAAGC CTATGGTTTTGAAGTGCCAGAAAGTATCATTGACATCTTCACTCATATT CGTAAAACTCACAACCAAGGTGTATTTGACGCTTATACAGATGAAATG AGAGCAGCTCGTAAAGCTGGTATCATTACTGGTCTGCCAGATGCATAC GGCCGTGGTCGTATTATCGGTGACTATCGTCGTGTAGCTTTATACGGT GCAGACTTCTTGATCCAAGATAAAAAGCTGCAACTGAAAAGCCTTGAA GTTGATTCCATGGAAGAAGACGTAATTCGTCTGCGTGAGGAAATTTCC GAGCAAATTCGTGCTTTGAATGAACTGAAACAAATGGCTGCTGAGCAC GGTTTTGATATTTCAAAACCTGCTAACAATGCGAAAGAAGCTTTCCAAT GGGTATATTTTGGATATCTTGCAGCAATCAAAGAACAAAATGGTGCAG CAATGTCCTTAGGACGTGTATCTTCCTTCCTTGATATTTACTCTCAACG TGATATGGAAGAGGGTACAATGACAGAAGAACAAGCTCAAGAACTGG TCGATCATTTTGTTATGAAATTGCGTATTGTTAAGTTCTTGCGCACACC TGACTATAATGAACTGTTCAGTGGAGATCCAACTTGGGTTACTGAGTC CATTGGTGGTATGTCCGTAACTGGAGAAACTCGGGTAACTAAAAACAG CTTCCGTTTCCTGCACACCCTCTACAATCTGGGACCTGCACCGGAAC CGAACCTGACTGTACTGTGGTCGGAAAAATTGCCTGAAGGCTTCAAAA AATATTGTGCGAAAGTTTCTATTGAAACTAGCTCCATTCAGTACGAGAA TGACGATCTGATGCGTCCGATTTATGGAGACGACTATGGTATTGCTTG TTGTGTGTCGGCAATGAAAATCGGTAAACAAATGCAATTCTTTGGTGC TCGTGCTAACTTGGCCAAAGCTCTGCTGTACGCAATTAACGGTGGTAA

Gene	Sequence 5'-> 3'
	AGATGAAAAATCTGGCGTCCAAGTAGGGCCTGAATTCCCAGCCATTA CAGGTGAAGTTCTGGATTATGAAGAAGTGTTGAGTCGCTTCAAACCTA TGATGGAATGGTTGGCAAAACTGTACATGAACACTTTGAATGTTATTC ACTACATGCATGATAAATATAGCTACGAACGCATTGAAATGGCGTTGC ATGATCGTGATATTCTCCGCACCATGGCTTGTGGTATTGCGGGTCTGT CAGTTGCTACCGACTCCTTGAGTGCCATTAAATTTGCAAAAGTTAAAC CAATCCGTAACGAAAAAGGGATTGCTGTTGACTTTGAAATTGAAGGCG AATACCCTTGCTACGGTAACAACGATGACCGTGTAGATAACATCGCTG TGGAACTGGTAGAGAGCTTCATGAGCATGATTCGCAAGCATAAAGCTT ATCGTAATGCGCTTCCTACTCAATCTGTACTGACTATTACTTCTAACGT GGTGTATGGTAAGAAAACAGGTACAACACCGGATGGACGTAAAGCAG GCGAACCGTTTGCACCAGGGGCAAACCCAATGCACGGCCGTGATAAA AAAGGTGCTTTGGCATCTCTGAGTTCCGTAGCCAAATTGCCTTATGAA CACAGTCTGGATGGTATCTCGAACACCTTCTCTATCGTGCCTAAAGCA CTGGGTAAAGAAGAAGAAACACGCAAAACAAACCTTGTTTCGATGATG GATGGCTATTTTGGAAGCCACGCACATCATCTGAATGTCAATGTGTTT GCTCGCGAACAACTGTTGGATGCAATGGAACACCCAGAAAACTATCC ACAGCTGACGATCCGTGTTTCCGGTTATGCCGTTAACTTTATCAAGCT GACTCGTGAACAACAATTGGATGTTATTAACCGTACTTTCCACGGTAC AATGTAA
ldh1	TTGCAAAAAATTGAGGTATTGTTGGCTGATGACAACCGGGAATTTACG AATTTGCTTGCTGAATATATTTCCGATCAGGAGGATATGGAAGTTACA GGAATCGCCTATAACGGCGAAGAAGTGCTCCAACGCATCGCAGAATC CCGCAACATACCTGATGTACTTATTTTAGATATTATTATGCCTCATCTG GATGGTCTCGGTGTATTGGAGCGTTTAAGAGAAATGAACCTGACTCCA CAGCCGAAAATCATTATGCTGACTGCATTCGGTCAAGAAAATATTACA CAGAGGGCCGTACAGCTCGGGGCATCTTATTATATTTTGAAGCCGTTT GACATGGAAGTGCTGGCTAACCGTGTTCGCCAATTAGTGGGGCCGCA GTTAGTTAGCAGCAGTCCGGTGACCATTTCATCCATGCGGTCCAACG TTGTACCTATGGGTAAGACGAAAAACCTGGATGCCAGTATCACGGCC ATCATACATGAAATTGGTGTACCAGCTCACATCAAGGGCTATCAATAT TTACGCGAAGCCATTACGATGGTGTACAATAATATCGAAATTTTGGGA GCCATCACCAAAACATTATATCCGGCAATCGCCGAAAAGTTCAAAACG ACGGCATCTCGCGTAGAACGCGCCATTCGTCATGCTATCGAGGTAGC ATGGACCCGTGGAAATATCGACAGCATTTCACACCTGTTCGGGTACA CCATTAATATCAGTAAGTCCAAACCAACCAATTCAGAGTTCATAGCGA TGGTAGCAGATAAGCTGAGAATTGAGAACAAGGTGTCCTGA
bdh	ATGCAAGCATTGAGATGGCATGGAGTAAAAGACTTACGATTGGAAAAC ATTGAGCAGCCCGCCGCTCTGGCAGGAAAAGTAAAAATCAAAGTAGA ATGGTGTGGCATTTGCGGAAGTGATCTTCACGAATATGTAGCAGGAC CAATCTTTATACCCCAAGATGCTCAGCATCCATTGACTGGCGAAAAAG CACCGATCGTAATGGGACATGAATTCTCCGGACAAGTCGTCGAAATC GGCGAAGGTGTAACCAAGATTCAGGTTGGCGACCGTGTCGTTGTAGA ACCGGTTTTTGCATGTGGAGAATGTGATGCATGTAAACAAGGCAAATA TAATCTTTGCGATAAAATGGGCTTCCTCGGTCTGGCAGGCGGCGGCG GTGGATTTTCTGAATATGTCGCAGCTGACGAGCACATGGTTCACAAAA TTCCAGAAAGCGTATCTTTCGAGCAAGGCGCATTGGTAGAGCCTTCG

Gene	Sequence 5'-> 3'
	GCCGTTGCTTTGTATGCTGTTCGTCAAAGCCAACTGAAAGTCGGCGAT AAAGCTGTAGTATTTGGCGCTGGTCCTATCGGATTGCTGGTTATTGAA GCGTTGAAAGCTTCGGGCGCATCTGAGATTTATGCAGTAGAGCTTTCT GAGGAGCGTAAAGCTAAAGCTGAAGAGTTGGGTGCTATTGTGCTTGA TCCTAAGACTTATGATGTGGTGCAAGAGCTGCACAAACGGACTAACG GTGGCGTAGATGTAGCCTATGAAGTCACTGGAGTACCTCCTGTGCTG ACTCAAGCTATTGAATCTACTAAAATCAGCGGACAAATCATGATCGTC AGCATTTTTGAAAAAGAAGCTCCGATTAAACCGAACAATATTGTCATGA AGGAACGCAATCTGACGGGTATTATCGGCTACCGTGATGTATTCCCG GCTGTTATCAGCCTGATGGAAAAAGGTTATTTCCCTGCCGACAAGCTG GTTACCAAACGTATTAAGCTAGAAGAAGTGATTGAGCAAGGTTTTGAA GGTCTCCTGAAAGAGAAAAATCAGGTTAAAATCCTGGTATCTCCGCAA GCTTAA

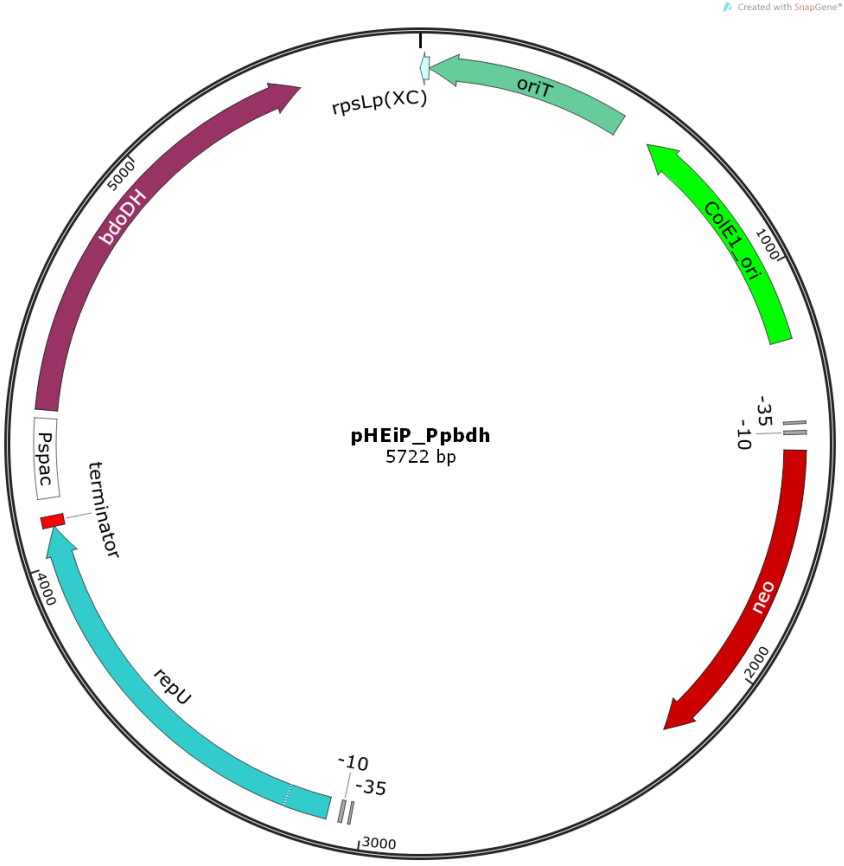


Figure S2: Plasmid map of pHEiP_Ppbdh. Conjugational plasmid was based on the pUB110 backbone and pCasPP (Rütering et al., 2017) and was constructed by replacing the Cas9-gRNA expression cassette by Pspac-*bdh* using Gibson isothermal assembly. Neomycin was used as a selection marker.

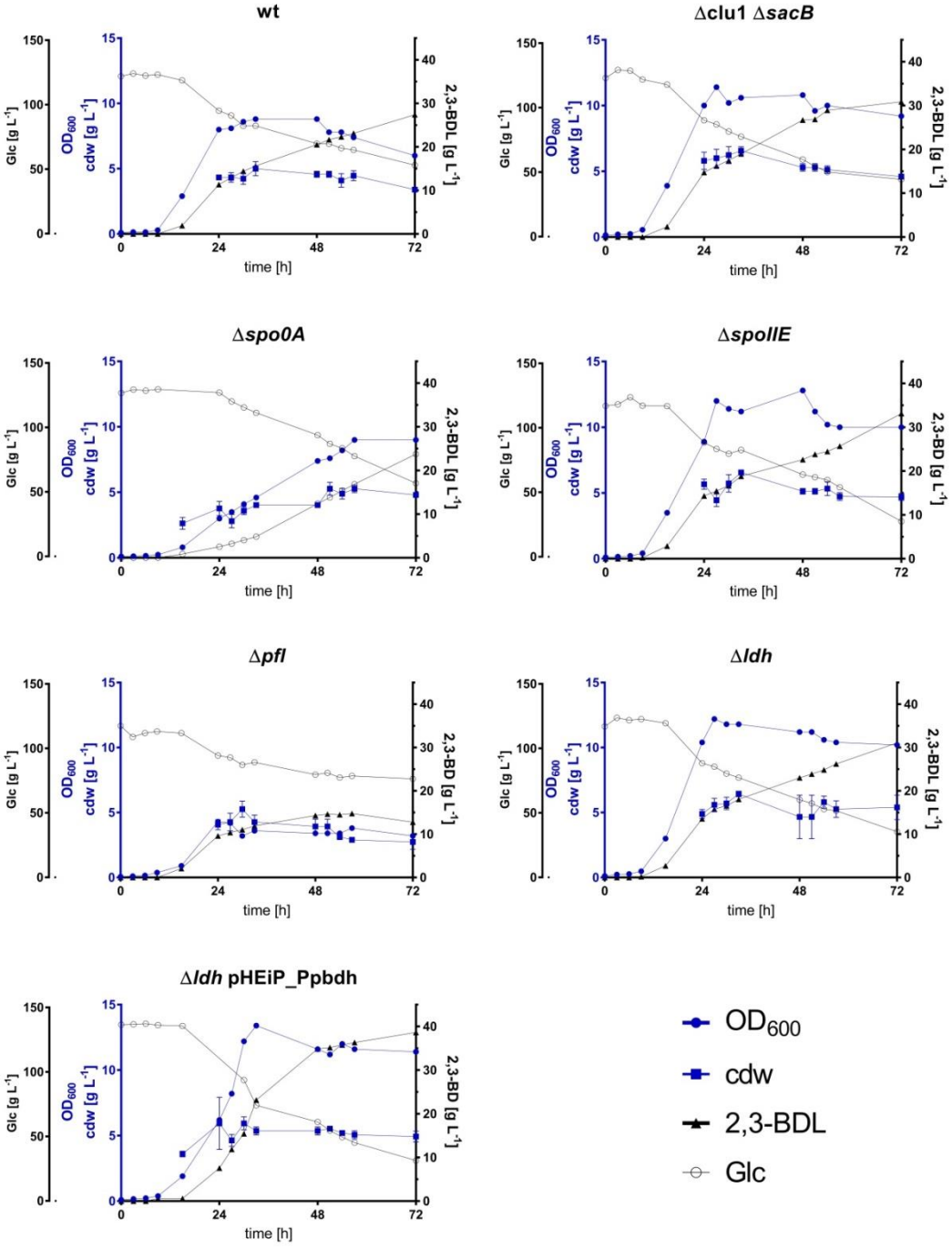


Figure S3: Overview of growth and 2,3-BDL production of wt and knock out strains of *P. polymyxa* DSM 365 from single batch fermentations. 2,3-BDL: 2,3-R-R-butanediol (black triangles), CDW: cell dry weight (blue rectangles), OD_{600}: optical density at 600 nm (blue circle); Glc: Glucose (unfilled circle)

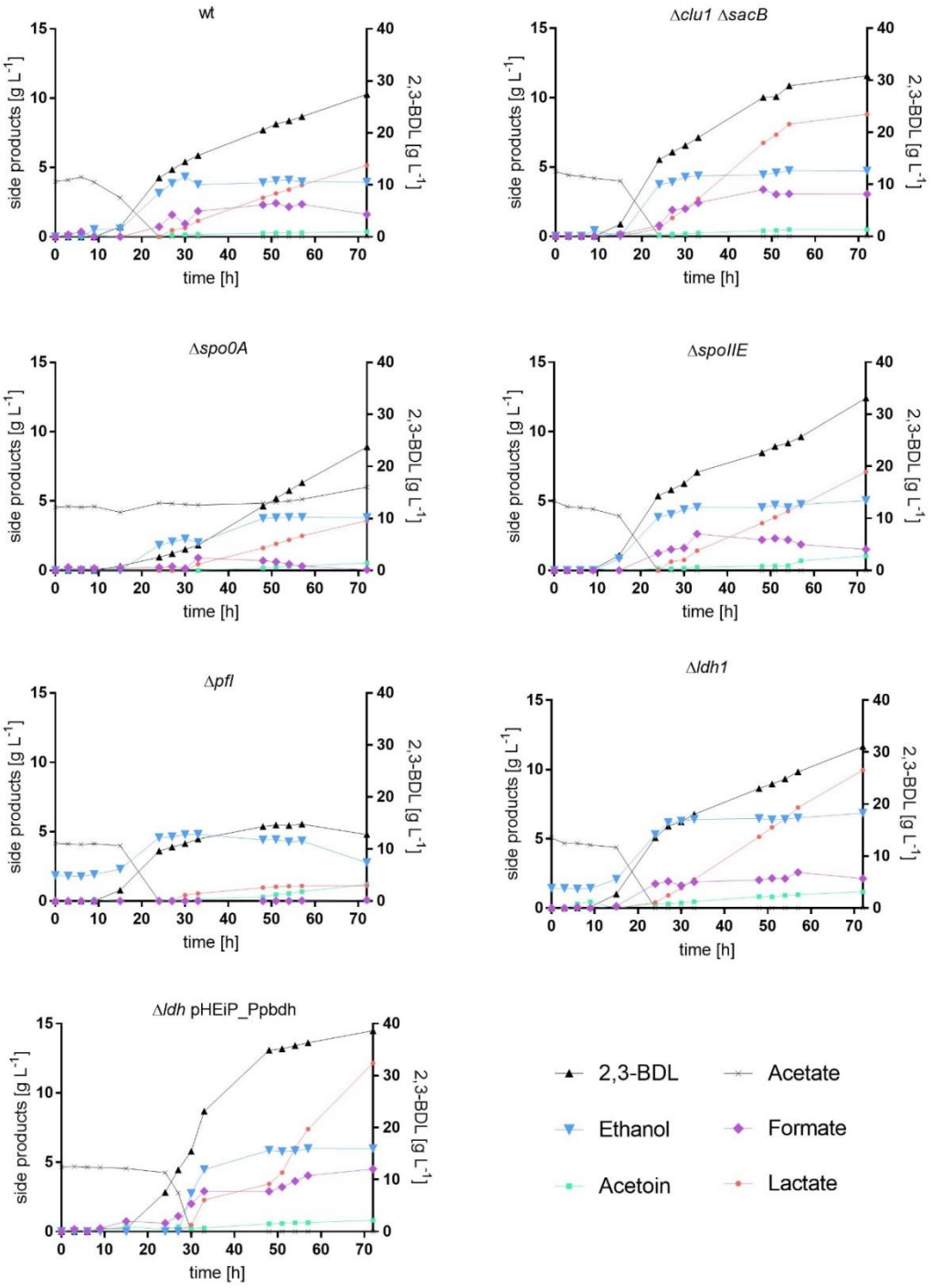


Figure S4: Overview of by-product formation of wt and knock out strains of *P. polymyxa* DSM 365 from single batch fermentations over the course of 72 h cultivation.

Table S4: Colony forming units of *P. polymyxa* wt and knockout variants with disrupted sporulation cascade after sporulation assay. Cultures grown in LB medium (37 °C, 280 rpm, 96 h) were subjected to heat treatment at 85 °C for 0, 30 or 45 min.

[cfu ml^{-1}]

Time [min]	wt	*Δspo0A*	*ΔspoIIE*
0	$1.00 \cdot 10^5$	$1.21 \cdot 10^6$	$1.54 \cdot 10^6$
30	$2.86 \cdot 10^4$	0	0
45	$3.78 \cdot 10^3$	0	0

References:

Rütering, M., Cress, B.F., Schilling, M., Rühmann, B., Koffas, M.A.G., Sieber, V., Schmid, J., 2017. Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. Synth. Biol. 2. https://doi.org/10.1093/synbio/ysx007

7.1.2. Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365

Table S 1: Bacterial strains used in this work.

Bacterial Strains	Genotype	Reference
E. coli S17-1	Conjugation strain; recA pro hsdR RP42 Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC 47055
P. polymyxa DSM 365	wild type	DSMZ
P. polymyxa DSM 365 *ΔpepC*	DSM 365 *ΔpepC*	This study
P. polymyxa DSM 365 *ΔpepD*	DSM 365 *ΔpepD*	This study
P. polymyxa DSM 365 *ΔpepE*	DSM 365 *ΔpepE*	This study
P. polymyxa DSM 365 *ΔpepF*	DSM 365 *ΔpepF*	(Rütering et al., 2017)
P. polymyxa DSM 365 *ΔpepG*	DSM 365 *ΔpepG*	This study
P. polymyxa DSM 365 *ΔpepH*	DSM 365 *ΔpepH*	This study
P. polymyxa DSM 365 *ΔpepI*	DSM 365 *ΔpepI*	This study
P. polymyxa DSM 365 *ΔpepJ*	DSM 365 *ΔpepJ*	(Rütering et al., 2017)
P. polymyxa DSM 365 *ΔpepK*	DSM 365 *ΔpepK*	This study
P. polymyxa DSM 365 *ΔpepL*	DSM 365 *ΔpepL*	This study
P. polymyxa DSM 365 *ΔpepQ*	DSM 365 *ΔpepQ*	This study
P. polymyxa DSM 365 *ΔpepT*	DSM 365 *ΔpepT*	This study
P. polymyxa DSM 365 *ΔpepU*	DSM 365 *ΔpepU*	This study
P. polymyxa DSM 365 *ΔpepV*	DSM 365 *ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepDJ*	DSM 365 *ΔpepD ΔpepJ*	This study
P. polymyxa DSM 365 *ΔpepFJ*	DSM 365 *ΔpepF ΔpepJ*	This study
P. polymyxa DSM 365 *ΔepsO*	DSM 365 *ΔepsO*	This study
P. polymyxa DSM 365 *Δfcl*	DSM 365 *Δfcl*	This study

Table S 2: Plasmids used and constructed in this work. Each plasmid was used for the transformation of E. coli S17-1 and furthermore for the conjugational transfer to P. polymyxa DSM 365. Before EPS production, each knock-out strain was cured from the specific plasmid resulting in the strains listed in Table S 1.

