Technische Universität München
Lehrstuhl für Mikrobiologie

# Investigation of xylose and arabinose metabolism in *Clostridium acetobutylicum* ATCC 824 and *Clostridium saccharobutylicum* NCP 262

Justyna Maria Lesiak

Vollständiger Abdruck der von der Fakultät für Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. E. Grill

Prüfer der Dissertation:

1. Univ.-Prof. Dr. W. Liebl
2. Univ.-Prof. Dr. R. F. Vogel

Die Dissertation wurde am 24.11.2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 03.02.2015 angenommen.

**Table of contents**

**Abbreviation list**

| | |
|---|---|
| °C | degree Celsius |
| µ | mikro |
| µF | mikro Farad |
| 2 x YTG | 2 x Yeast Tryptone Glucose medium |
| 5-FOA | 5-fluoroorotic acid |
| 5-FU | 5-fluorouracil |
| A | adenine |
| ACE | allelic exchange |
| ADP | adenosine-5'-diphosphate |
| AEBSF | 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride |
| Amp | ampicilin |
| AP | alkaline phosphatase |
| AraK | arabinose kinase |
| ATCC | American Type Culture Collection |
| ATP | adenosine-5'-triphosphate |
| BLAST | Basic Alignment Search Tool |
| bp | base pair |
| BSA | bovine serum albumin |
| C | cytosine |
| ca. | *circa* |
| cDNA | complementary deoxyribonucleic acid |
| CGM | Clostridial Growth Medium |
| cm | centimeter |
| Cm | chlorapmhenicol |
| CoA | coenzyme A |
| Crm | clarithromycin |
| dATP | deoxyriboadenosine-5'-triphosphate |
| dCTP | deoxyribocytosine-5'-triphosphate |

| | |
|---|---|
| dd | double distilled |
| DDT | dichlorodiphenyltrichloroethane |
| DELTA-BLAST | Domain Enhanced Lookup Time Accelerated BLAST |
| dGTP | deoxyriboguanine-5’-triphosphate |
| DHAP | dihydroxyacetone phosphate |
| DMSZ | Deutsche Stammsammlung von Mikroorganismen und Zellkulturen |
| DNA | deoxyribonucleic acid |
| DNase I | deoxyribonucleinase I |
| dNTP | deoxyribonucleotide-5’-triphosphate |
| dTTP | deoxyribotymidine-5’-triphosphate |
| EDTA | Ethylenediaminetetraacetic acid |
| EMP | Embden-Meyerhof-Parnas pathway |
| ENase | endonuclease |
| Erm | erythromycin |
| *et al.* | *et alteri* (and others) |
| ET buffer | elektroporation-transfer buffer |
| EtBR | ethydium bromide |
| ETM buffer | elektroporation-transfer (magnesium) buffer |
| EtOH | ethanol |
| Fd, Fdx | ferredoxin |
| g | gram |
| G | guanine |
| GC | gas chromatography |
| GOI | Gene Of Interest |
| h | hour |
| HA | homology arm |
| HCl | hydrochloric acid |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| Hg | inch of mercury |

| | |
|---|---|
| HsdM | methylation subunit of RM system |
| HsdR | restrictase subunit of RM system |
| HsdS | specificity subunit of RM system |
| JGI | Joint Genome Institute |
| k | kilo ($10^3$) |
| Kan | kanamycin |
| kb | kilo base |
| kJ | kilo joule |
| KOH | potassium hydroxide |
| kV | kilo volt |
| l | liter |
| LB | Luria-Bertani |
| Lm | lincomycin |
| log | logarithm |
| m | mili ($10^{-3}$) |
| M | molar (mol/l) |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| min | minute |
| MMVK | Minimal Medium For precultures |
| MOPS | 3-morpholinopropane-1-sulfonic acid |
| mRNA | messenger ribonucleic acid |
| MTase | methyltransferase of RM system |
| n | nano ($10^{-9}$) |
| NaAc | sodium acetate |
| $NAD^+$ | nicotinamide adenine dinucleotide, oxidized form |
| NADH | nicotinamide adenine dinucleotide, reduced form |
| NaOH | sodium hydroxide |
| NCBI | National Center for Biotechnology Information |
| NCIMB | National Collection of Industrial, Food and Marine Bacteria |

| | |
|---|---|
| NCP | National Chemical Products |
| No. | number |
| $OD_x$ | optical density at wavelength of x nm |
| *ori* | origin of replication |
| Ox | oxidized |
| P | phosphate |
| PABA | p-Aminobenzoic Acid |
| PBS | Phosphate buffered saline |
| PCR | polymerase chain reaction |
| PEP | phosphoenolpyruvate |
| pH | decimal logarithm of the reciprocal of the hydrogen ion activity |
| $P_i$ | inorganic phosphate |
| PKP | phosphoketolase pathway |
| Pkt | phosphoketolase |
| PLMM | Phosphate Limited Minimal Medium |
| PPP | pentose phosphate pathway |
| PTS | phosphotransferase system |
| $p_x$ | promotor of gene *X* |
| RACE | Rapid Amplification of cDNA End |
| RBS | ribosome binding site |
| RCM | Reinforced Clostridial Medium |
| Red | reduced |
| Rep | origin of replication (replicon) |
| RLM-RACE | 5′-RNA-linker-mediated RACE |
| RM system | Restriction – Modification Sysytem |
| RNA | ribonucleic acid |
| RNase | ribonuclease |
| Rnf | *Rhodobacter* nitrogen fixation |
| rpm | rotation per minute |

| | |
|---|---|
| RT | room temperature |
| RT-PCR | Reverse Transcription PCR |
| SDS | sodium lauryl sulfate |
| sec | second |
| SOE-PCR | Splicing by Overlap Extension PCR |
| Sp | spectinomycin |
| SSC | saline sodium citrate (buffer) |
| ssDNA | single stranded DNA |
| Strep-AP | Strep-Alkaline Phosphatase |
| T | thymindine |
| TAE | Tris-acetate-EDTA buffer |
| TE | Tris-EDTA |
| Tm | thiamphenicol |
| TRD | target recognition domain |
| Tris | tris(hydroxymethyl)aminomethane |
| UMP | uridine monophosphate |
| UV | ultraviolet |
| v/v | volume per volume |
| vol. | volume |
| W | watt |
| w/v | weight per volume |
| WT | Wild Type |
| XylB | xylulose kinase |
| Ω | ohm |

## 1. Introduction

The genus *Clostridium* is one of the largest bacterial genera of Gram-positive, endospore-forming, strictly anaerobic bacteria, bringing together many species of high medical and biotechnological importance (Wells & Wilkins, 1996).

*Clostridium* species can be found in human and animal gastrointestinal tracts, soils, sediments or decaying plant materials (Ma *et al.*, 2012; Rupnik *et al.*, 2009; Johnson *et al.*, 1997; Freier *et al.*, 1988). Several free-living species synthesize extracellular or membrane-binding enzyme complexes, capable of degrading cellulose or hemicelluloses, and the insolubility of plant polymers increases their accessibility to soil- inhabiting clostridia.

Cellulose and tightly bound hemicelluloses serve mainly to strengthen the structure of plant cell walls, but also play a role in protecting against numerous herbivorous unicellular organisms; hence the function of cellulose- and hemicellulose-degrading bacteria remains one of the most important in microbial communities. The cell walls of various plants are composed of different ratios of cellulose and hemicelluloses as well as of different types of glycans, yet in the majority the dominant hexose is glucose and the dominant pentoses are xylose and arabinose. Carbohydrate substrates from hydrolyzed hemicelluloses are converted to different intermediates, which can be used for the production of certain acids and solvents. The amount of solvents depends on the substrate used and is usually the highest with hexoses, particularly glucose (Jones & Woods, 1986). Clostridial solventogenic species can produce acetone, butanol and ethanol through ABE (acetone-butanol-ethanol) fermentation (Dürre & Bahl, 1996) in large scale fermentors and some of them, like *Clostridium acetobutylicum* are already being closely studied. For the others, namely *Clostridium saccharobutylicum*, methods for genetic manipulations are still needed, although they are already being investigated for biotechnological purposes (Ni *et al.* 2013).

The ability of *C. acetobutylicum* to synthesize industrially important solvents was discovered at the beginning of the 20$^{th}$ century by Chaim Weizmann and nas been explored since WWI, during which its acetone production proved to be very important in the munitions industry (Jones & Woods, 1986). Since then many other *Clostridium* species have been isolated and investigated for the purposes of industry (Linville *et al.*, 2013; Dredge *et al.*, 2011). After the WWII interest in the exploitation of solventogenic clostridia

decreased, but today the world again faces the problem of limited sources of fossil fuels and a need for new, renewable sources of energy. *Saccharomyces* species or engineered *Escherichia coli* are capable of producing bioethanol (Mattam & Yazdani, 2013; Branduardi *et al.*, 2013; McKee *et al.*, 2012), yet this bulk chemical is of a lesser interest for the fuel industry than biobutanol. Butanol is a more attractive alternative to gasoline, considering its higher calorific value and lower water absorption compared to ethanol and, most importantly, it does not require any changes to be introduced in engines (Dürre, 2007). Therefore, butanol produced via ABE fermentation by clostridia attracts much more attention as an attractive alternative to ethanol additives in gasoline. The clostridial production of bulk chemicals is not only cheaper than it used to be and already adapted to a large scale, but also does not contribute to the greenhouse effect. However, the preferred and most efficient source of energy for all these microbes is glucose, whose production is becoming more expensive and requires land for plantations. In times of global overpopulation this causes competition between the fuel and food industries. The non-edible parts of plants which can serve as cheap and sustainable feedstock are rich in cellulose and hemicelluloses, the second having a less organized structure composed of shorter chains, and hence can be much more easily hydrolyzed than highly crystallized cellulose (Jurgens *et al.*, 2012).

Nowadays, the research in biotechnology continues to develop new methods to use plant waste (biowaste) as a source of carbons for clostridial fermentation (Berezina *et al.*, 2009).

## 1.1 Characterization of solventogenic species used in this work

### 1.1.1. *Clostridium acetobutylicum* ATCC 824

*Clostridium acetobutylicum* ATCC 824 was isolated in the form of a mixture of solventogenic organisms from plant-derived food (corn meal) and has been used since 1916 for acetone-butanol-ethanol production. It is also called a ‘Weizmann Organism’ after its discoverer (Jones & Woods, 1986).

The complete genome of *C. acetobutylicum* has been sequenced (Nölling *et al.* 2001) and in addition to 3.9 Mb chromosome, the pSOL1, a 210 kb large megaplasmid, was identified. In total 4022 open reading frames were described. Many of the genes important for sporulation and solventogenesis are located on pSOL1 (Cornillot *et al.*, 1997), and therefore its loss results in the inability to form solvents or spores (Alsaker & Papoutsakis,

2005). In the genome of *C. acetobutylicum* the G+C content is 31%, similar to *C. saccharobutylicum* NCP 262.

### 1.1.2. *Clostridium saccharobutylicum* NCP 262

*Clostridium saccharobutylicum* NCP 262 (DSM 13864), formerly *C. acetobutylicum*, is one of the four industrial solvent-producing species (Johnson *et al.*, 1997; Keis *et al.*, 1995). It was first isolated in South Africa in 1936 from a soya bean field and used for industrial butanol fermentations. The *C. saccharobutylicum* genome of 5.1 Mb size has been sequenced by the Göttingen Genomics Laboratory and is available in the Doe Joint Genome Institute database (https://img.jgi.doe.gov) and recently also in the NCBI resources (Poehlein *et al.*, 2013). The total number of genes is 4532, with a G+C content of 29%.

As an industrial strain, *C. saccharobutylicum* has an important advantage over *C. acetobutylicum*, the model organism for solventogenic clostridia. The genes responsible for solventogenesis and sporulation are located on the chromosome of the microorganism, unlike *C. acetobutylicum* ATCC 824, which under certain conditions can easily degenerate by losing its mega plasmid, pSOL1, and the ability to produce solvents as a consequence (Cornillot *et al.*, 1997). The amount of research focused on *C. saccharobutylicum* has been increasing lately, mainly due to its potential as a good solvent producer (Ni *et al.*, 2013; Berezina *et al.*, 2009; Liew *et al.*, 2006) and the development of a methylation system for its transformation is highly desirable. This would allow a better understanding of the metabolism of the microorganism and the creation of new, high solvent yield-producing strains.

### 1.1.3. *Clostridium beijerinckii* NCIMB 8022

*Clostridium beijerinckii* NCIMB 8052 was isolated from plant-derived food (corn meal) and described as *Clostridium acetobutylicum* (Keis *et al.*, 1995, Johnson *et al.*, 1997). It has a large genome of around 6.0 Mb in size with a total number of open reading frames of 5290. The G+C content in the chromosome is 30%. Nowadays *C. beijerinckii* is considered a great hope for industry (Qureshi *et al.*, 2007, Lee *et al.*, 2007), and in contrast to many clostridial species no glucose catabolite repression has been observed (Ezeji *et al.*, 2007).

Subspecies of *C. beijerinckii* are ubiquitous in nature and can be found in soil and food-derived samples all over the world (O'Brien & Morris, 1971).

## 1.2. Restriction-modification systems of *Clostridium saccharobutylicum* NCP 262

The restriction-modification systems (RM) of unicellular organisms play the role of a simple defense system against exogenous DNA and are present in many microbes. They were discovered at the beginning of the 1950s and since then thousands of them have been found and described (Bickle & Krüger, 1993). Although these systems are composed of genetically diverse elements, they all play a very similar role. An RM system is typically composed of two opposing elements – the restriction enzyme (endodeoxyribonuclease, ENase) and the methylation enzyme (DNA-methyltransferase, MTase), as shown in figure 1.1. Genes encoding these enzymes are very often closely linked or overlapping. Both enzymes recognize specific, four- to eight-base-pair-long sequences of double-stranded (ds) deoxyribonucleic acids and some of the RM systems are capable of recognizing also single stranded (ss) DNA. In general the cleavage occurs once for each recognized sequence and in most cases happens without energy input. It can take place either within or close to a recognized sequence, or in a completely random region (Wilson & Murray, 1991).

All known RM systems have been classified into four types – type I, type II, type IIs and type III, which have been further subdivided into families. Systems that cannot be categorized into any of these types are simply grouped in a fifth, independent category of RM systems. Typical RM systems are presented in figure 1.1.

Type II RM systems are the simplest and most common in nature. The MTase and ENase act independently and cut or methylate precisely at recognition site of a symmetric, continuous or interrupted, sequence. Endonucleases form homodimers in most cases, which facilitates the cleavage of both strands at once. Type IIs systems, where the ‘s’ stands for shift cleavage, act similarly to type II enzymes, but their recognition sequences are asymmetric and uninterrupted. The endonuclease genes are approximately twice as large as in type II and act as monomers. In type III systems the M subunit can act separately, like a methyltransferase, or form a heterodimer with the R subunit to act as a methyltransferase or endonuclease. Cleavage occurs only on one strand and is generally close (around 25 nucleotides) to a recognition sequence (Wilson & Murray, 1991).

In some organisms, i.e. bacteriophages, only methyltransferases are encoded in the genome, and serve as means of protection against the restrictases of their host bacteria. *Bacillus* phage *φ*3T possesses a multispecific MTase, which has been used to develop a methylation system for ribonucleotide delivery into *Clostridium acetobutylicum* cells (Mermelstein & Papoutsakis, 1993).

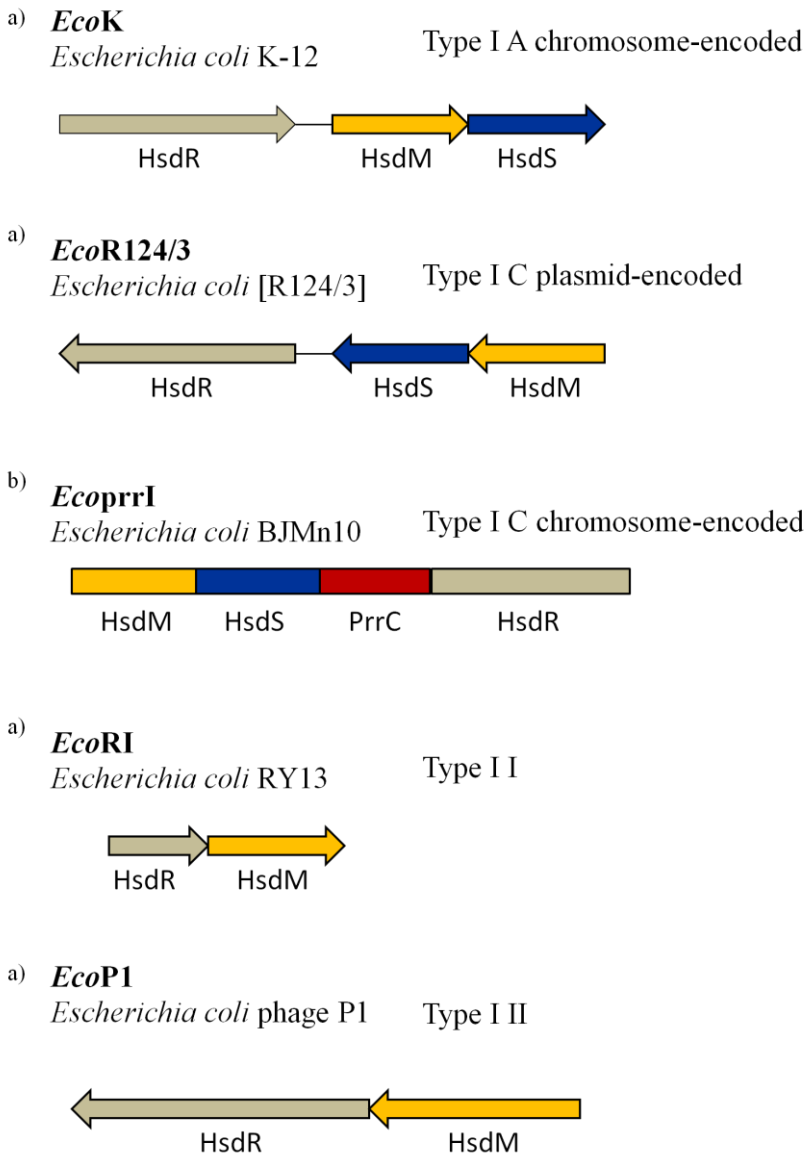



**Figure 1.1.** Schematic representation of type I, II and III Restriction-Modification systems. Type I systems are the most complex and can evolve easily, creating new specificities ( (a), after Tyndall *et al.*, 1994; (b), after Wilson & Murray, 1991). They are described in detail in section 1.2.1; HsdR, the restriction polypeptide; HsdM, the methylation polypeptide; HsdS, the specificity polypeptide; PrrC, a latent, phage T4 – induced anticodon nuclease.

Endonucleases could also appear as single genes, without methyltransferases, as a part of group I introns, recognizing asymmetric, degenerate and extended sequences of a range of around 20 base pairs. The insertion of an intron within these cut sequences changes their nucleotide structure and prevents the restrictase from further cleavage (Wilson & Murray, 1991).

The most closely studied solventogenic *Clostridium* species with an RM system is *C. acetobutylicum* (Nölling *et al.*, 2001), which has been studied from the early 1920s on, yet a functional methylation system was reported no earlier than over 80 years later, in 1993 (Mermelstein & Papoutsakis, 1993).

### 1.2.1. Type I systems, structure and characterization

Type I restriction-modification systems target specific DNA sequences and are the most diverse systems discovered so far. They encode one complex, multimeric enzyme (holoenzyme), composed of three subunits, whose behavior as an endonuclease or methyltransferase depends on the methylation state of the DNA. The hemimethylated DNA, appearing after DNA replication, is recognized by MTase and modified, while the unmethylated single- or double stranded DNA is cleaved by ENase in a random place (Murray, 2000; Wilson & Murray, 1991). The holoenzyme requires ATP for both its activities – as a restrictase or methyltransferase. Depending on the methylation status of the DNA, ATP stimulates either dissociation of the enzyme (methylated DNA), methylation of the second strand (hemimethylated DNA) or cleavage (non-methylated DNA). The enzyme consists of three subunits, HsdR, HsdM and HsdS, coded by *hsd* (host specificity for DNA) genes, all transcribed in the same direction. The MTase enzyme is composed of two HsdM and one HsdS subunits ($M_2S_1$), and an ENase enzyme is typically composed of an MTase and two HsdR subunits ($R_2M_2S_1$). In type I systems the restrictase (R) subunit is usually transcribed from a different promoter than the methylation (M) and specificity (S) subunits, with a promoter located upstream of the M subunit (Bickle & Krüger, 1993). These systems have been found to be the most adaptable to new specificities, most probably due to their unique modularity and complexity. The flexibility of HsdS and its

quick changes of specificity make type I RM a perfect defense system against fast evolving bacteriophages (Fig. 1.2) (Youell & Firman, 2012; Krüger & Bickle, 1983).

Although type I restriction enzymes recognize specific sequences of 6-7 base pairs, often separated by a non-specific spacer (Wilson & Murray, 1991), i.e. the sequence AACnnnnnnGTGC recognized by *Eco*KI (http://www.typei-rm.info), the cleavage itself may take place far from the recognition site and generally seems to be a random process. The modification always occurs within the recognized sequence at the N-6 group of adenine (Youell & Firman, 2012). The systems can be transferred between organisms without causing any harm to the recipient, which suggests a very tight control of their restriction activity, however it would be lethal to the host cell if the *hsdR* gene has been present and expressed prior to the transfer (Skrzypek & Piekarowicz, 1989).

Studies on the type IC enzyme *Eco*R124I have shown that, at low concentrations of the HsdR subunit, the restriction activity of the holoenzyme is relatively weak due to the poor stability of the $R_2M_2S_1$ complex. Even though the first HsdR subunit binds to trimeric MTase with high affinity, the attachment of the second subunit is much weaker and the complex dissociates easily. It is worth mentioning that the MTase-R complex shows modification activity on non-methylated DNA. This result stands in contrast to observations made on the type IA *Eco*KI enzyme, in which the $R_2M_2S_1$ complex appears to be very stable (Janscak *et al.*, 1998).

The methylation reaction differs between families, and while the family IA MTase prefers hemi-methylated DNA, the IB and IC families can methylate either hemimethylated or non-methylated DNA (in the absence of the ATP) (Youell and Firman, 2012; Price *et al.*, 1987).

The S subunit, often called a specificity subunit, contains two target recognition domains (TRDs) separated by two conserved domains (Fig 1.3) and is crucial for the recognition of the DNA sequence. TRD domains are responsible for binding to DNA, with TRD1 binding to 5’ part of the sequence and TRD2 binding to 3’ part of the sequence, and the conserved domains are important for protein-protein interaction. Due to the circular structure of the S subunit, TRDs can evolve quickly to bind new variations of sequences, which results in this RM system being extremely flexible. There are several ways to change the specificity of the S subunit described so far – swapping the TRDs sequences of related genes, unequal cross-over, affecting the spacer sequence, or the duplication of one TRD domain along

with the loss of the other one, leading to the recognition of the same fragment of sequence on both sense and antisense strands (Youell & Firman, 2012).

HsdS translation is controlled by a mechanism using an out-of-frame overlap of *hsdS* and *hsdM* genes (Youell & Firman, 2012). However, studies show the removal of the frameshift between the HsdS and HsdM genes of the *Eco*KI enzyme and the creation of a fusion protein does not affect the ability of the MTase to effectively methylate DNA (Roberts *et al.*, 2012).

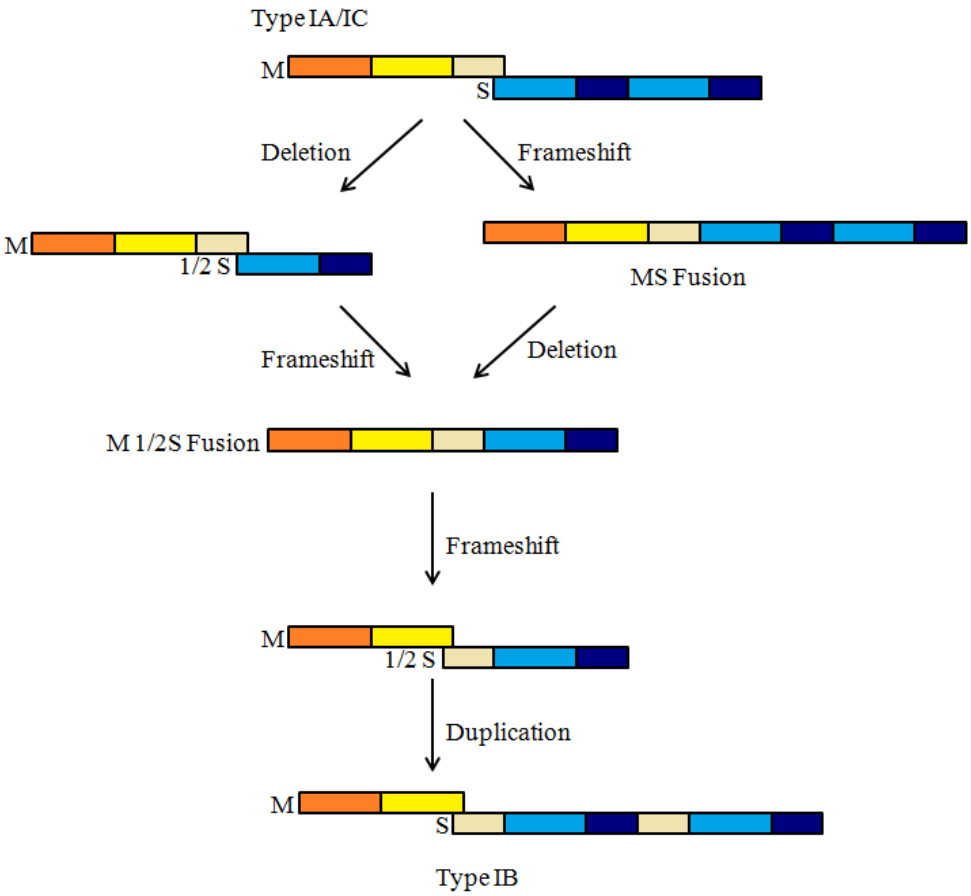



**Figure 1.2.** The possible evolutionary paths of the type IB RM methyltransferase from type IA/C RM methyltransferase (after Roberts *et al.*, 2012). The same color coding has been used to indicate the same structural domains during the mutagenesis process. The HsdM subunit is composed of a catalytic domain (yellow), N-terminal domain (orange) and C-terminal domain (grey). The HsdS subunit is composed of TRDs (target recognition domains; light blue) and conserved helical regions (dark blue). The process is initiated with the deletion of half of the S subunit, followed by the frameshift and fusion of the M and ½ S subunits, or first with the frameshift, followed by the fusion of the M and S subunits and then the deletion of one TRD and conserved helical region of the S subunit.

The M subunit is the second part of the methyltransferase enzyme. It contains a site for a cofactor and methyl donor, S-Adenosylmethionine (AdoMet), and an active site for DNA methylation and its sequence is more conserved than the S subunit (Murray, 2002).

The R subunit requires two co-factors, ATP and $Mg^{2+}$. It possesses a highly conserved region, called the motif X, which is crucial for DNA cleavage. Any mutations in that region can lead to a lack of restriction activity of the whole $R_2M_2S_1$ enzyme. The HsdR subunit can bind and nick single stranded DNA, but the cleavage of double-stranded DNA requires the recruiting of two R subunits, normally present on the whole holoenzyme (Youell & Firman, 2012). This theory was confirmed by Jindrova and colleagues (2005) in the work on cleavage mechanisms of type I restriction enzymes, showing how an *Eco*AI endonuclease mutant generated a high amount of nicked DNA, while the wild type enzyme could catalyze the cleavage of both strands.

### 1.2.2. Brief characterization of type I RM families A, B and C

Type I restriction – modification systems can be divided into five independent families, called IA, IB, IC, ID and IE, showing differences in antibody cross-reactivity, DNA hybridization pattern, subunit complementation and gene organization (Roberts *et al.*, 2012, Youell & Firman, 2012; Levitz *et al.*, 1990). In most cases the genes of the IA and IB families are located on a chromosome, while family IC members are encoded mainly on large conjugative plasmids (Youell & Firman, 2012).

The type IA family is the oldest and most closely studied. Many well-known enzymes, i.e. *Eco*KI, *Eco*BI or *Evo*DI, belong to this family. Its chromosome location and gene organization is similar to the IB family.

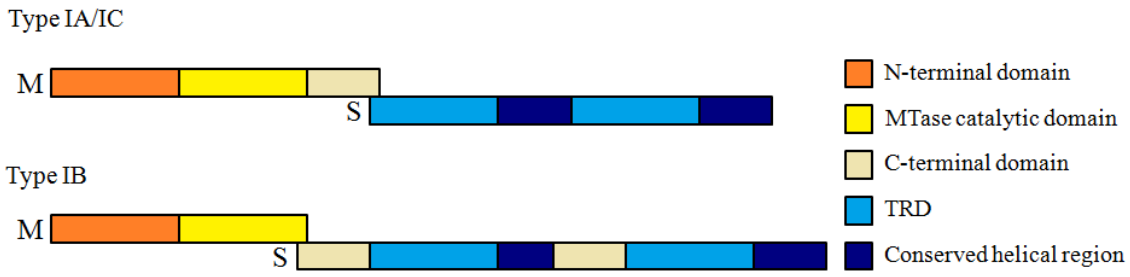



**Figure 1.3.** Comparison of type IA/IC and type IB restriction-modification methyltransferase enzymes (after Roberts *et al.*, 2012). Families IA and IC are closely related to each other, while type IB had evolved in different direction and presents different structure of both M and S subunits (compare with the figure 1.2).

Compared to other families, the IB family has a significantly smaller HsdM subunit and a much longer S subunit, with elongated conserved regions (Kannan *et al.*, 1989) as shown in figure 1.3.

Type IC family genes are very often plasmid-encoded, like the genes for *Eco*R124I, which are located on a large conjugative plasmid. The structural difference between the IC and IA or IB families is that in the IC family the *hsdR* and *hsdM-hsdS* operon order is reversed (Fig. 1.1). This gene organization causes additional expression of *hsdR* from the *hsdM-hsdS* promoter and in the case of the *Eco*R124I enzyme all three subunits are expressed and, in the presence of low amounts of R subunit, form a very unstable complex, dissociating from $R_2M_2S_1$to more stable $R_1M_2S_1$ (Kulik & Bickle, 1996).

