Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

Jesper Langlo

Development of a markerless tool for targeted deletion of chromosomal genes in the thermotolerant and methylotrophic bacterium *Bacillus methanolicus*

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Marta K. Irla and Trygve Brautaset
June 2020

NTNU
Norwegian University of
Science and Technology

Jesper Langlo

Development of a markerless tool for targeted deletion of chromosomal genes in the thermotolerant and methylotrophic bacterium *Bacillus methanolicus*

Master's thesis in Chemical Engineering and Biotechnology
Supervisor: Marta K. Irla and Trygve Brautaset
June 2020

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biotechnology and Food Science

NTNU
Norwegian University of
Science and Technology

Preface

This thesis marks the end of my master's in Chemical Engineering and Biotechnology at the Faculty of Natural Sciences, Norwegian University of Science and Technology (NTNU). The thesis has been written during the spring of 2020 under the supervision of Postdoctoral Fellow Marta K. Irla, in the Cell Factories group with Professor Trygve Brautaset as principal investigator. The experimental work was conducted at the Microbial Biotechnology Division under the Department of Biotechnology and Food Science (IBT).

In mid-March, the World Health Organization (WHO) declared the novel SARS-Cov2 virus a pandemic. As a result, the Norwegian Government announced the lock-down of all Norwegian universities, including NTNU. Fortunately, I was able to continue the experimental efforts from late April and finish most of the intended laboratory work.

Firstly, I would like to thank my supervisor Marta K. Irla, for letting me work on such an exciting topic and genuinely believe in my work. I am also truly grateful for her constant help, guidance, and encouragement throughout the project, besides always being available when needed. Secondly, I would like to thank my formal supervisor Professor Trygve Brautaset for his follow-up during the extent of this study and for always being available for any questions I may have. In final, my appreciation goes out to the other members of the Cell Factories group for welcoming me as part of the group and for their sharing of knowledge and support.

I also want to extend my gratitude to my family and friends. My fellow students, for making every day as enjoyable and uplifting as possible, and for making the five years in Trondheim a once in a lifetime experience. A special thanks to John L. Pedersen for his sharing of valuable experiences in the laboratory and in thesis writing. Some thanks also go to a good friend and tour mate, Julie C. Claussen, for all the enjoyable hikes and good conversations for the past year. Finally, a special thanks go out to my family for all their care and support.

Trondheim, June 2020
Jesper Langlo

Summary

The bacterium *Bacillus methanolicus* (*B. methanolicus*) serves as an exciting candidate within white biotechnology due to its thermophilic and methylotrophic traits, and for being able to produce value-added compounds from simple non-conventional feedstocks. Its applications have been shown by its ability to produce high yields of amino acids and the fact that it was metabolically engineered for the production of cadaverine, γ-aminobutyric acid, and acetoin. However, further insights into the metabolism of *B. methanolicus* and the construction of genetically stable production hosts require an extended genetic toolbox, especially for genome editing, which has not yet been established for the bacterium.

In this work, a genetic tool for carrying out markerless gene disruptions in the chromosomal DNA of *B. methanolicus* was established. The deletion of two genes emphasized its versatility; *upp* encoding uracil phosphoribosyltransferase, and *mtlD* encoding mannitol 1-phosphate 5-dehydrogenase. The established method is based upon the transfer of the suicide vector pDELxp-*oroP* into *B. methanolicus* through conjugation, followed by two subsequent homologous recombination events. pDELxp-*oroP* was constructed to include a *Kan*^{R} selection marker, an *oroP* counterselection system, and cloned regions upstream and downstream to the targeted gene. The first recombination event occurs between the homology regions of the conjugated suicide vector and the chromosomal DNA, to where the selection marker serves to detect the integrants. The second recombination event occurs intrachromosomally between the homology regions of the integrated vector and the chromosomal DNA, whereby the *oroP* counterselection system selects for isolates where the vector has been cured.

The deletion of *upp* and *mtlD* confirmed the specificity of the established method, as the DNA sequencing data displayed a seamless knockout of both genes. Characterizations of *upp* and *mtlD* deletions on the physiology of *B. methanolicus* were assayed through growth experiments on 5-fluorouracil (5-FU) and MtlD enzyme assay, both of which indicated reduced enzyme activities of the *upp* and *mtlD* encoded proteins. However, some residual enzyme activities were also found, suggesting that alternative enzymes with similar catalytic functions are present. For the deletion of *upp*, bioinformatics tools were used to investigate such alternative enzymes. Moreover, the complementation of *upp* resulted in a restored UPP phenotype and wild-type growth pattern, ultimately confirming its deletion and the tool's specificity only to alter the targeted region in the chromosomal DNA.

To conclude, as this work provides a tool for chromosomal gene deletions in *B. methanolicus*, further studies should be able to clarify the roles of the proposed overtaking enzymes and whether the tool's versatility is reflected by the deletion of other genes. Besides, the possibility of making targeted chromosomal deletions should facilitate the future establishment of production strains with desired traits of *B. methanolicus*.

Oppsummering

Bakterien *Bacillus methanolicus* (*B. methanolicus*) utgjør en svært interessant kandidat innen industriell bioteknologi, i hovedsak grunnet dens termofile og metylotrofiske egenskaper, og for at den kan produsere høy-verdi forbindelser fra enkle ukonvensjonelle råstoffkilder. Bakteriens bruksområder er blant annet vist gjennom dets evne til å produsere betydelige mengder aminosyrer, og forbindelser som kadaverin, gamma-aminobutansyre, og acetoin ved metabolsk prosjektering. Videre innsikt i metabolismen til *B. methanolicus* og konstruksjon av genetisk stabile produksjonsstammer krever imidlertid utviklingen av nye genetiske verktøy, herunder et verktøy for genomredigering, som fortsatt ikke er blitt etablert i bakterien.

I dette studiet ble et genetisk verktøy for utførelse av markørløse gendelesjoner i det kromosomale DNA til *B. methanolicus* etablert, og delesjonen av to gener underbygget dets allsidigheten; *upp* kodende for uracil fosforibosyltransferase, og *mtlD* kodende for mannitol 1-fosfat 5-dehydrogenase. Den etablerte metoden baserer seg på introduksjonen av selvmordsvektoren pDELxp-*oroP* i *B. methanolicus* ved konjugasjon, etterfulgt av to homologe rekombinasjonshendelser. pDELxp-*oroP* ble konstruert til å inneholde en Kan^R seleksjonsmarkør, et *oroP* kontraseleksjonssystem, og klonede regioner oppstrøms og nedstrøms til genet som skal slettes. Den første rekombineringshendelsen finner sted mellom homologe regioner i den konjugerte selvmordsvektoren og det kromosomale DNAet, hvorpå seleksjonsmarkøren selekterer for integranter. Den andre rekombineringshendelsen finner sted intrakromosomalt mellom den integrerte vektoren og det kromosomal DNAet, og *oroP* kontraseleksjonssytemet selekterer for isolater der vektoren har blitt fjernet.

Delesjonene av *upp* og *mtlD* bekreftet spesifisiteten til det etablerte verktøyet, ettersom DNA sekvensering viste en sømløs fjerning av de to genene. Karakteriseringer av de slettede genene på fysiologien til *B. methanolicus* ble undersøkt ved vekstforsøk på 5-fluorouracil (5-FU) og ved MtlD enzymanalyse, hvorpå de *upp* og *mtlD* kodede proteinene viste reduserte enzymaktiviteter. Gjenværende enzymaktivitet ble imidlertid også funnet, og indikerer tilstedeværelsen av andre enzymer med liknende katalytiske funksjoner. For delesjon av *upp* ble bioinformatikkverktøy derfor benyttet for å undersøke slike enzymer. Dessuten, komplementeringen av *upp* resulterte i en restaurert UPP fenotype og villtype vekstmønster, og bekreftet med det både delesjonen av *upp* og verktøyets spesifisitet for å kun medføre ønskede endringer i det kromosomale DNA.

Til slutt, ettersom dette arbeidet bidrar med et verktøy for å kunne utføre kromosomale gendelesjoner i *B. methanolicus*, bør videre studier avklare rollen til de foreslåtte overtakende enzymene til MtlD og UPP, og samtidig vurdere hvorvidt verktøyets allsidighet også er overførbart til andre gener. Dessuten, muligheten for å kunne gjøre målrettede kromosomale gendelesjoner bør også kunne fasilitere etableringen av produksjonsstammer av *B. methanolicus* med ønskede trekk i nær fremtiden.

Table of Contents

List of Tables

List of Figures

Abbreviations

Bp/kb	Basepair/ kilobasepair
Cas	CRISPR-associated proteins
Cm^r	Chloramphenicol selection marker gene
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
crRNA	CRISPR-derived RNAs
dCas9	Catalytically deactivated Cas9
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DSB	Double-stranded break
5-FO	5-fluoroorotate
5-FU	5-fluorouracil
gDNA/RNA	Genomic DNA/RNA
g.o.i.	Gene of interest
HDR	Homology-directed repair
Kan^R	Kanamycin selection marker gene
MDH	Methanol dehydrogenase
NAD	Nicotinamide adenine dinucleotide
NHEJ	Non-homologous end joining
ori	Origin of replication
oriT	Origin of Transfer
PAM	Protospacer adjacent motif
PCR	Polymerase Chain Reaction
repU	Plasmid replication protein gene
RNA	Ribonucleic acid
RO/DI water	Water deionized using reverse osmosis
RuMP	Ribulose monophosphate pathway
sgRNA	Single-guide RNA
Subsp.	Subspecies
TALENs	Transcription activator-like effector nucleases
tracrRNA	Trans-activating CRISPR RNA

Chapter 1

Introduction

1.1 The thermotolerant and methylotrophic *Bacillus methanolicus* as an interesting candidate in white biotechnology

Bacillus methanolicus (*B. methanolicus*) is a rod-shaped, spore-forming, obligate aerobic, thermophilic and facultatively methylotrophic gram-positive bacterium of the class *Bacilli* (Schendel et al. (1990); Arfman et al. (1992b)). The bacterium was first isolated from soil in 1990 and was found able to grow on methanol as the only carbon source at elevated temperatures up to $60\,^{\circ}\mathrm{C}$, with an optimal growth temperature at $50\,^{\circ}\mathrm{C}$. The first observations of the bacterium placed it among the thermotolerant genus of *Bacillus*, and because of its methylotrophic feature it was provided the species name of *methanolicus* (Arfman et al. (1992a)). The genome of *B. methanolicus* was fully sequenced and annotated by Irla et al. (2014), providing information about its methylotrophic metabolism and inherent metabolic pathways. Later research efforts have also made transcriptome, proteome, and metabolome data available for the organism (Irla et al. (2015); Müller et al. (2014); Carnicer et al. (2016)).

B. methanolicus is one of the most explored methylotrophic bacteria to present. One reason for this is because of its unique plasmid-dependent methylotrophy (Brautaset et al. (2004)). Being one out of few known bacteria harboring a plasmid-dependent methylotrophy, *B. methanolicus* serves as an interesting candidate for studying bacterial metylotrophy in general (Müller et al. (2015)). Methylotrophs use reduced one-carbon compounds (C_1) as carbon- and energy source (Anthony (1982)). However, *B. methanolicus* is regarded as a facultative methylotroph, meaning that besides methanol it is also able to utilize other compounds, such as glucose and mannitol as carbon- and energy source (Heggeset et al. (2012)). A newly discovered gene cluster in its genome indicated that it might also be able to use D-arabitol for growth. Additional growth studies also provided evidence for this, and the genes were found to be highly up-regulated upon growth on D-arabitol, indicating an arabitol inducible operon for D-arabitol utilization in

the bacterium (López et al. (2019).

The most characterized strain of *B. methanolicus*, MGA3, has been shown to carry the two plasmids pBM69 and pBM19, were the latter has been shown to include a methanol dehydrogenase (*mdh*) gene. The gene encodes a familiy III group NAD(P)-dependent alcohol dehydrogenase, MDH, which is made catalytically active by an activator protein, Act. This enzyme is responsible for catalyzing the oxidation of methanol into the intermediary metabolite formaldehyde (Arfman et al. (1991, 1992b, 1997); Irla et al. (2014)) (Kloosterman et al. (2002); Brautaset et al. (2004); Arfman et al. (1992b)). Pluschkell and Flickinger (2002) showed that a sudden increase in methanol concentration leads to accumulation of formaldehyde in the cell, indicating MDHs relatively low turnover number for high methanol concentrations. Besides, in addition to the chromosomal gene copies, the pBM19 plasmid was found to carry five genes (*glpX*, *fba*, *tkt*, *pfk*, and *rpe*) involved in the fructose-1,6-biphosphate aldolase-transaldolase type of the ribulose monophosphate pathway (RuMP), which is responsible for the assimilation of toxic formaldehyde in the cell (Heggeset et al. (2012)). The transcription levels of the aforementioned genes have been shown to greatly influence the methanol-tolerance of the cell (Jakobsen et al. (2006)), and hence its growth. Furthermore, transcriptome data have shown that these RuMP-related genes were up-regulated at most 40-fold upon growth on methanol versus mannitol, while the chromosomally encoded gene copies were presumably unaffected (Arfman et al. (1991); Jakobsen et al. (2006)). Curing of the plasmid resulted in no growth in minimal medium with methanol as sole C-source, indicating the importance of the plasmid for utilizing C_1 compounds for cell growth and maintenance. It is assumed that growth on feedstocks other than methanol causes pBM19 to become a burden on the cellular metabolism of *B. methanolicus* (Jakobsen et al. (2006)).

In the fructose-1,6-biphosphate aldolase- transaldolase type of RuMP, three molecules of formaldehyde are assimilated eventually leading to the formation of pyruvate. As displayed in Figure 1.1, the RuMP pathway begins with the assimilation of three molecules of formaldehyde by condensation with three molecules of ribulose-5-phosphate (Ru-5-P), catalyzed by hexulose phosphate synthase. The reaction yields three molecules of hexulose-6-phosphate (HPS), which are further isomerized in the reaction catalyzed by phosphate hexuloisomerase (PHI) forming three molecules of fructose-6-phosphate (F6P).

One out of the three F6P is then phosphorylated by phosphofructokinase (PFK) yielding fructose-1,6-bisphosphate (FBP). In the next phase, FBP is cleaved in the reaction catalyzed by fructose bisphosphate aldolase, forming glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). In the cell, a reversible interconversion between GAP and DHAP occurs regularly by the catalytic activity of triosephosphate isomerase (TPI). One molecule of DHAP is then converted by multiple enzymatic steps yielding one molecule of pyruvate. These steps will in net reduce one nicotinamide adenine dinucleotide (NAD+) and produce two adenosine

triphosphate (ATP). The remaining molecules of F6P, FBP, and GAP are eventually recycled to form three molecules of Ru-5-P, which further can assimilate additional formaldehyde in the cell (Heggeset et al. (2012); Arfman et al. (1989); Pluschkell (1999)).

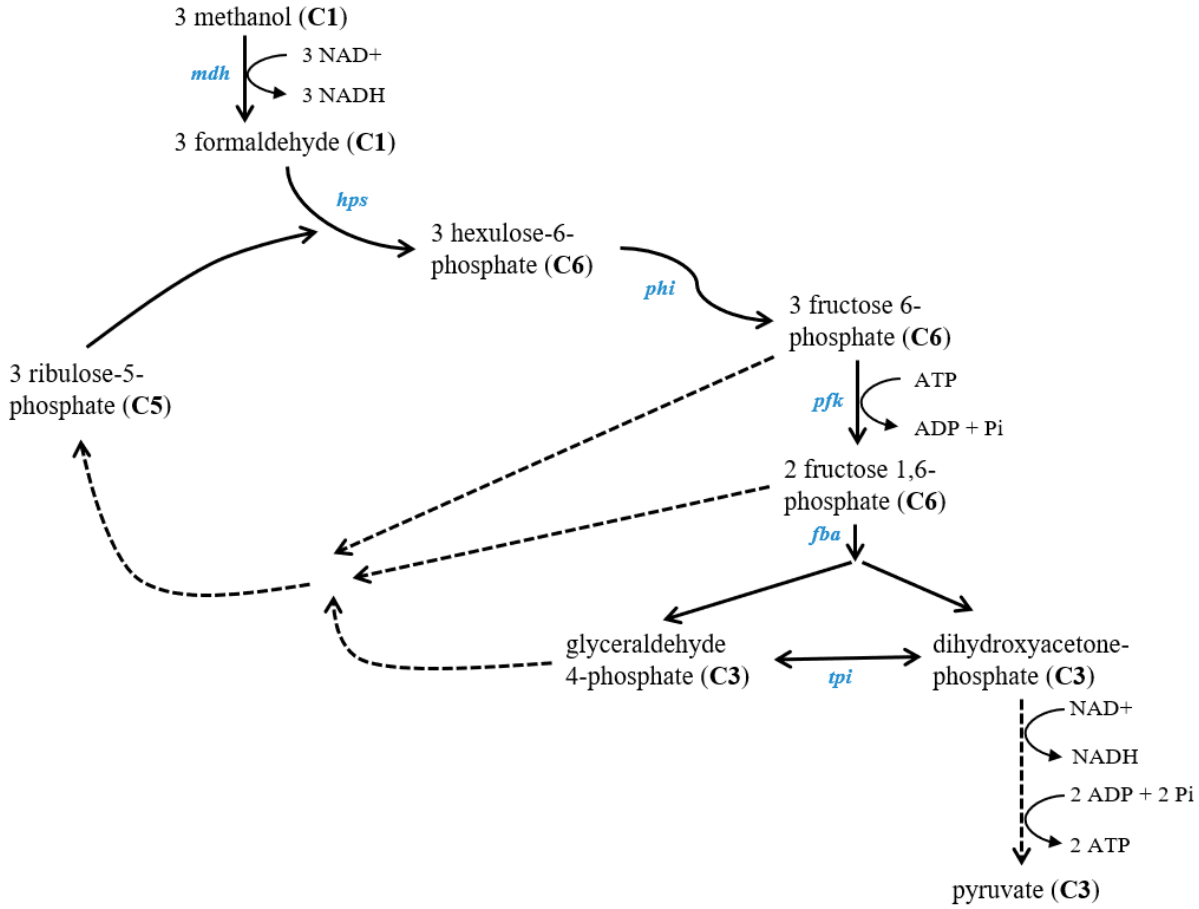


Figure 1.1: Assumed formaldehyde assimilation pathway in *B. methanolicus*, displaying the fructose-1,6-bisphosphate aldolase-transaldolase type of the ribulose monophosphate pathway (RuMP). The pathway results in the production of one mole of pyruvate from the glycolysis by the consumption of three moles of formaldehyde. Genes are displayed in blue and encode the following enzymes: *mdh* - methanol dehydrogenase, *hps* - hexulose phosphate synthase, *phi* - phosphate hexuloisomerase, *pfk* - phosphofructokinase, *fba* - fructose bisphosphate aldolase, *tpi* - triose phosphate isomerase. Dotted lines corresponds to multiple enzymatic steps which are excluded from the pathway.

The netto outcome resulting from the assimilation of three molecules of methanol is shown in Equation (1.1), ultimately reducing four NAD^+ and producing one ATP.

$$3CH_3OH + 4NAD^+ + Pi + ADP \rightarrow C_3H_4O_3 + 4NADH + ATP \tag{1.1}$$

The major interest in *B. methanolicus* as a microbial cell factory relates to its methylotrophic and thermophilic traits, besides being able to produce high amounts of L-glutamate ($59\,g/L$) naturally and L-lysine ($65\,g/L$) from classical mutants (Brautaset et al. (2007, 2010)). Studies have

also resulted in recombinant strains able to produce cadaverine, acetoin, and γ-aminobutyric acid (GABA) by metabolic engineering (Drejer et al. (2020); Naerdal et al. (2015); Irla et al. (2017)), hence emphasizing its potential as a potential microbial cell factory. More commonly industrially used microorganisms, such as *B. subtilis* are greatly explored with a wide and well-developed genetic toolbox for metabolic engineering, and simple requirements for cultivation (van Dijl and Hecker (2013); Perkins et al. (1999)). However, one of the main advantages of using *B. methanolicus* over other well-established cell factories is particularly related to its cultivation conditions. The risks of contamination are low due to cultivation on methanol, which is toxic to most other microorganisms, aside from the relatively high temperatures of fermentation. The latter implicate that even though fermentation generates some heat, which increase the temperature of the fermentation broth, the cooling costs are lowered as the bacterium is able to grow at elevated temperatures (Brautaset et al. (2007); Cotton et al. (2020)). Moreover, the use of methanol is one of the main reasons for the increased interest in synthetic methylotrophy (Antoniewicz (2019), and *B. methanolicus* in particular.

1.1.1 Methanol as feed-stock for industrial biotechnology applications

Today, value-added products are usually produced by chemical synthesis (Acevedo-Rocha et al. (2019); Yu et al. (2014)). These methods often accompany low yields and time-consuming procedures. Besides, the complexity of many compounds serve as major obstacles for their production by chemical synthesis. A solution is found in microorganisms being able to produce such compounds naturally (Du et al. (2011)). At present, the major drawback of using microorganisms for industrial production of chemicals is the widespread use of sugars and molasses as feedstock for fermentations, as these substrates compete with food industry for resources. The search for microorganisms able to use other non-food compounds for growth have thus gained an increased interest (Müller et al. (2015)).

Methanol is a C_1-compound which is not directly used in the production of human nutrition and is commonly a product of many industrial processes, especially the processing of natural gas (Cheng and Kung (1994); Methanol-Institute (No date)). Moreover, methanol is fully soluble in water and readily available in large quantities, estimated annually to over 1.8 metric tons (Methanol-Institute (2020)). To date, many projects are devoted to the establishment of inexpensive and sustainable methods for alternative ways of methanol production, in particular by hydrogenation of CO_2 (Borisut and Nuchitprasittichai (2019); Bukhtiyarova et al. (2017); An et al. (2009); Sánchez-Díaz (2017)). Besides, methanol has shown to provide high microbial growth, primarily when the RuMP pathway is utilized, which facilitates methanol conversion into biomass with an energetic efficiency of up to $50\,\%$ (Cotton et al. (2020)). However, since methanol is a reduced compound, even more than glucose, its consumption during fermentation requires high oxygen supply, thereby increasing the costs of oxygen supply and cooling as heat

is formed (Müller et al. (2015)). Nevertheless, the benefits of using methanol make it serve as an exciting and promising non-food feedstock for white biotechnology.

1.1.2 The genetic toolbox of *Bacillus methanolicus*

As being one of the most explored and studied thermophilic and methylotrophic bacteria, the use of *B. methanolicus* as a microbial cell factory is limited by a shortage in genetic tools for metabolic engineering. This limitation is linked to particularly methylotrophic bacteria in general, and is one of the main reasons why methylotrophic bacteria presently are not widely used for industrial biotechnology applications (Cotton et al. (2020)).

To meet the need for proper genetic tools in *B. methanolicus*, much of the research activity for this bacterium has been directed towards expanding its genetic toolbox. For instance, multiple reporter genes have been investigated for the examination of gene expression in the organism. Some of these involve *lacZ* (Andersen (2011)), *gfpuv* encoding green fluorescent protein and mCherry encoding red fluorescent protein (Irla et al. (2016), Irla. oral communication). Besides, both constitutive and inducible promoters, such as the methanol inducible dehydrogenase (*mdh*), the xylose inducible promoter from *Bacillus megaterium*, and the mannitol inducible promoter have been established for expression of recombinant gene expression in *B. methanolicus* (Irla et al. (2016); de Vries et al. (1992); Yi et al. (2018)). In their following efforts, Irla et al. (2016) have investigated compatible plasmids with theta- and rolling circle replication systems, enabling co-expression and simplifying over-expression of genes in the organism.

Molecular cloning work, in general, is usually conducted in a rapidly growing cloning host with a high transformation efficiency. As *B. methanolicus* has a doubling time of approximately 2 hours (Irla, oral communication) and a low transformation efficiency compared to other well-used bacteria, strains of *E. coli* are more commonly used. Therefore, as the construction of recombinant DNA has to be carried out extracellularly of the desired expression host, tools for efficient introduction of DNA into *B. methanolicus* is also preferred. Today, both methods for transformation and conjugation have been successfully established for this purpose (Jakobsen et al. (2006), Irla et al. (2019, unpublished)).

To present, most genetic tools of *B. methanolicus* are directed towards molecular cloning, and for the investigation and regulation of gene expression. Tools for repression and disruption of genes are on the other hand very limited. Such tools are nevertheless much needed for the investigation of gene functions and for establishing polite strains for industrial purposes, especially through pathway optimization such as pathway de-branching and by interupting enzymes involved in product degradation (Davy et al. (2017)). These developments may in turn allow for genome editing, which in certain cases are more preferred over plasmid-based genomic engineering as it does not require a selection pressure and permits implementation of lasting genetic

changes in the genome. A further contribution by Schultenkämper et al. (2019) have involved the use of CRISPR interference (CRISPRi) for sequence-specific repression of gene expression. With this tool, it should be possible to investigate the expression of genes under different conditions, enabling exploration of the resulting phenotypes and physiology of the cells in general. Still, in terms of genome editing, no genetic tools for chromosomal gene knockouts have yet been established in *B. methanolicus*.

1.2 Targeted deletion and repression of genes in chromosomal DNA of bacteria

Historically, the use of microorganisms for the production of valuable products has existed for thousands of years. In particular, Bacteria have gained increasing interest as industrial workhorses mainly due to their typically high growth rates, the fact that they contain multiple biosynthetic pathways of commercial interest and ease of cultivation and genetic engineering (Du et al. (2011)).

Being able to alter the expression of genes can be very useful when examining the function of genes in a bacteria, and the relationship between genotypes and resulting phenotypes (Vento et al. (2019); Qi et al. (2013)). However, to optimize biological pathways for the production of high-value products, a set of genetic tools must be available for the bacterium of interest. These tools should ease the genetic engineering of an organism and redesign the metabolic fluxes in a precise and efficient way. Previously, genetic modifications in organisms have commonly been made using classical mutagenesis; using mutagens to cause alterations in a genotype, hence affecting the resulting phenotype (Snustad and Simmons (2012)). One alternative, a modern and new approach, involves recruiting a living organism's biosynthetic machinery to change the expression of genes. These methods include RNA interference, Zinc fingers, and transcription activator-like effector nucleases (TALENs). Despite being very precise in gene expression regulation, these methods have their limitations, primarily related to their design and complexity, which makes the process very time-demanding and often highly expensive (Larson et al. (2013); Yao et al. (2018); Gaj et al. (2013)). Besides, these methods can result in unwanted off-target effects on gene expression (Larson et al. (2013); Kim and Rossi (2008)). The search for a more efficient, inexpensive, and specific genomic editing tool was found in the clustered regularly interspaced short palindromic repeats (CRISPR), and CRISPR-related proteins (Cas) system found inherent in some bacteria and archaea (Bhaya et al. (2011); Pickar-Oliver and Gersbach (2019)).

1.2.1 CRISPR/Cas system in bacteria and archaea

During evolution, bacteria and archaea have developed an adaptive immune system to encounter foreign and potential harmful nucleic acids originating from invading bacteriophages (Jinek et al. (2012); Yao et al. (2018); Hille et al. (2018)). Such an endogenous system of immunity has been found to be present in about 40% of all investigated bacteria (Godde and Bickerton (2006)), and usually, multiple variants can present in an organism (Wiedenheft et al. (2012)). The CRISPR/Cas system of bacteria and archaea works by integrating fragments of the foreign nucleic acid in proximity to a region of clustered regularly interspaced short palindromic repeats (CRISPR) in their genomes. Once transcribed, these genomic elements constitute CRISPR-derived RNAs (crRNAs), which are highly involved in the immune response and interaction with segments of foreign nucleic acids (protospacers), due to its complementary feature (Wiedenheft et al. (2012)). Besides, Cas proteins encoded by *cas* genes in *cas* operons are usually found close to a CRISPR array and work as endonucleases guided by the crRNAs (Makarova et al. (2011); Yao et al. (2018); Jinek et al. (2012)). The group of Cas proteins is very diverse, and only a few are common in most CRISPR/Cas systems. Besides, different Cas proteins have also been found to constitute different enzymatic functions, and the diversity and conservation of Cas proteins vary between species of bacteria (Makarova et al. (2018)).

In short, the CRISPR/Cas system for handling foreign nucleic acids works in three phases; spacer acquisition or adaption phase, biogenesis of crRNAs, and in final, the interaction with foreign protospacers (Yao et al. (2018); Jinek et al. (2012)).

In the first phase, protospacers are integrated into the chromosomal DNA next to a CRISPR array. Once integrated, these regions are known as spacers. The transcription of a CRISPR array by an RNA polymerase generates pre-crRNAs which are converted into mature crRNAs through enzymatic cleavage. The processing is done by specific Cas proteins and varies in different CRISPR/Cas systems. These mature crRNAs are capable of hybridizing with complementary parts of a protospacer sequence (Jinek et al. (2012); Hille and Charpentier (2016)). The binding of crRNAs to foreign nucleic acids guides bounded Cas proteins with endonuclease activity to targeted sequences for enzymatic cleavage, and induction of double-stranded breaks (DSB) in the target sequence (Gasiunas et al. (2012); Barrangou (2013)). The binding specificity of some Cas proteins is also determined by a short sequence downstream to the target site, known as the protospacer adjacent motif (PAM). Besides, the presence of a specific PAM sequence is required for a Cas protein to distinguish between the host DNA and foreign nucleic acids (Westra et al. (2013); Jinek et al. (2012); Qi et al. (2013)).

In response to a DSB in the cell, DNA repair can occur naturally by one of two competitive methods: homology-directed repair (HDR) or by non-homologous end joining (NHEJ) (Yao et al. (2018); Bowater and Doherty (2006)). In HDR, exogenous DNA can be introduced into the cut-site by homologous recombination, which is less efficient but very precise compared

to NHEJ. This happens as an introduced repair template with the desired insert is flanked by homologous regions upstream and downstream to the part were the DSB has occurred. However, the NHEJ type of DNA repair is not dependent on homologous regions, as DSBs, on the contrary, are directly ligated without the need for a repair strand. As a result, DSB repair by NHEJ is more efficient and error-prone than the HDR type, which, on the other hand, poses a more high-fidelity DNA repair system. Moreover, mutations at the target site resulting from the NHEJ repair system are not uncommon and may involve base insertions, deletions, and frameshift mutations (Addgene (No date); Pawelczak et al. (2017)).

In general, the CRISPR/Cas systems are divided into classes 1 and 2 based on their involvement of one or multiple Cas protein effectors. Class 1 systems use Cas protein effectors composed of multiple subunits to degrade foreign nucleic acids, and class 2 systems only use protein effectors constituting one single unit (Tang (2019); Shmakov et al. (2015); Cho et al. (2018)). For class 2 CRISPR/Cas systems, the maturation of crRNA usually requires the involvement of a trans-activating CRISPR RNA (tracrRNA), which binds to the crRNA by common complementary parts, forming an RNA-complex known as a guide RNA (gRNA) (Fujita and Fujii (2015)). The classification of CRISPR/Cas systems is further divided into six different types (I-VI) and 33 sub-types (Koonin et al. (2017)). However, this amount is expected to increase as research on CRISPR/Cas systems reveals new details by exploring other variants found in bacteria that yet have not been elucidated (Yao et al. (2018); Martynov et al. (2017); Makarova et al. (2011, 2019)).

1.2.2 Use of the CRISPR/Cas system for gene regulation and deletion

The use of CRISPR/Cas systems for genome editing relies on the action of two events. Firstly, it requires the recruitment of a Cas nuclease, which induces specific DSB at sequences complementary to the sgRNA in the genome. For specific genome editing, modifying the 20 nucleotide sequence of the crRNA complementary to the target sequence can be made, hence directing the Cas protein at specific target locations for initiation of DSBs. In the second part, the editing occurs as the host's DNA repair system causes alterations in the target sequence at the location of the DSB (Sander and Joung (2014)).

Two types of CRISPR/Cas systems; II-A CRISPR/Cas9 and V-A CRISPR/Cas12A, are at present the most elucidated (Yao et al. (2018)). Out of these, the natural type II CRISPR/Cas9 system derived from *Staphylococcus pyogenes* (*S. pyogenes*) is probably the most used CRISPR/Cas system for genomic engineering and regulation purposes, due to its simplicity (Yao et al. (2018); Sander and Joung (2014)). This system only needs an RNA-guided endonuclease called Cas9, and a crRNA-tracrRNA complex (Qi et al. (2013); Jinek et al. (2012); Chen et al. (2017)). The Cas9 induces DSB by its His-Asn-His (HDH) and RuvC domains. Both the HNH and the RuvC

domain pose catalytic nuclease activity, in which the first is assumed to be responsible for the cleavage of the target DNA that is complementary to the target-specific parts of the crRNAs, and the latter for the cleavage of the non-complementary DNA strand (Gasiunas et al. (2012); Jinek et al. (2012)). When re-engineering the Cas9, multiple loci can be targeted simultaneously. The CRISPR/Cas9 system of *S. pyogenes* has, among others, been previously used to introduce successful markerless gene deletions in *Escherichia coli* (*E. coli*) and *Streptococcus pneumoniae* (*S. pneumoniae*) (Yao et al. (2018); Jiang et al. (2013); Wang et al. (2015)), providing a basis for genomic engineering in other bacteria as well.

One main application of the CRISPR/Cas9 for genomic editing relies on the HDR method of DNA repair. When a DSB has occurred in the genome, HDR can be used for making precise changes in a target sequence (Xu et al. (2020)). Furthermore, despite being a valuable tool for genetic manipulations, the HDR method for genome editing has its limitations. Firstly as the HDR method has a relatively low efficiency (Xu et al. (2020)). Secondly, as the gRNA can have multiple partial homologous regions in the genome, which may result in off-target effects (Zhang et al. (2015)). Today, numerous bioinformatics tools have been made for predicting the occurrence of such off-target effects (Lin and Wong (2018); Peng et al. (2018)). The second application of CRISPR/Cas systems for genome editing uses the alternative type of DNA repair, NHEJ. Because NHEJ commonly introduces mutations in the target site, specific cleavage by Cas9 with subsequent DNA repair can be used to study the resulting phenotypes of a gene to determine its function in the cell (Pickar-Oliver and Gersbach (2019)). Besides, this method can also be used for targeted mutations to produce strains with preferred phenotypes. This is contrary to classical mutagenesis, which generates random mutations.

Despite constituting several beneficial properties for the host organism, the CRISPR/Cas9 system can also pose some difficulties. For instance, the need for a PAM sequence close to the target loci is usually required for a sgRNA-Cas9 complex to bind, limiting its use for target-specific cleavage. Research efforts to alter the specificity of the Cas9 is, therefore, a hot topic in CRISPR/Cas genome editing. Furthermore, the number of intracellular resources required for both the production and maintenance of the system are considered a heavy burden for the host cell. This also poses a challenge when introducing the system to new organisms exogenously. Moreover, high intracellular concentrations of Cas9 should also be avoided as this can be highly toxic, eventually leading to cell damage or death (Yao et al. (2018); Vento et al. (2019)).

In white biotechnology, CRISPR/Cas technology has been shown to enable inexpensive and simple optimization for the production of high-value products from microbial cell factories (Wang et al. (2015); Yao et al. (2018); Chen et al. (2017)). Besides, since CRISPR-technology can be used for a more in-depth insight into the metabolism of a wide variety of organisms and the function of their genes (Wang et al. (2015)), its use will probably be of great importance in the years to come.

1.2.3 Gene silencing using CRISPR interference

The advancements in CRISPR/Cas technology has led to a further developed method for gene silencing, known as CRISPR interference (CRISPRi). This novel use of the CRISPR/Cas9 machinery was first described by Qi et al. (2013), and uses a modified Cas9 protein which is catalytically deactivated (dCas9) by introduction of a mutation in the HNH- and RuvC- domains of Cas9 (Larson et al. (2013)).

In order to further simplify the system for genomic engineering purposes, the tracrRNA and the crRNA are linked into one single RNA molecule known as a single-guide RNA (sgRNA), exploiting the same combined activity as the two single RNAs independently constitute in a gRNA complex. The sgRNA then guides the dCas9 protein to a target site in the complementary DNA sequence by specific binding. As the Cas protein lacks nuclease activity, the binding does not result in DSB at the target locus. Instead, the binding will instead prevent transcription by the RNA polymerase leading to silencing of gene expression. Besides, the silencing effectiveness can be regulated by altering the 20 nucelotide part of the sgRNA sequence complementary to the target DNA, or by directing the binding to different parts of a locus. Due to the high specificity of the CRISPR/Cas system, CRISPRi can be used for specific and targeted gene silencing with a high specificity and efficiency (Jinek et al. (2012); Larson et al. (2013); Yao et al. (2018); Qi et al. (2013); Cho et al. (2018)).

As multiple sgRNAs present at the same time do not affect each other, CRISPRi can be used for silencing of multiple genes simultaneously and reversibly (Larson et al. (2013); Qi et al. (2013)). Furthermore, the simplicity of the system makes it possible to transfer it to different bacteria. It thus constitutes a powerful tool for gene regulation and pathway optimization in industrially interesting organisms (Cho et al. (2018)). The use is extended as CRISPRi does not make physical changes in the target DNA. This allows subsequent examination of the expression of multiple genes in the organism (Qi et al. (2013)).

However, the use of CRISPRi has in similarity to the CRISPR/Cas system its limitations on genome editing. As the CRISPR/Cas9 system for *S. pyogenes* is primarily used for CRISPRi, its requirements are also directly transferable. For instance, the need of a PAM region in close proximity to the target loci limits the potential target sites in the genome. Additionally, as the target specificity is determined by the region of the sgRNA homologous to the target sequence, partial homologous regions can be present in multiple parts of the genome, which conclusively can result in off-target effects (Larson et al. (2013). The method of gene silencing using CRISPRi has been established in multiple bacteria such as *Candida albicans* and *B. subtilis* (Wensing et al. (2019); Peters et al. (2016)). Recently the CRISPRi system has also been established in the bacterium *B. methanolicus*. Research efforts by Schultenkämper et al. (2019) introduced the CRISPR-dCas9 system derived from *S. pyogenes* into *B. methanolicus* by conjugation in order to silence the expression of the following genes; *spo0A* encoding a sporulation

protein, *mtlD* encoding a mannitol 1-phosphate 5-dehydrogenase, and the gene *katA* encoding a catalase. Silencing of all genes using CRISPRi yielded significantly lower enzyme activities, exemplifying the potential of CRISPRi in research on cell physiology in *B. methanolicus* and bacteria in general.