Plasmids	Description	Reference
pCasPP	*P. polymyxa* CRISPR-Cas9 genome editing plasmid; PsgsE-*cas9*; Pgapdh-off-target gRNA; *neo*; *oriT*; *repU*	(Rütering et al., 2017)
pCasPP H_pepC	*pepC* targeting knock-out plasmid containing repair template	This study
pCasPP H_pepD	*pepD* targeting knock-out plasmid containing repair template	This study
pCasPP H_pepE	*pepE* targeting knock-out plasmid containing repair template	This study
pCasPP H_pepF	*pepF* targeting knock-out plasmid containing repair template	(Rütering et al., 2017)

Plasmids	Description	Reference
pCasPPH_pepG	*pepG* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepH	*pepH* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepI	*pepI* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepJ	*pepJ* targeting knock-out plasmid containing repair template	(Rütering et al., 2017)
pCasPPH_pepK	*pepK* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepL	*pepL* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepQ	*pepQ* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepT	*pepT* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepU	*pepU* targeting knock-out plasmid containing repair template	This study
pCasPPH_pepV	*pepV* targeting knock-out plasmid containing repair template	This study
pCasPPH_epsO	*epsO* targeting knock-out plasmid containing repair template	This study
pCasPPH_fcl	*fcl* targeting knock-out plasmid containing repair template	This study

Table S 3: Oligonucleotides used in this study. Overhangs used for Golden Gate Assembly are depicted in lower case letters. Restriction sites used for cloning are underlined. For each KO construct, two sets of sgRNAs were designed, tested and listed below if successful knock-outs were obtained

Primer	Sequence 5'→3'
pepC_sgRNA1_fw	acgcGAGCATGGAAAATTTACCGG
pepC_sgRNA1_rev	aaacCCGGTAAATTTTCCATGCTC
pepC_sgRNA2_fw	aaacCCTTGATATGCTGCTGTCGT
pepC_sgRNA_2_rev	acgcACGACAGCAGCATATCAAGG
pepC_Harms_US_fw	GGCCTCTAGACCAAAATTCGATGAAGCAAGGAGAATGAGTGCG
pepC_Harms_US_rev_OE	CGCTTGAATGCATCCATCAATAAGCATCCCGTACCCCCTCTCAACCTGCCGTGG
pepC_Harms_DS_fw_OE	CCACGGCAGGTTGAGAGGGGGTACGGGATGCTTATTGATGGATGCATTCAAGCG
pepC_Harms_DS_rev	GGCCTCTAGACCATCTTTCACCCCGCAGCTGACTCG
pepC_KO_proof_fw	GCAACTCATGGAGCAGGCCAGTGCGACG
pepC_KO_proof_rev	GCTGTATGGTGTTTATTCTAGTAGTCCAAGG
pepD_sgRNA1_fw	acgcCAAAGTGTTATACATCACGG
pepD_sgRNA1_rev	aaacCCGTGATGTATAACACTTTG
pepD_sgRNA2_fw	acgcATGAAGCGGGAACATAACGG
pepD_sgRNA_2_rev	aaacCCGTTATGTTCCCGCTTCAT
pepD_Harms_US_fw	CCGGTCTAGAGGCTATTATGGGTCGCTACGTG
pepD_Harms_US_rev_OE	AGTCCAAGGATAGTTACTCATGTTCCCAATAAGCATCCTTGGAGAACAGC

Primer	Sequence 5'→3'
pepD_Harms_DS_fw_OE	TTCTCCAAGGATGCTTATTGGGAACATGAGTAACTATCCTTGGACT
pepD_Harms_DS_rev	AATTTCTAGACGTTACCGCGCCAAGACCTC
pepD_KO_proof_fw	CCGGATCAGCAGGACGGACG
pepD_KO_proof_rev	GCCCGCAACCGATATATCCC
pepI_sgRNA1_fw	acgcAGCATTATGAGCTAACCAAG
pepI_sgRNA1_rev	aaacCTTGGTTAGCTCATAATGCT
pepI_Harms_US_fw	aattTCTAGAGGCTGCGCTCTTTTTCTTCATC
pepI_Harms_US_rev_OE	GCAATACGAAATCCGTAGCGGGTGCCTCCTTCTTTAT
pepI_Harms_DS_fw_OE	AGGAGGCACCCGCTACGGATTTCGTATTGCGGCAAC
pepI_Harms_DS_rev	ttaaTCTAGAGGCTTCACTGTCCGTTCCCT
pepI_KO_proof_fw	GTTGCTGGCTACCATGTTCGG
pepI_KO_proof_rev	CCTTGCATTCGGTAGAAGTTCACG
pepK_sgRNA2_fw	acgcACAATCGGCTTCAGATACGG
pepK_sgRNA_2_rev	aaacCCGTATCTGAAGCCGATTGT
pepK_Harms_US_fw	aattTCTAGACATACGATTGGCAGCAAGCAG
pepK_Harms_US_rev_OE	TTTAACGCGAATGTCGGTTTCGCCGCCGCCTTTCTGTCA
pepK_Harms_DS_fw_OE	AAAGGCGGCGGCGAACCGACATTCGCGTTAAAGGATGAT
pepK_Harms_DS_rev	AATTTCTAGAGCTGCTCGACCTTTGCGACGGC
pepK_KO_proof_fw	GCTCTAATCTGTAGTGACGAGCAG
pepK_KO_proof_rev	CATCAAATTTGCGGGCATGGG
pepL_sgRNA1_fw	acgcGGAGGCTCCTATCTTCACAA
pepL_sgRNA1_rev	aaacTTGTGAAGATAGGAGCCTCC
pepL_sgRNA2_fw	acgcGTTATTGTCACACACCGATG
pepL_sgRNA_2_rev	aaacCATCGGTGTGTGACAATAAC
pepL_Harms_US_fw	AATTTCTAGAGCGCATCGTTCCATTGGACG
pepL_Harms_US_rev_OE	CCAGCTTCATTCCGTCTAAATCATCCTTTAACGCGAATGTCGG
pepL_Harms_DS_fw_OE	TAAAGGATGATTTAGACGGAATGAAGCTGGCTGTAATT
pepL_Harms_DS_rev	AATTTCTAGAGCGCCTGCCTGAATGAGTGT
pepL_KO_proof_fw	CGGAGAACATATCGACCCGCTC
pepL_KO_proof_rev	GGCCAGTTGTACGAGATCGACC
pepE_sgRNA1_fw	acgcGTCGATTCTCTGTATCAGGG
pepE_sgRNA1_rev	aaacCCCTGATACAGAGAATCGAC
pepE_Harms_US_fw	TTAATTTCTAGATCATGTACGCAATACGTGGCTCG
pepE_Harms_US_rev_OE	TCACTCTGCTCATCGGTTCCATCTTTCACCCCGCAGCTGA
pepE_Harms_DS_fw_OE	GGGTGAAAGATGGAACCGATGAGCAGAGTGAACATGTTCG
pepE_Harms_DS_rev	TTAATTTCTAGATGCCAGCGTGTAAGGAATCAGTAGC
pepE_KO_proof_fw	TGTTGGTCGGCGACATGACTTTAGT
pepE_KO_proof_rev	ATGGCTTTTCCGCCAGATGAATACC
pepG_sgRNA1_fw	actcCCATGCTAATGGAAAAACGG
pepG_gRNA1_rev	aaacCCGTTTTTCCATTAGCATGG
pepG_Harms_US_fw	AATTTCTAGAACGGGATATATCGGTTGCGGG

Primer	Sequence 5'→3'
pepG_Harms_US_rev_OE	ATAGCCGCGCGTTTCTCACCCCTGACCTCC
pepG_Harms_DS_fw_OE	GGTGAGAAACGCGCGGCTATCTGATCTGGATG
pepG_Harms_DS_rev	AATTAATCTAGACCGAAATGTTATGTCCACCGCCT
pepG_KO_proof_rev	CAATCAGCATCAGCATCATCGC
pepT_sgRNA1_fw	acgcAGAACGTGAGGAAAACCGTG
pepT_sgRNA1_rev	aaacCACGGTTTTCCTCACGTTCT
pepT_sgRNA2_fw	acgcTATGTCGTGGGATCGCATTG
pepT_sgRNA_2_rev	aaacCAATGCGATCCCACGACATA
pepT_Harms_US_fw	AATTTCTAGACCCGGATCATCGCTTTACAGT
pepT_Harms_US_rev_OE	TGCTGACCTGTTGGCGGGGTTTAGAGACCGCTGCG
pepT_Harms_DS_fw_OE	CTGAAGAAGCGCTAGGCCAACAGGTCAGCAGACAG
pepT_Harms_DS_rev	AATTTCTAGAGTCCTCCACGACTTCCTCCAG
pepT_KO_proof_fw	CGCAGCTCAAACGTGGGCTA
pepT_KO_proof_rev	GGGCTACTGCACAGCGAATC
pepU_sgRNA1_fw	acgcAGAAACCGAACATCTTCCTG
pepU_sgRNA1_rev	aaacCAGGAAGATGTTCGGTTTCT
pepU_sgRNA2_fw	acgcTTTACCGCGATACGATCCAG
pepU_sgRNA_2_rev	aaacCTGGATCGTATCGCGGTAAA
pepU_Harms_US_fw	AATTTCTAGAGGATGCTTATGGTCTGGTCATTACT
pepU_Harms_US_rev_OE	CACATTCACTTCCTGTCTGCTGACCTGTTGGC
pepU_Harms_DS_fw_OE	CAGCAGACAGGAAGTGAATGTGAAGCAGCTGAT
pepU_Harms_DS_rev	AATTTCTAGAGCACGCAGCTGATGAATTTTATCC
pepU_KO_proof_fw	GCGTATTATCTGGCTCCACGG
pepU_KO_proof_rev	GTTGTAAAGCTTAGTCGGCCTTGT
pepV_sgRNA1_fw	acgcTTACCCTGGAATCCGCATCG
pepV_sgRNA1_rev	aaacCGATGCGGATTCCAGGGTAA
pepV_Harms_US_fw	aattTCTAGAAAGAATCAAGAAGCTGCTTTTACAGG
pepV_Harms_US_rev_OE	TTTACCGACTGCATCCGCAATCCTTCCAACT
pepV_Harms_DS_fw_OE	ATTGCGGATGCAGTCGGTAAAAGCAGGATGCA
pepV_Harms_DS_rev	aattTCTAGAGCCAGTCCCCATATACCAATCG
pepV_KO_proof_fw	AGCGAACATGGCAAAGCTGG
pepV_KO_proof_rev	GGGGCATCCCTTACGTCATCT
fcl_sgRNA1_fw	acgcGATACATCTACCAACTTACG
fcl_sgRNA1_rev	aaacCGTAAGTTGGTAGATGTATC
fcl_sgRNA2_fw	acgcACAAACGAATGTGATTGATG
fcl_sgRNA2_rev	aaacCATCAATCACATTCGTTTGT
fcl_US_harms_fw	AATTAATCTAGAGACGTGATCTTAATGGATACGCCG
fcl_US_harms_rev_OE	CCACCGTTTTCTTTAATACGCATCCTTGGAGCCT
fcl_DS_harms_fw_OE	CGTATTAAAGAAAACGGTGGGGGAATTGG
fcl_DS_harms_rev	AATTTCTAGATCTCTTACACGACGACAGAAGC
fcl_Koproof_fw	AATCCGTCCGAGATGCTCAG
fcl_Koproof_rev	CCTTGCGTACAGCGAAGAGA

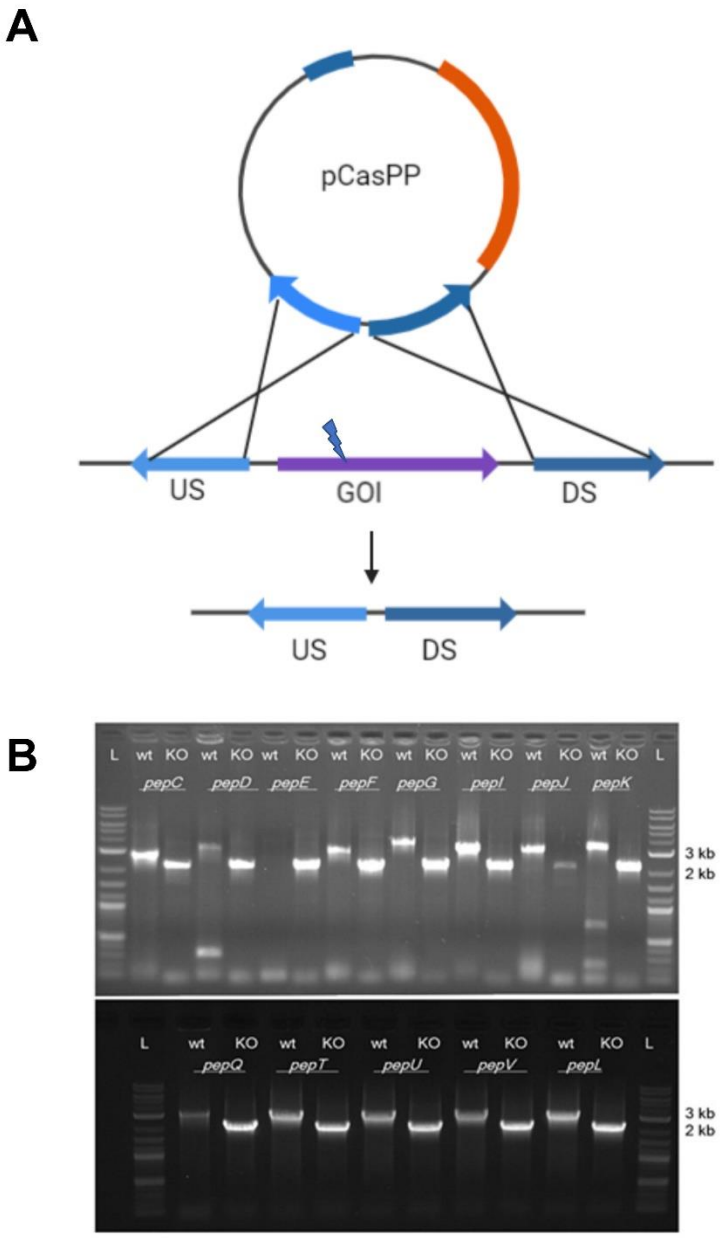


Figure S 1: A: Schematic overview of CRISPR-Cas9 mediated genome editing used for gene deletions in this study. A double strand break was induced by guiding a Cas9 effector to a gene of interest (GOI) which was repaired by providing a homologous repair template on the plasmid containing the upstream (US) and downstream (DS) sequence of the target gene. B: CRISPR-Cas9 mediated knock-outs of P. polymyxa DSM 365. The 1 % agarose gel shows the PCR products of the wildtype (wt) and knock-out mutants (KO) for genes of the paenan cluster (pepC – pepV). For cPCRs, primers (KO_proof in Table S3) located on the genome, outside of the flanking repair template were used for both amplicons resulting in ~ 2 kb amplicons for successful knock-outs, while the wildtype strain shows an amplicon between 2.9 kb – 3.5 kb depending on the size of the respective gene. L: NEB 1 kb Plus DNA ladder.

Table S4: Overview of EPS production in bioreactor cultivation. Genes encoding glycosyltransferases or Wzy polymerases (pepE, pepG) were individually knocked out and compared. Maximum growth rates (μmax) were determined during exponential growth phase

Strain	Titer [g/L]	Yield YPS [g_{EPS}/g_{Glc}]	μ_{max} [h^{-1}]
WT	5.024	0.135	0.508
ΔpepC	1.353	0.049	0.516
ΔpepD	2.326	0.049	n.d.
ΔpepE	2.104	0.060	0.502
ΔpepF	1.440	0.053	0.486
ΔpepG	7.138	0.123	0.511
ΔpepI	3.330	0.111	0.458
ΔpepJ	1.382	0.045	0.465
ΔpepK	1.851	0.046	0.468
ΔpepL	4.167	0.139	0.475
ΔpepQ	2.500	0.083	0.459
ΔpepT	2.017	0.081	0.611
ΔpepU	4.392	0.113	0.549
ΔpepV	1.591	0.029	0.506

n.d.: not determined due to lack of data points in exponential phase

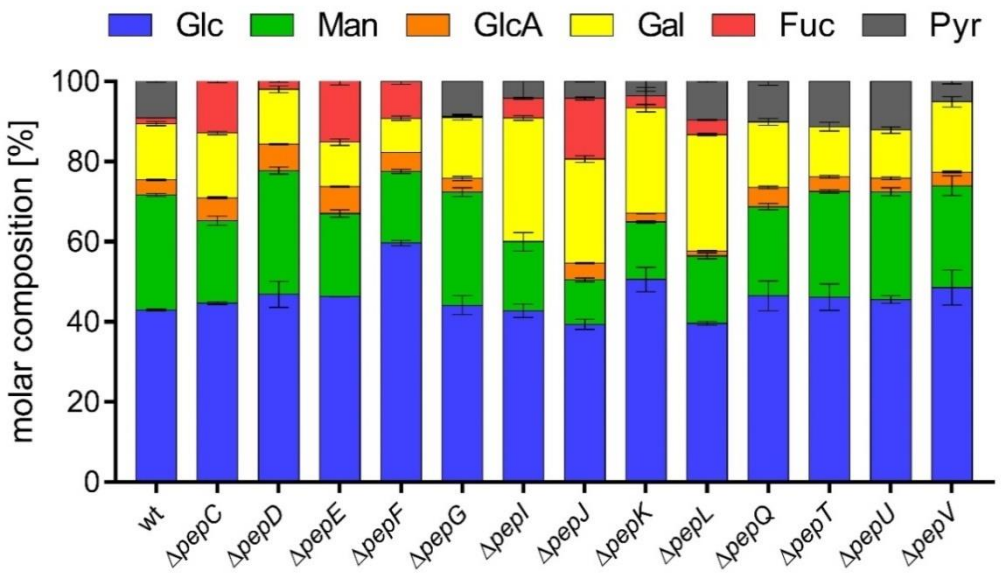


Figure S 2: Carbohydrate fingerprints of single target knock-outs of glycosyltransferases and Wzy-polymerases (pepE & pepG) of the paenan cluster of P. polymyxa DSM 365

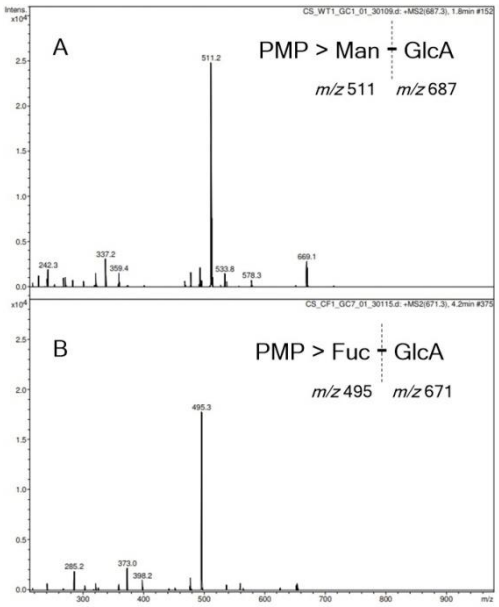


Figure S 3: MS/MS spectra of GlcA-dimers obtained in HT-PMP analysis. For both, the GlcA-Man dimer (A) and the GlcA-Fuc dimer (B), fragmentation resulted in the initial loss of m/z 176. Consequently, GlcA can be positioned at the non-reducing end of the dimer. Due to the early retention time of the GlcA-Hex dimer (A) at 1.8 min, the hexose was identified as mannose.

Table S 5: MS/MS fragmentation spectra of fucose containing oligosaccharides were analysed to determine the monomer sequence. EPS variants were obtained by partial hydrolysis of EPS variants from P. polymyxa DSM 365.

RT [min]	m/z	Oligosaccharide sequence
2.9	671	GlcA-Fuc-PMP
2.8	657	Hex-Fuc-PMP
3.4	833	Hex-GlcA-Fuc-PMP
5.3	995	Hex-Hex-GlcA-Fuc-PMP
6.1	995	Hex-GlcA-Fuc-Hex-PMP
7.8	1,317.5	Hex-GlcA-Fuc-Hex-GlcA-Fuc-PMP

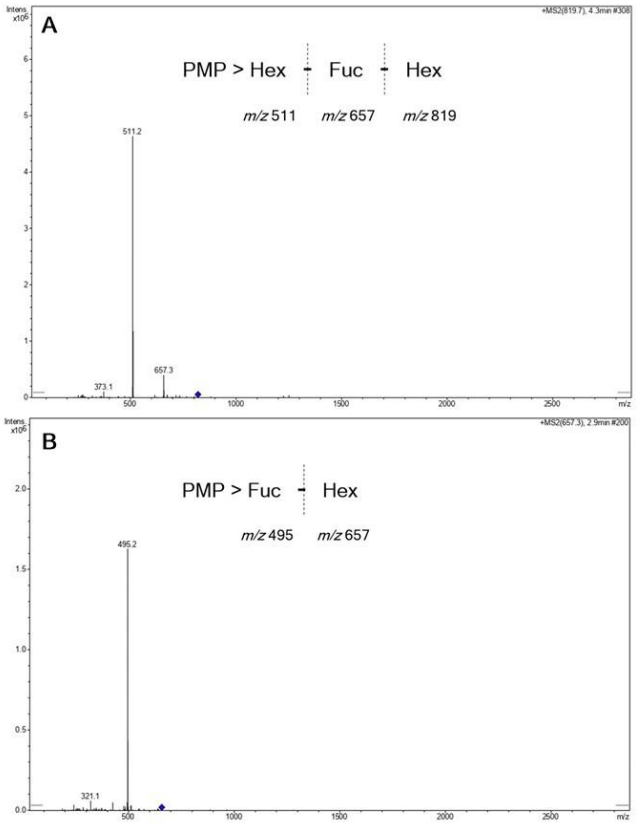


Figure S 4: MS-MS spectra of the major peaks obtained from lithium degradation of the EPS of P. polymyxa DSM 365 ΔpepFJ. A: Fragmentation of the peak with m/z 819 (4.3 min) indicates a Hex-Fuc-Hex trimer. B: Fragmentation of the smaller peak with m/z 657 (2.9 min) indicates a Hex-Fuc dimer.