## 1.3. Sugar metabolism in *Clostridium acetobutylicum* ATTC 824 and *Clostridium saccharobutylicum* NCP 266

### 1.3.1. Catabolic repression in bacteria

Many bacteria, including *C. acetobutylicum* and *C. saccharobutylicum*, would consume only one sugar at a time when introduced to medium containing multiple sugars. This specific metabolic regulation, called carbon catabolite repression, is common among various bacterial species (Servinsky *et al.*, 2010; Tangney *et al.*, 2003; Brückner *et al.*,

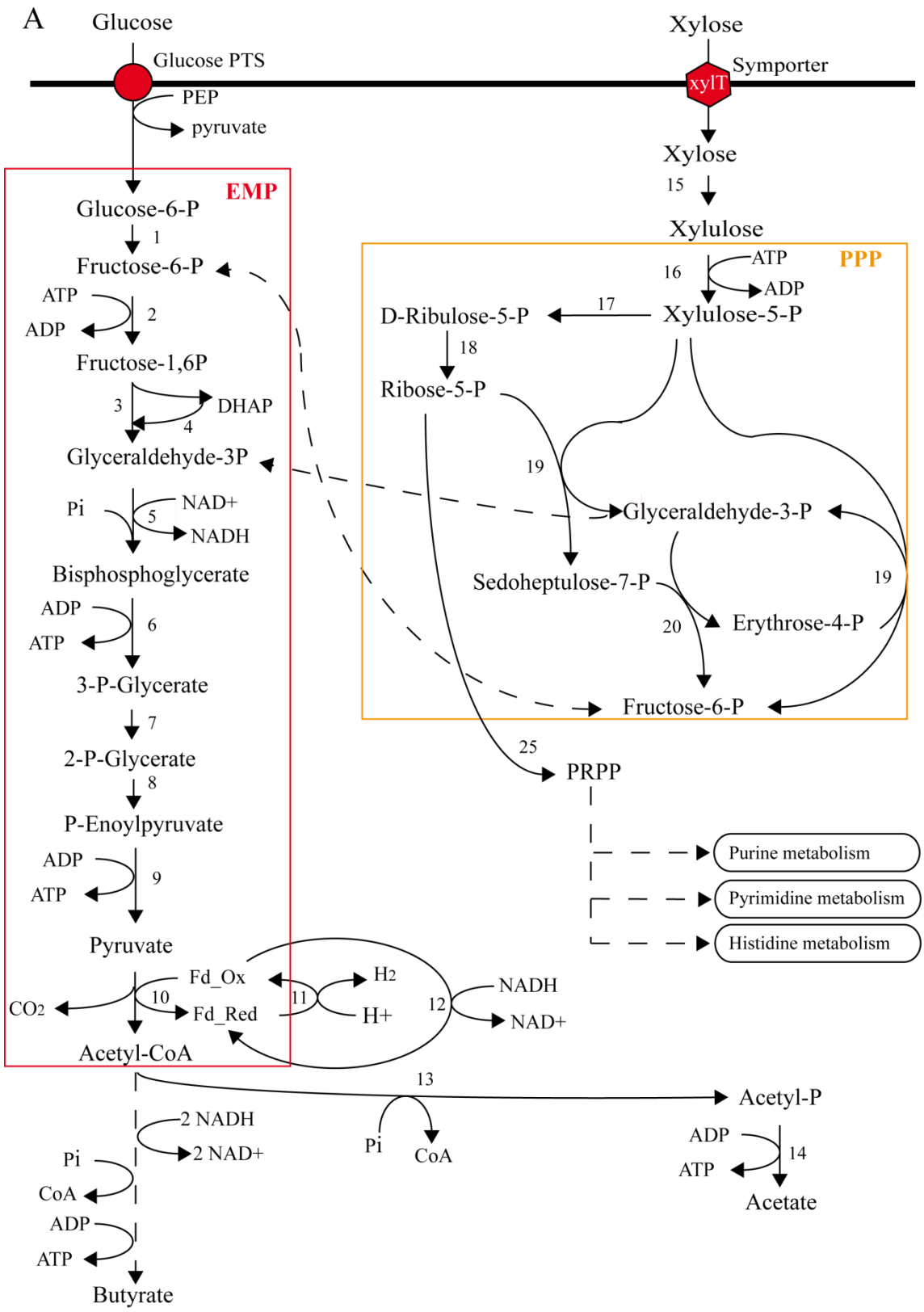
2002; Saier *et al.*, 1996), and in general microorganisms prefer monosaccharides to polysaccharides, hexoses to pentoses. In *C. acetobutylicum* the glucose inhibits the expression of genes responsible for xylose or arabinose metabolism even at high concentrations of these pentoses, and the repression lasts until all the glucose is consumed (Xiao *et al.*, 2011; Grimmler *et al.*, 2010). Similar behavior was described in *Bacillus subtilis* in the context of arabinose utilization (Sá-Nogueira & Mota, 1997). Diauxic growth was also observed in *Escherichia coli* grown on a mixture of glucose and lactose, where the organism would consume glucose prior to lactose, along with the strong inhibition of genes of lactose metabolism caused by the presence of glucose in the cell (Desai & Rao, 2010). In the past, studies have mostly focused on the induction of metabolic pathways in response to a single sugar. Elements of these pathways: the transporters, enzymes and helper proteins, are expressed under an often very tight control when a certain sugar or its intermediates in the downstream pathways were present in the medium (Saier *et al.*, 1996). These situations, however, are rarely if ever present in natural bacterial environments. Microorganisms are exposed to many carbon sources and the need to control their uptake and metabolism is crucial in conditions of constant competition for the most effective sources of energy.

### 1.3.2. Sugar metabolism in solventogenic clostridia

Knowledge of pentose metabolism in species of solventogenic clostridia is still very limited.

Although it known that many microorganisms use glucose as the most effective carbon and energy source, little is known about the metabolism of other simple sugars present in large amounts in plant biomass, for example D-xylose and L-arabinose.

Since plant material could serve as an easily accessible feedstock for biofuel-producing organisms, a more profound knowledge of their xylose and arabinose metabolism is needed. Some organisms, ABE-fermenting clostridia among them, can use these pentoses by hydrolyzing hemicelluloses – polysaccharides present along with cellulose in cell walls.


A
Glucose
Glucose PTS
PEP
pyruvate
Xylose
Symporter
xylT
Xylose
15
Xylulose
EMP
Glucose-6-P
1
Fructose-6-P
ATP
ADP
2
Fructose-1,6P
3
4
DHAP
Glyceraldehyde-3P
Pi
NAD+
5
NADH
Bisphosphoglycerate
ADP
ATP
6
3-P-Glycerate
7
2-P-Glycerate
8
P-Enoylpyruvate
ADP
ATP
9
Pyruvate
Fd_Ox
CO2
10
Fd_Red
11
H2
H+
12
NADH
NAD+
Acetyl-CoA
PPP
16
ATP
ADP
D-Ribulose-5-P
17
Xylulose-5-P
18
Ribose-5-P
19
Glyceraldehyde-3-P
Sedoheptulose-7-P
19
20
Erythrose-4-P
Fructose-6-P
25
PRPP
Purine metabolism
Pyrimidine metabolism
Histidine metabolism
13
Acetyl-P
2 NADH
2 NAD+
Pi
CoA
ADP
ATP
14
Acetate
Pi
CoA
ADP
ATP
Butyrate

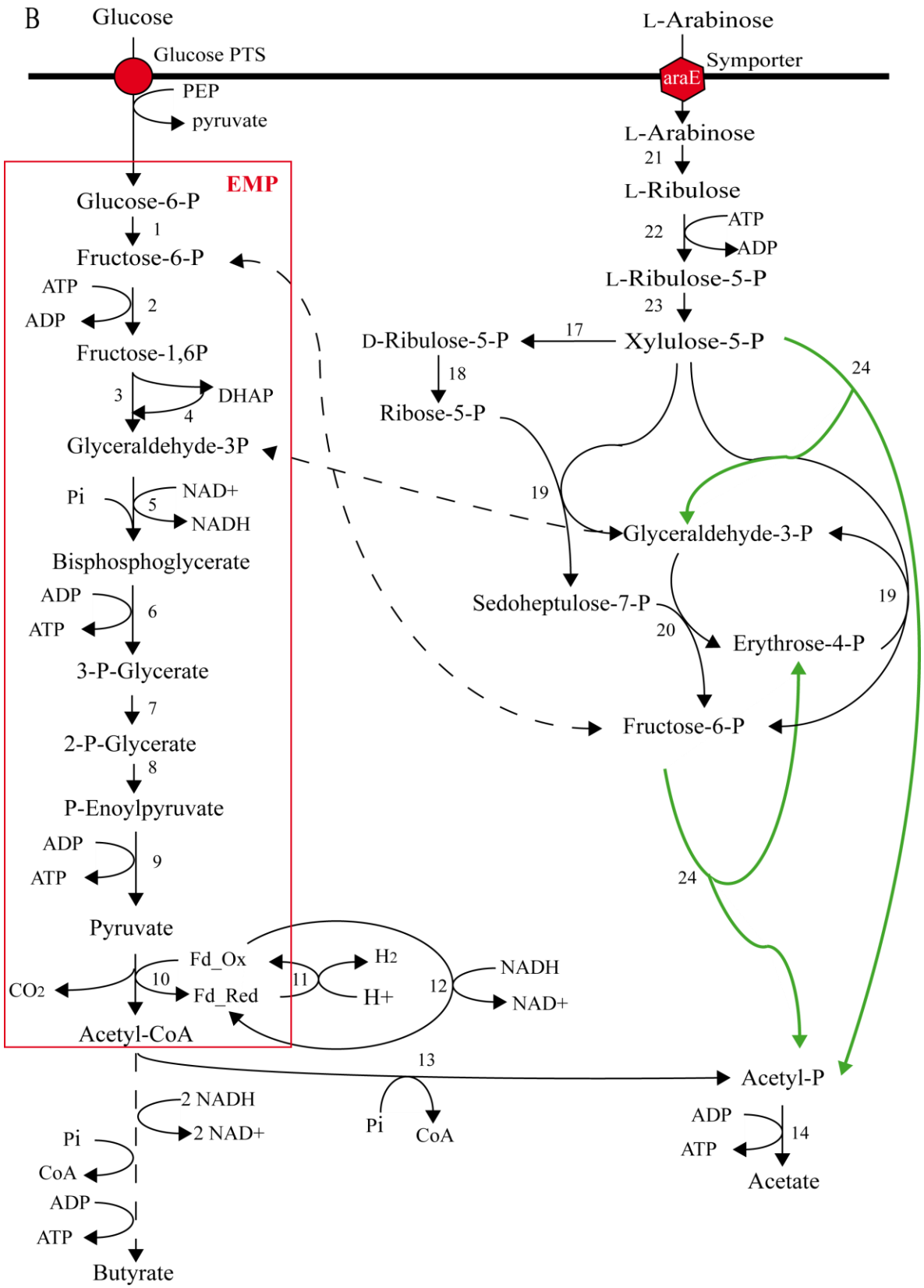



**Figure 1.4.** Schematic representation of xylose metabolism (A) via the pentose phosphate pathway (PPP) and arabinose metabolism (B) via the phosphoketolase pathway (PKP, as suggested by Servinsky *et al.*, 2012) and via the Embden-Meyerhof-Parnas (EMP) pathway. Enzymes taking part in a recently proposed phosphoketolase reaction as part of pentose degradation in

*C. acetobutylicum* are marked in green. Numbers by arrows indicate the enzymes of those pathways:

1, phosphoglucose isomerase; 2, phosphofructokinase; 3, fructose-bis-P aldolase; 4, triosephosphate isomerase; 5, glyceraldehyde-3-P dehydrogenase; 6, phosphoglycerate kinase; 7, phosphoglycerate mutase; 8, enolase; 9, pyruvate kinase; 10, pyruvate ferredoxin oxidoreductase; 11, hydrogenase; 12, $NAD^+$/NADH oxidoreductase; 13, phosphotransacetylase; 14, acetate kinase; 15, xylose isomerase; 16, xylulose kinase; 17, epimerase; 18, isomerase; 19, transketolase; 20, transaldolase; 21, arabinose isomerase; 22, ribulose kinase; 23, ribulose-5-phosphate epimerase; 24, phosphoketolase; 25, ribose-phosphate pyrophosphokinase; PRPP, phosphoribosyl pyrophosphate.

(after Servinsky *et al.*, 2012; Xiao *et al.*, 2011)

In general hexoses, such as glucose and mannose, are converted into pyruvate via the Embden-Meyerhoff-Parnas pathway (EMP), while pentoses are converted via the pentose phosphate pathway (PPP) and, among some microorganisms, such as heterofermentative lactic acid bacteria or *Bifidobacterium*, via the phosphoketolase pathway (PKP).

#### 1.3.2.1. Arabinose and xylose metabolism in *C. acetobutylicum* and *C. saccharobutylicum*

*C. acetobutylicum* and *C. saccharobutylicum* are capable of fermenting xylose and arabinose, and traditionally it was suggested both sugars are metabolized via the pentose phosphate pathway (PPP) (Zhang *et al.*, 2012; Xiao *et al.*, 2011; Grimmler *et al.*, 2010). During its non-oxidative branch there is firstly an interconversion of xylulose-5-P, D-ribulose-5-P and ribose-5-P and then the transfer of a glycoaldehyde or dihydroxyacetone group between sugar phosphates by two key enzymes, transketolase or transaldolase, respectively (Fig. 1.4 A). This part of the PPP is reversible, and by the action of ribose-phosphate pyrophosphokinase can lead to the formation of phosphoribosyl pyrophosphate (PRPP), an important intermediate for purine, pyrimidine and histidine metabolism (Karp *et al.*, 2005; Kanehisa *et al.*, 2014). L-Arabinose can enter the cell and be converted to L-ribulose by arabinose isomerase and then by ribulose kinase into L-ribulose-5-P. Ribulose-5-epimerase converts it to xylulose-5-P (X5P), which is metabolized through the PPP. The

conversion of arabinose to X5P requires one more step than the conversion of xylose, which should result in a growth advantage on xylose. However, for *Clostridium acetobutylicum* a much higher growth rate on arabinose than on xylose was observed (also in this study). The explanation for this phenomenon is not clear. It might be due to a lack of efficient xylose transport or the gene expression profile in *C. acetobutylicum*, especially the expression of transaldolase of the PPP (Servinsky *et al.*, 2012). Transcriptional analysis revealed the transketolase, transaldolase and epimerase genes are expressed with the same strength when the organism is growing on arabinose or on xylose, suggesting the differences in the metabolism of these two pentoses must result from another part of the pathway. It was suggested the crucial gene responsible for the much higher growth rate on arabinose is the phosphoketolase (*pkt*), CAC1343, and most recent studies have confirmed that CAC1343 is strongly induced by arabinose, but not xylose (Servinsky *et al.*, 2012). However, it has been shown recently the phosphoketolase pathway can also contribute up to 40% of the xylose catabolic flux in *C. acetobutylicum* (Liu *et al.*, 2012).

The phosphoketolase enzyme has two main functions, firstly it can convert fructose-6-P (F6P) and inorganic phosphate into erythrose-4-P and acetyl-P; and secondly, it converts xylulose-5-P (X5P) and inorganic phosphate into glyceraldehyde-3-P (G3P) and acetyl-P. Some of the phosphoketolase enzymes can act as bi-functional, promiscuous enzymes on both the intermediates – F6P and X5P (Meile *et al.*, 2001, Sánchez *et al.*, 2010; Kim & Gadd, 2008). The phosphoketolase pathway is associated mainly with *Bifidobacteria* and heterofermentative lactic acid bacteria, not with clostridia, and no phosphoketolase gene was identified in *C. saccharobutylicum* or *C. beijerinckii* (this study). The presence of the CAC1343 gene in the genome of *C. acetobutylicum* led to assumptions it could have been transferred via a horizontal gene-transfer event (Sánchez *et al.*, 2010).

All pentoses were thought to enter the Embden-Meyerhof-Parnas (EMP) pathway through the PPP, as glyceraldehye-3-P and fructose-6-P. Recent studies have revealed that during the growth of *C. acetobutylicum* on arabinose the *pkt* gene is strongly upregulated and the pentose metabolism might use the phosphoketolase pathway, instead of the PPP (Fig. 1.4 B), which would explain the advantage in growth rate compared to its growth on xylose (Servinsky *et al.*, 2012; Xiao *et al.*, 2011). Using a quantitative PCR reaction Zhang and colleagues have shown CAC1343 is 900-fold induced by arabinose compared to glucose and that its transcription is regulated by an arabinose regulator protein (AraR) (Zhang *et*

*al.*, 2012). With the same method it was shown that this gene was induced 3-fold by xylose, but 185-fold by arabinose. In this study they also showed the CAC1343 gene of *C. acetobutylicum* encodes a bi-functional phosphoketolase (Servinsky *et al.*, 2012).

### 1.3.2.2. Comparison of the pentose phosphate pathway and phosphoketolase pathway

The pathways of transport and metabolism of the main carbon sources are integrated into a central carbon metabolism (CCM). These pathways include, among others, the phosphotransferase system (PTS), the pentose phosphate pathway (PPP), the phosphoketolase pathway (PKP) or the tricarboxylic acid cycle (TCA) (Papagianni, 2012).

During the fermentation process for every 30 carbon molecules coming either from 5 glucose molecules or 6 pentose molecules that enter the EMP or PPP, 10 molecules of acetyl-CoA and acetyl-P are formed. However, if 6 molecules of pentose were metabolized via PKP, 12 molecules of acetyl-CoA and acetyl-P would be obtained. For EMP and PPP 30 carbon molecules would give the same yield of ATP, redox balance, and amount of G3P formed, and similar metabolic outputs should be produced during acidogenic growth. Servinsky and colleagues (2012) calculated the acetate:butyrate ratio during acidogenic growth on glucose, xylose and arabinose and demonstrated they were indeed similar in glucose- and xylose-grown cells (0.83 and 0.72, respectively). In the arabinose-grown culture the ratio was 1.95, indicating a higher production of acetate compared to butyrate (Servinsky *et al.*, 2012), and suggesting arabinose-derived carbon molecules can skip the EMP in the production of acetate (Fig. 1.4 B).

### 1.3.2.3. Xylose metabolism

D-xylose is a pentose sugar present in large amounts in plant biomass. There are several pathways of D-xylose metabolism known in nature: an oxido-reductase pathway present in eukaryotic organisms, the Weimberg pathway and Dahms pathway, both called oxidative pathways, and the fourth one - the isomerase pathway. The last three are typically used by prokaryotes (Wang R *et al.*, 2013 and 2011; Dahms, 1974; Weimberg, 1961). Solventogenic clostridia probably use the xylose isomerase pathway to convert the

aldose into D-xylulose, and this metabolism is being studied in detail in *C. acetobutylicum* (Jin *et al.*, 2014; Li *et al.*, 2013; Liu *et al.*, 2012; Xiao *et al.*, 2011). In other solventogenic clostridia, such as *C. beijerinckii* or *C. saccharobutylicum*, which are also known to produce high yields of biobutanol, the mechanism remains mostly unexplored.

The isomerase pathway converts D-xylose into D-xylulose using an enzyme called xylose isomerase (XylA). Next, the xylulose kinase (XylB) phosphorylates D-xylulose to yield D-xylulose-5-phospate (Grimmler *et al.*, 2010).

#### 1.3.2.4. Xylulose kinase (*xylB*) of *C. acetobutylicum* and *C. saccharobutylicum*

Xylulose kinase is an important enzyme of xylose metabolism in many organisms (Bunker *et al.*, 2013; Feng & Zhao, 2013; Cao *et al.*, 2013; Wang *et al.*, 2011), phosphorylating xylulose into xylulose-5-P (X5P). In the genomes of *C. acetobutylicum* and *C. saccharobutylicum* there are two *xylB* paralogs. According to studies on xylose metabolism in *C. acetobutylicum* two operons have been described – firstly, CAC1344 – CAC1349, independent from glucose repression, and secondly, CAC2610 – CAC2612, expressed only when all the glucose from the medium is depleted (Grimmler *et al.*, 2010). In recent studies it has been shown that the first operon is crucial for arabinose rather than xylose metabolism and this *xylB* gene was re-named *araK* (Zhang *et al.*, 2012). In *C. saccharobutylicum* two probable *xylB* genes are annotated, however no information on these genes or their function is available, and more details are needed for further strain improvements.

### 1.4. *Rhodobacter* nitrogen fixation (Rnf) system in clostridia

The *Rhodobacter capsulatus* nitrogen-fixating complex (Rnf) codes for a proton-translocating ferredoxin:$NAD^+$ oxidoreductase, involved in nitrogen metabolism through the conversion of energy (Biegel & Müller, 2010; Schmehl *et al.*, 1993). In the genus *Clostridium* it is composed of six subunits, RnfABCDGE and located in the cellular membrane (Fig 1.5). It is quite common among Gram-positive bacteria and clostridia, being present in *C. saccharobutylicum* (CSA00568 – CSA00573), yet not in *C. acetobutylicum* (Poehlein *et al.*, 2013; Tremblay *et al.*, 2013; Wang Y *et al.*, 2013 a; Müller *et al.*, 2008). In many bacterial species the Rnf complex is essential for nitrogen

fixation, hence the name, and plays the role of a system for electron transport to nitrogenase. Enzymes of the Rnf family are believed to be redox-driven ion pumps for $Na^+$ or $H^+$ transport across the membrane of a cell. In *R. capsulatus* and other nitrogen fixing bacteria the complex transports electrons from NADH to ferredoxin, with simultaneous proton transport across the membrane – from outside to inside (downhill), whereas in anaerobic bacteria the reaction is reversed – the transfer of positive ions to the outside of the cell is caused by electron transport from ferredoxin to $NAD^+$ (Tremblay *et al.*, 2012; Biegel *et al.*, 2011; Müller *et al.*, 2008; Li *et al.* 2006; Curatti *et al.*, 2005). In *Clostridium thermocellum* Rnf-like NADH:ferredoxin oxidoreductases were detected as part of $H_2$ generation pathways (Rydzak *et al.*, 2012).

Investigation of the Rnf complex mutants in different species of solventogenic *Clostrida* can give some insight into its role in energy conservation.

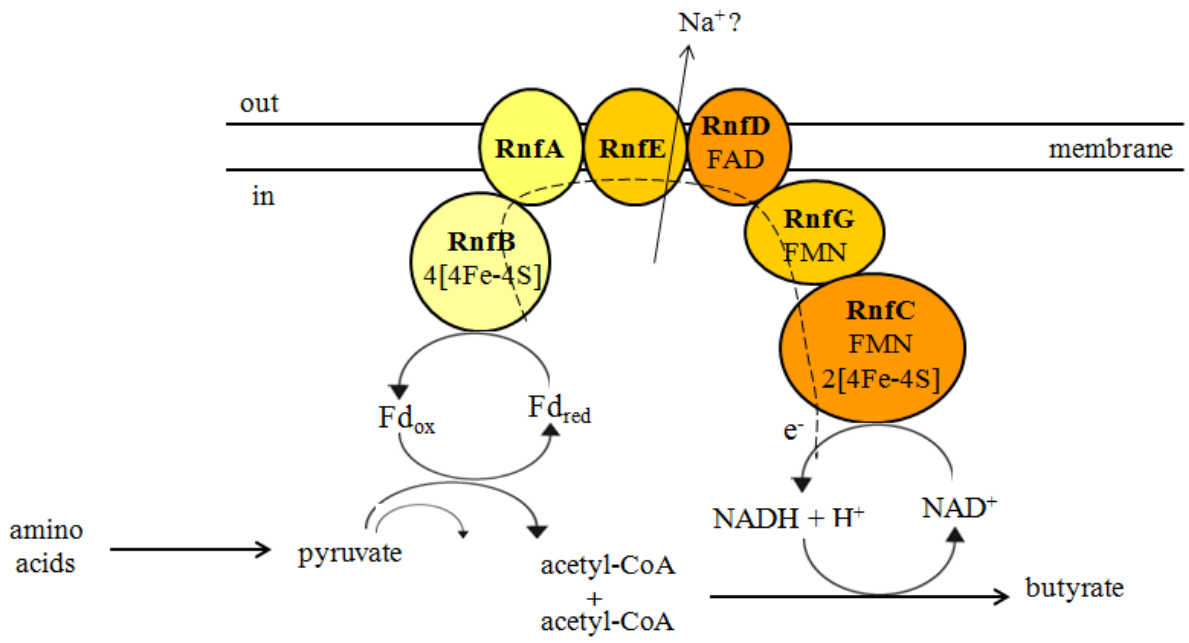



**Figure 1.5.** The schematic representation of the *Clostridium tetani* Rnf complex functions, showing the flow of electrons from reduced ferredoxin ($Fd_{red}$) through the Rnf complex to $NAD^+$, and therefore generating a positive ion gradient across the membrane (after Müller *et al.*, 2008). Similar subunits are marked with similar color coding; *ox*, oxidized; *red*, reduced.

## 1.5 Aim of this work

The main aim of this work was to analyze the xylose and arabinose metabolism in three major solventogenic *Clostridium* strains, *Clostridium acetobutylicum* ATCC 824, *C. saccharobutylicum* NCP 262 and *C. beijerinckii* NCIMB 8052. Based on the knowledge available at the beginning of the project along with the new data and experience collected during this work, not only was the role of arabinose and xylulose kinases investigated, but also the pentose and energy metabolism generally in all three strains. For these purposes various mutants in genes of different metabolic pathways were created. To this end, two methods were chosen, the ClosTron mutagenesis, which would disrupt a gene of interest by inserting an intron, and therefore could cause a polar effect on the genes located downstream, and the allelic exchange (ACE) method, which would produce a clean deletion, and markerless mutants, free of polar effect and which would allow multiple mutant creations. Additionally, a clean deletion system based on the *upp*-deficient strain for the allelic exchange in *C. acetobutylicum* should have been constructed prior to the creation of clean deletion mutants. Similarly, a methylation system to overcome the restrictases of the microorganism and an efficient way of introducing exogenous genetic material into host cells needed to be established prior to the application of the ClosTron method in *C. saccharobutylicum.* This part of the study was broadened by the investigation of the intron mutants of two restrictases of *C. saccharobutylicum*.

The ClosTron and clean deletion mutants of arabinose and xylulose kinases in *C. acetobutylicum* and *C. saccharobutylicum* and of phosphoketolase in *C. acetobutylicum* were analyzed regarding their phenotype, their ability to use pentoses as a carbon and energy source, their sugar consumption and fermentation profiles. Furthermore the gene transcription pattern of the wild type *C. acetobutylicum* in a xylose-supplemented, phosphate-limited continuous culture was analyzed and compared with results from previous experiments.

Additionally, a ClosTron mutant in the Rnf complex of *C. beijerinckii* was created to examine its phenotype and compare the results to previous studies on *Clostridium ljunghdalii*.

# 2. Materials and Methods

## 2.1. Strains and plasmids

All strains and plasmids used during this study are listed in table 2.1 and table 2.2.

**Table 2.1.** Strains used during this study

| Strain | Relevant characteristics | Reference |
|---|---|---|
| *Clostridium acetobutylicum* ATCC 824 | Wild type (WT) | Nölling, Breton *et al*. 2001 |
| *C. acetobutylicum ΔpyrE* | *C. acetobutylicum* ATCC 824_*ΔpyrE* | Heap *et al.*, 2012 |
| *C. acetobutylicum araK::int* | *C. acetobutylicum* ATCC 824_1344::*erm*B | This study |
| *C. acetobutylicum xylB::int* | *C. acetobutylicum* ATCC 824_2612::*erm*B | This study |
| *C. acetobutylicum pkt::1int* | *C. acetobutylicum* ATCC 824_1343::*ermB* | This study |
| *C. acetobutylicum pkt::2int* | *C. acetobutylicum* ATCC 824_1343::*ermB* | This study |
| *C. acetobutylicum ΔxylB* | *C. acetobutylicum* ATCC 824_Δ2612, Δ*pyrE* | This study |
| *C. acetobutylicum ΔaraKΔxylB* | *C. acetobutylicum* ATCC 824_Δ1344Δ261Δ*pyrE* | This study |
| *C. acetobutylicum araK::int* – pJL5.1 | *C. acetobutylicum* ATCC 824_1344::*ermB*, with complementation plasmid pJL5.1 | This study |

| | | |
|---|---|---|
| *C. acetobutylicum xylB::int* – pJL6.1 | *C. acetobutylicum* ATCC 824_2612::*ermB*, with complementation plasmid pJL6.1 | This study |
| *Clostridium saccharobutylicum* NCP 262 | Wild type (WT) | DSMZ 13864 |
| *C. saccharobutylicum hsdR1::int* | *C. saccharobutylicum* NCP 262_::00451::*ermB* | This study |
| *C. saccharobutylicum hsdR2::int* | *C. saccharobutylicum* NCP 262_02977::*ermB* | This study |
| *C. saccharobutylicum araK::int* | *C. saccharobutylicum* NCP 262_00774::*ermB* | This study |
| *C. saccharobutylicum xylB::int* | *C. saccharobutylicum* NCP 262_03335::*ermB* | This study |
| *C. saccharobutylicum araK::int* – pJL7.1 | *C. saccharobutylicum* NCP 262_00774::*ermB*, with complementation plasmid pJL7.1 | This study |
| *C. saccharobutylicum xylB::int* – pJL8.1 | *C. saccharobutylicum* NCP 262_03335::*ermB*, with complementation plasmid pJL8.1 | This study |
| *Clostridium beijerinckii* NCIMB 8052 | Wild type (WT) | NCIMB 8052 |
| *C. beijerinckii rnfC::int* | *Clostridium beijerinckii NCIMB 8052*_2449::*ermB* | This study |
| *E. coli* Top10 | F-, *mcr*A, φ80*lac*ZΔM15, Δ(*mrr-hsd*RMS-*mcr*BC), Δ*lac*X74, *rec*A1 *ara*D139, Δ(ara-leu)7697, *gal*U, *gal*K, *rps*L, ($Str^R$), | Invitrogen |

| | | |
|---|---|---|
| | *end*A1, *nup*G | |
| *E. coli* ER1793 | *mcrA*$^-$ *mcrBC*$^-$ *mrr*$^-$ | New England Biolabs |
| *E. coli* CA434 | HB101 carrying the IncPb conjugative plasmid, R702, Kan$^R$ | |
| *E. coli* Top10 – pJL1 | *E. coli* Top10 with pJL1 plasmid | This study |
| *E. coli* Top10 – pJL2 | *E. coli* Top10 with pJL2 plasmid | This study |
| *E. coli* – pAN1 | *E. coli* ER2275 with pAN1 plasmid | Mermelstein & Papoutsakis, 1993 |
| *E. coli* – pAN2 | *E. coli* ER2275 with pAN2 plasmid | Mermelstein & Papoutsakis, 1993 |
| *E. coli* ER2275 | *endA*1, *recA*1, Δ(*mcr-hsd-mrr*) | Prof. Dr. E. T. Papoutsakis; Evanston, USA |

**Table 2.2.** Plasmids used during this study

| Plasmid | Size (kbs) | Relevant characteristics | Reference |
|---|---|---|---|
| pACYC184 | 4.245 | Expression vector in *E. coli*, p15A1, Tc$^R$ Cm$^R$ | DSMZ |
| pUC19 | 2.686 | Expression vector in *E. coli*, Amp$^R$ | Vieira & Messing, 1982 |
| pAN1 | 7.000 | oriR, φ3tI, p15A, Cm$^R$ | Mermelstein & Papoutsakis 1993 |
| pAN2 | 7.000 | oriR, φ3tI, p15A, Tet$^R$ | Mermelstein & Papoutsakis 1993 |

| | | | |
|---|---|---|---|
| pCH1 | 4.328 | ori repL, ori pBR322, $Amp^R$, $Erm^R$ | Held, PhD thesis, 2013 |
| pCLF1-1 | 4.878 | *Flp*, $Cm^R$ | Nigel Minton, University of Nottingham |
| pMTL-ME3 | 3.561 | Modular plasmid for clean deletion in *Clostridium*, $Cm^R$, *Csp pyrE* | Ehsaan, PhD thesis, 2013 |
| pMTL-ME6X | 4.930 | Modular plasmid, $p_{fdx}$, $Cm^R$ | Ehsaan, PhD thesis, 2013 |
| pJet | 2.974 | ori pMB1,$_pT7$, $Amp^R$ | Thermo Scientific |
| pJL1 | 5.733 | Derived from pACYC184, hsdMSI, $_pT7$, $Tc^R$ | This study |
| pJL2 | 6.314 | Derived from pACYC184, hsdMSII, $_pT7$, $Tc^R$ | This study |
| pJL3 | 4.540 | pMTL-ME3-derived vector for clean deletion of *C. acetobutylicum araK* | This study |
| pJL4 | 4.519 | pMTL-ME3-derived vector for clean deletion of *C. acetobutylicum xylB* | This study |
| pJL5.1 | 6.604 | pMTL-ME6X-derived vector for *C. acetobutylicum araK::int* complementation; $p_{fdx}$, | This study |
| pJL6.1 | 6.178 | pMTL-ME6X-derived vector for *C. acetobutylicum xylB::int* complementation; $p_{fdx}$, | This study |
| pJL7.1 | 6.357 | pMTL-ME6X-derived vector for *C. saccharobutylicum araK::int* complementation; $p_{fdx}$, | This study |
| pJL8.1 | 6.288 | pMTL-ME6X-derived vector for | This study |

| | | | |
|---|---|---|---|
| | | *C. saccharobutylicum xylB::int* complementation; $p_{fdx}$, | |
| pJL9-520 | 5.192 | pCH1-derived vector for clean deletion of *C. acetobutylicum upp* | This study |
| pJL10 | 4.922 | pCH1-derived vector for clean deletion of *C. acetobutylicum upp* | This study |
| pJL11 | 5.522 | pCH1-derived vector for clean deletion of *C. acetobutylicum upp* | This study |
| pJL11-RecA | 6.723 | pCH1-derived vector for clean deletion of *C. acetobutylicum upp*; *recA* | This study |
| pMADK*upp*$^-$ | 8.868 | pMADK derivative with homologous regions for the deletion of the uracil phosphoribosyltransferase gene (*upp*), $Amp^R Cm^R Erm^R$ | Kraußе, PhD thesis, 2012 |
| pMTL007C-E2 | 8.925 | *E. coli-Clostridium* shuttle vector with the Ll.ltrB group II intron containing an *ermB* cassette activated by retrotransposition, $Cm^R$ | Heap *et al.*, 2010 |
| pMTL007C-E2::*Cac araK*-260\|261a | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. acetobutylicum araK* gene (CAC1344) | This study |
| pMTL007C-E2::*Cac pkt.1*-1887\|1888s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. acetobutylicum pkt* gene (CAC1343), rear part | This study |
| pMTL007C-E2::*Cac pkt.2*-621\|622s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. acetobutylicum pkt* gene | This study |

| | | | |
|---|---|---|---|
| | | (CAC1343), middle part | |
| pMTL007C-E2::*Cac upp*-84\|85a | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. acetobutylicum upp* gene (CAC2879) | This study |
| pMTL007C-E2::*Cac xylB*-843\|844s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. acetobutylicum xylB* gene (CAC612) | This study |
| pMTL007C-E2::*Csa araK*-1104\|1105s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. saccharobutylicum araK* gene (CSA00774) | This study |
| pMTL007C-E2::*Csa xylB*-1143\|1144s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. saccharobutylicum xylB* gene (CSA03335) | This study |
| pMTL007C-E2::*Csa_hsdR1*-2565\|2566s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. saccharobutylicum hsdR1* gene (CSA00451) | This study |
| pMTL007C-E2::*Csa_hsdR2*-2220\|2221s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. saccharobutylicum hsdR2* gene (CSA02977) | This study |
| pMTL007S-E2 | 9.196 | *E. coli-Clostridium* shuttle vector with the Ll.ltrB group II intron containing an *ermB* cassette activated by retrotransposition, $Sp^R$ | Heap *et al.*, 2010 |
| pMTL007S-E2::*Cbei rnfC*-783\|784s | ~11.500 | pMTL007C-E2-derived plasmid with re-targeted region for *C. beijerinckii rnfC* gene | This study |

| | | | |
|---|---|---|---|
| | | (CBEI2449) | |
| pMTL82151 | 5.254 | pBP1, $Cm^R$, ColE1 + tra, MSC | Heap *et al.*, 2009 |
| pMTL83151 | 4.476 | pCB102, $Cm^R$, ColE1 + tra, MSC | Heap *et al.*, 2009 |
| pMTL84151 | 6.297 | pCD6, $Cm^R$, ColE1 + tra, MSC | Heap *et al.*, 2009 |
| pMTL85151 | 3.729 | pIM13, $Cm^R$, ColE1 + tra, MSC | Heap *et al.*, 2009 |

## 2.2. Media, buffers and supplements

Media, buffers and solutions were dissolved in distilled or double-distilled water and autoclaved for 20 min at 121°C, at a pressure of 2 bars if not otherwise indicated. Thermosensitive solutions, vitamins and antibiotics stocks were filtered (Spritzenvorsatzfilter, 0.45 µm, Sartorius) into sterile flasks.