One of the prerequisites for a commercial microbial cell factory is that its genome is genetically stable. The current use of CRISPRi for genome editing is considered less stable than alternative methods which physically make permanent alterations in the genome (Schultenkämper et al. (2019)). Other challenges are related to the dCas9 originating from *S. pyogenes*, which has a reduced activity at high temperatures above $42\,^{\circ}\mathrm{C}$ (Mougiakos et al. (2017a); Schultenkämper et al. (2019)). Low thermostability in combination with the toxicity of Cas9/dCas9 at high levels, limits the use of the CRISPR/Cas9 system from *S. pyogenes* in thermophilic production hosts. Therefore, establishing genetic tools to control the activity of Cas9/dCas9, constitutes a bottleneck for future use of CRISPR/Cas in genome editing. The search for novel CRISPR/Cas systems in thermophile bacteria could provide an additional contribution to the present CRISPR-technologies potentially providing more reliable and thermostable Cas9 enzymes (Mougiakos et al. (2017b)).

An alternative to CRISPR/Cas systems for introduction of sustained genomic changes in an organism is to exploit a phenomenon that occurs naturally in bacteria, called homologous recombination. Like CRISPR/Cas, this method is very specific as it is based on events of recombination between homologous parts of DNA sequences in a cell. Because of the broad diversity of bacteria, genomic editing using homologous recombination can thus serve as a preferred alternative to CRISPR/Cas and CRISPRi in some bacterial species (Schultenkämper et al. (2019)).

1.2.4 Homologous recombination based gene deletion

A smooth and efficient method for gene replacement and deletion of genes in *E. coli* was first described by Hamilton et al. (1989). This method leads to alterations in the genome based on the outcome of two subsequent homologous recombination events. A flowchart illustrating these two steps is displayed in Figure 1.2.

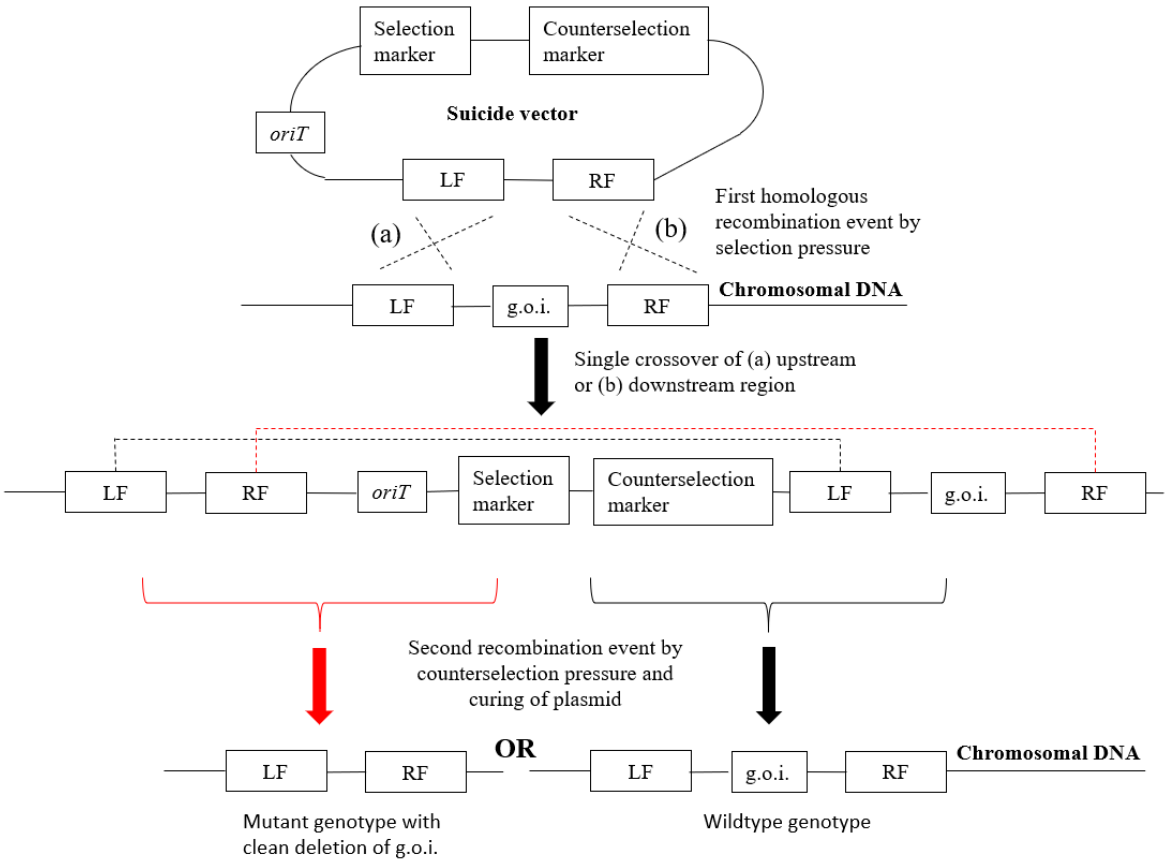


Figure 1.2: Flowchart of a markerless and seamless deletion of the gene of interest (*g.o.i.*). In the first homologous recombination event, recombination between homologous regions of a suicide vector and its corresponding homologous regions in the chromosomal DNA occurs either by single crossover at the left flank (LF) or the right flank (RF), eventually integrating the plasmid in the chromosome. The selection marker selects for isolates that have integrated the plasmid. Following the integration of the plasmid, a second recombination event occurs between the homologous regions within the chromosome due to selection pressure in the presence of a counterselection compound. This results in either the retaining of the wildtype genotype or deletion of the *g.o.i* and plasmid curing of the cell.

Before the first recombination event, an integrative plasmid unable to replicate within the host organism and containing homologous regions flanking the gene of interest (g.o.i.), has to be constructed. Such a plasmid is known as a suicide vector. It lacks an origin of replication (*ori*) necessary for replication in the host cell and can, hence, persist only by integration into the host chromosome. The homologous regions are obtained by cloning of the upstream and downstream regions of the g.o.i. (Huang and Wilks (2017); Hamilton et al. (1989)). The size of these flanking regions for efficient recombination may vary in different organisms. Still, studies in *Xylella fastidiosa* and corresponding experimental data have given indications that regions of about 100 bps to 1 kb are optimal for the recombination to occur, and that the recombination efficiency may vary by the size of the regions (Kung et al. (2013)).

The cloning and assembly of the suicide vector are done in a cloning host different from the target organism. In brief, the vector includes a selection marker, a counterselection marker, and

an origin of transfer (*oriT*) for bacterial conjugation. A counterselectable marker is a gene that encodes a protein whose activity strongly inhibits cell growth in the presence of a counterselection compound. Such a gene can thus be highly unfavorable for the cell when present in the genome. The counterselection marker is meant to impose selection pressure in the presence of a counterselection compound and for curing of the integrated plasmid in the gDNA of the cell. Following the construction of a suitable suicide vector, it is commonly introduced from a donor cell to a recipient cell by conjugation, for gene deletion through homologous recombination. Alternatively, the introduction of the plasmid can be obtained by other means using transformation by electroporation or heat-shock (Bosma et al. (2015)).

Recombination using a selection marker

Homologous recombination (HR) occurs naturally in bacteria, and can be involved in DNA repair through the HDR system or in horizontal gene transfer (Bhatti et al. (2016); Kuzminov (1999); Ochman et al. (2000)). Since the occurrence of homologous recombination is rather low, transformation methods with high efficiency are needed as only a small percentage of the transformants will integrate the suicide vector into its chromosomal DNA. Alternatively, the expression of certain bacteriophage proteins, including those encoded by *Red* genes of the bacteriophage λ, have shown to increase the homologous recombination efficiency in bacteria, such as *E. coli*, by homologous recombination of double-stranded DNA (Liu (2003); Mosberg et al. (2010)). As a basis for targeted chromosomal gene knockouts and deletions in *B. methanolicus*, transfer of the conjugative plasmid pCasPP by conjugation into *B. methanolicus* has been shown to provide high transformation efficiencies, and its presence in the cells has been verified by PCR amplification using specific primers (Irla et al. (2019, unpublished)). The pCasPP plasmid hence serves as a good starting point for the construction of suicide vectors in the bacterium.

After the introduction of a suicide vector into the recipient cell, a homologous recombination (single crossover) event occurs between the (a) upstream (LF) or the (b) downstream (RF) homologous regions of the gene of interest (g.o.i.). This eventually results in the integration of the vector into the host chromosome. As this event also integrates the selection marker, supplementation of its corresponding antibiotic can be used to select clones where the HR has occurred. This event is displayed as the first step in Figure 1.2, and correspond to the initial step of an allelic exchange or gene knockout in the chromosome.

Recombination using a counterselection marker

In the second step, a second homologous recombination event (single crossover) is prompted by the supplementation of a counterselection compound. As Figure 1.2 shows, this recombination event results in either of two events; the integrated plasmid along with the gene of interest

is cured from the cell, or the cell restores its wildtype genotype. In the case of the latter, the gene of interest remains in the chromosomal DNA. Thus, a counterselection marker selects for away transformants that have undergone only a single homologous recombination event, and therefore have retained the counterselection marker in their chromosomal DNA. Using a toxic inducible counterselection system makes it possible to activate the second recombination event by the introduction of an inducer. As second recombination events may be of low frequency, the use of counterselection markers simplifies the selection of positive mutants (Reyrat et al. (1998)). Another advantage of using a counterselection marker is that curing of the integrated plasmid makes it possible to use the same selection marker and counterselection marker for sequential deletion of genes in the chromosomal DNA. This is preferred as the availability of suitable selection markers is often very limited, especially for those that are useable in thermophile bacteria (Drejer et al. (2018); Zeldes et al. (2015)). A system that does not allow for plasmid curing and removal of the selection marker means that multiple selection markers have to be used if multiple loci are targeted, hence limiting the system for further genomic editing applications (Zyl et al. (2019)).

There are many counterselection markers used for counterselection-based recombination for gene deletion. Some of these involve *sacB* and *lacY*. *SacB* encodes levansucrase, which converts sucrose to the toxic compound levrans, which results either in inhibition of growth or lysis of the cell (Gay et al. (1985)). *LacY* encodes a lactose permease, making the cell sensitive to the lactose analogue *O*-nitrophenyl-β-D-thiogalactoside (tONPG) (Berman and Beckwith (1979)). The choice of a suitable counterselection marker for a specific genus of bacteria depends on various aspects, economic and physical, concerning both the counterselection gene and its corresponding counterselection compound. Besides, a small counterselection marker gene is often beneficial for molecular cloning purposes (Barkan et al. (2011)). When considering the counterselection agent, its toxicity to humans and the concentration needed to induce the second homologous recombination event also must be considered.

1.3 Background for this work

1.3.1 Establishing a basis for a conjugative suicide vector

The plasmid used as a basis for the creation of suicide vectors in this study was the pCasPP plasmid, constructed by Rütering et al. (2017) for use in *Paenibacillus polymyxa* (*P. polymyxa*). pCasPP was found suited for conjugation into bacteria of the class *Bacilli*. Preliminary research has shown that the plasmid also is applicable for conjugation into *B. methanolicus* with higher transformation efficiencies than conventionally used electroporation (Irla, unpublished (2019)). Therefore, it was reasonable to predict that the use of suicide vectors for homologous recombination in *B. methanolicus* is possible.

The plasmid includes an *oriR* for replication in *E. coli*, kanamycin resistance gene (Kan^R) and the *repU* gene of the plasmid pUB110, originating from *Staphylococcus aureus* (Leclere et al. (2005)). The selection marker gene encodes a thermostable enzyme, and kanamycin was used as a thermostable antibiotic. Furthermore, the plasmid includes an *oriT* with a RP4 origin of transfer and *gapdh* promoter for constitutive sgRNA expression obtained from plasmid pCRISPomyces-2, and the *Cas9* encoding gene under the control of the *sgsE* promoter from *Geobacillus stearothermophilus* (Irla (Oral communication), Rütering et al. (2017)). To construct a basis plasmid for use in gene deletion, the backbone of pCasPP without the CRISPR/Cas9 cassette and *repU* gene encoding the plasmid replication protein was amplified by polymerase chain reaction (PCR). The counterselection gene *oroP* was then introduced into the plasmid backbone under the control of a xylose inducible promoter (P_{xylR}) (Irla, Oral communication).

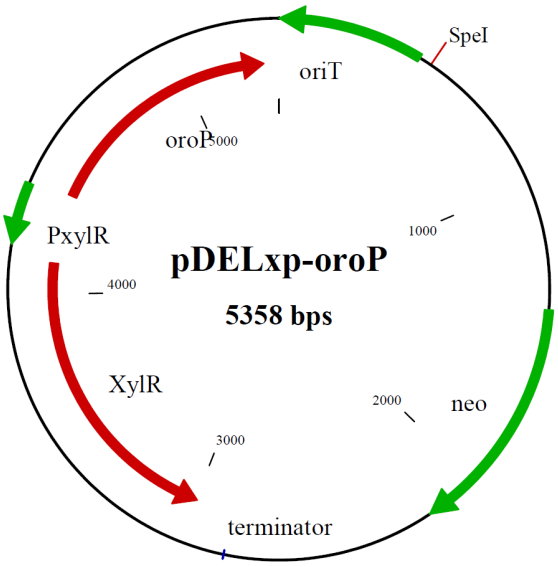


Figure 1.3: Plasmid pDELxp-*oroP* including a *oriT*, a selection marker for neomycin/kanamycin (*neo*), and *oroP* as a counterselection marker under the control of a xylose inducible promoter (P_{xylR}). The restriction site *Spe*I is used as cloning site for homologous regions flanking the gene to be deleted.

For the subsequent introduction of cloned homologous regions flanking the gene of interest, the restriction site *Spe*I was designated for the linearization of the plasmid. A figure of the resulting plasmid, named pDELxp-*oroP*, is displayed in Figure 1.3.

1.3.2 Investigation of counterselection markers

Ahead of this study, multiple counterselection markers for use in *B. methanolicus* were investigated by Leonie Benninghaus and Carsten Haupka in a joint study between NTNU and UniBi, under the supervision by Marta Irla (Benninghaus (2018b,a)). The investigated counterselection markers involved the *oroP* gene encoding an orotate transporter, *codA* and *codB* encoding a

cytosine permease and deaminase, *lacZ* encoding *β-galactosidase*, and *sacB* encoding levansucrase. A table summarizing the findings for the investigated counterselection markers is shown in Table 1.1.

Table 1.1: Counterselection markers previously investigated for construction of suicide vectors for targeted chromosomal gene knockouts in *B. methanolicus*.

Counterselection Marker	Gene product	Description/Usage
oroP	Orotate transporter	Transports toxic antimetabolite 5-FU into the cell
codA	Cytosine permease	Uptake of toxic cytosine analogue 5-fluorocytosine
codB	Cytosine deaminase	Metabolization of toxic cytosine analogue 5-fluorocytosine
lacZ	β-galactosidase	Cleaves lactose analogue X-Gal yielding a blue-colored dimer of substituted indoles
sacB	Levansucrase	Converts sucrose to toxic metabolite levans

The orotate transporter is necessary for the transport of orotate into the cell, which can be used as a source of pyrimidine and nitrogen utilization in the organism (Defoor et al. (2007); Milstein and Bekker (1976)). The introduction of the orotate analog and anti-metabolite 5-fluoroorotate (5-FO) is toxic to the cell (Defoor et al. (2007)). Growth on 5-FO can, therefore, be used as a measure of the activity of the gene, or as a counterselection marker providing selection pressure for the second recombination event previously described (Section 1.2.4). The cytosine permease and deaminase encoded by the genes *codA* and *codB*, are responsible for the uptake and utilization of cytosine in the cell. The introduction of a toxic anti-metabolite to cytosine, 5-fluorocytosine (5-FC), will result in conversion into the toxic metabolite 5-fluorouracil. In previous studies conducted in *Gluconobacter* subsp., the presence of both genes has shown to be necessary for efficient gene deletion by homologous recombination (Kostner et al. (2013)). The *sacB* gene derived from *Bacillus subtilis* (*B. subtilis*) encodes levansucrase, which catalyzes the reaction of hydrolysis of sucrose and the synthesis of levan, the latter being highly toxic to the cell, even in small amounts (Gay et al. (1985)). The fourth of the investigated counterselection markers, *lacZ*, originates from the *lac* operon present in *E. coli*. The gene encodes the enzyme β-galactosidase, which cleaves galactoside (lactose) to galactose and allolactose. This counterselection marker is widely used in what is called the ”blue-white screening technique´´. For this purpose, colonies containing the *lacZ* gene will turn blue in the presence of the colorless

counter-selective compound 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal). In contrast, colonies lacking the gene will be colorless, as X-Gals itself is colorless. The blue color results from enzymatic hydrolysis of X-Gal by β-galactosidase, which yields galactose and a blue-colored dimerized compound of the remaining substituted indole parts of X-Gal (Juers et al. (2012)). Furthermore, the X-Gal counterselection system has previously been successfully established for the bacterium *Bacillus smithii* (*B. smithii*) (Bosma et al. (2015).
Based on multiple factors such as toxicity of the counterselection compound to humans, the concentration needed for a growth-inhibiting effect, the price of the compound, and its gene size, the *oroP* counterselection marker was chosen as the most promising counterselection marker out of the investigated. This counterselection marker was found to have a half-maximal inhibitory concentration (IC_{50}) of $6.8\,\mu g/ml$ (Benninghaus (2018b,a)).

1.4 Salvage pathway of pyrimidines and *upp* importance

The importance of an extensive genetic toolbox in *B. methanolicus* can be used for the elucidation of metabolic pathways, which hitherto are not fully understood. One such pathway is the pyrimidine salvage pathway, which is vital for the re-use of pyrimidines in the bacterium.

1.4.1 Pyrimidine salvage pathway

Nucleotides are diverse molecules that constitute the basic components of essential processes and molecules in a living organism. They are precursors of the nucleic acids RNA and DNA, and important energy carriers in the cell, like adenosine triphosphate (ATP) and cofactors such as nicotinamide adenine dinucleotide (NAD) and flavin adenine dinucleotide (FAD) (Snustad and Simmons (2012); Nelson et al. (2017)). Therefore, nucleotides are utterly crucial for a cell's reproduction and growth. As a result, an organism depends on having synthetic pathways for nucleotide production to accommodate these needs. Nucleotide production is known to originate from two different types of pathways; *de novo* synthesis from nucleotide precursors, or salvage pathways where nucleic acids are degraded to their constituting components. The degradation of nucleotides in a cell by salvage pathways releases purines and pyrimidines, which can be used to synthesize new nucleotides in the cell (Nelson et al. (2017); Martinussen et al. (1994)). This "re-use " of pyrimidines is very beneficial for the organism, as the *de novo* biosynthesis of pyrimidines usually requires more energy, salvage pathways, thus being more energetically favorable (Moffatt and Ashihara (2002)).

Some microorganisms have been shown to be able to use exogenous sources of pyrimidines for nucleotide synthesis by salvage pathways (Martinussen et al. (1994); Turnbough and Switzer (2008)). Which of the pyrimidines a cell can take up and utilize depends on its metabolic pathways, and varies among different genus of microorganisms. For pyrimidine salvage pathways

in *B. subtilis*, exogenous uracil can be metabolized by the action of two genes; *upp* encoding uracil phosphoribosyltransferase (UPRTase) and *pyrR* encoding bifunctional protein PyrR (Martinussen et al. (1995)). The latter primarily functions as a repressor of the *pyr* operon, which is responsible for the biosynthesis of pyrimidines. The activity and binding on mRNA upstream of the genes in the *pyr* operon containing genes responsible for pyrimidine biosynthesis is regulated by intracellular pools of uridine monophosphate (UMP), causing attenuation of transcription when the intracellular concentration is exceeding a certain level (Turner et al. (1994); Swttzer et al. (1998)). The *pyrR* gene has also been shown to encode a phosphoribosyltransferase with a similar function as the *upp* encoding protein, albeit with lower activity. This is surprising, as the gene has been found to only pose low partial homology to other known phosphoribosyltransferases (Martinussen et al. (1995)).

Common for both *upp* and *pyrR* is their ability to convert uracil and 5-phosphoribosyl α-1-pyrophosphate (PRPP) into UMP and pyrophosphate by phosphoribosylation (Martinussen et al. (1995)). Exploration of pyrimidine salvage pathways in *Mycobacterium tuberculosis* (*M. tuberculosis*) and *Lactococcus lactis* (*L. lactis*) have also been shown to metabolize uracil by the action of *upp* and *pyrR*, indicating their importance for bacterial pyrimidine salvage in general (Singh et al. (2015); Martinussen et al. (1994, 1995)). Furthermore, enzyme assays in *B. subtilis* by Martinussen et al. (1995) have shown that a mutation in the *pyrR* gene resulted in only a minor change in enzymatic activity of the phosphoribosyltransferase activity, indicating that it only constitute a minor role in conversion of uracil relative to that of the *upp* gene. These results give it reason to believe that *upp* serves as one of the most important and essential participants for uracil metabolism in *B. subtilis*.

1.4.2 Alternative routes of pyrimidine metabolization

Later research efforts on pyrimidine salvage pathways in *M. tuberculosis*, *B. subtilis* and *L. lactis* have given indications of an alternative salvage pathway of pyrimidines (Singh et al. (2015); Martinussen et al. (1994, 1995)). The UPRTase has also displayed some affinity to the toxic uracil-analog 5-fluorouracil (5-FU). When metabolized, 5-FU works by inhibiting thymidylate synthase (TS). TS is necessary for the conversion of deoxyuridine monophosphate (dUMP) to thymidine monophosphate (dTMP), needed for the synthesis of DNA (Wigmore et al. (2010)). From growth studies in a *upp* deletion strain of *B. subtilis* on 5-FU by by Martinussen et al. (1995), resistance to concentrations up to $5\,\mathrm{mg/ml}$ 5-FU were observed. Higher concentrations were shown to be lethal and resulted in cell death. Consequently, it was proposed that uracil must be metabolized by either the catalytic activity of *pyrR* or other salvage pathways not formerly known to be present in *B. subtilis* (Martinussen et al. (1995)). Furthermore, a dual mutant strain of *upp* and *pyrR* showed that exogenous uracil nevertheless was incorporated into the cell, meaning that another pathway has to account for metabolization of uracil in the

organism. Studies of *L. lactis* and in *Lactobacillus plantarum* (*L. plantarum*) have shown that uracil also could be salvaged by the enzymatic catalysis of pyrimidine phosphorylase (*deoA*) and thymidine kinase, eventually forming dUMP (Martinussen and Hammer (1994); Arsène-Ploetze et al. (2006)). However, the presence of *deoA* or a gene serving a similar function has not been identified as present in *B. subtilis*.

1.5 Aims of the project

By making use of a new and efficient method for horizontal gene transfer through conjugation in *B. methanolicus*, this work aimed to establish a genetic tool for markerless gene deletions in the bacterium using suicide vectors. The suicide vectors are constructed to include a Kan^R selection marker, an *oroP* counterselection marker, and cloned upstream and downstream regions flanking the gene to be deleted. By the introduction of gene-specific suicide vector into a desired strain of *B. methanolicus*, this study intended to show that the deletion of a targeted gene can be made by the occurrence of two subsequent homologous recombination events. The method enables re-use of the selection marker and counterselection marker genes, thereby only requiring the cloning of homologous flanking regions to the targeted gene into the vector. Therefore, suicide vectors with different flanking regions were also made to investigate the smallest size needed for optimal recombination efficiency.

To display the full potential of the tool, three chromosomal genes were targeted for the construction of multiple mutant strains; *mtlD* (BMMGA3_01075) encoding mannitol 1-phosphate 5-dehydrogenase (EC 1.1.1.17), *upp* (BMMGA3_16035) encoding an uracil phosphoribosyltransferase (EC 2.4.2.9), and *aroE* (BMMGA3_12130) encoding a shikimate dehydrogenase (EC 1.1.1.25). The deletion of the respective genes was chosen as they are easy to characterize phenotypically, and the assumed importance in their respective metabolic pathways. For the deletion of *upp* and *mtlD*, the corresponding deletion strains were assayed by growth experiments and enzyme assays. Furthermore, by establishing a deletion strain, this study also aimed to investigate if the wild-type phenotype could be restored by plasmid-based complementation of a gene. A complementation study was assumed to provide information on whether a physiologically observed change in enzyme function was resulting from the deletion itself or by disruption of regions upstream and downstream to the gene.

Moreover, to put the observed enzyme function into context, another aim was to investigate the deleted gene's metabolic context using bioinformatics approaches. This methodology intended to provide an example of how a combined *in silico* and *in vitro* approach can be used to investigate gene function in *B. methanolicus*, hence emphasizing the importance of the established tool.

Chapter 2

Materials and methods

2.1 Growth media and solutions

The composition of growth media and solutions used in the present study are listed in Appendix A.

2.2 Bacterial strains, plasmids, oligonucleotides and growth conditions

The bacterial strains and plasmids used in this study are presented in Table 2.1 and Table 2.2, respectively. DH5α (Stratagene) was used as a general cloning host for the construction of recombinant plasmids. Following construction, the correctness of the inserts was confirmed by DNA sequencing. Plasmids were then introduced into the donor strain *E. coli* S17-1 strain (ATCC 47055) for subsequent conjugation. In all cases, *B. methanolicus* (ATCC 51375) was used as the recipient cell for conjugation and expression host. Oligonucleotides used in the present study are listed in Appendix B.

The cultivation of *E. coli* strains were conducted in liquid LB-medium (225 rpm) or LB-medium supplemented with agar (LA-plates) at $37\,^{\circ}\mathrm{C}$, unless stated otherwise. Strains of *B. methanolicus* MGA3 were, on the other hand, cultivated at $50\,^{\circ}\mathrm{C}$ and 200 rpm in liquid SOB or MVcMY minimal medium with 200 mM methanol supplemented ($\mathrm{MVcMY}_{200\mathrm{mM\ (MeOH)}}$), or SOB agar plates, using a New Brunswick™ Innova 42R shaker incubator (Eppendorf). Preparations of electrocompetent *B. methanolicus* was made using SOB medium supplemented with sucrose (SOBsuc). This preparation could also be made using SOB medium, as sucrose does not pose any additional positive effects.

An appropriate antibiotic was supplemented when needed to ensure the retainment of plasmids harboring an antibiotic marker gene and for positive selection. For *E. coli*, a concentration of

$50\,\mu g/ml$ kanamycin (Kan) or $25\,\mu g/ml$ chloramphenicol (Cm) was used. For the selection of positive transformants after conjugation between *E. coli* S17-1 carrying vector pDELxp-*oroP* and *B. methanolicus* MGA3, a final concentration of $25\,\mu g/ml$ Kan was used. To apply selection pressure for the second homologous recombination event *B. methanolicus* MGA3 was plated out on SOB plates containing $80\,\mu g/ml$ 5-fluoroorotate (5-FO) and 1% xylose. Cultivations of strains harboring plasmid pTH1mp and its respective derivates were supplemented with $5\,\mu g/ml$ Cm for *B. methanolicus* and $30\,\mu g/ml$ Cm for *E. coli*. The progress of growth of bacteria in liquid cultures was monitored by measuring the optical density (OD) at $600\,nm$ (OD_{600}), using Biowave Cell Density Meter CO8000 (BIOCHROM US BE).

Table 2.1: Bacterial strains used in the present study.

Strain	Description	Reference
E. coli DH5α	General cloning host	Stratagene
E. coli S17-1	*recA*, RP4 origin of transfer, chromosomally integrated RP4::Mu. Donor strain for conjugation of suicide vectors	ATCC 47055
B. methanolicus MGA3	Wild-type strain.	ATCC 53907
B. methanolicus MGA3 Δ*upp*	MGA3 strain with deletion of *upp* gene	Heid (2020), unpublished
B. methanolicus MGA3 Δ*mtlD*	MGA3 strain with deletion of *mtlD* gene	This study
B. methanolicus MGA3 Δ*upp*Δ*mtlD*	MGA3 strain with deletion of the genes *upp* and *mtlD*	This study

Table 2.2: Plasmids used in the present study.

Plasmid	Description	Reference
pTH1mp	Cm^R; derivate of pTH1mp-*lysC* for gene expression under control of the *mdh* promoter. The *lysC* gene was replaced with multiple cloning sites. Shuttle vector in *B. methanolicus* and *E. coli*.	Irla et al. (2016)
pTH1mp-*upp*	Cm^R; pTH1mp derivate for expression of *upp* from *B. methanolicus* MGA3 under control of the *mdh* promoter.	This study
pCasPP	Neo^R; *P. polymyxa* genome editing vector, derivative of puB110 (ATCC 37015).	Rütering et al. (2017)
pDELxp-*oroP*	Kan^R, *B. methanolicus* genome editing vector, derivative of pCasPP, *oroP* counterselection marker under regulation of P_{xyl}.	Irla, unpublished
pDELxp-*oroP*-Δ*mtlD*	Kan^R; pDELxp-*oroP* derivate. Contains 1000 bps homologous flanking upstream and downstream regions to chromosomal *mtlD* gene of *B. methanolicus.*	Heid (2020), unpublished
pDELxp-*oroP*-Δ*upp*	Kan^R; pDELxp-*oroP* derivate. Contains 1000 bps homologous flanking upstream and downstream regions to chromosomal *upp* gene of *B. methanolicus.*	Heid (2020), unpublished
pDELxp-*oroP*-Δ*aroE*250	Kan^R; pDELxp-*oroP* derivate. Contains 250 bps homologous flanking upstream and downstream regions to chromosomal *aroE* gene of *B. methanolicus.*	This study
pDELxp-*oroP*-Δ*aroE*500	Kan^R; pDELxp-*oroP* derivate. Contains 500 bps homologous flanking upstream and downstream regions to chromosomal *aroE* gene of *B. methanolicus.*	This study
pDELxp-*oroP*-Δ*aroE*1000	Kan^R; pDELxp-*oroP* derivate. Contains 1000 bps homologous flanking upstream and downstream regions to chromosomal *aroE* gene of *B. methanolicus.*	This study

2.3 Storing of bacterial strains

Glycerol stocks were made of all bacterial strains used in the present study when stored over an extended period. Initially, overnight cultures of colonies inoculated from agar plates were aliquoted and mixed in cryo-tubes with glycerol to a final concentration of 25-30% (vol/vol). LB medium and $MVcMY_{200mM\ (MeOH)}$ supplemented with suitable antibiotics when needed, were used for overnight cultures of *E. coli* and *B. methanolicus* strains, respectively. Once mixed with glycerol, the strains were stored at -80 °C for later use. For short-term storage, colonies of *E. coli* on agar plates were stored at 4 °C, while colonies of *B. methanolicus*, on the other hand, were stored at room temperature (25 °C). Plates were sealed tight to prevent them from drying out. After 7-10 days of storage, colonies were regularly picked and transferred to new fresh agar plates for prolonged storage.

2.4 Isolation of genomic DNA from crude cell extract

After introducing alterations in the chromosomal DNA, a rapid yet efficient method to harvest genomic DNA (gDNA) from cell material is needed for screening purposes and DNA sequencing.

The procedure of isolating gDNA from colonies, was in this study based on a protocol by Lõoke et al. (2011), albeit with some modifications. In the first step, a colony was picked from an agar plate, or 150 µl of a liquid overnight cell culture harvested and resuspended in 100 µl of a 200 mM lithium acetate (LiOAc) containing 1% sodium dodecyl sulfate (SDS). The mixture was then incubated on a heating block at 70 °C for 5 minutes for the lysis of the cells to occur. After incubation, 300 µl of a 96% ethanol solution was added, and the mixture gently mixed by vortexing. This was followed by centrifugation (15000 rpm, 3 minutes), and the supernatant discarded. The remaining DNA was washed using 5 µl of 70% ethanol and dried (70 °C, 10 minutes). Finally, the DNA was resuspended in 50-60 µl of sterile H_2O, and the solution centrifuged (15000 rpm, 60 seconds). 1 µl of the supernatant containing the gDNA was then used as a template for PCR-based applications.

2.5 Measuring DNA concentrations

In molecular biology, small volumes in the nano or micro quantity are commonly used. Thus, differences in the concentration of DNA can hence inflict significant differences in the amount of the applied DNA. Besides, many established protocols used in the present study, such as methods of transformation, Gibson assembly, and polymerase chain reactions, require specific amounts of DNA for optimal efficiency. Therefore, when needed, the concentration of DNA

samples was measured spectrophotometrically using NanoDrop® One/One^C Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific).

The measurements were performed by applying 1-2 µl of the respective sample to the sample pedestal, and by comparing it to a blank (reference) sample. The spectrophotometer then applies UV light through the samples with wavelengths of 230, 260, and 280 nm. DNA concentrations were measured at 260 nm. The calculated ratios of 260/280 nm and 260/230 nm were used to determine the presence of proteins and other contaminants in the samples, besides providing data on the purity of the DNA itself (ThermoFisher (2017)). The measured light intensities of the blank sample and the DNA sample are then used to calculate the absorbance of the DNA sample, which can be directly associated with the DNA concentration using Beer's law (Appendix F). The sample absorbance is calculated as the negative inversely logarithm of the ratio between the sample intensities of the DNA sample on the blank sample (ThermoFisher (2017)).

2.6 Molecular cloning techniques

Molecular cloning constitutes a set of techniques for heterologous gene expression of recombined DNA (Snustad and Simmons (2012)). A standard method for constructing recombinant DNA is based on the cloning of one or more DNA fragments of interest, such as a gene, into plasmids. The specific fragments used for cloning purposes are commonly generated by polymerase chain reaction amplification, using the genome as template.

Centrifugation of sample volumes above $1.5\,\mathrm{ml}$ were carried out using Centrifuge 5430 R (Eppendorf), while Centrifuge 5424 (Eppendorf) were used for smaller sample volumes.

2.6.1 Polymerase chain reaction

Polymerase chain reaction (PCR) is a useful tool in molecular genetics for amplifying DNA *in vitro*. This method uses synthetic and specific primers together with a thermostable DNA polymerase in a mixture of reaction buffer and deoxynucleotide triphosphates (dNTPs). The execution of PCR was done using a Mastercycler Nexus X2 (Eppendorf) or C1000 Touch Thermal Cycler (Bio-Rad).

Two DNA polymerases were used in PCR amplifications of the present study; CloneAmp™ HiFi DNA polymerase (Takara Bio) and GoTaq® DNA polymerase (Promega). CloneAmp™ DNA polymerase is a high fidelity enzyme that was primarily used for the cloning of genes and homologous flanking regions of the genes targeted for deletion, besides the screening of putative mutants of *B. methanolicus* MGA3 when necessary. GoTaq® DNA polymerase was exclusively used for the screening of clones and mutants of *E. coli* DH5α, *E. coli* S17-1 and

B. methanolicus MGA3 by colony PCR. In cases of difficulty using GoTaq® DNA polymerase for colony PCR, CloneAmp™ DNA polymerase was instead used for screening of clones and mutants. Furthermore, PCR was also used for the amplification of templates used for DNA sequencing.

The mixture compositions for CloneAmp™ DNA polymerase (25 µl reaction) and GoTaq® DNA polymerase (12.5 µl reaction) are listed in Table 2.3 and Table 2.4, respectively. The reaction mixture used for PCR with GoTaq® DNA polymerase consists of a green loading dye, making it possible to apply samples directly to an agarose gel for gel electrophoresis after PCR amplification.

Table 2.3: General reaction mixture for PCR using CloneAmp™ DNA polymerase (Takara).

Component	Volume
CloneAmp™ Hifi PCR Premix	12.5 µl
Forward primer (10 µM)	0.5 µl
Reverse primer (10 µM)	0.5 µl
DNA template	< 100 ng
Sterile deionized water	To 25 µl

Table 2.4: Reaction mixture for PCR using GoTaq® DNA Polymerase (Promega). Template DNA could also be colonies from plate directly transferred to the reaction mixture without gDNA isolation.

Component	Volume
5X Green GoTaq® Reaction Buffer	2.5 µl
PCR nucleotide mix	0.25 µl
Forward primer (10 µM)	0.63 µl
Reverse primer (10 µM)	0.63 µl
GoTaq® DNA polymerase	0.06 µl
Template DNA	< 0.125 µg
Nuclease-free water	To 12.5 µl

The PCR reaction is divided into three main steps: denaturation, annealing, and elongation. During denaturation, the hydrogen bonds between the two complementary DNA strands are broken, yielding two single-stranded DNA molecules. The initial denaturation step is often longer than the following one as the whole initial DNA template is separated into two single DNA strands. In contrast, subsequent denaturation steps will only separate the newly synthesized and smaller double-stranded DNA templates. In the next step, annealing, specific primers hybridize with their complementary parts of the single-stranded DNA molecules. The annealing

temperature varies by the set of primers used for the amplification. Routinely, for CloneAmp™ DNA polymerase the annealing temperature was set to $55\,^{\circ}\mathrm{C}$. In other cases, and when $55\,^{\circ}\mathrm{C}$ was found unsuited for the PCR using CloneAmp™ DNA polymerase, the optimal annealing temperatures were set to $3\,^{\circ}\mathrm{C}$ lower than the melting temperatures of the primers used. When found appropriate, the optimal annealing temperatures were found by gradient PCR and gel electrophoresis. The annealing temperature providing the highest concentration of a PCR product, shown by its fluorescence intensity of an expected band on the gel staining, was used for subsequent PCR reactions. In the final step, elongation occurs as the DNA polymerase binds to the primers and synthesize new complementary DNA, yielding two identical double-stranded DNA molecules out of the original one. This three-step process is then repeated for 30-35 cycles, followed by a final elongation step for 5 minutes. This last step ensures that all DNA is present as double-stranded molecules.