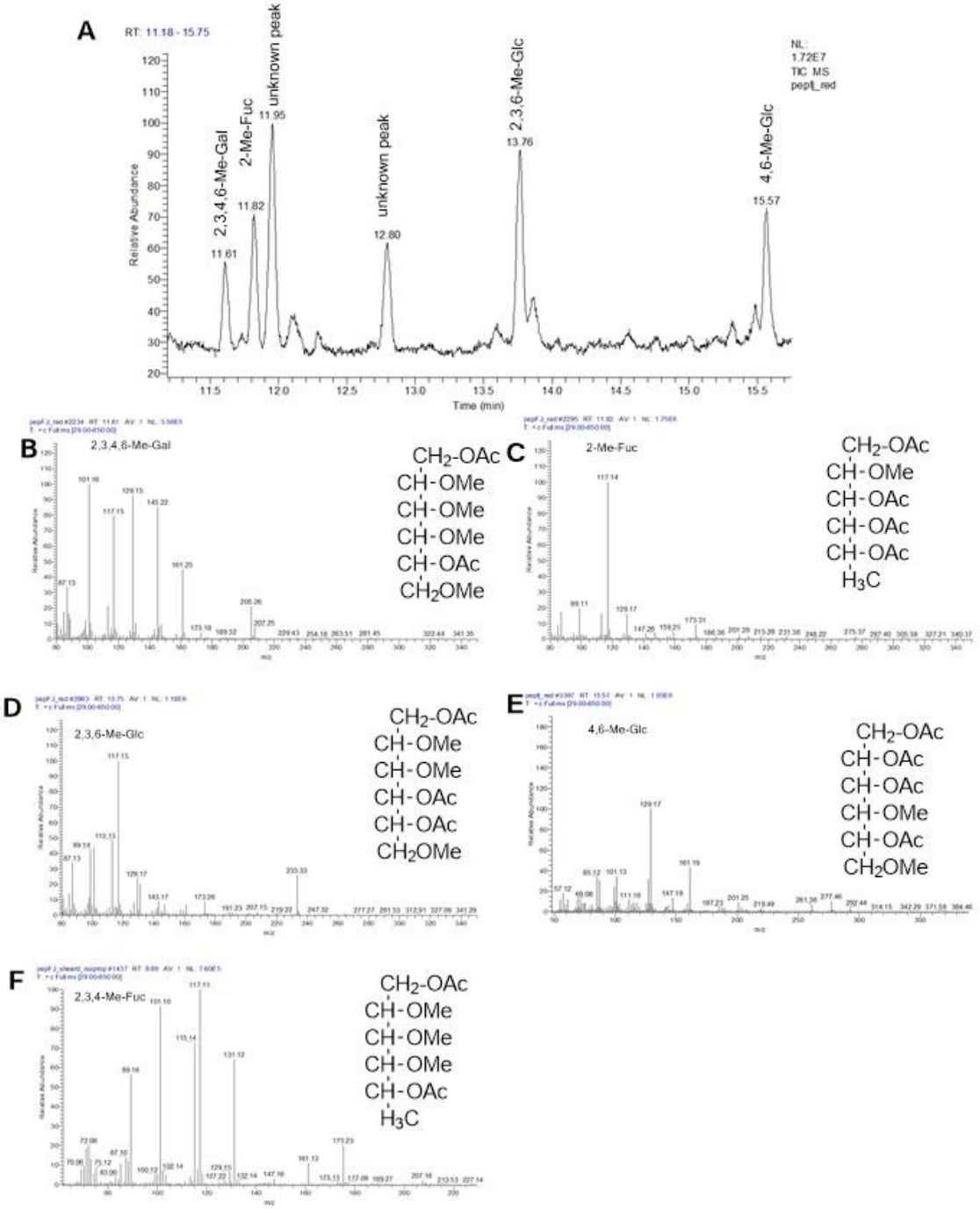


Figure S 5: Methylation analysis of the EPS obtained by the P. polymyxa DSM 365 ΔpepFJ variant. The sample was reduced prior to methylation to enable detection of GlcA A) Five partially methylated alditol acetates (PMAAs) were detected and identified based on their retention times and MS spectra. Retention time standards were prepared as previously described (Sassaki et al., 2005). MS fragmentation of unknown peaks (before 11.50 min, 11.95 min and 12.80 min) did not correspond to PMAAs. B) MS spectrum of 2,3,4-Me-Gal. C) MS spectrum of 2-Me-Fuc. D) MS spectrum of 2,3,6-Me-Glc. E) MS spectrum of 4,6-Me-Glc. F) MS spectrum of 2,3,4-Me-Fuc obtained from a mechanically sheared and heat precipitated methylation sample to remove protein impurities. 2,3,4-Me-Gal: 1,5-acetyl-2,3,4,6-methyl-galactitol etc.

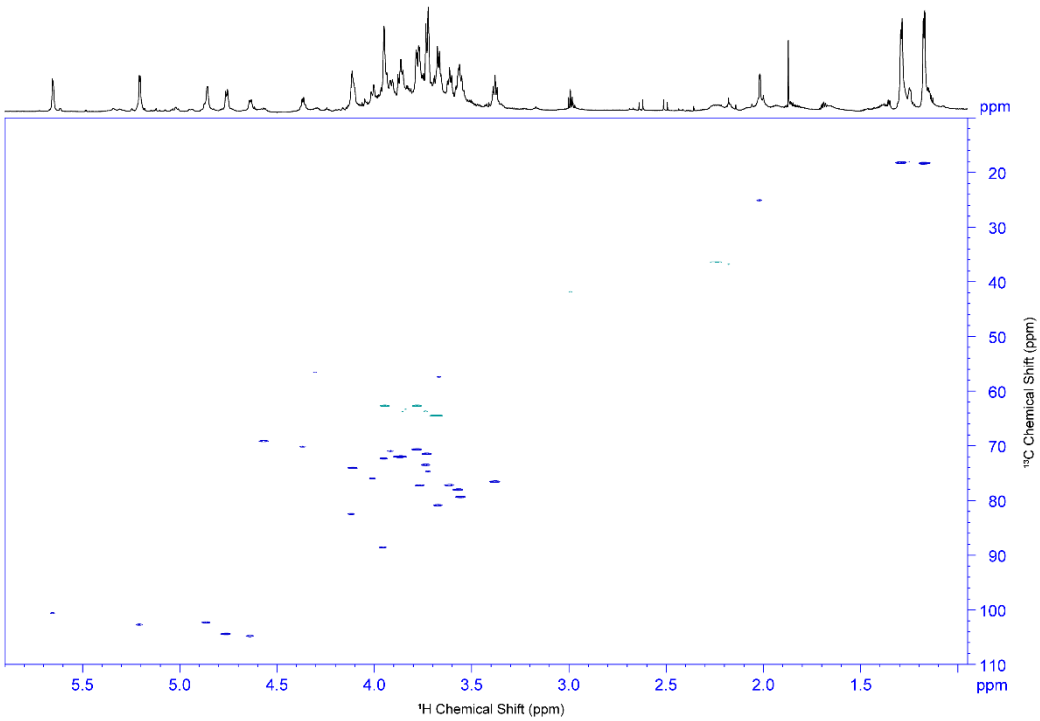


Figure S 6: ^{1}H-^{13}C HSQC spectrum with multiplicity editing of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

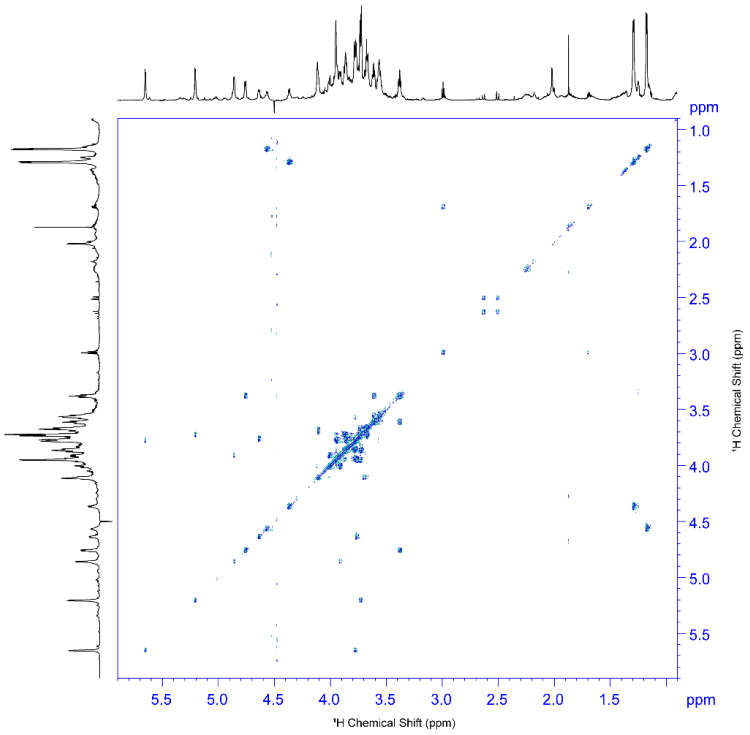


Figure S 7: ^{1}H-^{1}H COSY spectrum acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm).

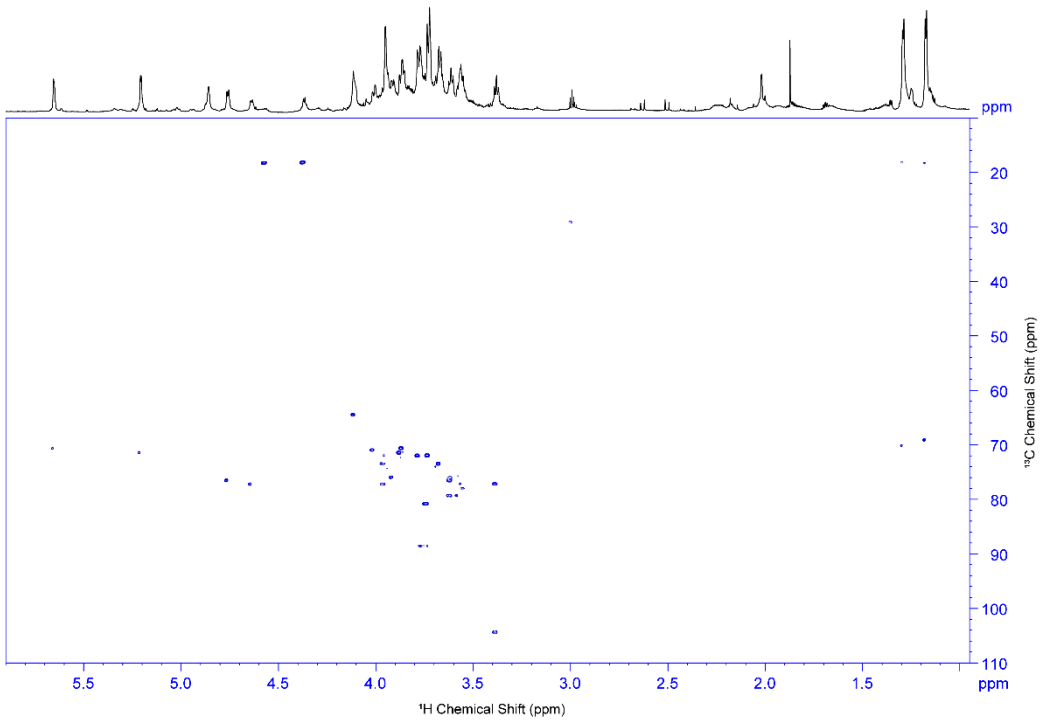


Figure S 8: ^{1}H-^{13}C H2BC spectrum of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

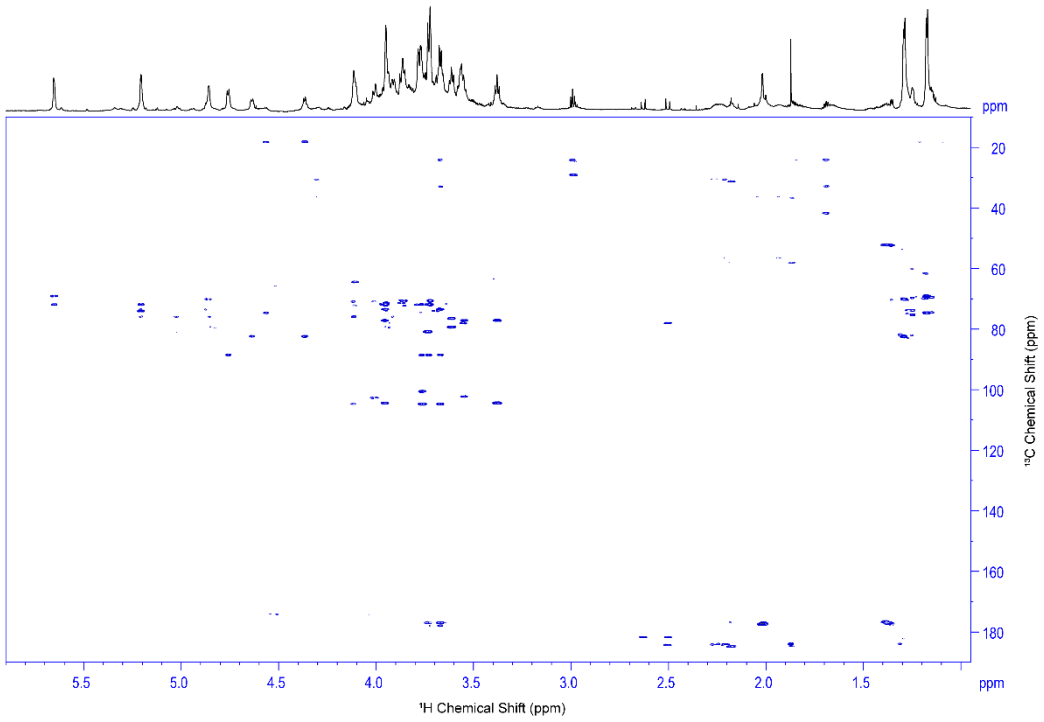


Figure S 9: ^{1}H-^{13}C HMBC spectrum of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

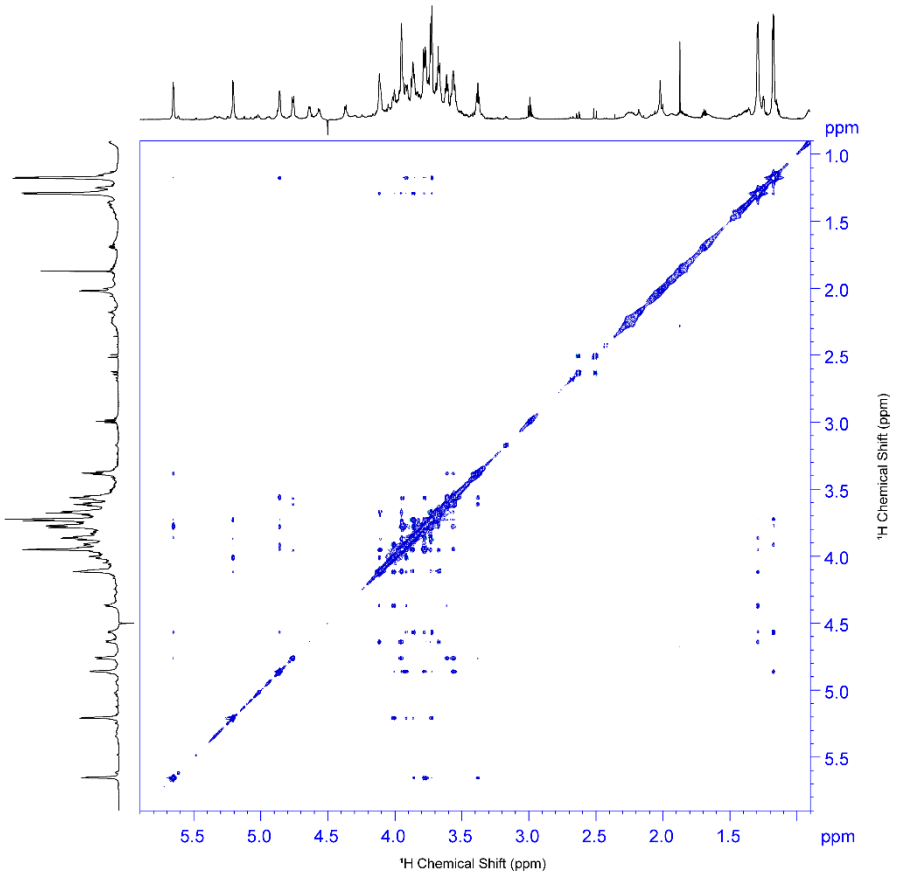


Figure S 10: ^{1}H-^{1}H NOESY spectrum of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm).

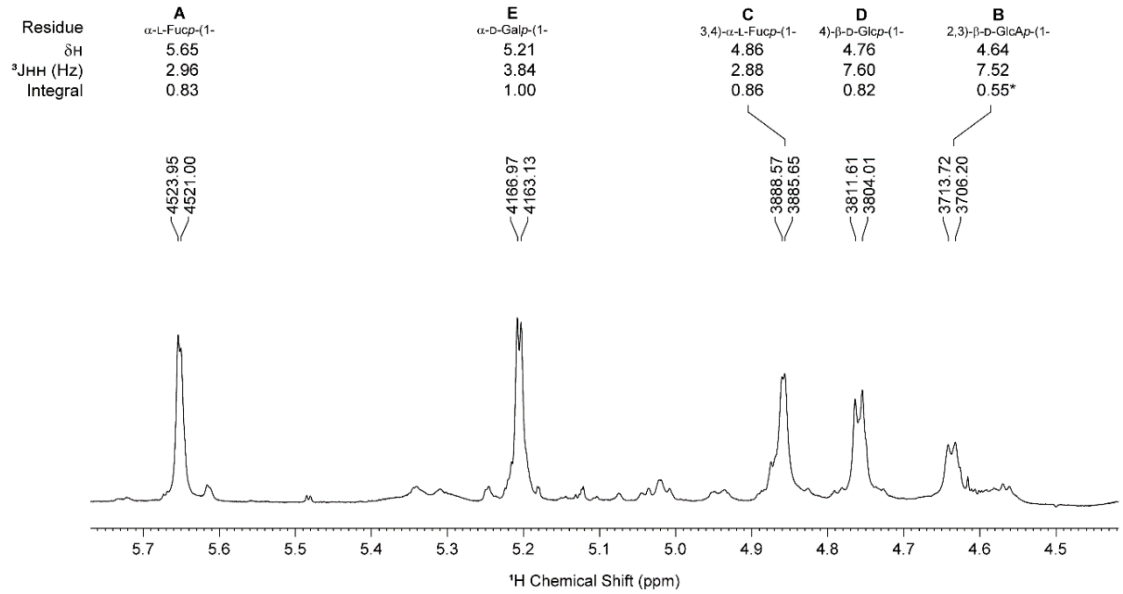


Figure S 11: ^{1}H spectrum of acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift referenced to residual water signal (4.50 ppm). Spectrum shows anomeric signals and are annotated with the residue, proton chemical shift in ppm (δ_H), the calculated coupling constant in Hz ($^{3}J_{HH}$), and integral. *this signal resides close to the residual water signal and is affected by water suppression.

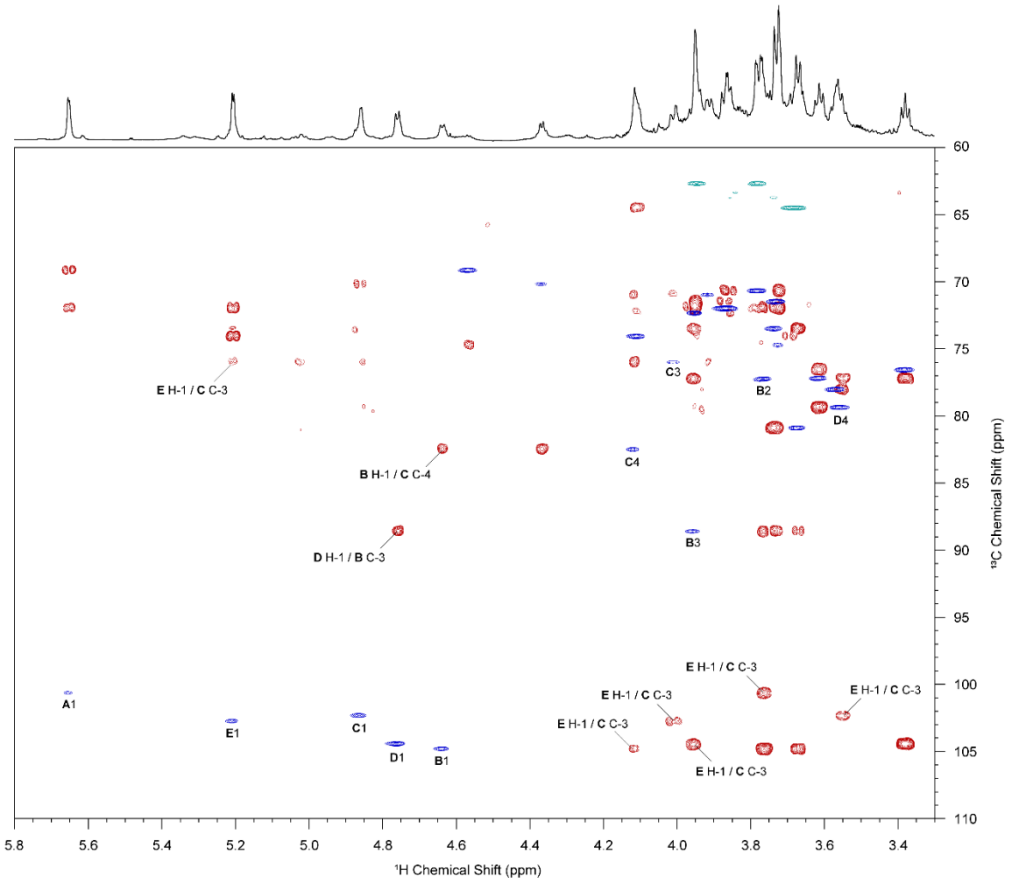


Figure S 12: Overlay of HSQC and HMBC of acid hydrolysed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. HSQC spectrum is coloured in blue and cyan, HMBC spectrum is coloured in red, projection across the top shows the ^{1}H spectrum. Inter-residue correlations that were used to establish partial structure are annotated. Bold letter indicates residue, and number indicate relevant carbon or proton position. Acquired in D_2O (99.9%;D), at 50 °C , 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS. Key: **A** = α-L-Fuc*p*-(1-, **B** = 2,3)-β-D-GlcA*p*-(1-, **C** = 3,4)-α-L-Fuc*p*-(1-, **D** = 4)-β-D-Glc*p*-(1-, **E** = α-D-Gal*p*-(1-.