Solid media were prepared with the addition of 1.5% (w/v) agar and autoclaved.

### 2.2.1 Complex media

Complex (rich) media used for culture maintenance and genetic manipulations were dissolved in distilled water ($H_2O$). Media used for growth curves, gas chromatography experiments and enzymatic assays were dissolved in double-distilled water ($ddH_2O$). Antibiotics were added to maintain plasmids or mutant strain selection after the autoclaving step.

LB (Luria-Bertani)-Medium (Sambrook 2001)

| | |
|---|---|
| Tryptone | 10 g |
| Yeast extract | 5 g |
| NaCl | 10 g |
| $H_2O$ | to 1000 ml |

Modified 2 x YTG-Medium

| | |
|---|---|
| Glucose | 7.5 g |
| Tryptone | 16 g |
| Yeast extract | 10 g |
| NaCl | 5 g |
| $H_2O$ | to 1000 ml |

CGM (Clostridial growth medium, Hartmanis and Gatenbeck, 1984, modified)

| | |
|---|---|
| Glucose | 50 g |
| $KH_2PO_4$ | 0.75 g |
| $K_2HPO_4$ | 0.75 g |
| $(NH_4)_2SO_4$ | 2 g |
| $MgSO_4$ x 7 $H_2O$ | 0.71 g |
| $MnSO_4$ x $H_2O$ | 0.01 g |
| $FeSO_4$ x 7 $H_2O$ | 0.01 g |
| NaCl | 1 g |
| Yeast extract | 5 g |
| Asparagine | 2 g |
| $H_2O$ | to 1000 ml |
| pH | 6.9 |

The pH was set using NaOH before autoclaving. Glucose, asparagine and iron sulphate solutions were sterile filtered and added after autoclaving.

Reinforced Clostridial Medium (Difco, BD)

The RCM was always prepared according to the recommendations of the manufacturer. There was no need to establish the pH and the medium was already supplemented with carbon sources. If needed, the medium was additionally supplemented with glucose as indicated in the experiment.

### 2.2.2 Synthetic media, vitamin and salt solutions

Minimal (synthetic) media

All minimal media were dissolved in double distilled water ($ddH_2O$). Vitamins and additional solutions were added according to the requirements of the recipe or the requirements of a strain. Antibiotics were added to maintain plasmids or for mutant strain selection.

MM-MES (Mineral Medium MES-based) (synthetic medium, modified)

| | |
|---|---|
| Glucose | 60 g |
| $(NH_4)_2SO_4$ | 2 g |
| $MgSO_4$ x 7 $H_2O$ | 0.22 g |
| $KH_2PO_4$ | 0.55 g |
| $K_2HPO_4$ | 0.55 g |
| $FeSO_4$ x 7 $H_2O$ | 0.011 g |
| PABA | 8 mg/ml |
| Biotin | 0.1 mg/ml |
| MES | 21.3 g |
| $ddH_2O$ | to 1000 ml |
| pH | 6.6 |

The pH was set using NaOH and sterile filtered, not autoclaved.

For spore formation an optional addition of 2.3 ml of acetate was added.

In the standard MM-MES recipe the pH was set using 30% $NH_4OH$ and there was no supplementation with $(NH_4)_2SO_4$. However, ammonium hydroxide is highly volatile and worse growth in the standard medium was observed, therefore this change was proposed as a solution for the problem.

Vitamin solutions for MM-MES

Biotin solution, 1000 x concentrated

| | |
|---|---|
| Biotin | 10 mg |
| $ddH_2O$ | 100 ml |

P-Aminobenzoic Acid solution (PABA), 500 x concentrated

| | |
|---|---|
| PABA | 400 mg |
| $ddH_2O$ | 100 ml |

Vitamin solutions should be prepared in advance and kept at 4°C.

CBM (Clostridial Basal Medium) (O'Brien & Morris, 1971, modified)

| | |
|---|---|
| $MgSO_4$ x 7 $H_2O$ | 200 mg |
| $MnSO_4$ x $H_2O$ | 7.58 ml of a 1 mg/ml stock |
| $FeSO_4$ x 7 $H_2O$ | 10 ml of a 1 mg/ml stock |
| PABA | 1 ml of a 1 mg/ml stock |
| Biotin | 20 µl of a 0.1 mg/ml stock |
| Thiamine-HCl | 1 ml of a 1 mg/ml stock |
| Casein hydrolysate (enzymatic only, Oxoid) | 4.0 g |
| $ddH_2O$ | to 800 ml |

For 1.5% agar 15 g of bacteriological agar was added at this stage.

The medium was sterilized by autoclaving and supplemented with sterile filtered:

| | |
|---|---|
| Glucose (1% final) | 10 g |
| $KH_2PO_4$ | 10 ml of a 50 mg/ml stock |
| $K_2HPO_4$ | 10 ml of a 50 mg/ml stock |
| $ddH_2O$ | to 200 ml |

At this stage the pH should have reached ~6.9.

For sporulation or solventogenesis assays and batch cultures 5% (w/v) glucose was used, and 5g/l of $CaCO_3$ was added to the medium prior to autoclaving.

This medium was used mainly to induce sporulation of *Clostridium* strains and for spore stock preparations.

All continuous culture experiments were prepared in a phosphate-limited medium.

Sugars, vitamins, $FeSO_4$ x 7 $H_2O$ solution and all salt solutions were sterilized using a 0.45 µm filter and added after autoclaving.

MMVK Medium (Minimal Medium for precultures)

| | |
|---|---|
| Glucose | 20 g |
| $CaCO_3$ | 1 g |
| $KH_2PO_4$ | 1 g |
| $K_2HPO_4$ x 3 $H_2O$ | 1 g |
| $(NH_4)_2SO_4$ | 2 g |
| $MgSO_4$ x 7 $H_2O$ | 0.1 g |
| $MnSO_4$ x $H_2O$ | 0.01 g |
| $FeSO_4$ x $H_2O$ | 0.015 g |
| NaCl | 0.01 g |
| $Na_2MoO_4$ x $H_2O$ | 0.01 g |
| $CaCl_2$ x $H_2O$ | 0.01 g |
| Biotin | 0.1 mg |
| Thiamin - HCl | 2 mg |
| p-Aminobenzoic Acid | 2 mg |
| Resazurin (0.1% (w/v)) | 1 ml |
| $Na_2S_2O_4$ | 0.035 g |
| dd$H_2O$ | to 1000 ml |

PLMM (Phosphate-limited minimal medium for fermentors, modified

| | |
|---|---|
| Carbon source | 600 g |
| $KH_2PO_4$ | 1.05 g |
| $(NH_4)_2SO_4$ | 30 g |
| $MgSO_4$ x 7 $H_2O$ | 1.5 g |
| $MnSO_4$ x $H_2O$ | 0.225 g |
| $FeSO_4$ x $H_2O$ | 0.225 g |
| NaCl | 0.15 g |
| $Na_2MoO_4$ x $H_2O$ | 0.15 g |
| $CaCl_2$ x $H_2O$ | 0.15 g |
| Biotin | 1.5 mg |
| Thiamin - HCl | 0.03 g |
| p-Aminobenzoic Acid | 0.03 g |
| dd$H_2O$ | to 15 l |

The pH was adjusted to a value of 2.0 with $H_2SO_4$.

Medium supplements:

Salt solution for minimal media 100 x concentrated

| | |
|---|---|
| $MgSO_4$ x 7 $H_2O$ | 1 g |
| NaCl | 0.1g |
| $Na_2MoO_4$ x 2 $H_2O$ | 0.1 g |
| $CaCl_2$ x 2 $H_2O$ | 0.1 g |
| $MnSO_4$ x $H_2O$ | 0.15 g |
| $ddH_2O$ | to 100 ml |

Vitamin solution for minimal media, 100 x concentrated

| | |
|---|---|
| Biotin | 10 mg |
| Thiamin – HCl | 200 mg |
| p-Aminobenzoic acid | 200mg |
| $ddH_2O$ | to 1000 ml |

$FeSO_4$ x 7 $H_2O$ solution (15 mg/ml)

| | |
|---|---|
| $FeSO_4$ x 7 $H_2O$ | 150 mg |
| $ddH_2O$ | to 10 ml |

The pH was adjusted to a value of 2 with $H_2SO_4$

$Na_2S_2O_4$ solution (35 mg /ml)

$Na_2S_2O_4$ was used as a reduction agent for residues of oxygen remaining in the medium and was added at the very end of the media preparation.

| | |
|---|---|
| $Na_2S_2O_4$ | 175 mg |
| $ddH_2O$ | to 5 ml |

$Na_2S_2O_4$ solution should be stored anaerobically.

All solutions were sterile filtered and stored at 4°C.

### 2.2.3. Additional solutions

Sugar solutions (50%)

| | |
|---|---|
| Glucose | 250 g |
| $ddH_2O$ | 500 ml |
| Autoclaved | |

| | |
|---|---|
| Xylose | 250 g |
| $ddH_2O$ | to 500 ml |
| Autoclaved | |

| | |
|---|---|
| Arabinose | 250 g |
| $ddH_2O$ | to 500 ml |
| Autoclaved | |

| | |
|---|---|
| Fructose | 250 g |
| $ddH_2O$ | to 500 ml |
| Sterile filtered | |

| | |
|---|---|
| Ribose | 250 g |
| $ddH_2O$ | to 500 ml |
| Sterile filtered | |

The glucose, xylose and arabinose were autoclaved in normal conditions for 15 minutes instead of 20 minutes. Only crystal clear solutions were used as sugar stock solutions, as darkening indicated sugar degradation due to high temperature, which alters its concentration. Fructose and ribose were sterile filtered using 45 µm filters.

All sugars were dissolved in double distilled water heated up to 60°C.

### 2.2.4. Antibiotics and additives

| **Antibiotics** | **Stock solution** | **Working concentrations** | |
|---|---|---|---|
| | | *E. coli* | Clostridia |
| Ampicilin | 100 mg/ml in $H_2O$ | 100 µg/ml | – |
| Chloramphenicol | 25 mg/ml in EtOH (90% (v/v)) | 25 µg/ml | – |
| Thiamphenicol | 15 mg/ml in EtOH (90% (v/v)) | - | 15 µg/ml |
| Spectinomycin | 750 mg/ml in $H_2O$ | - | 750 µg/ml |
| Tetracycline | 10 mg/ml in EtOH (70% (v/v)) | 10 µg/ml | – |
| Kanamycin | 50 mg/ml in $H_2O$ | 50 µg/ml | – |
| Clarithromycin | 5 mg/ml in EtOH (90% (v/v)) | – | 5 µg/ml |
| Erythromycin | 5 mg/ml in EtOH (90% (v/v)) | – | 5 µg/ml |
| Lincomycin | 5 mg/ml in EtOH (90% (v/v)) | – | 5 µg/ml |
| Polymyxin B | 40 mg/ml in $H_2O$ | 40 µg/ml | – |
| Cefoxitin | 10 mg/ml in $H_2O$ | 10 µg/ml | – |
| Colistin | 10 mg/ml in $H_2O$ | 10 µg/ml | – |
| D-cycloserine | 25mg/ml in $H_2O$ | 25 µg/ml | – |
| 5-fluorouracil (5-FU) | 200 mg/ml In DMSO | – | 200 µg/ml |
| 5-fluoroorotic acid (5-FOA) | 1 g/ml in DMSO | – | 2 mg/ml |

Antibiotic solutions in 70%-100% alcohol can be used without filtration; all other solutions should be sterile filtered. Antibiotic stocks must be kept at -20°C in aliquots to avoid repetitive freeze/thaw cycles.

### 2.2.5. Commonly used buffers

PBS buffer

| | |
|---|---|
| NaCl | 8 g |
| KCl | 0.2g |
| $KH_2PO_4$ | 0.24 g |
| $Na_2HPO_4$ | 1.44 g |
| $ddH_2O$ | to 1000 ml |

20x TAE Buffer (DNA Gel electrophoresis)

| | |
|---|---|
| Tris | 96.9 g |
| Acetic Acid (100%) | 22.8 g |
| EDTA | 3.8 g |
| $H_2O$ | to 1000 ml |

TE Buffer (pH 8.0)

| | |
|---|---|
| Tris | 10 mM |
| EDTA | 1 mM |

Tris Buffer (pH 8.0)

| | |
|---|---|
| Tris | 1 M |

The pH was set with HCl.

## 2.3. Cultivation methods

During the cultivation of liquid cultures of *E. coli* and *Clostridium* strains samples for optical density determination were taken at every time point and measured using the Spectrophotometer (Ultrospec 2100 *pro*, Amersham Biosciences). For this machine a believable range was set between 0.1 – 1.0, and therefore dilutions were made if necessary. For clostridial growth curves, samples for pH measurements were also gathered (non-diluted) and measured using a pH Meter (pH 720 WTW series, inoLab).

Three technical replicates were made for the optical density measurement of all samples taken from hungate or batch culture experiments. The OD values were measured three times for each sample; the average value was calculated and used as the right optical density value of a culture. The measurements were considered significant if the difference between the average and the measured OD values was below 0.005, and only then were they used for a graphic representation of an experiment.

Similarly, three technical replicated were made for the pH measurements, and the results were considered significant if the difference between the average and the measured pH values was below 0.02.

### 2.3.1. Cultivation of *E. coli* strains

All *E. coli* cultures were prepared in an LB medium with appropriate antibiotics at 37°C with shaking. Fresh cultures for chemocompetent cell preparation or conjugation were inoculated with overnight cultures to the $OD_{600}$ of ~0.1. For plasmid isolation overnight cultures were used. Colonies kept on solid media were re-streaked onto a fresh plate every two weeks.

For strain preservation glycerol stocks (30% glycerol, 70% LB medium) were prepared from 2 ml of overnight cultures, spun down at 5000 rpm for 5 min and resuspended in 1 ml of glycerol-LB solution.

### 2.3.2. Cultivation of *Clostridium* strains

All three *Clostridium* strains required different liquid and solid media for cultivation and maintenance. A short list of strains and their cultivation requirements for strain maintenance are presented below:

| | **Medium** | |
|---|---|---|
| **Organism** | Liquid | Solid |
| *C. acetobutylicum* ATCC 824 | CGM | CGM |
| *C. saccharobutylicum* NCP 262 | CGM | 2xYTG |
| *C. beijerinckii* NCIMB8052 | RCM | RCM or 2xYTG |

Heat shock procedure

The heat shock procedure is a simple way to induce a germination process of clostridial spores. The spore suspension (100-200 µl, depending on the stock culture quality) should be incubated for 10 minutes at 80°C for *C. acetobutylicum* or *C. beijerinckii* and 10 minutes at 70°C for *C. saccharobutylicum*. After incubation the total suspension volume should be plated on agar plates with no antibiotic selection and left at 37°C for 2-4 days.

Hungate and batch cultures

Clostridial cultures for DNA isolation were prepared in 5 ml of an appropriate, rich medium in anaerobic hungate tubes. The precultures for batch cultures were prepared in 5 ml of anaerobic, rich or minimal medium (in hungates), or in 20 ml of anaerobic medium (in 100 ml bottles), and then used to inoculate the appropriate amount of medium in anaerobic bottles, so the starting $OD_{600}$ value was between ~0.1 – 0.25. For standard growth curves in fermentation profile experiments 60 – 100 ml of medium was used.

Continuous fermentation

Continuous culture is an alternative for batch culture experiments, enabling more detailed and precise observation of strain behavior during the acidogenesis and solventogenesis and the switch between both. Throughout the entire fermentation process fresh medium is constantly added to the fermentation flask, while the same amount of culture is washed away and collected in the so called 'waste bottle'. Therefore the main limitation on the duration of the continuous culture is the volume of fresh medium. The pH value is constantly monitored by a built-in electrode and could be changed as necessary.

Continuous cultures were preparedin a small scale fermentor with a volume of 15 l. The experiment was designed to investigate the strain behavior in precisely described and maintained conditions, with all parameters constantly monitored. The pH values were constant, except for the switch from acidogenesis to solventogenesis.

The cultures were started from spores by a heat shock and the first liquid culture was prepared in a rich medium (CGM). The two following liquid cultures were done in an anaerobic minimal medium for precultures (MMVK), and later introduced into a big, 15 l fermentor flask with a phosphate limited medium (PLMM) in a Biostat B PLUS fermentor system (Sartorius). Around 10-12 h after inoculation of the PLMM medium the culture had reached an $OD_{600}$ value high enough to start the fermentation experiment, and from this moment the $OD_{600}$ values and sugar concentration were measured every 24 hours. The dilution rate was set at 0.075 $h^{-1}$. The pH was maintained at pH 5.7 by the addition of 2M KOH. When the optical density reached stable values and cells entered steady acidogenic growth and the first two 13 ml sample for RNA isolation had been taken, the switch to a solventogenic state was induced by the discontinuation of the KOH addition. At each pH shift – at the pH values of 5.5, 5.2, 4.9, 4.7 and 4.5, two 13 ml samples for RNA isolation were taken. A steady solventogenic growth was reached when the optical density values stabilized, the last two samples for RNA isolation were taken (pH 4.5) and the fermentation was considered finished (Grimmler *et al.*, 2011). The phosphate-limited continuous cultures were performed as described previously (Bahl *et al.*, 1982; Fischer *et al.*, 2006). For the fermentation profile, supernatants from 2 ml of culture were frozen and stored at -20°C. For microarray analysis, cell pellets from 10 ml of culture were collected by

centrifugation at -20°C, 9000 rpm (Hettich Universal 320R, Sartorius) for 5 min and immediately frozen in liquid nitrogen. Cell pellets were stored at -80°C.

Plate and PCR tests for confirmation of pSOL1 presence in *C. acetobutylicum* colonies

To confirm the presence of the pSOL1 plasmid in *C. acetobutylicum* ATCC 824 single mutant or wild type colonies were streaked out on plates containing starch (RCM-agar) and incubated for at least 24 h at 37°C. If the pSOL1 was still present in cells; the starch would be degraded by the alpha-amylase, a large extracellular protein encoded by *amyP* gene, located on the megaplasmid. The reaction with Lugol's iodine would therefore result in a large halo appearing around colonies (Fig. 2.1). Each mutant strain was tested for the presence of the megaplasmid before preparing the spore stock solution and after each experiment, to exclude the degradation of strain, resulting in the loss of sporulation and solvent production ability (Sabathé *et al.*, 2002).

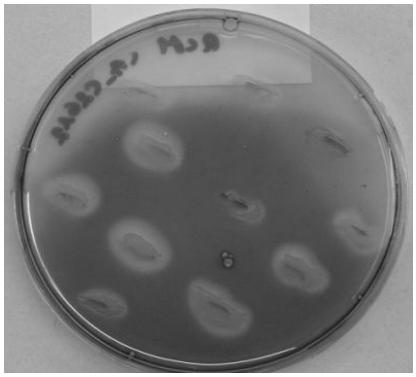


**Figure 2.1.** Reaction between Lugol's solution and starch present in Reinforced Clostridial Agar (RCM-agar). Bright halos around streaked colonies of *C. acetobutylicum araK::int* mutant indicate the degradation of starch and hence the presence of pSOL1; the non-degraded starch present in agar was dyed with iodine and became dark-violet in color.

As an alternative a molecular check with PCR using a pSOL check primer pair (Tab. 9.5) on the amylase P gene (CAP0098) was performed in DNA samples isolated from investigated strains. A reaction would give a product of a size of 1159 bps.

Strain determination based on the sequencing of variable regions

Sequencing results were BLASTed against all clostridial nucleotide sequences in the NCBI or JGI databases to determine the species of *Clostridium*.. For this purpose we used a specially designed primer pair, 161V and 630R (Tab. 9.5). The sequencing was done using a 630R primer.

### 2.3.3. Counting cells using a Thoma Chamber

*E. coli* and clostridia cells were counted using a Thoma Chamber (Celeromics) according to the protocol of the manufacturer, always in 3 independent repeatitions.

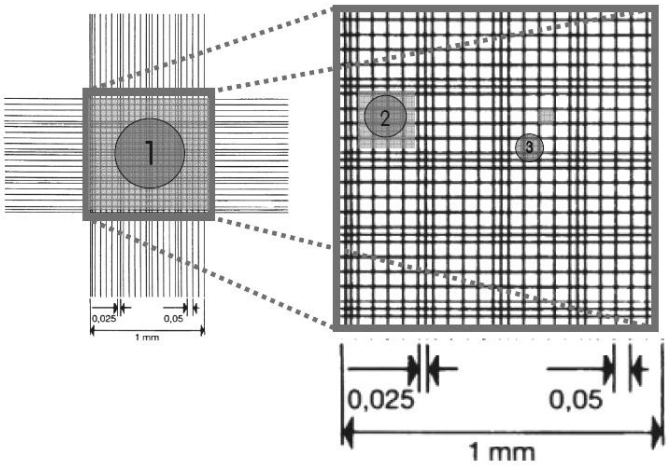



**Figure 2.2.** Thoma Chamber structure and dimensions for easy calculations. More detailed information is given at the Celeromics website (www.celeromics.com).

## 2.4. DNA and RNA manipulation methods

### 2.4.1. DNA isolation and preparation

#### 2.4.1.1. Plasmid isolation from *E. coli*

All plasmids from *E. coli* strains were isolated using an AxyPrep Plasmid Miniprep Kit (Axygen) according to the protocol of the manufacturer. Instead of eluant, double-distilled water was used.

#### 2.4.1.2. Chromosomal DNA isolation from *Clostridium* using a Master Pure DNA Purification Kit

Chromosomal DNA of *Clostridium* strains, for templates for standard checks or cloning PCR reactions, was isolated using a Master Pure™ Complete DNA and RNA Purification Kit (Epicentre) with minor modifications.

Around 600 µl of the overnight culture were centrifuged and the pellet was resuspended in 150 µl of TE buffer. The 30 µl of lysozyme (20 mg/µl) were added and samples were incubated at 37°C for 30 min. Then, 1 µl of proteinase K (5 mg/ml) was added along with 180 µl of 2xT&C Lysis Solution (Master Pure™ kit) and the samples were incubated at 65°C for 15 min. After the incubation, the samples were cooled down on ice and 1 µl of RNase A (10 mg/ml) was added prior to a 30 minute-long incubation at 37°C. Afterwards 210 µl of MPC Protein Precipitation Reagent (Master Pure™ kit) were added and the samples were vortexed for 10 s and centrifuged for 10 min at 10 000 rpm, 4°C. Supernatant was transferred to a new collection tube and 500 µl of pre-cooled isopropanol were added. The samples were kept for 5 min at -20°C and centrifuged again as previously. The Supernatant was gently removed and the pellet washed with 500 µl of pre-cooled 70% (v/v) ethanol. Samples were centrifuged for 2 min, at 10 000 rpm, 4°C, and the supernatant was again gently removed. The pellet was left in an open collection tube at 37°C until dry, and resuspended in ~200 µl of double-distilled water. The DNA concentration and purity were checked on an ND1000 Spectrophotometer (peQLab, Erlangen) at a wavelength of 260 nm.

### 2.4.1.3. Isolation of ultra pure chromosomal DNA from *Clostridium*

Ultra pure and high quality DNA was isolated from overnight cultures prepared in 5ml CGM hungates with no antibiotic selection. This DNA was mainly used for Southern Blots.

Around 2 ml of well-grown overnight culture were centrifuged at 11 000 rpm for 2 min. The pellet was washed once in 1 ml of K-Phosphate buffer (0.1 M, pH 7.5) and resuspended in 1 ml of TE buffer. Next, 100 µl of lysozyme (40 mg/ml) were added and samples were incubated at 37°C for 1 h. 70 µl of 10% (w/v) SDS were added and mixed carefully; subsequently 65 µl of EDTA (0.5 M, pH 8.0) and 5 µl of Tris-HCl (1 M, pH 7.5) were added. Afterwards, 30 µl of proteinase K (5 mg/ml) were added and the samples were incubated at 37°C for 1 h. Next, 180 µl of sodium perchlorate solution (5 M) were added and mixed carefully. The samples were centrifuged for 10 min at 10 000 rpm. The upper phase was transferred into a new collection tube, and 500 µl of phenol along with 500 µl of chloroform-isoamyl alcohol (24:1 [v/v]) were added and vortexed for 10 seconds. The samples were centrifuged as previously and again the upper phase was transferred into a new collection tube. 1ml of isopropanol was added and gently mixed by inverting the tube for 2-3 min and centrifuged for 10 min at 10 000 rpm. The DNA pellet was washed with 500 µl of 70% (v/v) ethanol and centrifuged as previously. The pellet was incubated at 37° C for several minutes to remove residues of ethanol and solvents, and then resuspended in 100 µl of nuclease-free double distilled water. Next, 5 µl of RNase A (10 mg/ml) were added and incubated at 37° C for at least 30 min. DNA quality and RNA degradation were checked by visualization on the 0.8% (w/v) agarose gel.

### 2.4.1.4. Gel electrophoresis in agarose gels

DNA fragments were separated in 1% (w/v) agarose gels in a PowerPac 1000 (BioRad) electrophoresis machine. Standard gels were running for 35 minutes in 120 V and a constant current (A). Being an acid, the DNA migrates to the positive electrode with a speed that correlates to its size. Therefore large-sized DNA is always detectable close to the wells, while small-sized fragments migrate fast and are usually located close to the

rear part of the gel. Afterwards the gels were stained in an ethydium bromide (EtBr) bath for 10-15 minutes and visualized in the AlphaImager MINI (Biozym) camera.

Purification of DNA fragments from the agarose gel

DNA fragments needed for cloning or sequencing were isolated from the gel using the Wizard DNA Clean-Up System (Promega) according to the protocol of the manufacturer. The DNA concentration and purity was checked on an ND1000 Spectrophotometer (peQLab, Erlangen) at a wavelength of 260 nm.

### 2.4.1.5. Measurement of DNA concentration

The concentration of all DNA samples, including DNA treated with restrictases or other DNA-modifying enzymes was measured on the NanoDrop ND1000 Spectrophotometer (peQLab, Erlangen) at a wavelength of 260 nm. The absorbance of the DNA solution was compared to the absorbance of its eluant (blank sample). The NanoDrop also gives information on the purity of DNA solution by measuring the absorption at a wavelength of 280 nm and calculating the 260/280 ratio. The DNA solution is considered purest with a 260/280 ratio of 1.80 (Sambrook *et al.*, 1989).

### 2.4.1.6. Sequencing

Sequencing reactions were prepared only on double-purified DNA samples, being either plasmid DNA or gel-extracted PCR products and performed by the Eurofins MWG Operon company (Eurofins Scientific). DNA samples were prepared in 1.5 ml collection tubes in a volume of at least 15 µl with 3 µl of 5 pmol primer. The concentration of DNA samples differed depending on its size and type, according to the guidelines given by the Eurofins company, shown below:

| | Plasmid DNA | PCR product | | |
|---|---|---|---|---|
| | | < 300 bps | 300 – 1000 bps | >1000 bps |
| **Concentration(ng/µl)** | 50 - 100 | 2 | 5 | 10 |

### 2.4.2. *In silico* DNA manipulations

For all necessary *in silico* DNA manipulations Clone Manager Professional 9 Software (Sci-Ed Software, Cary, NC) was used. The software was designed to help with multiple DNA manipulations, such as cloning, restriction analysis, primer design, PCR reactions and graphic map drawing. It allows sequence files to be edited, genes or specific sequences in those files to be found, genes to be translated and open reading frames to be found. The *in silico* DNA manipulations were performed prior to plasmid creation, PCR reactions, and the analysis of the sequencing results and also for primer design, or to predict the sizes of detectable DNA bands in the Southern Blot experiments.

### 2.4.3. Polymerase Chain Reaction (PCR)

#### 2.4.3.1. Standard PCR

Phusion PCR

Phusion polymerase (New England Biolabs) is a high-fidelity polymerase used for ligations and sequencing. The time of extension was set according to the length of PCR product, with 15-30s per 1kb, and a final extension lasting 10 minutes.

PCR mix:

| | |
|---|---|
| 5 x Phusion buffer | 10 µl |
| dNTPs (10mM) | 1 µl |
| Primer forward | 4 µl |
| Primer reverse | 4 µl |
| DNA template | 50 – 200 ng |
| Phusion polymerase | 0.5 µl |
| Water, nuclease free | to 50 µl |

Phire PCR

The Phire polymerase (Thermo Scientific) has no proof-reading ability and therefore was used only for check PCRs or colony PCRs. It is much quicker than Phusion polymerase and the extension time is 10-15 seconds per 1 kb, with the final extension lasting 10 minutes.

PCRmix:

| | |
|---|---|
| 5 x Phire buffer | 10 µl |
| dNTPs (10mM) | 1 µl |
| Primer forward | 4 µl |
| Primer reverse | 4 µl |
| DNA template | 50 – 200 ng |
| Phire polymerase | 0.5 µl |
| Water, nuclease free | to 50 µl |

### 2.4.3.2. SOE-PCR

SOE-PCR (Splicing by Overlap Extension) was used only to construct a re-targeted region for ClosTron mutagenesis. For that purpose 4 different primers (Tab. 9.3) in different concentrations were used in one reaction:

| Primer name | concentration |
|---|---|
| IBS | 100 µM |
| EBS1d | 100 µM |
| EBS2 | 20 µM |
| EBS Universal | 20 µM |

For a 25 µl reaction 2 µl of each primer should be used along with 1 µl of the template (pMTL007 plasmid) and an appropriate reading-roof polymerase (Phusion). The cycling conditions are given below:

| | | | |
|---|---|---|---|
| Lid | | 110°C | |
| | | 94°C | 2 min |
| 35 cycles | | 94°C | 15 sec |
| | | 55°C | 0.5 min |
| | | 72°C | 0.5 min |
| | | 72°C | 10 min |

### 2.4.3.3. Colony PCR

The colony PCR was performed when a quick check of many clones was needed. A small amount of the bacterial colony was picked using a sterile 10 µl tip or sterile toothpick and resuspended in 5µl of double-distilled water in a PCR collection tube. Next, the water-colony mix was heated up to 100°C for 5 minutes and 20 µl of properly prepared PCR master mix was added. For colony PCRs either Phire or Dream Taq Green Master Mix were used. For Phire polymerase the PCR mix was prepared as described previously, and the preparation of the Dream Taq Green Master Mix PCR is described below.