In the case of colony PCR, cells were preferably picked and resuspended in the PCR mixture while extending the initial denaturation step of the PCR reactions to 10 minutes for the cells to lysis. Alternatively, DNA was first isolated from the cells, see Section 2.9, before being used as a template in the PCR mix. When DNA was used as a template, the initial denaturation step used was reduced to 10 seconds.

2.6.2 Digestion of DNA by restriction enzymes

For the construction of recombinant plasmids, a vector has to be linearized by cutting at specific restriction sites. Usually, a plasmid used for cloning and heterologous gene expression consists of short DNA segments with multiple cloning sites (MCS). Thus, the linearization of the desired plasmid using restriction endonucleases (restriction enzymes) is a commonly used procedure. Restriction endonucleases enzymes that recognize and cut at specific restriction sites, usually independently of other restriction enzymes. Hence, multiple restriction enzymes can often simultaneously be used if they share the same activation temperature and reaction buffer. Digestion by restriction enzymes can result in blunt or staggered cuts. As DNA is complementary and anti-parallel, staggered cuts at such sites will give sticky ends providing overhangs useful when ligating multiple overlapping fragments together. Using multiple restriction enzymes simultaneously can be preferable to reduce the effect of re-ligation between the ends of the linearized plasmid. Under conditions deviating from optimal, restriction enzymes can, in some cases, pose a phenomenon known as star-activity, referring to an alteration in the enzyme specificity (NEB (No date)).

All reagents and enzymes used for restriction digestion in the current study were provided by New England Biolabs (NEB). The reaction mixture was made according to the online tool NEBcloner (NEB). A general mixture for digestion by restriction enzymes consisted of a template DNA, a 10X CutSmart® Buffer, and one or two restriction enzymes diluted with nuclease-free

water to a final reaction volume of 50 µl (Table 2.5). The digestion-mixture was incubated at the optimal activation temperature provided by NEBCloner for 1 hour. Were the enzymes have been shown not to confer any star-activity for the optimal set-up, the mixture was incubated overnight.

Table 2.5: Reaction mixture for digestion of DNA by two restriction enzymes in a total reaction volume of 50 µl.

Component	Volume
DNA	1 µg
10X CutSmart® Buffer	5 µl
Restriction enzyme 1	1 µl
Restriction enzyme 2	1 µl
Nuclease-free water	to 50 µl

As part of this study, the linearization of plasmid pDELxp-*oroP* was found suited using the restriction enzyme *Spe*I. The plasmid pTH1mp was, on the other hand, linearized using restriction enzymes *Pci*I and *Bam*HI. All three enzymes produce sticky ends, as displayed in Figure 2.1.

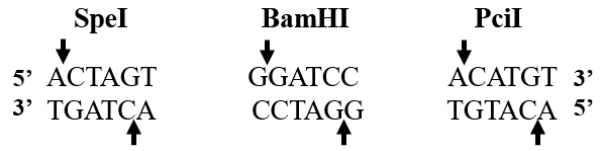


Figure 2.1: Cut sites of restriction enzymes *Spe*I, *Pci*I, and *Bam*HI. The direction of the DNA strands are indicated by 3'- and 5'-ends. All three enzymes results in sticky ends.

2.6.3 Gel electrophoresis

In molecular biology, the possibility of effectively separating and identifying the approximate size of DNA fragments is a valuable tool. The method is based on the physiology of the DNA itself, being evenly distributed by negative charges along the two complementary strands. When applying a current over a stained agarose gel containing a DNA sample, the DNA will begin migrating from the negative towards the positive pole. Due to pore sizes of the gel, the DNA fragments will be separated based on their length; the smallest ones travel the farthest distance. By adding a sample containing multiple DNA fragments of known sizes, namely a DNA ladder, the size of DNA fragments present can be estimated.

In the current study, gels were made of 0.8% (w/v) agarose in Tris-Acetate-EDTA (TAE) buffer containing a nucleic acid fluorescent dye; either 20 µl GelRed (Biotium) or 20 µl GelGreen (Biotium) per 400 ml of agarose-TAE solution. By applying UV light to the DNA bound to the fluorescent particles makes it possible to visualize both the relative amounts of the DNA

fragments and their sizes in the gel. As DNA ladder, the GeneRuler™ 1 kb Plus DNA Ladder (Thermo Fisher Scientific) was used, shown in Appendix C. In all cases, except for when GoTaq® DNA Polymerase (Promega) was used, Loading dye (NEB) was added to the samples before the gel electrophoresis. All gels were carried out at 100 V for 35-45 minutes in TAE buffer and power-supplied by PowerPac^{TM} Basic Power Supply (Bio-Rad). For subsequent imaging of the gel stainings, the Molecular Imager ChemiDoc™ XRS+ (Bio-Rad) and the compatible ImageLab™ software (Bio-Rad) was used.

2.7 Gibson assembly for construction of suicide vectors and complimenting plasmids

The construction of suicide vectors for the deletion of chromosomal genes and plasmids for complementation analysis of *B. methanolicus* MGA3 was carried out using Gibson assembly. Gibson assembly is a rapid and efficient method for joining multiple DNA fragments with complementary and overlapping ends in one isothermal step. The composition of the reaction mixture (Table 2.6) and the protocol for the assembly of DNA fragments was done accordingly to a protocol described by Gibson et al. (2009).

For the assembly, an aliquot of $15\,\mu l$ Gibson assembly Master Mix (1.5X) was used. The master mix contains a reaction buffer, a 5'-exonuclease, a DNA polymerase, and a DNA ligase. In the first step, the 5'-exonuclease degrades the DNA strands at their 5'-end, making a single-stranded overlap between two complementary DNA sequences. An elevation in temperature then ensures optimal activities of the involved enzymes. Upon annealing, the DNA polymerase synthesizes the remaining nucleotides in 5' $\rightarrow$ 3' direction, before the DNA ligase seals the nicks between the synthesized DNA and the original DNA fragments (Gibson et al. (2009)).

The preparation of the mixture was made in the following order. Firstly, the isothermal reaction buffer (IRB) was thawed on ice. The 5'-exonuclease, DNA polymerase, and DNA ligase were then added to the mixture from a freezer block and diluted with pre-chilled MQ-water. The solution was then aliquoted into tubes containing $12.5\,\mu l$ each, and stored at -$20\,^{\circ}C$ for later use. For the assembly of multiple DNA fragments, the master mix only required the adding of a linearized vector and the DNA fragments to be inserted. For the reactions to occur, the samples were incubated in a thermocycler at $50\,^{\circ}C$ for 60 minutes, when 2-3 fragments were assembled.

Table 2.6: Gibson assembly Master Mix (1.5X) stock solution including a 5'- exonuclease, DNA polymerase, and a DNA ligase, suspended in a solution of reaction buffer and deionized water.

Component	Volume
Isothermal reaction buffer (IRB)	320 µl
T5 exonuclease	0.54 µl
Phusion DNA polymerase	20 µl
Taq DNA ligase	160 µl
Deionized water	700 µl

2.8 Preparation of competent cells and transformation

The preparation of chemically competent cells responsive to the introduction of extracellular DNA by heat-shock transformation was made similarly for both strains of *E. coli* DH5α and S17-1, based on a protocol by Green and Rogers (2013). The preparation of electrocompetent cells of *B. methanolicus* MGA3 was conducted in another way, as the transformation of exogenous DNA was carried out by electroporation according to a protocol by Jakobsen et al. (2006).

Preparation of chemically competent *E. coli* cells

A pre-culture of the desired *E. coli* strain was prepared in 10 ml of Psi medium and the culture cultivated over-night (37 °C, 225 rpm). The next day, 1 ml of the pre-culture was reinoculated into 100 ml of fresh Psi medium and cultivated (37 °C, 225 rpm) for approximately 2.5 hours to an OD_{600} between 0.4-0.43. When the desired OD_{600} was reached, the cells were put on ice for 15 minutes. Following incubation, the cells were transferred to a pre-chilled sterile centrifugation tube (50 ml), and the tubes centrifuged (5000 g, 5 minutes, 4 °C). The supernatant was then discarded, and the cells gently resuspended in 40 ml of TFB1 (2x20 ml for each tube) and incubated on ice for 5 minutes. The mixture was then centrifuged (4000 rpm, 5 minutes, 4 °C), the supernatant discarded, and the remaining cell pellet gently resuspended in 6 ml of TFB2. The resuspended cell solution was then aliquoted into Eppendorf tubes (100 µl each), and flash-frozen using liquid N_2. The tubes of competent cells were then stored at −80 °C for later use.

Transformation of chemically competent *E. coli* by heat shock

Chemically competent *E. coli* cells transformed by heat shock transformation. For each transformation, a vial of the desired chemically competent *E. coli* cells were thawed on ice. Then, 10 pg - 100 ng of plasmid DNA or 2 µl of a Gibson reaction mix was added to each vial and gently mixed with the cells by pipetting. The mixture was placed on ice and incubated for 30 minutes.

Following incubation, the cells were heat-shocked for 45 seconds at $42\,°C$ and immediately after transferred to ice for two additional minutes. $1\,ml$ of LB medium was then added to each of the vials, and the cells incubated ($37\,°C$, 225 rpm) for 1 hour. After incubation, $100\,µl$ of the cells were plated out on pre-warmed ($37\,°C$) selection agar plates of LA supplemented with a suited antibiotic. The remaining culture was also spin down and plated similarly. Plates were incubated overnight at $37\,°C$. Colonies were then picked for screening by colony PCR, and positive clones inoculated to LB supplemented with suitable antibiotics. Plasmids were then isolated, and unconfirmed inserts sent for DNA sequencing. Finally, glycerol stocks were made of the confirmed positive clones.

Preparation of electrocompetent *B. methanolicus*

In the first step for the preparation of electrocompetent *B. methanolicus*, $25\,mL$ of SOB medium was aliquoted into a baffled flask and pre-heated to $50\,°C$. The pre-warmed medium was inoculated with a chosen *B. methanolicus* strain from a frozen glycerol stock and grown over-night for approximately 16 hours at $50\,°C$ and 200 rpm. The following day, reinoculation of the culture into fresh SOB medium was carried out to an initial OD_{600} of 0.1 was measured and incubated ($50\,°C$, 200 rpm). After 3-4 hours of growth, the cultures were reinoculated to 4 x $100\,mL$ of pre-warmed SOB medium to OD_{600}=0.05 and incubated under the same growth conditions to a measured OD_{600} of 0.18–0.30. When the expected OD_{600} was reached, the cultures were transferred to $50\,mL$ falcon tubes, and the cells harvested by centrifugation (10 min, 5000 rpm, $25\,°C$). Cell pellets from each culture were then resuspended in $4.5\,mL$ electroporation buffer (EPB), and the cell suspensions from two tubes combined. The washing step was afterward repeated using $9\,mL$ of EPB. Following centrifugation, supernatants were discarded, and the cells resuspended in the remaining media. Suspensions were then aliquoted into sterile Eppendorf tubes ($100\,µl$) and stored at $-80\,°C$ until further use.

Transformation of electrocompetent *B. methanolicus* by electroporation

A vial of $100\,µl$ electrocompetent cells of the desired *B. methanolicus* MGA3 strain was thawed on ice and mixed with $1\,µg$ of plasmid DNA in an Eppendorf tube. The mixture was then transferred to a pre-cooled electroporation cuvette (0.2 cm gap, Bio-Rad) and incubated on ice for 30 minutes. For each transformation, $12.5\,ml$ SOB medium was prepared in $125\,ml$ flasks and pre-warmed to $50\,°C$. A suitable antibiotic was supplemented if the strain harbored a plasmid with a selection marker. Cells were then electroporated ($200\,\Omega$, $25\,µF$, $12.5\,kV/cm$) using Gene Pulser Electroporation System (Bio-Rad). Cells were then immediately resuspended in $1\,ml$ of pre-warmed SOB medium ($50\,°C$) and transferred to its corresponding culture flask. Following inoculation, flasks were cultivated ($50\,°C$, 200 rpm) for additional 6 hours, before $1\,ml$ and $10\,ml$ of the cell cultures harvested by centrifugation (4000 rpm, 5 min, $40\,°C$). The supernatants were discarded. Cell pellets were resuspended in the remaining medium and plated out on pre-warmed ($50\,°C$) SOB plates supplemented with suitable antibiotics and cultivated

overnight ($50\,^{\circ}\mathrm{C}$).

By the appearance of colonies, five distinct colonies on each plate were transferred to an overnight culture of $25\,\mathrm{ml}$ $\mathrm{MVcMY}_{200\mathrm{mM\ (MeOH)}}$ medium with a suitable antibiotic (3x$25\,\mathrm{ml}$ and 2x$12.5\,\mathrm{ml}$) and cultivated ($50\,^{\circ}\mathrm{C}$, 200 rpm) until OD_{600}=1. At the same time, screening of the respective colonies was carried out using a specific set of primers. Finally, glycerol stocks of positive transformants were made from the overnight cultures and stored for later use at $-80\,^{\circ}\mathrm{C}$.

2.9 Purification of PCR products and restriction-digestion products, and plasmid preparation

After PCR amplification and digestion with restriction enzymes, the DNA was purified and concentrated before further use. Common for these methods is based on purifying DNA by binding it to a membrane, with following washing steps to remove unwanted contaminants. Purification of PCR products and vectors was done using DNA Clean and Concentrator™ - 5 (Zymo Research, Cat no. D4004). Plasmid DNA was extracted from overnight cultures and purified using Wizard® Plus SV Miniprep DNA Purification System (Promega, Cat no. A1460). When found suited, DNA fragments were also cut out from a gel stainings containing GelGreen (Biotium), isolated and purified using QIAquick® Gel Extraction Kit (Qiagen, Cat no. 28704 & 28706), according to the supplied protocol from Qiagen. All purified DNA samples were eluted in distilled $\mathrm{H_2O}$ and stored at $-20\,^{\circ}\mathrm{C}$. In all cases unless stated otherwise, the samples were centrifuged at 13,000 rpm.

2.9.1 DNA cleaning and concentration

The sample DNA was transferred to a $1.5\,\mathrm{ml}$ microcentrifuge tube and 2-5 volume-equivalents of DNA Binding Buffer was added, according to Table 2.7. The solution was briefly mixed by vortexing and transferred to a Zymo-Spin Column in a Collection Tube. Binding of DNA to the membrane was made by centrifugation (30 seconds). The flow-through was then discarded, and the DNA washed by the addition of $200\,\mu\mathrm{l}$ DNA Wash Buffer to the column, followed by centrifugation (30 seconds). The wash step was repeated once more before the DNA was eluted by transferring the Zymo-Spin Column to a $1.5\,\mathrm{ml}$ microcentrifuge tube with the addition of $14\,\mu\mathrm{l}$ of DNA Elution Buffer and centrifuged (30 seconds). For improved yield, the solution was incubated at $50\,^{\circ}\mathrm{C}$ for 2 minutes before centrifugation.

Table 2.7: Amount of DNA Binding Buffer volume-equivalents per unit of volume of the sample DNA.

Sample material	DNA Binding Buffer : Sample Volume
Plasmid DNA	2:1
PCR product, DNA fragment	5:1

2.9.2 Plasmid isolation and purification

$10\,\mathrm{ml}$ of an overnight culture ($15\,\mathrm{ml}$) of the bacterial strain with the desired plasmid was pelleted by centrifugation (7000 rpm, 5 minutes). The supernatant was then discarded, and the cells resuspended in $250\,\mu\mathrm{l}$ of Cell Resuspension Solution. Cells were lysed by the addition of $250\,\mu\mathrm{l}$ Cell Lysis Buffer and briefly mixed by inverting the tube four times. Following lysis, $10\,\mu\mathrm{l}$ of Alkaline Protease Solution was added to degrade proteins, such as nucleases, that might degrade the extracted DNA. The solution was again mixed by inverting the tube four times, with subsequent incubation at room temperature for 5 minutes. After incubation, $350\,\mu\mathrm{l}$ of Neutralization Solution was added, and the solution inverted four times to mix. Binding of plasmid DNA performed by centrifugation (1 minute) after transferring the cleared lysate to a Spin Column in a Collection Tube. The flow-through from the Collection Tube was discarded. Washing of Plasmid DNA was made by subsequently adding $750\,\mu\mathrm{l}$ and $250\,\mu\mathrm{l}$ of Wash Solution before centrifuging the solution for 1 and 2 minutes, respectively. The resulting flow-through was then discarded from the Collection Tube, and the Spin Column transferred to a $1.5\,\mathrm{ml}$ microcentrifuge tube.

Elution of plasmid DNA was in final conducted by adding $60\,\mu\mathrm{l}$ of sterile nuclease-free water to the Spin Column and incubated ($50\,^{\circ}\mathrm{C}$, 2 minutes). The plasmid DNA was in final collected by centrifugation (1 minute), and stored at -20 $^{\circ}\mathrm{C}$ for later use.

2.9.3 DNA extraction from agarose gel

Following gel electrophoresis, DNA fragments were excised from the agarose gel and resuspended in three volume equivalents of Buffer QC and incubated ($50\,^{\circ}\mathrm{C}$, 10 minutes) for dissolving the gel. Then $10\,\mu\mathrm{l}$ of 3M sodium acetate (pH 5) and isopropanol corresponding to 1 volume-equivalent of the gel was added to the sample mixture. Binding of DNA was performed by transferring the solution to a QIAquick spin column in a collection tube and centrifuge (1 minute). Flow-through was discarded, and the spin column reassembled with the collection tube. The DNA was washed in a two-step process. Firstly by adding $500\,\mu\mathrm{l}$ of Buffer QC and centrifugation (1 minute). Secondly, by adding of $750\,\mu\mathrm{l}$ of Buffer PE, incubated at room temperature (4 minutes) and centrifuged (1 minute). Flow-throughs were discarded. Elution

of membrane-bound DNA was made by transferring the spin column to a $1.5\,\mathrm{ml}$ centrifugation tube (Eppendorf) and applying $30\,\mu\mathrm{l}$ of deionized water to the membrane. The tube was then incubated (2 minutes, $50\,^{\circ}\mathrm{C}$), and DNA eluted by centrifugation (1 minute). The concentration of eluted DNA was then measured (Section 2.5), and the DNA stored at $4\,^{\circ}\mathrm{C}$.

2.10 DNA sequencing and multi-alignment analysis

DNA sequencing

Inserts of constructs made by Gibson Assembly were DNA sequence to verify their presence and correctness. PCR amplicons of confirmations of deleted genes in chromosomal DNA were also sequenced. All DNA sequencing samples were mixed according to the sample requirements for "LightRun Tubes" (Eurofins Genomics and GATC), with each sample tube consisting of a primer (forward or reverse), besides the DNA to be sequenced. The samples were then analyzed by Eurofins Genomics and GATC.

For analyzing the samples, the Sanger sequencing method modified using a thermostable DNA polymerase was used. The method is based upon similar principles as for PCR, albeit without amplification of the DNA as only one primer is used for each reaction. In addition to a general PCR reaction mixture (see Section 2.6.1), a small amount of 2',3'-dideoxynucleoside triphosphate chain-terminators (ddNTPs) of all four DNA bases are also added, each labeled with a specific fluorescent dye. In the extension phase, binding of a ddNTP will terminate the further extension of the template by the DNA polymerase as no 3'-OH end is available. After conducting the DNA polymerization reaction in multiple cycles, multiple synthesized chains of different lengths are made, each labeled with a specific fluorescent dye. Finally, the nucleotide sequence of the amplified DNA can be obtained by separating the fragments with polyacrylamide capillary gel electrophoresis based on their chain length and monitoring the detected fluorescence by subjection to a laser scanner (Eurofins Genomics, Snustad and Simmons (2012)).

Multi-alignment of DNA sequences

DNA sequencing results were analyzed by multi-alignment analysis of the sequences towards a template sequence in *Benchling* [Biology Software, 2020, Retrieved from `https://benchling.com`], to determine their correctness. Template sequences were made by cloning *in silico* using CloneManager [Bioinformatics tool, Sci-Ed] and corresponded to the desired DNA sequence. A multi-alignment analysis was used for comparing a query sequence to a template sequence to determine the degree of correctness and possible mutations present. The alignments were visualized in Benchling, which displays red-colored bases when a mismatch is present. Promising sequences were also multi-aligned using CloneManager [Bioinformatics tool, Sci-Ed], as this

tool was found more useful for detecting possible mutations in the sequenced DNA. Furthermore, all multi-alignments presented in Appendix E of this thesis were exported from CloneManager.

2.11 Conjugation and homologous recombination events

Gene deletions in *B. methanolicus* MGA3 strains were obtained by first constructing a suicide vector containing homologous regions upstream and downstream to the g.o.i. The recipient cell, *B. methanolicus* MGA3, was then subjected for conjugation of the vector with a donor strain of *E. coli* S17-1 harboring the suicide plasmid.

The process of conjugation and homologous recombination between the introduced suicide vector and the gDNA was based on Rütering et al. (2017). In the following, a modified procedure is presented, including the following steps:

- Pre-cultures ($25\,\mathrm{ml}$) of the donor and recipient strains were prepared by the inoculating desired strains of *B. methanolicus* and *E. coli* S17-1 in SOB medium and LB-medium, respectively. A suitable antibiotic was supplemented in order to conserve the suicide vector.
- The following day, pre-cultures were sub-cultured 1:100 into $50\,\mathrm{ml}$ of fresh media and incubated 4 hours until the early exponential phase of growth.
- After cultivation, $9\,\mathrm{ml}$ and $4.5\,\mathrm{ml}$ of the recipient strain were mixed together with $3\,\mathrm{ml}$ and $1.5\,\mathrm{ml}$ of the donor strain, respectively.
- Cells were then harvested by centrifugation (8,000 x g), the supernatant discarded, and the cells dropped on non-selective SOB plates and incubated overnight ($40\,^{\circ}\mathrm{C}$).
- The next day, cells from each plate were collected using sterile pipette tips, and resuspended each in $500\,\mu\mathrm{l}$ of 0.9% NaCl before being plated out on selective SOB plates with $25\,\mu\mathrm{g/ml}$ Kan. This step selected for the clones where the first recombination event had occurred. Plates were then incubated ($50\,^{\circ}\mathrm{C}$, 24-72 hours) until colonies appeared.
- In the next step, colonies were picked and transferred to SOB plates supplemented with $80\,\mu\mathrm{g/ml}$ of 5-FO and xylose (1%) to select for clones where the second recombination step and plasmid curing had occurred. The absence of antibiotics in this step releases the selection pressure retaining the plasmid.
- Colonies were picked, and gDNA isolated for colony PCR (initial denaturation step, 20 minutes) using specific primers for verification of deletion. The PCR products were subjected to gel electrophoresis to confirm their expected sizes.

- Mutants with presumably clean deletions were picked and streaked out on non-selective SOB plates to obtain single colonies. Plates were cultivated overnight at $50\,^{\circ}\mathrm{C}$.
- Screening of colonies harboring a clean deletion was carried out by colony PCR and gel electrophoresis
- For mixed strains of mutant and wild-type genotype, the two last steps were repeated until colonies with only a clean deletion were obtained.

2.12 Growth studies of *B. methanolicus* MGA3

Growth experiments carried out to evaluate the influence of either external factors or genomic modifications on the physiology of chosen *B. methanolicus* strains. Cell growth was estimated as a function of cell densities by measuring the OD_{600} at two-hour intervals. All cultivations were done in technical triplicates to calculate mean cell densities and the variance of growth as standard deviations. The procedure for calculating standard deviations is described in Appendix F.

B. methanolicus wild-type, *B. methanolicus* Δ*upp*, *B. methanolicus* Δ*upp* (pTH1mp), and *B. methanolicus* Δ*upp* (pTH1mp-*upp*) strains were cultivated in $\mathrm{MVcMY}_{200\mathrm{mM\ (MeOH)}}$ minimal medium supplemented with various concentrations of toxic antimetabolite 5-fluorouracil (5-FU). Pre-cultures were re-inoculated to main cultures to an initial OD_{600} of 0.2. An example calculation showing the cell culture needed for the re-inoculation is described in Appendix F. Cell growth was then monitored until late stationary phase of growth was reached. This was indicated by measuring the approximately same OD_{600} of a culture for two subsequent sampling points.

2.13 Quantification of total protein content in crude extracts using Bradford assay

The quantification of total protein amounts in cell crude extracts was based on the Bradford protein assay method (Bradford (1976)) and carried out according to the Lit33 instruction manual by Bio-Rad. The associated 5X Protein Assay Dye Reagent Concentrate (Bio-Rad. Cat. no 500-0006) was used as a dye for the proteins. It included the compound Coomassie®Brilliant Blue G-250, which differs in the color corresponding to the amount of proteins present in the solution. A 1X dye was made by diluting the concentrate 1:5 with Milli-Q Ultrapure water (Merck), which contains a reduced ion content, hence facilitating protein stability.

For the colorimetric determination of protein content in cell crude extracts, a calibration curve with the measured absorbance of 7 known Bovine Serum Albumin (BSA, NEB, Lot. 0071204)

concentrations was made. A table of the dilutions and the calibration curve is presented in Appendix G. Samples were prepared by adding $20\,\mu l$ of BSA or cell crude extracts to $1\,ml$ of 1X Protein Assay Dye Reagent Concentrate (Bio-Rad. Cat. no 500-0006) in a cuvette (1 ml), and incubating the mixtures for 10 minutes at room temperature. OD_{595} were measured using Helios Epsilon Spectrophotometer (Thermo Fisher Scientific). Measured absorbance of the samples with known concentrations was then plotted as a function of protein concentration, and a linear fit line was made. The linear correlation was eventually used to estimate protein concentrations in cell crude extracts from measured OD_{595}. See an example calculation in Appendix G.

2.14 MtlD enzyme assay

The procedure for performing a MtlD enzyme assay is based on a protocol described by Schultenkämper et al. (2019), and involves two parts; first, the preparation of cell crude extract, then followed by the enzyme assay itself. The enzyme activity was monitored by oxidation of the co-factor nicotinamide adenine dinucleotide hydride (NADH). NADH facilitates the reversible enzymatic reaction in which D-mannitol 1-phosphate is formed by the consumption of the substrate β-D-fructose 6-phosphate (F6P), catalyzed by the *mtlD* encoded enzyme mannitol 1-phosphate 5-dehydrogenase. The supplementation of the substrate F6P to the reaction mixture initiates the reaction.

The enzyme activities (U/ml) were calculated from the slopes of the decrease in absorbance of the baseline (background) and reaction activities for each technical triplicate, corresponding to the decrease of NADH per minute. The tangent slopes of the baselines were subtracted from the measured slope of the reaction activities to exclude the expected use of NADH by other enzymes present in the cell crude extracts. The new slope values, representatively for the reaction of interest, were then related to NADH concentrations using Beer's law and the enzyme activities calculated (See Appendix F). The enzyme specific activities (U/mg) were then calculated by dividing the calculated enzyme activities (U/ml) with the respective protein concentrations estimated for each cell crude extract using Bradford protein assay (Bio-Rad). See Appendix F for calculations of enzyme specific activity.

Preparation of cell extract

Initially, $40\,ml$ of pre-warmed ($50\,°C$) MVcMY medium was prepared in a baffled flask and inoculated with the desired *B. methanolicus* strain using an inoculation loop from a frozen glycerol stock. The culture was cultivated overnight until the exponential phase of growth (16 hrs, final OD_{600} of 0.5 - 4.5), and reinoculated (OD_{600}=0.2) into technical triplicates ($40\,ml$) of pre-warmed ($50\,°C$) MVcMY media supplemented with 50 mM mannitol. Adding of mannitol

was made to induce the mannitol inducible promoter (P_{mtlR}) for expression of the *mtlD* gene (Irla et al. (2016)). Each main culture was then grown for at least 6 hours to an OD_{600} of 0. 8-1.0 was measured, corresponding to two doubling times. $20\,\mathrm{ml}$ of cell culture was then harvested by centrifugation (4,000 rpm, 10 min, $4\,^{\circ}\mathrm{C}$) and washed two times using pre-chilled $50\,\mathrm{mM}$ TRIS-HCl buffer (pH 7.5). After this step, cell pellets were stored at $-80\,^{\circ}\mathrm{C}$ for later use. In advance of the enzyme assay, cells were thawed on ice and resuspended in $1\,\mathrm{ml}$ of $50\,\mathrm{mM}$ TRIS-HCl buffer (pH 7.5). Cell lysates was then obtained by sonication using Fisherbrand Sonic Dismembrator (FB-505) with the following settings: 25% amplitude, 2 seconds pulses, and 1-second pause, in repeating cycles for a total of 5 minutes. Cell debris was collected by centrifugation ($14.000\,\mathrm{g}$, 1 hour, $4\,^{\circ}\mathrm{C}$), and the supernatants collected into separate tubes. Finally, their concentrations were determined using Bradford assay (Section 2.13).

Enzyme assay

A reaction mixture was prepared, containing $850\,\mu\mathrm{l}$ of $50\,\mathrm{mM}$ TRIS-HCl buffer (pH 7.5) and $50\,\mu\mathrm{l}$ of $5\,\mathrm{mM}$ NADH (giving a final concentration of approximately 0.25 mM), and pre-warmed to $50\,^{\circ}\mathrm{C}$. $50\,\mu\mathrm{l}$ of crude extract was then added to the reaction mixture, and the baseline monitored for 3 minutes. Afterward, the reaction was initiated by the addition of $50\,\mu\mathrm{l}$ of F6P, giving a final concentration of the substrate of $10\,\mathrm{mM}$ per $1\,\mathrm{ml}$ reaction volume. The reaction was then monitored at $50\,^{\circ}\mathrm{C}$ and 340 nm for an additional 3 minutes. Absorbances were measured using Cary 100 UV-Vis Spectrophotometer, with the compatible software WinUV Kinetics Version 3.00 (Cary). For samples of cell crude extracts showing a non-linear decrease in OD_{340} as a function of time, dilutions were made, and the assay repeated. Dilutions of the samples were important for estimating the slope of change in OD_{340} as a function of time, in order to determine the enzyme activities of the samples.

Chapter 3

Results

3.1 Construction of plasmids for deletion of chromosomal genes by homologous recombination or plasmid-based complementation

To develop a genetic tool for the deletion of chromosomally-encoded genes using homologous recombination, multiple suicide vectors targeting the genes of interest (g.o.i.) were constructed. A flowchart describing this process is presented in Figure 3.1. These suicide vectors contain homologous regions both upstream and downstream to the g.o.i. in the chromosomal DNA to enable homologous recombination events to occur. As a basis for the construction, the conjugative plasmid pDELxp-*oroP* created by Irla (unpublished, 2019) was used. Specific primers targeting the upstream and downstream regions of the g.o.i. were used to amplify the flanking regions of the suicide vector using PCR and were designed to include the 15 first and last basepairs of the gene. Gene-specific suicide vectors were then made by cloning of the PCR amplified flanking regions into the *Spe*I digested vector backbone of pDELxp-*oroP* using Gibson assembly. For efficient assembly, primers used for the amplification of the flanking regions included 20-22 basepairs homologous to the 5' and 3' ends of the linearized pDELxp-*oroP* vector backbone.

The physiological effects of the gene knockouts were investigated by growth experiments and enzyme assays. Furthermore, to determine if an altered phenotypic response of the cells was due to the deletion of the g.o.i., and not from interruptions in the regions upstream or downstream of the target loci, plasmid-based complementation of a mutant strain was performed. For complementation of the knocked-out genes, cloning of the respective genes was obtained through PCR-based amplification of the target chromosomal DNA. Genes were then cloned into the rolling circle-replicating plasmid pTH1mp, where the expression could be regulated by the strong constitutive methanol dehydrogenase promoter (P_{mdh}, abbreviated *mp*) promoter (Irla

et al. (2016)). Following construction, plasmids were transformed into *B. methanolicus* and new physiological tests carried out.

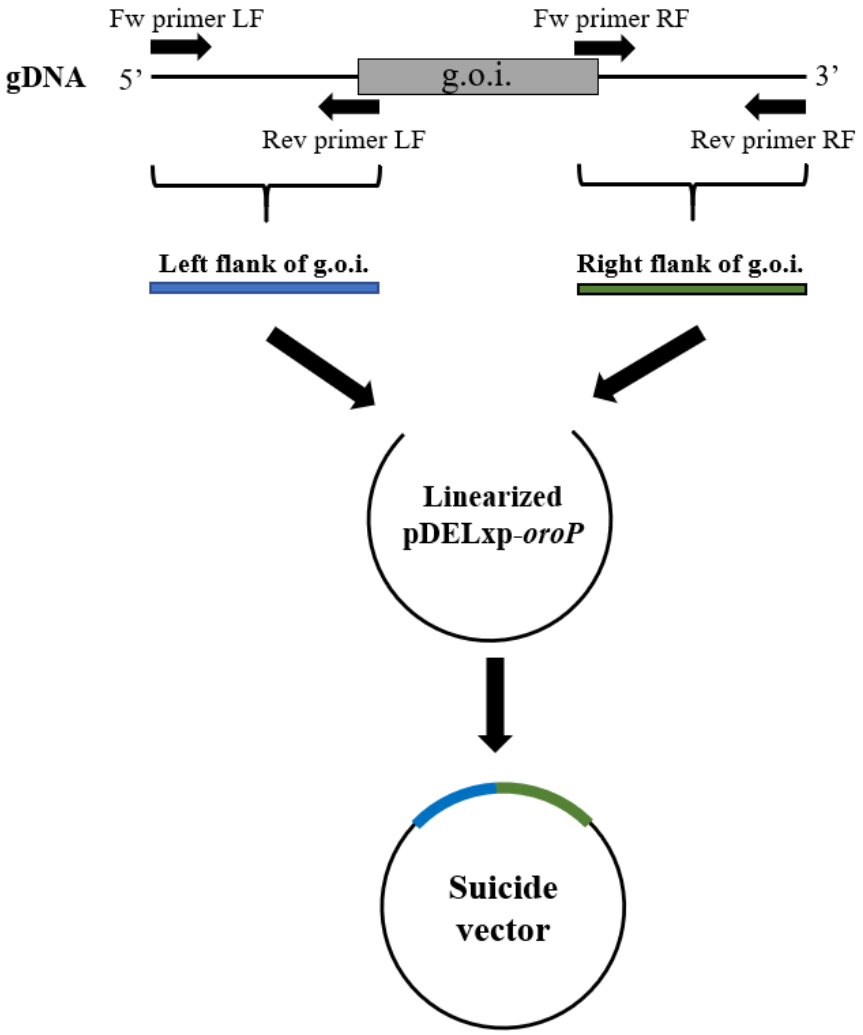


Figure 3.1: Flowchart showing the construction of a pDELxp-*oroP* suicide vector. Regions upstream (blue) and downstream (green) of the gene of interest (g.o.i.) were amplified by PCR using specific primers. "Fw"-forward, "Rev"-reverse. The amplified regions were then assembled with the linearized pDELxp-*oroP* vector using Gibson assembly, to construct a suicide vector with flanking regions homologous to the gene to be deleted.

3.1.1 Construction of suicide vectors for homologous recombination

Suicide vectors were made for the deletion of the chromosomal genes *upp*, *mtlD*, and *aroE*. The linearized vector basis, pDELxp-*oroP* was restriction digestion with *Spe*I, yielding a linear DNA molecule of 5.3 kb in length. The digestion was verified by gel electrophoresis, as shown in Figure 3.2a. The DNA ladder is displayed in well 1, and the circular and linear plasmid in wells 2 and 3, respectively. Construction of suicide vectors for deletion of *upp* and *mtlD* with 1000 bps flanking regions was conducted in preliminary work by Heid and Irla (unpublished).

Contrarily, the construction of suicide vectors for the deletion of *aroE* was part of this study.

For construction of the suicide vector pDELxp-*oroP*-Δ*aroE*, used for the deletion of the chromosomal gene *aroE*, flanking regions to the gene were PCR amplified from *B. methanolicus* gDNA of sizes: 250 bps, 500 bps, and 1000 bps. The construction of suicide vectors with three different sizes of homologous regions flanking the gene was carried out to determine the effect of the size of the homologous regions on the recombination efficiency leading to the chromosomal integration of the plasmid. PCR amplification of flanking regions of 250 bps was made using primers DEL47/DEL50, 500 bps regions using primers DEL47/DEL49, and for 1000 bps flanks the primers DEL47/DEL48 were used. Sizes of the amplified regions were confirmed by gel electrophoresis, shown in Figure 3.2b. Well 1 displays the DNA ladder, wells 2 and 5 display the amplified 250 bps flanks, wells 3 and 6 display the 500 bps flanks, and wells 4 and 7 display the 1000 bps flanks.

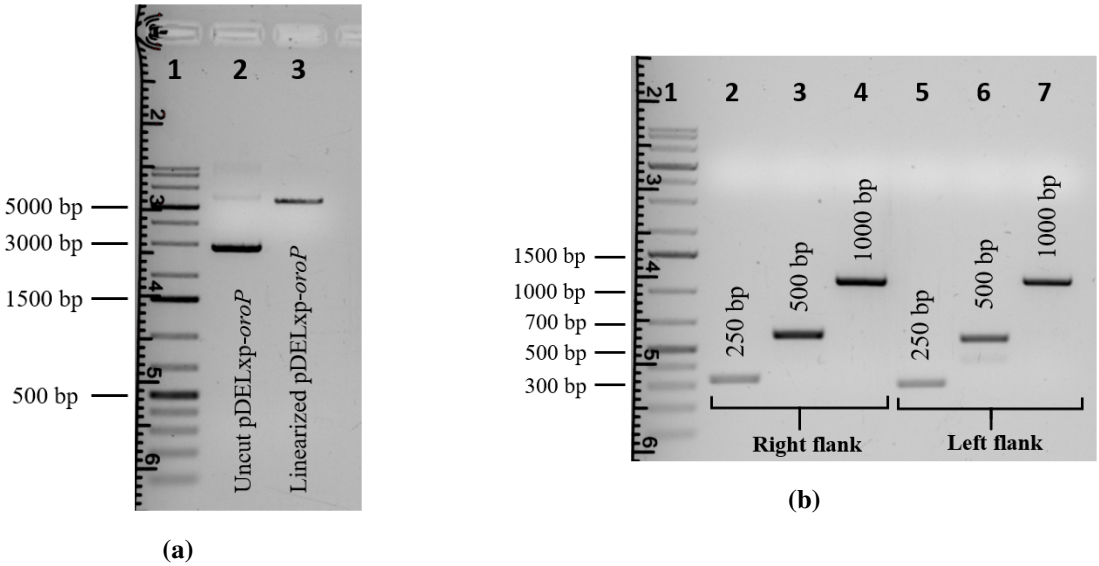


Figure 3.2: 0.8% agarose gel pictures. Gel electrophoresis carried out at 100 V, 40 minutes. DNA ladder displayed in well 1. (**a**) Gel picture of circular (well 2) and *Spe*I digested pDELxp-*oroP* (well 3). (**b**) PCR amplified flanking regions of sizes 250 bps, 500 bps and 1000 bps to *aroE* gene (wells 2-7).