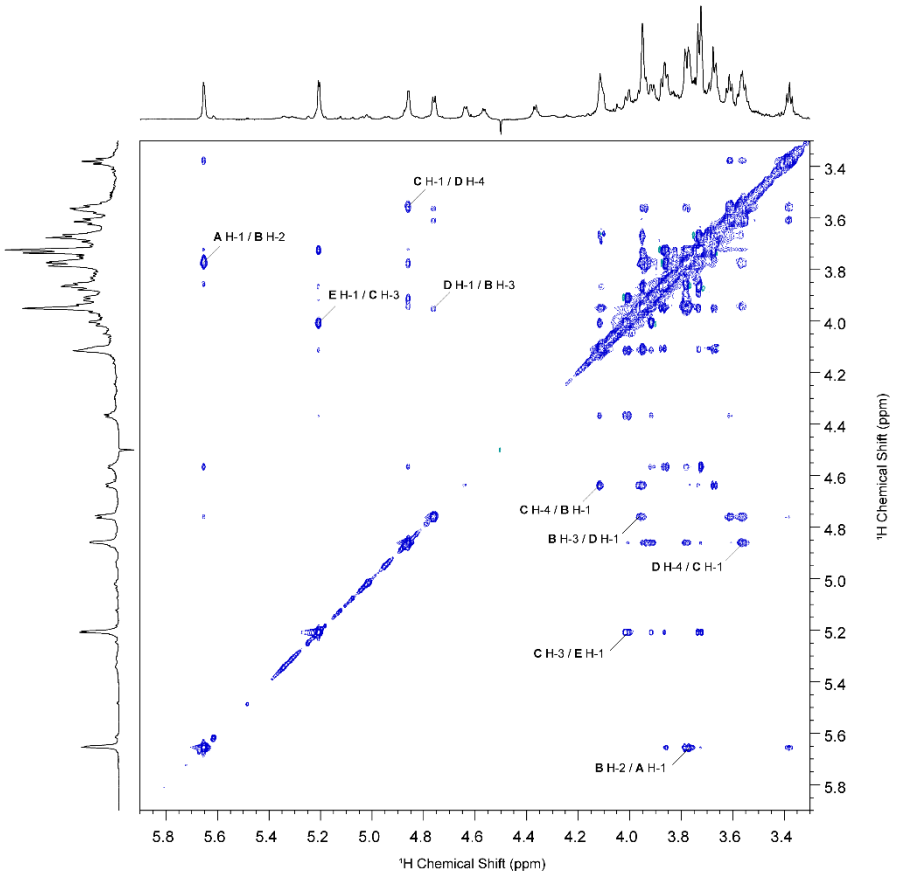


Figure S 13: ^{1}H-^{1}H NOESY spectrum acid hydrolyzed EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Key inter-residue correlations that confirm the position of glycosidic linkages are annotated with residue in bold letter and proton position. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm). Key: **A** = α-L-Fuc*p*-(1-, **B** = 2,3)-β-D-GlcA*p*-(1-, **C** = 3,4)-α-L-Fuc*p*-(1-, **D** = 4)-β-D-Glc*p*-(1-, **E** = α-D-Gal*p*-(1-
.

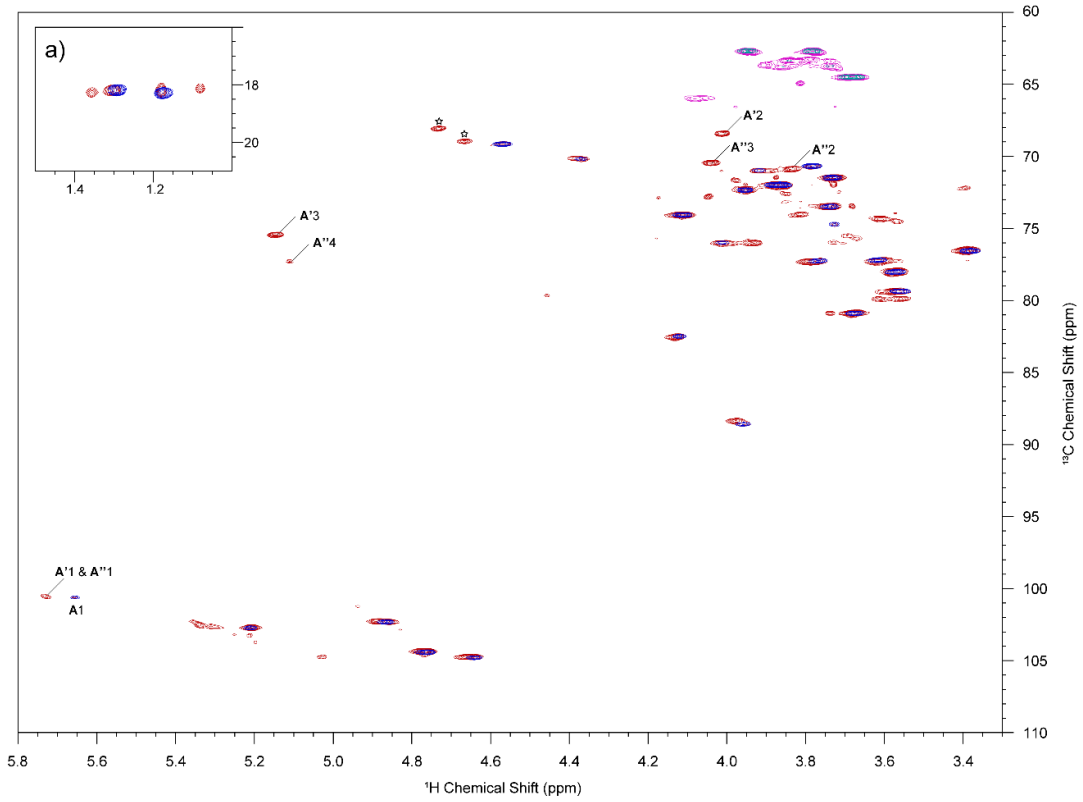


Figure S 14: Overlay of HSQC of acid hydrolyzed (blue and cyan signals) and sheared (red and pink signals) EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Methyl region is show in panel a). Signal are annotated with acetylated residues. Stars show unassigned position-5 fucose signals expected to belong to **A'** and **A''** residues. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS. Key: **A** = α-L-Fuc*p*-(1-, **A'** = α-L-Fuc*p*-3-O-Ac-(1-, **A''** = α-L-Fuc*p*-4-O-Ac-(1-.

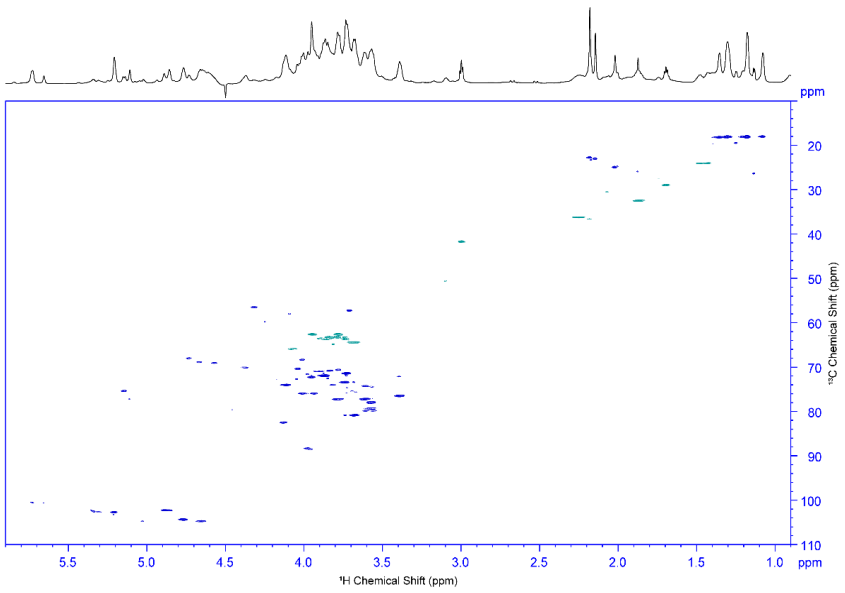


Figure S 15: ^{1}H-^{13}C HSQC spectrum with multiplicity editing of mechanically sheared EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

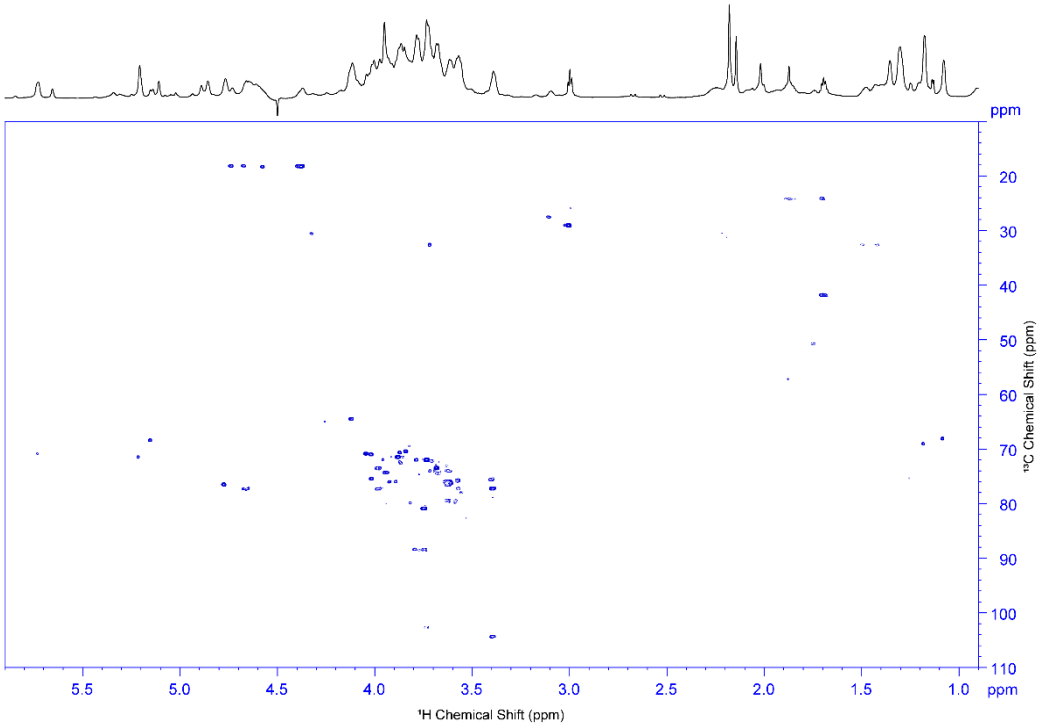


Figure S 16: ^{1}H-^{13}C H2BC spectrum of mechanically sheared EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

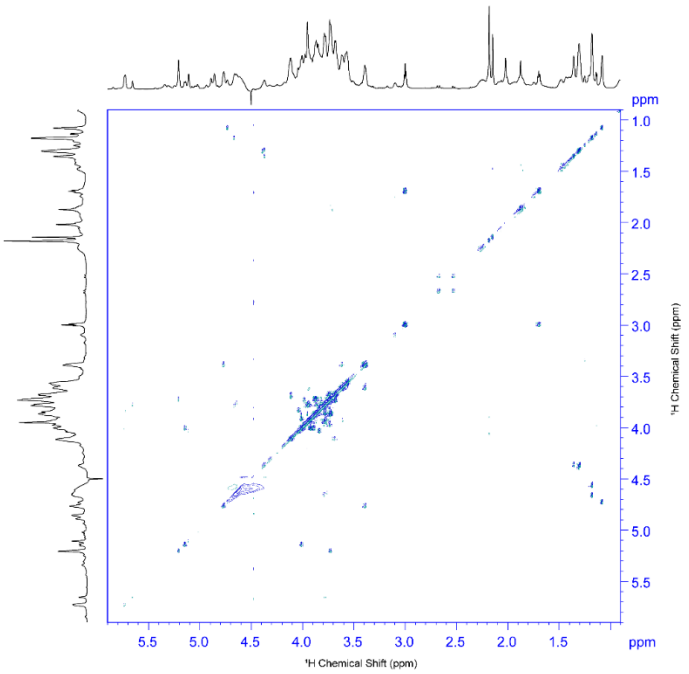


Figure S 17: ^{1}H-^{1}H COSY spectrum mechanically sheared EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800 MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm).

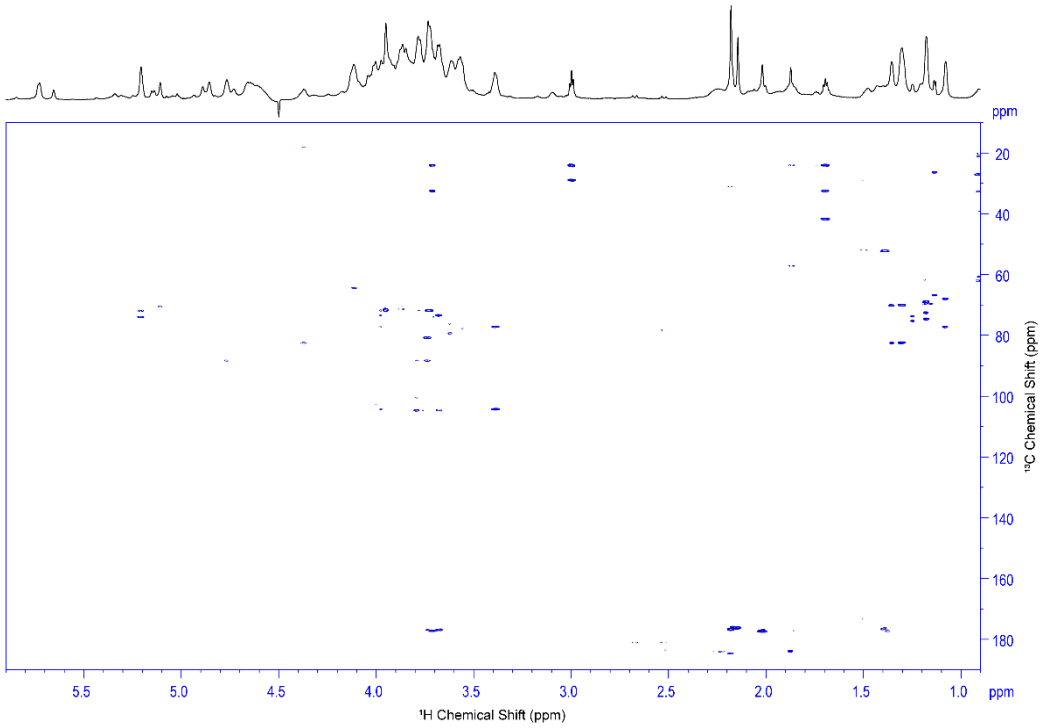


Figure S 18: ^{1}H-^{13}C HMBC spectrum of mechanically sheared EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift internally referenced to the residual water signal (4.50 ppm) and ^{13}C chemical shift referenced indirectly to DSS.

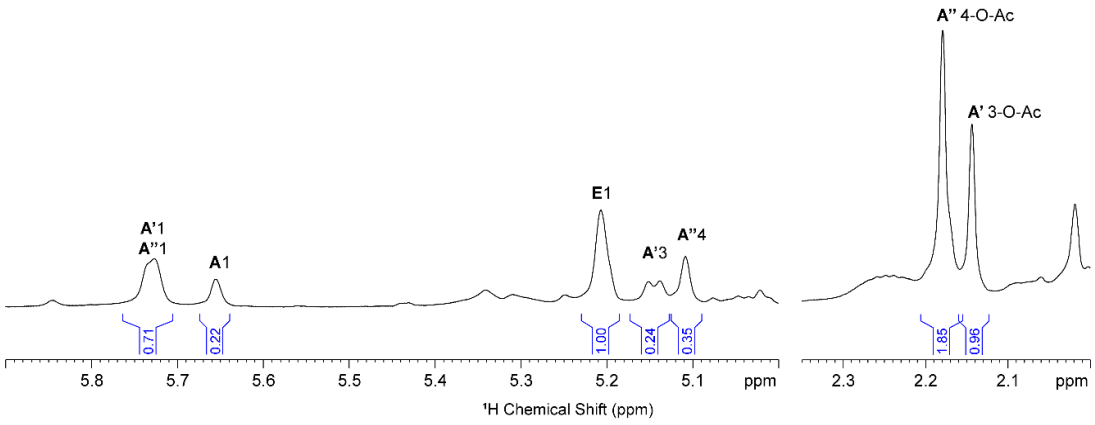


Figure S 19: ^{1}H spectrum of mechanically sheared EPS produced by *P. polymyxa* DSM 365 *ΔpepFJ*. Spectrum shows key signals annotated with residue for anomeric signals (in the region 5.0-6.0 ppm) and methyl signals (~2.1-2.2 ppm) for O-Ac. Integrals that were used to establish degree of acetylation are shown below the spectrum (in blue). Acquired in D_2O (99.9%;D), 50 °C, 800MHz, ^{1}H chemical shift referenced to residual water signal (4.50 ppm).

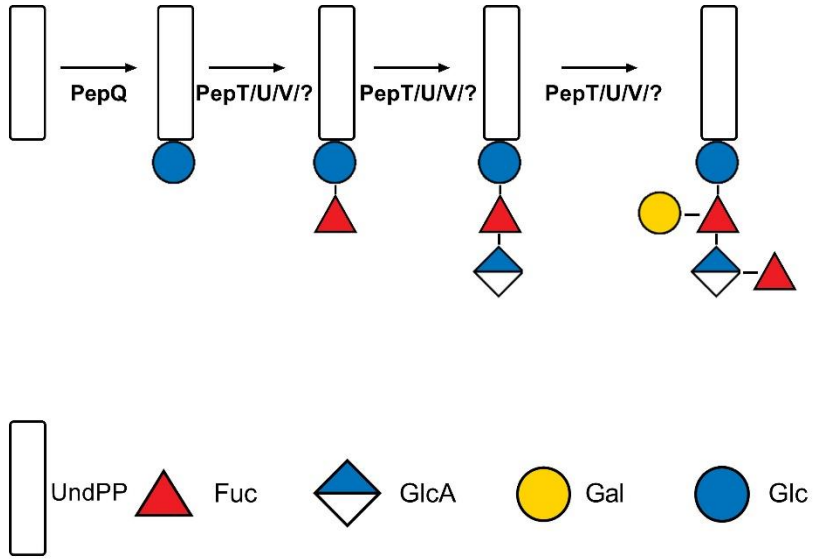


Figure S 20: Schematic overview of the assembly of the repeating unit of paenan II. PepQ transfers a glycosyl residue to a membrane bound undecaprenyl-pyrophosphate lipid anchor (UndPP). The glycosyltransferases PepT, PepU, PepV as well as a yet unknown glycosyltransferase catalyse the stepwise assembly of the repeating unit, followed by the translocation to the periplasm by a Wzx flippase (PepR) and the consequent polymerization by the Wzy protein PepG.

Supplemental Methods

NMR acquisition parameters and processing

Acquisition parameters for each experiment were as follows. 1D proton spectrum, pulse sequence: noesygppr1d, spectral width: 10 ppm or 12 ppm, number of scans: 16. HSQC spectrum, pulse sequence: hsqcedetgpsisp2.3, spectral width: 120 ppm (f1) 10 ppm (f2), time do-main: 512 (f1) 1k (f2), number of scans: 16, dummy scans: 16. H2BC spectrum, pulse sequence: h2bcetgpl3pr, spectral width: 180 ppm (f1) 10 ppm (f2), time domain: 1k (f1) 2k (f2), number of scans: 16, dummy scans: 16. HMBC spectrum, pulse sequence: hmbcetgpl3nd, spectral width: 180 ppm (f1) 10 ppm (f2), time domain: 400 (f1) 2k (f2), number of scans: 32, dummy scans: 16. DQF-COSY spectrum, pulse sequence: cosydfph, spectral width: 10 ppm (f1) 10 ppm (f2), time domain: 1k (f1) 4k (f2), number of scans: 8 or 16, dummy scans: 16. NOESY spectrum, pulse sequence: noesyesgpph, spectral width: 10 ppm (f1) 10 ppm (f2), time domain: 1k (f1) 4k (f2), number of scans: 16, dummy scans: 64, mixing time (d8): 80 ms. Spectra were processed with exponential window multiplication, a Fourier transform, phase correction, and baseline correction.