Dream Taq Green PCR

Dream Taq Green PCR Master Mix (Thermo Scientific) is a twice-concentrated master mix containing the Dream Taq DNA polymerase, the PCR buffer being mixed with a green dye, $MgCl_2$ and dNTPs. The dye allows direct loading of the PCR mixture on the gel and does not interfere with the PCR reaction. The polymerase requires 1 minute per 1 kb for extension and the final extension lasts 10 minutes.

PCR mix:

| | |
|---|---|
| Dream Taq Green PCR Master Mix (2x) | 25 µl |
| Primer forward | 4 µl |
| Primer reverse | 4 µl |
| Template DNA | 50 – 200 ng |
| Water, nuclease free | to 50 µl |

### 2.4.3.4. RT-PCR

The RT-PCR (reverse transcription PCR) was performed either as part of the reverse transcription and the cDNA labeling step of the DNA microarray experiment or as part of the promoter mapping with the RLM-RACE method. More details are given in sections devoted to DNA microarray and RLM-RACE experiments. The RT-PCR was also used as a molecular check of the RNA quality prior to the reverse transcription step for the microarray DNA procedure.

RT-PCR mix (OneStep RT-PCR Kit, Qiagen):

| | |
|---|---|
| 5 x buffer* | 5 µl |
| dNTPs* | 1 µl |
| Enzyme-mix* | 1 µl |
| RNA template | 0.6 µg |
| Primer forward | 3 µl |
| Primer reverse | 3 µl |
| Water, nuclease-free | to 25 µl |

*OneStep RT-PCR Kit (Qiagen).

Cycler program

| | | | |
|---|---|---|---|
| Lid | | 110°C | |
| | | 50°C | 30 min |
| | | 95°C | 15 min |
| | | 96°C | 2 min |
| 30 cycles | { | 96°C | 0.5 min |
| | | 55°C | 0.5 min |
| | | 72°C | 3.5 min |
| | | 72°C | 10 min |

### 2.4.3.5. Probe-generating PCR for Southern Blot

To generate a Southern Blot probe a standard Phusion PCR reaction was used and the product was biotin-labeled with the Biotin Chromogenic Detection Kit (Thermo Scientific). In this work the Southern Blot technique was used only to confirm the single *ermB* integration event and the probe was generated using the RAM primers (Tab. 9.5), was 900 bps in size.

### 2.4.4. Southern Blot

The DNA for the Southern Blot was isolated using phenol-chloroform extraction, and its quality was checked on 1% (w/v) agarose gel. Only DNA of high quality with no signs of degradation was used for the Southern blot experiments. To obtain bands of detectable sizes, different restriction enzymes (frequent cutters) were used for digestion of the DNA prior to blotting procedures, and a list of the enzymes used for different mutants is given in the Results section.

Digestion reactions were set for overnight and the amount of DNA used for blotting was always between 30-40 µg. As a marker the Biotinylated 2-Log DNA Ladder (0.1-10 kb), New England Biolabs, was used, unless otherwise stated. The digested genomic DNA amount used for the blotting experiment varied between 30 and 40 µg, and the quality and efficiency of digestion was checked on 0.8 % (w/v) agarose gel prior to the blotting procedures.

### 2.4.4.1. Preparation of Southern Blot probes

Labeling of the probe

| | |
|---|---|
| DNA template (100 ng – 1 µg) | 10 µl |
| Decanucleotide in 5x reaction buffer | 10 µl |
| dd$H_2O$ water | 24 µl |

The reaction mixture was vortexed and spun down for 5 sec. The tube was then incubated in boiling water for 5 min and cooled down on ice.

Next, 5 µl of biotin labeling mix and 1 µl of Klenow fragment were added to the reaction mixture, shaken well and spun down. Samples were incubated from 1h to overnight at 37°C, and the reaction was stopped by the addition of 1 µl of 0.5 M EDTA. A labeled probe was stored at -20°C, and was reused multiple times. Prior to each use an incubation of the probe in boiling water for 10 min was required.

### 2.4.4.2. Gel and blotting procedure

<u>Gel and vacuum</u>

Digested DNA samples were separated on 0.8% (w/v) agarose gel at low speed (80V) for several hours. A nylon membrane (Nylon Membranes, positively charged, Roche)

was cut to a size approximately 1 cm larger on each side than a plastic mask and placed on a blotting plate under the plastic mask (model 785 Vacuum Blotter, BioRad). Fully developed gel was placed on the top of the mask, and the vacuum pressure (CVC 2000, Vacuumbrand) was adjusted to 5 Hg.

Blotting procedure

Depurination of the gel was done with depurination solution for 10 min, until the bromophenol – blue (used as a component of a loading dye) turned yellow. The gel was then washed twice with double-distilled water for 1 min. The Gel was denatured for two 15-min-long periods with denaturation solution, until the yellow color of the dye component turned blue again and washed twice with double-distilled water for 1 min. The Neutralization step was done with a neutralization buffer for 15 min, also twice, and then the gel was washed briefly with water before proceeding to the transfer step.

The DNA was transferred from the gel onto the nylon membrane in the 20 x SSC buffer for 2 h and fixed on it by exposing the membrane to UV light for 1 min.

### 2.4.4.3. Hybridization and detection procedure

Prehybridization=hybridization solution

| | |
|---|---|
| 5x SSC | 12.5ml |
| 1% (w/v) blocking stock solution | 5.0ml |
| 0.1% (w/v) N-laurylsarcosine | 0.5ml |
| 0.02% (w/v) SDS | 0.2ml |

The SDS solution was pre-warmed to RT to avoid precipitation prior to the hybridization step. The hybridization solution can be kept for a longer time at -20°C and should be dissolved at 60°C.

Pre-hybridization was done with the hybridization=prehybridization solution (without a probe) in the oven at 42°C for 2-4 h. A biotin – labeled probe was denatured at 100°C for 5 min, cooled on ice and mixed with the hybridization solution (the final concentration of the probe = ~25-100 ng/ml). The probe was hybridized overnight at 42°C. After the incubation step the hybridization buffer with the probe was poured off into a fresh collection tube and was stored at -20°C. The membrane was washed twice with buffer 1 at RT for 10 min and also twice with buffer 2 for 20 min at 65°C. The washing buffers were poured off and the membrane was dried before proceeding to the detection of biotin-labeled probe step.

Detection of biotin-labeled probe

The membrane was washed in 30 ml of blocking/washing buffer for 5 min and blocked with a blocking solution for 30 min. Next, it was incubated in 20 ml of Strep-AP solution for 30 min and washed twice for 15 min in 60 ml of blocking/washing buffer. The membrane was incubated for 10 min in 20 ml of detection buffer prior to the detection reaction and carried in 10 ml of freshly prepared substrate solution in complete darkness for 15 min to overnight. The reaction was stopped by washing the membrane in distilled water and the membrane was dried in the air.

Reagents used for Southern Blot:

Depurination solution autoclaved

0.25 M HCl

Denaturation solution autoclaved

1.5 M NaCl

0.5 M NaOH

Neutralization buffer autoclaved

0.5 M Tris-HCl

3 M NaCl

pH 7.5 set with HCl

Transfer solution (20 x SSC) autoclaved

3 M NaCl

0.3 M $Na_3$citrate

pH 7.0 set with HCl

0.5 M EDTA

pH 8.0

Buffer 1 autoclaved

2x SSC

0.1% (w/v) SDS

Buffer 2 autoclaved

0.1x SSC

0.1% (w/v) SDS

10 % (w/v) N-Laurylsarcosine solution autoclaved

200 ml Blocking/Washing solution

| | |
|---|---|
| 10x buffer | 20ml |
| MQ water | 180 ml |

50 ml Blocking solution

| | |
|---|---|
| Blocking/Washing Solution | 50 ml |
| Blocking powder | 0.5 g |

Stir at 60°C to dissolve.

20 ml Streptavidin

| | |
|---|---|
| Blocking solution | 20 ml |
| Streptavidin-Conjugate | 4 µl |

Prepare freshly before use.

30 ml detection buffer

| | |
|---|---|
| 10 x buffer | 3 ml |
| $ddH_2O$ | 27 ml |

10 ml substrate solution

| | |
|---|---|
| 50x BCIP/NBT | 0.2 ml |
| Detection buffer | 9.8 ml |

Prepare freshly before use.

### 2.4.5. Cloning methods

#### 2.4.5.1. DNA modification with restriction enzymes and dephosphorylases

For all restriction analysis during this study Fast Digest enzymes (Thermo Scientific) were used. For all modifications of DNA the enzymes from either Thermo Scientific (conventional and/or Fast Digest) or from New England Biolabs were used. In all cases the digestion or double-digestion reactions were performed according to the instructions of the manufacturer.

For dephosphorylation of plasmid ends prior to the ligation step the FastAP enzyme (Thermo Scientific) was added to the digestion mix or cleaned plasmid solution and incubated for 45 min at 37°C.

Before proceeding to the ligation or sequencing steps the DNA was cleaned using Wizard® SV Gel and PCR Clean-Up System (Promega).

#### 2.2.5.2. Ligation

Ligations were performed in two variants, small and large, depending on the amount of DNA used. Ligation mixtures were left for 2 hours at 22°C or overnight at 16°C, and deactivated at 75°C for 5 minutes prior to transformation.

| | **Vector DNA** | **Insert DNA** | **Total DNA** |
|---|---|---|---|
| Small variant | 40 ng | 120 ng | 160 ng |
| Large variant | 80 ng | 300 ng | 380 ng |

## 2.5. Genetic manipulations of clostridia

### 2.5.1. DNA transfer

#### 2.5.1.1. Chemocompetent *E. coli* cells and transformation method

Chemocompetent *E. coli* cells

For the preparation of chemocompetent cells 10 ml of LB medium were inoculated with 100 µl of overnight culture and left shaking at 37°C until the $OD_{600}$ reached a value of between 0.5-0.6. Next, the cells were centrifuged at 6 000 rpm at 0°C for 5 min and washed in 5 ml of 100 mM $CaCl_2$ solution (pre-cooled to 0°C). The cell pellet was resuspended in 1 ml of 100 mM $CaCl_2$ solution and 200 µl of sterile 87% (v/v) glycerin were added. The solution of chemocompetent cells was divided into 100 µl aliquots and frozen before use. The prepared aliquots were stored at -80°C.

Chemocompetent *E. coli* cells buffer

| | |
|---|---|
| $CaCl_2$ x 2 $H_2O$ | 1.47 g |
| $ddH_2O$ | to 100 ml |

Sterile-filtered, kept at 4°C.

Transformation of chemocompetent *E. coli*

The frozen *E. coli* cells, which had been divided into aliquots, were placed on ice for 10 min. Next, 50-100 ng of purified plasmid or the total volume of ligation mixture was added and the cells were incubated on ice for 25 min. Heat shock was performed at 42°C for 60 sec, and then the mixture was cooled down on ice for 2 min. Next, 400 µl

of LB medium was added and gently mixed. The cells were incubated at 37°C for 1 h and plated on selective plates.

### 2.5.1.2. Electrocompetent *Clostridium acetobutylicum* cells and standard transformation method

Electrocompetent *Clostridium acetobutylicum* cells

Prior to the procedure a new culture from a spore stock solution was started by heat shock treatment as described in section 2.3.2.1.

A heavy loop of cells was inoculated into 5 ml of CGM cultures and a series of dilutions was left overnight. The next day 50 ml of CGM were inoculated with total growth from a pre-culture still showing active growth. The cells were collected when the $OD_{600}$ reached 0.6-0.8 and centrifuged down at 4°C for 10 min, 5 000 rpm. The cell suspension was kept on ice between the centrifugation steps and all buffers were pre-cooled to 0°C before use. The cell pellet was washed with 20 ml of ETM buffer and centrifuged as previously. Next, the cell pellet was washed in 10 ml of ET buffer and resuspended in 3 ml of ET buffer. The cell suspension was divided into aliquots of 600 µl, and each aliquot was mixed with 8-12 ng of purified DNA, transferred to cold 4 mm electroporation cuvettes and left on ice for 1 min; prior to this step, the cuvettes had been kept at -20°C for 2 h. The electroporation machine (a GenePulser II with Pulse Controller Plus and Capacitance Extender Plus, BioRad) was set at 50 µF, 600 Ω, 1.8 kV, and the electroporation time was considered acceptable when between 10-30 ms. Afterwards the electroporation cells were transferred immediately to 3 ml of CGM and incubated on a shaker for 4 h. Then the cells were collected by centrifugation for 5 min, 5 000 rpm at room temperature, and the total growth was plated on selective plates and incubated for 2 – 4 days.

The plasmid DNA was cleaned of any salt residues on MF - Membrane Filters (Milipore) prior to electroporation for at least 45min.

Electroporation buffers:

ETM Buffer

| | |
|---|---|
| Saccharose | 27.7 g |
| $Na_2HPO_4$ x $H_2O$ | 0.032 g |
| $NaH_2PO_4$ x $H_2O$ | 0.18 g |
| $MgCl_2$ x 6 $H_2O$ | 12.8 g |
| $ddH_2O$ | to 300 ml |

ET Buffer

| | |
|---|---|
| Saccharose | 27.7 g |
| $Na_2HPO_4$ x $H_2O$ | 0.032 g |
| $NaH_2PO_4$ x $H_2O$ | 0.18 g |
| $ddH_2O$ | to 300 ml |

Buffers were sterile-filtered and kept at 4°C in anaerobic conditions, preferably in anaerobic flasks for no longer than 3 months.

### 2.5.1.3. Modified electroporation protocol for *Clostridium acetobutylicum*

A fresh batch of *C. acetobutylicum* cells was prepared as described above, from the spore stock solution.

A heavy loop of cells was inoculated into 10 ml of CGM cultures and a series of dilutions was left overnight. Next, 50 ml of CGM medium was inoculated with the precultures still showing active growth to a starting optical density of 0.02 and cultured until the $OD_{600}$ reached 0.20-0.35. The cells were harvested quickly by centrifuging at

8 000 rpm for 15 min at room temperature. Next, the cells were washed in 20 ml of ETM buffer and centrifuged (8 000 rpm, 15 min, RT). The pellet was resuspended in 3 ml of ET buffer. Aliquots of 0.6 ml were mixed with salt-free DNA in standard cuvettes (4 mm gap) and left on ice for 10-15 min. The electroporation machine settings were the same as in the standard protocol.

The cell mixture was transferred directly after electroporation to 4 ml of CGM broth, pre-warmed to 37°C and left for at least 3 hours at 37°C, and shaken at 180 rpm.

#### 2.5.1.4. DNA transfer to *Clostridium saccharobutylicum* and *Clostridium beijerinckii* via the triparental conjugation

*Clostridium saccharobutylicum*

Cells of the recipient and the donors were harvested at the beginning of the exponential growth phase, when the $OD_{600}$ values were between 0.6 and 0.8. Next, 1 ml of *E. coli* CA434 and the *E. coli* – pJL1 or *E. coli* – pJL2 strain, transformed with the plasmid which was supposed to be conjugated into *C. saccharobutylicum* cells, were washed once in PBS and transferred immediately to an anaerobic cabinet. Pellets were dissolved in 200 µl of recipient culture, plated as 25 µl drops on 2 xYTG plates without any antibiotic selection and left overnight. The whole growth was harvested from the plate and resuspended in 1ml of anaerobic PBS and plated on 2 xYTG plates supplemented with thiamphenicol (Tm) and a counter-selection against *E. coli* donors – colistin. The first colonies can be seen after approximately 12 hours of incubation at 37°C.

*Clostridium beijerinckii* was conjugated using the same protocol, with the alternative use of D-cycloserine or polymyxin-B instead of colistin, and selective plates were supplemented with spectinomycin (Sp) instead of thiamphenicol, due to *C. beijerinckii's* natural resistance to Tm. *Clostridium beijerinckii* does not require methylation of the exogenous DNA, therefore the CA434 strain was used as a donor along with *E. coli* Top10 bearing the plasmid of interest.

### 2.5.2. ClosTron mutant creation

The ClosTron is a mutagenesis system based on the group II intron of *Lactobacillus lactis* (Ll.ltrB group II intron) containing an *ermB* cassette activated by retrotransposition (Heap *et al.*, 2009; Heap *et al.*, 2007). Modular ClosTron plasmids, pMTL007, bear additionally an antibiotic-resistance gene, either for chloramphenicol/thiamphenicol (pMTL007C-E2), or spectinomycin (pMTL007S-E2) for *E. coli* and *Clostridium* selections of the mutants (Fig. 2.3 A). The group II introns are active in a broad host range, and they insert into a specific, recognized site by base-pairing between the target DNA site and the intron RNA sequence (Fig. 2.3 B). This process has a high accuracy, and therefore provides a trustworthy method for mutant creation. Using a simple algorithm (Perutka *et al.*, 2004) one can design primers that can create a specific re-targeted region for precise ClosTron mutation. The algorithm is provided free-of-charge at www.ClosTron.com. Using 3 specific primers (the IBS, EBS1d, EBS2 and the universal EBS primer) a re-targeted region can be created with the SOE-PCR method. The correct re-targeted introns can be confirmed with pMTL2 check fwd and rev primers (Tab. 9.3) and cloned into the ClosTron vector, the pMTL007C-E2 or pMTL007S-E2 plasmid, using endonuclease *Hind*III and *BsrG*I.

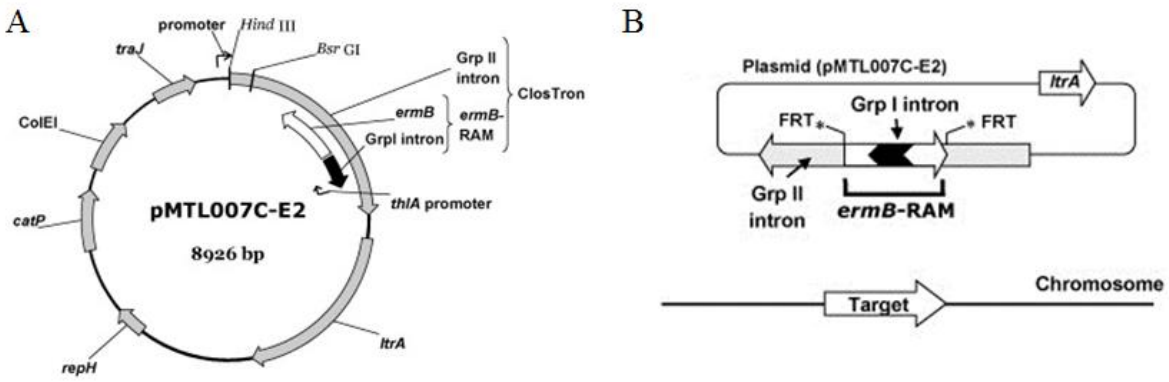



**Figure 2.3. (A)** ClosTron vector pMTL007C-E2 with chloramphenicol/thiamphenicol resistance gene. All modules of the plasmids: the intron module, the Gram-positive (*repH*) and Gram-negative (*colEI*) replicons, the antibiotic resistance gene and the additional *traJ* module are indicated in the figure. **(B)** The group II intron with an internal ErmB-RAM cassette containing the group I intron, which jumps out of the cassette after intron insertion into the targeted DNA region (from Cartman *et al.*, 2010).

For the molecular mutant check two PCR reactions were performed on DNA samples isolated from overnight single-colony cultures. The RAM fwd and rev primers (Tab. 9.5) amplify part of the intron cassette and give a product of 900 bps if the integration event took place or of 1.2 kb for the intron located on the plasmid. A larger band is produced in a PCR reaction also when the plasmid is still present in clostridial cells; therefore prior to this PCR check a complete loss of plasmid is required. The reaction on gene-specific primers produces a 2.1 kb larger product (the size of intron and *ermB* cassette) compared to the wild type if the integration event occurred.

*Clostridium acetobutylicum* ATCC 825

*Clostridium acetobutylicum* cells were transformed and plated on thiamphenicol selective plates. Single colonies were restreaked once onto a second plate containing Tm and freshly grown colonies were restreaked onto selective plates with erythromycin or clarithromycin. Erythromycin/clarithromycin-resistant colonies were again restreaked onto fresh selective plates and single colonies were picked for mutant screening.

Pure mutant colonies were restreaked several times on thiamphenicol-free plates before proceeding to the mutant analysis and checked for plasmid loss (PCR or plating on thiamphenicol test).

*Clostridium beijerinckii* NCIMB 8025

The procedure for ClosTron mutant creation in *C. beijerinckii* is generally the same as for *C. acetobutylicum*, with two minor changes. *C. beijerinckii* is naturally resistant to thiamphenicol, therefore another antibiotic – spectinomycin was used to select cells bearing plasmid – pMTL007S-E2 (Heap *et al.*, 2010). The plasmid was transferred into *C. beijerinckii* cells according to the principles of triparental conjugation (2.5.1.4). All other procedures were done as described previously for *C. acetobutylicum*.

*Clostridium saccharobutylicum* NCP 262

This study is the very first work to describe mutant creation and gene transfer into the *C. saccharobutylicum* NCP 262 strain. The ClosTron mutant creation protocol had to be adjusted for this particular strain. The re-targeted regions for *C. saccharobutylicum* were synthesized by DNA 2.0 (Menlo Park, CA). After conjugation of the *C. saccharobutylicum* wild type strain with an *E. coli* donor bearing plasmid of interest (2.5.1.4 section), thiamphenicol-resistant colonies were selected as described in the protocol for *C. acetobutylicum*. ClosTron mutants were selected on lincomycin, an erythromycin analogue from the lincosamides group, as described in the *C. acetobutylicum* section, and the *ermB* gene provided resistance also against this particular antibiotic, and checked for the correct integration.

### 2.5.3. *ErmB* marker removal

The ClosTron technique is simple and effective; however it leaves an antibiotic resistance marker in the genome of the bacterium and precludes the creation of another gene disruption in the mutant strain. This is an important limiting factor; therefore a specific flippase-system was created to remove the *ermB* cassette. The *flp* recognition sites (FRT) are located on either side of the ErmB – RAM cassette, which would be removed from the host DNA when the functional flippase gene is present (Fig. 2.4). This leads to the sensitivity of the mutant to erythromycin, and the procedure can be repeated many times, giving a possibility to create multiple ClosTron mutants (Kuehne & Minton, 2012; Heap *et al.*, 2010; Cartman *et al.*, 2010; Heap *et al.*, 2007).

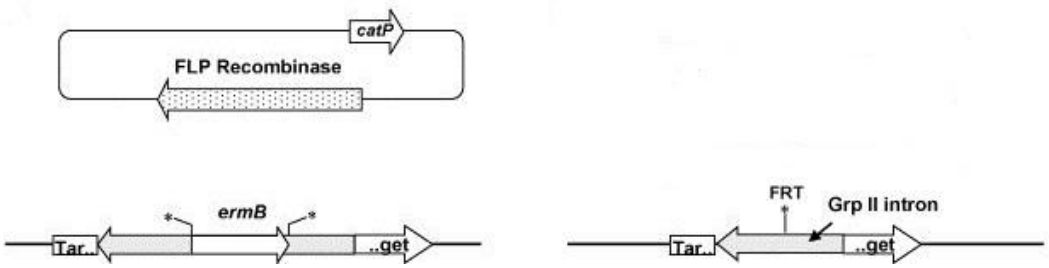



**Figure 2.4.** Representation of the *ermB* marker removal with Flp recombinase-mediated step. On the left: the modular plasmid bearing functional FLP recombinase gene is introduced into ClosTron mutant host cells, resulting in the excision of the *ermB* cassette, located between two *flp* recognition sites, here indicated with stars (*). On the right: genomic DNA of the ClosTron mutant with the ermB cassette removed; FRT, *flp* recognition site (from Cartman *et al.*, 2010).

To remove the *ermB* marker a special plasmid containing the *flp*-recombinase gene (pCLF1-1) was transferred into the *C. saccharobutylicum* mutant strain and cells were plated onto selective plates with thiamphenicol and incubated anaerobically at 30°C for 2 – 3 days, until the colonies were large enough to inoculate the liquid medium cultures. Next, 10 hungates, each with 4ml of CGM broth containing thiamphenicol, were inoculated with single colonies from the plate and incubated overnight at 30°C. Several dilutions of the broth – $10^{-2}$, $10^{-3}$, $10^{-4}$ and $10^{-6}$ were plated on thiamphenicol-containing plates and left for 3 days at 30°C. Exactly 50 colonies were replica-plated on Tm and Erm plates. Colonies appearing only on Tm plates, but not on Erm plates were PCR-checked for intron *ermB* cassette loss, using the RAM primers and primers binding to either side of the targeted regions.

This procedure was applied only to *C. saccharobutylicum* mutant strains.

### 2.5.4. Gene knock-out through allelic exchange

The allelic exchange is a method of markerless gene knock-out, and for *Clostridium acetobutylicum* an orotate phosphoribosyl transferase (*pyrE)*-deficient strain has been created (Ng *et al.*, 2013; Heap *et al.*, 2012). Markerless gene knock-out allows the deletion of genes or their fragments without affecting up- or downstream genes or regulatory elements, which is important for the investigation of single genes located at the beginning or in the middle of the operons. The *pyrE* gene encodes an

enzyme which catalyzes the transformation of fluoroorotic acid (FOA) into a highly toxic compound, so only *pyrE* negative cells are FOA resistant, but at the same time they require uracil supplementation. The pMTL-ME3 plasmid (Tab. 2.2) containing an altered version or only the START and STOP codons of a gene to be deleted (the gene of interest, GOI) with two 0.5 – 1.0 kb-long homologous sequences on either side of the GOI, is equipped with a functional copy of *pyrE* gene, providing an easy selection for first recombinants.

To obtain a clean deletion plasmid, a clean deletion cassette needed to be created. To this end two homology arms, left (LHA) and right (RHA) were designed and amplified using LHA fwd with LHA rev and RHA fwd with RHA rev primer pairs (Fig. 2.5 A). Next, equal concentrations of cleaned PCR products were used as a DNA template to create the clean deletion cassette with LHA fwd and RHA rev primers. Since LHA rev and RHA fwd overlap, amplified products of LHA and RHA will have homologous regions and eventually would give one product during the final PCR reaction (Fig. 2.5 B). The clean deletion cassette was then digested with appropriate restrictases and cloned into the pMTL-ME3 modular plasmid backbone (Tab. 2.2). Correct clones were checked by sequencing prior to the electroporation of *Clostridium*.

After the introduction of the pMTL-ME3 with an appropriate deletion cassette into a *C. acetobutylicum ΔpyrE* strain the transformants were selected on thiamphenicol-supplemented plates. Although in *pyrE*-deficient strains the synthesis of uracil is interrupted, because of the non-functional orotate phosphoribosyl transferase and supplementation of the media with uracil is required (Heap *et al.*, 2012), it is not necessary as long as a functional copy of the *pyrE* gene on the pMTL-ME3 plasmid is present in the clostridial cells. The fastest growing colonies were picked up and PCR-checked for the first recombination event, using appropriate primer pairs (Tab. 9.4). Confirmed first recombinants were plated on a non-selective medium and incubated for 4 days. Next, the total growth was harvested, dissolved in anaerobic PBS and plated on a minimal medium containing 5-FOA in a series of dilutions. Colonies were replica-plated onto mediums with and without thiamphenicol, and Tm-sensitive clones (the loss of resistance to Tm indicates the loss of plasmid) were identified and PCR-checked for the second recombination event. The loss of vector and the resistance to FOA is a sign of the second recombination event, leading to either a new mutant strain or a primary wild type strain.

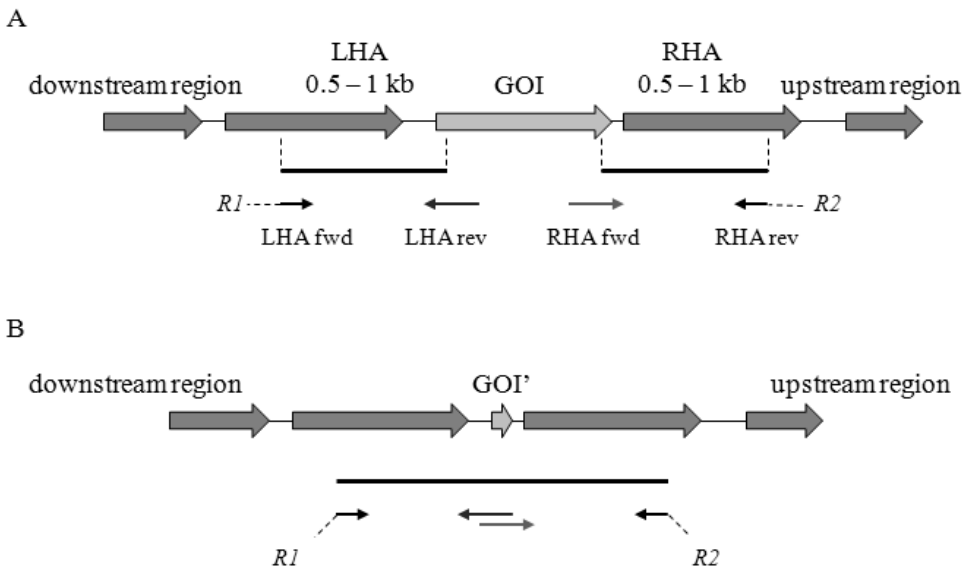



**Figure 2.5. (A)** The schematic representation of the clean deletion cassette creation and the wild type strain genome; GOI, gene of interest; LHA, left homology arm; RHA, right homology arm; *R1*, restrictase 1; *R2*, restrictase 2. The arrows indicate primers used to amplify the left and the right homology arms, which overlap with the START and STOP codons of the GOI. **(B)** The representation of the created clean deletion cassette and the genome region with the deleted gene of interest. Arrows indicate primers used to create LHA and RHA. Here, it is shown clearly how LHA rev and RHA fwd overlap for the purpose of the clean deletion cassette creation.

## 2.6. RNA isolation and manipulation

### 2.6.1. RNA isolation from *Clostridium* strains

The material for RNA isolation was collected from continuous cultures, and all solutions were autoclaved twice for the total removal of RNases.

RNA preparation with micro dismembrator (Sartorius Mikro-Dismembrator U Laboratory Ball Mill)

A frozen cell pellet was resuspended in 1 ml of TE buffer and centrifuged for 1 min at 13 000 rpm, 4°C. The pellet was resuspended in 300 µl of TE buffer and stored on ice.

The shaking flask along with a grinding ball was placed in liquid nitrogen and left until cold. The complete cell suspension was pipetted into a cooled flask filled with nitrogen and ground at 1600 rpm for 3 min. The powder was then resuspended in an RLT buffer and transferred into a sterile falcon.

RNA purification with RNeasy midi kit (Qiagen)

Bacterial lysate in an RTL buffer was centrifuged for 5 min at 9 000 rpm, at room temperature and the supernatant was transferred into a new sterile falcon. Next, 2.8 ml of frozen (-20°C) 100% ethanol were added, mixed well and left on ice for 5 min. Supernatant mixture was applied onto an RNeasy midi column placed in a 15 ml collection tube and centrifuged for 10 min at 9 000 rpm, RT. The maximal loading capacity of the column is 4 ml, therefore the procedure was repeated if needed. The flow-through was discarded; the membrane was washed with 4 ml of RW1 buffer and centrifuged for 5 min at the same speed. Next, the membrane was washed twice with 2.5 ml of RPE buffer and centrifuged for 3 min, at 9 000 rpm, RT. An additional centrifugation step for 5 min was done to make sure the membrane was completely dry before proceeding to the elution step. Next, 150 µl of nuclease-free water, pre-warmed to 50°C, were applied onto the dried membrane and incubated at room temperature for 5 min. Samples were centrifuged for 5 min at 9 000 rpm, RT. The elution step was repeated, and the total eluant (~300µl) was transferred into a sterile 1.5 ml collection tube.

DNA degradation

DNase mix

| | |
|---|---|
| Eluate with RNA | 300 µl |
| 5 x DNase buffer | 85 µl |
| DNase (recombinant,RNase free, Roche) | 15 µl |

The mixture was incubated at 37°C for 3 hours. After this step it is possible to stop the protocol and store samples at -80°C or proceed to the next step.

Purification and precipitation of RNA

To purify the RNA from digested DNA particles 1 volume of acidic phenol was added to the sample, vortexed for 10 seconds and centrifuged for 3 min, 13 000 rpm at 4°C. The upper phase was removed into a new collection tube. 1 volume of chloroform/isoamyl alcohol (24/1) was added, vortexed and centrifuged as above. The upper phase was again transferred into a new collection tube, where a 1:10 volume of 3.3 M sodium acetate (pH 5.0) and 2.5 volume of absolute ethanol (-20°C cold) were added and gently mixed. The RNA was precipitated overnight at -70°C.