The Gibson Assembly of each plasmid was then made by cloning of the right and left flanks of the same length with the linearized pDELxp-*oroP* vector, yielding inserts of 500 bps, 1000 bps, and 2000 bps in size. Plasmid maps of all three suicide vectors are presented in Appendix D; pDELxp-*oroP*-Δ*aroE*250, pDELxp-*oroP*-Δ*aroE*500, and pDELxp-*oroP*-Δ*aroE*1000, for regions of 250 bps, 500 bps, and 1000 bps flank sizes, respectively. Constructs were verified by gel electrophoresis using specific primers and transformed into chemically competent *E. coli* DH5α cells and plated on LA plates with Kan for positive selection. Primers DEL50/DEL54, DEL53/DEL49, and DEL51/DEL48, were then used for colony PCR in order to select for clones with inserts. The results of colony PCR are shown in Figure 3.3 and displays the expected insert

sizes of 500 (wells 1, 2, 3, 5), 1000 (wells 6 to 9) and 2000 bps (wells 10 to 13) for pDELxp-*oroP*-Δ*aroE*250, pDELxp-*oroP*-Δ*aroE*500, and pDELxp-*oroP*-Δ*aroE*1000, respectively. The DNA ladder is shown in well "L".

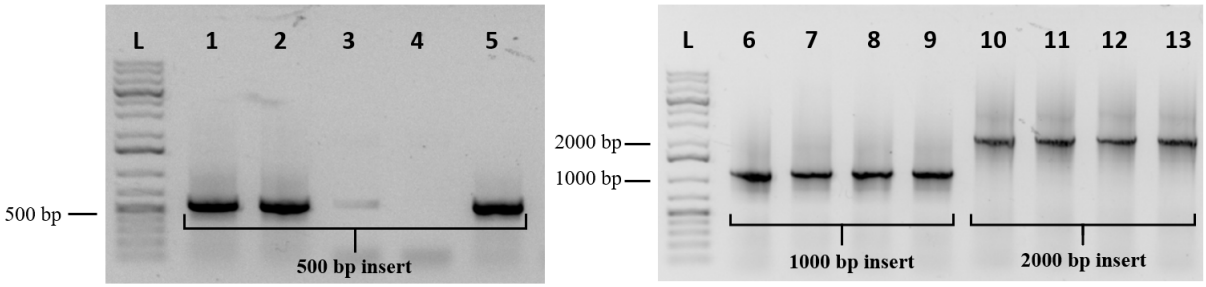


Figure 3.3: Gel electrophoresis carried out at 100 V, 40 minutes. 0.8% agarose gel pictures showing colony PCR products of *E. coli* DH5α clones with pDELxp-*oroP*-Δ*aroE* suicide vectors. DNA ladder is shown in well L. Positive clones with inserts of 500 bps are shown in wells 1, 2, 3, and 5. Positive clones with inserts 500 bps and 1000 bps are displayed in wells 6-9 and 10-13, respectively.

Plasmids were then isolated from overnight cultures of positive clones, and their inserts verified by DNA sequencing. Primers DEL50/DEL54, DEL53/DEL49, and DEL51/DEL48 were used for confirming the inserts of 500 bps, 1000 bps, and 2000 bps, respectively. Verified constructs were then transformed into chemically competent *E. coli* S17-1 and grown overnight on selective LA plates with Kan for positive selection. Positive transformants were confirmed by colony PCR using the same primers as for DNA sequencing, and the plasmids introduced into *B. methanolicus* MGA3 by conjugation.

3.1.2 Construction of a plasmid for complementation experiment

The construction of a complementation plasmid for overexpression of the chromosomally deleted gene *upp* was done by general plasmid-based cloning. *upp* was amplified from gDNA of *B. methanolicus* MGA3 using primers DEL57/DEL58. The cloning vector pTH1mp was isolated from an overnight culture of *E. coli* DH5α (pTH1mp) and digested with restriction enzymes *Pci*I and *Bam*HI over-night. Both of the corresponding restriction sites are placed directly downstream of the *mdh* promoter, hence controlling the expression of the introduced gene (see Figure 3.4).

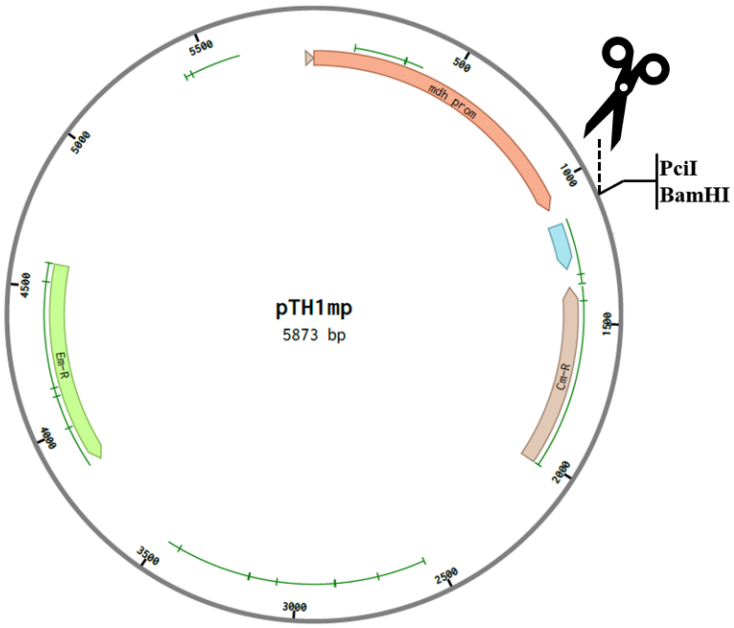


Figure 3.4: Plasmid map of pTH1mp (5873 bps) with restriction sites *Pci*I and *Bam*HI. The plasmid harbors a constitutive methanol dehydrogenase promoter (*mdh*) and a chloramphenicol selection marker (*Cm-R*)). The scissors indicates the cleavage point for linearization of the plasmid.

Successful restriction digestion of pTH1mp and PCR amplification of *upp* using primers DEL57/DEL58 were confirmed by gel electrophoresis, and the DNA fragments were subsequently purified. As displayed in Figure 3.5, the amplified *upp* gene (630 bps) containing overhangs yields a band of almost 685 bps in length (well 1) and the digested pTH1mp about 5.8 kb in size (well 3). Due to the different conformations of the plasmid, the linear one appears much larger than the undigested circular plasmid (well 2).

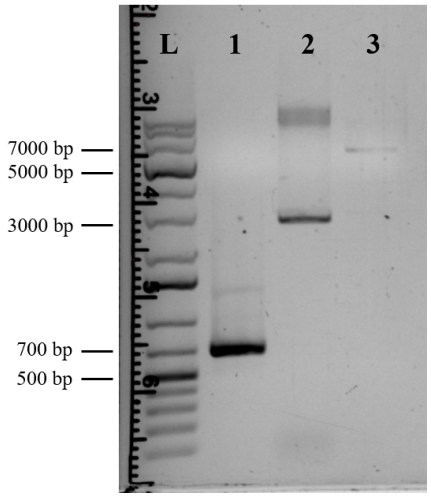


Figure 3.5: 0.8% agarose gel carried out at 100 V, 35 minutes. Gel picture showing *Pci*I and *Bam*HI digested pTH1mp (well 3, 5.8 kb) and PCR amplified *upp* (well 1, 685 bps). DNA ladder is displayed in well L and the circular PTH1mp vector is shown in well 2.

upp was then cloned into the vector backbone of pTH1mp using Gibson assembly and transformed into chemically competent *E. coli* DH5α cells. Transformants were plated on selective plates overnight, and colonies were picked for colony PCR using primers VPJF/VPJR and subjected to gel electrophoresis. Plasmids of positive clones were isolated from over-night cultures and verified by DNA sequencing, using the same set of primers. Correct constructs from sequencing were transformed into the Δ*upp* and wild-type strain of *B. methanolicus* MGA3 by electroporation. Besides, wild-type and Δ*upp* strains harboring the empty vector pTH1mp vector were prepared in the same manner.

3.2 Deletion of *upp*

3.2.1 Establishing a Δ*upp* deletion strain

The targeted chromosomal deletion of *upp* in *B. methanolicus* MGA3 using the formerly constructed suicide vector pDELxp-*oroP*-Δ*upp* was done in preliminary experimental efforts by Heid and Irla (unpublished, 2019). As part of this study, the confirmation of deletion was achieved by colony PCR using confirmation primers DEL29/DEL30. The amplified PCR product was subjected to gel electrophoresis (well 3) for confirmation of the clean deletion, see Figure 3.6. Simultaneously, the resulting amplicon from PCR using the same primers and gDNA of *B. methanolicus* MGA3 wild-type as a template was applied as reference (well 2). The difference between the DNA bands resulting from the wild-type and putative Δ*upp* MGA3 strain was estimated to the expected size of 630 bps, indicating a clean deletion of *upp*.

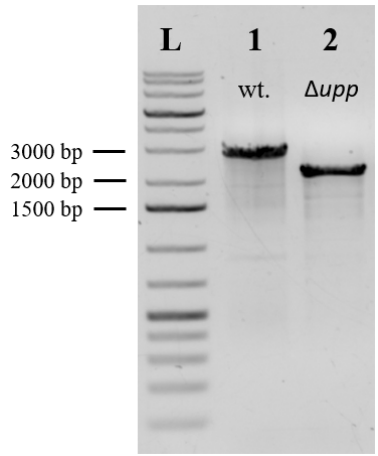


Figure 3.6: 0.8% agarose gel. Gel electrophoresis carried out at 100 V, 35 minutes. Gel picture after colony PCR for confirmation of *upp* deletion. Well L contains the DNA ladder and well 1 and 2 displays the amplified PCR products from gDNA of the wild-type strain and the Δ*upp* mutant strain, respectively.

Furthermore, for additional verification of a clean deletion strain, the PCR product presented in well 2 was sent for DNA sequencing, using the same primers as for the colony PCR. The obtained sequences were multi-aligned towards a template sequence of the clean deletion of *upp* in the chromosomal DNA. From this alignment, it was eventually confirmed that *upp* successfully had been deleted and that no interruptions have been made directly upstream or downstream of the gene loci within the sequenced region. Furthermore, the streaking of the putative Δ*upp* strain on SOB agar plates with $25\,\mu g/ml$ Kan yielded no colonies after three days of incubation, indicating the absence of a Kan^R gene. Thus, these results indicate the successful curing of the integrated plasmid and the *upp* gene.

3.2.2 Proposed uracil salvage pathway in *B. methanolicus*

As part of this study, one of the genes targeted for disruption was *upp*, encoding a uracil phosphoribosyltransferase. This enzyme is assumed to possess an essential function in the utilization of exogenous uracil in bacteria. This enzyme's physiological properties were investigated in the present study by growth experiments, carried out in minimal medium supplemented with different concentrations of 5-FU (Section 3.2.4). As the pyrimidine salvage pathways to present has not been elucidated in *B. methanolicus*, other potentially involved genes and enzymes of the pyrimidine salvage pathway were explored using bioinformatics approaches. Furthermore, this elucidation of related genes intended to provide information that also could be used to explain the physiological changes upon deletion of *upp*.

The presence of genes encoding for enzymes with similar functions between species of bacteria gives insights into the putative mechanisms of pyrimidine and uracil metabolism in *B. methanolicus*. By comparing the characterized pyrimidine salvage genes found in different species of bacteria, *M. tuberculosis* and *B. subtilis*, with the available genome sequence of *B. methanolicus* (MGA3) (Irla et al. (2014)), multiple similar genes were discovered. From the pyrimidine metabolism map from KEGG: Kyoto Encyclopedia of Genes and Genomes (Bioinformatic resource, Kanehisa (2000)), the putative uracil metabolism-related genes in *B. methanolicus* was found to support the synthesis of both RNA and DNA from exogenous uracil. The gene *pyrP* encoding a membrane-bound uracil permease, the enzyme responsible for uptake of extracellular uracil into the cell (Turner et al. (1994)), was also found present in *B. methanolicus*.

A comparison of the genes involved in uracil metabolism in *B. subtilis* subspecies (subsp.) 168 and the genome of *B. methanolicus* revealed at least nine similar genes, involved in subsequent reactions after the uptake of uracil from the environment. The amino acid sequences of the respective enzymes were compared using Protein BLAST (NCBI), and their calculated percentage identity is summarized in Table 3.1. Together with these findings, a proposed outline of the uracil metabolism pathway in *B. methanolicus* (Figure 3.7) was made by comparison of the

genome of *B. methanolicus* with other elucidated pyrimidine salvage pathways, as the one in *M. tuberculosis* (Singh et al. (2015)) and *L. plantarum* (Arsène-Ploetze et al. (2006)). The highest percentage identities were found for the genes *upp* (86.12%) and *pyrH* (86.67%) in which their encoded enzymes are assumed to be involved in the conversion of uracil to UMP.

Table 3.1: Genes associated with the uracil salvage pathway in *B. subtilis* subsp. 168 and *B. methanolicus* MGA3. A BLAST protein search between the amino acid sequences of the encoded proteins from the two strains yielded their similarities by percentage identity.

Gene	Enzyme	Locus tag *B. methanolicus* MGA3	Locus tag *B. subtilis* subsp. 168	EC Number	% Identity
upp	Uracil phosphoribosyltransferase	BMMGA3_16035	BSU36890	2.4.2.9	86.12
pyrR	Pyrimidine operon attenuation protein/ Uracil phosphoribosyltransferase	BMMGA3_05780	BSU15470	2.4.2.9	74.59
pyrH	Uridylate kinase	BMMGA3_06265	BSU16510	2.7.4.22	86.67
cmk	Cytidylate kinase	BMMGA3_10855	BSU22890	2.7.4.25	68.47
ndk	Nucleoside diphosphate kinase	BMMGA3_10755	BSU22730	2.7.4.6	70.95
tk	Thymidine kinase	BMMGA3_16090	BSU37060	2.7.1.21	76.04
tmk	TMP kinase	BMMGA3_00180	BSU00280	2.7.4.9	66.33
pdp	Pyrimidine-nucleoside phosphorylase	BMMGA3_11090	BSU39400	2.4.2.2	76.44
udk	Uridine kinase	BMMGA3_12210	BSU27330	2.7.1.48	81.04

In the proposed pathway (Figure 3.7), uracil is readily converted to UMP in the reaction catalyzed by the phosphoribosyltransferases encoded by *upp* and *pyrR*. UMP is further converted to uridine diphosphate (UDP) by the activity of a phosphotransferase, encoded by the genes *cmk* and *pyrH*. UDP can, in turn, be converted to deoxy-UDP (dUDP) by the catalytic activity of ribonucleotide reductase. A pyrimidine-nucleoside phosphorylase encoded by *pdp* was found to potentially catalyze the conversion of deoxyuridine (dUrd) to uracil.

A thymidine kinase encoded by *tk* and a TMP kinase was also proposed for the conversion of dUrd to deoxyuridine monophosphate (dUMP) and dUDP, respectively. In turn, UDP and dUDP are further converted to their corresponding nucleoside triphosphates by nucleoside diphosphate kinase. Multiple subsequent enzymatic steps eventually incorporate these compounds into RNA and DNA synthesis. Based on this study, no genes responsible for the direct formation of deoxyuridine (dURD) from uracil was found in *B. methanolicus*.

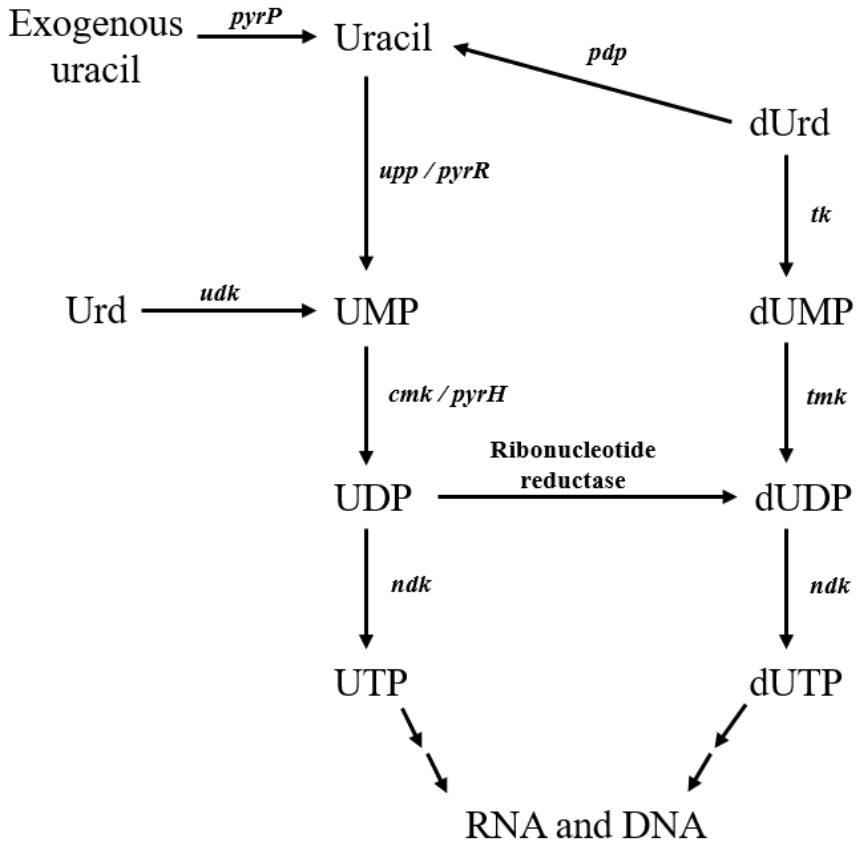


Figure 3.7: Proposed uracil salvage pathway in *B. methanolicus*. Arrows indicate the preferred direction of the reaction. Genes and proposed enzyme: *upp* - uracil phosphoribosyltransferase, *pyrR* - pyrimidine operon attenuation protein/ uracil phosphotransferase, *cmk* - CMP kinase, *pyrH* - UMP kinase, *ndk* - nucleoside diphosphate kinase, *pdp* - pyrimidine-nucleoside phosphorylase, thymidine kinase (*tk*), *tmk* - TMP kinase, *udk* - uridine kinase. The salvage pathway of uracil is associated with the production of RNA and DNA in the cell. Figure adapted from Singh et al. (2015).

The primary utilization pathway of uracil was found to include proteins with enzymatic uracil phosphoribosyltransferase activity encoded by the genes *upp* and *pyrR*. Interestingly, the gene *udp* encoding uridine phosphorylase was found absent in *B. methanolicus*, making it less likely that formed UMP results from an alternative uracil metabolic pathway than the one using *upp*/*pyrR*. A protein BLAST of the UPP enzyme in *B. subtilis* subsp. 168 towards the corresponding enzyme in *B. methanolicus* gave an identity score of 86.12%. From the high degree of homology, it can be suggested that these enzymes catalyze the same reactions. In *B. subtilis*, *upp* encoded UPP was found responsible for the major conversion of exogenous uracil (Martinussen et al. (1995)). It is, therefore, likely that the deletion of *upp* in *B. methanolicus* will pose a great reduction in exogenous uracil uptake and utilization.

Alternative pathways for uracil utilization in *B. methanolicus*

Martinussen et al. (1995) showed that a *B. subtilis* Δ*upp* strain displayed sensitivity to the 5-

FU above specific concentrations and the incorporation of uracil into its nucleic acids. These results indicate an alternative pathway, able to take over the main catalytic activity of the UPP. Based on these findings, it is reasonable to expect a similar process to occur in *B. methanolicus*. Therefore, alternative enzymes to *upp* and *pyrR* for the conversion of uracil were also investigated.

A pyrimidine-nucleoside phosphorylase (EC 2.4.2.2) encoded by *pdp*, catalyzes the reaction that converts uridine and phosphate to uracil and α-D-ribose 1-phosphate, and has previously been described in *Geobacillus stearothermophilus* by Saunders et al. (1969). The presence of this gene was also found in the genome of *B. methanolicus* (Table 3.1). However, the affinity and turnover number of this enzyme to uracil in *B. methanolicus* has not yet been studied. For this reason, the reaction converting dUrd to uracil in Figure 3.7 is displayed as a non-reversible reaction. From a search in the KEGG database (Kanehisa (2000)), two additional *B. methanolicus* genes were also found interesting. These genes, *doeD* (EC 2.4.2.1) and *punA* (EC 2.4.2.1), encode both assumed purine-nucleoside phosphorylases. The *punA* and *deoD* encoded proteins have previously been described in the protozoan parasite *Plasmodium falciparum* for being able to convert uridine to uracil (Cui et al. (2010)). However, the reversibility of the reaction has not been fully clarified, nor is it known whether these enzymes have the same catalytic activity in *B. methanolicus*.

In final, the *udp* gene encoding uridine phosphorylase (EC 2.4.2.3), also associated with the conversion of uridine to uracil, was investigated. Formerly this gene has previously have been investigated in bacteria such as *Shewanella oneidensis* and *E. coli* (Mordkovich et al. (2013); Leer et al. (1977)). However, this gene was not found in the genome of *B. methanolicus*, so no further analysis of the function of this gene was made.

3.2.3 Genomic neighborhood of *upp* and effect on homologous recombination in Δ*upp* strains

The genomic neighborhood of a gene represents the other genes located close to its loci in the genome. Bacterial genes are commonly organized in operons, and many are usually co-expressed. Thus, promoters and other transcription regulators could potentially be affected upon the deletion of a gene. Therefore, the elucidation of a gene's genomic neighborhood can serve as a good starting point in search of physiological explanations. As this study intends to target the *upp* gene for deletion, the exploration of its genomic neighborhood aimed to provide useful insights that could explain the observed phenotypic traits of its corresponding mutant strains.

Based on the annotated genome of *B. methanolicus* (Irla et al. (2014)), the genes directly upstream and downstream of *upp* were found from a search in the genome (GeneBank:CP007739.1)

using ”Genome search”, a tool provided by National Center for Biotechnology Information (NCBI). Four genes 4-5 kb upstream and downstream of the gene were carefully examined (Figure 3.8), and the findings are presented in Table 3.2. Table 3.2 displays the MGA3 gene annotated names, and the assumed encoded proteins and protein functions.

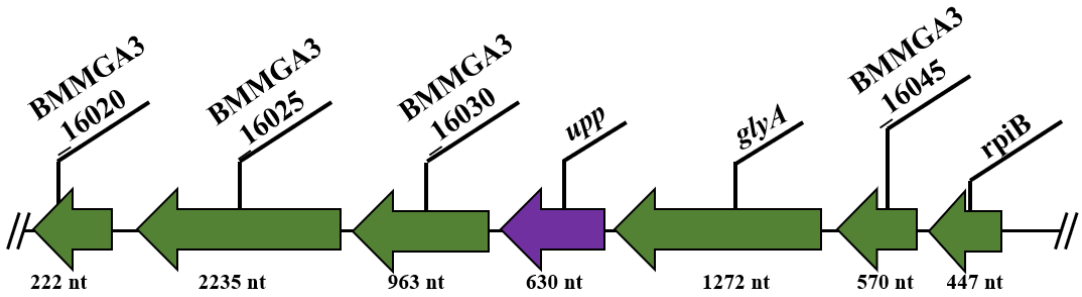


Figure 3.8: Genomic neighborhood of *upp* in *B. methanolicus*. The four genes upstream and downstream of *upp* are displayed, and all genes are located in the same direction. The nucleotide (nt) size of the genes are also presented.

The annotations of protein functions were obtained using the basic local alignment search tool (BLAST) for protein alignments (BLASTP) (Altschul (1997)), provided by NCBI. Using the amino acid sequences of the found genes neighboring *upp* (Table 3.2) as query sequences in BLASTP, this tool aligns the sequences towards a non-redundant protein sequence database of known proteins to check for sequence similarities. From this alignment, a GenBank report was obtained for the match with the highest percentage identity, corresponding to the annotated protein in *B. methanolicus*. GenBank accession numbers and references to the presented annotations are shown in Table 3.2.

However, the proteins encoded by three of the investigated genes were not previously annotated, including; BMMGA3_16025, BMMGA3_16015, and BMMGA3_16045. The annotations of these proteins were made automatically by NCBI using the Prokaryotic Genome Annotation Pipeline (PGAP), using the Reference Sequence (RefSeq) database (Tatusova et al. (2016)). RefSeq includes an extensive collection of amino acid sequences, among others, and can be used to annotate uncharacterized proteins by homology to other characterized proteins and gene prediction algorithms. The resulting annotations are also included in Table 3.2.

Table 3.2: The genomic neighborhood of *upp* gene in gDNA of *B. methanolicus* MGA3. The table displays four genes directly upstream and downstream of the gene, the predicted names of the encoded proteins and their assumed function in the organism.

Gene	Protein	Assumed Function	Reference/GenBank Accession
		Upstream of *upp*	
BMMGA3_16030	NERD domain-containing protein	Nuclease involved in DNA processing	Grynberg (2004)/AIE61559.1
BMMGA3_16025	S8 family serine peptidase	Cleavage of peptide bonds	AIE61558.1
BMMGA3_16020	AtpZ/AtpI family protein	Transmembrane transportation of ATP	Hicks et al. (2003)/AIE61557.1
BMMGA3_16015	ATP synthase F0 subunit I	Production of ATP	AIE61556.1
		Downstream of *upp*	
glyA	Serine hydroxymethyltransferase	Conversion of serine to glycine and tetrahydrofolate to 5,10-methylenetetrahydrofolate	Irla et al. (2015)/AIE61561.1
BMMGA3_16045	TIGR01440 family protein	Uncharacterized	AIE61562.1
rpiB	Ribose 5-phosphate isomerase B	Conversion of ribose 5-phosphate to ribulose-5-phosphate	Heggeset et al. (2012)/AIE61563.1
BMMGA3_16055	HAMP domain-containing protein	Involved in chemotaxis	Aravind and Ponting (1999)/AIE61564.1

Searching upstream of the *upp* gene in *B. methanolicus* genome and BLASTP revealed a gene (BMMGA3_16030) encoding a NERD domain-containing protein, assumed to function as a nuclease in the processing of DNA. Neighboring, a gene (BMMGA3_16025) encoding an S8 family serine peptidase was located. Furthermore, the two following genes, which were closest to *upp*, BMMGA3_16020 and BMMGA3_16015, were found to encode an AtpZ/AtpI family protein involved in transmembrane transportation of ATP and an ATP synthase F0 subunit I involved in the production of ATP, respectively.

None of these genes were assumed to directly influence the growth of the cells or the recombination efficiency to a great extent. Downstream of *upp* the genes *glyA* and *rpiB* were found, encoding a serine hydroxymethyltransferase and a ribose-5-phosphate isomerase B, respectively. Ribose-5-phosphate isomerase B converts ribose-5-phosphate into ribulose-5-phosphate, thus contributing to the latter's recycling upon methylotrophic growth through the RuMP pathway. The serine hydroxymethyltransferase converts serine to glycine and tetrahydrofolate (THF) to 5,10-methylene-THF. Furthermore, a gene (BMMGA3_16045) encoding an uncharacterized TIGR01440 family protein and a gene (BMMGA3_16055) assumed to encode a HAMP domain-containing protein was found. The latter is assumed to be involved in chemotaxis.

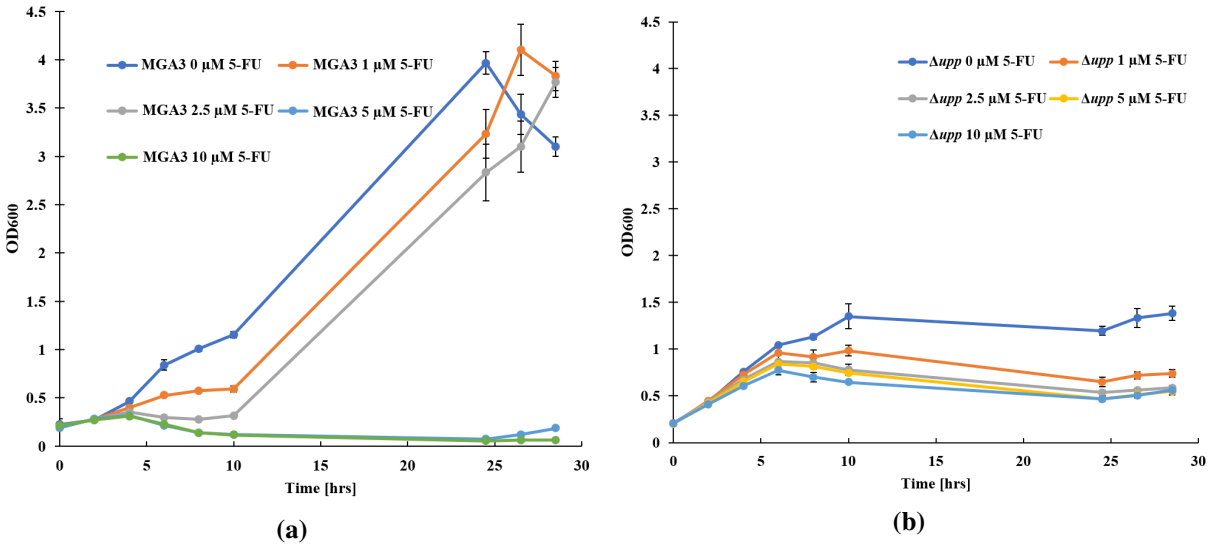
Putative promoters in the *upp* gene

The localization of *glyA* only 220 nucleotides downstream to *upp* (Figure 3.8) makes it reasonable to think that regulatory elements within the *upp* gene can influence the transcription of this gene. Using the Prediction of BROM – bacterial promoters online tool (Solovyev (2011)) two putative σ^{70} promoters within the *upp* gene were found positioned at positions 340 and 653 from the 5' end of *upp*. Furthermore, the positions of their -10 and -35 elements indicate that these putative promoters can regulate genes downstream of *upp*. Hence, the proximity of the last predicated promoter to the *glyA* gene, makes this the most probable transcription regulator out of the two predicted promoters.

3.2.4 Growth experiments of *Bacillus methanolicus* show increased tolerance towards the uracil analog 5-fluorouracil, upon deletion of *upp*

Characterization of the established *B. methanolicus* MGA3 Δ*upp* strain was conducted phenotypically by growth experiments in $\text{MVcMY}_{200\text{mM (MeOH)}}$ minimal medium supplemented the toxic uracil analog 5-FU. The *upp* gene is assumed to encode the principal converter of exogenous uracil in the pyrimidine salvage pathways of *B. methanolicus*. Hence, it was assumed that the deletion of *upp* would reduce the ability of *B. methanolicus* to utilize 5-FU, and consequently increasing its tolerance to this compound.

The toxicity of 5-FU has hitherto not been fully determined in *B. methanolicus*. For this reason, the growth experiment was carried out with the wild-type and deletion strains using varying concentrations of 5-FU: 0 µM, 1 µM, 2.5 µM, 5 µM, and 10 µM. All cultivations were performed in technical triplicates to calculate the sample standard deviation of the cultivations under similar conditions. The growth was monitored spectroscopically, and the mean OD_{600} and corresponding standard deviations for the technical triplicate cultivations are for both *B. methanolicus* strains plotted against time (hours). The plots are presented in Figure 3.9a and Figure 3.9b, for the wild-type and Δ*upp* strain, respectively. The standard deviations are represented as error bars. As both figures show, for cultivation without 5-FU, the growth of the wild-type strain is more rapidly increasing than the mutant strain. In this condition, the final OD_{600} of the wild-type strain (3.97) was measured to be almost three times higher than the final OD_{600} of the Δ*upp* strain (1.38).



(a) **(b)**

Figure 3.9: Growth pattern of *B. methanolicus* MGA3 ((**a**)) wild-type and ((**b**)) Δ*upp* strains. Mean OD_{600} of technical triplicate shake flasks is plotted against time (hours). $\text{MVcMY}_{200\text{mM (MeOH)}}$ minimal media supplemented with different concentrations of 5-FU was used. Standard deviations are displayed as error bars.

When the wild-type strain was grown on $2.5\,\mu\mathrm{M}$ 5-FU, the growth was strongly hindered in the beginning, indicating the lethal effect of 5-FU. Furthermore, the observed maximum OD_{600} overnight (24.5 hrs) under the conditions of 1 and $2.5\,\mu\mathrm{M}$ concentrations of 5-FU are shown to be seven to ten-fold higher than at 10 hrs. This is higher than the expected growth from the calculated doubling times (Table 3.3), which may result from the adaption of the wild-type strain to 5-FU.

By comparing the growth of the two strains at different concentrations of 5-FU, it becomes clear that the wild-type strain confers a higher sensitivity to 5-FU in general, compared to the Δ*upp* strain. From a further inspection of the calculated specific growth rates of the two strains (Table 3.3), the Δ*upp* strain is shown to be only slightly affected by the presence of 5-FU. For cultivation without 5-FU the specific growth rate of the Δ*upp* strain was calculated to 0.27 hr^{-1}, and 0.22 hr^{-1} for cultivation with $10\,\mu\mathrm{M}$ of 5-FU.

By assumption of a constant growth rate in the exponential phase, the doubling times (t_d) of each strain and conditions were then computed using Equation (3.1).

$$t_d = \frac{ln2}{\mu} \tag{3.1}$$

In the absence of 5-FU, the doubling times were estimated to be nearly the same for the two strain; 2.47 hrs for the wild-type strain and 2.57 hrs for the Δ*upp* strain. However, the decrease in specific growth rates for both strains upon increased concentrations of 5-FU varies significantly, as shown in Figure 3.10. The specific growth rate of the mutant strain decreases linearly according to an increased concentration of 5-FU. For the wild-type strain, the effect of 5-FU is most prominent as the medium is supplemented $1\,\mu\mathrm{M}$ of 5-FU, significantly reducing the specific growth rate from $0.28\,\mathrm{hr}^{-1}$ to $0.16\,\mathrm{hr}^{-1}$.

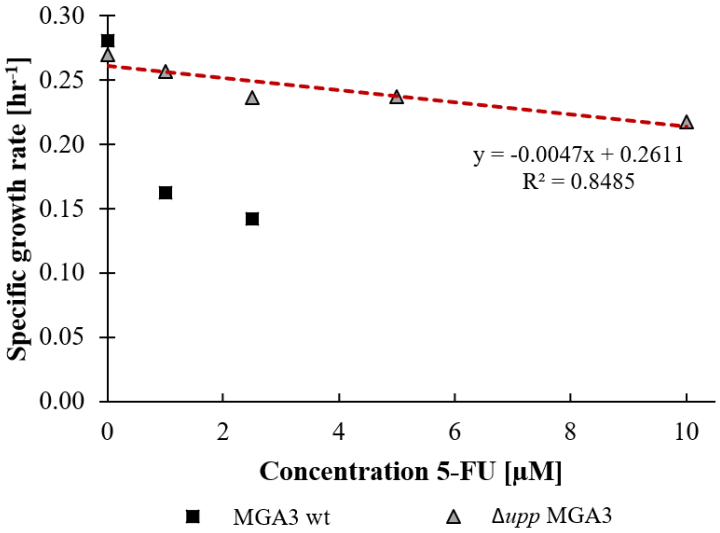


Figure 3.10: Specific growth rates for *B. methanolicus* wild-type and Δ*upp* strains as function of 5-FU concentration. An linear trend line is shown for the decrease in specific growth rates for Δ*upp* strain by increased 5-FU concentration: y=-0.0047x+0.2611. Coefficient of determination calculated to R^2=0.8485.

Table 3.3: Experimental results from growth experiment of B. methanolicus MGA3 wt. and Δ*upp* mutant strains upon growth on 5-FU. The calculated specific growth rates, doubling times and the highest measured OD_{600} for each strain under its respective conditions are included. "n.g." corresponds to no growth.

Strain	Concentration 5-FU (µM)	Specific growth rate (hr^{-1})	Doubling time (hr)	Maximum OD_{600}
MGA3 wt	0	0.28	2.47	3.97
MGA3 wt	1	0.16	4.26	4.10
MGA3 wt	2.5	0.14	4.87	3.77
MGA3 wt	5	n.g.	n.g.	n.g.
MGA3 wt	10	n.g.	n.g.	n.g.
Δ*upp* MGA3	0	0.27	2.57	1.38
Δ*upp* MGA3	1	0.26	2.7	0.98
Δ*upp* MGA3	2.5	0.24	2.93	0.87
Δ*upp* MGA3	5	0.24	2.92	0.84
Δ*upp* MGA3	10	0.22	3.19	0.77

The toxicity threshold limiting the growth of the wild-type strain was observed at $5\,\mathrm{mM}$ 5-FU, as shown in Figure 3.11. Under these conditions, no growth of the wild-type strain was observed. The difference in the growth of the wild-type strain when growing on $0\,\mu\mathrm{M}$ and $5\,\mu\mathrm{M}$ of 5-FU is displayed in Figure 3.12a. On the other hand, as Figure 3.12b shows, the growth of the Δ*upp* strain under the different concentrations of 5-FU is almost unaffected.