Supplemental References

Rütering, M., Cress, B.F., Schilling, M., Rühmann, B., Koffas, M.A.G., Sieber, V., Schmid, J., 2017. Tailor-made exopolysaccharides—CRISPR-Cas9 mediated

genome editing in *Paenibacillus polymyxa*. Synth. Biol. 2. https://doi.org/10.1093/synbio/ysx007

Sassaki, G.L., Gorin, P.A.J., Souza, L.M., Czelusniak, P.A., Iacomini, M., 2005. Rapid synthesis of partially O-methylated alditol acetate standards for GC–MS: some relative activities of hydroxyl groups of methyl glycopyranosides on Purdie methylation. Carbohydr. Res. 340, 731–739. https://doi.org/10.1016/j.carres.2005.01.020

7.1.3. CRISPR-Cas9 driven structural elucidation of the heteroexopolysaccharides of *Paenibacillus polymyxa* DSM 365

Table S 1: Strains used in this work

Bacterial Strains	Genotype	Reference
E. coli S17-1	Conjugation strain; recA pro hsdR RP42 Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC 47055
P. polymyxa DSM 365	wild type	DSMZ
P. polymyxa DSM 365 *ΔpepC*	DSM 365 *ΔpepC*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepD*	DSM 365 *ΔpepD*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepF*	DSM 365 *ΔpepF*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepI*	DSM 365 *ΔpepI*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepJ*	DSM 365 *ΔpepJ*	(Rütering et al., 2017)
P. polymyxa DSM 365 *ΔpepK*	DSM 365 *ΔpepK*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepL*	DSM 365 *ΔpepL*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepQ*	DSM 365 *ΔpepQ*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepT*	DSM 365 *ΔpepT*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepU*	DSM 365 *ΔpepU*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepV*	DSM 365 *ΔpepV*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepCQ*	DSM 365 *ΔpepC ΔpepQ*	This study
P. polymyxa DSM 365 *ΔpepFJ*	DSM 365 *ΔpepF ΔpepJ*	(Schilling et al., 2020)
P. polymyxa DSM 365 *ΔpepQF*	DSM 365 *ΔpepQ ΔpepF*	This study
P. polymyxa DSM 365 *ΔpepTUV*	DSM 365 *ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepITUV*	DSM 365 *ΔpepI ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepKTUV*	DSM 365 *ΔpepK ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔpepLTUV*	DSM 365 *ΔpepL ΔpepT ΔpepU ΔpepV*	This study
P. polymyxa DSM 365 *ΔepsO*	DSM 365 *ΔepsO*	(Schilling et al., 2020)

Table S 2: Plasmids used and constructed in this work. Each plasmid was used for the transformation of *E. coli* S17-1 and furthermore for the conjugational transformation of *P. polymyxa* DSM 365. Before EPS production, each knock-out strain was cured from the plasmids resulting in the strains listed in Table S 1.

Plasmids	Description	Reference
pCasPP	*P. polymyxa* CRISPR-Cas9 genome editing plasmid; PsgsE-*cas9*; Pgapdh-off-target gRNA; *neo*; *oriT*; *repU*	(Rütering et al., 2017)
pCasPPH_pepC	*pepC* targeting knock-out plasmid containing repair template	(Schilling et al., 2020)
pCasPPH_pepQ	*pepQ* targeting knock-out plasmid containing repair template	(Schilling et al., 2020)
pCasPPH_pepTUV	Knock-out plasmid containing repair template for the simultaneous triple knock-out of *pepT*, *pepU* and *pepV*	This study

Table S 3: Oligonucleotides used in this study. Overhangs used for Golden Gate Assembly are depicted in lower case letters. Restriction sites used for cloning are underlined. For each knock-out construct, two sets of sgRNAs were designed, tested and listed below if successful knock-outs were obtained

Primer	Sequence 5'→3'
pepC_KO_proof_fw	GCAACTCATGGAGCAGGCCAGTGCGACG
pepC_KO_proof_rev	GCTGTATGGTGTTTATTCTAGTAGTCCAAGG
pepTUV_sgRNA1_fw	acgcAGAACGTGAGGAAAACCGTG
pepTUV_sgRNA1_rev	aaacCACGGTTTTCCTCACGTTCT
pepTUV_sgRNA2_fw	acgcTATGTCGTGGGATCGCATTG
pepTUV_sgRNA_2_rev	aaacCAATGCGATCCCACGACATA
pepTUV_Harms_US_fw	AATTTCTAGACCCGGATCATCGCTTTACAGT
pepTUV_Harms_US_rev_OE	TTTACCGACTGGGGTTTAGAGACCGCTGCG
pepTUV_Harms_DS_fw_OE	AGTCTAAACCCCGGTAAAAGCAGGATGCA
pepTUV_Harms_DS_rev	AATTTCTAGAGCCAGTCCCCATATACCAATCG
pepTUV_KO_proof_fw	CGCAGCTCAAACGTGGGCTA
pepTUV_KO_proof_rev	GGGGCATCCCTTACGTCATCT
pepQ_KO_proof_fw	GGAATCGGCTTGGCGTTTTTGCTGG
pepQ_KO_proof_rev	CCGTATACCTCATAATCCTTCGACAGC

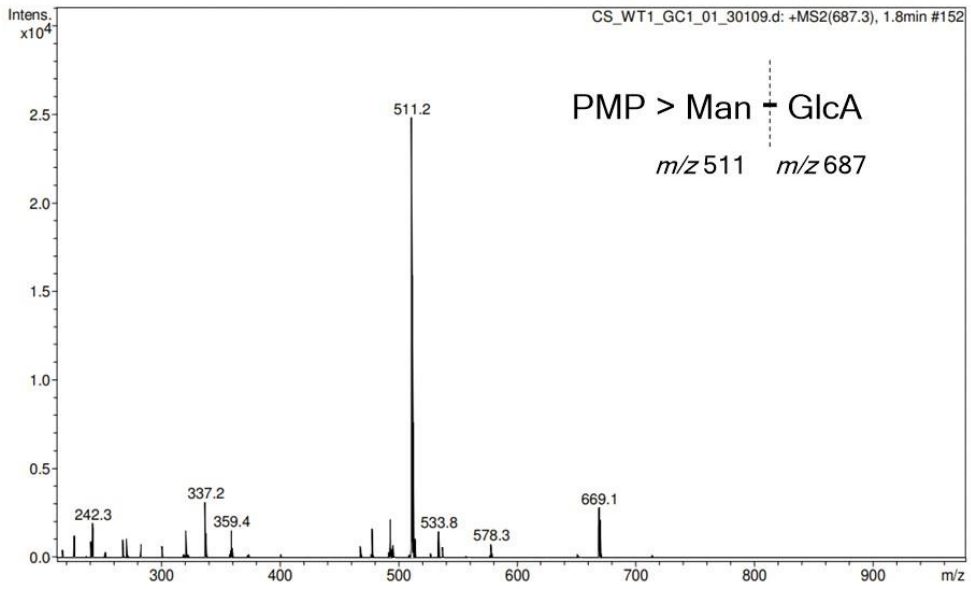


Figure S 1: MS/MS fragmentation of a GlcA-Hex dimer (*m/z* 687). Early retention time at 1.8 min and fragmentation indicate mannose as the residue positioned at the reducing end of the dimer.

Table S 4: Overview of different oligosaccharide fragments obtained by partial hydrolysis of EPS from *P. polymyxa* DSM 365 and mutant variants. MS/MS spectra were evaluated to determine the sequence of each oligosaccharide. Retention times of each fragment are given in brackets. Fragments associated with the pyruvate containing polysaccharide paenan I are colored blue. Fragments corresponding to the fucose containing EPS paenan II are colored red and fragments specifically attributed to paenan III yellow. x: fragment detected; *: traces detectable.

Strain	PMP>Hex (1.9) *m/z* 511	PMP>Hex-Hex (2.9) *m/z* 673	PMP>Hex-Hex (3.1) *m/z* 673	PMP>Hex-Hex-Hex (3.4) *m/z* 835	PMP>Hex-Hex-Hex (3.7) *m/z* 835	PMP>Hex-Hex-Hex (4.2) *m/z* 835	PMP>Hex-Hex-Hex (4.9) *m/z* 835	PMP>Hex-Hex-Hex-Hex (6.5) *m/z* 997	PMP>Hex-Hex-Hex-Hex (7.0) *m/z* 997	PMP>Hex-Hex-Hex-Hex (7.9) *m/z* 997	PMP>Hex-Hex-Hex-Hex-Hex (7.0) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex (7.8) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex (8.1) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex-Hex (8.5) *m/z* 1321	PMP>Hex^{Pyr} (2.3) *m/z* 581	PMP>Hex-Hex^{Pyr} (3.3) *m/z* 743	PMP>Hex-Hex-Hex^{Pyr} (4.1) *m/z* 905	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex (6.2) *m/z* 1229	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex (7.2) *m/z* 1067	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex-Hex (7.8) *m/z* 1229	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex-Hex (8.1) *m/z* 1391	PMP>Hex-Hex-Hex-Hex-Hex^{Pyr}-Hex (8.2) *m/z* 1391	PMP>Fuc-GlcA (2.9) *m/z* 671	PMP>Fuc-Hex (2.8) *m/z* 657	PMP>Hex-Fuc-Hex (3.4) *m/z* 819	PMP>Fuc-GlcA-Hex (3.4) *m/z* 833	PMP>Fuc-GlcA-Hex-Hex (5.2) *m/z* 995	PMP>Fuc-Hex-GlcA-Hex (5.3) *m/z* 833	PMP>Hex-Fuc-GlcA-Hex (6.1) *m/z* 833	PMP>Fuc-GlcA-Hex-Fuc-GlcA-Hex (7.8) *m/z* 1317	PMP>Hex-GlcA-Hex *m/z* 511	PMP>Hex-GlcA-Hex-Hex (6.5) *m/z* 511	PMP>Hex-Hex-GlcA-Hex (7.0) *m/z* 511
WT	x	x	x	x		x	x	x		x					x	x	x	x	x	x					x						x	x	x
ΔepsO	x	x									x	x											x		*	x	x	x		x			
ΔpepD	x		x																				x			x	x						
ΔpepE	x	x	x																				x			x	x			x			
ΔpepF	x			x																			x	x		x	x	x		x			
ΔpepFJ	x																						x			x	x			x			
ΔpepI	x			x											x			x					x			x	x		*				
ΔpepITUV	x	x	x	x		x			x						x	x	x	x	x	x		x											
ΔpepJ	x	x													x					x			x			x	x			x			
ΔpepK	x			x							x				x	x		x		x			x			x	x						
ΔpepKTUV	x	x	x	x		x	x	x	x				x	x	x	x		x	x	x		x											
ΔpepL	x	x	x	x		x		x	x			x	x		x	x	x	x	x	x	x	x	x		*		x						
ΔpepLTUV	x	x	x	x		x	x	x	x				x		x	x	x	x	x	x														

Strain	PMP>Hex (1.9) *m/z* 511	PMP>Hex-Hex (2.9) *m/z* 673	PMP>Hex-Hex (3.1) *m/z* 673	PMP>Hex-Hex-Hex (3.4) *m/z* 835	PMP>Hex-Hex-Hex (3.7) *m/z* 835	PMP>Hex-Hex-Hex (4.2) *m/z* 835	PMP>Hex-Hex-Hex (4.9) *m/z* 835	PMP>Hex-Hex-Hex-Hex (6.5) *m/z* 997	PMP>Hex-Hex-Hex-Hex (7.0) *m/z* 997	PMP>Hex-Hex-Hex-Hex (7.9) *m/z* 997	PMP>Hex-Hex-Hex-Hex-Hex (7.0) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex (7.8) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex (8.1) *m/z* 1159	PMP>Hex-Hex-Hex-Hex-Hex-Hex (8.5) *m/z* 1321	PMP>Hex^{Pyr} (2.3) *m/z* 581	PMP>Hex-Hex^{Pyr} (3.3) *m/z* 743	PMP>Hex-Hex-Hex^{Pyr} (4.1) *m/z* 905	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex (6.2) *m/z* 1229	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex (7.2) *m/z* 1067	PMP>Hex-Hex-Hex-Hex^{Pyr}-Hex-Hex (7.8) *m/z* 1229	PMP>Hex-Hex-Hex-Hex-Hex^{Pyr}-Hex-Hex (8.1) *m/z*	PMP>Hex-Hex-Hex-Hex-Hex^{Pyr}-Hex (8.2) *m/z*	PMP>Fuc-GlcA (2.9) *m/z* 671	PMP>Fuc-Hex (2.8) *m/z* 657	PMP>Hex-Fuc-Hex (3.4) *m/z* 819	PMP>Fuc-GlcA-Hex (3.4) *m/z* 833	PMP>Fuc-GlcA-Hex-Hex (5.2) *m/z* 995	PMP>Fuc-Hex-GlcA-Hex (5.3) *m/z* 833	PMP>Hex-Fuc-GlcA-Hex (6.1) *m/z* 833	PMP>Fuc-GlcA-Hex-Fuc-GlcA-Hex (7.8) *m/z*	PMP>Hex-GlcA-Hex *m/z* 511	PMP>Hex-GlcA-Hex-Hex (6.5) *m/z* 511	PMP>Hex-Hex-GlcA-Hex (7.0) *m/z* 511
ΔpepQ	x	x	x	x		x		x	x			x	x		x	x	x	x	x	x	x	x	x								*	x	
ΔpepQF	x																																
ΔpepT	x	x		x					x						x	x		x		x													
ΔpepU	x	x		x		x			x	x					x	x	x	x	x	x		x											
ΔpepV	x	x	x	x							x				x	x		x	x	x		x											

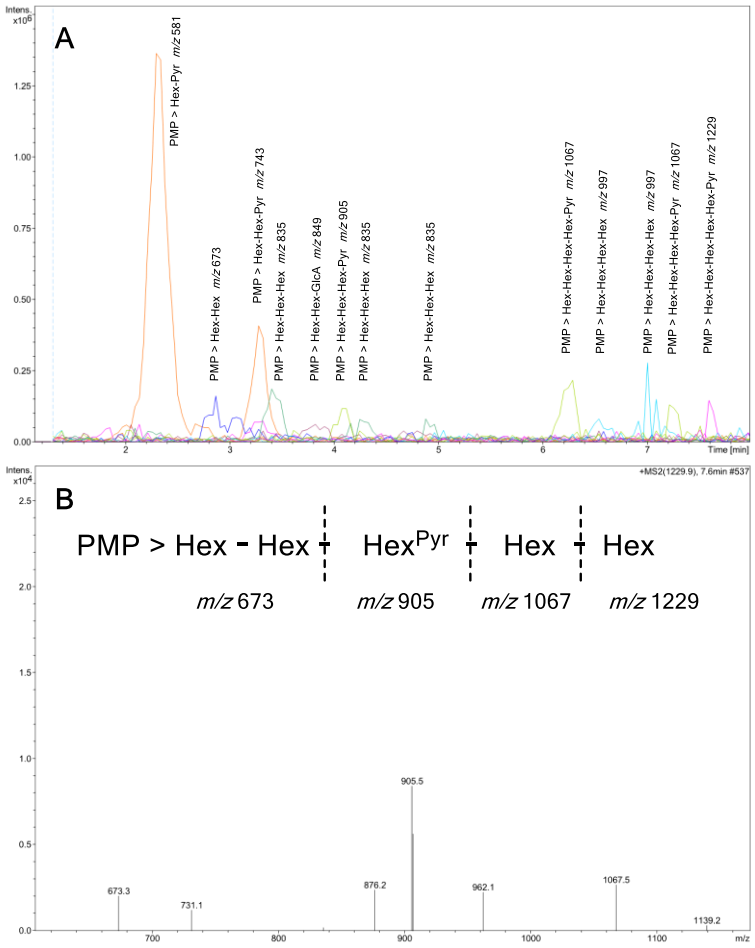


Figure S 2: A: Exemplary chromatogram of oligosaccharides from EPS obtained from the wildtype strain. The polymer was subjected to mild hydrolysis conditions (2 M TFA, 90 °C, 105 min), derivatized with PMP and analyzed via HPLC-UV-MS. B: For each obtained fragment (exemplarily shown for the largest fragment containing pyruvate with *m/z* 1,229), the sequence of the oligosaccharide was obtained by evaluation of the MS spectra.

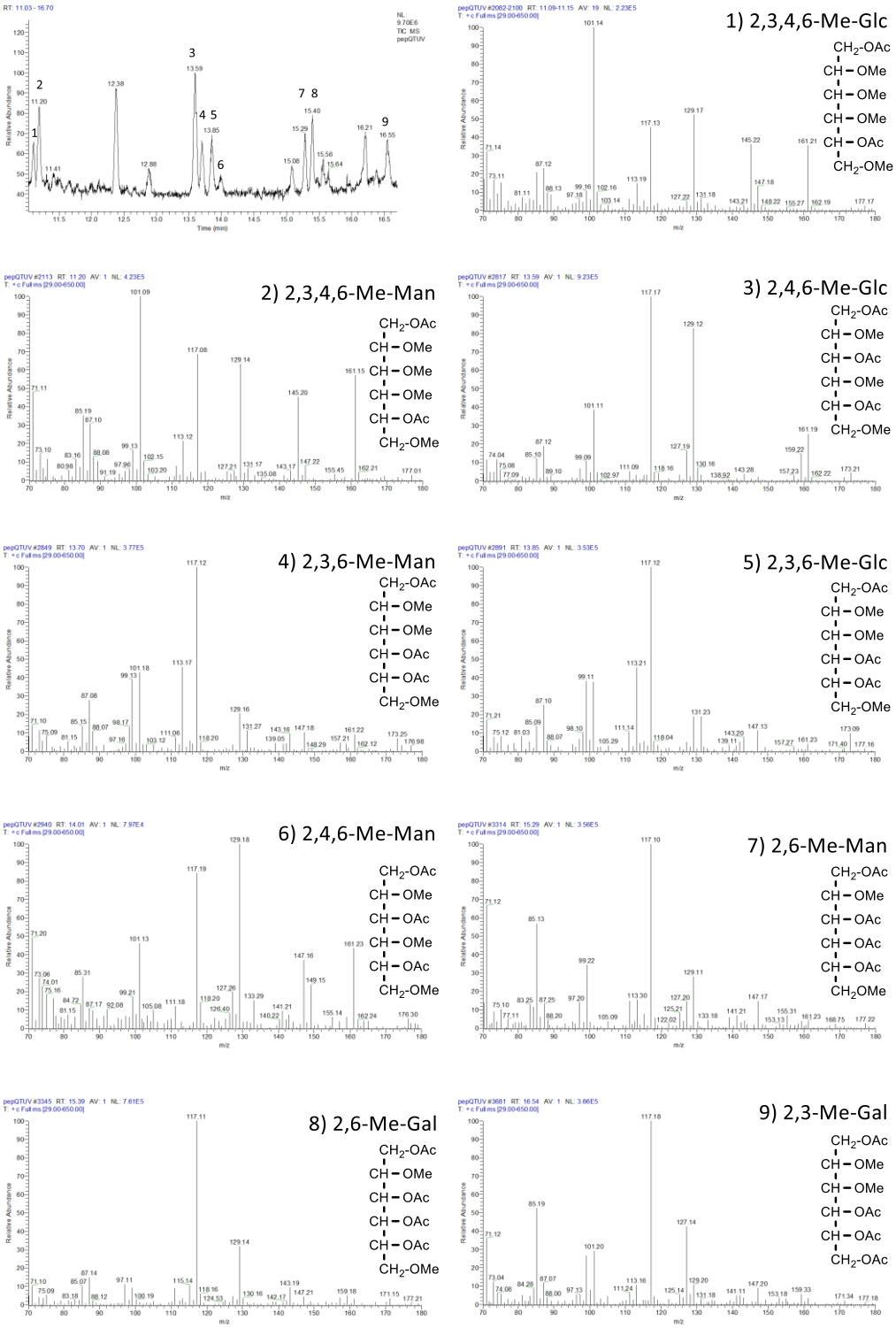


Figure S 3: Exemplary chromatogram of methylation analysis of the EPS obtained from *P. polymyxa* DSM 365 containing paenan I and III. Unlabeled peaks showed MS-spectra not corresponding to partially methylated alditol acetates (PMAAs). PMAAs were identified via retention time standards and MS spectra as described by Sassaki et al. (2005) (Sassaki et al., 2005).

Supplemental References:

Rütering, M., Cress, B.F., Schilling, M., Rühmann, B., Koffas, M.A.G., Sieber, V., Schmid, J., 2017. Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. Synth. Biol. 2. https://doi.org/10.1093/synbio/ysx007

Sassaki, G.L., Gorin, P.A.J., Souza, L.M., Czelusniak, P.A., Iacomini, M., 2005. Rapid synthesis of partially O-methylated alditol acetate standards for GC–MS: some relative activities of hydroxyl groups of methyl glycopyranosides on Purdie methylation. Carbohydr. Res. 340, 731–739. https://doi.org/10.1016/j.carres.2005.01.020

Schilling, C., Rühmann, B., Klau, L., Aachmann, F., Schmid, J., Sieber, V., 2022. Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365. Carbohydrate Polymers.278.