RNA cleaning

The precipitated RNA was centrifuged in a centrifuge, pre-cooled to 4°C, for 30 min at 13 000 rpm and gently washed with 1 ml of 70% (v/v) ethanol (at –20°C). The samples were centrifuged for 10 min at 13 000 rpm, at 4°C and the ethanol was carefully discarded. The pellet was dried for 1h at 37°C in a tightly closed sterile beaker. The dry pellet was resuspended in 30 µl of nuclease-free water and incubated on ice until it was fully dissolved (1 – 2 h).

The cleaned RNA should be stored at -70°C

Buffers:

Tris-HCl (1 M, pH 8)

| | |
|---|---|
| Tris | 121.1 g |
| $ddH_2O$ | to 1000 ml |

EDTA (0.5 M, pH 8)

| | |
|---|---|
| $Na_2EDTA$ x 2 $H_2O$ | 46.5 g |
| $ddH_2O$ | to 250 ml |

TE buffer (pH 8)

| | |
|---|---|
| Tris-HCl (1 M, pH 8) | 10 ml |
| EDTA (0.5 M, pH 8) | 2 ml |
| $ddH_2O$ | to 1000 ml |

5 x DNase Buffer

| | |
|---|---|
| $MgSO_4$ | 25 mM |
| NaAc | to 500 mM |

NaAc (3.3 M, pH 5)

| | |
|---|---|
| NaAc | 408.1 g |
| $ddH_2O$ | to 1000 ml |

The pH should be adjusted with acetic acid.

RNA quantification and visualization on agarose gel

The quality of RNA was checked by visualization on gel (1% (w/v) agarose gel) and by using the RT-PCR technique. To this end, 1:10 dilutions of RNA samples were prepared and 1-2 µl were loaded on the agarose gel. For the RT-PCR 1 µl of the diluted RNA sample was used, and the reaction was performed as described previously (2.4.3.4 section).The concentration of the RNA in the samples was measured on the NanoDrop ND1000 Spectrophotometer (peQLab, Erlangen).

### 2.6.2. RLM-RACE

The 5' RLM-RACE (First Choice RLM-RACE Kit, Invitrogen™) method was used for promoter mapping. Since one of the last steps, the random primer RT-PCR, often resulted in unspecific bands during the first nested PCR (the outer PCR) it was changed for a direct RT-PCR (OneStep RT-PCR kit, Qiagen) on the outer 5'RLM-RACE PCR primers for the trouble-generating probes. Apart from that modification all procedures were done according to the manufacturer's specifications. All the primers used for RLM-RACE are listed in the table 9.7.

## 2.7. DNA Microarray

### 2.7.1. Reverse transcription and labeling of cDNA

RNA labeling

Primer hybridization mix:

| | |
|---|---|
| RNA template | 25 µg |
| Hexamer primer (5 µg/µl) | 4 µl |
| Water, nuclease-free | 10 µl |

Labeling reactions were prepared twice for each dye, 2 x Dy3 and 2 x Dy5 (GE Healthcare Life Sciences).

Samples were mixed carefully and annealed with primers at 70°C for 10 min.

Reverse transcription

For the reverse transcription procedure the labeled samples were placed on ice to stop the reaction and the rest of the reaction mix was added as follows:

| | |
|---|---|
| 5 x first strand buffer* | 4 µl |
| 0.1 M DDT * | 2 µl |
| dCTP nucleotide mix (4 mmol dCTP, 10 mmol dATP, dTTP and dGTP) | 2 µl |
| dCTP Cy-Dye-labeled nucleotide** (Cy3 or Cy5) | 1 µl |
| SuperScript reverse transcriptase* (200 U/µl) | 1 µl |

* SuperScript III Reverse Transcriptase, Invitrogen

**Cy3-dCTP and Cy5-dCTP, GE Healthcare Life Sciences

Samples were mixed carefully and incubated at 42°C for 3 hours. For purification the samples were placed on ice. The labeled samples were protected from light at all times.

### 2.7.2. Purification of labeled cDNA

Prior to the purification of the labeled cDNA step, all RNA residues were removed by adding 2 µl of 2.5 M NaOH to each labeled sample, and mixed carefully. Samples were incubated at 37°C for 15 min. Next, the reaction was neutralized by adding 10 µl of 2 M acid-free HEPES and vortexing. Before proceeding to the QIAquick PCR Purification Kit (Qiagen), the samples were spun down briefly. The cleaned cDNA was eluted with 35 µl of nuclease-free water.

For the microarray experiment the cleaned Cy3- and Cy5-dye labeled samples were mixed together and incubated for 5 min at 100°C. After 4 min the TF (Tom Freeman) buffer was added up to 150 µl and vortexed briefly. Prior to loading to the hybridization station the samples were placed on ice.

Tom-Freeman Hybridization buffer (Fitzpatrick, Johnston *et al.* 2005)

| | |
|---|---|
| Formamid (deionized) | 20 ml |
| 50x Denhart's solution | 5.0 ml |
| 20x SSC | 12.5 ml |
| 100 mM Na-Pyrophosphate | 0.5 ml |
| 1 M Tris (pH 7.4) | 2.5 ml |
| 10 % (w/v) SDS | 0.5 ml |
| HLPC-$H_2O$ | to 9.0 ml |

### 2.7.3. Quantification of marked cDNA

The quantification of labeled cDNA could be done either by a NanoDrop check or by a wavelength scan.

For Cy3 quantification a 500 – 580 nm scan was used with the maximal extinction being 150 000 $mol^{-1}cm^{-1}$ (at 550 nm).

For Cy5 quantification a 600 – 700 nm scan was used with the maximal extinction being 250 000 $mol^{-1}cm^{-1}$ (at 650 nm).

The absorbance was determined at 550 nm and 650 nm and the content of incorporated fluorescent dye was calculated using the following formula:

Cy3 or Cy5 in a sample (in pmols) = $(A \times Z \times F \times 10^{12}) / E$

A = Absorption value of Cy3 (550 nm) or Cy5 (650 nm)

Z =volume of labeled cDNA, (µl)

F = dilution factor

E = extinction coefficient (Cy3 = 150000 x 1 $mol^{-1}cm^{-1}$) or

(Cy5 = 250000 x 1 $mol^{-1}cm^{-1}$)

### 2.7.4. Hybridization of labeled cDNA

For the hybridization of the labeled cDNA (target) to nucleic acids (probes) immobilized on the surface of a microarray chip, a Tom Freeman hybridization buffer was used.

The solution was sterile-filtered prior to use.

For the hybridization of the targets 80 pmol of each labeled cDNA were prepared as described in section 2.7.2. Samples were injected into the hybridization chamber and the hybridization was done automatically according to the *C. acetobutylicum* ATCC 825-optimized protocol using an Automated Slide Processor.

### 2.7.5. Scanning and Analysis of DNA microarray using GenePix Pro 6.0

The scanning of the fluorescent signal of Cy3 and Cy5 dyes was done using a GenePix 4000B scanner (Axon Instruments, Union City) and hybridization data were generated using GenePix Pro 6.0 software. Genes located on the microarray that are not regulated under the growth conditions used were used as standards to determine the strength of the pre-scan (PMT 635 nm – Cy5, PMT 532 nm – Cy3). Therefore the ratio of their fluorescence intensity should be 1:1. Another measurement determining the strength of the pre-scan was of the microarray background. The settings for the main scan were as follows:

Pixel size: 10 µm

Lines to average: 1

Focus position: 0 µm

Scan area: ca.1000 x 7000 Pixel

Before the analysis of the scanned microarray chips a special mask with annotated *Clostridium acetobutylicum* ATCC 825 genes was put on the scanned picture. The mask

covers the individual spots for the most significant genes of the organism and was created using specially developed software (Ehrenreich, unpublished). Each spot on the chip has an annotation shown in a text format. The fluorescent values of the background, the standard deviation of the background and the two dyes, the ratio of medians, the ratio of means and the ratio of regression were calculated automatically by GenePix Pro 6.0 software and exported to a text file. The results could also be viewed in the form of a scatter plot. During the analysis of the data the ratio of medians should stay equal to 1 and the normalization factor should be around 1.

Data were exported to a text file and saved as an Excel file (Microsoft Office).

### 2.7.6. Evaluation of the transcriptional data

The normalized data in the Excel file (Microsoft Office, Microsoft) were analyzed using specific transcription criteria. Only the spots complying with several filter criteria were considered relevant for further analysis in order to reduce technical artifacts. The values for the fluorescence of both dyes and the values for the fluorescence of both dyes minus the standard deviation of the background were filtered to be equal to or above zero value. The ratio of medians, the ratio of means and the regression of the ratio were filtered to be equal to or lower than 30%. These operations remove all methodological artifacts and leave trustworthy data for analysis. Genes were considered significantly regulated if their values, shown as logarithms to the basis of 2, were equal to or higher than 1.6 for upregulation, or equal to or lower than -1.6 for downregulation.

## 2.8. Analysis of bacterial culture products and assay procedures

### 2.8.1. Gas chromatography

The quality and quantity of the fermentation products were analyzed using the Shimadzu GC – 2010 on a Stabilwax DA column. Nitrogen was used as a carrier gas. Samples from growth curves or continuous cultures were centrifuged for 15 min at

maximum speed at 15°C prior to the GC samples preparation. To ensure adequate and trustworthy measurements two different standards were used during every GC run – and the external standard, consisting of the most common acids and solvents in three different concentration variants, and the internal standard, the 0.5% (w/v) 1-propanol, added directly to the measured samples. Prior to the GC analysis of prepared samples, the concentrations of the internal and external standards were analyzed. If both standards were in correct concentrations, the GC run results were considered significant with no need to make technical replicates.

GC water

500 ml of double-distilled sterile water with a pH value adjusted to 2.0 using HCl.

GC sample preparation

| | |
|---|---|
| GC water | 350 µl |
| 1 – propanol (0.5% [w/w]) | 50 µl |
| Supernatant | 100 µl |
| Total volume | 500 µl |

Standard preparation

0.5% Standard mix

| | |
|---|---|
| Acid mix (1% [w/w in GC water]) | 200 µl |
| Solvent mix (1% [w/w in GC water]) | 200 µl |
| Total volume | 400 µl |

The 0.05% and 0.005% standard mixes were prepared using 200 µl of 0.5% standard mix and 800 µl of GC water.

GC standard samples (0.05%, 0.005% and 0.0005%)

| | |
|---|---|
| GC water | 800 µl |
| 1 – propanol (0.5% [w/w]) | 100 µl |
| 0.5% - 0.005% standard mix | 100 µl |
| Total volume | 1000 µl |

1 – Propanol internal standard

0.5% (w/w) solution

| | |
|---|---|
| GC water | 9.0 g |
| 1 – propanol (5% [w/w in GC water]) | 1.0 g |

Data were analyzed using GC Solution Software (Schimadzu).

### 2.8.2. Assay procedures

The assay procedures were used to measure the concentrations of monosaccharides in minimal or rich media in order to determine the effectiveness of the consumption of a given carbon source.

Sugar concentrations were estimated in samples taken at chosen time points on the growth curves of batch cultures or continuous cultures. Samples were centrifuged at 4°C for 15 minutes at maximum speed and either used immediately or stored at -20°C. Dilutions were made in double-distilled water if necessary.

Three technical replicates were made for the measurements of glucose, xylose and arabinose concentration in all samples taken from hungate or batch culture experiments. The absorbance of the solutions was measured three times for each sample; the average value was calculated and used as the right absorbance value. The measurements were considered significant if the difference between the average and the measured OD values was below 0.005, and only then were they used for a graphic representation of an experiment.

### 2.8.2.1. D-Glucose concentration

Determination of D-glucose concentration (D-glucose Assay Kit, GOPOD Format, Megazyme).

The assay procedure (GOPOD format) allows the specific estimation of the amount of free D-glucose in any solution with a simple glucose oxidase/peroxidase reaction:

$$\text{D-glucose} + O_2 + H_2O \xrightarrow{\text{(glucose oxidase)}} \text{D-gluconate} + H_2O_2$$

$$2\ H_2O_2 + \text{p-hydroxybenzoic acid} + \text{4-aminoantipyrine} \xrightarrow{\text{(peroxidase)}} \text{quinoneimine dye} + 4\ H_2O$$

The resulting color stays stable at room temperature for at least two hours according to the manufacturer's guidelines.

All the procedures were done as stipulated by the manufacturer.

### 2.8.2.2. D-Xylose concentration

Determination of D-xylose (D-Xylose Assay Kit, Megazyme, Ireland).

The assay uses a simple way to measure the amount of D-xylose by the measurement of NADH formed in a reaction with the sugar and β-xylose dehydrogenase (β-XDH) at pH 7.2. The amount of NADH is stoichiometric with the amount of β-D-xylose present in

the reaction solution. The conversion of α-D-xylose is catalyzed by xylose mutarotase (XMR):

$$\alpha\text{-D-xylose} \xrightarrow{XMR} \beta\text{-D-xylose}$$

$$\beta\text{-D-xylose} + NAD^+ \xrightarrow{\beta\text{-XDH}} \text{D-xylonic acid} + NADH + H_2$$

All the procedures were done as stipulated by the manufacturer.

### 2.8.2.3. L-Arabinose concentration

Determination of L-arabinose (L-Arabinose/D-galactose Assay Kit, Megazyme, Ireland).

The principle of this assay is very similar to the previous one. The β-galactose dehydrogenase (β-GalDH) can recognize only the β-anomeric form of L-arabinose, which causes changes in the absorbance of samples containing the α- and β-form. The galactose mutarotase (GalM) turns the alpha form of L-arabinose into the beta form prior to the dehydrogenase reaction:

$$\alpha\text{-L-arabinose} \xleftrightarrow{(GalM)} \beta\text{-L-arabinose}$$

$$\beta\text{-L-Arabinose} + NAD^+ \xleftrightarrow{(\beta\text{-GalDH})} \text{L-arabionic acid} + NADH + H^+$$

The amount of NADH formed in this reaction is stoichiometric with the amount of L-arabinose present in the sample.

All the procedures were done as stipulated by the manufacturer.

### 2.8.2.4. Analysis of protein concentration

The concentration of proteins in the crude extract was determined using a method developed by Bradford (Bradford, 1976). This colorimetric assay is based on the absorbance differences of the dye Coomasie Brilliant Blue G-250, which under acidic conditions changes from the red, ‘free’ form to blue when bound with assayed proteins. The maximum absorbance spectrum of the blue form is set at 595 nm. Any increase in

the absorbance at this wavelength is proportional to the amount of protein bound to the Coomasie Brilliant Blue dye. As the protein concentration standard, an assay with fixed amounts of BSA (Bovine Serum Albumin, Sigma-Aldrich) was done. The Bradford assay was performed as described previously (Bradford, 1976).

Three technical replicates were made for the protein concentration measurements. The absorbance of the solutions was measured three times for each sample; the average value was calculated and used as the right absorbance value. The measurements were considered significant if the difference between the average and the measured OD values was below 0.005, and only then were they used for a graphic representation of an experiment.

## 2.9. Standard techniques in protein handling

### 2.9.1. Cell disruption

Disruption of the cell membrane was done using ultrasound (Dr. Hielscher, UP 200 s). For this experiment 45 ml of *Clostridium* culture were grown until the $OD_{600}$ reached values between 1.5-2.0. Cells were collected by centrifuging at maximum speed for 10 min, dissolved in 10 ml of wash buffer and centrifuged again. The pellet was resuspended in 2 ml of wash buffer and incubated for 30 min with 30 µl of lysozyme (20 mg/ml) at 37°C. The cell suspension was cooled on ice prior to the sonication step.

During the cell membrane disruption the cell suspension was kept on ice at all times to prevent the probe boiling and ultrasound was applied until the suspension became clear. The disrupted cells were centrifuged down for 2 min at maximum speed in a centrifuge pre-cooled to 4°C, and the supernatant was transferred into a new sterile collection tube with glycerol (final concentration 50%) and 1:100 BSA solution (10 mg/ml), separated into 25 µl aliquots and stored at -20°C for further experiments.

Wash buffer

| | |
|---|---|
| Monopotassium phosphate | 21 mg |
| EDTA | 44 mg |
| NaCl | 44 mg |
| AEBSF Stock Solution (1 mM) | 3 ml |
| $ddH_2O$ | to 15 ml |

Store at 4°C.

TNE buffer

| | |
|---|---|
| Tris | 30 mg |
| EDTA | 15 mg |
| NaCl | 15 mg |
| $ddH_2O$ | to 5 ml |

### 2.9.2. In vitro DNA methylation with native methyltransferases

Samples were placed on ice until thawed. Next, the plasmid DNA was supposed to be methylated and the following substrates were added to each previously prepared 25 µl aliquot:

| | |
|---|---|
| TNE buffer | 50 µl |
| S-adenosylmethionine stock solution (0.8 mM, ThermoScientific) | 10 µl |
| BSA stock solution (10 mg/ml) | 1 µl |
| Plasmid DNA (50-100 ng/µl) | 10 µl |

The reaction mixture was incubated at 37°C for 16 hours, and the DNA was cleaned with a phenol/chloroform extraction as described in section 2.4.1.3.

# 3. Results

## 3.1. The *upp*-based clean deletion system

The *upp* gene (CAC2879) of *Clostridium acetobutylicum* ATCC 824 codes for uracil phosphoribosyltransferase, an enzyme that transforms uracil and phosphoribosylpyrophosphate into uridine monophosphate (UMP). The *upp* gene is located between the ABC-type iron (III) transport system, ATPase component (CAC2878) and the ribose 5-phosphate isomerase, *rpiB* (CAC2880; Fig. 3.1.1).

As long as a cell possesses a functional copy of the *upp* gene, it remains sensitive to a toxic base analog, 5-fluorouracil (5-FU), and cannot grow in its presence. Disrupting the gene results in the creation of a strain resistant to 5-FU that would require the constant supplementation of uracil. Both these features of an *upp*-deficient strain were used to design a markerless gene deletion/insertion system (Soucaille, Figge & Croux, 2008). A very similar system, based on the disruption of the *pyrE* gene (CAC0027), which encodes the orotate phosphoribosyltransferase, was created soon after (Heap *et al.*, 2012). The Δ*pyrE*- and the Δ*upp*-based methods are founded on the same principle and both strains require uracil supplementation. The only difference between them is the selection agent; 5-fluoroorotic acid (5-FOA) for Δ*pyrE* strains and 5-fluorouracil (5-FU) for Δ*upp* strains. Since the *upp* system created is a patented strain and the *pyrE* system was not fully developed at the beginning of this work, it was important to obtain a new, available *C. acetobutylicum upp*-deficient strain for the purposes of this study.

The 5-fluorouracil (5-FU) is a toxic uracil analog, which, when converted by uracil phosphoribosyltransferase (*upp*) to 5-fluoro-dUMP, inhibits the function of thymidylate synthase and, as a result, DNA replication. This leads to the death of a cell. Disrupting or removing the *upp* gene is the key to using the 5-FU as a selection marker for allelic exchange genome modifications. The inhibitory concentrations of the chemical were established previously by Kraußе (PhD thesis, 2012). To construct the *upp*-defective strain of *Clostridium acetobutylicum* ATCC 824 several clean deletion plasmids were created. In this work an attempt was made to obtain a ClosTron *upp* mutant, by disrupting the uracil phosphoribosyltransferase gene with an *ermB*-containing intron.

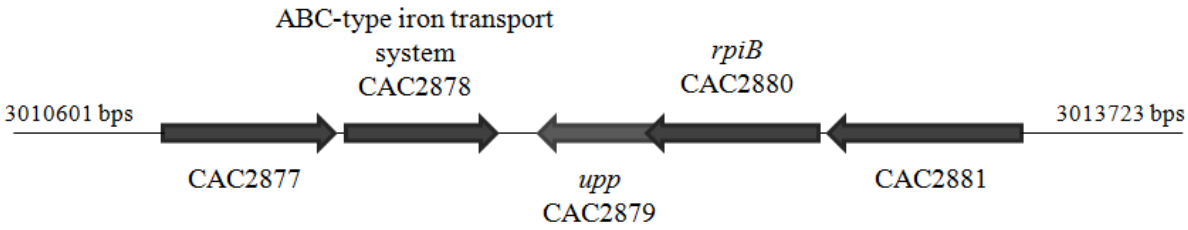



**Figure 3.1.1.** Graphic representation of the *C. acetobutylicum* ATCC 824 genome region containing the *upp* gene. The STOP codon of the *rpiB* gene overlaps the START codon of the *upp* gene. The genes are described in detail in the text.

### 3.1.1. General principles of *C. acetobutylicum Δupp* strain creation

Several plasmids containing two homolog regions upstream and downstream of the *upp* gene, called the left homology arm (LHA) and the right homology arm (RHA), respectively, were created. Each plasmid had different regions of different lengths and some of them were overlapping the *upp* gene. After the transformation of *C. acetobutylicum* wild type with the chosen plasmid the fastest growing colonies were isolated and inspected by PCR check for the first integration event, presented schematically in figure 3.1.2. Colonies with the antibiotic resistant gene integrated along with the plasmid in their genome were expected to grow faster. The first integrants would then be plated on minimal medium plates with the addition of 5-FU and uracil for the double integration event selection. The second integration and loss of the plasmid would be confirmed by the PCR method. All primers used in this study are listed in the table 9.1.

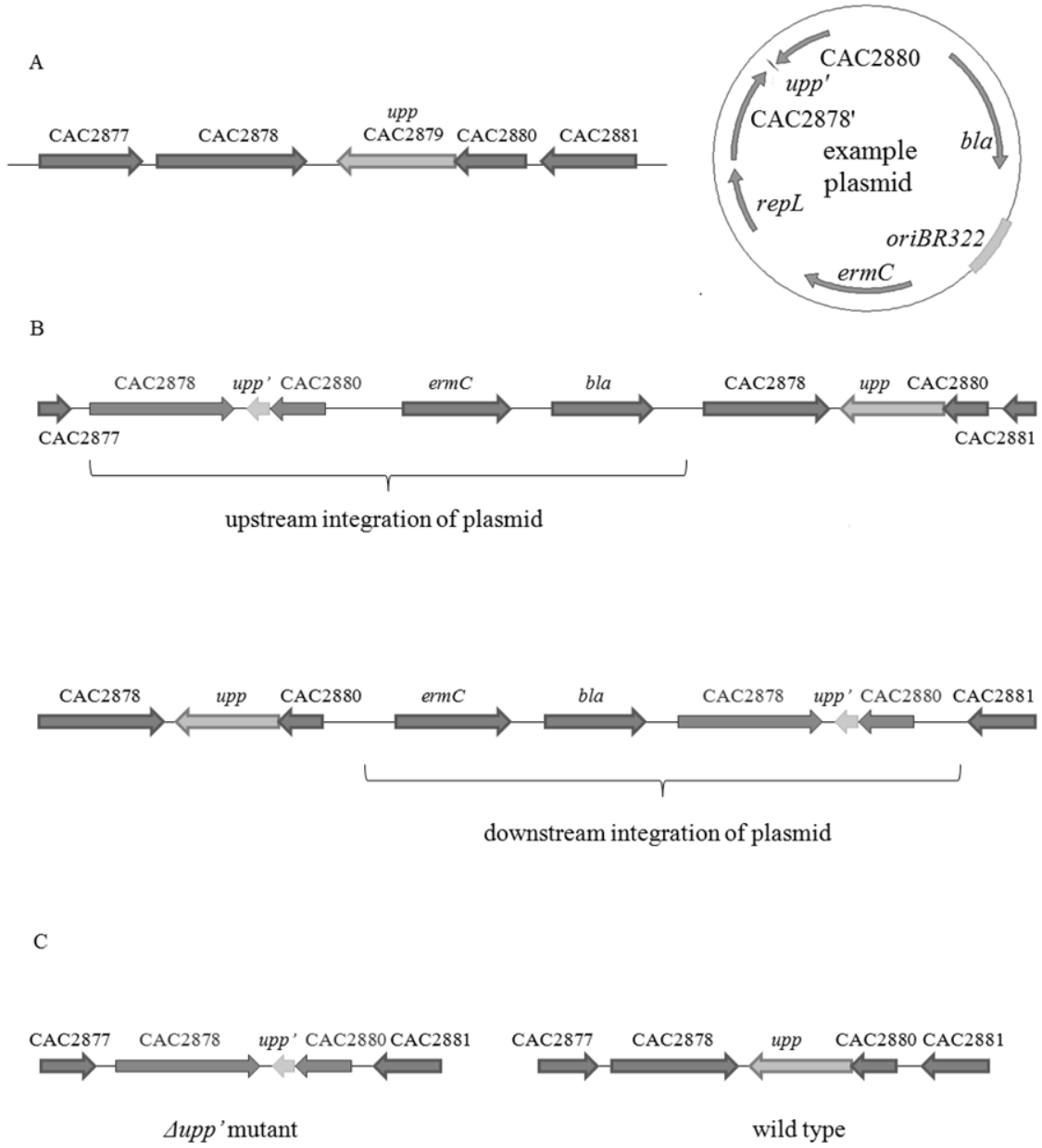



**Figure 3.1.2. (A)** Schematic representation of an integration event of an exemplary plasmid into the genome region up and downstream of the *upp* gene. **(B)** Two possible ways of integration exist – in the region of the left homology arm (part of CAC2878 gene) or right homology arm (part of CAC2880 gene). **(C)** During the second recombination event the plasmid is removed from the genome, leaving either the disrupted (on the left) or the wild type (on the right) copy of the *upp* gene.

### 3.1.2. Use of the pMADK*upp*$^-$ plasmid to create *C. acetobutylicum Δupp* strain

The pMADK*upp*$^-$ plasmid creation has been described previously (Kraußе,PhD thesis, 2012), and during this study the pMADK*upp*$^-$ transformants were screened to find the first integrants using the *upp* down *BamH*I fwd and check *upp* fwd primer pair and the *upp* up *Hind*III rev and check *upp* rev primer pair (Kraußе, PhD thesis, 2012). The deleted fragment should be 630 bps in size. The first pair, *BamH*I fwd and check *upp* fwd would give a product of 2438 bps for the wild type strain or 1808 bps if the first integration occurs, and the second pair, *upp* up *Hind*III rev and check *upp* rev would give a product of 2320 bps for the wild type and 1690 bps for the first integrant. The primers are listed in the table 9.1. Over 100 clones resistant to Crm were examined, using both primer pairs, for upstream and downstream integration, however no first integrants were found and therefore alternative deletion plasmids were created.

### 3.1.3. Creation of pJL9-520

The pMADK*upp*$^-$plasmid was designed to delete the whole *upp* gene, including START and STOP codons. However, the downstream gene, ribose-5-phosphate isomerase B (*rpiB*, CAC2880) and *upp* gene overlap in their START/STOP codon sequences (Fig. 3.1.1). To avoid possible problems caused by this particular feature of the homologous regions, a new plasmid was created to delete only that part of the *upp* gene which would affect the *upp* function. The pJL9-520 was designed to have homology arms (HA) of around 520 bps. The *upp* START and STOP codons were part of HA, and therefore would also stay preserved in the mutant strain sequence. The pJL9 plasmid was created using the backbone of the pCH1 plasmid (Tab. 1.2).

The left (LHA) and right (RHA) homology arms were constructed using the 520 LHA fwd with 520 LHA rev primers and 520 RHA rev with 520 RHA fwd pair, respectively. Cleaned PCR products of both reactions were used as a template in one fusion reaction mix to amplify the 520-deletion cassette of 1009 bps size, with 520 LHA fwd and 520 RHA rev. Next, the *BamH*I and *Hind*III restrictases were used for the cloning of the amplified fusion cassette into a pCH1-backbone, giving a pJL9-520 of a size of 5192 bps (Fig. 3.1.3). All primers used are listed in the table 9.1.

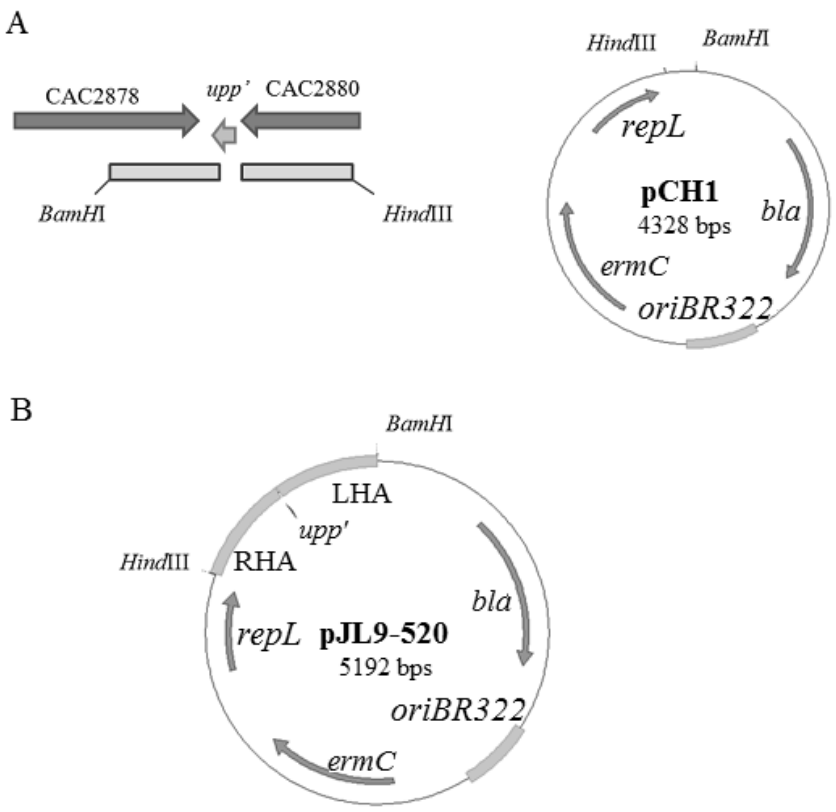



**Figure 3.1.3. (A)** Schematic representation of the right and left homology arms for clean deletion of the *upp* gene (on the left), the pCH1 plasmid used as a backbone (on the right), **(B)** and the map of the created pJL9-520 plasmid. The most important features are described in the figure. *RepL*, origin of replication for clostridia; *oriBR322*, origin of replication for *E. coli*.

The pJL9-520 plasmid was introduced into *C. acetobutylicum* WT and the fastest growing colonies (possible 1st integrants) were re-streaked three times on clarithromycin selective plates. Next, liquid overnight cultures of the 5 largest colonies from the last re-streak were prepared on CGM-Crm agar plates. Isolated DNA samples were PCR-checked for the first integration event using the *upp* check up fwd and rev primers for upstream integration and the *upp* check down fwd and rev primers for downstream integration (Tab. 9.1). The wild type fragment for upstream PCR primers is 1790 bps and for downstream PCR primers – 1636 bps. The mutant-derived band is 623 bps smaller. Within 5 clones one downstream integrant was found (Fig. 3.1.4). As this clone grew extremely poorly on the plate an attempt was made to restore its growth from the liquid culture. However, no growth in a fresh liquid culture was observed, and therefore the total volume of the original liquid culture used for DNA extraction was

spun down and plated onto a CGM selective plate. Only 8 clones appeared on the plate and they were all screened for up- and downstream integration, yet no positive results were obtained. The reason behind these problems remains unclear; however it was also observed in the following experiments.

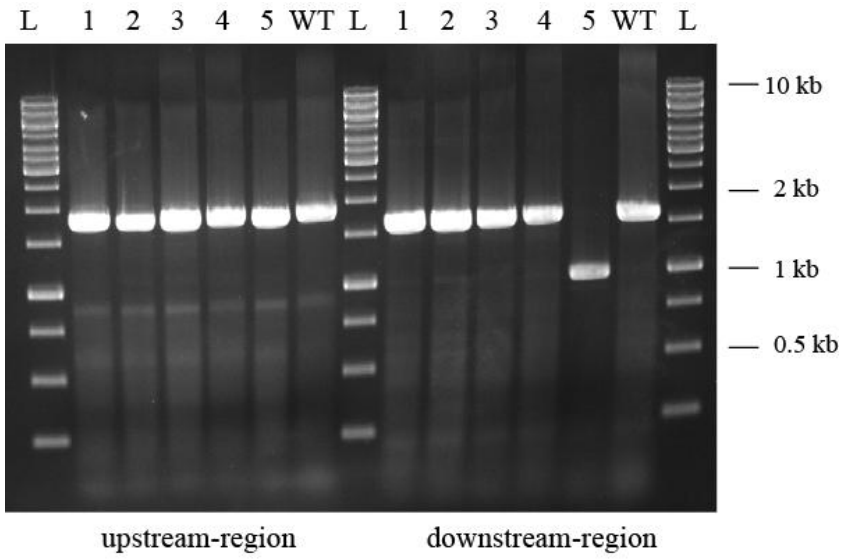



**Figure 3.1.4.** Check for the up- and downstream integration of pJL9-520 plasmid into *C. acetobutylicum* genome. The DNA from five fastest-growing colonies was isolated for a molecular check of upstream (left) and downstream (right) integration event. The PCR reaction on the DNA isolated from the colony no. 5 produced a band suggesting there was an integration event in the region downstream of the *upp* gene. L, 1 kb ladder; 1 – 5, five colonies tested; WT, wild type DNA.