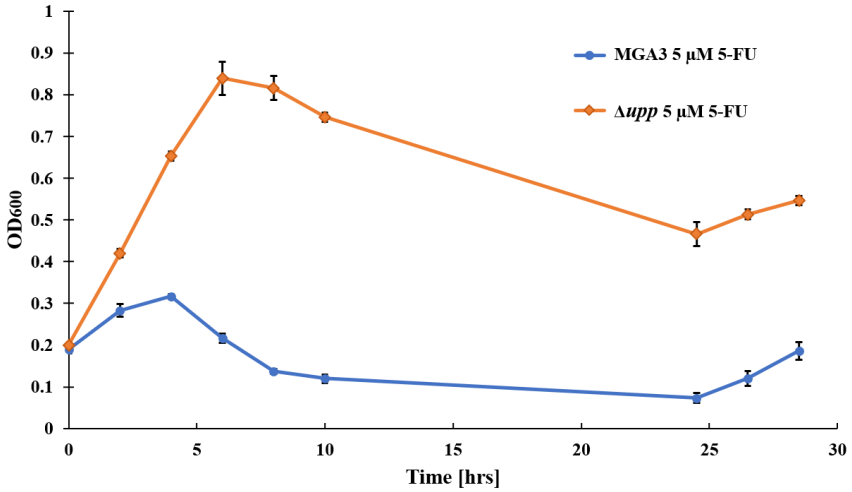


Figure 3.11: Effect of 5-FU on growth of *B. methanolicus* MGA3 wild-type and Δ*upp* strains on growth in $\text{MVcMY}_{\text{200mM (MeOH)}}$ minimal medium. Error bars correspond to standard deviations.

As shown in Table 3.3, the Δ*upp* strain reaches a lower final OD_{600} than the MGA3 wild-type strain, even when no 5-FU is added to the growth medium. To test if a similar growth behavior occurs when the *upp* gene is not entirely deleted, a mutant strain resulting from a single nucleotide polymorphism (SNP) in the *upp* gene was used. The strain denoted Δ*upp* EID was created in preliminary research by Drejer (unpublished) and was obtained from a selective adaption of *B. methanolicus* MGA3 wild-type strain plated on SOB with $10\,\mu\text{M}$ 5-FU. Colonies formed were shown to be resistant to 5-FU because of a frameshift mutation resulting in an early stop codon (Appendix E).

To test for growth, MGA3 wild-type, Δ*upp*, and the Δ*upp* EID strains were inoculated in $\text{MVcMY}_{\text{200mM (MeOH)}}$ media supplemented with $0\,\mu\text{M}$ and $5\,\mu\text{M}$ of 5-FU, and cultivated in technical triplicates for 31 hours. Mean values and standard deviations of the monitored OD_{600} of each technical triplicate as a function of time are presented in Figure 3.12a and Figure 3.12b for the wild-type and the mutant strains, respectively. The observed effect of 5-FU for the wild-type strain was similar to the previously described growth experiment, as the growth rapidly halted in the presence of $5\,\mu\text{M}$. Besides, as Figure 3.12b displays, a similar growth pattern of the two mutant strains was observed, both displaying a higher resistance to 5-FU than the wild-type strain. Notably, out of the two mutant strains, Δ*upp* EID showed a higher specific growth rate in the exponential phase. Moreover, from Figure 3.12b, both mutant strains (Δ*upp* and Δ*upp* EID) are shown to reach the same maximum OD_{600} of approximately 1.2.

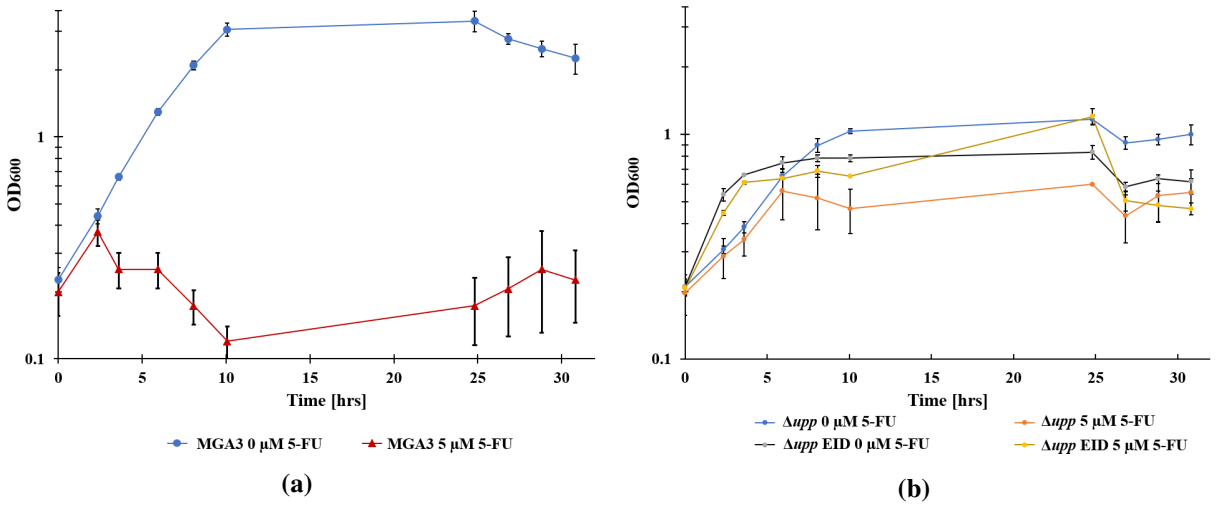


Figure 3.12: Growth curves showing mean OD_{600} of technical triplicate cultivations in $MVcMY_{200mM\ (MeOH)}$ minimal medium as function of time on a semi-logarithmic axis. Standard deviation is shown by error bars. **(a):** Growth curve of *B. methanolicus* MGA3 wild-type on 0 and $5\,\mu M$ of 5-FU. **(b):** Growth curve of the *B. methanolicus* MGA3 Δ*upp* and Δ*upp* EID strains cultivated on 0 and $5\,\mu M$ of 5-FU.

3.2.5 Complementation of *upp* in mutant strain shows regained sensitivity to 5-FU

Complementation of the gene *upp* by plasmid-based overexpression was expected to restore the wild-type phenotype and hence the sensitivity towards 5-FU. To test this growth experiment on 5-FU was performed. The strains *B. methanolicus* MGA3 and Δ*upp B. methanolicus* MGA3 harboring one of the two plasmids pTH1mp or pTH1mp-*upp* were cultured in technical triplicates using $MVcMY_{200mM\ (MeOH)}$ minimal medium.

The media were supplemented $5\,\mu g/ml$ Cm for retaining the plasmids, and the strains cultivated under conditions with $0\,\mu g/ml$ or $5\,\mu g/ml$ of 5-FU. Growth was monitored by a sampling interval of 2 hours for 25 hours, until stationary phase of growth. Growth curves are presented for all strains and growth conditions in Figure 3.13. The sampling points are shown by the mean value and corresponding standard deviations as error bars for the technical triplicates.

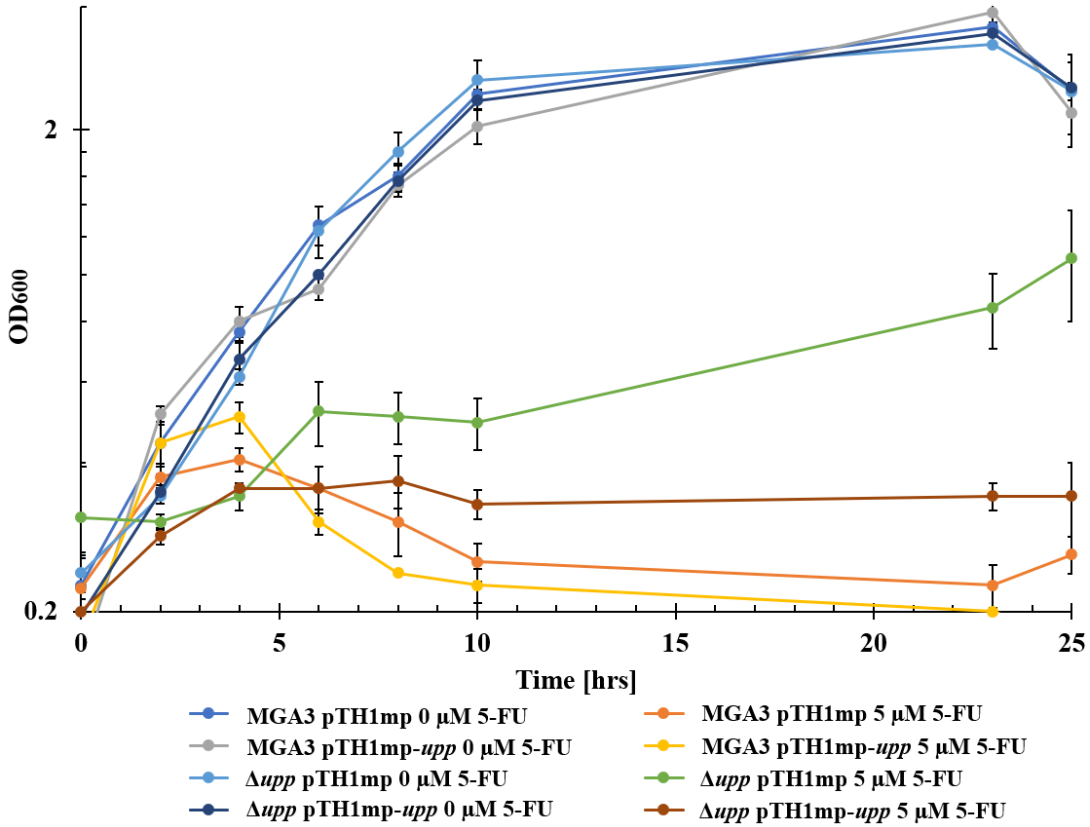


Figure 3.13: Growth curves of *upp* complementations study showing *B. methanolicus* MGA3 (pTH1mp/-*upp*) and MGA3 Δ*upp* (pTH1mp/-*upp*) for growth in $\text{MVcMY}_{200\text{mM (MeOH)}}$ minimal medium supplemented $0\,\mu\text{M}$ or $5\,\mu\text{M}$ of 5-FU. Data points are presented as the mean OD_{600} for each technical triplicate cultivation and standard deviations are displayed as error bars.

The exponential phases of growth for the technical triplicate cultivations under each respective condition were estimated from the semi-logarithmic plot in Figure 3.13. Based on these, the corresponding specific growth rates (μ) and doubling times were calculated. These values and measured maximum OD_{600} for each strain and condition are summarized in table Table 3.4. In general, a higher maximum OD_{600} were measured for strains cultivated without 5-FU supplementation in comparison to cultivation with 5-FU supplementation. Following previous observations, no growth was observed for the MGA3 wild-type strain harboring pTH1mp or pTH1mp-*upp*, when the media was supplemented $5\,\mu\text{M}$ 5-FU. As expected, this concentration of 5-FU shows no lethal effect on the Δ*upp* strain, albeit being strongly inhibiting, indicated by a low specific growth rate of $0.05\,\text{hr}^{-1}$. In the absence of 5-FU, the calculated specific growth rates indicate a more moderate growth for the wild-type strains and the mutants strains harboring pTH1mp-*upp* then the strains harboring the empty vector. Furthemore, from the growth curves in Figure 3.13 it is also observed that that the final OD_{600} of the Δ*upp* (pTH1mp) is higher than the previously growth experiments of the Δ*upp* in Table 3.3. The reason for this growth pattern is, however, difficult to explain.

Table 3.4: Calculated specific growth rates, doubling times and maximum OD_{600} of the *B. methanolicus* wild-type and Δ*upp* strains harboring plasmid pTH1mp or pTH1mp-*upp* cultivated under different conditions of 5-FU. Cultivations made in technical triplicates."n.g." corresponds to no observed growth.

Strain	5-FU (µM)	Specific growth rate (µ) (hr^{-1})	Doubling time (hrs)	Max OD_{600}
MGA3 pTH1mp	0	0.28	2.43	3.3
MGA3 pTH1mp	5	n.g.	n.g.	n.g.
MGA3 pTH1mp-*upp*	0	0.26	2.69	3.5
MGA3 pTH1mp-*upp*	5	n.g.	n.g.	n.g.
Δ*upp* MGA3 pTH1mp	0	0.32	2.18	3.0
Δ*upp* MGA3 pTH1mp	5	0.05	14.65	1.1
Δ*upp* MGA3 pTH1mp-*upp*	0	0.25	2.83	3.2
Δ*upp* MGA3 pTH1mp-*upp*	5	n.g.	n.g.	n.g.

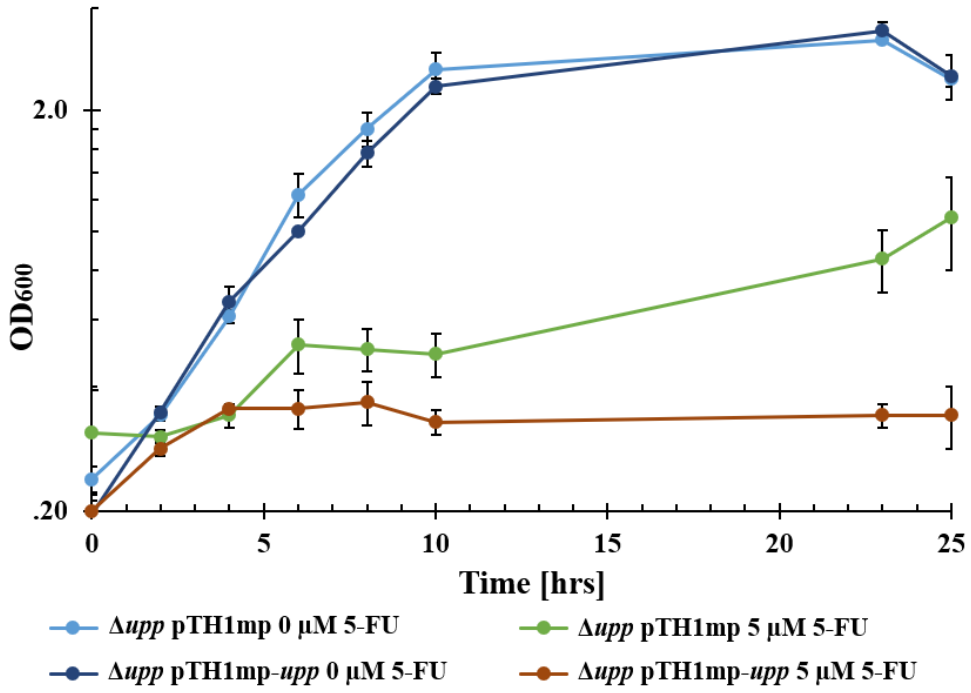


Figure 3.14: Growth curve of the Δ*upp* mutant strain of *B. methanolicus* MGA3 with and without the complementation of *upp*, when grown on $MVcMY_{200mM\ (MeOH)}$ minimal medium supplemented with 0 µM or 5 µM of 5-FU. Data points are presented as the mean OD_{600} for each triplicate cultivation. Standard deviations are displayed as error bars.

Figure 3.14 presents the growth curves of only the Δ*upp* strain harboring pTH1mp or pTH1mp-*upp* under the two conditions of 5-FU. In supplementation of 5 µM 5-FU, the effect of overexpressing *upp* in the Δ*upp* mutant strain results in no observable growth. Hence, the Δ*upp*

(pTH1mp-*upp*) strain has a lower growth rate and maximum OD_{600} than the control strain, Δ*upp* (pTH1mp), indicating the restored sensitivity to 5-FU for the *upp* overexpressed Δ*upp* strain. Nevertheless, this result matches the observed growth of the wild-type strain under the same condition, ultimately confirming the successful deletion of *upp* in the strain.

3.3 Deletion of *mtlD*

3.3.1 Successful deletion of *mtlD* shows versatility of the genetic tool

The second gene targeted for disruption in *B. methanolicus* was the mannitol 1-phosphate 5-dehydrogenase encoding gene, *mtlD*. An effective and targeted gene deletion tool that is applicable for subsequent deletion of multiple genes is seen as a great advantage as it proves its versatility. The deletion of the *mtlD* gene was carried out in the *B. methanolicus* MGA3 wild-type and Δ*upp* mutant strains, the latter eventually resulting in a double-deletion mutant strain.

The suicide vector pDELxp-*oroP*-Δ*mtlD* was simultaneously introduced by conjugation with *E. coli* S17-1 strain harboring the plasmid to the wild-type and Δ*upp* mutant strains of *B. methanolicus*. After the second selective plating of transconjugants on selective SOB-plates with xylose and 5-FO, colonies of putative mutants were screened by colony PCR using the primers SEMF/SEMR. These primers were constructed to bind outside the 1000 bps regions flanking the *mtlD* gene, and was hence suited for verifying its deletion. The PCR products of possible clean deletions were verified by gel electrophoresis. The gel pictures are presented in Figure 3.15, and displays a DNA fragment of 2.5 kb in size for the putative deletion of *mtlD*. Single deletion candidates are shown in wells 1 and 2 (Figure 3.15a), and double deletion candidates in wells 3 to 6 (Figure 3.15b). The corresponding PCR product resulting from the wild-type strain contains the *mtlD* gene (1161 bps) and is displayed as a band of approximately 3.7 kb in size (well R). An additional band is shown for the strain in well 1 and matching in size to the wild-type strain, thus indicating a mixed strain of the wild-type and mutant strains. To ensure that the other applied PCR products were resulting from pure deletion strains, gel electrophoresis was run with the PCR samples in overload to identify any weak bands. From the corresponding gel pictures, the absence of the wild-type DNA band was confirmed for the candidates in wells 2 to 6. Out of these, the single deletion candidate in well 2 and the double deletion candidates in wells 4 and 6 were used further.

Presumed positive clones were cultivated over-night, and their gDNA isolated. The gDNAs were then used as templates for PCR amplification to confirm the deletion, using primers SEMF/SEMR. Due to the unspecific binding of the primers (Figure 3.15), the band with the correct size of 2.5 kb was isolated from an agarose gel and purified. The confirmed deletion was

then ultimately verified by DNA sequencing using primers SEMF/SEMR and DEL39/DEL40. A multi-alignment analysis was then conducted to verify clean deletions of *mtlD*.

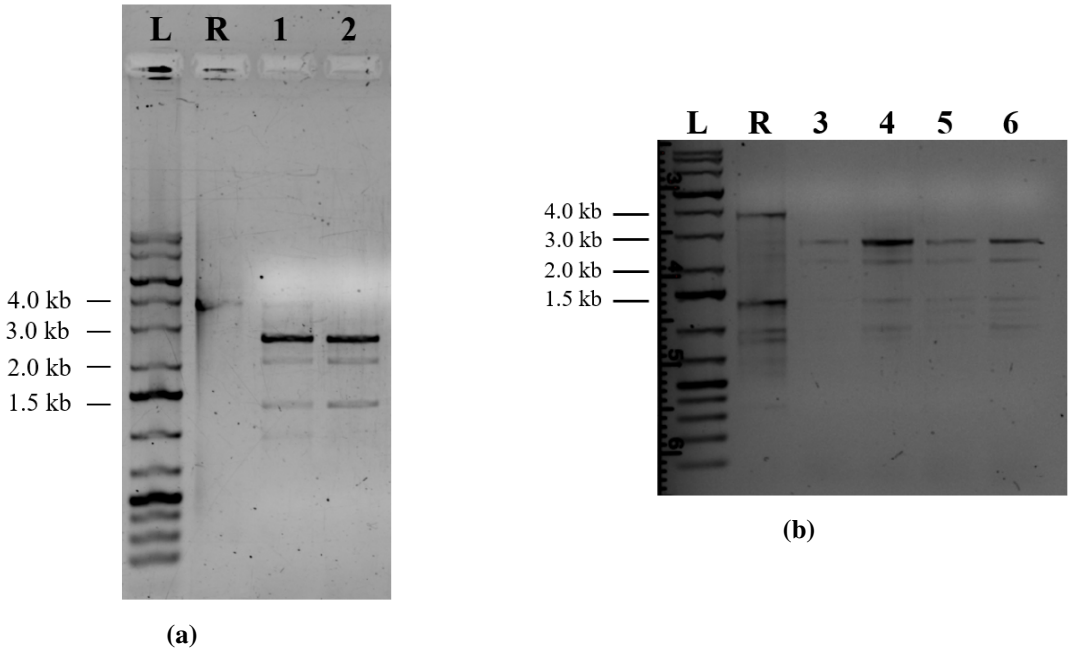


Figure 3.15: Gel Pictures from 0.8% agarose gel staining, displaying deletion of *mtlD*. Gel electrophoresis run at 100 V, 35 min. Well L displays the DNA ladder, well R the reference PCR product using the primers SEMF/SEMR for amplification from the wild-type strain. **(a):** Gel picture for confirmation of *mtlD* deletion in MGA3 wild-type strain. Wells 1 and 2 display the PCR products resulting from two putative Δ*mtlD* deletion strains. **(b):** Gel picture for confirmation of *mtlD* deletion in Δ*upp B. methanolicus* MGA3. Wells 3 to 6 display four Δ*upp* candidates with a clean deletion of *mtlD*.

The obtained sequences from DNA sequencing showed that only the primer DEL39 was annealing close enough to the target loci to verify the deletion of *mtlD*. Furthermore, the sequencing results also displayed that only the 15 bps of each end of the gene that were included in the homology flanks of the suicide vector, were remaining in the gDNA. From multi-alignments of the obtained sequences, the clean deletion of *mtlD* was then ultimately verified in both the wild-type and Δ*upp* strains (Appendix E).

3.3.2 Deletion of *mtlD* shows reduced mannitol-1-phosphate 5-dehydrogenase activity

An enzyme assay was conducted to assess the physiological effect of chromosomal deletion of *mtlD* on the reversible enzymatic conversion of D-fructose 6-phosphate to D-mannitol 1-phosphate, catalyzed by the *mtlD* encoded mannitol 1-phosphate 5-dehydrogenase. Furthermore, as the method for gene deletion described in this study permits the deletion of the gene as a whole, the enzyme assay was also set to reveal if other enzymes in *B. methanolicus* can take over the catalytic reaction upon the deletion of *mtlD*. The reaction requires the oxidation of the

co-factor nicotinamide adenine dinucleotide hydride (NADH) to occur. In contrary to NAD+, NADH absorbs light at 340 nm, making it possible to monitor the occurrence of the enzymatic reaction by spectrophotometry.

The calculated enzyme specific activities (U/mg) for the MGA3 wild-type and Δ*mtlD*Δ*upp* strains are presented in Figure 3.16, and were obtained by the following. The reaction mixtures were prepared by combining cell crude extracts of each technical triplicate from *B. methanolicus* wild-type and Δ*upp* mutant strains with TRIS-HCl buffer (pH 7.5) and NADH. Reaction mixtures were then pre-heated ($50\,°\mathrm{C}$) to mimic cellular conditions of *B. methanolicus* before being monitored spectrophotometrically at $340\,\mathrm{nm}$, using TRIS-HCl as a blank. As NADH is a commonly used co-factor for many enzymes, the system was calibrated by monitoring the baseline of the cell crude extracts for three minutes. The reaction was then initiated by adding the substrate F6P, and the absorbances monitored for three more minutes. Cell crude extracts were diluted to assure a linear decrease in monitored absorbance, which was more suitable for calculating enzyme activities. Raw data of the monitored baselines and reactions, along with the estimated total protein contents and enzyme specific activities (U/mg) of the cell crude extracts are presented in Appendix H. Furthermore, an example calculation for calculation of enzyme specific activity is presented in Appendix F.

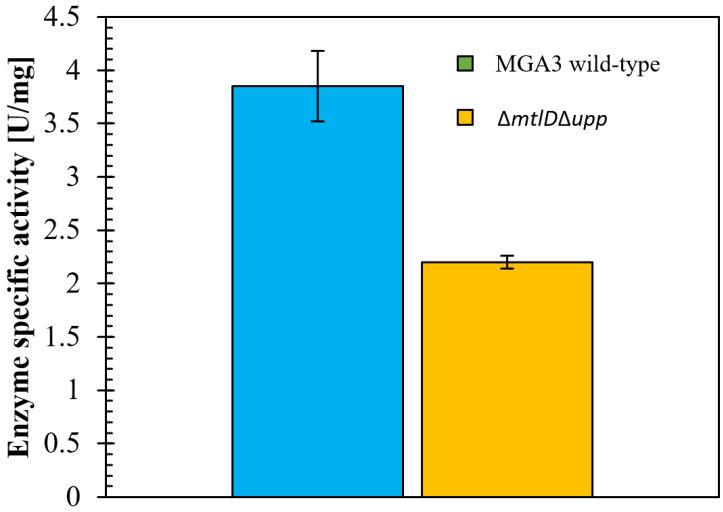


Figure 3.16: Bar chart showing the mean enzyme specific activities of *B. methanolicus* wild-type and Δ*mtlD*Δ*upp* strains cultivated in technical triplicates. Errors bars display calculated standard deviations.

The mean specific activity of the wild-type strain was calculated to $3.9 \pm 0.3\,\mathrm{U/mg}$, and $2.20 \pm 0.06\,\mathrm{U/mg}$ for the Δ*mtlD*Δ*upp* strain. The third sample for the technical triplicate of the mutant strain displayed almost 50% higher enzyme-specific activity than the remaining two samples(Appendix H), which on the other hand displayed almost the same value of approximately 2.2 U/mg. Because of this deviation, the mean enzyme-specific activity of the Δ*mtlD*Δ*upp*

strain was calculated after omitting the third sample value. The enzyme specific activity was then estimated to be significantly reduced by 43% upon deletion of *mtlD*.

3.4 Deletion of *aroE*

3.4.1 Targeting *aroE* for optimization of the established protocol

The shikimate dehydrogenase encoded by *aroE* has an essential role in the biosynthesis of aromatic amino acids and other vital biomolecules in bacteria through the shikimate pathway (Herrmann and Weaver (1999)). Shikimate dehydrogenase catalyzes the reversible reaction on the formation of shikimic acid from 3-dehydroshikimate by reduction of nicotinamide adenine dinucleotide phosphate ($NADP^+$) to NADPH. As a further contribution to the present study, the deletion of *aroE* aimed to optimize the established protocol by examining the effects of the flanking regions to the gene for the occurrence of homologous recombination.

For this purpose, three suicide vectors flanking the *aroE* gene upstream and downstream were constructed using Gibson assembly with 250 bps, 500 bps, and 1000 bps flanking regions. Specific primers were used to amplify the flanking regions from gDNA of *B. methanolicus* (See Appendix B). Plasmid maps for the three suicide vectors are shown in Appendix D. The suicide vectors were then introduced into *B. methanolicus* MGA3 by conjugation with *E. coli* S17-1 strains harboring the suicide vectors. The conjugations were done in technical triplicates with volume ratios of 9+3 ml and 4.5+1.5 ml of the recipient and donor cell volumes, respectively, and the mixed cells were dropped on non-selective SOB plates and incubated over-night ($40\,^{\circ}\mathrm{C}$). Following this, cells were resuspended in NaCl and plated on selective SOB plates supplemented with Kan ($25\,\mu\mathrm{g/ml}$) and incubated at $50\,^{\circ}\mathrm{C}$ for two days until colonies were formed.

A colony count was assumed to reflect the recombination efficiency between the introduced suicide vector and the homologous regions flanking the gene targeted for deletion. However, many of the plates were showing overgrowth. For this study's purpose, only colonies that were appearing opaque and convex were counted, thereby excluding most of the present colonies.

Figure 3.17 presents the mean counted colonies for each ratio between donor and recipient strains and the corresponding standard deviations. The figure also shows a plate of the morphology of the counted colonies (Sub-figure **(a)**). The numbers in parenthesis represent the ratio in milliliters of the recipient and donor strain, in their respective order. As the bar charts show, the number of colonies are in general over ten times as high for the conjugation made with a volume ration of 9+3 of recipient and donor cells, respectively. As shown in sub-figure **(b)** and **(c)**, the 1000 bps flanking regions results in a higher number of counted colonies for both volume ratios.

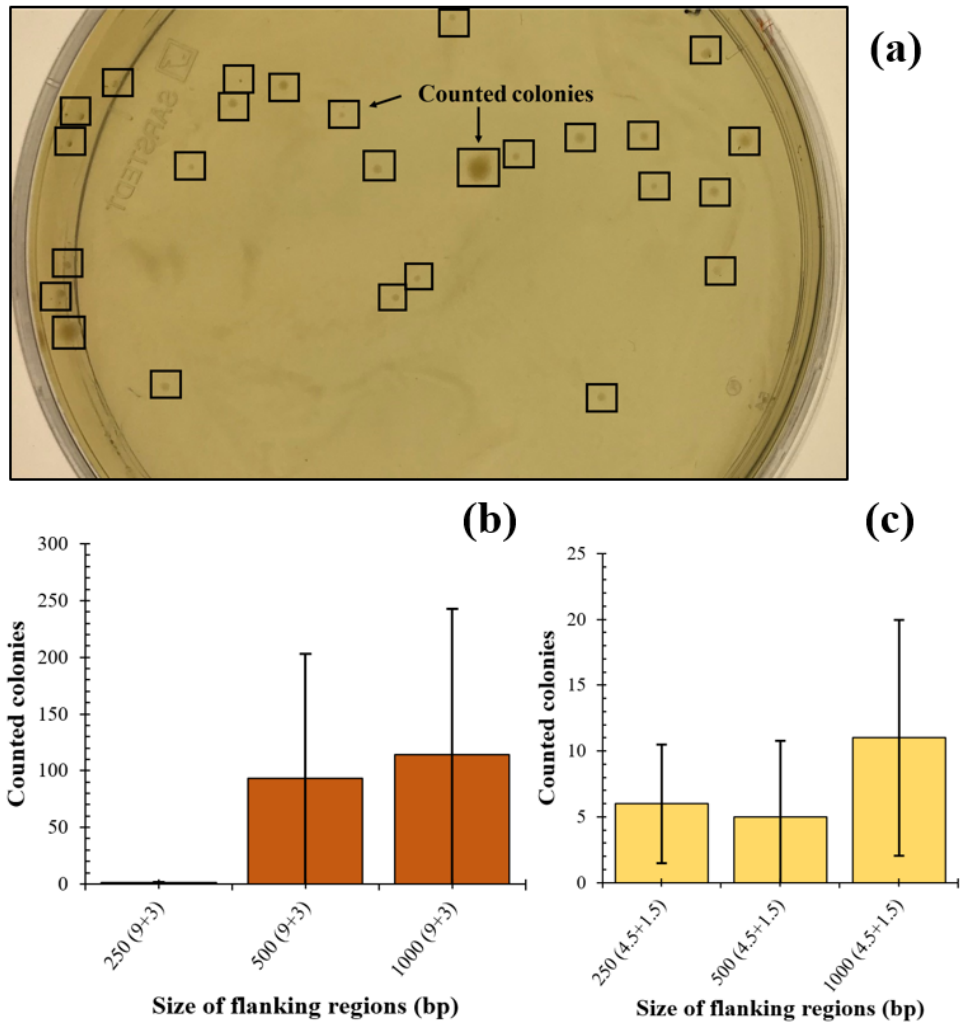


Figure 3.17: **(a)** Selective SOB plate showing the colony morphology counted for determination of flank size impact on recombination efficiency. All colonies displayed in the black boxes of sub-figure **(a)** were counted. **(b & c):** Bar charts of the number of MGA3 colonies resistant to Kan after conjugation with *E. coli* S17-1 strains carrying suicide vectors with homology flanking regions differing in size. The suicide vectors contained flanking regions of either 250 bps, 500 bps, or 1000 bps. Numbers in parenthesis correspond to the ratio between the recipient and donor strain in ml, in the respective order. The average values and standard deviations are presented.

Furthermore, the size of flanking regions seems to provide less amount of colonies for smaller sizes. Yet, the measured standard deviations are significantly large due to significant variations in the measurements. For instance, the conjugation of the suicide vector with 1000 bps flanks and 9+3 volume ratio display an average colony count of 114 with a standard deviation of 129. Similar observations are shown for the other counted plates as well. Overall, these results indicate that there is no significant difference in the recombination efficiency by varying the size of the flanking regions of *aroE*.

Chapter 4

Discussion

A standard and powerful method for determining the function of a gene in the metabolism is to study physiological changes upon its deletion. Such deletion is often accompanied by the complementation of the same particular gene to confirm whether the observed physiological changes only result from the deletion of the gene and not from alterations elsewhere in the genome. However, this approach favors a tool for creating gene deletions, which previously has not been established in *B. methanolicus*.

4.1 Protocol development for conjugation and homologous recombination-based gene deletion

The main aim of this study was to establish a novel markerless genetic tool for carrying out targeted gene deletions in the thermophilic and methylotrophic bacterium *B. methanolicus*. A tool for gene silencing by Schultenkämper et al. (2019) using the CRISPR/dCas9 system was newly established for the bacterium. However, it is limited in its use, partially due to plasmid dependency, possible off-target effects, the toxicity of dCas9 in high concentrations, and residual activity of proteins encoded by silenced genes. The establishment of a homologous recombination-based method is expected to provide more genetically stable expression hosts, by providing genome editing with permanent genetic alterations. Gene deletion via homologous recombination is a valuable tool for gene characterization in *B. methanolicus* using an reverse genetics approach, here applied for the deletion of the genes *upp* and *mtlD*.

Because the frequency of homologous recombination was expected to be low in *B. methanolicus*, it was considered important to provide a transformation protocol resulting in high quantities of cells harboring the suicide vector from the beginning. The transfer of the pDELxp-*oroP* suicide vector from the *E. coli* S17-1 strain to the *B. methanolicus* strains by conjugation displayed under the conditions of this study a high transformation efficiency, as numerous colonies

were present after the first homologous recombination event. Thus, indicating that conjugation indeed occurs at a high level. Based on this, the main focus for optimizing the protocol for gene deletions presented in this study was shifted to optimize the frequency of the homologous recombination events.

Although a 100% identity between homologous regions is not required for a recombination event to occur, it is conceivable that these regions' size may have an impact on its occurrence. Studies in *Xylella fastidiosa* (*X. fastidiosa*) and *Cryptococcus neoformans* (*C. neofromans*) (Kung et al. (2013); Nelson (2003)) have noted the importance of the size of the homology regions for efficient recombination. For *X. fastidiosa*, the optimal size was found for the range of 100-1000 bps, with an exponential increase with increased flank sizes. In *C. neoformans*, however, an optimal size of the homology regions were observed from 300 bps and above in size. Both examples serve to display that the efficiency of homologous recombination varies among bacteria and needs to be determined experimentally for each species. Therefore, three suicide vectors with varying sizes of the flanking regions (250 bps, 500 bps, 1000 bps) were constructed to investigate the smallest adequate size for homologous recombination in *B. methanolicus*. The findings of this study imply, however, no significant difference between the estimated recombination efficiencies. Due to the remarkably high calculated uncertainties, these findings should be interpreted with caution. As only the colonies who displayed a convex morphology were counted, this experiment can be proposed repeated as part of further studies. Preferentially by diluting the cell cultures before plating them out on the agar plates, and including all colonies present. An interesting extension can also determine the frequency of the second homologous recombination event resulting in the mutant genotype. However, based on this study, thoughtful suggestions can be made that using the smallest size of 250 bps flanking regions for developing suicide vectors might be preferred. Prominently, because the use of smaller cloned fragments is more convenient for PCR-related applications and DNA sequencing. The attempts of sequencing the confirmation of *mtlD* deletion in this study exemplifies this, as only for one primer (DEL39), the sequenced region was long enough to cover the target genome region.

The *oroP* counterselection system used for this tool enables the curing of the integrated plasmid and its selection and counterselection marker genes. It is hence well suited for the deletion of sequential genes. Mainly because the retainment of selection markers in the chromosomal DNA could hinder further deletions performed with the same tool using the same marker genes. Because the variety of suited selection and counterselection markers are limited, the re-use of such markers poses a great advantage with the established tool. Using the same suicide vector as a basis further simplifies the system, as the only requirements are the cloning of flanking regions into the vector to carry out gene deletions. Consequently, the pDELxp-*oroP* suicide vector presented in this study can be used to carry out sequential gene deletions in *B. methanolicus*, which were ultimately proven by the establishment of a double deletion strain of *upp* and *mtlD*.

One of the main limitations encountered for the deletion of *upp* and *mtlD* was the necessity of widespread screening of integrants that had undergone second homologous recombination, in order to identify strains harboring a deletion. This issue relates to the *oroP* counterselection system, which does not select for deletion mutants generated during the second event of homologous recombination. Consequently, the screening for deletion mutants was experienced time-consuming, as the presence of mixed strains of mutant and reverted strains also was expected. This concern was experimentally shown as only a smaller fraction of the putative gene deletion mutants was proven to be pure strains by colony PCR screening. However, this challenge was encountered by the screening of numerous colonies from a streaked plate of a mixed strain. This strategy commonly resulted in a pure strain after only a few re-streakings.

The deletion of *upp* and subsequent analysis through growth experiments on 5-FU also showed that the genetic information at the target loci could be decisive for whether a knockout is feasible without causing other genomic changes. By growth experiments, the function of the encoded protein of the mutated gene of the Δ*upp* EID strain was observed hampered, in similarity to the strain where the *upp* gene had been fully deleted. Furthermore, the growth in control conditions was hampered in comparison to the wild-type strain. To understand the reason for the observed growth behavior, possible polar effects were studied through bioinformatics tools. The reason for this is that many bacterial genes are known to be co-regulated in operons. The disruption of a gene can hence result in polar effects affecting the expression of other downstream genes. For the elucidation of the genomic neighborhood of *upp*, two putative σ^{70} promoters were predicted inside the *upp* gene sequence. However, the confirmation of these promoters could not be made by the study of growth for the Δ*upp* EID strain, as it displayed a similar growth pattern as the Δ*upp* strain.

Nevertheless, it is possible to propose that the elucidation of a gene's genomic neighborhood is an integral part of targeting a gene for deletion, as it can be used to predict polar effects. This also attributes to the construction of the suicide vectors. To prevent codon frameshift resulting in interruption of transcription by RNA polymerase, suicide vectors may be constructed to leave a short fragment of the gene in the chromosomal DNA. In this study, this was done by cloning of 15 bps upstream and downstream of the gene to be deleted as part of the flanking regions. From the DNA sequencing of *mtlD* mutants, only these in total 30 bps of the gene were remaining in the gDNA, emphasizing the feature of the established tool to provide seamless deletions. Based on this, one may assume that the occurrence of unwanted codon frameshifts will be minimal.