7.1.4. Rheological characterization of artificial paenan compositions produced by *Paenibacillus polymyxa* DSM 365

Table S 1: Bacterial strains used in this study

Bacterial Strains	Produced paenan variant	Reference
P. polymyxa DSM 365	I & II & III	DSMZ
P. polymyxa DSM 365 *ΔepsO*	I (depyruvated) & II &III	Schilling et al. (2022a)
P. polymyxa DSM 365 *ΔpepF*	II &III	Schilling et al. (2022a)
P. polymyxa DSM 365 *ΔpepFJ*	II	Schilling et al. (2022a)
P. polymyxa DSM 365 *ΔpepQF*	III	Schilling et al. (2022a)
P. polymyxa DSM 365 *ΔpepUL*	I	Schilling et al. (2022b)
P. polymyxa DSM 365 *ΔpepLJ*	I & II	Schilling et al. (2022b)
P. polymyxa DSM 365 *ΔpepQTUV*	I & III	Schilling et al. (2022b)

Table S 2: Concentration and conductivity of different EPS variants used for rheological measurements

Paenan variant		Concentration [% m]	Conductivity mS cm^{-1}
I & II & III	aq.	1.02	0.73
	0.5 % NaCl	1.02	9.07
I & II & III (depyruvated)	aq.	1.01	1.01
	0.5 % NaCl	1.01	9.98
I	aq.	1.00	1.12
	0.5 % NaCl	1.00	11.76
II	aq.	1.00	0.87
	0.5 % NaCl	1.00	9.55
III	aq.	1.00	0.92
	0.5 % NaCl	1.00	10.02
I & II	aq.	1.01	0.78
	0.5 % NaCl	1.01	9.55
I & III	aq.	1.01	1.06
	0.5 % NaCl	1.01	9.02
II & III	aq.	1.00	0.77
	0.5 % NaCl	1.00	9.30

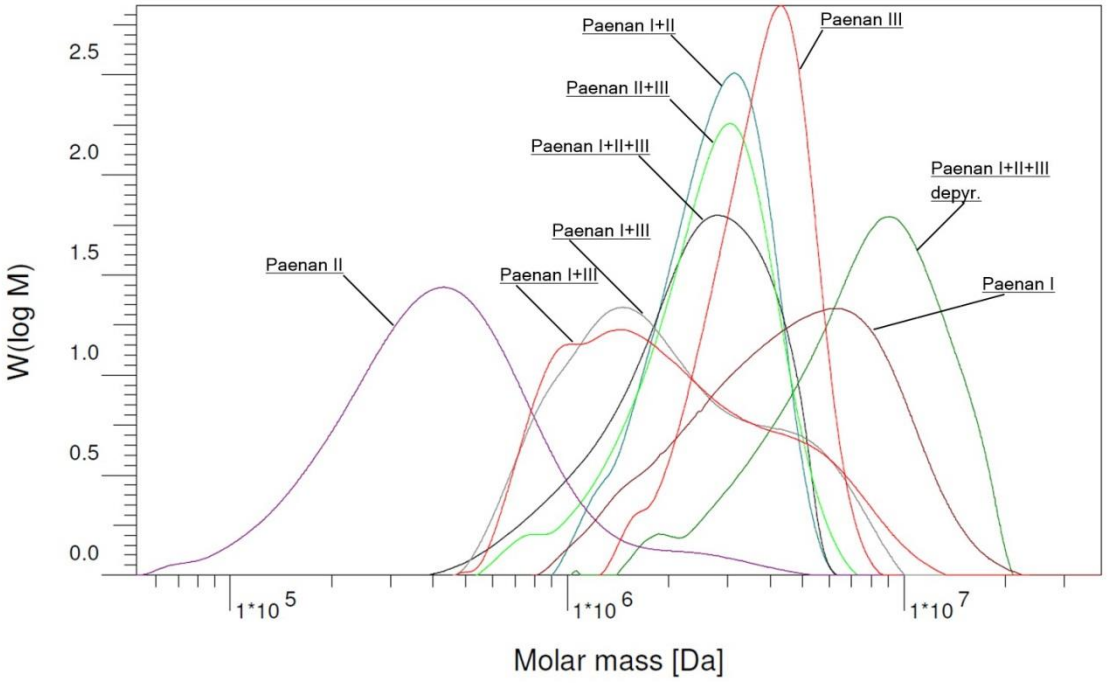


Figure S 1: Molar weight distribution of different EPS mixtures obtained from knock-out variants of *P. polymyxa*.

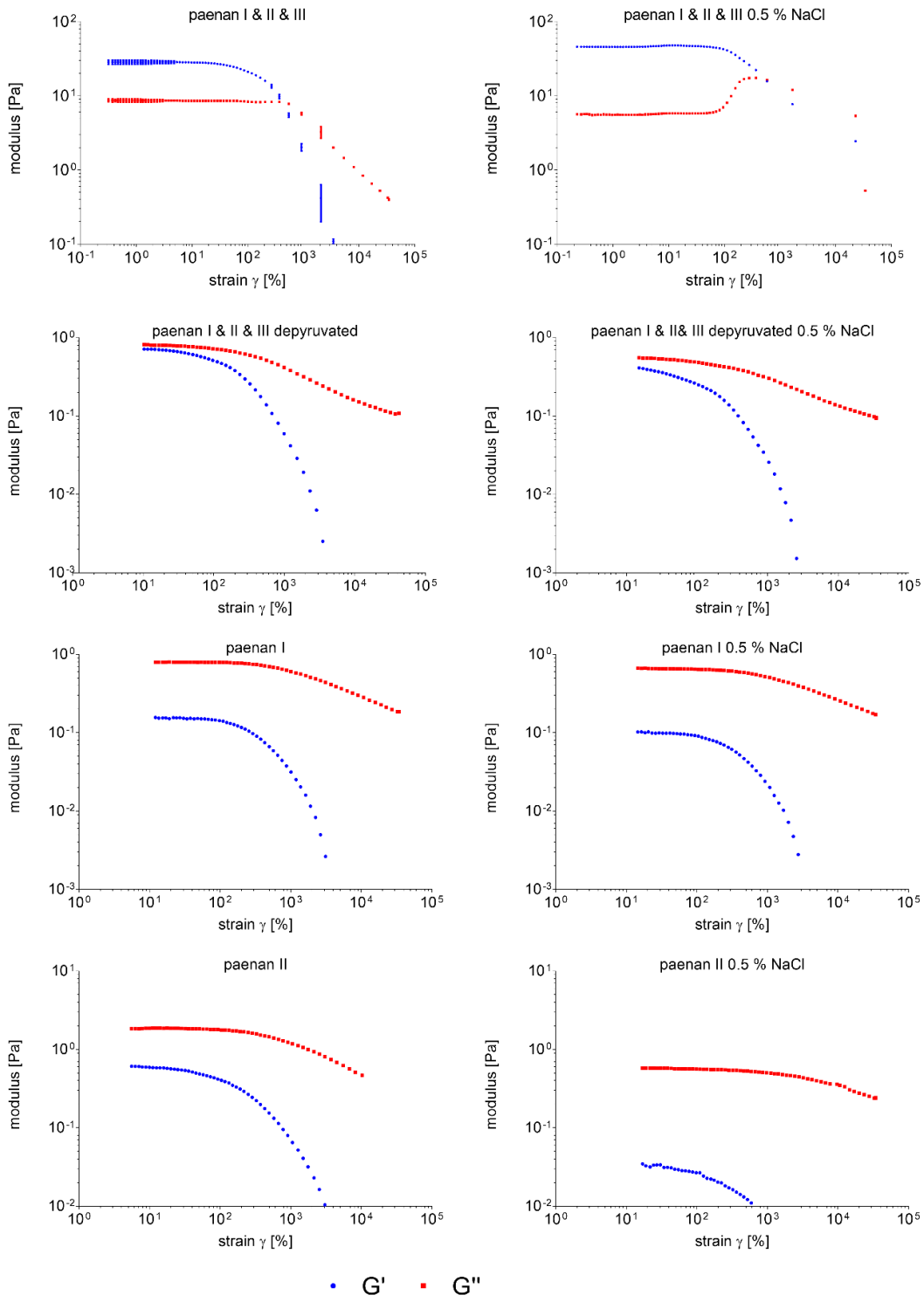


Figure S 2: Amplitude sweeps of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. G' ●: storage modulus; G'' ■: loss modulus

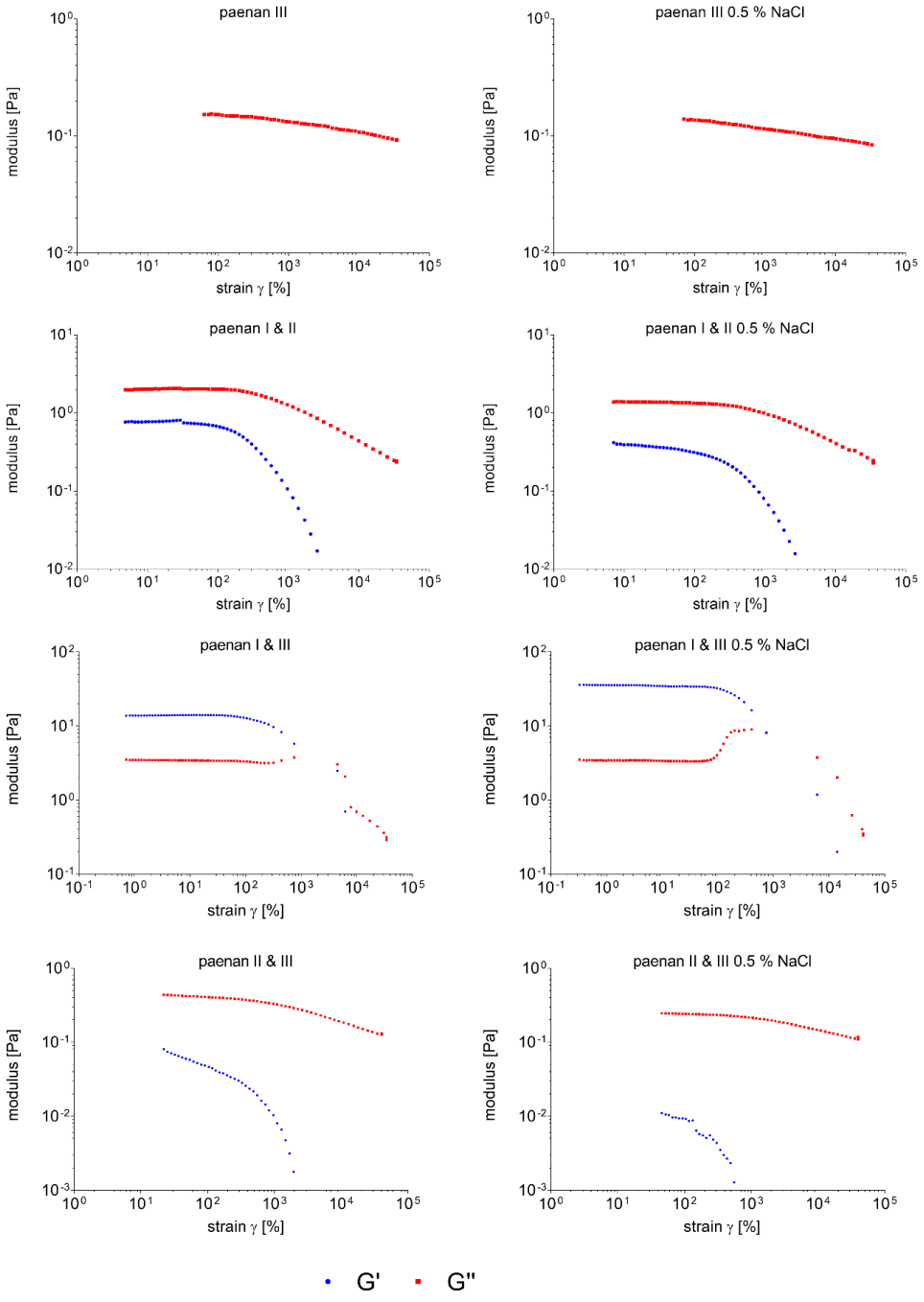


Figure S 3: Amplitude sweeps of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. G' ●: storage modulus; G'' ■: loss modulus

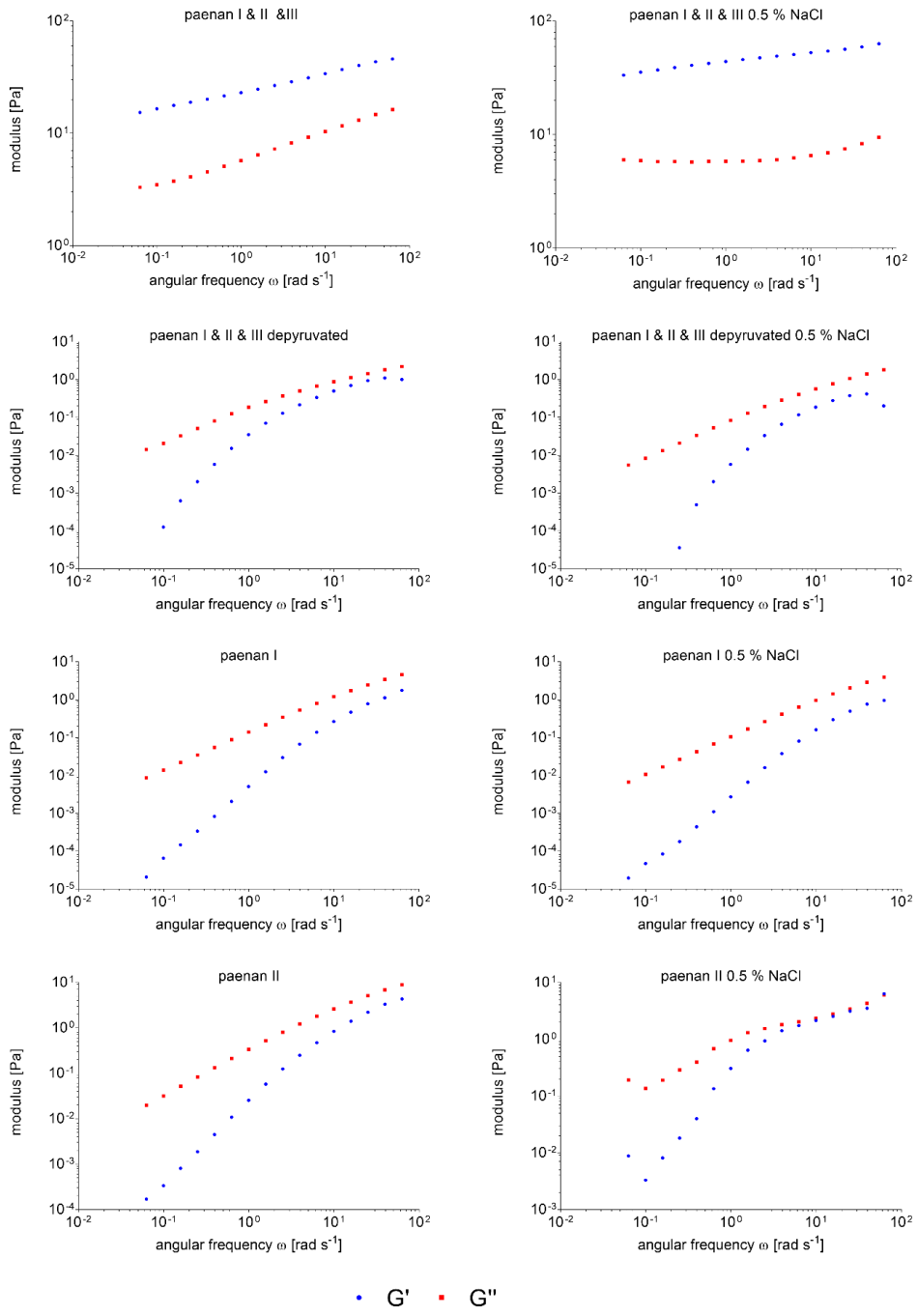


Figure S 4: Frequency sweeps of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. Due to a limited LVE region of paenan III in the applied conditions, this variant was not measured in frequency sweeps, G' ●: storage modulus; G'' ■: loss modulus

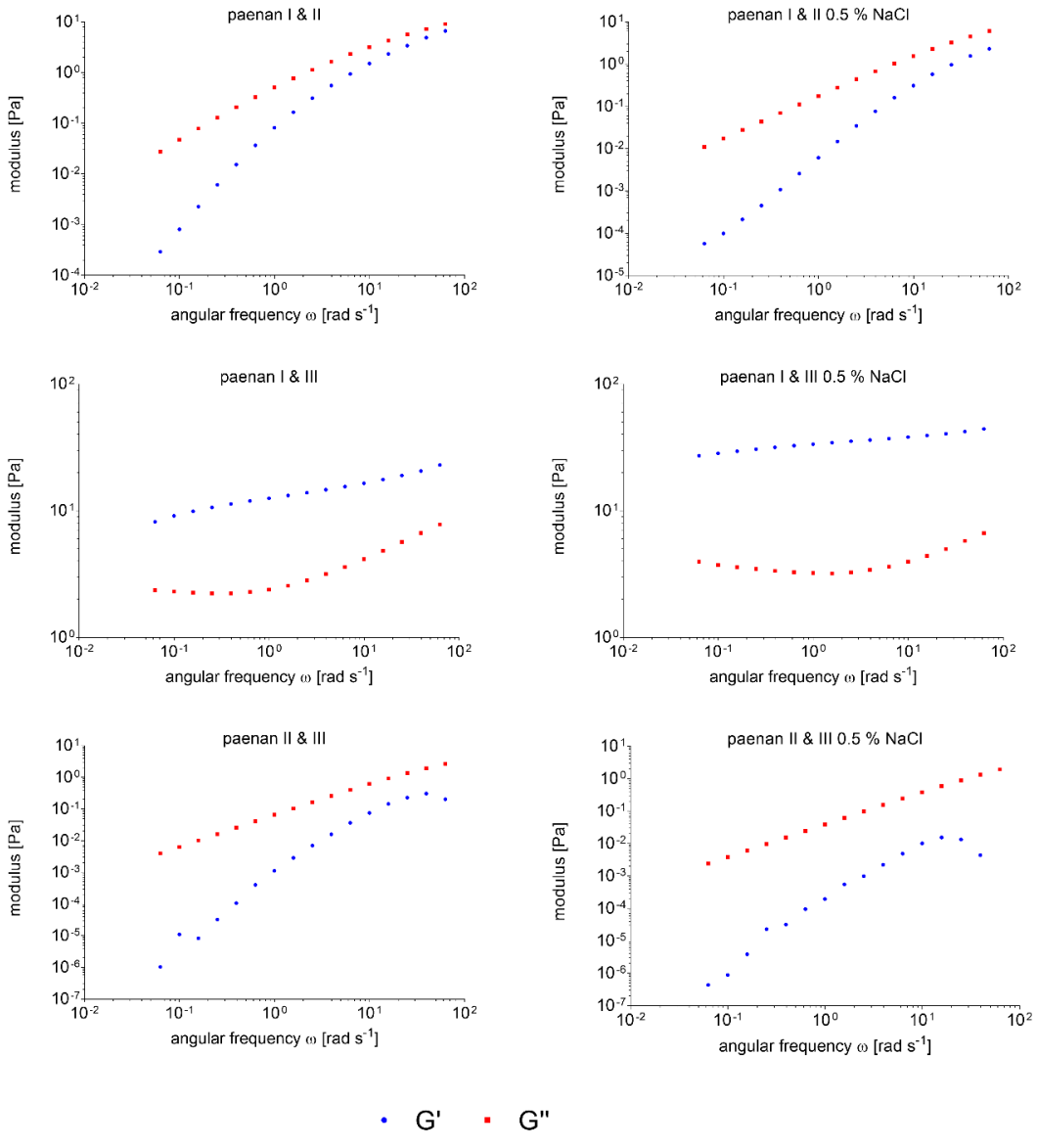


Figure S 5: Frequency sweeps of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. Due to a limited LVE region of paenan III in the applied conditions, this variant was not measured in frequency sweeps. G' ●: storage modulus; G'' ■: loss modulus

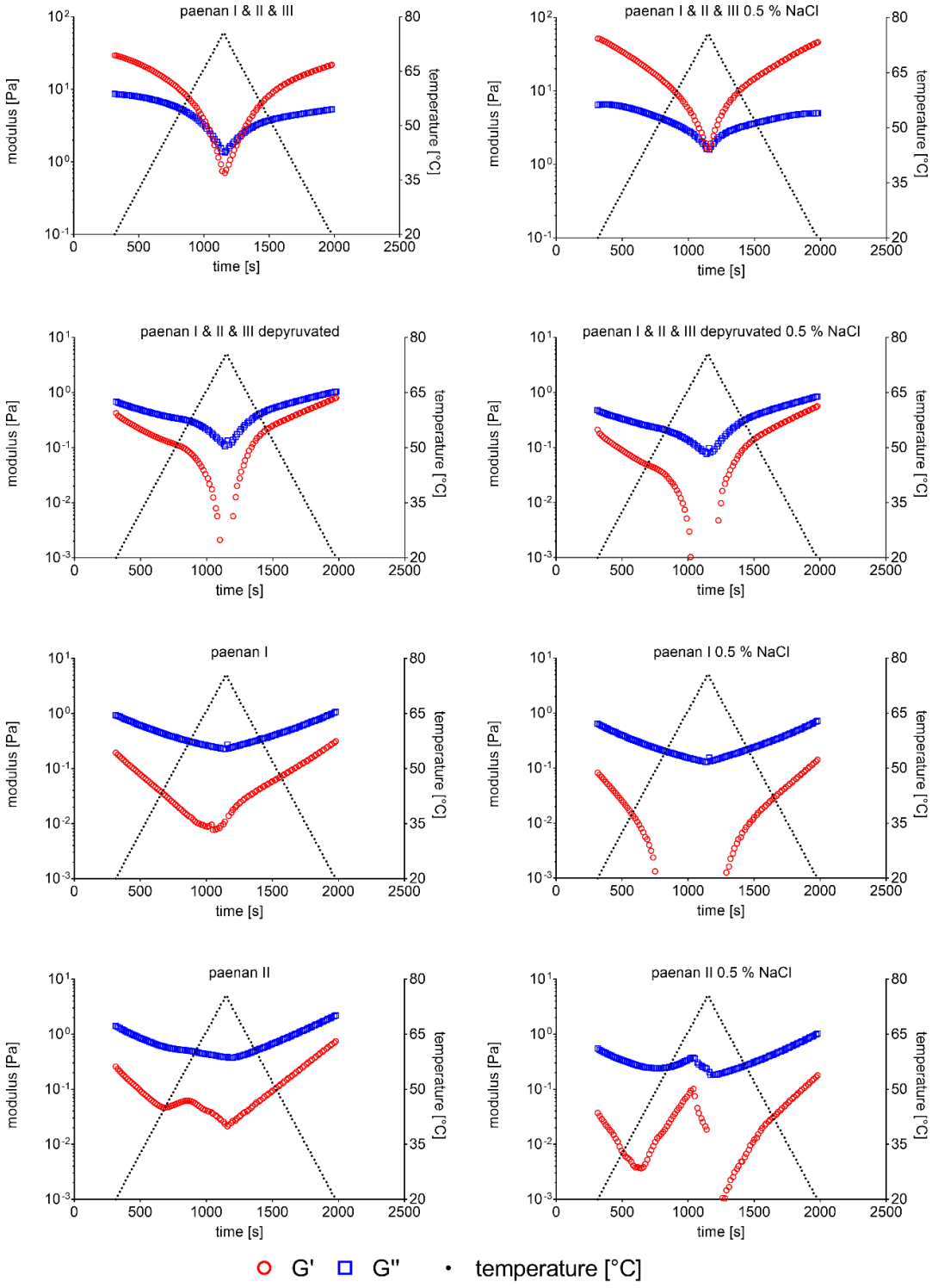


Figure S 6: Temperature sweeps from 20-75 °C of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. G' ●: storage modulus; G'' ■: loss modulus