Therefore another approach needed to be developed and new plasmids were designed in the way that has not been used in previous experiment or in the work of Kraußе (2012).

### 3.1.4. Creation of the pJL10 and pJL11

In previous experiments the integration of the *upp*-deletion plasmids into the *C. acetobutylicum* genome was tested, however, without any positive results. Therefore

the homology regions were modified and located partially or completely ‘inside’ the *upp* gene (Fig. 3.1.5 A and B).

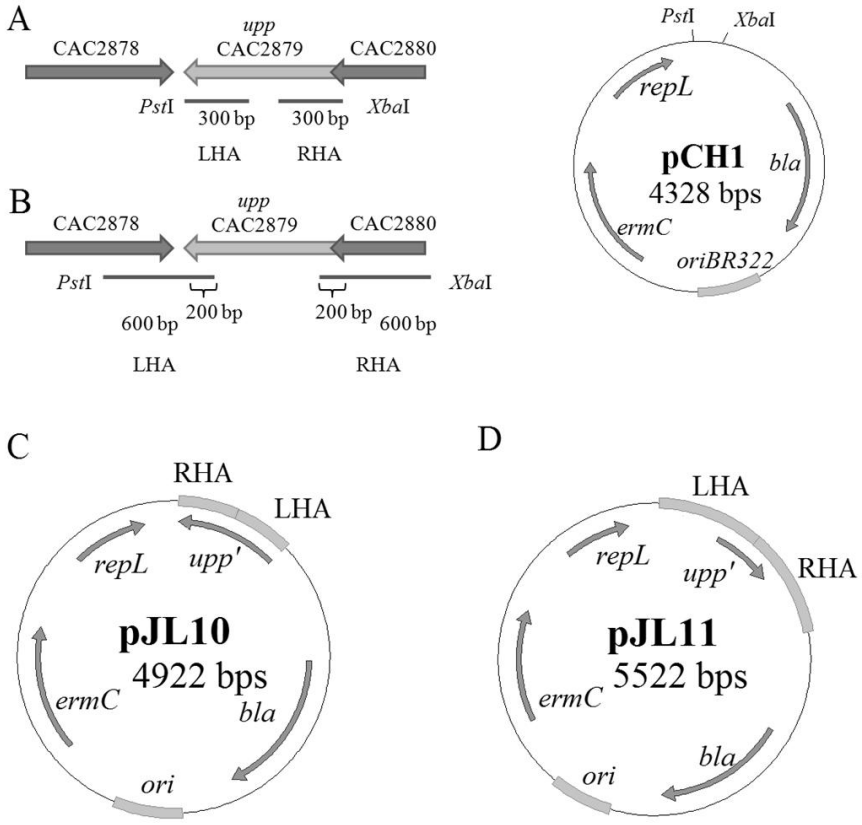



**Figure 3.1.5.** Schematic representation of the creation of two deletion cassettes, **(A)** composed of 300 base-pair-long homology arms hidden within the *upp* gene and **(B)** overlapping the *upp* gene only with 200 bps. Restriction enzymes used for cloning of the cassettes into a pCH1-derived backbone (on the right) are shown in the graphic. The Cloning procedures resulted in two new clean deletion plasmids: **(C)** pJL10 and **(D)** pJL11. *RepL*, origin of replication for clostridia; *oriBR322/ori,* origin of replication for *E. coli*.

The left (LHA) and right (RHA) homology arms were constructed using the LHA *upp* in fwd with LHA *upp* in rev primers and RHA *upp* in fwd with RHA *upp* in rev pair, respectively. Cleaned PCR products of both reactions were combined in one fusion reaction mix and used to amplify the *upp* inside-deletion cassette of 624 bps in size with LHA *upp* in fwd and RHA *upp* in rev primers. For cloning purposes the *Xba*I and *Pst*I restriction enzymes were used to digest the insert and the pCH1 plasmid, creating the pJL10 plasmid of a size of 4922 bps. The pJL11 was created likewise, using the LHA

*upp* ins-out fwd with LHA *upp* ins-out rev and RHA *upp* ins-out fwd with RHA *upp* ins-out rev primer pairs to create the LHA and RHA, respectively. The cassette was amplified using the LHA *upp* ins-out fwd and RHA *upp* ins-out rev primers, and it was cloned into the pCH1 backbone using the *Xba*I and *Pst*I restrictases. The inside-out clean deletion cassette was of a size of 1224 bps, and pJL11 was of a size of 5522 bps. All primers are listed in table 9.1.

When pJL10 and pJL11 were used, 30 bps and 230 bps regions from the centre of the *upp* gene would be deleted, respectively. In most of the experiments, the first integrants were detected with the PCR performed using one primer that would bind to the homologous region on one side of the *upp* gene and the primer binding to the genomic region on the opposite side of the *upp* gene. However, with the standard technique for PCR product visualization on the agarose gel it would be impossible to observe the difference between the wild type and mutant size in a case of pJL10-derived deletion, and therefore a different way to find the 1$^{st}$ integrant had to be designed. For this purpose the plasmid-genome primer pairs were used, and the product would be created only if the integration event occurred. This method, however, has one main disadvantage, which is the lack of any possibility to run a positive control in the PCR check.

Since pJL10 and pJL11 were made on the pCH1 backbone, pCH1 check fwd and rev primers were used (Held, PhD thesis, 2012) and a new pair of primers for the genome check, *upp* genome check fwd and *upp* genome check rev, was designed (Tab. 9.1).

For the pJL10 and pJL11 transformants multiple fast growing colonies were examined with the PCR method, yet no integration event was observed (Fig. 3.1.6). In some cases the PCR created unspecific bands, although they were much larger or smaller than expected.

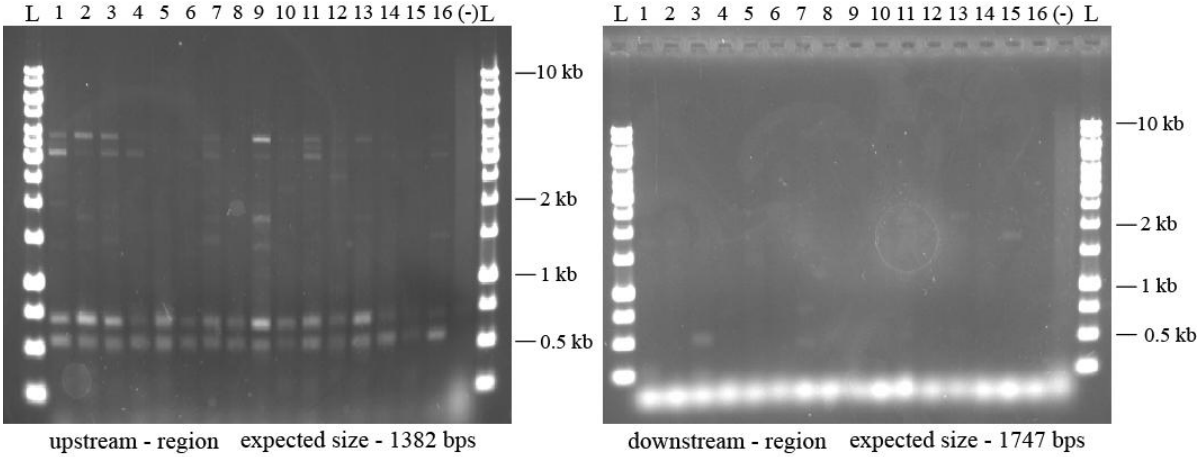



**Figure 3.1.6.** Upstream and downstream integration of the pJL10 plasmid into the *C. acetobutylicum* genome. Sixteen colonies of *C. acetobutylicum* strain transformed with pJL10 were checked for upstream (right) and downstream (left) recombination. L, 1kb DNA ladder; (-), negative control, template-free; 1 – 16, sixteen colonies chosen for the molecular check.

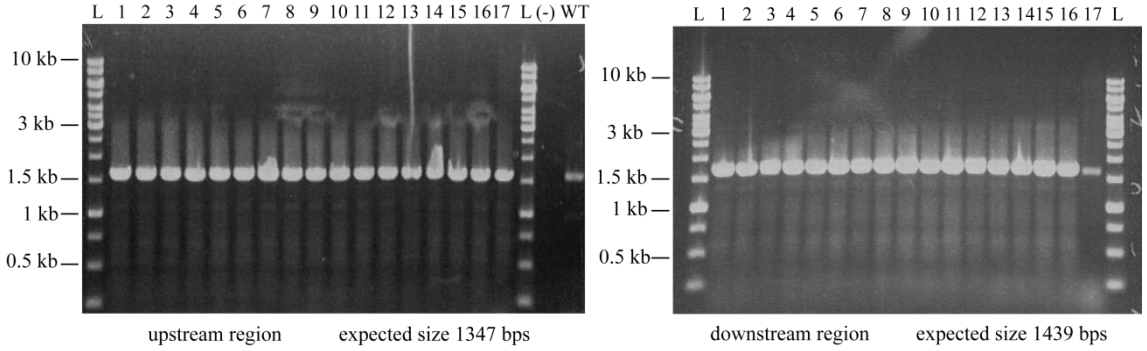



**Figure 3.1.7.** Upstream and downstream integration of the pJL11 plasmid into the *C. acetobutylicum* genome. The DNA of 17 colonies was PCR-checked for upstream (right) and downstream (left) recombination. The predicted wild type size for the upstream primer set is 1577 bps (1347 bps mutant) and for the downstream primer set is 1669 bps (1439 bps mutant). L, 1kb DNA ladder; (-), negative control, template-free; WT – wild type DNA used as a template; 1 – 17; seventeen colonies chosen for the molecular check.

To check the integration event for pJL11, one genome-binding primer and one primer binding to the region of the homology arm were used. For upstream integration the *upp* genome fwd and RHA upp insout rev were used, and for the downstream event the *upp* genome rev and LHA *upp* insout fwd were used (Tab. 9.1). In both experiments no first integrants were detected.

### 3.1.5. ClosTron *upp* mutant creation attempt

To create the *upp* mutant, multiple approaches of allelic exchange were established, yet the ClosTron method was never tested. To test whether this method would prove more efficient, the pMTL007C-E2::*upp*-84|85a plasmid was created as described (Materials and Methods sections 2.4.3.2 and 2.5.2; the primers are listed in the Tab. 9.3) and introduced into the *C. acetobutylicum* WT strain. Multiple Tm-resistant colonies were obtained and re-streaked onto Crm-containing CGM plates supplemented with uracil. This procedure was repeated several times and eventually 20 colonies were obtained on CGM+Crm+uracil plates. All were checked with the RAM primers (Tab. 9.5), which is a typical procedure to assure intron integration and to eliminate false positive clones. Out of twenty screened DNA samples fifteen RAM-integrated clones were detected, but none of them gave a positive signal in the PCR check on the *upp* genome fwd and Clostron *upp* check rev primers (Fig. 3.1.8). This would suggest the integration event took place somewhere else in the genome, which later was found out to be very rare, but possible, also for introns targeting other genes. Primer sequences are listed in the Tab. 9.1.

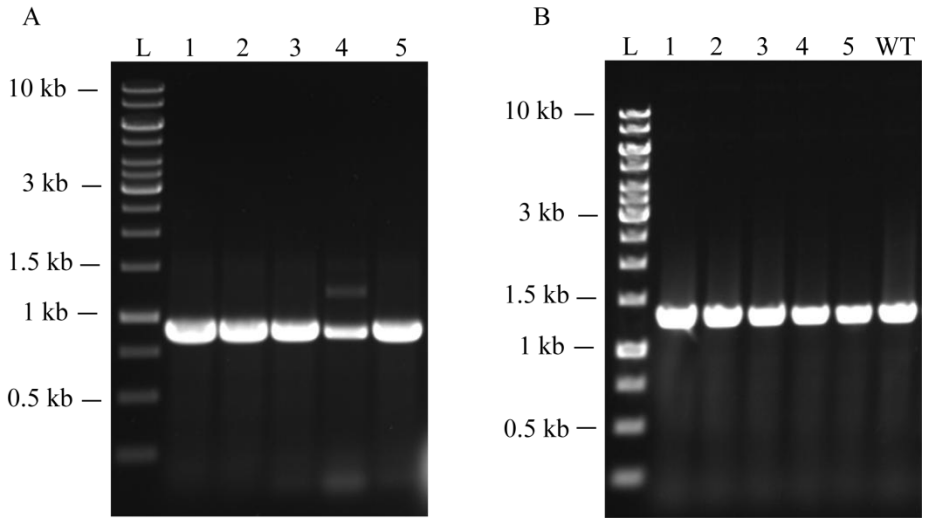



**Figure 3.1.8**. The ClosTron integration of the *upp* re-targeted intron into the *C. acetobutylicum* genome. Here, the integration event was checked on (A) RAM primers and (B) genome primers on the DNA of five chosen colonies. For the genome check the wild type (WT) size should be 1307 bps and the mutant size around 3400 bps. L, 1kb DNA ladder; WT, wild type DNA sample; 1 – 5, five colonies chosen for the molecular check.

### 3.1.6. Application of different methods to improve the allelic exchange

#### 3.1.6.1. Different cultivation conditions

During the attempts to create the *upp* mutant different conditions were applied to improve the integration of plasmids into the genome. Multiple approaches designed for this purpose included testing the impact of more re-streaks and a longer incubation time for the fastest-growing colonies in the liquid medium (for pJL9-520), cultivation in minimal (MES) and rich (CGM) mediums (for pJL9-520, pJL10, pJL11), multiple transfers to minimal medium hungates with different pH values (for pJL10, pJL11) or supplementation of 5-FU (for pJL10, pJL11). Examples are presented in figure 3.1.9. The DNA isolated from cultures subjected to these methods was checked for signs of plasmid integration to establish whether any of these conditions might favor allelic exchange in *C. acetobutylicum* cells.

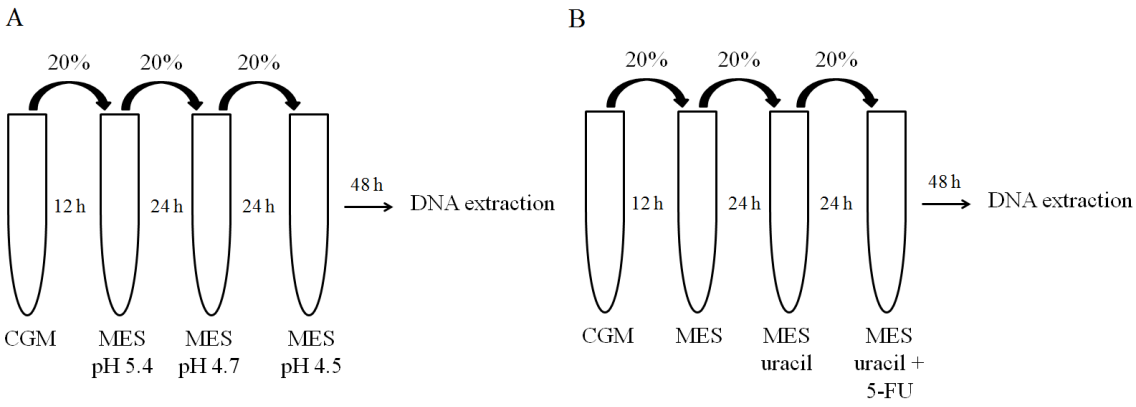



**Figure 3.1.9.** Schematic representation of two chosen cultivation methods applied to improve the recombination event in *C. acetobutylicum*. Media with different **(A)** pH values were used to simulate low pH stress and **(B)** 5-FU supplementation was applied to select for the *upp* disruptants. 5-fluorouracil is light-sensitive and very unstable in high temperatures, therefore a growth observed in hungates with 5-FU supplementation without molecular signs of plasmid integration was probably caused by an insufficient concentration of the selective chemical in the medium.

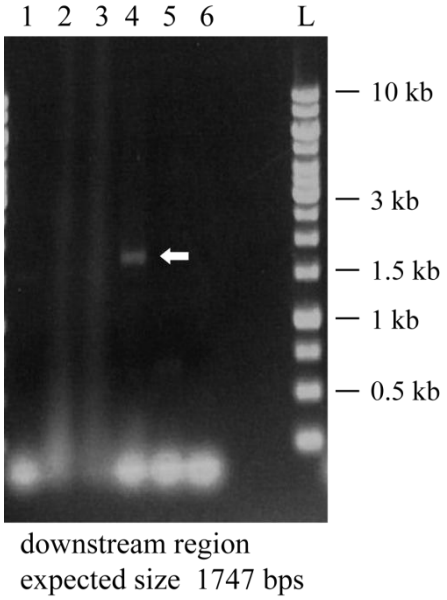



**Figure 3.1.10.** Downstream integration of the pJL10 plasmid into the *C. acetobutylicum* genome present in the sample treated with 5-FU to induce the integration event. The DNA samples were isolated from six pJL10–transformed colonies grown in an MES-uracil medium supplemented with 5-FU. One colony (no. 4, indicated by the arrow) produced a band of the expected size. L, 1 kb DNA ladder; 1 – 6, six colonies chosen for the molecular check.

Out of multiple colonies of *C. acetobutylicum* transformed with pJL10 and treated with 5-FU, only six grew in the presence of the chemical, and were checked by PCR. One DNA sample produced a band of the expected size, 1747 bps, and this sample was used for further tests (Fig. 3.1.10). However, as in previous experiments it was not possible to maintain this sample for multiple re-streaks, regardless of the uracil supplementation. All primers were used as described in previous sections and are listed in the table 9.1.

Although it seems obtaining the first recombinant in this particular region is difficult, it is strongly recommended to continue applying the methods described and testing them along with new deletion cassettes and different concentrations of uracil supplementation.

### 3.1.6.2. Cloning of the *recA* gene into pJL11

The RecA protein is a part of the DNA repair and recombination pathway. It plays a crucial role in homologous recombination by creating synapsis between two

complementary regions of DNA, because of its multiple DNA binding sites. The pJL11 was used as a backbone to create the pJL11-RecA plasmid bearing a copy of *recA* along with a 157 bp-long upstream sequence supposedly containing its promoter region. The *recA* region of *C. acetobutylicum* (CAC1815) of 1230 bps in size was amplified using RecA-pJL11 fwd and rev primers (Tab. 9.1) and cloned into pJL11 with *Kpn*I and *EcoR*I restrictases (Fig. 3.1.11).

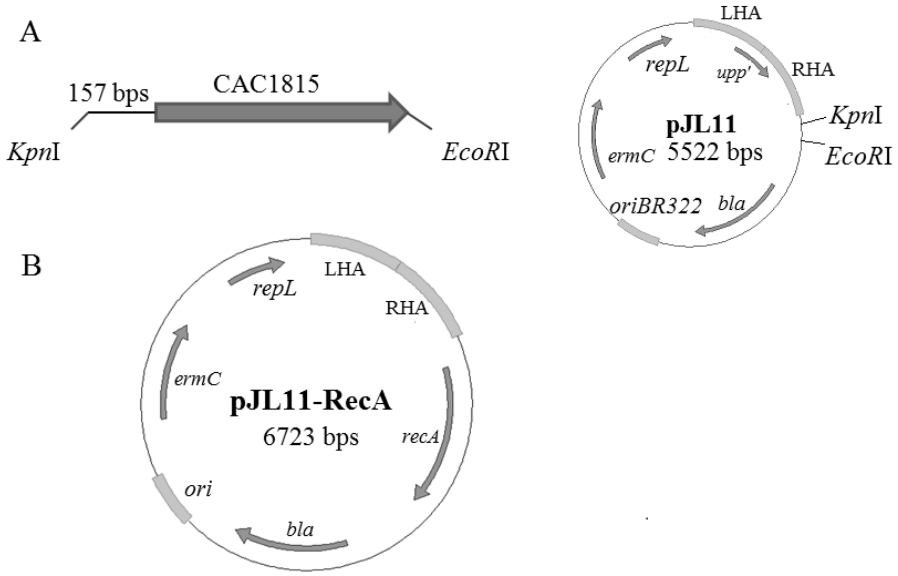



**Figure 3.1.11. (A)** Schematic representation of the amplified *recA* cassette with restrictases used for the cloning procedure (on the left) and the pJL11 plasmid with *KpnI* and *EcoR*I restriction sites indicated (on the right). **(B)** Cloning of the *recA* cassette into pJL11 resulted in the pJL11-RecA plasmid. *RepL*, origin of replication for clostridia; *oriBR322/ori*, origin of replication for *E. coli*.

At the time of the creation of the pJ11-RecA plasmid a fully developed and well-described *pyrE*-based system was received from the Nigel Minton group of the University of Nottingham (Ehsaan, PhD thesis, 2013), and therefore the *C. acetobutylicum ΔpyrE* strain was used to create clean deletion mutants, while work on the Δ*upp* mutant construction was postponed.

### 3.1.7. Development of the new electroporation protocol to overcome troubles with difficult transformations of *C. acetobutylicum*

In this work some of the plasmids were found to be exceptionally difficult to transfer into *C. acetobutylicum* cells with the standard electroporation protocol. Therefore a new, more sensitive method was required. A rapid, more efficient protocol for electroporation of *Clostridium perfringens* was described recently (Jiráskova *et al.*, 2005). Based on this protocol and our observations the standard protocol was modified as described in the Material and Methods section 2.5.1.3. Cells were harvested at a lower optical density value ($OD_{600}$ 0.20 – 0.35) and were washed only once in an ETM buffer. Additionally, the cells mixed with ultra-pure, salt-free DNA were left for 10 – 15 minutes on ice.

For lower amounts of DNA (around 3µg) 9 colonies per 1 µg of plasmid were obtained, and for higher amounts of DNA (around 7 µg) 12 colonies per 1 µg of plasmid were obtained. The values given are the average numbers from 3 independent experiments on the same plasmid DNA (pMADK*upp*$^-$).

More clones were obtained with this new method, especially when large plasmids (over 10 kb) were used, when compared to the standard electroporation protocol. It was also observed the new protocol works significantly better mainly for those experiments in which no transformants were obtained using the standard protocol, while in the other cases the electroporation efficiencies were broadly similar.

### 3.1.8. Establishing the inhibitory concentrations of 5-fluorouracil for *Clostridium saccharobutylicum* NCP 262

*The C. saccharobutylicum* NCP 262 genome contains the uracil phosphoribosyltransferase gene, CSA00515 (Poehlein *et al.*, 2013), and therefore it would be possible to create an *upp*-deficient strain for clean deletion mutation in this organism. For this purpose it was necessary to determine the inhibitory concentration of the selective chemical, 5-fluorouracil (5-FU). Tests were performed in a liquid medium and on agar plates, using ten different concentrations of 5-FU, and repeated three times. Here the average results of three independent experiments are shown, with standard

deviation values. It is shown *C. saccharobutylicum* is sensitive to 5-FU, and is able to grow only at low concentrations (10 µg/ml - 50 µg/ml) after several days of cultivation in a rich liquid medium, and not able to grow in higher concentrations at all (Tab. 3.1.1 A). Therefore, it is proposed the inhibitory concentration for the solid and liquid cultures is 150 µg/ml.

**A**.

| 5-FU (µg/ml) | 0 | 10 | 25 | 50 | 100 | 150 | 200 | 250 | 300 | 350 | 400 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| growth | 1d±0 | 4d±0 | 4d±0 | 6d±1 | X | X | X | X | X | X | X | X |

**B.**

| 5-FU (µg/ml) | 0 | 10 | 25 | 50 | 100 | 150 | 200 | 250 | 300 | 350 | 400 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| number of colonies | un. | 250 ±12.49 | 50 ±4.04 | 50 ±5.86 | X | X | X | X | X | X | X | X |

**Table 1.3.1. (A)** A 5-fluorouracil resistance test of the wild type *C. saccharobutylicum* in increasing amounts of 5-FU in a CGM liquid medium, given in µg/ml. The letter 'd' indicates the number of days after which we observed active growth. The inocula were composed of $3x10^5$ cells. **(B)** A 5-fluorouracil resistance test of the wild type *C. saccharobutylicum* in increasing amounts of 5-FU, given in µg/ml, presented as the number of colonies noticeable after 1 week of incubation on 2 xYTG plates in the dark. Un., uncountable; X, no growth observed.

## 3.2. Restriction-modification systems of *Clostridium saccharobutylicum* NCP 262

### 3.2.1. Type I restriction–modification systems recognition

*C. saccharobutylicum* NCP 262 has two restriction-modification systems annotated during the strain sequence analysis (Poehlein *et al.*, 2013). The analysis of their gene organization suggests they are both restriction-modification (RM) type I systems, with the *hsdR* gene, coding for a restriction endonuclease R subunit, and hsdM-hsdS subunits, forming together methyltransferase. The first of two RM systems

(RM1) identified in *C. saccharobutylicum* consists of three genes: the restriction subunit (*hsdR1*, CSA00451), the methylation subunit (*hsdM1*, CSA00452) and the specificity subunit (*hsdS1*, CSA00453). The second RM system (RM2) contains all three subunits: *hsdR2* (CSA02977), *hdsM2* (CSA02981) and *hdsS2* (CSA02979) and two hypothetical genes, CSA02978 and CSA2980. The structure of *C. saccharobutylicum* RM systems is shown in figure 3.2.1.

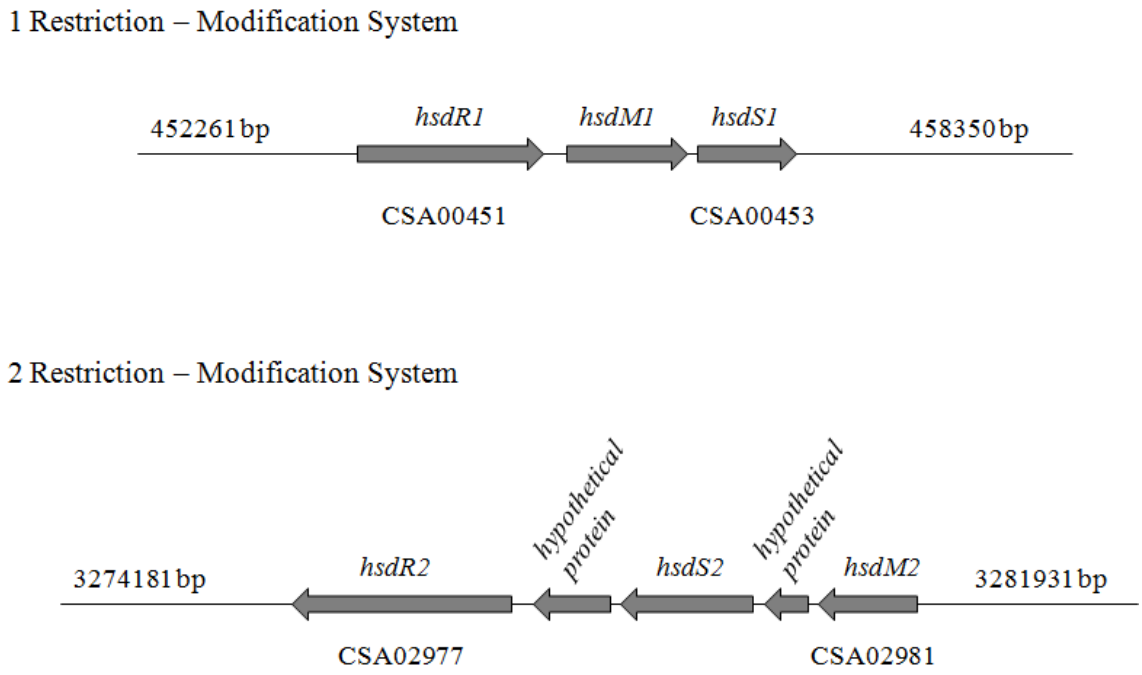



**Figure 3.2.1.** Structure of the restriction-modification operons of *C. saccharobutylicum* NCP 262. The first operon is composed only of the restrictase (HsdR) and methyltransferase subunits (HsdM and HsdS), while the second also contains two hypothetical proteins of unknown function. The genes are described more profoundly in the text.

Type I RM systems are divided into five independent families, called IA, IB, IC, ID and IE. All of them show differences in antibody cross-reactivity, DNA hybridization patterns, subunit complementation, existenceg only within a family but not between families, and gene organization (Roberts *et al.*, 2012). In the first restriction-modification system of *C. saccharobutylicum* the *hsdR*1 gene is followed by the *hsdM1* and *hsdS1* genes, which is typical for the IA or IB family. In the second RM the *hsdMS2* operon precedes the *hsdR2* gene, which is typical for the IC family (Kulik & Bickle,

1996). The gene order and size of the respective M and S subunits affirmed the assignment of RM1 to family IA and RM2 to family IC (Kulik and Bickle, 1996).

To confirm this theory a protein-protein BLAST (Basic Local Alignment Search Tool) and DELTA-BLAST experiments were done. The DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST) is a new algorithm that constructs a position-specific scoring matrix (PSSM) based on the results from a Conserved Domain Database search and then uses it to search a sequence database (Marchler-Bauer *et al.*, 2013; Boratyn *et al.*, 2012; Schäffer *et al.*, 2001; Atlschul *et al.*, 1997). Using this algorithm to blast the HsdR1 amino acid sequence against the *Escherichia coli* taxid it was found that, among sequences of classified restriction enzymes, the *EcoKI* R protein, belonging to the IA family, resulted in the most significant alignment (max identity 21%, query coverage 85%). For the *hsdR2* peptide sequence the best alignment was the *EcoR124II* R protein, a member of the IC family of restriction enzymes (max identity 17%, query coverage 91%). For this experiment the *E. coli* peptides were chosen due to the components of their RM systems being well-known and well described.

The main reason to investigate the methylation systems of *C. saccharobutylicum* is to overcome the activity of their restrictases and enable the possibility to transform the organism with exogenous DNA. Despite the biotechnological importance of this solventogenic microbe (Ni *et al.*, 2013), no successful transformation has yet been described. This might be attributed to the two restriction-modification (RM) systems. The development of an efficient conjugation or transformation protocols is a prerequisite for genetic engineering of *C. saccharobutylicum* NCP 262. *In vitro* DNA methylation is simple and fast, although mostly inefficient. Creation of an *in vivo* methylation system based on the *E. coli* strain requires more time and resources, but is more reliable once adapted.

### 3.2.2. Exogenous DNA methylation

#### 3.2.2.1. *In vitro* DNA methylation

Along with the development of an *in vivo* methylation system for *C. saccharobutylicum* an attempt was made to establish a method to modify plasmid DNA *in vitro* with putative enzymes of the organism. The protocol was based on a method of Alegre and colleagues (Alegre *et al.*, 2004), created for another Gram-positive organism, *Lactobacillus plantarum*.

A crude extract from *C. saccharobutylicum* cells grown to the $OD_{600}$ value of 2.0 was isolated using ultrasound, and cell membrane disruption was checked under the microscope. It was observed that the most effective way to disrupt the cells was a 30-min-long incubation with 30µl lysozyme (20 mg/ml) at 37°C and 5 sonications, each lasting 5 minutes, with a 1-minute pause for cooling the probe down .

The standard curve to determine the protein concentration in the crude extract from the cells was done using the Bradford assay on fixed concentrations of bovine serum albumin (BSA) and the results are shown in figure 3.2.2.