4.2 Elucidation of *upp* role in pyrimidine metabolism

In this project, the established gene deletion tool's versatility was emphasized by an attempt to delete three chromosomally bounded genes in *B. methanolicus*. The first of these genes,

upp encoding a uracil phosphoribosyltransferase, appears as the first successful gene deletion achieved by the method proposed herein. Based on the assumed function of UPP as a converter of exogenous uracil, the deletion-based characterization of the UPP function meant to provide insights into the uracil salvage pathway of the bacterium.

In this study, characterization of the *upp* gene is presented as a sequence of experimental procedures that ultimately confirmed the successful deleted of the gene but also provided insights into the function of this gene. This work can be divided into three main parts, the first being a genotypic confirmation of deletion using DNA sequencing. Then, a physiological confirmation of the deletion by growth experiments in minimal medium supplemented with (and without) 5-FU. Lastly, complementation of *upp*, to finally confirm that the gene was deleted and that the observed physiological changes only were resulting from the deletion of the gene.

The growth experiments conducted on the uracil analog 5-FU displayed in all cases of this study a high lethality for the wild-type strain of *B. methanolicus* MGA3, showing that just $1\,\mu\mathrm{M}$ of 5-FU had detrimental effects on its growth. The Δ*upp* strain, on the other hand, displayed a much higher resistance to 5-FU. This finding is consistent with expectations and discoveries from the bioinformatics elucidation of uracil salvage pathways in *B. methanolicus*, suggesting the importance of the encoded uracil phosphoribosyltransferase as the main converter of exogenous uracil to UMP in the organism.

Surprisingly, growth studies also revealed almost 66% reduction of the maximum OD_{600} of the Δ*upp* mutant strain compared to the wild-type. A similar finding was observed for the Δ*upp* EID strain, suggesting that inactivation of *upp* either by deletion or frameshift mutation has similar effects on the physiology of *B. methanolicus*.

Accordingly, the deletion of *upp* can be hypothesized to influence the growth of *B. methanolicus* in minimal medium, using methanol as a carbon source. By examining the genomic neighborhood of *upp*, two genes identically oriented to *upp* were found directly downstream of the gene. One of these, the *rpiB* gene encoding a ribose-5-phosphate isomerase B is involved in the regeneration part of the RuMP cycle. Interference with this gene might reduce the ability to assimilate formaldehyde and, consequently, reduce the organism's growth. The second found gene, *glyA*, encodes a serine hydroxymethyltransferase. This enzyme is associated with the synthesis of other important biomolecules, such as purines. The location of *glyA* directly downstream (220 nt) to *upp*, suggests that this gene can be affected by the *upp* deletion through unpredicted polar effects. Interestingly, a search in the region of the *upp* gene predicted two putative σ^{70} promoters. If these were found to regulate the expression of *glyA*, this could explain the observed growth pattern of the mutant strain. Besides identifying promoters, other regulatory elements, such as transcriptional binding sites, may be useful to predict whether the deletion of a gene may result in undesirable polar effects.

In contrast to these findings, the growth experiment after complementation of *upp* in the Δ*upp* mutant strain showed a different behavior with the previous results. This experiment displayed that the wild-type growth pattern was almost completely restored for the Δ*upp* strain through complementation with the plasmid-based expression of the *upp* gene, resulting in a maximum OD_{600} of 3.2 in the absence of 5-FU, compared to 3.3 for the wild-type. Besides, the sensitivity to 5-FU was significantly increased for the *upp* overexpressed strain, as the growth was observed to be significantly inhibited by 5-FU.

When comparing to the wild-type strain, a drop in OD_{600} after the first doubling is displayed (Figure 3.13) as steeper for the wild-type strain than for the *upp* overexpressed mutant strain. Furthermore, the drop in OD_{600} is higher when overexpressing *upp* in the wild-type strain, as both a plasmid-bound and chromosomal-bound *upp* gene is present. These results can indicate that the chromosomal expression of *upp* is more significant than the plasmid-based *upp* gene.

The empty vector control of the mutant strain was found to behave like the Δ*upp* strain. This strain also displayed a similar growth pattern and maximum OD_{600} of 3.0 as the wild-type for growth in the absence of 5-FU. Thereby, indicating that the growth of the *upp* deletion strain harboring the plasmid pTH1mp is not affected compared to the wild-type strain. An explanation for these contradictory results to the other growth experiments, which showed reduced growth by deletion of *upp* compared to the wild-type strain, is difficult to make without further experimental efforts. However, this may indicate that *upp* does not affect the expression of nearby genes to any significant degree.

Another important finding arises from the observation that the Δ*upp* strain was exhibiting a linear decrease of the specific growth rate by increased concentrations of 5-FU. By cultivating the Δ*upp* strain harboring pTH1mp on $5\,\mu M$ of 5-FU, the growth displayed a low specific growth rate of $0.05\,hr^{-1}$ and growth was still observed after 25 hours. These findings imply that the Δ*upp* strain holds some sensitivity to 5-FU. A previous study conducted for *Lactococcus lactis* (*L. lactis*) by Martinussen and Hammer (1994) encountered similar observations. By plating a *upp* mutant *L. lactis* on minimal plates with differing concentrations of 5-FU, Martinussen and Hammer (1994) claimed that only some small colonies were formed at high concentrations due to growth inhibition. By ^{14}C labeling uracil, they were also able to detect exogenous uracil incorporation into nucleic acids. Based on the presumed importance of the *upp* gene in the pyrimidine salvage pathway, a possible explanation of the finding could be the presence of alternative enzymes available that can convert uracil when *upp* expression is disrupted. A transcriptome analysis (Irla et al. (2015)) revealed the presence of the *pyrR* gene in the genome of *B. methanolicus*. Studies in *B. subtilis* by Turner et al. (1994) revealed that the *pyrR*-encoded enzyme to a limited degree confers uracil phosphoribosyltransferase activity. Moreover, an amino acid identity of 74.59% between the *pyrR* encoded proteins of the two *Bacillus* strains suggests that the catalytic activity is similar. Furthermore, Martinussen et al. (1995) discovered that a

double deletion strain of *upp* and *pyrR* of *B. subtilis* still posed sensitivity to 5-FU. Because of the similar findings in this study, efforts were also made for the *in silico* elucidation of the uracil salvage pathway in *B. methanolicus*.
Three possible candidates were found in addition to *pyrR*; *pdp* putatively encoding a pyrimidine-nucleoside phosphorylase, and the genes *punA* and *deoD* putatively encoding purine-nucleoside phosphorylases. The exact function of these enzymes is yet to be determined for *B. methanolicus* and to determine whether these enzymes have any significant affinity for uracil as a substrate. However, with the genetic tool developed in this study, the characterization of these genes should be possible. Furthermore, the proposed uracil salvage pathway that emerges as part of this work should constitute a good foundation for further studies on the uracil salvage pathway in *B. methanolicus*.

4.3 Deletion of *mtlD* and proof of versatility of the genetic tool

As part of this study, the deletion of *mtlD* was successfully created in the wild-type MGA3 and Δ*upp* strains of *B. methanolicus* and the deletion verified through DNA sequencing and enzyme assay. The process of obtaining pure strains harboring a *mtlD* deletion was found to be laborious, as it was difficult to obtain clean deletion strains that were not contaminated by the wild type strain. Multiple re-streakings of the mixed wild-type/Δ*upp* colonies was therefore needed. Furthermore, another issue was encountered, namely nonspecific annealing of the primers SEMF/SEMR leading to nonspecific PCR products during colony PCR. As a result, the expected band of approximately 2.5 kb had to be extracted from the agarose gel before being sent for DNA sequencing, which ultimately verified the deletion of the gene.

Only the 15 base pairs from each side of the *mtlD* gene remained in the genome, and no perturbations were found either directly upstream or downstream of the gene, according to the sequencing data. This observation is in line with expectations, as the flanking regions cloned into the pDELxp-*oroP* vector contained the same base pairs of the gene. This finding indicates that the homologous recombination events are accurate and result only in the deletion of the targeted sequence without off-target effects. This specificity serves as an essential advantage of the established gene deletion tool. Unspecificity is considered to be one of the major drawbacks of other conventional gene disruption or gene repression strategies using classical mutagenesis or the CRISPR-Cas system. Furthermore, the *mtlD* deletion in the Δ*upp* strain also shows that the method has its versatility as multiple genes can be deleted in the same strain.

Due to the limited duration of this study, an *mtlD* enzyme assay was only conducted for the Δ*mtlD*Δ*upp* strain. Because *upp* is associated with the uracil salvage pathway of *B. methano-*

licus, UPP was not considered as essential under the growth conditions for preparations of the enzyme assay. Therefore, the phenotypical observations can be expected to be equivalent to both the *mtlD* single deletion and double deletion strains. Nevertheless, due to the validity of the results shown by this enzyme assay, performing the assay for the single deletion strain of *mtlD* may be relevant as part of further studies. This issue is emphasized as one of the samples for the technical triplicate of the Δ*mtlD*Δ*upp* strain was omitted from the sampling data due to a deviation of approximately 50% from the other calculated activities.

The enzyme assay for mannitol 1-phosphate 5-dehydrogenase carried out as part of this study was based upon the assumption that the protein concentration and enzyme activity could be estimated from a cell crude extract and not from a solution of purified enzyme. Therefore, the estimated enzyme-specific activities were calculated as relative enzyme activities in reference to total protein concentration in the cell crude extracts. Because the total protein concentration can be assumed to change almost insignificantly by the deletion of *mtlD*, the *mtlD* enzyme assay results were expected to correlate well with the actual change in enzyme activity of mannitol 1-phosphate 5-dehydrogenase. Furthermore, the reaction monitored the oxidation of the coenzyme NADH to NAD^+ and H^+, occurs naturally in numerous other reactions of a cell. To account for background activities in the samples, the activity slopes of monitored absorbances for the reactions were correlated to the measured baseline slopes from monitoring the oxidation of NADH before the addition of F6P.

The enzyme specific activity of mannitol 1-phosphate 6-dehydrogenase was estimated to decrease by 43% compared to the wild-type strain, corresponding to a reduction from 3.9 ± 0.2 U/mg to 2.2 ± 0.06 U/mg. The most surprising about this finding is that the activity yet is measured considerably high, meaning that the oxidation of NADH still occurs after the deletion of *mtlD*. This result is also consistent with a similar finding by Schultenkämper et al. (2019), where the expression of *mtlD* in *B. methanolicus* was silenced using CRISPRi, resulting in a 50% decrease of the enzyme specific activity. However, the CRISPRi method for gene repression is associated with some residual activity of the *mtlD* gene. In preliminary experimental efforts by Heid (2020), a similar assay was carried out for a *B. methanolicus mtlD* deletion strain, which provided similar findings to the results of this study. Both yielding enzyme specific activities in the same magnitude and residual enzyme activity. Collectively, the remaining coinciding activity may be attributed to the presence of other enzymes that can convert fructose 6-phosphate by the oxidation of NADH, or the presence of a mixed culture of the mutant and wild-type strains. The latter assumption is, albeit less reasonable, as these findings arise from three independently performed studies and that the deletion was confirmed by sequencing.

López et al. (2019) have described the first characterization of D-arabitol as a carbon source in *B. methanolicus* and noted the importance of an arabitol phosphate dehydrogenase encoded by

the gene *atlD* (BMMGA3_RS07345) for arabitol utilization. Interestingly, in cell crude extracts obtained for cells grown on mannitol, this enzyme was found active (0.02±0.00 U/mg). In the contrarily, no activity was found from cells grown on methanol. This suggests that the expression of *atlD* may be regulated by mannitol or its derivates, hence indicating the presence of an alternative dehydrogenase for the enzyme assay conducted in this study. In *Corynebacterium glutamicum* (*C. glutamicum*), gene deletion experiments of the *mtlD* encoding a mannitol 2-dehydrogenase was shown to be essential for growth on D-mannitol and D-arabitol (Laslo et al. (2011)). Besides, López et al. (2019) did show that the complementation of *atlD* in a *mtlD* disrupted *C. glutamicum* strain displayed restored growth on arabitol. These results serve as an indication that a single dehydrogenase present can utilize both D-mannitol and D-arabitol metabolites in some bacterial species. A brief prediction of whether the arabitol phosphate dehydrogenase might have an affinity to F6P as a substrate can be made from structural similarities. Both F6P and D-arabitol 1-phosphate are pentoses containing a phosphate group and numerous polar groups. Hence, it can be speculated that the arabitol phosphate dehydrogenase also has some affinity to fructose 6-phosphate, leading to the remaining activity detected during the MtlD enzyme assay.

A broader affinity of arabitol phosphate dehydrogenase is, on the other hand, inconsistent with the characterization of the enzyme in *Enterococcus avium* (*E. avium*) by Povelainen et al. (2003), where the enzyme was shown to display a limited substrate affinity. This study pointed out only D-arabitol 1-phosphate and D-arabitol 5-phosphate as potential substrates, which results in the formation of the two products xylulose 5-phosphate and ribulose 5-phosphate, respectively. Moreover, the amino acid identity between the two arabitol dehydrogenases between *B. methanolicus* and *E. avium* was found to be 51% (Guil-López et al. (2019)), indicating some homology between the two enzymes. However, to verify whether the arabitol phosphate dehydrogenase from *B. methanolicus* has affinity to F6P, the *atlD* gene can be characterized according to the methodology in this study, using the presented tool for gene deletion and gene characterization.

As described by Schultenkämper et al. (2019), the silencing of *mtlD* using CRISPRi in *B. methanolicus* did not affect growth when using methanol as the sole carbon source, in line with expectations. Moreover, the targeting of *mtlD* did not fully halt growth on mannitol either. It was proposed that this result could have two putative reasons; there might be an alternative mannitol 1-phosphate 5-dehydrogenase present, or the repression by CRISPRi on *mtlD* was insufficient. In the latter case, quantitative reverse transcription PCR (qRT-PCR) revealed residual activity of *mtlD*. Nonetheless, these results do not exclude the presence of another dehydrogenase participating in mannitol utilization either. As part of future studies, it may be useful to study whether the Δ*mtlD* single and double deletion mutants established as part of this study can grow on either D-mannitol or D-arabitol. Such an experiment could provide information on

whether the strains are pure mutant strains and, ultimately, whether the *B. methanolicus mtlD* gene is essential for growth on these two substrates.

4.4 Recommendations for further work

The presented genetic tool for chromosomal gene deletions by homologous recombination demonstrates the feasibility of successful deletions by the deletion of the two genes *upp* and *mtlD* in *B. methanolicus*.

However, the Δ*mtlD* single deletion strain was to the extent of this study only verified using DNA sequencing. Further research efforts should proceed on investigating if similar enzyme specific activities to the *mtlD* double deletion strain can be obtained from an enzyme assay and whether the restoring of a native *mtlD* phenotype is possible by complementation studies. The latter also applies to the *mtlD* double deletion strain. Like the findings for the *upp* deletion, a restored phenotype will prove that the only genomic change by the method arises from the deletion of the gene. This study also aimed towards elucidating the role of *upp* gene in the uracil utilization pathway of *B. methanolicus*. As growth experiments showed unexpected sensitivities of an *upp* deletion strain to 5-FU, intriguing questions regarding the extent of enzymes able to take over the catalytic function of the *upp* were made. A similar assumption was made for *mtlD*, as residual activities were observed for both *upp* and *mtlD* in this study. Therefore, a quantitative reverse transcription PCR (qRT-PCR) analysis can be devoted to further studies. This experiment can reveal if the observed residual enzyme activities arise from the enzymes encoded by *upp* or *mtlD*. In combination with expression levels of other proteins in the *B. methanolicus* proteome, these results can substantiate the claim of alternative enzymes that may take over the catalytic activity of proteins encoded by the deleted genes.

At last, additional studies should also be made for the deletion of *aroE*. Firstly, as the deletion of a triple deletion strain will ultimately show that sequential deletions are possible with the established method. Secondly, as multiple suicide vectors with different sizes of the flanking regions to *aroE* were constructed in this study, further experiments should be carried out to investigate the effects of the size on the homologous recombination efficiency. Finally, as the deletion of *aroE* can contribute to emphasize the specificity of the homologous recombination effects for seamless deletion.

Chapter 5

Conclusion

This study established a tool for targeting markerless gene deletions in the chromosomal DNA of *B. methanolicus*, by the action of two homologous recombination-based events. Transfer of the suicide vector pDELxp-*oroP* from donor strain *E. coli* S17-1 to a recipient strain of *B. methanolicus* MGA3 via conjugation was shown as an efficient method for horizontal gene transfer between the two strains. Furthermore, the used *oroP* counterselection system for this tool provided both gene deletion and curing of the integrated plasmid along with the selection and counterselection marker genes, thereby allowing for the sequential deletion of genes in the same strain.

The versatility of the established tool was proven by the seamless deletion of *upp* and *mtlD*, eventually resulting in single and double mutant strains. The resulting phenotypes of the deletion strains were assayed through growth experiments and enzyme assays. In both cases, the encoded proteins of *upp* and *mtlD* displayed decreased activity. However, residual activity was also observed, indicating other enzymes present with a similar catalytic activity of the proteins encoded by the deleted genes. Moreover, verification of the deletion was ultimately shown for *upp* by plasmid-based complementation of *upp* resulting in a restored wild-type phenotype. Finally, the effect on homologous recombination efficiency was investigated by varying the flanking homology regions' size to the targeted gene. However, due to large variations in this experiment, no conclusions could be made.

Bibliography

Acevedo-Rocha, C.G., Gronenberg, L.S., Mack, M., Commichau, F.M., Genee, H.J., 2019. Microbial cell factories for the sustainable manufacturing of b vitamins. *Current Opinion in Biotechnology* **56**, 18–29.

Addgene, No date. *CRISPR Guide*. `http://www.addgene.org/guides/crispr/#Disrupt`. Source: Addgene (Accessed on 03/27/2020).

Altschul, S., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.

An, X., Zuo, Y., Zhang, Q., Wang, J., 2009. Methanol synthesis from CO2 hydrogenation with a cu/zn/al/zr fibrous catalyst. *Chinese Journal of Chemical Engineering* **17**, 88–94.

Andersen, E.M.S., 2011. *Genetiske verktøy i Bacillus methanolicus* . Master's thesis. Norwegian University of Science and Technology. http://hdl.handle.net/11250/245757.

Anthony, C., 1982. *The biochemistry of methylotrophs*. 1 ed., Academic Press. 431 pp.

Antoniewicz, M.R., 2019. Synthetic methylotrophy: Strategies to assimilate methanol for growth and chemicals production. *Current Opinion in Biotechnology* **59**, 165–174.

Aravind, L., Ponting, C.P., 1999. The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins. *FEMS Microbiology Letters* **176**, 111–116.

Arfman, N., Beeumen, J., de Vries, G., Harder, W., Dijkhuizen, L., 1991. Purification and characterization of an activator protein for methanol dehydrogenase from thermotolerant bacillus spp. *The Journal of biological chemistry* **266**, 3955–60.

Arfman, N., Dijkhuizen, L., Kirchhof, G., Ludwig, W., Schleifer, K.H., Bulygina, E.S., Chumakov, K.M., Govorukhina, N.I., Trotsenko, Y.A., White, D., Sharp, R.J., 1992a. Bacillus methanolicus sp. nov., a new species of thermotolerant, methanol-utilizing, endospore-forming bacteria. *International Journal of Systematic Bacteriology* **42**, 439–445.

Arfman, N., Hektor, H.J., Bystrykh, L.V., Govorukhina, N.I., Dijkhuizen, L., Frank, J., 1997. Properties of an NAD(h)-containing methanol dehydrogenase and its activator protein from bacillus methanolicus. *European Journal of Biochemistry* **244**, 426–433.

Arfman, N., de Vries, K.J., Moezelaar, H.R., Attwood, M.M., Robinson, G.K., van Geel, M., Dijkhuizen, L., 1992b. Environmental regulation of alcohol metabolism in thermotolerant methylotrophic bacillus strains **157**, 272–278.

Arfman, N., Watling, E.M., Clement, W., van Oosterwijk, R.J., de Vries, G.E., Harder, W., Attwood, M.M., Dijkhuizen, L., 1989. Methanol metabolism in thermotolerant methylotrophic bacillus strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme. *Archives of Microbiology* **152**, 280–288.

Arsène-Ploetze, F., Nicoloff, H., Kammerer, B., Martinussen, J., Bringel, F., 2006. Uracil salvage pathway in lactobacillus plantarum: Transcription and genetic studies. *Journal of Bacteriology* **188**, 4777–4786.

Barkan, D., Stallings, C.L., Glickman, M.S., 2011. An improved counterselectable marker system for mycobacterial recombination using galK and 2-deoxy-galactose. *Gene* **470**, 31–36.

Barrangou, R., 2013. CRISPR-cas systems and RNA-guided interference. *Wiley Interdisciplinary Reviews: RNA* **4**, 267–278.

Benninghaus, L., 2018a. *Development of tools for B. subtilis based on construction of expression vectors and screening of candidates of counter selection markers*. Project protocol, University of Bielefeld and Norwegian University of Science and Technology.

Benninghaus, L., 2018b. *Experimental analysis of markerless gene deletion and cloning vectors in thermophilic Bacillus methanolicus*. Bielefeld University, Germany. Bachelor thesis.

Berman, M.L., Beckwith, J., 1979. Use of gene fusions to isolate promoter mutants in the transfer RNA gene tyrT of escherichia coli. *Journal of Molecular Biology* **130**, 303–315.

Bhatti, A., Masood, Q.F., Khaleeq, T., Shah, A.R., 2016. homologous recombination. `https://www.britannica.com/science/homologous-recombination`. Source: Encyclopædia Britannica (Accessed on 03/30/2020).

Bhaya, D., Davison, M., Barrangou, R., 2011. CRISPR-cas systems in bacteria and archaea: Versatile small RNAs for adaptive defense and regulation. *Annual Review of Genetics* **45**, 273–297.

Borisut, P., Nuchitprasittichai, A., 2019. Methanol production via CO2 hydrogenation: Sensitivity analysis and simulation—based optimization. *Frontiers in Energy Research* **7**.

Bosma, E.F., van de Weijer, A.H.P., van der Vlist, L., de Vos, W.M., van der Oost, J., van Kranenburg, R., 2015. Establishment of markerless gene deletion tools in thermophilic bacillus smithii and construction of multiple mutant strains. *Microbial Cell Factories* **14**.

Bowater, R., Doherty, A.J., 2006. Making ends meet: Repairing breaks in bacterial DNA by non-homologous end-joining. *PLoS Genetics* **2**, e8.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248–254.

Brautaset, T., Jakobsen, Ø.M., Degnes, K.F., Netzer, R., Nærdal, I., Krog, A., Dillingham, R., Flickinger, M.C., Ellingsen, T.E., 2010. Bacillus methanolicus pyruvate carboxylase and homoserine dehydrogenase i and II and their roles for l-lysine production from methanol at 50°c. *Applied Microbiology and Biotechnology* **87**, 951–964.

Brautaset, T., Jakobsen, O.M., Flickinger, M.C., Valla, S., Ellingsen, T.E., 2004. Plasmid-dependent methylotrophy in thermotolerant bacillus methanolicus. *Journal of Bacteriology* **186**, 1229–1238.

Brautaset, T., Jakobsen, Ø.M., Josefsen, K.D., Flickinger, M.C., Ellingsen, T.E., 2007. Bacillus methanolicus: a candidate for industrial production of amino acids from methanol at 50°c. *Applied Microbiology and Biotechnology* **74**, 22–34.

Bukhtiyarova, M., Lunkenbein, T., Kähler, K., Schlögl, R., 2017. Methanol synthesis from industrial CO2 sources: A contribution to chemical energy conversion. *Catalysis Letters* **147**, 416–427.

Carnicer, M., Vieira, G., Brautaset, T., Portais, J.C., Heux, S., 2016. Quantitative metabolomics of the thermophilic methylotroph bacillus methanolicus. *Microbial Cell Factories* 15.

Chen, W., Zhang, Y., Yeo, W.S., Bae, T., Ji, Q., 2017. Rapid and efficient genome editing in staphylococcus aureus by using an engineered CRISPR/cas9 system. *Journal of the American Chemical Society* **139**, 3790–3795.

Cheng, W.H., Kung, H.H., 1994. *Methanol Production and Use (Chemical Industries*. 1 ed., CRC Press, USA.

Cho, S., Shin, J., Cho, B.K., 2018. Applications of CRISPR/cas system to bacterial metabolic engineering. *International Journal of Molecular Sciences* **19**, 1089.

Cotton, C.A., Claassens, N.J., Benito-Vaquerizo, S., Bar-Even, A., 2020. Renewable methanol and formate as microbial feedstocks. *Current Opinion in Biotechnology* **62**, 168–180.

Cui, H., Ruda, G.F., Carrero-Lérida, J., Ruiz-Pérez, L.M., Gilbert, I.H., González-Pacanowska, D., 2010. Exploring new inhibitors of plasmodium falciparum purine nucleoside phosphorylase. *European Journal of Medicinal Chemistry* **45**, 5140–5149.

Davy, A.M., Kildegaard, H.F., Andersen, M.R., 2017. Cell Factory Engineering. *Cell Systems* **4**, 262–275.

Defoor, E., Kryger, M.B., Martinussen, J., 2007. The orotate transporter encoded by oroP from lactococcus lactis is required for orotate utilization and has utility as a food-grade selectable marker. *Microbiology* **153**, 3645–3659.

van Dijl, J.M., Hecker, M., 2013. Bacillus subtilis: from soil bacterium to super-secreting cell factory. *Microbial Cell Factories* **12**, 3.

Drejer, E., Hakvåg, S., Irla, M., Brautaset, T., 2018. Genetic tools and techniques for recombinant expression in thermophilic bacillaceae. *Microorganisms* **6**, 42.

Drejer, E.B., Chan, D.T.C., Haupka, C., Wendisch, V.F., Brautaset, T., Irla, M., 2020. Methanol-based acetoin production by genetically engineered bacillus methanolicus. *Green Chemistry* **22**, 788–802.

Du, J., Shao, Z., Zhao, H., 2011. Engineering microbial factories for synthesis of value-added products. *Journal of Industrial Microbiology & Biotechnology* **38**, 873–890.

Fujita, T., Fujii, H., 2015. Applications of engineered DNA-binding molecules such as TAL proteins and the CRISPR/cas system in biology research. *International Journal of Molecular Sciences* **16**, 23143–23164.

Gaj, T., Gersbach, C.A., Barbas, C.F., 2013. ZFN, TALEN, and CRISPR/cas-based methods for genome engineering. *Trends in Biotechnology* **31**, 397–405.

Gasiunas, G., Barrangou, R., Horvath, P., Siksnys, V., 2012. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* **109**, E2579–E2586.

Gay, P., Le Coq, D., Steinmetz, M., Berkelman, T., Kado, C.I., 1985. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. *Journal of Bacteriology* **164**, 918–921.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., Smith, H.O., 2009. Enzymatic assembly of dna molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345.

Godde, J.S., Bickerton, A., 2006. The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution* **62**, 718–729.

Green, R., Rogers, E.J., 2013. Transformation of chemically competent e. coli, in: *Methods in Enzymology*. Elsevier, pp. 329–336.

Grynberg, M., 2004. NERD: a DNA processing-related domain present in the anthrax virulence plasmid, pXO1. *Trends in Biochemical Sciences* **29**, 106–110.

Guil-López, R., Mota, N., Llorente, J., Millán, E., Pawelec, B., Fierro, J., Navarro, R.M., 2019. Methanol synthesis from CO2: A review of the latest developments in heterogeneous catalysis. *Materials* **12**, 3902.

Hamilton, C.M., Aldea, M., Washburn, B.K., Babitzke, P., Kushner, S.R., 1989. New method for generating deletions and gene replacements in escherichia coli. *Journal of Bacteriology* **171**, 4617–4622.

Heggeset, T.M.B., Krog, A., Balzer, S., Wentzel, A., Ellingsen, T.E., Brautaset, T., 2012. Genome sequence of thermotolerant bacillus methanolicus: Features and regulation related

to methylotrophy and production of l-lysine and l-glutamate from methanol. *Applied and Environmental Microbiology* **78**, 5170–5181.

Heid, C., 2020. *Development of a novel method for markerless gene deletion via homologous recombination*. Praxissemesterbericht, University of Bielefeld and Norwegian University of Science and Technology.

Herrmann, K.M., Weaver, L.M., 1999. The Shikimate Pathway. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**, 473–503.

Hicks, D.B., Wang, Z., Wei, Y., Kent, R., Guffanti, A.A., Banciu, H., Bechhofer, D.H., Krulwich, T.A., 2003. A tenth atp gene and the conserved atpI gene of a bacillus atp operon have a role in mg2 uptake. *Proceedings of the National Academy of Sciences* **100**, 10213–10218.

Hille, F., Charpentier, E., 2016. CRISPR-cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150496.

Hille, F., Richter, H., Wong, S.P., Bratovič, M., Ressel, S., Charpentier, E., 2018. The biology of CRISPR-cas: Backward and forward. *Cell* **172**, 1239–1259.

Huang, W., Wilks, A., 2017. A rapid seamless method for gene knockout in pseudomonas aeruginosa. *BMC Microbiology* **17**.

Irla, M., Heggeset, T.M.B., Nærdal, I., Paul, L., Haugen, T., Le, S.B., Brautaset, T., Wendisch, V.F., 2016. Genome-based genetic tool development for bacillus methanolicus: Theta- and rolling circle-replicating plasmids for inducible gene expression and application to methanol-based cadaverine production. *Frontiers in Microbiology* **7**.

Irla, M., Nærdal, I., Brautaset, T., Wendisch, V.F., 2017. Methanol-based -aminobutyric acid (GABA) production by genetically engineered bacillus methanolicus strains. *Industrial Crops and Products* **106**, 12–20.

Irla, M., Neshat, A., Brautaset, T., Rückert, C., Kalinowski, J., Wendisch, V.F., 2015. Transcriptome analysis of thermophilic methylotrophic bacillus methanolicus MGA3 using RNA-sequencing provides detailed insights into its previously uncharted transcriptional landscape. *BMC Genomics* **16**, 73.

Irla, M., Neshat, A., Winkler, A., Albersmeier, A., Heggeset, T.M., Brautaset, T., Kalinowski, J., Wendisch, V.F., Rückert, C., 2014. Complete genome sequence of bacillus methanolicus MGA3, a thermotolerant amino acid producing methylotroph. *Journal of Biotechnology* **188**, 110–111.

Jakobsen, O.M., Benichou, A., Flickinger, M.C., Valla, S., Ellingsen, T.E., Brautaset, T., 2006. Upregulated transcription of plasmid and chromosomal ribulose monophosphate pathway genes is critical for methanol assimilation rate and methanol tolerance in the methylotrophic bacterium bacillus methanolicus. *Journal of Bacteriology* **188**, 3063–3072.

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**, 233–239.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.

Juers, D.H., Matthews, B.W., Huber, R.E., 2012. LacZ -galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Science* **21**, 1792–1807.

Kanehisa, M., 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **28**, 27–30.

Kim, D.H., Rossi, J.J., 2008. RNAi mechanisms and applications. *BioTechniques* **44**, 613–616.

Kloosterman, H., Vrijbloed, J.W., Dijkhuizen, L., 2002. Molecular, biochemical, and functional characterization of a nudix hydrolase protein that stimulates the activity of a nicotinoprotein alcohol dehydrogenase. *Journal of Biological Chemistry* **277**, 34785–34792.

Koonin, E.V., Makarova, K.S., Zhang, F., 2017. Diversity, classification and evolution of CRISPR-cas systems. *Current Opinion in Microbiology* **37**, 67–78.

Kostner, D., Peters, B., Mientus, M., Liebl, W., Ehrenreich, A., 2013. Importance of codB for new codA-based markerless gene deletion in gluconobacter strains. *Applied Microbiology and Biotechnology* **97**, 8341–8349.

Kung, S.H., Retchless, A.C., Kwan, J.Y., Almeida, R.P.P., 2013. Effects of DNA size on transformation and recombination efficiencies in xylella fastidiosa. *Applied and Environmental Microbiology* **79**, 1712–1717.

Kuzminov, A., 1999. Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda. *Microbiol. Mol. Biol. Rev.* **63**, 751–813.

Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., Qi, L.S., 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8**, 2180–2196.

Laslo, T., von Zaluskowski, P., Gabris, C., Lodd, E., Ruckert, C., Dangel, P., Kalinowski, J., Auchter, M., Seibold, G., Eikmanns, B.J., 2011. Arabitol metabolism of corynebacterium glutamicum and its regulation by AtlR. *Journal of Bacteriology* **194**, 941–955.

Leclere, V., Bechet, M., Adam, A., Guez, J.S., Wathelet, B., Ongena, M., Thonart, P., Gancel, F., Chollet-Imbert, M., Jacques, P., 2005. Mycosubtilin overproduction by bacillus subtilis BBG100 enhances the organism's antagonistic and biocontrol activities. *Applied and Environmental Microbiology* **71**, 4577–4584.

Leer, J.C., Hammer-Jespersen, K., Schwartz, M., 1977. Uridine phosphorylase from escherichia coli. physical and chemical characterization. *European Journal of Biochemistry* **75**, 217–224.

Lin, J., Wong, K.C., 2018. Off-target predictions in CRISPR-cas9 gene editing using deep learning. *Bioinformatics* **34**, i656–i663.

Liu, P., 2003. A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Research* **13**, 476–484.

Lõoke, M., Kristjuhan, K., Kristjuhan, A., 2011. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* **50**, 325–328.

López, M.G., Irla, M., Brito, L.F., Wendisch, V.F., 2019. Characterization of d-arabitol as newly discovered carbon source of bacillus methanolicus. *Frontiers in Microbiology* **10**.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J.M., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR–cas systems. *Nature Reviews Microbiology* **9**, 467–477.

Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., Horvath, P., Moineau, S., Mojica, F.J.M., Scott, D., Shah, S.A., Siksnys, V., Terns, M.P., Venclovas, Č., White, M.F., Yakunin, A.F., Yan, W., Zhang, F., Garrett, R.A., Backofen, R., van der Oost, J., Barrangou, R., Koonin, E.V., 2019. Evolutionary classification of CRISPR–cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology* **18**, 67–83.

Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2018. Classification and nomenclature of CRISPR-cas systems: Where from here? *The CRISPR Journal* **1**, 325–336.

Martinussen, J., Andersen, P.S., Hammer, K., 1994. Nucleotide metabolism in lactococcus lactis: salvage pathways of exogenous pyrimidines. *Journal of Bacteriology* **176**, 1514–1516.

Martinussen, J., Glaser, P., Andersen, P.S., Saxild, H.H., 1995. Two genes encoding uracil phosphoribosyltransferase are present in bacillus subtilis. *Journal of bacteriology* **177**, 271–274.

Martinussen, J., Hammer, K., 1994. Cloning and characterization of upp, a gene encoding uracil phosphoribosyltransferase from lactococcus lactis. *Journal of Bacteriology* **176**, 6457–6463.

Martynov, A., Severinov, K., Ispolatov, I., 2017. Optimal number of spacers in CRISPR arrays. *PLOS Computational Biology* **13**, e1005891.

Methanol-Institute, 2020. *Production of methanol*. Available from: `https://www.methanol.org/production/`. Source: Methanol Institute (Accessed on 03/18/2020).

Methanol-Institute, No date. *How is Methanol Produced*. `https://www.methanol.org/methanol-production/`. Source: Methanol Institute (Accessed on 03/22/2020).

Milstein, O.A., Bekker, M.L., 1976. Utilization of exogenous pyrimidines as a source of nitrogen by cells of the yeast rhodotorula glutinis. *Journal of Bacteriology* **127**, 1–6.

Moffatt, B.A., Ashihara, H., 2002. Purine and pyrimidine nucleotide synthesis and metabolism. *The Arabidopsis Book* **1**, e0018.

Mordkovich, N.N., Safonova, T.N., Manuvera, V.A., Veiko, V.P., Polyakov, K.M., Alekseev, K.S., Mikhailov, S.N., Popov, V.O., 2013. Physicochemical characterization of uridine phosphorylase from shewanella oneidensis MR-1. *Doklady Biochemistry and Biophysics* **451**, 187–189.

Mosberg, J.A., Lajoie, M.J., Church, G.M., 2010. Lambda red recombineering in escherichia coli occurs through a fully single-stranded intermediate. *Genetics* **186**, 791–799.

Mougiakos, I., Bosma, E.F., Weenink, K., Vossen, E., Goijvaerts, K., van der Oost, J., van Kranenburg, R., 2017a. Efficient genome editing of a facultative thermophile using mesophilic spCas9. *ACS Synthetic Biology* **6**, 849–861.

Mougiakos, I., Mohanraju, P., Bosma, E.F., Vrouwe, V., Bou, M.F., Naduthodi, M.I.S., Gussak, A., Brinkman, R.B.L., van Kranenburg, R., van der Oost, J., 2017b. Characterizing a thermostable cas9 for bacterial genome editing and silencing. *Nature Communications* **8**.

Müller, J.E.N., Litsanov, B., Bortfeld-Miller, M., Trachsel, C., Grossmann, J., Brautaset, T., Vorholt, J.A., 2014. Proteomic analysis of the thermophilic methylotroph Bacillus methanolicus MGA3. *Proteomics* **14**, 725–737.

Müller, J.E.N., Heggeset, T.M.B., Wendisch, V.F., Vorholt, J.A., Brautaset, T., 2015. Methylotrophy in the thermophilic bacillus methanolicus, basic insights and application for commodity production from methanol **99**, 535–551.

Naerdal, I., Pfeifenschneider, J., Brautaset, T., Wendisch, V.F., 2015. Methanol-based cadaverine production by genetically engineered bacillus methanolicus strains. *Microbial Biotechnology* **8**, 342–350.

NEB, No date. Star activity. `https://international.neb.com/tools-and-resources/usage-guidelines/star-activity`. (Accessed on 06/29/2020).