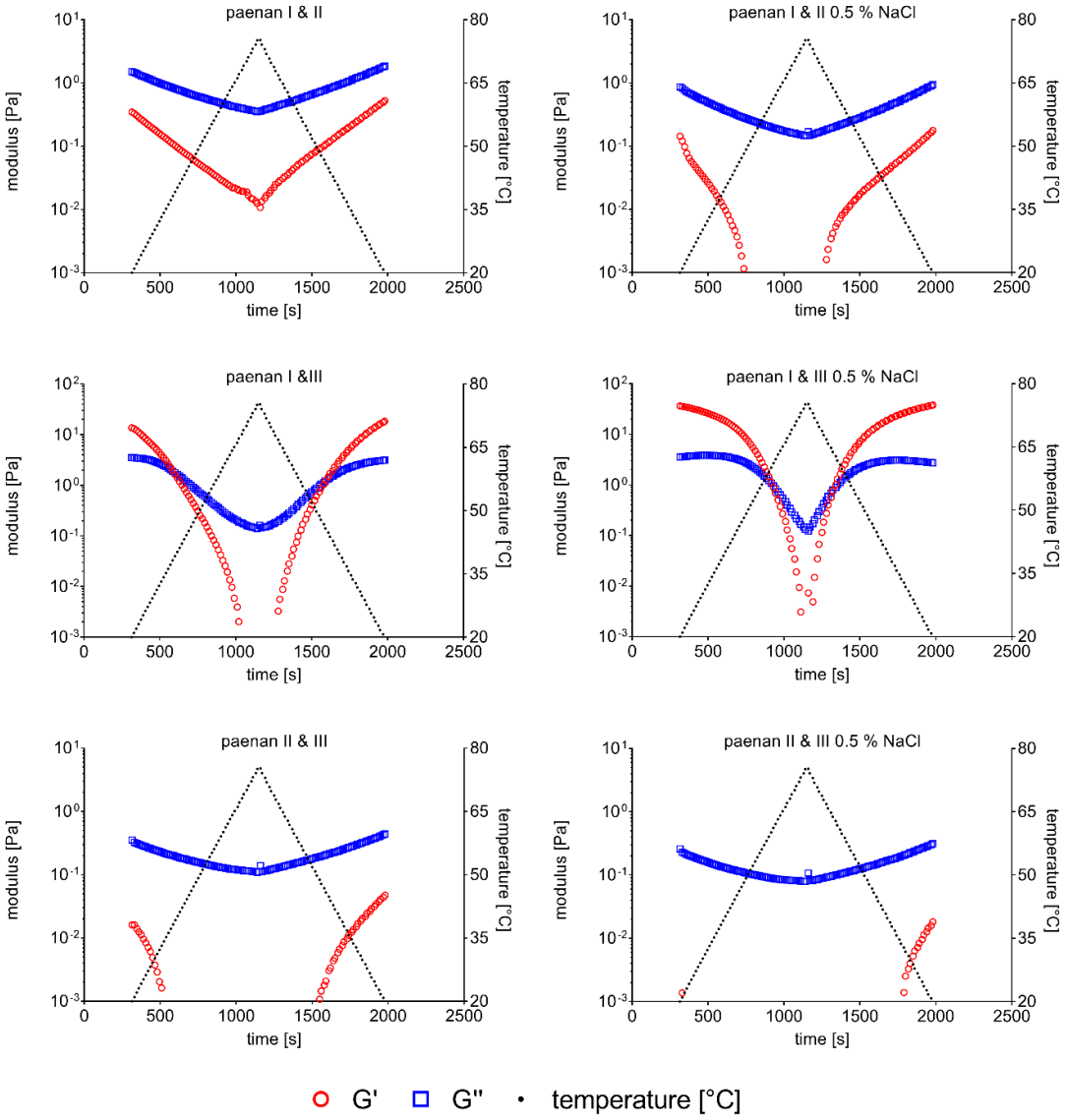


Figure S 7: Temperature sweeps from 20-75 °C of 1 % solutions of indicated EPS variants produced by *P. polymyxa* DSM 365 in the aqueous solutions and in the presence of 0.5 % NaCl. G' ●: storage modulus; G'' ■: loss modulus

Table S 3: Thixotropic recovery of paenan variants after a three stage oscillatory shear test. Thixotropic recovery was calculated after 30/90/180 s based on G' relative to initial values determined within the LVE range. n.d.: not determined for this variant due to limited LVE range

		Thixotropic recovery [%]		
Group	**solution**	**30 s**	**90 s**	**180 s**
Paenan I & II & III	aq.	79.3	83.5	85.6
	0.5 % NaCl	82.2	84.7	86.8
Paenan I & II & III (depyruvated)	aq.	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.
Paenan I & II	aq.	97.7	101.9	99.1
	0.5 % NaCl	115.2	114.8	114.4
Paenan I & III	aq.	81.9	91.6	96.4
	0.5 % NaCl	106.4	110.7	112.2
Paenan II & III	aq.	114.1	119.3	121.8
	0.5 % NaCl	174.0	146.9	159.2
Paenan I	aq.	103.2	103.4	103.5
	0.5 % NaCl	107.2	109.6	108.1
Paenan II	aq.	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.
Paenan III	aq.	n.d.	n.d.	n.d.
	0.5 % NaCl	n.d.	n.d.	n.d.

Supplemental References

Schilling, C., Klau, L., Rühmann, B., Aachmann, F., Schmid, J., Sieber, V., 2022a. Structural elucidation of the fucose containing polysaccharide of *Paenibacillus polymyxa* DSM 365. Prep. Manuscr.

Schilling, C., , Klau, L., Rühmann, B. Aachmann, F., Schmid, J., Sieber, V., 2022b. CRISPR-Cas9 driven structural elucidation of the heteroexopolysaccharides from *Paenibacillus polymyxa* DSM 365. Prep. Manuscr.

7.1.5. A novel prokaryotic CRISPR-Cas12a based tool for programmable transcriptional activation and repression

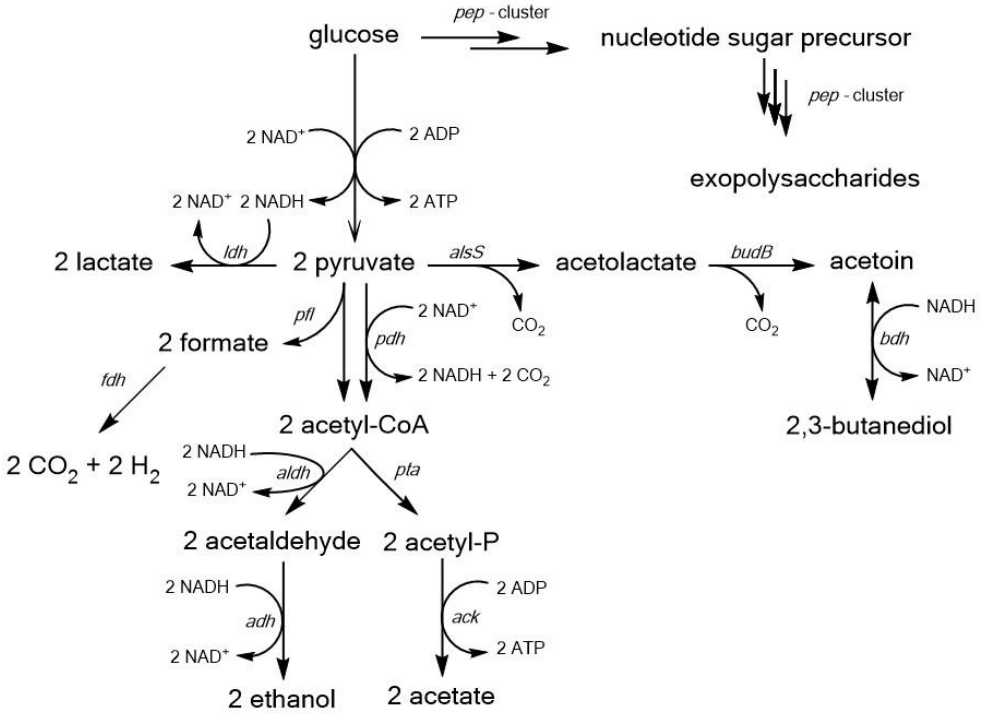


Figure S 1: Overview of the mixed acid pathway and butanediol biosynthesis of *Paenibacillus polymyxa* in microaerobic conditions. Two mols of NADH are formed during glycolysis that need to be regenerated in order to maintain redox balance. Only 1 mol NADH is converted to NAD+ in the 2,3-BDL pathway. Therefore, other redox neutral end products such as lactate or ethanol compete for the intermediate pyruvate. Furthermore, glucose is utilized for the production of exopolysaccharides (EPS). High molecular weight biopolymers are assembled by multiple glycosyltransferases encoded in the *pep*-cluster from nucleotide sugar precursor substrates. Adapted from Schilling et al., (2020)[1].

Table S 1: Bacterial strains used in this study

Bacterial Strains	Genotype	Reference
Escherichia coli S17-1	Conjugation strain; recA pro hsdR; RP42Tc::Mu-Km::Tn7 integrated into the chromosome	ATCC 47055
P. polymyxa DSM365	wild type	DSMZ
P. polymyxa DSM365 *Δldh1*	DSM365 *Δldh1*	Schilling et al., (2020)
P. polymyxa DSM365 Δ*pepCQ*	DSM365 *ΔpepC*	This study
P. polymyxa DSM365 *ΔpepQ*	DSM365 *ΔpepQ*	This study
P. polymyxa DSM365 *ΔpepCQ*	DSM365 *ΔpepC ΔpepQ*	This study

In addition to the strains listed above, each plasmid listed in Table S 2 was used for the transformation of *E. coli* S17-1 as well as for the conjugational transformation of *P. polymyxa* DSM 365.

Table S 2: Plasmids used in this study

Plasmids	Description	Reference
pCasPP	*P. polymyxa* CRISPR-Cas9 genome editing plasmid	Rütering et al (2017)
pCasPPH_pepC	pepC targeting knock out plasmid containing repair template	This study
pCasPPH_pepQ	pepQ targeting knock out plasmid containing repair template	This study
pCRai	PsgsE-AsdCas12a; Pgapdh-off target gRNA; neo; oriT, repU	This study
pCRaiGFP	PsgsE-AsdCas12a; Pgapdh-off target gRNA;neo; oriT, repU; PsgsE-sfGFP	This study
pCRaiGmR_soxS	PsgsE-AsdCas12a-soxS; Pgapdh-off target gRNA; neo; oriT, repU; PsgsE_mRFP-US-PsgsE-sfGFP	This study
pCRaiGFP_GltC	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - off target gRNA	This study
pCRaiGFP_GltC_a1	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a1	This study
pCRaiGFP_GltC_a2	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a2	This study
pCRaiGFP_GltC_a3	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a3	This study
pCRaiGFP_GltC_a4	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a4	This study
pCRaiGFP_sig70	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - off target gRNA	This study

Plasmids	Description	Reference
pCRaiGFP_sig70_a1	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a1	This study
pCRaiGFP_sig70_a2	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a2	This study
pCRaiGFP_sig70_a3	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a3	This study
pCRaiGFP_sig70_a4	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a4	This study
pCRai_GFP_rpoD_a1	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a1	This study
pCRai_GFP_rpoD_a2	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a2	This study
pCRai_GFP_rpoD_a3	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a3	This study
pCRai_GFP_rpoD_a4	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a4	This study
pCRai_GFP_rpoD	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - off target gRNA	This study
pCRai_GFP_soxS_a1	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a1	This study
pCRai_GFP_soxS_a2	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a2	This study
pCRai_GFP_soxS_a3	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a3	This study
pCRai_GFP_soxS_a4	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a4	This study
pCRai_GFP_soxS	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - off target gRNA	This study
pCRai_GFP_CRP_a1	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a1	This study
pCRai_GFP_CRP_a2	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a2	This study
pCRai_GFP_CRP_a3	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a3	This study
pCRai_GFP_CRP_a4	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - sgRNA_a4	This study
pCRai_GFP_CRP	Test plasmid for CRISPRa targeting US region of PsgsE-sfGFP - off target gRNA	This study
pCRaiGFP_soxS_sfGFP_T1	Test plasmid for sfGFP repression sgRNA_T1	This study
pCRaiGFP_soxS_sfGFP_T2	Test plasmid for sfGFP repression sgRNA_T2	This study
pCRaiGFP_soxS_sfGFP_T3	Test plasmid for sfGFP repression sgRNA_T3	This study
pCRai_pepQ_T1	Plasmid for repression of pepQ sgRNA T1	This study
pCRai_pepCQ_T1	Plasmid for dual repression of pepQ and pepC	This study

Plasmids	Description	Reference
pCRai_soxS_ldhmultibdh1	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 1)	This study
pCRai_soxS_ldhmultibdh2	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 2)	This study
pCRai_soxS_ldhmultibdh3	Multiplex knock-down of repression of ldh2, ldh3, ldh4 and activation of Pbdh (sgRNA 3)	This study
pCRaiGmR_soxS_mT2	PsgsE-mRFP-US-PsgsE_sfGFP expression plasmid with gRNA for mRFP repression (sgRNA T2)	This study
pCRaiGmR_soxS_dualT2	PsgsE-mRFP-US-PsgsE_sfGFP expression plasmid with gRNA for mRFP repression (sgRNA T2) and GFP activation (sgRNA a1)	This study
pCRaiGmR_soxS_GFPa1	PsgsE-mRFP-US-PsgsE_sfGFP expression plasmid with gRNA and GFP activation (sgRNA a1)	This study

Table S 3: Oligonucleotides and primers used in this study. Overhangs used for Golden Gate Assembly or Gibson isothermal assembly are depicted in lower case. Restriction sites are underlined.

Name	Sequence 5'->3'	Comment
soxS_fw	ATGTTGGTCTCAACGCTGATGTCCCATCAGAAAATTATTCAGGATC	Cloning of activator domain
soxS_rev	TAAAAGGTCTCGTACCATTTACAGGCGGTGGCGATAATCG	Cloning of activator domain
sig70_fw	ATGTTGGTCTCAACGCTGATGATAGAGGACACACCATTCCGT	Cloning of activator domain
sig70_rev	TAAAAGGTCTCGTACCATTCATTGCCATTCCTCCTCCTTCC	Cloning of activator domain
cAMP_fw	ATGTTGGTCTCAACGCTGATGCATCCCTATTATCGTTCATTAGAAAAAA	Cloning of activator domain
cAMP_rev	TAAAAGGTCTCGTACCATTTAATGCGGATCACTACGCAGCAA	Cloning of activator domain
GltC_fw	ATGTTGGTCTCAACGCTGGTGGAATTACGACAGTTACAATACTTCATGAA	Cloning of activator domain
GltC_rev	TAAAAGGTCTCGTACCATTCACTGCCCTAAATCCGCTTTTG	Cloning of activator domain
RpoD_fw	ATGTTGGTCTCAACGCTGGTGCTGTATCCTTCCATTGATGAAATG	Cloning of activator domain
RpoD_rev	TAAAAGGTCTCGTACCATTTATTCTTCTTCGCCGCTTTCAAGT	Cloning of activator domain

Name	Sequence 5'->3'	Comment
As_CRISPR_fw	GGGGATACGCTAATTTCTACTCT TGTAGATAAGTCTTCTCAGCCG	Construction of pCRai
As_CRISPR_rev	AAATCCAGATGGAGTATGTCTTC ACCGGTGGAAAGCG	Construction of pCRai
AsCRISPR_vec_fw	CACCGGTGAAGACATACTCCATC TGGATTTGTTCAGAACGC	Construction of pCRai
AsCRISPR_vec_rev	GTAGAAATTAGCGTATCCCCTTT CAGATACTCGCAC	Construction of pCRai
enAsCPF1_fw	TTAGGCTTTTACTTAATGACACAG TTTGAAGGCTTTACGAATCTG	Construction of pCRai
enAsCPF1_rev	CGTAGATCTGAATTCTTATTACCC GAGACCTACCCAATGCG	Construction of pCRai
enAsCPF1_Vector.fw	GGTCTCGGGTAATAAGAATTCAG ATCTACGCGTTCCCCGC	Construction of pCRai
enAsCPF1_Vector.REV	TTCAAACTGTGTCATTAAGTAAAA GCCTAAAATCCCCCTTCGTT	Construction of pCRai
GFP_vecint_rev	gcaacgcggccttttacggTTCCTGGCC ATATGACGATCCTCCTTACCTCTC ATTG	sfGFP plasmid cloning
GFP_vecint_fw	gcccgcaaccggcgcatcaagcccgccGAC TAGCAATATGAAACACGGAAAAA ATCAAGC	sfGFP plasmid cloning
sgsE-a1_fw	agatTTTATCTCACATAATAGGGCT	Test gRNA for activation of PsgsE-sfGFP
sgsE-a1_rev	gagtAGCCCTATTATGTGagatAA	Test gRNA for activation of PsgsE-sfGFP
sgsE-a2_fw	agatAGTCTATATCAATCGGTAAC	Test gRNA for activation of PsgsE-sfGFP
sgsE-a2_rev	gagtGTTACCGATTGATATAGACT	Test gRNA for activation of PsgsE-sfGFP
sgsE-a3_fw	agatATCCTCATATTTTCCTAGTA	Test gRNA for activation of PsgsE-sfGFP
sgsE-a3_rev	gagtTACTAGGAAAATATGAGGAT	Test gRNA for activation of PsgsE-sfGFP
sgsE-a4_fw	agatCATCGATGGCGACATTGATA	Test gRNA for activation of PsgsE-sfGFP
sgsE-a4_rev	gagtTATCAATGTCGCCATCGATG	Test gRNA for activation of PsgsE-sfGFP
sfGFP1_T1_fw	agatCGTGCGTGGCGAGGGTGAA G	Test gRNA for repression of PsgsE-sfGFP

Name	Sequence 5'->3'	Comment
sfGFP1_T1_rev	gagtCTTCACCCTCGCCACGCACG	Test gRNA for repression of PsgsE-sfGFP
sfGFP_T2_fw	agatCCATTAGTTGCGTCACCTTC	Test gRNA for repression of PsgsE-sfGFP
sfGFP_T2_rev	gagtGAAGGTGACGCAACTAATGG	Test gRNA for repression of PsgsE-sfGFP
sfGFP_T3_fw	agatAGCTCAATGCGGTTTACCAG	Test gRNA for repression of PsgsE-sfGFP
sfGFP_T3_rev	gagtCTGGTAAACCGCATTGAGCT	Test gRNA for repression of PsgsE-sfGFP
mRFP_T_fw	agatAAAGTTCGTATGGAAGGTTC	gRNA for repression of PsgsE-mRFP
mRFP_T_rev	gagtGAACCTTCCATACGAACTTT	gRNA for repression of PsgsE-mRFP
sfGFP_dual_fw	agatTTATCTCACATAATAGGGCTTAATTTCTACTC	Test gRNA for repression of PsgsE-mRFP and simultaneous activation of PsgsE-sfGFP
sfGFP_dual_rev	acaaGAGTAGAAATTAAGCCCTATTATGTGAGATAA	Test gRNA for repression of PsgsE-mRFP and simultaneous activation of PsgsE-sfGFP
mRFP_dual_fw	ttgtAGATAAAGTTCGTATGGAAGGTTC	Test gRNA for repression of PsgsE-mRFP and simultaneous activation of PsgsE-sfGFP
mRFP_dual_rev	gagtGAACCTTCCATACGAACTTTATCT	Test gRNA for repression of PsgsE-mRFP and simultaneous activation of PsgsE-sfGFP
GGA_sfGFP_fw	AATTGGTCTCAAACTCAATATGAAACACGGAAAAAATCAAGCAG	Cloning of PsgsE-mRFP-US-PsgsE-sfGFP expression cassette
GGA_sfGFP_rev	AATTGGTCTCACTAGTATGACGATCCTCCTTACCTCTCATTG	Cloning of PsgsE-mRFP-US-PsgsE-