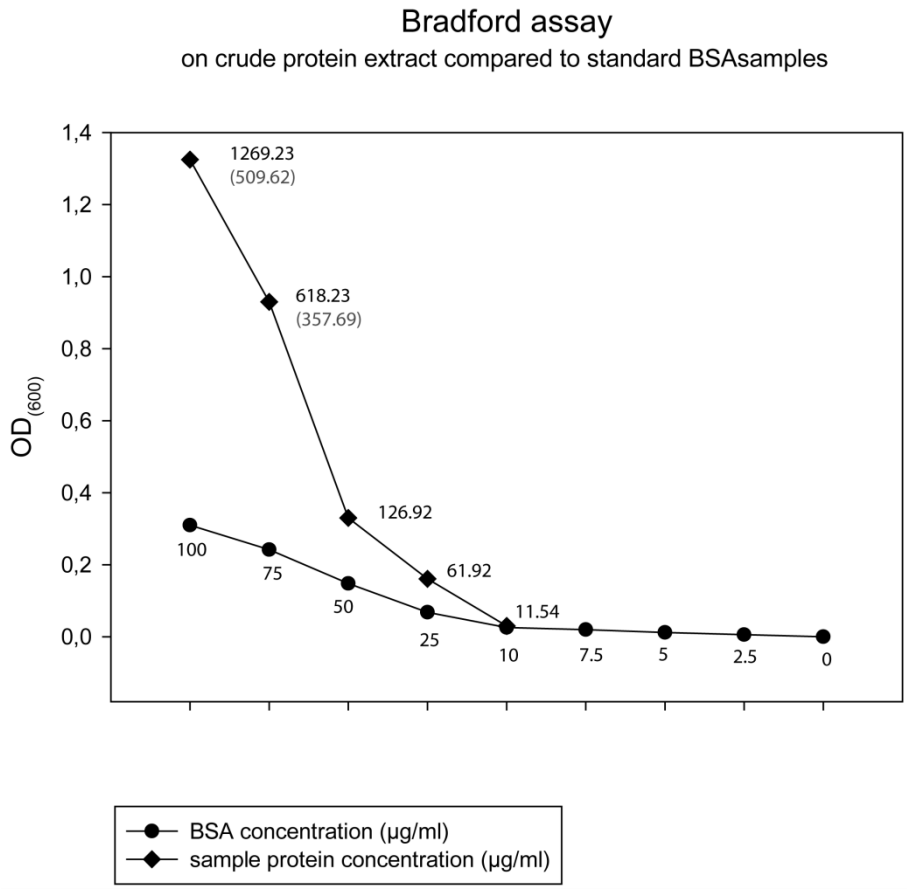



**Figure 3.2.2.** Standard protein concentration curve of the Bradford assay made on BSA samples compared with the protein curve made on samples from the crude protein extract from 45 ml overnight culture of *C. saccharobutylicum* cells, centrifuged and resuspended in 2 ml of the buffer. After the Bradford assay the two samples with the highest concentrations of crude protein extract were diluted twice for measurement purposes, and the non-computed values are given in brackets. The results shown come from the representative experiment. The Bradford assay was repeated three times and all results were similar; for each optical density measurement three technical replicates were made, as described in section 2.8.2.4.

The DNA of the pMTL007C-E2 plasmid was incubated with crude protein extract and a buffer mix for the *in vitro* methylation, as described in the Materials and Methods section. Next, the methylated plasmid was used in a standard *Clostridium* transformation protocol (described in Materials and Methods). In the meantime a new transformation protocol was developed (Results, section3.1.6) and tested for *C. saccharobutylicum*. However, neither of the protocols gave the expected results. At the same time only several single colonies were observed on the non-selective plates in

the positive transformation test, which would suggest high sensitivity of the microorganism to the conditions of either of the transformation protocols.

Along with the *in vitro* methylation protocol an *in vivo E. coli*-based system was developed, and considering the troubles with the *in vitro* procedure, more attention was focused was on the latter one, which proved to be very efficient in later experiments.

### 3.2.2.2. *In vivo Escherichia coli*-based DNA methylation system

#### 3.2.2.2.1. Construction of pJL1 and pJL2

To create the *in vivo* methylation system the native methyltransferases of *C. saccharobutylicum* were used. Since no information on the activity and specificity of either of the enzymes was provided, two methylation plasmids: pJL1 and pJL2 containing MTase1 and MTase2, respectively, were created and their methylation efficiencies were compared.

The genes for methylation and specificity subunits of the RM1 operon were amplified with the 010 primers from the genomic DNA of *C. saccharobutylicum* (all primers used for pJL1 and pJL2 plasmids creation are listed in table 9.2) and cloned into pUC19 using the *PstI/XbaI* sites, resulting in the plasmid pJL010. Respectively, methylation and specificity subunits of RM2 were amplified with 020 primers and cloned into pUC19 using *PstI/XbaI* and *XbaI/EheI* sites resulting in the plasmid pJL020.

The T7 promoter from the pJet plasmid (Thermo Scientific) was cloned into pJL010 and pJL020 using the *PciI/PstI* sites and the JetT7 primers resulting in pJL011 and pJL021 (Fig. 2.3.2). Next, the pACYC184 plasmid backbone was amplified using the PACYC_backbone primers, introducing *SpeI/SmaI* restriction sites. This fragment was ligated with the methylation cassettes, including the T7 promoter, amplified by the 011 and 021 primers. Then the construct was amplified using the JL-Tet primers and closed by the introduced *KpnI* site, resulting in pJL1 and pJL2 (Fig. 3.2.3).

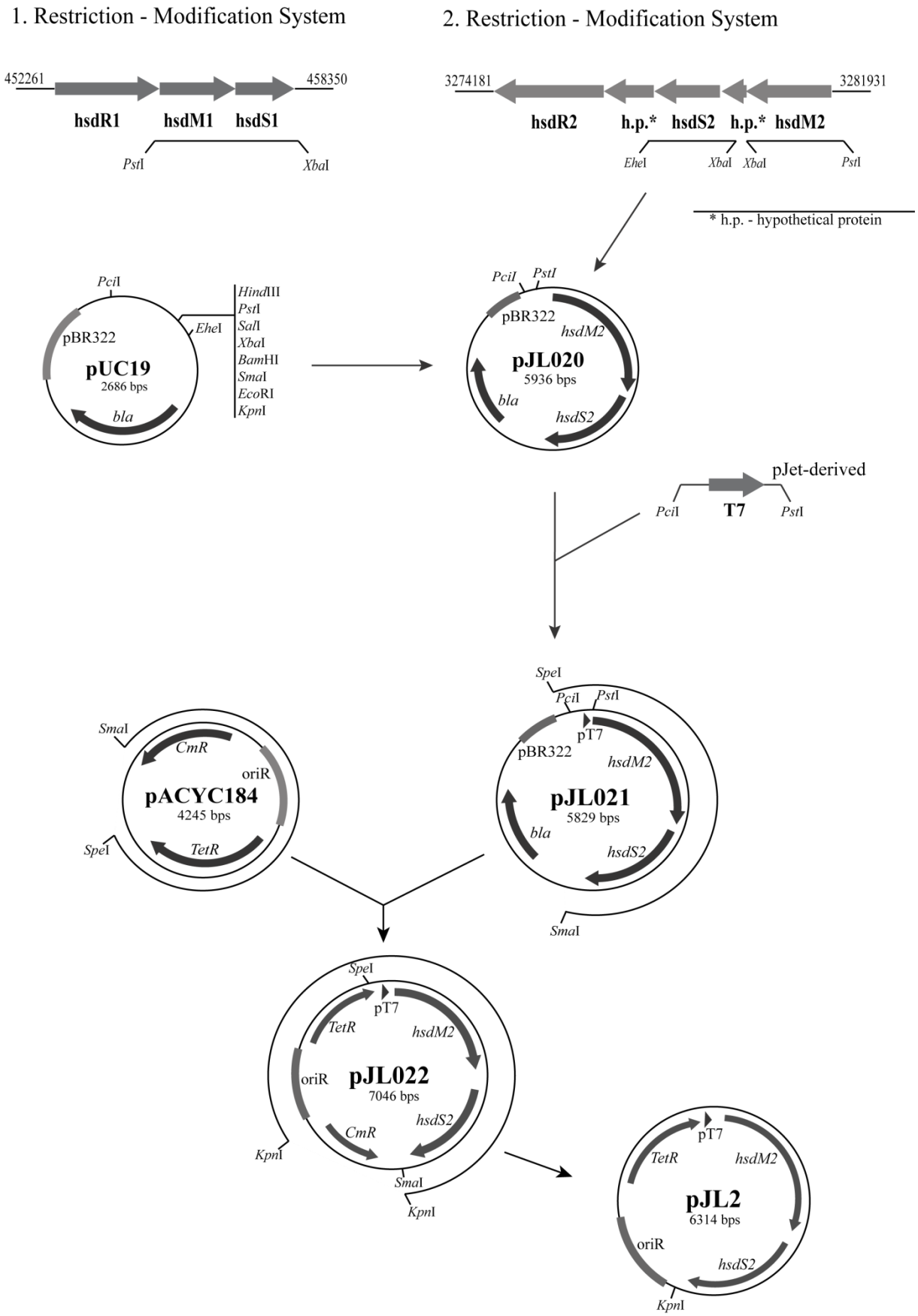



**Figure 3.2.3.** Schematic representation of the creation of the pJL1 and pJL2 methylation plasmids. The structure of both restriction-modification operons of *C. saccharobutylicum* is presented (at the top) and the methylation cassettes regions are indicated with solid black lines

shown below the graphic representation of the genes, along with the restriction enzymes used for cloning purposes. Below, the pUC19 plasmid, used as a backbone for the pJL1 and pJL2, is shown. The figure shows in detail the process of pJL2 plasmid construction. The plasmid pJL1 was made accordingly.

#### 3.2.2.2.2. *E. coli* host strains for pJL1 and pJL2 plasmids

Initially, both plasmids were supposed to be propagated in the Top10 strain, and later introduced into the *E. coli* strain with T7-RNA polymerase in the genome, i.e. the BL21(DE3) strain, to assure the effective expression of the MTase cassettes from the T7 promoter. However, during preliminary tests of the conjugation protocol it was observed that plasmid DNA was methylated and introduced into *C. saccharobutylicum* cells also when the Top10 strain was a host for the methylation plasmids. The results were repeated in several tests, and therefore the non DE3 lysogen strains were also considered as future hosts for methylation plasmids. Two possible explanations of this phenomenon exist – either the expression is driven by a native MTase promoter, or by the promoter of a Tet-resistance gene located upstream on the plasmid. An attempt to detect a native promoter using the BPROM algorithm (SoftBerry) was made, however none were found within the first 200 nucleotides of the cloned MTase region. However, it must be underlined that the BPROM algorithm is designed to search for *E. coli* promoter regions mainly, and may not recognize some of the clostridial regulatory DNA elements.

To this end, pJL1 and pJL2 were introduced into three different strains of *E. coli* hosts – the Top10, a standard cloning and plasmid propagation strain; the ER1793, an *mcrA*-*BC* and *mrr*-deficient strain, which is a close relative of the ER2275 strain, used as a host for pAN1 and pAN2 methylation plasmids for *C. acetobutylicum* (Mermelstein & Papoutsakis, 1993); and the INV110, a *dam* and *dcm*-deficient strain (Tab. 2.2). All three hosts were tested for their methylation and conjugation efficiency. It was established that the first two *E. coli* strains proved to be very good hosts for the purpose of this work, showing fast and efficient growth. The third strain, INV110, when tested grew significantly slower in the presence of selective antibiotics than the two other

strains. For all further conjugation experiments, the fastest growing Top10 strain was used.

### 3.2.3. Modification of *C. saccharobutylicum* cells with plasmid DNA

#### 3.2.3.1. Efficiency of conjugation

Since nothing is yet known about the sequence specificity of the restriction systems of *C. saccharobutylicum*, both methyltransferases (MTases) from the organism were tested for their ability to overcome the activity of restrictases. To this end, two methylation plasmids pJL1 and pJL2 (Fig. 3.2.3) were constructed, expressing the methyltransferases of RM1 and RM2 in an *E. coli* Top10 strain.

*C. saccharobutylicum* was conjugated with an *E. coli* strain bearing a random plasmid to test the efficiency of the newly-created modification systems. For this purpose the pMTL007C-E2 plasmid was used, as this *E. coli-Clostridium* shuttle vector with the Ll.ltrB group II intron is widely used in the creation of ClosTron mutants. The conjugation method was performed as described in section 2.5.1.4 of Material and Methods. Multiple colonies resistant to thiamphenicol were obtained when using either the pJL1 or pJL2 plasmid, confirming a successful transfer of pMTL007C-E2 into *C. saccharobutylicum* cells. It was possible to restreak the colonies multiple times, which implied no loss of plasmid. For molecular proof a PCR check was performed by amplifying a specific region of pMTL007C-E2, using the RAM primers (Tab. 9.5) on the DNA extracted from five randomly picked clostridial colonies, transformed with the plasmid methylated by either pJL1 or pJL2. All of the tested samples produced a band of the expected size (Fig. 3.2.4 A). Furthermore, 5 µl of total DNA extracted from one of the five colonies modified with each methylation variant (pJL1 or pJL2), which had all been checked with PCR, were used to transform the *E. coli* Top10 cells. The pMTL007C-E2 particles should be present in the total DNA extracted from the *C. saccharobutylicum* transconjugants, therefore this procedure should result in Tm-resistant, plasmid-bearing *E. coli* cells. As expected, it was possible to observe colonies on selective plates, and a randomly-picked *E. coli* transformant from each methylation variant was used for plasmid DNA extraction and restriction analysis with *Eco*RI.

Digestion products were visualized on 1% agarose gel; the pMTL007C-E2 plasmid should give four bands of the following sizes: 6244 bps, 2313 bps, 324 bps and 26 bps (Fig. 3.2.4 B).

Due to the conjugative transfer of the DNA, requiring the presence of plasmid-bearing *E. coli* cultures, there is always a risk that plasmid DNA extracted from Tm-resistant clostridial cultures could originate from *E. coli* contamination. To rule out this possibility 200µl of each previously PCR and *Eco*RI-checked clostridial cultures were plated on the LB+Cm plates and left overnight in aerobic conditions, as a negative control. The plates remained clear after 5 days of incubation at 37ºC, proving there was no *E. coli* contamination of the clostridial cultures.

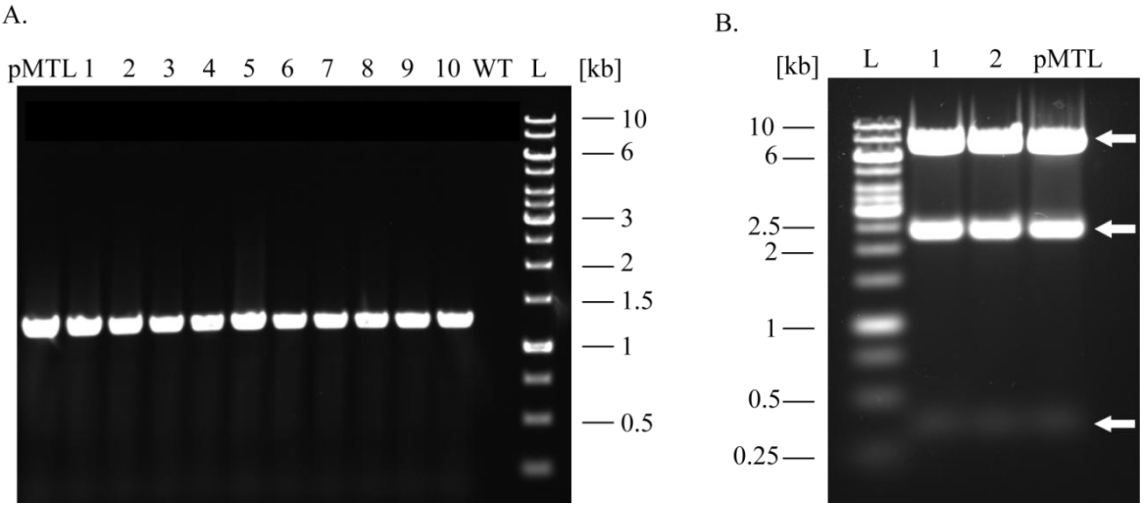



**Figure 3.2.4. (A)** The PCR test on RAM primers confirming the presence of pMTL007C-E2 in conjugated *C. saccharobutylicum* cells. Lanes: pMTL; control PCR product on empty pMTL007C-E2; 1 – 5, PCR on the total DNA isolated from five *C. saccharobutylicum* colonies modified with the pJL1-methylated pMTL007C-E2 plasmid; 6 – 10, PCR on the total DNA isolated from five *C. saccharobutylicum* colonies modified with the pJL2-methylated pMTL007C-E2 plasmid; WT , *C. saccharobutylicum* wild type control; L, 1kb DNA ladder. **(B)** *Eco*RI restriction analysis of plasmid DNA isolated from *E. coli* transformed previously with the total DNA extracted from conjugated *C. saccharobutylicum* cultures. Three visible bands are indicated by the arrows. Lanes: L, 1kb ladder, NEB; 1, colony modified with the pJL1-methylated pMTL007C-E2 plasmid; 2, colony modified with the pJL2-methylated pMTL007C-E2 plasmid; pMTL, pMTL007C-E2 control.

The conjugational approach creates less challenging conditions for the clostridial host cells, but at the same time it is much slower and more laborious than the electroporation method. Several attempts to electroporate *C. saccharobutylicum* with the newly developed methylation system were made, however with no positive result, and only a few clostridial colonies were spotted on non-selective plates. Therefore for all future experiments, the conjugation protocol was used to introduce plasmid DNA into *C. saccharobutylicum* cells.

### 3.2.3.2. Different origins of replication for *C. saccharobutylicum*

In order to assess the usefulness of four widely used origins of replication for clostridia, four modular plasmids methylated with pJL1 were introduced into *C. saccharobutylicum* and replica-plated on 2 x YTG plates with and without any selective antibiotic.

The stability of the four plasmids: pMTL85151, pMTL82151, pMTL83151 and pMTL84151 (Tab. 3.2.1) bearing four Gram-positive origins of replication from pIMP13, pBP, pCB102 and pCD6 respectively were tested. In three independent repeats of this experiment it was observed that the origins from pIM13 and pBP were the most stable as the plasmids pMTL85151 and pMTL82151 were still present in the tenth subculture, while the plasmids pMTL83151 and pMTL84151, containing origins from pCB102 or pCD6, were lost, at the latest, in the seventh subculture in the absence of antibiotic pressure (Tab. 3.2.1 A and B).

### 3.2.4. Analysis of the *C. saccharobutylicum* NCP 262 restrictases

#### 3.2.4.1. Creation of ClosTron mutants in *hsdR1* and *hsdR2* restrictase genes

To test the importance of RM1 and RM2 in the restriction of *C. saccharobutylicum*, the ClosTron mutants *hsdR1::int* and *hsdR2::int* were constructed by knocking-out each of the restrictases. One of the characteristic traits of type I RM-systems is that the M and S subunits are transcribed from a different promoter than the R subunit. Therefore, a disruption of the *hsdR* gene should not affect the activity of its MTase (Wilson & Murray 1991).

ClosTron plasmids with introns targeting *hsdR1* and *hsdR2* were designed using the www.ClosTron.com website (Perutka *et al.*, 2004), synthesized by DNA2.0 (Menlo Park, CA) and methylated with pJL1. Mutants were created as described previously (Heap *et al.*, 2010), except for the usage of lincomycin instead of erythromycin during the mutant selection step. The correct mutations were verified by PCR (Fig. 3.2.5 A) and sequencing. Neither the *hsdR1* nor the *hsdR2* mutation affected the growth rate of *Clostridium* strains (Fig. 3.2.6). The single integration event of the ClosTron group II intron in each mutant was confirmed by Southern blotting (Fig. 3.2.5 B); the *hsdR1::int* and *hsdR2::int* mutants genomic DNA were digested with *HindIII* restrictase and should give visible bands of around 2.4 kb and 3.1 kb, respectively.

Phenotypes of the mutants and the WT strain were characterized according to their ability to receive DNA by conjugation with unmethylated and *in vivo* methylated pMTL85151 plasmid DNA (Tab. 3.2.3 A and B).

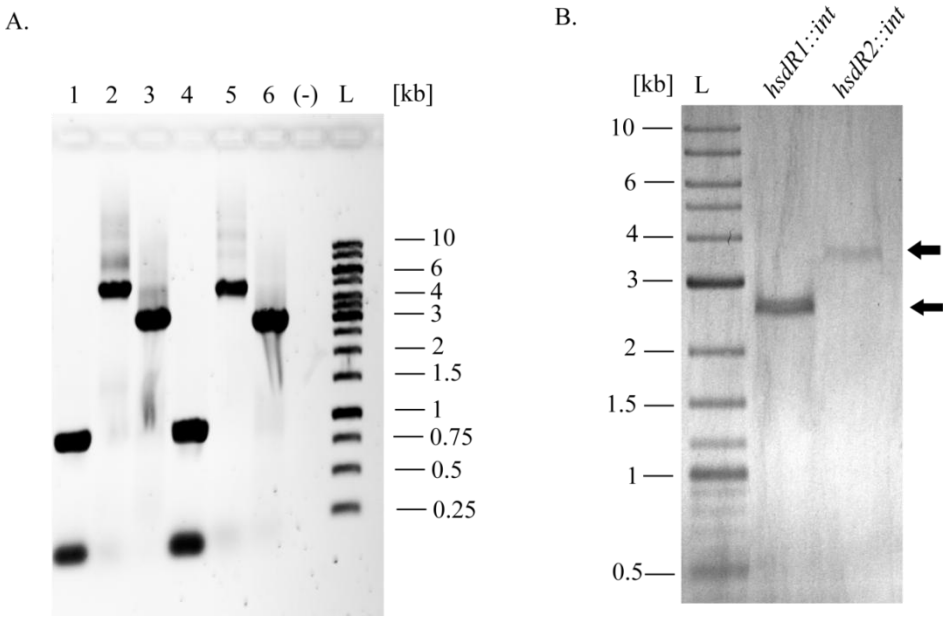



**Figure 3.2.5. (A)** Integration of the *hsdR1* and *hsdR2*-targeting intron into the *C. saccharobutylicum* genome. The correct integration of the intron was determined by the PCR reaction on mutant and wild type samples using different primer sets. Lanes: 1, RAM primers, *C. saccharobutylicum hsdR1::int* sample; 2, *hsdR*1 gene primers, *hsdR1::int* sample; 3, *hsdR*1 gene primers, *C. saccharobutylicum* wild type sample; 4, RAM primers, *hsdR2::int* sample; 5, *hsdR*2 gene primers, *hsdR2::int* sample; 6, *hsdR*1 gene primers, *C. saccharobutylicum* wild type sample; (-), negative control; L, 1 kb DNA ladder. The primers used for the molecular check are listed in table 9.5. **(B)** Southern Blot showing single integration event of the *hsdR1*and *hsdR2*-targeting intron into the *C. saccharobutylicum* genome on *HindIII* -digested mutant DNA; L, biotinylated 2-log DNA ladder NEB. Visualized mutant bands are indicated by black arrows.

**A** **2xYTG**

| | pMTL85151 | pMTL82151 | pMTL83151 | pMTL84151 |
|---|---|---|---|---|
| **Restreak number** | Number of colonies | | | |
| 1 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 2 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 3 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 4 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 5 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 6 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 7 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 8 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 9.7 ± 0.58 |
| 9 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 9.7 ±0.58 |
| 10 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 9.7 ±0.58 |

**B** **2xYTG+Tm (15 µg/ml)**

| | pMTL85151 | pMTL82151 | pMTL83151 | pMTL84151 |
|---|---|---|---|---|
| **Restreak number** | Number of colonies | | | |
| 1 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 2 | 9.7 ±0.58 | 10.0 ± 0 | 10.0 ± 0 | 10.0 ± 0 |
| 3 | 8.7 ±1.53 | 9.0 ±1.0 | 9.7 ±0.58 | 10.0 ± 0 |
| 4 | 7.0 ±1.0 | 8.7 ±0.58 | 9.7 ±0.58 | 9.7 ±0.58 |
| 5 | 5.0 ±1.0 | 6.7 ±1.15 | 9.3 ±0.58 | 9.3 ±1.15 |
| 6 | 3.0 ±1.73 | 5.0 ±1.73 | 8.3 ±1.15 | 9.0 ±1.0 |
| 7 | 0.0 ± 0 | 2.7 ±0.58 | 7.3 ±1.52 | 8.3 ±1.53 |
| 8 | 0.0 ± 0 | 0.0 ± 0 | 6.7 ±0.58 | 7.3 ±0.58 |
| 9 | 0.0 ± 0 | 0.0 ± 0 | 5.7 ±0.58 | 7.3 ±0.58 |
| 10 | 0.0 ± 0 | 0.0 ± 0 | 5.3 ±0.58 | 6.7 ±0.58 |

**Table 3.2.1.** The analysis of the stability of four different Gram-positive origins of replication in *C. saccharobutylicum* cells: pMTL85151, pMTL82151, pMTL83151 and pMTL84151 on the (A)

non-selective and (B) selective 2 x YTG agar plates. The table shows the number of single cell-derived colonies visible on the plate, and the starting number of colonies equaled 10. The decreasing number of colonies appearing on the selective plates suggested plasmid loss. The standard thiamphenicol concentration was applied, and the numbers of colonies visible on selective and non-selective plates were counted after every restreak. The numbers given in the table are the average of the scores of three independent experiments, with standard deviation values.

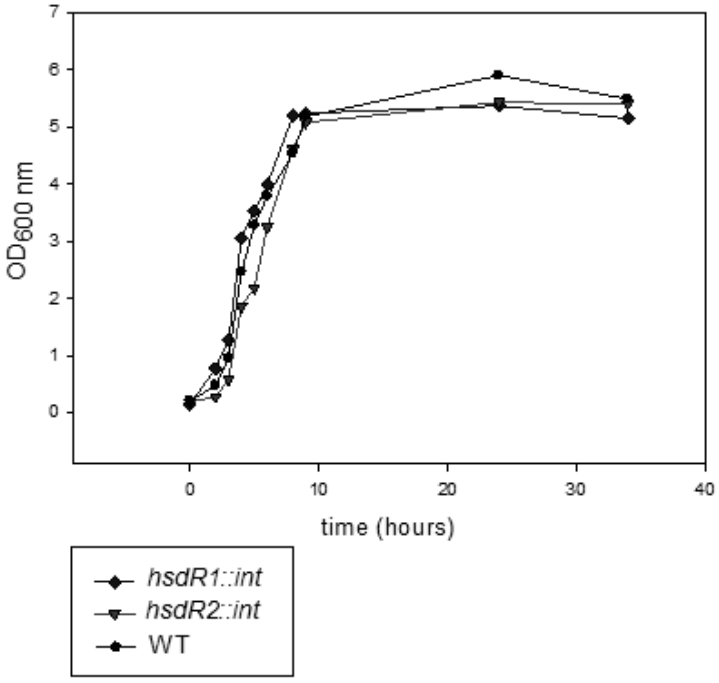



**Figure 3.2.6.** Growth curves of *C. saccharobutylicum* wild type and two restrictase mutant strains, *hsdR1::int* and *hsdR2::int* in rich medium batch culture experiments. The results shown come from the representative experiment. The growth curve experiment was repeated three times and all results were similar; for each optical density measurement three technical replicates were made, as described in section 2.3.

The markerless *C. saccharobutylicum* strain with disrupted restrictase genes is important for further easy mutant creation. To this end a pCLF1-1 plasmid, bearing the flippase gene (2.5.3, Materials and Methods), was used to remove the *ermB* cassette from the *hsdR1::int* and *hsdR2::int* mutants. However, despite multiple attempts it was not possible to obtain a markerless mutant, and similar problems were described previously (Heap *et al.*, 2014).

### 3.2.4.2. Phenotype analysis of *C. saccharobutylicum hsdR1::int* and *hsdR2::int* mutants

Since nothing is yet known about the specificity of the restriction systems of *C. saccharobutylicum*, the *in vivo* methylation by the methyltransferases (MTases) from the organism was used to overcome the restrictase activity. To this end, two methylation plasmids pJL1 and pJL2 were constructed (Fig. 3.2.3), expressing the methyltransferases of RM1 and RM2 in an *E. coli* Top10 strain, as *in vivo* methylation systems. For approximately $10^7$ *C. saccharobutylicum* WT cells used as recipients on average $1.1 \times 10^{-4}$ and $1.7 \times 10^{-3}$ transconjugants per recipient cell were obtained after methylation in donor strains bearing pJL1 (MTase1) and pJL2 (MTAse2) respectively, (Tab. 3.2.2) whereas no transconjugants could be observed without the *in vivo* methylation step.

**A**

**unmethylated donor DNA**

| recipient | |
|---|---|
| **WT** | 0 |
| ***hsdR1::int*** | $1.13 \times 10^{-4} \pm 6.13 \times 10^{-6}$ |
| ***hsdR2::int*** | $3.24 \times 10^{-4} \pm 8.54 \times 10^{-6}$ |

**B**

**methylated donor DNA**

| recipient | methylation plasmid | |
|---|---|---|
| | **pJL1** | **pJL2** |
| **WT** | $1.1 \times 10^{-4} \pm 3.5 \times 10^{-5}$ | $1.7 \times 10^{-3} \pm 3.65 \times 10^{-5}$ |
| ***hsdR1::int*** | $1.25 \times 10^{-3} \pm 8 \times 10^{-4}$ | $4.59 \times 10^{-3} \pm 1 \times 10^{-3}$ |
| ***hsdR2::int*** | $8.04 \times 10^{-4} \pm 6.42 \times 10^{-5}$ | $4.24 \times 10^{-4} \pm 2 \times 10^{-4}$ |

**Table 3.2.2.** Representative conjugation efficiencies for *C. saccharobutylicum* wild type, *hsdR1::int* and *hsdR2::int* strains. *E. coli* transformed with the plasmid pMTL85151 was used as the donor in the presence of *E. coli* CA434 as the helper strain. Donor DNA was **(A)** unmethylated or **(B)** *in vivo* methylated using either pJL1 or pJL2. The efficiency values are given as a number of conjugants per recipient cell with standard deviation numbers given, after 12 hours of incubation at 37ºC and resuspension in 1ml of anaerobic PBS.

When *in vivo* methylation was done using *E. coli* Top10 containing pJL1 with the *hsdMS1* methyltransferase, the number of colonies that successfully received pMTL85151 by conjugation and became thiamphenicol-resistant was one order of magnitude higher than in the wild type, if the *hsdR1::int* mutant was used as a recipient, and eight times higher if the *hsdR2::int* mutant was used. When an *E. coli* Top10 containing pJL2 for *in vivo* methylation was used as a donor there was a threefold difference in the number of conjugants between the wild type and the *hsdR1::int* mutant in favor of the latter one. We also observed less conjugants by one order of magnitude for the *hsdR2::int* recipient compared to the *hsdR1::int* (Tab. 3.2.2 B).

No conjugant colonies were detected for the WT recipient when unmethylated DNA was used. However, conjugants can be observed for both restrictase mutants as recipients, with higher rates for the *hsdR2::int* mutant (Tab. 3.2.2 A).

## 3.3. Sugar and pentose metabolism in solventogenic clostridia

### 3.3.1. Characterization and comparison of sugar metabolism in three solventogenic *Clostridium* strains, *C. acetobutylicum* ATCC 824, *C. saccharobutylicum* NCP 262 and *C. beijerinckii* NCIMB 8052

Solventogenic *Clostridium* strains are known to be able to metabolize simple plant-derived sugars, like glucose, xylose, arabinose, fructose or ribose. Different strains can produce different types of industrially important solvents and acids. This work focused on the ABE (acetone, butanol and ethanol) fermentation in three major solventogenic *Clostridium* strains, and the role of two kinases, the arabinose and the xylose kinase, in the metabolism of simple pentoses was investigated.

To create a background for the more profound analysis of the wild type and mutant strains of solventogenic clostridia, a basic knowledge about the phenotype of the wild type *Clostridium acetobutylicum, C. saccharobutylicum* and *C. beijerinckii* in the most suitable conditions, namely a rich medium supplemented with glucose, was needed to create the control growth curves. *C. acetobutylicum* and *C. saccharobutylicum* were cultured in the same liquid medium (CGM with 5% glucose), however *C. beijerinckii* required more complex conditions and was cultured in 2 x YTG supplemented with 2% glucose, or in RCM with additional glucose supplementation, if needed. In this experiment general observations were made, and it was noticed that *C. saccharobutylicum* was capable of growing to higher optical density values than *C. acetobutylicum*, and its regeneration time after innoculation was much shorter, when compared to the two other strains, regardless of the state of preculture. Although in further experiments *C. acetobutylicum* was observed to be the most resistant to oxygen exposure, it was also the most liable to degenerate due to pSOL1 loss, which affects the solvent production and sporulation abilities of the strain (Xu Zheng, personal communication). *C. beijerinckii* was the most sensitive to oxygen stress and relatively difficult to handle. The best conditions to cultivate this strain were the Reinforced Clostridial Medium (RCM) for both solid and liquid cultures or modified 2 xYTG, with higher sugar supplementation (2%), for liquid cultures only.