Nelson, D.L., Cox, M.M., Lehninger, A.L., 2017. *Lehninger principles of biochemistry*. Seventh edition ed., W.H. Freeman and Company ; Macmillan Higher Education, New York, NY : Houndmills, Basingstoke.

Nelson, R., 2003. Sequence length required for homologous recombination in cryptococcus neoformans. *Fungal Genetics and Biology* **38**, 1–9.

Ochman, H., Lawrence, J.G., Groisman, E.A., 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304.

Pawelczak, K.S., Gavande, N.S., VanderVere-Carozza, P.S., Turchi, J.J., 2017. Modulating DNA repair pathways to improve precision genome engineering. *ACS Chemical Biology* **13**, 389–396.

Peng, H., Zheng, Y., Zhao, Z., Liu, T., Li, J., 2018. Recognition of CRISPR/cas9 off-target sites through ensemble learning of uneven mismatch distributions. *Bioinformatics* **34**, i757–i765.

Perkins, J.B., Sloma, A., Hermann, T., Theriault, K., Zachgo, E., Erdenberger, T., Hannett, N., Chatterjee, N.P., II, V.W., Jr, G.R., Hatch, R., Pero, J., 1999. Genetic engineering of bacillus subtilis for the commercial production of riboflavin. *Journal of Industrial Microbiology and Biotechnology* **22**, 8–18.

Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H., Koo, B.M., Marta, E., Shiver, A.L., Whitehead, E.H., Weissman, J.S., Brown, E.D., Qi, L.S., Huang, K.C., Gross, C.A., 2016. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493–1506.

Pickar-Oliver, A., Gersbach, C.A., 2019. The next generation of CRISPR–cas technologies and applications. *Nature Reviews Molecular Cell Biology* **20**, 490–507.

Pluschkell, S.B., 1999. *Characterization and mathematical modeling of growth and glutamic acid production by bacillus methanolicus MGA3*. Ph.D. thesis. University of Minnesota. Ann Arbor, Mich.

Pluschkell, S.B., Flickinger, M.C., 2002. Dissimilation of [13c]methanol by continuous cultures of bacillus methanolicus MGA3 at 50 °c studied by 13c NMR and isotope-ratio mass spectrometry. *Microbiology* **148**, 3223–3233.

Povelainen, M., Eneyskaya, E.V., Kulminskaya, A.A., Ivanen, D.R., Kalkkinen, N., Neustroev, K.N., Miasnikov, A.N., 2003. Biochemical and genetic characterization of a novel enzyme of pentitol metabolism: d-arabitol-phosphate dehydrogenase. *Biochemical Journal* **371**, 191–197.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., Lim, W.A., 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183.

Reyrat, J.M., Pelicic, V., Gicquel, B., Rappuoli, R., 1998. Counterselectable markers: Untapped tools for bacterial genetics and pathogenesis. *Infection and Immunity* **66**, 4011–4017.

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**.

Sánchez-Díaz, Á., 2017. MefCO2 – synthesis of methanol from captured carbon dioxide using surplus electricity (EU-h2020). *Impact* **2017**, 6–8.

Sander, J.D., Joung, J.K., 2014. CRISPR-cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32**, 347–355.

Saunders, P.P., Wilson, B.A., Saunders, G.F., 1969. Purification and comparative properties of a pyrimidine nucleoside phosphorylase from bacillus stearothermophilus. *Journal of Biological Chemistry* **244**, 3691–3697. Publisher: American Society for Biochemistry and Molecular Biology.

Schendel, F.J., Bremmon, C.E., Flickinger, M.C., Guettler, M., Hanson, R.S., 1990. L-lysine production at 50 degrees c by mutants of a newly isolated and characterized methylotrophic bacillus sp. **56**, 963–970.

Schultenkämper, K., Brito, L.F., López, M.G., Brautaset, T., Wendisch, V.F., 2019. Establishment and application of CRISPR interference to affect sporulation, hydrogen peroxide detoxification, and mannitol catabolism in the methylotrophic thermophile bacillus methanolicus. *Applied Microbiology and Biotechnology* **103**, 5879–5889.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., Koonin, E.V., 2015. Discovery and functional characterization of diverse class 2 CRISPR-cas systems. *Molecular Cell* **60**, 385–397.

Singh, V., Brecik, M., Mukherjee, R., Evans, J.C., Svetlíková, Z., Blaško, J., Surade, S., Blackburn, J., Warner, D.F., Mikušová, K., Mizrahi, V., 2015. The complex mechanism of antimycobacterial action of 5-fluorouracil. *Chemistry & Biology* **22**, 63–75.

Snustad, D.P., Simmons, M.J., 2012. *Genetics*. sixth ed., John Wiley & Sons, Inc., Singapore. International Student Version.

Solovyev, V., 2011. *Automatic Annotation of Microbial Genomes and Metagenomic Sequences* . Nova Science Publishers. pp. 61–78. Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies (Ed. R.W. Li).

Swttzer, R.L., Turner, R.J., Lu, Y., 1998. *Regulation of the Bacillus subtilis Pyrimidine Biosynthetic Operon by Transcriptional Attenuation: Control of Gene Expression by an mRNA-Binding Protein*, in: Progress in Nucleic Acid Research and Molecular Biology. Elsevier, pp. 329–367.

Tang, L., 2019. Exploring class 1 CRISPR systems. *Nature Methods* **16**, 1079–1079.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., Ostell, J., 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research* **44**, 6614–6624.

ThermoFisher, S.I., 2017. NanoDrop One User Guide. ThermoFisher. https://www.thermofisher.com/order/catalog/product/ND-ONE-W/ND-ONE-W. Revision A.

Turnbough, C.L., Switzer, R.L., 2008. Regulation of pyrimidine biosynthetic gene expression in bacteria: Repression without repressors. *Microbiology and Molecular Biology Reviews* **72**, 266–300.

Turner, R.J., Lu, Y., Switzer, R.L., 1994. Regulation of the bacillus subtilis pyrimidine biosynthetic (pyr) gene cluster by an autogenous transcriptional attenuation mechanism. *Journal of Bacteriology* **176**, 3708–3722.

Vento, J.M., Crook, N., Beisel, C.L., 2019. Barriers to genome editing with CRISPR in bacteria. *Journal of Industrial Microbiology & Biotechnology* **46**, 1327–1341.

de Vries, G.E., Arfman, N., Terpstra, P., Dijkhuizen, L., 1992. Cloning, expression, and sequence analysis of the bacillus methanolicus c1 methanol dehydrogenase gene. *Journal of Bacteriology* **174**, 5346–5353.

Wang, Y., Zhang, Z.T., Seo, S.O., Choi, K., Lu, T., Jin, Y.S., Blaschek, H.P., 2015. Markerless chromosomal gene deletion in clostridium beijerinckii using CRISPR/cas9 system. *Journal of Biotechnology* **200**, 1–5.

Wensing, L., Sharma, J., Uthayakumar, D., Proteau, Y., Chavez, A., Shapiro, R.S., 2019. A CRISPR interference platform for efficient genetic repression in candida albicans. *mSphere* **4**.

Westra, E.R., Semenova, E., Datsenko, K.A., Jackson, R.N., Wiedenheft, B., Severinov, K., Brouns, S.J.J., 2013. Type i-e CRISPR-cas systems discriminate target from non-target DNA through base pairing-independent PAM recognition. *PLoS Genetics* **9**, e1003742.

Wiedenheft, B., Sternberg, S.H., Doudna, J.A., 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338.

Wigmore, P.M., Mustafa, S., El-Beltagy, M., Lyons, L., Umka, J., Bennett, G., 2010. *Effects of 5-FU*, in: Chemo Fog. Springer New York, pp. 157–164.

Xu, K., Segal, D.J., Zhang, Z., 2020. Editorial: Precise genome editing techniques and applications. *Frontiers in Genetics* **11**.

Yao, R., Liu, D., Jia, X., Zheng, Y., Liu, W., Xiao, Y., 2018. CRISPR-cas9/cas12a biotechnology and application in bacteria. *Synthetic and Systems Biotechnology* **3**, 135–149.

Yi, J., Lee, J., Sung, B.H., Kang, D.K., Lim, G., Bae, J.H., Lee, S.G., Kim, S.C., Sohn, J.H., 2018. Development of bacillus methanolicus methanol dehydrogenase with improved formaldehyde reduction activity. *Scientific Reports* **8**.

Yu, H., Tang, H., Xu, P., 2014. Green strategy from waste to value-added-chemical production: efficient biosynthesis of 6-hydroxy-3-succinoyl-pyridine by an engineered biocatalyst. *Scientific Reports* **4**.

Zeldes, B.M., Keller, M.W., Loder, A.J., Straub, C.T., Adams, M.W.W., Kelly, R.M., 2015. Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Frontiers in Microbiology* **6**.

Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S., Yang, S.H., 2015. Off-target effects in CRISPR/cas9-mediated genome engineering. *Molecular Therapy - Nucleic Acids* **4**, e264.

Zyl, W.F.V., Dicks, L.M.T., Deane, S.M., 2019. Development of a novel selection/counter-selection system for chromosomal gene integrations and deletions in lactic acid bacteria. *BMC Molecular Biology* **20**.

Appendix A: Media and solutions

The media and solutions used for the purpose of this study are presented below. When not stated otherwise, media and solutions were sterilized by autoclavation. Agar plates were made by supplementing $15\,\mathrm{g}$ of agar to either LB or SOB medium before autoclaving. Autoclavations were carried out at $121\,^{\circ}\mathrm{C}$ for 20 minutes.

LB/LA medium

Component	Mass (g/L)	Volume (ml/L)
Bacto tryptone	10	
Yeast extract	5	
NaCl	10	
Agar (for making of LA)	15	
Ion free water		To 1000 ml

SOB medium

Component	Mass (g/L)	Volume (mL)
Difco SOB medium (Cat. No. 244310)	28	
Ion free water		To 1000

SOBsuc medium

Component	Mass (g/L)	Volume (ml/L)
SOB medium	28	
Sucrose	85	
Ion free water		To 1000

Psi-medium

Component	Mass (g/L)	Volume (ml/L)
Bacto tryptone	20	
Yeast extract	5	
$MgSO_4$ x $7H_2O$	10.2	
Ion free water		To 1000 ml

TFB1 and TFB2 buffers

TFB1:

Component	Mass (g/200 ml)	Volume (ml/200 ml)
CH_3COOK	0.59	
RbCl	2.42	
$CaCl_2$	0.29	
$MnCl_2$	2.0	
Glycerol		30
Ion free water		To 200

pH was adjusted to 5.8 using diluted acetic acid, followed by sterile filtration.

TFB2:

Component	Mass (g/100 ml)	Volume (ml/ 100 ml)
MOPS	0.21	
RbCl	0.12	
$CaCl_2$	1.1	
Glycerol		15
Ion free water		To 100

pH was adjusted using NaOH to 6.5, followed by sterile filtration.

MvCM High Salt Buffer 10x

Component	Concentration (M)	Mass (g/500 ml)	Volume (ml/500 ml)
K_2HPO_4	0.1175	20.5	
NaH_2PO_4	0.054	7.5	
$(NH_4)_2SO_4$	0.08	10.6	
Ion free water			To 500 ml

MVcM/MVcMY Medium

Component	Volume (ml/L)	Mass (g/L)
MvCM High Salt Buffer 10x	100	
Yeast Extract		0.250
Ion free water	890	
Vitamins	0.1	
Trace metals	0.1	
Methanol	0.811	
MgSO4	0.1	
Antibiotics, when necessary		

Trace metals stock solution

Component	Concentration (mM)	Volume (ml/L)	Mass (g/L)
$FeSO_4$ x $7H_2O$	20		5.56
$CuCl_2$ x $2H_2O$	0.16		0.03
$CaCl_2$ x $2H_2O$	50		7.35
$CoCl_2$ x $6H_2O$	0.17		0.04
$MnCl_2$ x $4H_2O$	50		9.90
$ZnSO_4$ x $7H_2O$	1.00		0.29
7 Na_2MoO_4 x $2H_2O$	0.20		0.05
H_3BO_3	0.50		31
High molarity HCl	199	80	72.72
Ion free water		To 1000	

Mixture was sterilized by sterile filtration.

1M $MgSO_4$ stock solution

6.15 g of $MgSO_4\,x\,7\,H_2O$ ($246.47\,\mathrm{g/L}$ was mixed with water (25 ml) to give a 1M $MgSO_4$ solution.

$$25\,\mathrm{ml} \rightarrow \frac{246.47\,\mathrm{g/L}}{1000\,\mathrm{ml/L}} \cdot 25\,\mathrm{ml} \approx 6.15\,\mathrm{g}.$$

Electroporation buffer (EPB)

Components	Mass (g/250 ml)	Volume (ml/250 ml)
HEPES buffer	0.06	
PEG8000 buffer	62.5	
Ion free water		To 250

Solution pH adjusted to pH 7.0 using HCl or NaOH and sterilie filtrated.

1M Mannitol stock solution

Component	Mass (g/50 ml)	Volume (ml/50 ml)
D-mannitol	9.11	
Ion free water		To 50

50 mM TRIS-HCl solution

Component	Mass (g/L)	Volume (ml/L)
Tris	6.06	
Ion free water		To 1000

pH was adjusted to 7.5 using HCl, and sterile filtrated.

5 mM NADH solution

Component	Mass (g/L)	Volume (ml/L)
NADH (Sigma-Aldrich)	0.018	
Tris-HCl buffer, pH 7.5		5

The NADH was made immediately prior to use, and was stored on ice until use.

200 mM Fructose-6-Phosphate

Component	Mass (g/2 ml)	Volume (ml/2 ml)
Fructose 6-phosphate	0.12	
Tris-HCl buffer, pH 7.5		2

50 mg/ml kanamycin stock solution

A stock solution of kanamycin was prepared by dissolving kanamycin sulphate ($582.6\,\mathrm{g/mol}$) in deionized water to a final concentration of $50\,\mathrm{mg/ml}$, followed by sterile filtration. The stock solutions were stored at $-20\,^{\circ}\mathrm{C}$ for later use.

25 mg/ml chloramphenicol stock solution

A stock solution of chloramphenicol was made by dissolving $25\,\mathrm{mg}$ of chlroamphenicol salt ($323.13\,\mathrm{g/mol}$) in $1\,\mathrm{ml}$ of 96% ethanol. The solution was mixed until dissolved and sterile filtrated. Aliquotes stored at -20 $^{\circ}\mathrm{C}$ for later use.

Appendix B: List of primers

The primers used in the present study are listed below. ”Fw“ and ”Rev“ are abbreviations for forward and reverse primers, respectively.

Table 1: List of primers used in the present study.

Primer name	Sequence (5’-3’)	Description
DEL47	TTTCCTCCCAGATGTTTCAAATC CCCTATAACCCCAAATAA	Fw amplification of right flanking region to *aroE*
DEL48	TTACGGTTCCTGGCCACTAGGCG CTGTTTTCGATTAAA	Rev amplification of right flanking region to *aroE* (1000 bps downstream)
DEL49	TTACGGTTCCTGGCCACTAGTAC CGCTTGATGATGGTAAAT	Rev amplification of right flanking region to *aroE* (500 bps downstream)
DEL50	TTACGGTTCCTGGCCACTAGCAA CAACCATGTAGGTGAAAT	Rev amplification of right flanking region to *aroE* (250 bps downstream)
DEL51	ACCGGCGCATCAAGCCCGCCGA CGCCTAAAGTATGCTGATAA	Fw amplification of left flanking region to *aroE* (1000 bps upstream)
DEL52	TTGAAACATCTGGGAGGAAA	Rev amplification of left flanking region
DEL53	ACCGGCGCATCAAGCCCGCCGA GCAATCAGAAGATGTCCATA	Fw amplification of left flanking region to *aroE* (500 bps upstream)
DEL54	ACCGGCGCATCAAGCCCGCCGA ATGATTTGAACGATGTCTGCACCTGTA	Fw amplification of left flanking region to *aroE* (250 bps upstream)
DEL55	TTTATAATAATTTCCCGCATTTC	*aroE* confirmation of deletion fw upstream primer
DEL56	TTAACATGGGAGGATTAATTTG	*aroE* confirmation of deletion rev downstream primer
DEL57	TAAACAATTACATAAATAGGAGGTAGTACAT ATGGCAAAAGTTTATGTTTTTGATCATCCAC	Fw amplification of *upp*
DEL58	TAGACCTATGGCGGGTACCATAT GTTATTTCGTTCCAAATAATCGGTCTCCT	Rev amplification of *upp*
SEMF	ACTTTGGCTCGGCTTTATTG	*mtlD* confirmation of deletion fw upstream primer
SEMR	CACTTCTGTAATAAGATAAC	*mtlD* confirmation of deletion rev downstream primer
DEL29	CTCAGGAAAATGTGGGGTATG	*upp* confirmation of deletion fw upstream primer
DEL30	TCAACTTCAGCGGAGTTCA	*upp* confirmation of deletion rev downstream primer
VPJF	TCTAATCCTTCTAAAAAATATAAT TTAGAAAACTAAG	Fw primer pTH1mp
VPJR	GGTGCGGGCCTCTTCGCTATTACG	Rev primer pTH1mp
DEL39	CGAAGTGAATATGTCGTACATC	Fw primer for sequencing of *mtlD*
DEL40	TTCCGCTGCAGGATTTAATG	Rev primer for sequencing of *mtlD*

Appendix C: DNA ladder

The 1 Kb Plus DNA Ladder (Thermo Fisher Scientific) includes 15 DNA fragments of different sizes, used as reference to determine sizes of applied DNA to an agarose gel (0.8%) after gel electrophoresis.

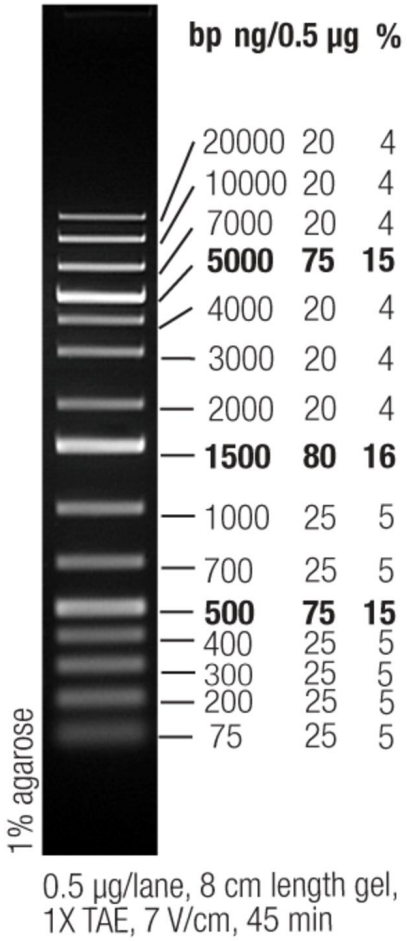


Figure 1: 1 KB Plus DNA Ladder were used as reference for estimations of fragment sizes of DNA subjected to gel electrophoresis. Picture obtained from Thermo Fisher Scientific.

Appendix D: Plasmid maps

Complementation plasmids

The complementation plasmid pTH1mp-*upp* for complementation of *upp* in Δ*upp B. methanolicus* MGA3 strain was constructed *in silico* using Clone Manager 9 (Sci-Ed). The visualization of the plasmid was made in *Benchling* [Biology software] (`https://benchling.com/`), and is shown in Figure 2. The figure shows the *upp* gene downstream of the *mdh* promoter, besides the chlormaphenicol resistance gene (abbreviated Cm-R).

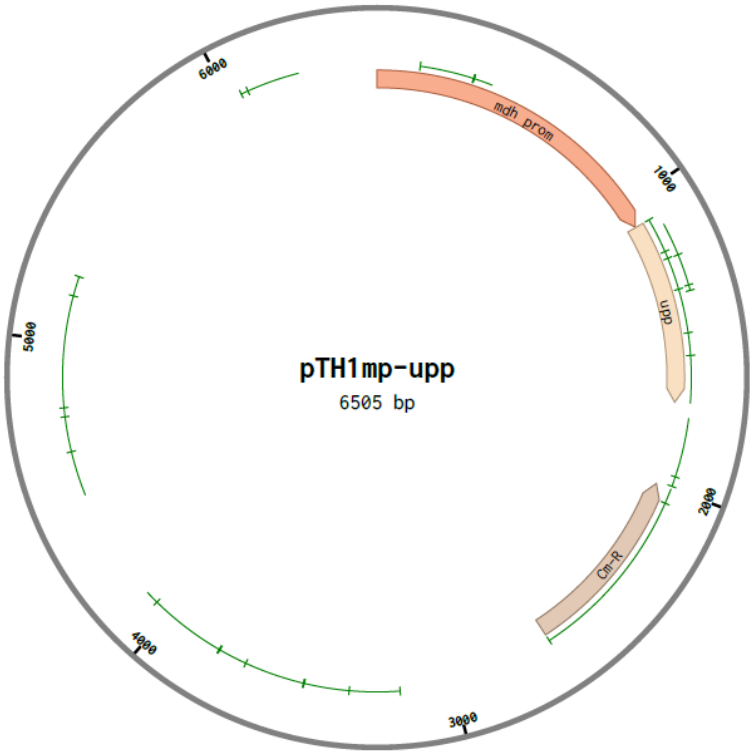
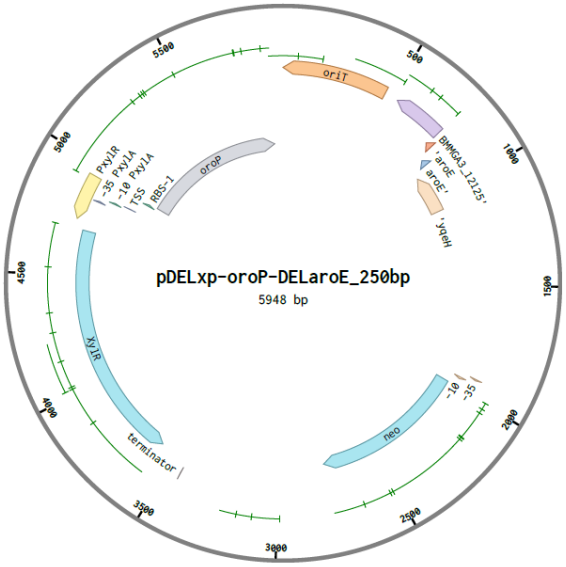


Figure 2: Plasmid pTH1mp-*upp*. *upp* gene is under the control of the methanol dehydrogenase promoter (MDH). Chloramphenicol (Cm-R) selection marker gene.

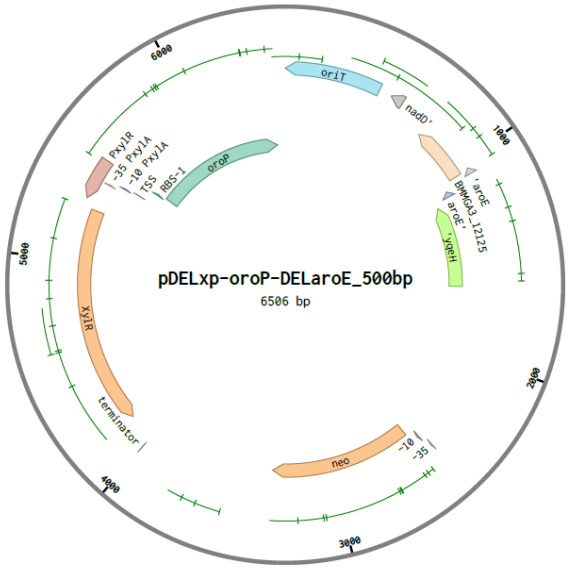
Suicide vectors

Plasmid maps of suicide vectors created for the deletion of *aroE* are presented in Figure 3.

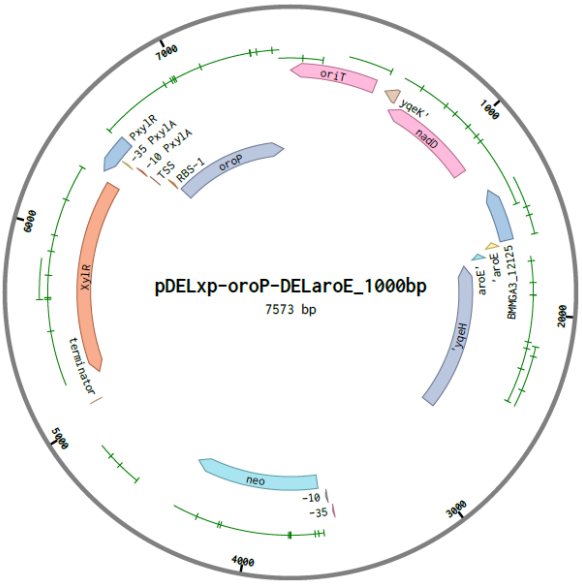
Suicide vectors for deletion of *aroE*



(a) pDELxp-*oroP*-Δ*aroE*250.



(b) pDELxp-*oroP*-Δ*aroE*500.



(c) pDELxp-*oroP*-Δ*aroE*1000.

Figure 3: Constructed suicide vectors of pDELxp-*oroP* for deletion of *aroE*. **(a)** displays the suicide vector with 250 bps homologous flanking regions to *aroE*, **(b)**, the corresponding suicide vector with 500 bps flanks, and **(c)** the suicide vector with 1000 bps flanks.

Appendix E: DNA sequencing results

All multi-alignments presented in this thesis were made in Clone Manager 9 (Sci-Ed Software), and are presented in 5'-3' direction. The template sequence is in all alignments presented as the upper line, while the query sequences are presented below. Numbers between the sequence names and nucleotide sequences correspond to their first nucleotide number in the alignment for each sequence. Green highlights correspond to a match in nucleotide between the query sequences and the template sequence. Hyphens (-) indicate a mismatch between the sequences, or no nucleotide in the query sequence.

Sequence of frameshift mutation in *upp* from selective adaptation on 5-FU

Preliminary research efforts by Drejer (unpublished) on selective plating of *B. methanlicus* MGA3 on plates with 10 µM 5-FU gave a mutation in nucleotide 6 (A) inframe of *upp* from the first ATG start codon in the sequence below. The mutation results in a frameshift, eventually making codon 11 a stop codon (TAA). "N" corresponds to amino acids that were unidentifiable.

NNNNNNANNNTTGCCATGGCAAAGTTTATGTTTTTGATCA
TCCACTAATTCAGCACAAACTTACATATATTCGGGATAAA
AATACAGGTACGAAGGAATTCCGCGAACTTGTCGATGAAG
TCGCAACATTAATGGCATTTGAAATCACCCGCGACATGCC
TCTTGAAGAAATTGAAATTGAAACACCAGTAAGCTCAACA
AAATCAAAAGTATTATCCGGCAAAAAACTAGGAATCGTTC
CGATCTTGCGGGCAGGCATCGGAATGGTGGACGGCATATT
AAAACTGATTCCAGCTGCAAAAGTTGGTCATATCGGTCTA
TATCGCGATCCTGAAACTTTAAAGCCGGTTGAATATTATG
TAAAGCTTCCGAGCGATGTAGAAGAAAGAGACTTCATCGT
CGTCGACCCAATGCTGGCAACAGGCGGCACTGCCGTTGAA
GCGATAAACTCACTGAAAAAACGCGGCGCAAAACACATTA
AATTCATGTGCCTCATTGCAGCTCCTGAAGGAGTAGAAGT
CGTAAAGAAAGCCCATCCGGACGTAGATATTTATATCGCA
GCCCTTGATGAAAAACTAAATGACCACGGCTATATCGTAC
CTGGACTTGGAGATGCAGGAGACCGATTATTTGGAACGAA
ATAAGGAAAAAGAAGGTTAATTTTAAAAACAGAAAGCCGG
CCTATAGGGGGGTCGGNAAAA

DNA multi-alignment of confirmation of *upp*

Multi-alignment of amplicons resulting from PCR amplification of Δ*upp B. methanolicus* strain using primers DEL29/DEL30. ””upp_deletion“ corresponds to the template sequence and displays the gDNA with only the 15 bps of each end of the *upp* gene remaining. Query sequences ”FwdDELupp“ and ”RevDELupp“ correspond to the forward and reverse sequencing results.

```
upp deleted          961 atgagctgatttcgaatgggttttgcagacagggcatttccatgttccgaaaatacggat
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1021 catcgggagagaatgacatttggggcattgaacacctttgataatttccgaactgttaat
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1081 tccaaatgtctgaaggatctcaggaaaatgtggggtatgttctcttaataaaagctgatt
02JB41_FwdDELupp       1 ----------------------------------------------------aaa-ctga-t
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1141 tatttttggatttctttgttactaattatttccttattatacttaagttcaaattctgc
02JB41_FwdDELupp       9 tatttttggatttctttgttactaattatttccttattatacttaagttcaaattctgc
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1201 aatcctttcaagcagttggtcgacaggaataattctttgaaagatttgttgattcccctc
02JB41_FwdDELupp      69 aatcctttcaagcagttggtcgacaggaataattctttgaaagatttgttgattcccctc
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1261 tgttgttttcaaaatggttttcgtttggctgattacgacaagatattcaattggcaatat
02JB41_FwdDELupp     129 tgttgttttcaaaatggttttcgtttggctgattacgacaagatattcaattggcaatat
02JB42_RevDELUpp    1145 ------------aatgg-------------------------------------------

upp deleted         1321 gttaatttttcgttttcccagccatacttttagctgctgtttctggcgtttcgcctgact
02JB41_FwdDELupp     189 gttaatttttcgttttcccagccatacttttagctgctgtttctggcgtttcgcctgact
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1381 aagagggtttggaaatccttcttcgttattttcaattatccgaattagctgtttagaata
02JB41_FwdDELupp     249 aagagggtttggaaatccttcttcgttattttcaattatccgaattagctgtttagaata
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1441 tttgtctaaaaacaaagttccggaatagtttttcacttctataatgagggcaaaacttgg
02JB41_FwdDELupp     309 tttgtctaaaaacaaagttccggaatagtttttcacttctataatgagggcaaaacttgg
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1501 agaaagaatgagagcatcaatttgaaaaggctgatcatgcaatagacgcagatcaaaaaa
02JB41_FwdDELupp     369 agaaagaatgagagcatcaatttgaaaaggctgatcatgcaatagacgcagatcaaaaaa
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1561 aatataatatttatttttcatttaaatacttgagatagtaagcgagttcttcttcgccacg
02JB41_FwdDELupp     429 aatataatatttatttttcatttaaatacttgagatagtaagcgagttcttcttcgccacg
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1621 aaaaccggctgttcgcttattttgatcatcttctatttttcgttctcttatgatgatcttt
02JB41_FwdDELupp     489 aaaaccggctgttcgcttattttgatcatcttctatttttcgttctcttatgatgatcttt
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1681 aggaattcttcttaataatgcttttgtttgtttaattctgatcggttcttgaagttcttt
02JB41_FwdDELupp     549 aggaattcttcttaataatgcttttgtttgtttaattctgatcggttcttgaagttcttt
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         1741 tagaatcaatataaccactccttttatagttatttcgtatttaaattctatgtgggtttct
02JB41_FwdDELupp     609 tagaatcaatataaccactccttttatagttatttcgtatttaaattctatgtgggtttct
02JB42_RevDELUpp         ------------------------------------------------------------
```

```
upp deleted        1801 atgattcctgcaaattattaaaatgatttaccggctaaagttctttttggatccaggttt
02JB41_FwdDELupp    669 atgattcctgcaaattattaaaatgatttaccggctaaagttctttttggatccaggttt
02JB42_RevDELUpp        ------------------------------------------------------------

upp deleted        1861 caagctcttaaacgaaatttatgtttttatcgaccggttttggaatttttggaccagttt
02JB41_FwdDELupp    729 caagctcttaaacgaaatttatgtttttatcgaccggttttggaatttttggaccagttt
02JB42_RevDELUpp        ------------------------------------------------------------

upp deleted        1921 tttgatttataaatcgaccacttttgggatttatcgaccagtttttcgttttatcaacca
02JB41_FwdDELupp    789 tttgatttataaatcgaccacttttgggatttatcgaccagtttttcgttttatcaacca
02JB42_RevDELUpp        ------------------------------------------------------------

upp deleted        1981 ctttccgggtcttatcgaccagtttttcgttttatcaaccactttccggattatatcaac
02JB41_FwdDELupp    849 ctttccgggtcttatcgaccagtttttcgttttatcaaccactttccggattatatcaac
02JB42_RevDELUpp        ------------------------------------------------------------

upp deleted        2041 cactttctttttatcgaccactttccgagtcttatcgaccactttcttttatagtccact
02JB41_FwdDELupp    909 cactttctttttatcgaccactttccgagtcttatcgaccactttcttttatagtccact
02JB42_RevDELUpp        ------------------------------------------------------------

upp deleted        2101 ttcaaatttattaaacttgttattaaattgttataaaaaattaaagttaataagccgacc
02JB41_FwdDELupp    969 ttcaaatttattaaacttgttattaaattgttataaaaaattaaagttaataagccgacc
02JB42_RevDELUpp   1140 --------------------------------tat-aaaaattaaagttaataagccgacc

upp deleted        2161 catgataggccggctttctgtttttaaaattaaccttctttttccttatttcgttccaaa
02JB41_FwdDELupp   1029 catgataggccggctttctgtttttaaaattaaccttctttttccttatttcgttccaaa
02JB42_RevDELUpp   1112 catgataggccggctttctgtttttaaaattaaccttc-ttttccttatttcgttccaaa

upp deleted        2221 ataaacttttgccatggcaatctctccttttttctgtttcataacacgctgttgatgaaaa
02JB41_FwdDELupp   1089 ataaacttttgccatggcaatctctccttttttctg-ttcataacacgctgttgatg-aaa
02JB42_RevDELUpp   1053 ataaacttttgccatggcaatctctccttttttctgtttcataacacgctgttgatgaaaa

upp deleted        2281 aaatacgaactacacacttcgtctaattttacagaaaaaaagtgcatagagcaacaagga
02JB41_FwdDELupp   1147 aaatacgaactac-cacttcgtctaattttacaaaaaaaaagtgca--------------
02JB42_RevDELUpp    993 aaatacgaactacacacttcgtctaattttacagaaaaaaagtgcatagagcaacaagga

upp deleted        2341 gcaaaaactgtaatcgaaacgtcaaaatcgatattgctttaccgctgtaaagtatgaggg
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    933 gcaaaaactgtaatcgaaacgtcaaaatcgatattgctttaccgctgtaaagtatgaggg

upp deleted        2401 ccagttcctcaatgcttgaaagcattgaggaaccggccccgtcgagtgaacatattagag
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    873 ccagttcctcaatgcttgaaagcattgaggaaccggccccgtcgagtgaacatattagag

upp deleted        2461 ttctggatataaagtaaatttgctggttaaagcttcaacacgctttcttgcttcatcaag
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    813 ttctggatataaagtaaatttgctggttaaagcttcaacacgctttcttgcttcatcaag

upp deleted        2521 ctttgcttcgtcttcatggttcttaagggtaaagccgataatcgacgcaatttcgtccat
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    753 ctttgcttcgtcttcatggttcttaagggtaaagccgataatcgacgcaatttcgtccat

upp deleted        2581 ttcttcaagaccgaaaccgcgggatgttacagccggtgtacctatgcggataccgcttgt
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    693 ttcttcaagaccgaaaccgcgggatgttacagccggtgtacctatgcggataccgcttgt

upp deleted        2641 tacaaacgggctttcaggatcaaacggaatggtgtttttgtttacagttattccaatttc
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    633 tacaaacgggctttcaggatcaaacggaatggtgtttttgtttacagttattccaatttc

upp deleted        2701 atcaagaacttttttctgccactttttccggttaagcctagtgaacgaacatcgacaagaag
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    573 atcaagaacttttttctgccactttttccggttaagcctagtgaacgaacatcgacaagaag

upp deleted        2761 taagtggttgtcagttcctccggaaacaaggttaagtccttctttctgaagtccttctgc
02JB41_FwdDELupp        ------------------------------------------------------------
02JB42_RevDELUpp    513 taagtggttgtcagttcctccggaaacaaggttaagtccttctttctgaagtccttctgc
```

```
upp deleted         2821 aagccgcttcgcgttggcaaccacatttttcgcatattctttaaagctgtcctgcaatac
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     453 aagccgcttcgcgttggcaaccacatttttcgcatattctttaaagctgtcctgcaatac

upp deleted         2881 ttcgccaagtgcaacggctttagccgcaattacatgcatgagcggaccgccttgaatgcc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     393 ttcgccaagtgcaacggctttagccgcaattacatgcatgagcggaccgccttgaatgcc

upp deleted         2941 agggaaaattgatttatcaatcttctgtgcaaactcttgtttgcaaaggatcattccgcc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     333 agggaaaattgatttatcaatcttctgtgcaaactcttgtttgcaaaggatcattccgcc

upp deleted         3001 gcgcggaccgcgcaatgttttgtgggttgtagttgtcacaaaatcagcatatgggactgg
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     273 gcgcggaccgcgcaatgttttgtgggttgtagttgtcacaaaatcagcatatgggactgg

upp deleted         3061 gctttgatgcaggcctgctgccactaaaccggcaatatgggccatatcgaccatcaaata
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     213 gctttgatgcaggcctgctgccactaaaccggcaatatgggccatatcgaccatcaaata

upp deleted         3121 agctccgacttcatcggcaatttcccggaatcgtttaaagtcaatttcacgaggataagc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp     153 agctccgacttcatcggcaatttcccggaatcgtttaaagtcaatttcacgaggataagc

upp deleted         3181 gctcgcacccgcaacgatcaatttaggcttatgttctcgtgctttttcaaggacatcatc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp      93 gctcgcacccgcaacgatcaatttaggcttatgttctcgtgctttttcaaggacatcatc

upp deleted         3241 aaaattaatccggtgtgttttttcatctaccccatattcaacaaagttatattgaactcc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp      33 aaaattaatccggtgtg-ttttttcatctacccct--------------------------

upp deleted         3301 gctgaagttgaccggacttccatgcgtcaaatggccgccgtgggacagattcatgccaag
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3361 cactgtatcgccaggctttagcaccgcataataaaccgccatgtttgcctgagcaccgga
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3421 gtgaggttgaacattgacgtgttcagcaccgaaaatttgcttagcacggtcccgtgcaag
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3481 gttttccactacatcgacatgttcacaaccgccataataacgccggcccggataaccttc
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3541 cgcatatttattcgttaataccgaaccttgtgcttccataaccgcttcactgacaaagtt
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3601 ttcggaagcgattaattcgatttttgtccgctgccgctttaactcatcttgaattgcttg
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3661 aaatacttgttcatcctgttgcgctaagtgcttcaaatcgttatcctccttttttatttac
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3721 attcataaaatatgtatcctttagaaaattcttttttatttttatcacgttagcttgatta
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------

upp deleted         3781 aaagaagaacaaaaaagcctttgctccctgcaaaggcttcctgtctaggattggtaaatt
02JB41_FwdDELupp         ------------------------------------------------------------
02JB42_RevDELUpp         ------------------------------------------------------------
```