Name	Sequence 5'->3'	Comment
		sfGFP expression cassette
GGA_mRFP_fw	AATTGGTCTCATCGACCTGCATACTAGCCTGTTACAGGCATATTCATATCAATGTC	Cloning of PsgsE-mRFP-US-PsgsE-sfGFP expression cassette
GGA_mRFP_rev	CTTTGGTCTCAAGTTAAGCACCGGTGgagtGACGAC	Cloning of PsgsE-mRFP-US-PsgsE-sfGFP expression cassette
pepQ_T1 fw	agatCACCAGCAGTGCAGACAATC	gRNA for repression of pepQ
pepQ_T1_rev	gagtGATTGTCTGCACTGCTGGTG	gRNA for repression of pepQ
pepCQ_T1_fw	agatCCTGACATGACAGCCTCGCCTAATTTCTACT	gRNAs for dual repression of pepCQ
pepCQ_T1_rev	caagAGTAGAAATTAGGCGAGGCTGTCATGTCAGG	gRNAs for dual repression of pepCQ
pepCQ2_T1_fw	cttgTAGATCACCAGCAGTGCAGACAATC	gRNAs for dual repression of pepCQ
pepCQ_T1_rev	gagtGATTGTCTGCACTGCTGGTGATCTA	gRNAs for dual repression of pepCQ
bdh1_fw	cttgTAGATGACACTCATTCTGTGGTATA	gRNA cloning for Pbdh activation
bdh1_rev	gagtTATACCACAGAATGAGTGTCATCTA	gRNA cloning for Pbdh activation
bdh2_fw	cttgTAGATTTCTTGTCTTTGCTTCAATT	gRNA cloning for Pbdh activation
bdh2_rev	gagtAATTGAAGCAAAGACAAGAAATCTA	gRNA cloning for Pbdh activation
bdh3_fw	gagtGCTCGTTACTTTTATACAAA	gRNA cloning for Pbdh activation
bdh3_rev	gagtTTTGTATAAAAGTAACGAGCATCTA	gRNA cloning for Pbdh activation
sfGFP_qPCR_fw	CCCTATTCTGGTGGAACTGGATGG	qPCR primer
sfGFP_qPCR_rev	CAGTAGTACAGATGAACTTCAGCGTC	qPCR primer
gyrA_qPCR_fw	GAGATATGGCCGCTGCGATG	qPCR primer
gyrA_qPCR_rev	GCTCTCTTCACCATCGTAGTTCGG	qPCR primer
mRFP_qPCR_fw	TACCTGAAACTGTCCTTCCCGG	qPCR primer
mRFP_qPCR_rev	GTAGATGAACTCACCGTCTTGCAGG	qPCR primer
ldh3_qPCR_fw	GAGTCATTGGATCA GA CGTTGC	qPCR primer
ldh3_qPCR_rev	AACTCGGAGTCACCGTGTTCTCC	qPCR primer
ldh2_qPCR_fw	GATTATCGGGGTTGGACGCATTGG	qPCR primer

Name	Sequence 5'->3'	Comment
ldh2_qPCR_rev	CAAGGTTGTGAATGGAATGACTTCCTTG	qPCR primer
ldh4_qPCR_fw	GAATGAGCCGGACCTGGATGCG	qPCR primer
ldh4_qPCR_rev	CAATCCTCGACTGGATCGGTGTTC	qPCR primer
bdh_qPCR_fw	AGATGCTCAGCATCCATTGACTGG	qPCR primer
bdh_qPCR_rev	CAACGACACGGTCGCCAACCTG	qPCR primer

Table S 4: spacer sequences and protospacer adjacent motifs (PAM) of each gRNA used in this study

Target	Spacer 5'-3'	PAM 5'-3'
sgsE-a1	TTTATCTCACATAATAGGGCT	TTTC
sgsE-a2	AGTCTATATCAATCGGTAAC	TTTC
sgsE-a3	ATCCTCATATTTTCCTAGTA	TTTC
sgsE-a4	CATCGATGGCGACATTGATA	TTTC
sfGFP_T1	CGTGCGTGGCGAGGGTGAAG	TTTC
sfGFP_T2	CCATTAGTTGCGTCACCTTC	TTTA
sfGFP_T3	AGCTCAATGCGGTTTACCAG	TTTC
mRFP	AAAGTTCGTATGGAAGGTTC	TTTC
pepQ_	CACCAGCAGTGCAGACAATC	TTTC
pepC	CCTGACATGACAGCCTCGCC	TTTA
bdh1	TGACACTCATTCTGTGGTAT	TTTA
bdh2	TTTCTTGTCTTTGCTTCAATT	TTTC
bdh3	GCTCGTTACTTTTATACAAA	TTTG
ldh2_T1	CGATTGGCAGCACAGTTTCC	TTTC
ldh2_T2	CAGTTATGGATCTCTTTGCT	TTTA

Target	**Spacer 5’-3’**	**PAM 5’-3’**
ldh3_T1	ATCACAAATGGACTGATTAA	TTTC
ldh3_T2	TACCATATACGTTACAATGT	TTTG
ykgE_T1	ATCCCAATAGCCGCTGTTAT	TTTC
ykgE_T2	TTCCAGCCTCCGTGCTGAAT	TTTG
YkgF_T1	CGTTACGCAAACGCTCTGTA	TTTC
ykgF_T2	TACAAATAGATTTAAGTAAT	TTTC
ykgG_T1	ATGAAGCCACGCTTCATGGG	TTTG
ykgG_T2	TGCGCGATCCAGTCTGCGGT	TTTG
off-target	AAGTCTTCTCAGCCGCTACA	TTTV

Table S 5: nucleotide sequences of activator domains used in this study

Activator	**Protein Class**	**Nucleotide sequence (5’-3’)**
SoxS (324 bp)	AraC/XylS family	ATGTCCCATCAGAAAATTATTCAGGATCTTATCGCATG GATTGACGAGCATATTGACCAGCCGCTTAACATTGATG TAGTCGCAAAAAAATCAGGCTATTCAAAGTGGTACTTG CAACGAATGTTCCGCACGGTGACGCATCAGACGCTTG GCGATTACATTCGCCAACGCCGCCTGTTACTGGCCGC CGTTGAGTTGCGCACCACCGAGCGTCCGATTTTTGATA TCGCAATGGACCTGGGTTATGTCTCGCAGCAGACCTT CTCCCGCGTTTTCCGTCGGCAGTTTGATCGCACTCCC AGCGATTATCGCCACCGCCTGTAA
GltC (939 bp)	LysR family	GTGGAATTACGACAGTTACAATACTTCATGAAGGTAGC GCAAAAAGAACATGTCACTCAGGCAGCCGAGGAGCTA CATGTGGCGCAATCTGCGGTAAGTCGGCAAATTCATC AGCTTGAAGAGGAGCTGGGAGTTAATCTTTTTATGCAA AAGGGTCGCAATTTGCAGCTTACGCCAGTGGGACAGC TATTTTGCAAACGGGTTGAGACGATAATAAAAGATCTA GAGCGAGCTGTTTTGGAGGTTCATGAGTTTTTGGATCC TGAGGGGGGCGAAATCCGCATCGGTTTTCCGCACAGT TTGGGGATTCATCTTATTCCTACGGTAGTGGCTGAATT CCGAAAGCGGTATCCCAATGTTAAATTTAGATTTAAAC

Activator	**Protein Class**	**Nucleotide sequence (5'-3')**
		AGGGGATGTATCCGAGTTTAATTCGTGATGTGTTAGCG GGTGAGGTGGACTTGGCTTTTGTATCTCCTTTTCCCGA TCGTCATGATCATGTGGAGGGGGATGTGGTGCTGACC GAGGAGTTGTTTGCTGTTCTCCCGCCCAATCATCCTTT AGCTACGGCCAAACATATCACGCTGAGTCAGCTTAAAG GAGAAAAATTTATTCTGTTCAGAGATGGATATTCGTTAC GTCCGATTGTCTGGCAGGCTTGTCTGGAGGCAGGATT TACGCCAGATATCGCTTTTGAAGGTGAAGAGACAGATA CGATTCGCGGATTAGTTGCGGCTGGTATGGGCGTTAG CCTTTTGCCTGAAATGGCGTTGTTTCAGACCAACCCGC TTCAACCTGCAAGAGTGTCAATTGTTGATCCAGAGGTC ACGCGAACTATAGGCTTGATTCATCGCAAAGATGACAA GCTACCATTAGTTGCAAAGTCATTCCGGACTTTTTTGTT GCAATATTTTGGTCTTAAAGGTAGTGTTGCTACTCCGA ATGGAAGTGCAAAAGCGGATTTAGGGCAGTGA
RpoD (666 bp)	Sigma factor	ATGATAGAGGACACACCATTCCGTCAGGAGGGTGTGT CCTTTTTAAATTATTTTTTCGAAATATGCAGGGGAAGGA AATTTCCTGGCGTCTATAGTATTGGGAAGGGGGAGAG CGTCATTGATATAGAGAGAATCATTGAACGAATACGGT CTAGTGACCGGCAGGCGTTCCGTGAGATCGTAGAGCT GTACAGCAAGCATGTATTTCATATCGCCTATTCGGTGC TGCATGACAGCAAGGAAGCAGAAGATGCCGCGCAGGA GGCATTCGTTCAGGTGTACAAGTCTCTCCCCCAGTACC GGAACGAAGGTTTTAAAACGTGGCTGAGCCGAATTGC GCTCCACAAAGCGCTGGATATTAAGCGTAAACAGGAC AGACGGCCCGCAGAGCTAATAGACGTCGCACAATCCC TTGTACAGCTCCCTTCCCGGGATGAAGATGTGCTGGC TCGTCTCATCCGCGAGGAGCAAATTAAAGAACTATCAC AAAAAATCGCTCAGTTGCCCCAGCAACATCGGGATATT ATTCAAGCCTACTATATGCAGGGCAAAACCTACGACCA AATTGCAGAAGAGACGCAGGTAGCTTTAAAAACGGTA GAGTCGAGACTCTACCGGGCGCGACTATGGATTCGAA ACCACTGGAAGGAGGAGGAATGGCAATGA
RpoZ (204 bp)	DNA directed RNA polymerase subunit	GTGCTGTATCCTTCCATTGATGAAATGATGAAAAAGGC AGATAGTAAGTATTCGTTGGTTGTAGCGGCTTCACGTC GTGCCAGACAGTTGCGCGAGGGTGAAAAGACGACCCT CAAGAATCCTAAATCCCACAAACAGGTTGGAGTAGCTT TGGAAGAGATCTATGCAGATCATCTTCTACTTGAAAGC GGCGAAGAAGAATAA
CRP (582 bp)	cAMP receptor	ATGCATCCCTATTATCGTTCATTAGAAAAAATGGCTCAT CATGCCTCGATTATAGAGCAGGATTGGAACAGGTTTGT AGAGCAAACCCGTGTGAAGCATGTTCAGCAGGGAACG GTTCTCATTGAAACGGGTGAGCCTGTGAAACATGCCTA

Activator	Protein Class	Nucleotide sequence (5'-3')
		TTTTTGTTCAGAGGGATTATTCCGTTTATTTTATACATTA GTAGATGGCAGAGAATATAATGTTGGTTTTTCGCCAGA AGATGATTACGTAACCTCTTATGGGGCTATGATCCAAG GAGAACCATCCACTTTTTCGATTGAGGCCATGGAAGAT TCGGTTGTCATTGAAATTCCATATCACGTATTGAAGGA GCTTATGGATACAAGCCATATGTGGGAAAGATTCGTCC GTAAAAGTGTAGAAAGACTTTATATTCGAAAAGAGGAA CGGGAGCGCGAATTATTGTATCTTTCCGCTAAGGAGC GTTATCATGCTTTTCTTCTCAAATACCCCGGATTGGAC AAACGAGTTGCCCAATATCATATAGCTTCGTACATAGG CGTCTCCCCTGTATCACTTAGTCGTTTGCTGCGTAGTG ATCCGCATTAA

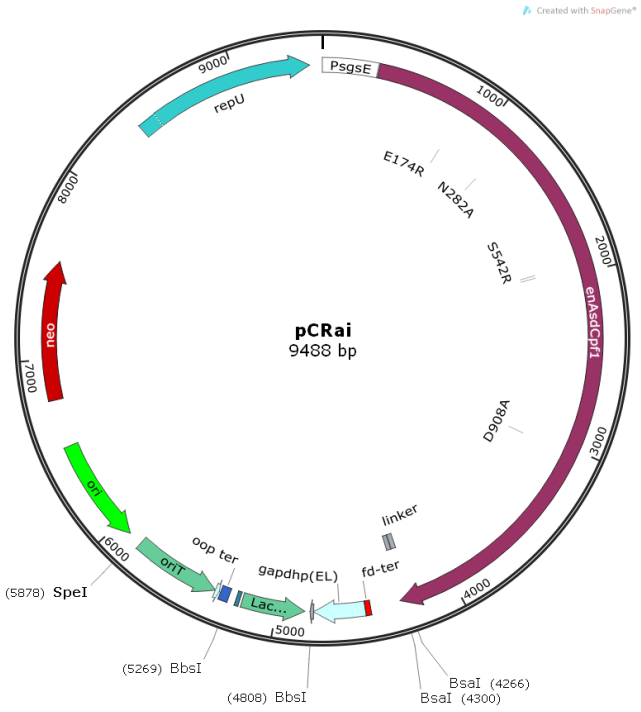


Figure S 2: Plasmid map of pCRai. The conjugational plasmid based on the pCasPP[2] construct was cloned by Gibson isothermal Assembly by replacing the Cas9-gRNA expression cassette with an engineered catalytically inactive variant of Cas12a[3] (E174R, N282A, S542R, K548R, D908A) and an appropriate CRISPR-array. Variable activator domains can be cloned by replacing a BsaI-cassette linked via a 10 aa flexible linker. gRNAs can be cloned by Golden Gate Assembly using BbsI by replacing a *lacZ* expression cassette.

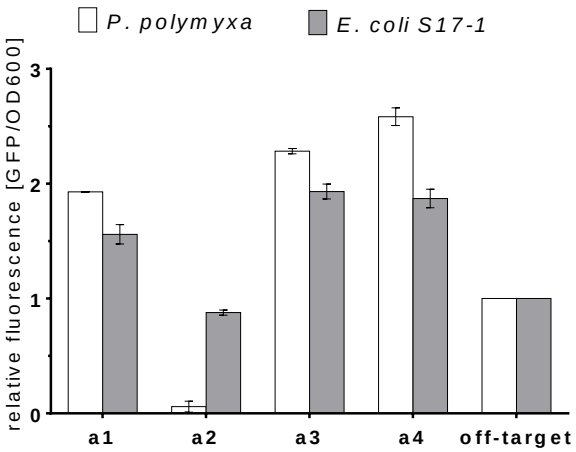


Figure S 3: Relative fluorescence levels of *P. polymyxa* and *E. coli* S17-1 harboring pCRaiGFP_soxS variants. 4 different gRNAs (a1-a4) were expressed and normalized fluorescence levels (Ex. 488 nm Em. 515 nm) were compared to a strain expressing an off-target spacer. Spacers a1, a3 and a4 showed increased GFP fluorescence in both *P. polymyxa* and *E. coli* S17-1 demonstrating similar effects in both, Gram-positive and Gram-negative host organisms.

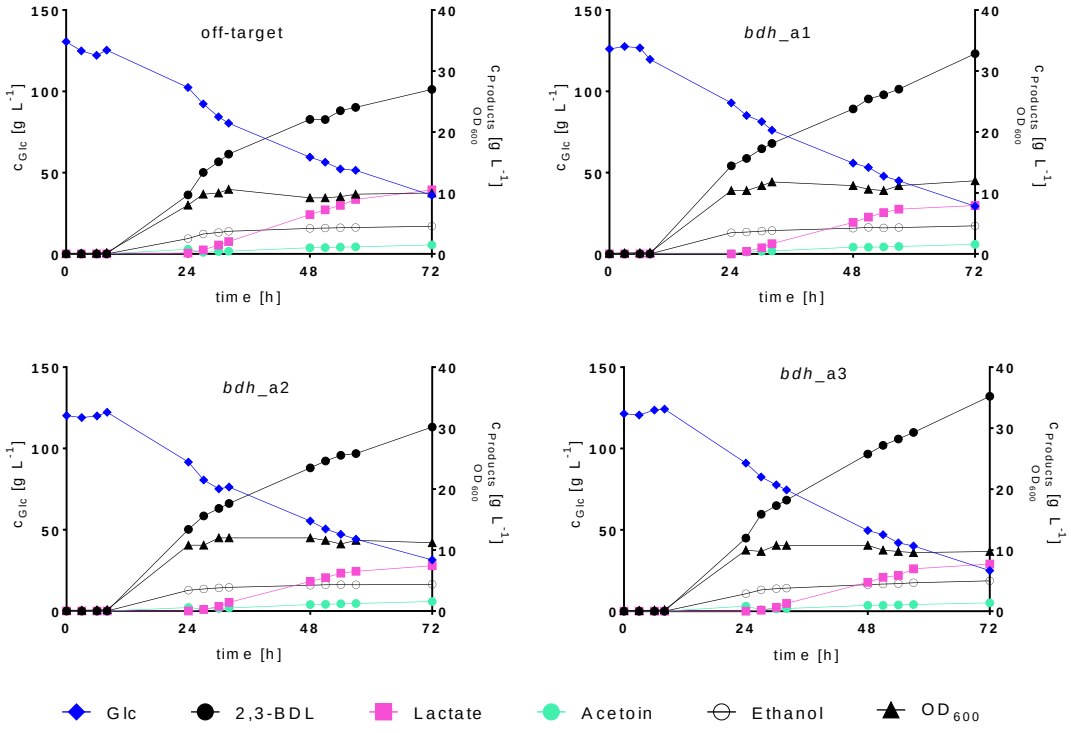


Figure S 4: Fermentation profile of 2,3-BDL fermentations. Overview of fermentation profiles of single batch fermentations. *P. polymyxa* DSM 365 was transformed with pCRai_soxS encoding either off-target gRNAs or a multiplex CRISPR-array targeting the open reading frame of three lactate dehydrogenases and the upstream region of P_{bdh} with three gRNAs (bdh_a1-a3) individually. Lactate production was reduced by ~ 20 % in all variants expressing the corresponding gRNAs. Glc: Glucose; 2,3-BDL: 2,3-*R,R*-butantediol

Table S 5 Final product titers and 2,3-BDL yield (g_{BDL} g_{Glc}^{-1}) of 2,3-BDL fermentations. *P. polymyxa* DSM 365 was transformed with pCRai_soxS encoding either off-target gRNAs or a multiplex CRISPR-array targeting the open reading frame of three lactate dehydrogenases and the upstream region of P_{bdh} with three gRNAs (bdh_a1-a3) individually. Values represent mean and deviation of biological duplicates after 72 h cultivation in microaerobic conditions.

	2,3-BDL	**Ethanol**	**Acetoin**	**Lactate**	**Formate**	**Y_{PS}**
off-target	27.46 ± 0.47	4.21 ± 0.33	1.34 ± 0.14	10.30 ± 0.20	0.34 ± 0.34	0.29 ± 0.01
bdh_a1	34.30 ± 1.45	4.84 ± 0.23	1.50 ± 0.08	7.69 ± 0.26	0.72 ± 0.03	0.34 ± 0.00
bdh_a2	32.44 ± 2.26	4.64 ± 0.25	1.42 ± 0.15	7.74 ± 0.30	0.72 ± 0.04	0.34 ± 0.00
bdh_a3	34.74 ± 0.50	4.89 ± 0.10	1.22 ± 0.08	7.81 ± 0.08	0.67 ± 0.07	0.35 ± 0.02

Supplemental references:

(1) Schilling, C., Ciccone, R., Sieber, V., and Schmid, J. (2020) Engineering of the 2,3-butanediol pathway of *Paenibacillus polymyxa* DSM 365. *Metab. Eng. 61*, 381–388.

(2) Rütering, M., Cress, B. F., Schilling, M., Rühmann, B., Koffas, M. A. G., Sieber, V., and Schmid, J. (2017) Tailor-made exopolysaccharides—CRISPR-Cas9 mediated genome editing in *Paenibacillus polymyxa*. *Synth. Biol. 2*.

(3) Kleinstiver, B. P., Sousa, A. A., Walton, R. T., Tak, Y. E., Hsu, J. Y., Clement, K., Welch, M. M., Horng, J. E., Malagon-Lopez, J., Scarfò, I., Maus, M. V., Pinello, L., Aryee, M. J., and Joung, J. K. (2019) Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol. 37*, 276–282.

7.2. List of figures

7.3. List of tables

8. Curriculum vitae

Christoph Schilling

Geburtstag: 14.11.1991

in Graz, Österreich

seit 11/2020	wissenschaftlicher Mitarbeiter Technische Universität München Lehrstuhl für Chemie Biogener Rohstoffe
07/2017 - 10/2020	Doktorand Technische Universität München Lehrstuhl für Chemie Biogener Rohstoffe
03/2017 - 06/2017	wissenschaftlicher Mitarbeiter Austrian Center for Industrial Biotechnology (ACIB) Wiltschi Group, Graz
05/2016 - 01/2017	Diplomarbeit Evolva SA Basel
02/2015 - 10/2015	iGEM Manchester-Graz, Team Member
10/2014 - 02/2017	Studium der Biotechnologie mit Auszeichnung Diplomingenieur (Dipl. Ing.) Technische Universität Graz
07/2014 - 03/2015	Studentischer Mitarbeiter Technische Universität Graz Christian-Doppler Labor
10/2011 - 09/2014	Studium der Molecularbiology mit Auszeichnung Bachelor of Science (B.Sc.) Technische Universität Graz
08/2010 - 04/2011	Zivildienstleistender Altenheim Gratkorn
09/2001 - 06/2010	Matura mit ausgezeichnetem Erfolg Bundesgymnasium Rein, Steiermark