DNA multi-alignment showing deletion of *mtlD* in *B. methanolicus* MGA3:

The sequencing results for confirmation of deletion of *mtlD* in Δ*mtlD B. methanolicus* MGA3 strain, using primer DEL39 is shown below. The sequence was multi-aligned towards the *mtlD* gene and the wild-type gDNA. The mutant strain include 15 bps overhangs due to the design of the homologous flanking regions of the suicide vector, including 15 bps of the gene both upstream and downstream.

```
mtlDtemplate      10801 gtacatccttctttcacaatcttaaggcttcaaaagccgcttcgagtaaaggcgatggac
mtlD                    ------------------------------------------------------------
DELmtld              10 -----------------------aggcttca-aagccgcttcgagtaaaggcgatggac

mtlDtemplate      10861 caatctgaaatggaagcagaaagcatccttttattattatcacctcagcatttatcaagc
mtlD                    ------------------------------------------------------------
DELmtld              45 caatctgaaatggaagcagaaagcatccttttattattatcacctcagcatttatcaagc

mtlDtemplate      10921 gagggattagaggtcttaagttttatcagcgcattaattattgaaaatgaacagagcata
mtlD                    ------------------------------------------------------------
DELmtld             105 gagggattagaggtcttaagttttatcagcgcattaattattgaaaatgaacagagcata
```

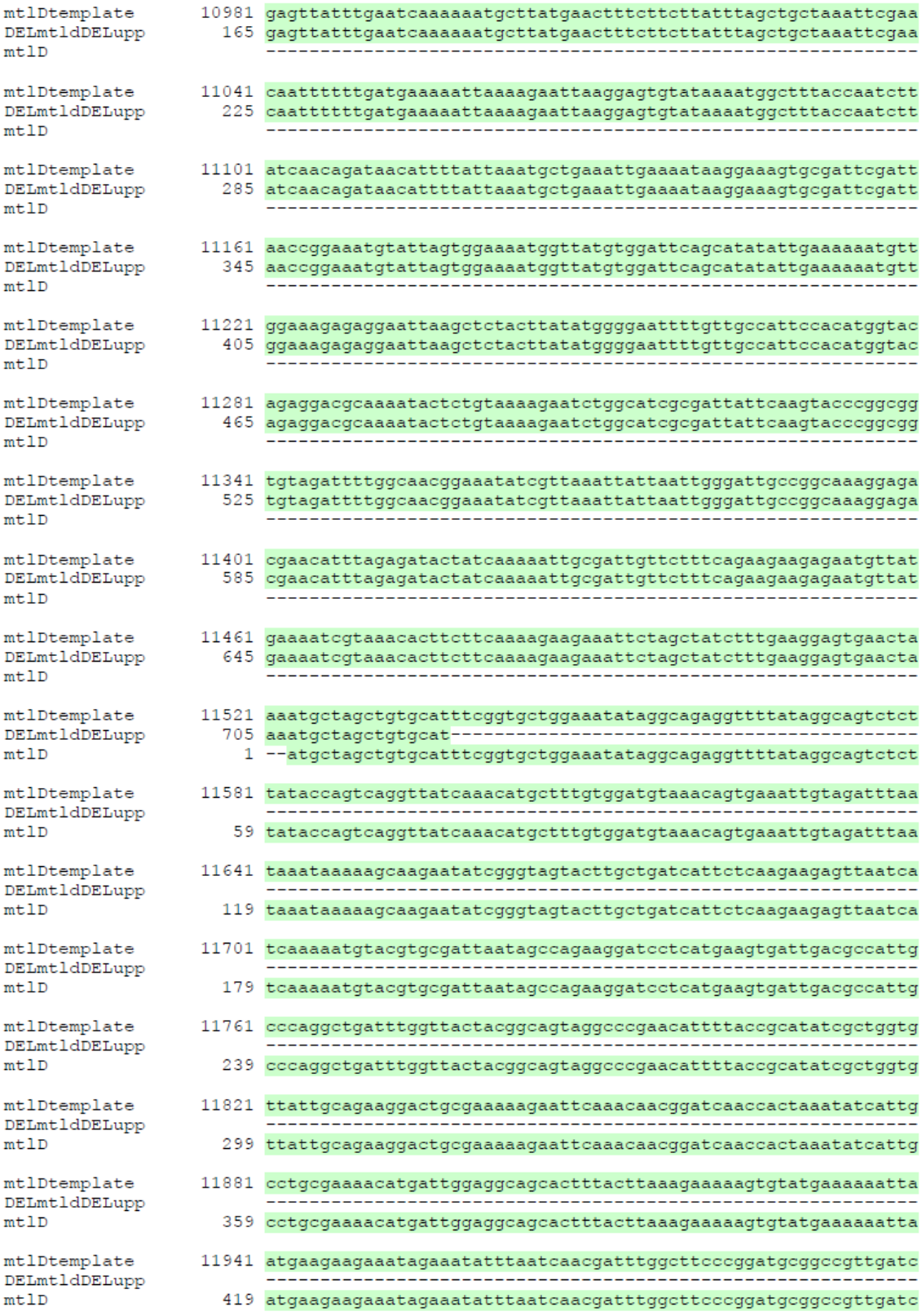
```
mtlDtemplate      10981 gagttatttgaatcaaaaaatgcttatgaactttcttcttatttagctgctaaattcgaa
mtlD                  ------------------------------------------------------------
DELmtld             165 gagttatttgaatcaaaaaatgcttatgaactttcttcttatttagctgctaaattcgaa

mtlDtemplate      11041 caatttttttgatgaaaaattaaaagaattaaggagtgtataaaatggctttaccaatctt
mtlD                  ------------------------------------------------------------
DELmtld             225 caatttttttgatgaaaaattaaaagaattaaggagtgtataaaatggctttaccaatctt

mtlDtemplate      11101 atcaacagataacattttattaaatgctgaaattgaaaataaggaaagtgcgattcgatt
mtlD                  ------------------------------------------------------------
DELmtld             285 atcaacagataacattttattaaatgctgaaattgaaaataaggaaagtgcgattcgatt

mtlDtemplate      11161 aaccggaaatgtattagtggaaaatggttatgtggattcagcatatattgaaaaaatgtt
mtlD                  ------------------------------------------------------------
DELmtld             345 aaccggaaatgtattagtggaaaatggttatgtggattcagcatatattgaaaaaatgtt

mtlDtemplate      11221 ggaaagagaggaattaagctctacttatatggggaattttgttgccattccacatggtac
mtlD                  ------------------------------------------------------------
DELmtld             405 ggaaagagaggaattaagctctacttatatggggaattttgttgccattccacatggtac

mtlDtemplate      11281 agaggacgcaaaatactctgtaaaagaatctggcatcgcgattattcaagtacccggcgg
mtlD                  ------------------------------------------------------------
DELmtld             465 agaggacgcaaaatactctgtaaaagaatctggcatcgcgattattcaagtacccggcgg

mtlDtemplate      11341 tgtagattttggcaacggaaatatcgttaaattattaattgggattgccggcaaaggaga
mtlD                  ------------------------------------------------------------
DELmtld             525 tgtagattttggcaacggaaatatcgttaaattattaattgggattgccggcaaaggaga

mtlDtemplate      11401 cgaacatttagagatactatcaaaaattgcgattgttctttcagaagaagagaatgttat
mtlD                  ------------------------------------------------------------
DELmtld             585 cgaacatttagagatactatcaaaaattgcgattgttctttcagaagaagagaatgttat

mtlDtemplate      11461 gaaaatcgtaaacacttcttcaaaagaagaaattctagctatctttgaaggagtgaacta
mtlD                  ------------------------------------------------------------
DELmtld             645 gaaaatcgtaaacacttcttcaaaagaagaaattctagctatctttgaaggagtgaacta

mtlDtemplate      11521 aaatgctagctgtgcatttcggtgctggaaatataggcagaggttttataggcagtctct
mtlD                  1 --atgctagctgtgcatttcggtgctggaaatataggcagaggttttataggcagtctct
DELmtld             705 aaatgctagctgtgcat----------------------------------------

mtlDtemplate      11581 tataccagtcaggttatcaaacatgctttgtggatgtaaacagtgaaattgtagatttaa
mtlD                 59 tataccagtcaggttatcaaacatgctttgtggatgtaaacagtgaaattgtagatttaa
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11641 taaataaaaagcaagaatatcgggtagtacttgctgatcattctcaagaagagttaatca
mtlD                119 taaataaaaagcaagaatatcgggtagtacttgctgatcattctcaagaagagttaatca
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11701 tcaaaaatgtacgtgcgattaatagccagaaggatcctcatgaagtgattgacgccattg
mtlD                179 tcaaaaatgtacgtgcgattaatagccagaaggatcctcatgaagtgattgacgccattg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11761 cccaggctgatttggttactacggcagtaggcccgaacattttaccgcatatcgctggtg
mtlD                239 cccaggctgatttggttactacggcagtaggcccgaacattttaccgcatatcgctggtg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11821 ttattgcagaaggactgcgaaaaagaattcaaacaacggatcaaccactaaatatcattg
mtlD                299 ttattgcagaaggactgcgaaaaagaattcaaacaacggatcaaccactaaatatcattg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11881 cctgcgaaaacatgattggaggcagcactttacttaaagaaaaagtgtatgaaaaaatta
mtlD                359 cctgcgaaaacatgattggaggcagcactttacttaaagaaaaagtgtatgaaaaaatta
DELmtld                 ------------------------------------------------------------

mtlDtemplate      11941 atgaagaagaaatagaaatatttaatcaacgatttggcttcccggatgcggccgttgatc
mtlD                419 atgaagaagaaatagaaatatttaatcaacgatttggcttcccggatgcggccgttgatc
DELmtld                 ------------------------------------------------------------
```

```
mtlDtemplate      12001 gaattgtgccaaatcaatcgaatgaagacaaattaatggtgatggttgaaccgttttatg
mtlD                479 gaattgtgccaaatcaatcgaatgaagacaaattaatggtgatggttgaaccgttttatg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12061 aatgggctgtagaccaaacaaaaattgttggggaaaaaccacctgtcgagggaattacgt
mtlD                539 aatgggctgtagaccaaacaaaaattgttggggaaaaaccacctgtcgagggaattacgt
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12121 ttgtagacgatttaaagccttacattgaaagaaaactatttaccgttaacactggacatg
mtlD                599 ttgtagacgatttaaagccttacattgaaagaaaactatttaccgttaacactggacatg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12181 cagttgcagcttattttggctattatgcgggcatccaaacaatcaatgaagcaatggaaa
mtlD                659 cagttgcagcttattttggctattatgcgggcatccaaacaatcaatgaagcaatggaaa
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12241 gagaagatatcaagactcttgtagagaaaactcttcaagaaacaggcttgctgctggtga
mtlD                719 gagaagatatcaagactcttgtagagaaaactcttcaagaaacaggcttgctgctggtga
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12301 aaaaatataactttgagatggaaaagcatcaagaatacatttctaaaattgtcaatcgct
mtlD                779 aaaaatataactttgagatggaaaagcatcaagaatacatttctaaaattgtcaatcgct
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12361 tctcaaatccgtatatttcagatgaagtaacaagggtgggccgttctccaatccgtaagc
mtlD                839 tctcaaatccgtatatttcagatgaagtaacaagggtgggccgttctccaatccgtaagc
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12421 ttggaccgaatgaccggttaataagcccggctaaacaatttgcagaaatcgttggcgatc
mtlD                899 ttggaccgaatgaccggttaataagcccggctaaacaatttgcagaaatcgttggcgatc
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12481 agccggtatacttagcaaaaagcattgcagcagctttactatatgattatccacaagatg
mtlD                959 agccggtatacttagcaaaaagcattgcagcagctttactatatgattatccacaagatg
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12541 cggaagcggtagaaatacaagaaagcatcaagaataatggattagaattaaccattgaaa
mtlD                1019 cggaagcggtagaaatacaagaaagcatcaagaataatggattagaattaaccattgaaa
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12601 aatatactcagctggacagtgattcaaatttagcaaaattaattgtagaacaatacaatt
mtlD                1079 aatatactcagctggacagtgattcaaatttagcaaaattaattgtagaacaatacaatt
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12661 catttaagggcgggaaattataaaaattattccgctgcaggatttaatgaattaagtaag
mtlD                1139 catttaagggcgggaaattataa----------------------------------------
DELmtld              722 --------ggcgggaaattataaaaattattccgctgcaggatttaatgaattaagtaag

mtlDtemplate      12721 atttaaagttttaacaatgctttttgctggaatgtaattacaagataataattcggttc
mtlD                    ------------------------------------------------------------
DELmtld              774 atttaaagttttaacaatgctttttgctggaatgtaattacaagataataattcggttc

mtlDtemplate      12781 tagatcgaagaaaaattaggcatgatcctaatttttcttcgatttttttttgtctgttttt
mtlD                    ------------------------------------------------------------
DELmtld              834 tagatcgaagaaaaattaggcatgatcctaatttttcttcgatttttttttgtctgttttt

mtlDtemplate      12841 tggaatgaggggggaaaaggggaggaagttttaaagatttgaataaaataatctttttt
mtlD                    ------------------------------------------------------------
DELmtld              894 tggaatgaggggggaaaaggggag-------------------------------------

mtlDtemplate      12901 gtcgttgttcgaacttgtaatcatagcctcaaacttaactttatttaagtctgaaaccag
mtlD                    ------------------------------------------------------------
DELmtld                 ------------------------------------------------------------

mtlDtemplate      12961 agtatataaaagctcgaacaactttggattacaattcatatattaatcaatctttgttaa
mtlD                    ------------------------------------------------------------
DELmtld                 ------------------------------------------------------------
```

DNA multi-alignment showing clean deletion of *mtlD* in Δ*mtlD*Δ*upp B. methanolicus* MGA3 strain:

The obtained sequencing result for confirmation of deletion of *mtlD* in *B. methanolicus* Δ*mtlD*Δ*upp* using primer DEL39 is shown below. The sequence was multi-aligned towards the *mtlD* of *B. methanolicus* and a wild-type gDNA sequence of *B. methanolicus* MGA3 (”mtlDtemplate“),

but including a 15 bps overlap from then ends of the gene.



```
mtlDtemplate      12001 gaattgtgccaaatcaatcgaatgaagacaaattaatggtgatggttgaaccgttttatg
DELmtldDELupp           ------------------------------------------------------------
mtlD                479 gaattgtgccaaatcaatcgaatgaagacaaattaatggtgatggttgaaccgttttatg

mtlDtemplate      12061 aatgggctgtagaccaaacaaaaattgttggggaaaaaccacctgtcgagggaattacgt
DELmtldDELupp           ------------------------------------------------------------
mtlD                539 aatgggctgtagaccaaacaaaaattgttggggaaaaaccacctgtcgagggaattacgt

mtlDtemplate      12121 ttgtagacgatttaaagccttacattgaaagaaaactatttaccgttaacactggacatg
DELmtldDELupp           ------------------------------------------------------------
mtlD                599 ttgtagacgatttaaagccttacattgaaagaaaactatttaccgttaacactggacatg

mtlDtemplate      12181 cagttgcagcttattttggctattatgcgggcatccaaacaatcaatgaagcaatggaaa
DELmtldDELupp           ------------------------------------------------------------
mtlD                659 cagttgcagcttattttggctattatgcgggcatccaaacaatcaatgaagcaatggaaa

mtlDtemplate      12241 gagaagatatcaagactcttgtagagaaaactcttcaagaaacaggcttgctgctggtga
DELmtldDELupp           ------------------------------------------------------------
mtlD                719 gagaagatatcaagactcttgtagagaaaactcttcaagaaacaggcttgctgctggtga

mtlDtemplate      12301 aaaaatataactttgagatggaaaagcatcaagaatacatttctaaaattgtcaatcgct
DELmtldDELupp           ------------------------------------------------------------
mtlD                779 aaaaatataactttgagatggaaaagcatcaagaatacatttctaaaattgtcaatcgct

mtlDtemplate      12361 tctcaaatccgtatatttcagatgaagtaacaagggtgggccgttctccaatccgtaagc
DELmtldDELupp           ------------------------------------------------------------
mtlD                839 tctcaaatccgtatatttcagatgaagtaacaagggtgggccgttctccaatccgtaagc

mtlDtemplate      12421 ttggaccgaatgaccggttaataagcccggctaaacaatttgcagaaatcgttggcgatc
DELmtldDELupp           ------------------------------------------------------------
mtlD                899 ttggaccgaatgaccggttaataagcccggctaaacaatttgcagaaatcgttggcgatc

mtlDtemplate      12481 agccggtatacttagcaaaaagcattgcagcagctttactatatgattatccacaagatg
DELmtldDELupp           ------------------------------------------------------------
mtlD                959 agccggtatacttagcaaaaagcattgcagcagctttactatatgattatccacaagatg

mtlDtemplate      12541 cggaagcggtagaaatacaagaaagcatcaagaataatggattagaattaaccattgaaa
DELmtldDELupp           ------------------------------------------------------------
mtlD               1019 cggaagcggtagaaatacaagaaagcatcaagaataatggattagaattaaccattgaaa

mtlDtemplate      12601 aatatactcagctggacagtgattcaaatttagcaaaattaattgtagaacaatacaatt
DELmtldDELupp           ------------------------------------------------------------
mtlD               1079 aatatactcagctggacagtgattcaaatttagcaaaattaattgtagaacaatacaatt

mtlDtemplate      12661 catttaagggcgggaaattataaaaattattccgctgcaggatttaatgaattaagtaag
DELmtldDELupp       722 --------ggcgggaaattataaaaattattccgctgcaggatttaatgaattaagtaag
mtlD               1139 catttaagggcgggaaattataa----------------------------------

mtlDtemplate      12721 atttaaagttttaacaatgctttttttgctggaatgtaattacaagataataattcggttc
DELmtldDELupp       774 atttaaagttttaacaatgctttttttgctggaatgtaattacaagataataattcggttc
mtlD                    ------------------------------------------------------------

mtlDtemplate      12781 tagatcgaagaaaaattaggcatgatcctaatttttcttcgatttttttttgtctgttttt
DELmtldDELupp       834 tagatcgaagaaaaattaggcatgatcctaatttttcttcgatttttttttgtctgttttt
mtlD                    ------------------------------------------------------------

mtlDtemplate      12841 tggaatgaggggggaaaaggggaggaagttttaaagatttgaataaaataatctttttttt
DELmtldDELupp       894 tggaatgaggggggaaaaggggaggaagttttaaagatttgaataaaataatctttttttt
mtlD                    ------------------------------------------------------------

mtlDtemplate      12901 gtcgttgttcgaacttgtaatcatagcctcaaacttaactttatttaagtctgaaaccag
DELmtldDELupp       954 gtcgttgttcgaacttgtaatcatagcctcaaacttaactttatttaagtctgaaaccag
mtlD                    ------------------------------------------------------------

mtlDtemplate      12961 agtatataaaagctcgaacaactttggattacaattcatatattaatcaatctttgttaa
DELmtldDELupp      1014 agtatataaaagctcgaacaactttggattacaattcatatattaatcaatctttgttaa
mtlD                    ------------------------------------------------------------
```

Appendix F: Calculations

Calculation of reinoculation volumes

Reinoculation volumes from pre-cultures to main cultures were calculated from Equation (1), which can be rearranged as Equation (2). "C_1" and "C_2" correspond to the initial and final concentration, respectively. "V_1" and "V_2" is the initial and final volumes, respectively. As OD_{600} is a measure of cell density, these equations are also applicable for ODs as a measure of concentration.

$$C_1 V_1 = C_2 V_2 \tag{1}$$

$$V_1 = \frac{C_2}{C_1} \times V_2 \tag{2}$$

For a $40\,\mathrm{ml}$ main culture with initial OD_{600} of 0.2 for the main culture, the needed volume to be re-inoculated from the pre-culture with OD_{600}=C_1 corresponds to Equation (3).

$$V_1 = \frac{0.2}{C_1} \times 40\,\mathrm{ml} \tag{3}$$

Example calculation

For reinoculation into six main cultures of each $40\,\mathrm{ml}$ with an average OD_{600} equal to 1 of the pre-cultures, the volume needed for reinoculation corresponds to 48 ml (Equation (4)).

$$V_1 = \frac{0.2}{1.0} \times 6 \times 40\,\mathrm{ml} = 48\,\mathrm{ml}. \tag{4}$$

Standard deviation

The standard deviation is measured as a quantity of uncertainty in a data set. It describes the spreading (variation) of the data obtained under the same conditions and can be calculated as the deviation from the calculated sample mean value. All standard deviations presented as part of this study was calculated using the function "STDAV.P" in Microsoft® 365 Excel (2020). The sample standard deviation is calculated as the square root of the sample variance, s^2. In turn, the sample variance is calculated from the formula shown in Equation (5). "n" is the number of values, x_i is the sample value, $\overline{x}$ is the mean sample value.

$$s^2 = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{(n-1)}} \tag{5}$$

Calculating enzyme specific activity

Calculations of the specific activity of *mtlD* encoded mannitol 1-phosphate 5-dehydrogenase, was carried out in accordance with Beer's law which relates the measured absorbance to concentration, as presented in Equation (6). "A" correponding to the measured absorbance, "C" the concentration of the solution, "l" the path length, and "ϵ" the molar absorptivity.

$$\mathrm{A} = \epsilon \cdot \mathrm{C} \cdot \mathrm{l} \tag{6}$$

The Beer's law can further be rearranged to display the concentration as a function of absorbance, as shown in Equation (7).

$$\mathrm{C} = \frac{\mathrm{A}}{\epsilon \cdot \mathrm{l}} \tag{7}$$

Example calculation

The first cell crude extract of *B. methanolicus* MGA3 wild-type technical triplicate cultures gave a slope of -0.1261 min^{-1} as the reaction was initiated. The corresponding baseline was measured to -0.0081 min^{-1}. The change in absorbance as function of time for the mannitol 1-phosphate 5-dehydrogenase catalyzed reaction was then calculated as the absolute difference of the two measured slopes, yielding 0.118 min^{-1}. For NADH, ϵ is a constant with a known value of 6220 $\mathrm{Lmol}^{-1}\mathrm{cm}^{-1}$. The optical path length of the cuvette used was 1 cm. The concentration can then be calculated from Equation (7), yielding:

$$\mathrm{C} = \frac{|-0.1261 - (-0.0081))|\,\mathrm{min}^{-1}}{\mathrm{L} \cdot \mathrm{min}} = 18.97\,\mu\mathrm{mol/L} \cdot \mathrm{min} \tag{8}$$

The reaction was carried out in a total volume of 1 ml with a 1:10 dilution of $50\,\mu$l cell crude extract. This gives the enzyme activity of the cell crude extract to be:

$$\Rightarrow \mathrm{C} = 0.0187\,\mu\mathrm{mol/min} \cdot \left(\frac{1000\,\mu\mathrm{l}}{5\,\mu\mathrm{l}}\right)\mathrm{ml} = 3.74\,\mu\mathrm{mol}, \mathrm{min}^{-1}, \mathrm{ml}^{-1} \tag{9}$$

The enzyme specific activity is a measure correlated to the amount of total proteins present in the crude extract and is found from dividing the enzyme activity by the measured protein concentration of the cell crude extract estimated using Bradford assay (see Appendix G):

$$\Rightarrow \mathrm{C} = \frac{3.74\,\mu\mathrm{mol}, \mathrm{min}^{-1}\mathrm{ml}^{-1}}{905.46\,\mu\mathrm{g/ml}} = 4.13 \cdot 10^{-3}\mu\mathrm{mol/min}, \mu\mathrm{g} = 4.13\,\mu\mathrm{mol/min}, \mathrm{mg} \quad (10)$$

In terms of enzyme units, the calculated enzyme specific activity corresponds to 4.13 U/mg.

Appendix G: Bradford assay

Dilutions of BSA

The dilutions of BSA used for the making of a calibration curve was proposed in the Quick Start® Bradford Protein Assay instruction manual (Bio-Rad), and are presented in Table 2.

Table 2: Dilutions of BSA made for making a calibration curve.

Tube nr.	Final protein concentration (ug/ml)	Absorbance (494 nm)
1	2000	1.74
2	1500	1.49
3	1000	1.22
4	750	0.81
5	500	0.80
6	250	0.48
7	125	0.20
8	0	0.00

Calibration curve for protein quantification

The calibration curve made from measured absorbances at 595 nm for the different concentrations of BSA is shown in Figure 4. Data points of measured OD_{595} versus known sample concentrations were plot using Microsoft® Excel® (Microsoft Corporation, spreadsheet software).

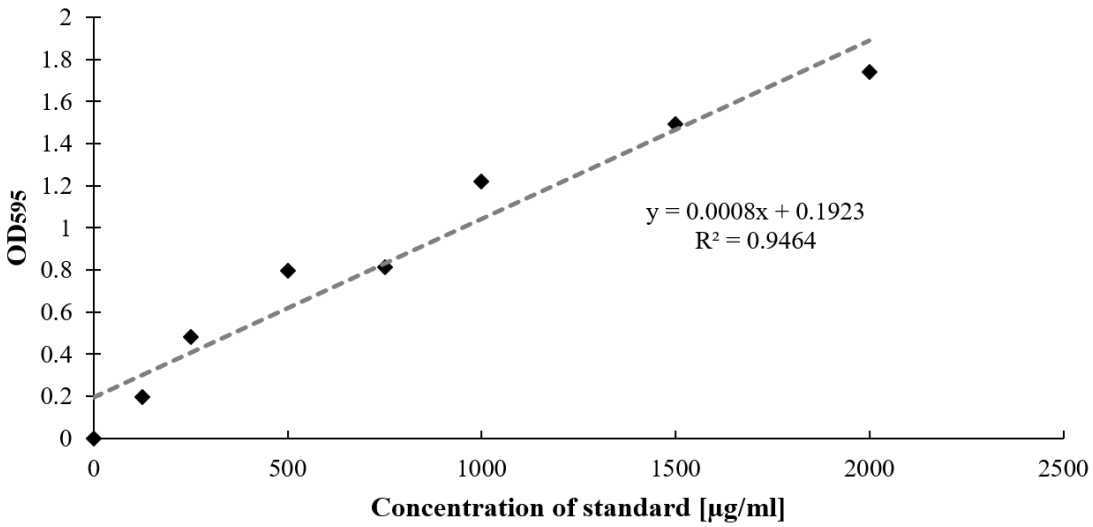


Figure 4: Calibration curve of BSA for Bradford protein assay. Absorbances were measured for each BSA concentration at 595 nm. A linear trend line fitted between the data-points is displayed as y=0.0008 + 0.1923, with a coefficient of determination (R^2) value of 0.9464.

From the linear fitted trend line, an expression relating optical density (absorbance) to concentration ($\mu g/ml$) could be made (Equation (11)). The coefficient of determination was calculated to R^2=0.9464. This equation was eventually used for the estimation of protein content in the cell crude extracts prepared for the *mtlD* enzyme assay.

$$C = \frac{OD_{595} - 0.1923}{0.0008} \tag{11}$$

Example calculation

Measuring an OD_{595} of 0.29 of a cell crude extract diluted 1:10, will estimate the total protein concentration from Equation (11) to be

$$C = \frac{0.29 - 0.1923}{0.0008} \cdot 10 = 1221\,\mu g/ml.$$

Appendix H: Activity data for *mtlD* enzyme assay

Activity data (absorbance versus time) was obtained as monitored OD_{340} for 3 minutes. Activity data was processed in Microsoft® Excel® (Microsoft Corporation, Spreadsheet Software) by plotting absorbance versus time. An example is presented below in Figure 5 for the first sample of *B. methanolicus* MGA3. From the slope of the curves, the change in absorbance was estimated from the formula obtained by a linear regression of the values, giving a slope of -0.2353 min^{-1}. The slopes of the baselines were also estimated similarly.

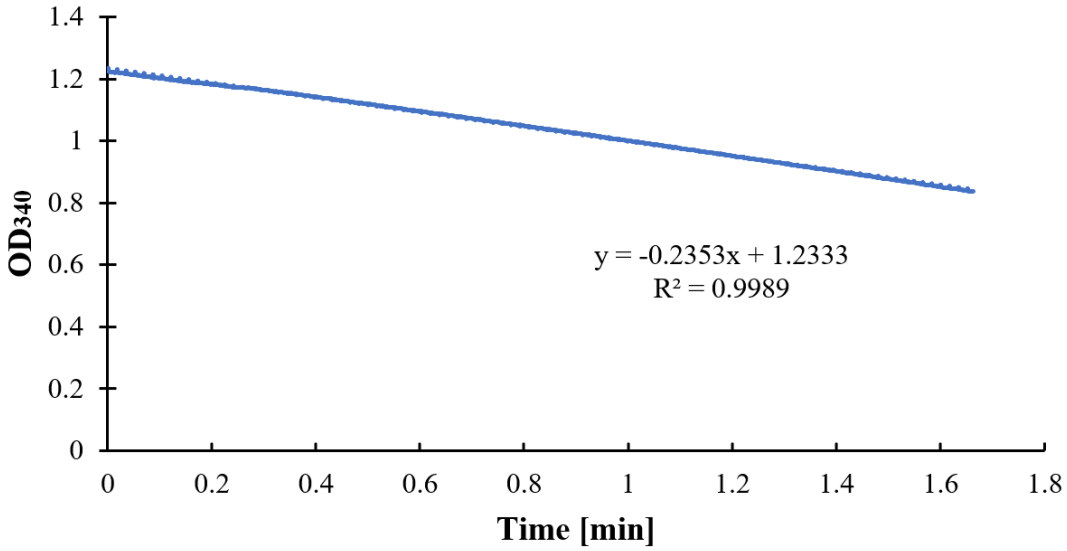


Figure 5: Activity data plot for the first *B. methanolicus* MGA3 sample in *mtlD* enzyme assay. A linear relation between OD_{340} and time (min) was obtained from linear regression: y=-0.2353x+1.2333, R^2=0.9989. The coefficient of determination (R^2) was found to be 0.9989.

The estimated slopes (min^{-1}) for the *B. methanolicus* MGA3 and Δ*mtlD* MGA3 technical triplicates are displayed in Table 3.

Table 3: Measured background and enzyme activity in terms of slopes from linear regression of absorbances monitored as function of time. The table displays slopes (min^{-1}) of baselines and activities of *B. methanolicus* MGA3 and Δ*mtlD* technical triplicates.

Strain	Baseline (Abs/min)	Activity (Abs/min)
MGA3_1	-0.008	-0.126
MGA3_2	-0.011	-0.083
MGA3_3	-0.009	-0.046
Δ*mtlD* MGA3_1	-0.009	-0.235
Δ*mtlD* MGA3_2	-0.011	-0.203
Δ*mtlD* MGA3_3	-0.011	-0.141

The estimated total protein concentrations (mg/ml), enzyme activities (U/ml), and specific enzyme activities (U/mg) are listed in Table 4. See Appendix G for determining protein concentrations and Appendix F for calculating enzyme activities. Table 4 also displays the mean enzyme specific activities and standard deviations of the technical triplicates.

Table 4: Measured concentrations of technical triplicate cell crude extracts of *B. methanolicus* MGA3 wildtype strain and Δ*mtlD*Δ*upp* mutant strain, estimated from Bradford assay. Underscores corresponds to each triplicate of the investigated strains. Enzyme activity in units (U/ml) was calculated using Beer's law. Enzyme specific activities (U/mg) are also included.

Strain	Protein concentration (mg/ml)	U/ml (µmol/ml·min)	U/mg (µmol/mg ·min)
MGA3_1	0.91	3.80	4.19
MGA3_2	0.59	2.32	3.95
MGA3_3	0.35	1.20	3.41
Δ*mtlD*Δ*upp*_1	0.64	1.46	2.26
Δ*mtlD*Δ*upp*_2	0.58	1.24	2.14
Δ*mtlD*Δ*upp*_3	0.25	0.81	3.23

Plots of baseline and activity data for the technical triplicates of *B. methanolicus* MGA3 and Δ*mtlD* are shown in Figure 6 and Figure 7, respectively. Crude extracts of the wildtype strain was diluted 1:10 and the Δ*mtlD* strain diluted 1:2 before being added to their respective reaction mixtures. The data are only plotted for the first 2 minutes, despite that the reaction was performed for 3 minutes each. However, the linear trend is clearly displayed.

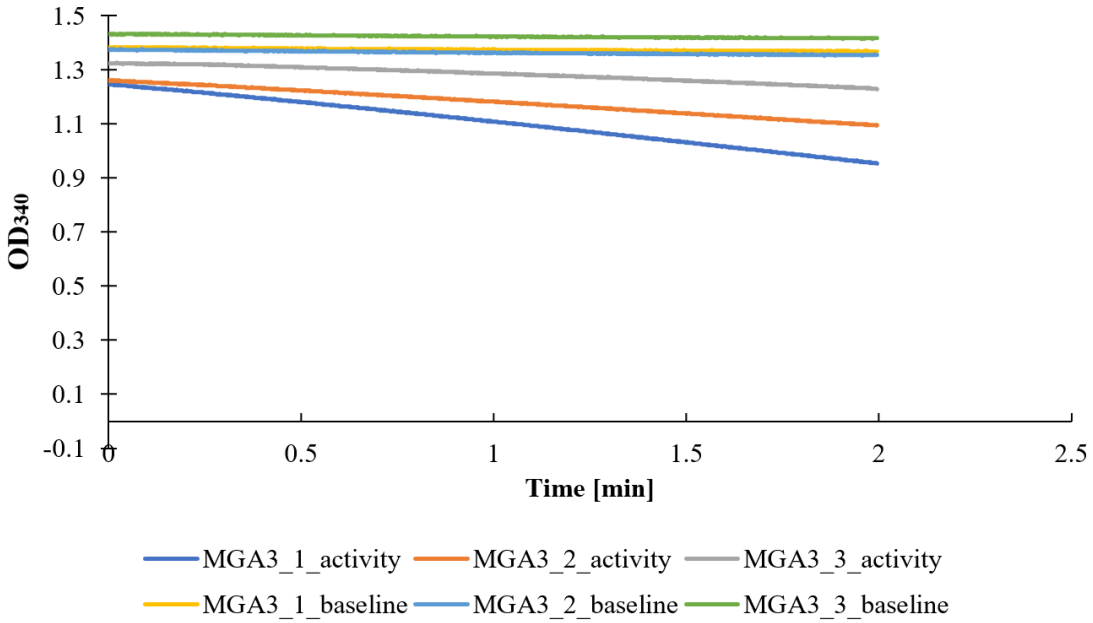


Figure 6: Continous activity data plots for technical triplicate *B. methanolicus* MGA3 in MtlD enzyme assay. Figure displays data for both baseline and reaction activity.

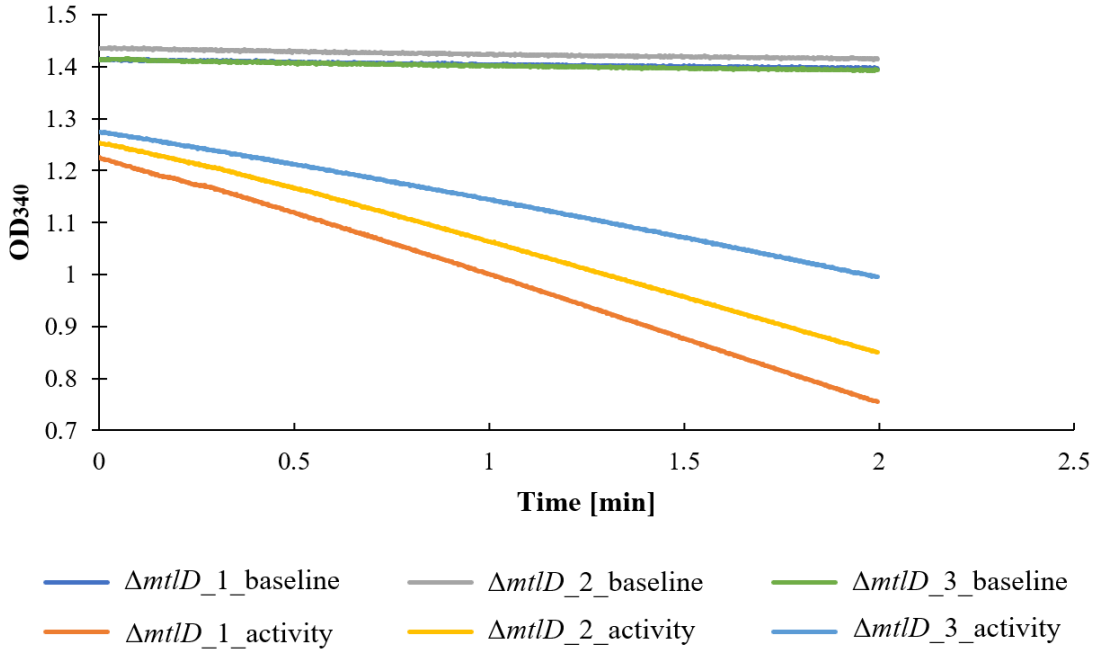


Figure 7: Continous activity data plots for technical triplicate of *B. methanolicus* MGA3 Δ*mtlD* in MtlD enzyme assay. Figure displays data for both baseline and reaction activity.

Jesper Langlo

Development of a markerless tool for gene deletions in *Bacillus methanolicus*

NTNU
Norwegian University of
Science and Technology