The ability of all three solventogenic strains to grow in a rich medium supplemented with glucose, fructose, ribose, xylose or arabinose was tested. To this end a hungate

experiment was designed, in which the strains' behavior in 5 ml of rich medium supplemented with different sugars could be observed. Generally, this experiment only gives information on the growth ability of the organisms under certain conditions, but when conducted carefully from a small innoculum or single colony, it can also produce information on how fast the strain could grow. The same amount of preculture: 0.5 ml, which was 10% of the total medium volume was used for all experiments to observe if the strains are capable of using all the sugars mentioned as carbon and energy sources. For *C. acetobutylicum* and *C. saccharobutylicum* the CGM medium was used, and for the *C. beijerinckii* the 2 x YTG medium was required. All wild type strains showed comparable growth in hungates, when compared to each another, confirming their ability to ferment glucose, fructose, arabinose, xylose and ribose.

To test their ability to ferment glucose in the rich medium (CGM for *C. acetobutylicum* and *C. saccharobutylicum* or 2x YTG for *C. beijerinckii*), batch cultures were supplemented with 200 mM of a sugar, and out of three experiments one was chosen as representative (Fig. 3.3.1).

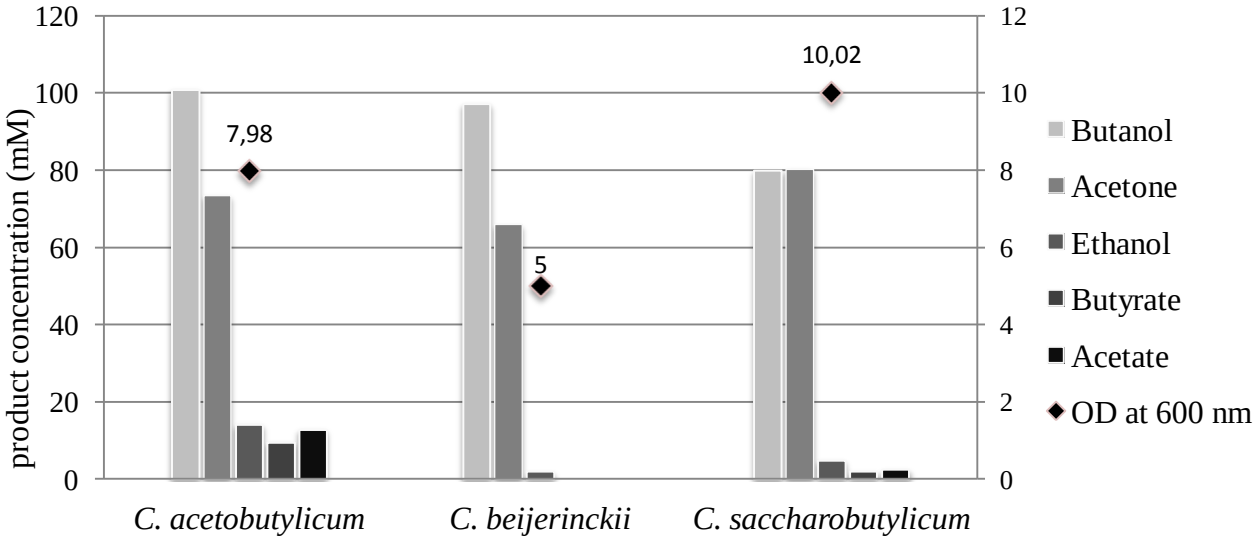



**Figure 3.3.1.** Comparison of the fermentation profile of three major solventogenic *Clostridium* strains: *C. acetobutylicum* ATCC 824, *C. beijerinckii* NCIMB 8024 and *C. saccharobutylicum* NCP 262 in batch cultures of rich medium (as indicated in the text) supplemented with 200 mM glucose. The samples were collected 78h after inoculation time, solvent and acids concentrations were measured by gas chromatography, as described in section 2.8.1. The

$OD_{600}$ values are indicated above as diamond points; for each optical density measurement three technical replicates were made, as described in section 2.3. The results shown come from one representative experiment. The experiments were repeated three times, giving similar results.

### 3.3.2. Characterization of *Clostridium acetobutylicum* ATCC 824 and *Clostridium saccharobutylicum* NCP 262 arabinose and xylulose kinase mutants

#### 3.3.2.1. Arabinose and xylose operons in *C. saccharobutylicum*

During the sequencing of the *C. saccharobutylicum* genome two clusters of presumed xylose metabolism genes were annotated (Poehlein *et al.*, 2013). Using simple nucleotide or protein BLAST (Basic Local Alignment Search Tool) and DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST) experiments, sequences of genes from these two clusters were compared to *Clostridium acetobutylicum* ATCC 824 *xylB* and *araK* kinases, to find possible xylulose and arabinose kinases (Marchler-Bauer *et al.*, 2013; Boratyn *et al.*, 2012; Schäffer *et al.*, 2001; Atlshul *et al.*, 1997). For *C. acetobutylicum araK* the BLAST algorithm gave the highest score with CSA00774 (query cover 100%, identity 84%) and for *xylB* the highest score was produced by CSA03335 (query cover 100%, identity 82%). The DELTA-BLAST algorithm was used to confirm these results and to look for highly conserved regions. The up- and downstream genes were analyzed and two clusters containing genes responsible for arabinose and xylose metabolism were described (Fig. 3.3.2). The first cluster most probably consists of an operon, CSA00700 – CSA00775, containing a transcriptional regulator, L-ribulose-5-phosphate-4-epimerase, transaldolase (*tal*), transketolase (*tkt*), sugar kinase and hypothetical protein, and L-arabinose isomerase, located upstream (CSA00769). The entire second cluster forms possibly one operon, CSA03335 – CSA03329, consisting of xylulokinase (*xylB*), transcriptional regulator or sugar kinase (as annotated), *tal*, *tkt*, hypothetical protein and L-fucose isomerase, with promoter region mapped upstream of the CSA03335 gene (Fig. 3.3.2 and Fig. 3.3.14). Using the nucleotide and protein BLAST each gene of both *C. saccharobutylicum* clusters was analyzed and compared to the genes of *C. acetobutylicum*. In the first cluster most of the genes were aligned with the

*C. acetobutylicum* genome in a protein blast, although some of them produced significant alignments also in nucleotide blast. It is further specified whether the results were obtained with a nucleotide or protein blast experiment. The CSA00769 (L-arabinose isomerase) is homologous to CAC1342, the *araA* gene (protein blast, 98% coverage, 58% identity). The following gene, annotated as a transcriptional regulator is homologous with the *araR* gene of *C. acetobutylicum* (protein blast, 98% coverage, 37% identity) and the L—ribulose-5-phosphate-4-epimerase gene (CSA00771) is a homolog of CAC1341 – *araD*, (protein blast, 100% coverage, 61% identity). For both transaldolase and transketolase the best alignments were found in the nucleotide blast for *C. acetobutylicum* CAC1347 and CAC1348, respectively. The last gene of the first cluster, annotated as a hypothetical protein (CSA00775), was blasted against a nucleotide sequence of clostridia in the NCBI database and all matches were described as D-xylose-proton symporters, also in *C. acetobutylicum* (CAC1345). In the NCBI database the CSA00775 gene is annotated as a putative metabolite transport protein YwtG and it is possible it could be involved in arabinose transport.

Out of the second cluster the nucleotide blast of the CSA03330 gene, annotated as L-fucose isomerase, was matched with the CAC2610 gene, which is also an L-fucose isomerase (nucleotide blast, query coverage 99% and identity 82%). Interestingly, for the hypothetical protein (CSA03331) the best *C. acetobutylicum* alignment was the CAC2611 (nucleotide blast, coverage 99%, identity 82%), which is also annotated as a hypothetical protein. Transketolase and transaldolase were aligned to *C. acetobutylicum* CAC1348 (nucleotide blast, 98% query coverage, 73% identity) and CAC1347 (nucleotide blast, 93% query coverage, 75% identity) genes, respectively (also *tkt* and *tal*). The CSA03334 gene, in the JGI database annotated as a transcriptional regulator or sugar kinase and in the NCBI database as a putative xylose repressor, was aligned to the CAC3673, *xylR* gene (protein blast, 96% query coverage, 36% identity).

Both transketolases of *C. saccharobutylicum* matched the same *tkt* gene of *C. acetobutylicum*. The CSA00773 gene shows 74%, and CSA03332 73% identity with the CAC1348 gene. As for the transaldolases; CSA0333 matched *C. acetobutylicum* CAC1347 with 75% and CSA00772 with 74% identity.

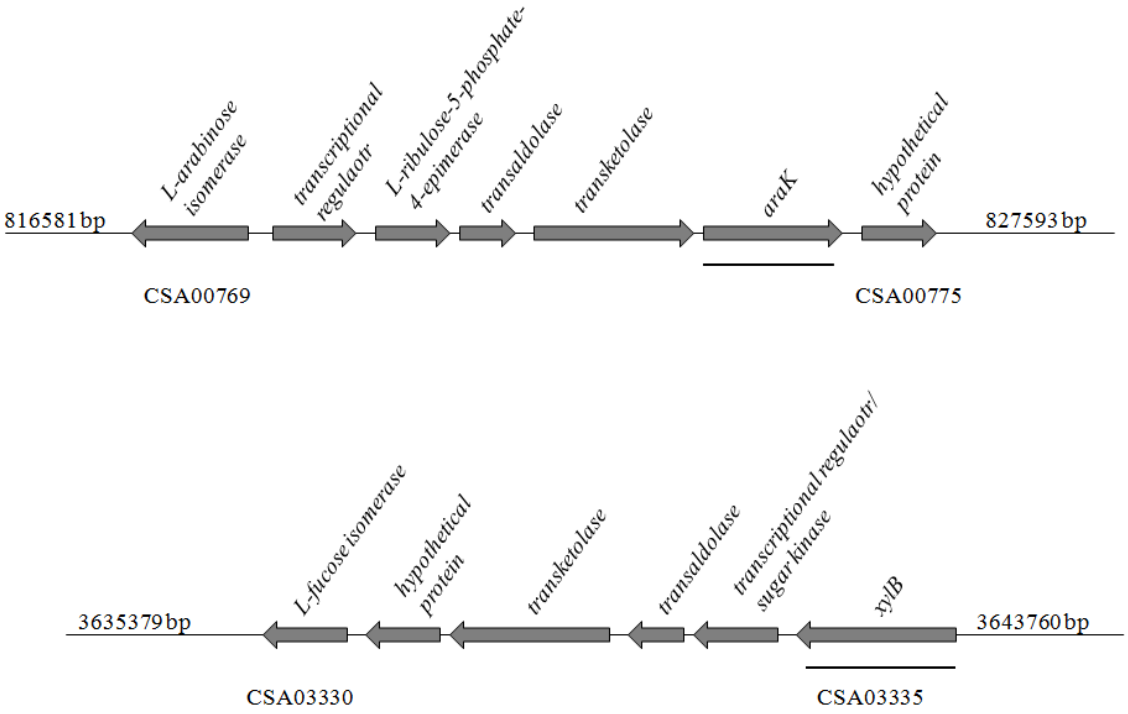



**Figure 3.3.2.** Schematic representation of two clusters of genes involved in the arabinose and xylose metabolism in *Clostridium saccharobutylicum* NCP 262. The kinases are underlined, and the first and last genes of each cluster are indicated with their JGI numbers. Abbreviation of the gene names are explained in the text.

### 3.3.2.2. Arabinose and xylose operons in *C. acetobutylicum* ATCC 824

Two presumed xylose operons of *C. acetobutylicum* ATCC 824, CAC1344 – CAC1349 and CAC2612 – CAC2610, were annotated previously (Grimmler *et al.*, 2010), and in both NCBI and JGI databases CAC2612 and CAC1344 genes are annotated as xylulose kinases. Recent studies, however, point out that the first operon could be involved in arabinose metabolism and experiments performed during this study confirm this theory. For the purposes of this work the CAC1344 – CAC1349 operon was named the arabinose operon, and the CAC2612 – CAC2610 operon the xylose operon. In the analysis of the arabinose operon multiple genes involved in the arabinose metabolism were found and are shown in figure 3.3.3.

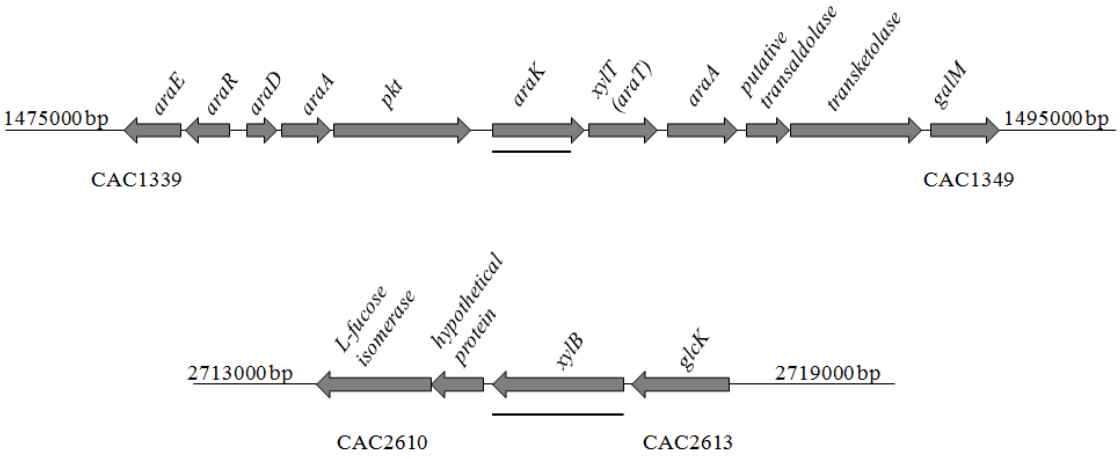



**Figure 3.3.3.** Schematic representation of two clusters of genes involved in the arabinose and xylose metabolism of *Clostridium acetobutylicum* ATCC 824. Both kinases, *araK* and *xylB*, are underlined, and first and last genes of clusters are indicated with their NCBI numbers. Abbreviation of the gene names are explained in the text.

The first cluster, composed of genes involved in the arabinose metabolism, opens with the *araE* gene (CAC1339), coding for a possible sugar-proton symporter, and is followed by *araR*, the transcriptional regulator from the LacI family. The next three genes possibly form an operon, with the *araD* gene (L-ribulose-5-phosphate-4-epimerase), *araA* (arabinose isomerase) and *pkt* (phosphoketolase), with the promoter located upstream of the *araD* gene, CAC1341 (Fig. 3.3.19).

The arabinose operon starts with *araK* (CAC1344), an arabinose kinase, which is followed by *xylT*, annotated as a D-xylose proton symporter, *araA* (L-arabinose isomerase), a putative transaldolase, transketolase and *galM* (CAC1349), coding for aldose-1-epimerase.

The xylose operon consists of three genes, the *xylB* (CAC2612), which encodes a xylulose kinase, a hypothetical protein and L-fucose isomerase (CAC2610). Downstream of the *xylB* there is one additional gene that may be involved in xylose metabolism, the *glcK* – a transcriptional regulator from the NagC/ XylR family, and is present in the same KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway as the CAC2612 gene (Kanehisa *et al.*, 2014).

### 3.3.2.3. ClosTron mutant creation of the *araK* and *xylB* genes of *C. acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262

To investigate the xylose and arabinose metabolism in *Clostridium acetobutylicum*, the ClosTron mutants in two kinases of previously described operons (Grimmler *et al.*, 2010): the *araK* and *xylB*, and a double clean deletion mutant in both genes were constructed. Each kinase gene is located in a pentose metabolism-related gene cluster (Fig. 3.3.3), and for a long time it remained unclear which one is responsible for xylose and which for arabinose fermentation. Additionally, genes homologous to *C. acetobutylicum araK* and *xylB* in *C. saccharobutylicum* (Fig. 3.3.2) were disrupted using the ClosTron method and the phenotypes of mutants of both species were compared.

ClosTron plasmids were created as described previously in Materials and Methods (section 2.5.2) and are listed in table 2.2. For *C. acetobutylicum* appropriate primers (Tab. 9.3) were designed at the www.ClosTron.com web site, using the Perutka algorithm (Perutka *et al.*, 2004) and re-targeted introns were cloned into the pMTL007C-E2 plasmid. For *C. saccharobutylicum* re-targeted introns were synthesized by DNA2.0 (Menlo Park, CA) and sent already-cloned into the pMTL007C-E2 plasmid. The list of all the primers used for each mutant check is attached to the Supplementary Data section (Tab. 9.5). The list of plasmids with integration regions are listed in table 2.2. All mutants were screened with the PCR technique for the correct region of integration (Fig. 3.3.4) and checked with Southern Blot to assure only one integration event (Fig. 3.3.5). For all ClosTron mutants the sizes of the bands produced in the PCR reaction were around 2.1 kb larger than the wild type template-derived product. The ClosTron check for *C. acetobutylicum* was done using the *Cac araK* check primers (1656 bps product) and *Cac xylB* check primers (1558 bps product); for *C. saccharobutylicum* the following primers were used: *Csa araK* check (1693 bps product) and *Csa xylB* check (657 bps product). The molecular check for the *araK* clean deletion was made using *Cac araK* clean deletion check primers on the WT (3266 bps product) and *C. acetobutylicum ΔaraK* (1667 bps product) strains. The molecular check for the *xylB* clean deletion was done using *Cac xylB* clean deletion check primers on the

WT (3548 bps product), *C. acetobutylicum ΔxylB* and *C. acetobutylicum ΔaraKΔxylB* (2051 bps product) templates. It is noticeable that the PCR on the *C. acetobutylicum c araK* template with *Cac araK* clean deletion check primers produced a band of larger size than expected and this phenomenon was observed multiple times, however the sequencing reactions always resulted in the sequence of a correctly constructed mutant.

Additionally, all *C. acetobutylicum* strains were PCR checked for pSOL1 presence using a pSOL check primer pair, producing a band of around 1.2 kb, and exemplary PCR reactions on RAM primers were made for *C. saccharobutylicum* ClosTron mutants. The primers are listed in table 9.5.

For the blotting experiment different restrictases were used. The *C. acetobutylicum araK::int* mutant genomic DNA was digested with the *NcoI* restrictase and produced a visible band of a size of around 6.1 kb. The *C. acetobutylicum xylB::int*, *C. saccharobutylicum araK::int* and *C. saccharobutylicum xylB::int* genomic DNA were digested with the *HindIII* enzyme, producing detectable DNA fragments of around 3.9 kb, 2.6 kb and 6.2 kb in size, respectively. The RAM cassette was visible as a band of a size of 900 bps, and the pMTL007C-E2 was visualized as a band around 8.9 kb large. Additionally, all *C. acetobutylicum* strains were PCR-checked for pSOL1 presence (Fig. 3.3.4).

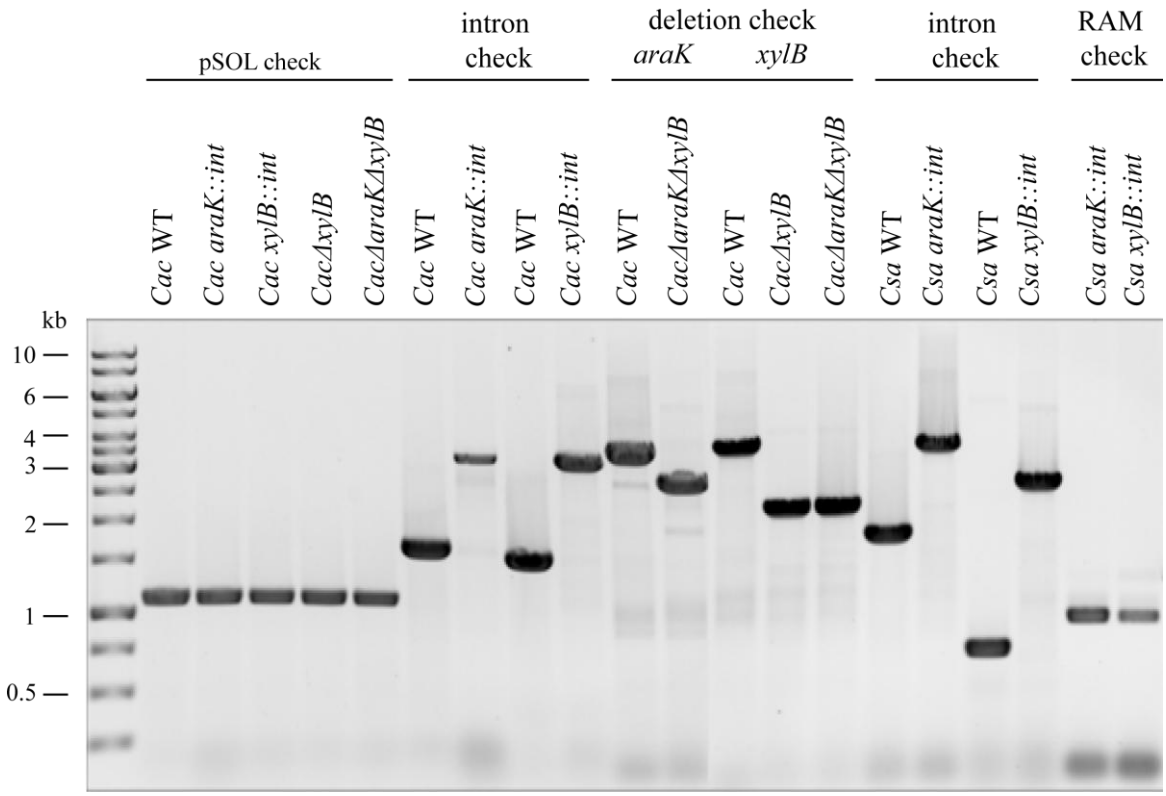



**Figure 3.3.4.** Visualization of the integration of introns targeting the *araK* and *xylB* genes of *C. acetobutylicum* and *C. saccharobutylicum* and the representation of the clean deletion of the *araK* and *xylB* genes of *C. acetobutylicum* on 1% agarose gel. *Cac, C. acetobutylicum; Csa, C. saccharobutylicum*. For every mutant sample a control reaction was performed on wild type DNA using the same primer pair, and the products were shown in juxtaposition on the gel. For all *C. acetobutylicum* strains a special additional PCR check for pSOL1 presence was done, producing a band of around 1.2 kb, and exemplary RAM-primers PCR were made for *C. saccharobutylicum* ClosTron mutants. All primers used for the mutant check are listed in table 9.5.

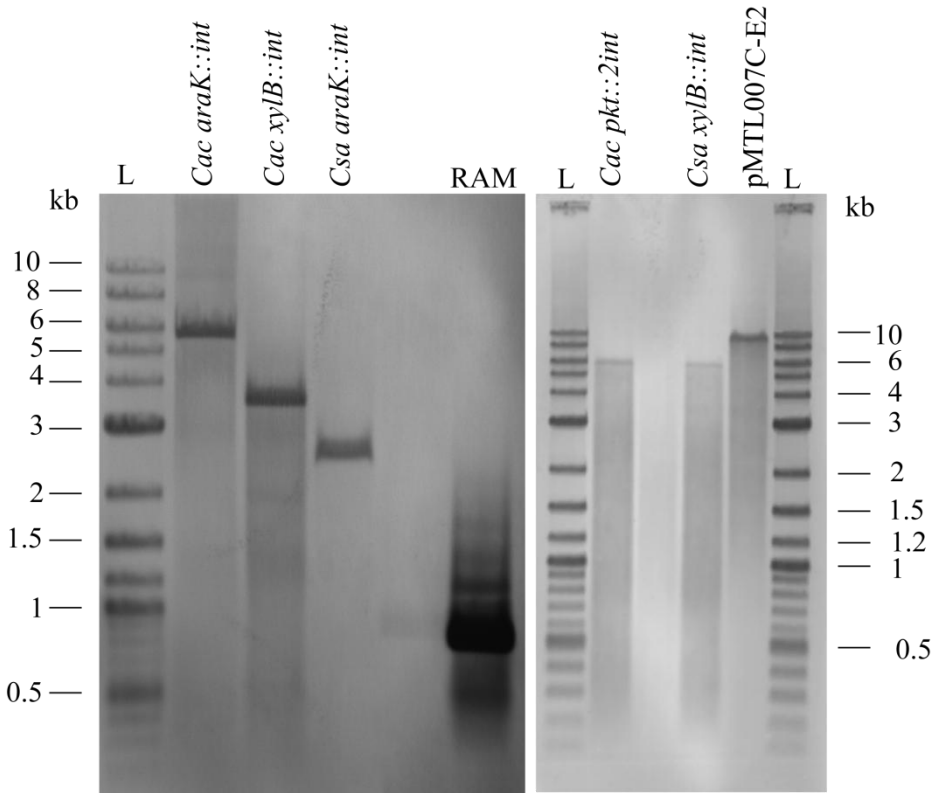



**Figure 3.3.5.** Visualization of the Southern Blot showing the single integration event of the *araK, xylB* and *pkt*-targeting introns into the *C. acetobutylicum* and *C. saccharobutylicum* genomes. Two positive controls were performed, on an amplified RAM cassette (900 bps) for small-sized DNA molecules and on pMTL007C-E2 (8.9 kb) for large-sized DNA molecules. *Cac, C. acetobutylicum; Csa, C. saccharobutylicum.* The probe was designed to target the 900 bps long RAM cassette, present in each mutant DNA and on the pMTL007C-E2 plasmid. Restriction enzymes and band sizes are described in the text.

### 3.3.2.4. Creation of pJL3 and pJL4 plasmids for the clean deletion of the *araK* and *xylB* genes of *C. acetobutylicum*

The double clean deletion mutant *C. acetobutylicum ΔaraKΔxylB* was created using the *pyrE*-deficient strain, according to the principles described in Materials and Methods (section 2.5.4). To create left and right homology arms for the *araK* gene, the LHA *araK* fwd with LHA *araK* rev and RHA *araK* fwd with RHA *araK* rev primers were used respectively (Tab. 9.4). To amplify the whole cassette the LHA *araK* fwd and RHA *araK* rev primer pair was used, and the cleaned product was cloned into the pMTL-ME3 backbone using *EcoR*I and *NcoI* restrictases (Fig. 3.3.6 A). To amplify the

left and right homology arms for the *xylB* gene, the LHA *xylB* fwd with LHA *xylB* rev and RHA *xylB* fwd with RHA *xylB* rev primer pairs were used respectively (Tab. 9.4). The cassette was created using LHA *xylB* fwd and RHA *xylB* rev and cloned into the pMTL-ME3 plasmid backbone using *Eco*RI and *XbaI* restrictases (Fig. 3.3.6 B). The flanking regions (homology arms, HA) of genes of interest (GOI) were about 0.5 kb large, including the gene START and STOP codons. He correct cassettes were PCR-checked and sequenced using pMTL-ME3 check fwd and rev primers (Tab. 9.4). First, the *C. acetobutylicum ΔxylB* mutant was created and then used as a basal strain for the *araK* gene removal to create the double mutant. The mutants were screened by PCR to isolate pure mutant culture (Fig. 3.3.2), using *Cac araK* clean deletion check and *Cac xylB* clean deletion check primer pairs.

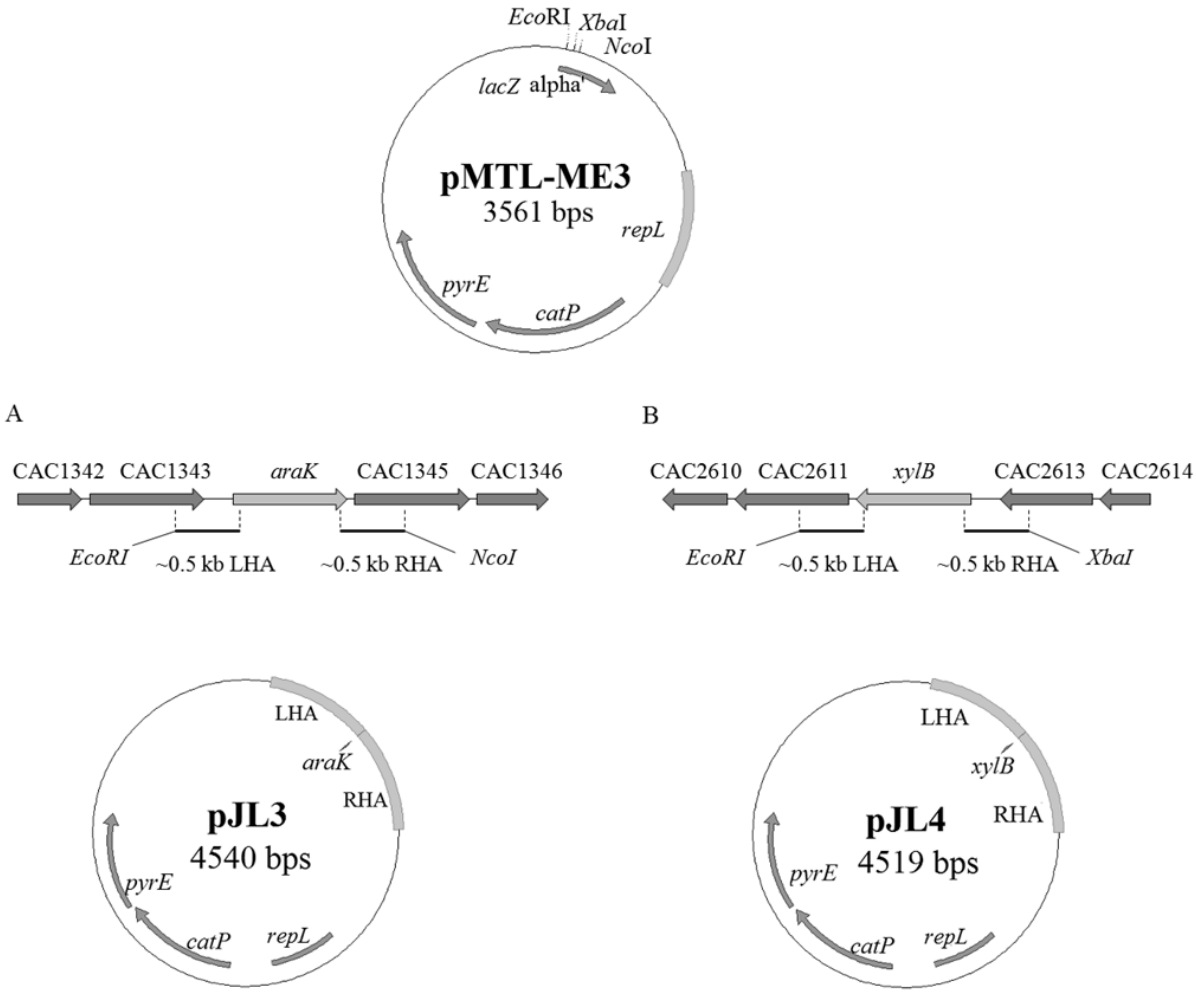



**Figure 3.3.6.** Schematic representation of the creation of (A) pJL3 and (B) pJL4 plasmids for the clean deletion of *C. acetobutylicum araK* and *xylB*, respectively. The structure of genome

regions neighboring the *araK* and *xylB* genes is shown and the homologous regions are indicated with solid black lines, along with the restriction enzymes used for cloning purposes. The backbone plasmid, pMTL-ME3, is shown above. The *repL* gene is an origin for replication for Gram-positive bacteria from pIM13 and the *pyrE* gene is a functional copy of the *pyrE* gene of *C. sporogenes.*

### 3.3.2.5. Characterization of *C. acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262 *araK* and *xylB* mutants

#### 3.3.2.5.1. *Clostridium acetobutylicum* ClosTron and clean deletion mutants

*Clostridium acetobutylicum* ATCC 824 has two operons, CAC1344 – CAC1349 and CAC2610 – CAC2612, which were found to be upregulated in the presence of xylose (Grimmler *et al.*, 2010). Previously both were annotated as xylose operons, however recent studies suggest the CAC1344 – CAC1349 group genes are involved rather in arabinose than xylose metabolism (Servinsky *et al.*, 2012 and 2010). Both operons contain kinase genes; their impact on the pentose metabolism was investigated in this work. For this purpose intron mutants in *xylB* (CAC2612) and *araK* (CAC1344) genes were created. To study the behavior of mutants in both *araK* and *xylB* genes a clean deletion double mutant was created using the *pyrE*-deficient strain (Heap *et al.*, 2012).

Hungate tests

The ability of mutant strains to grow on glucose, fructose and ribose was tested in a hungate test and compared to the results obtained in the previous hungate experiment on the wild type strain. Both kinases are not known to be involved in the utilization of these particular monosaccharides, but since these sugars are important components of the plant-derived mass, it was necessary to investigate whether their metabolism would not be affected by *araK* or *xylB* disruption. In principle, the hungate growth test gives information on the ability of a strain to grow under given conditions, i.e. a particular

sugar supplementation of the medium. In all the tested *C. acetobutylicum* strains no visible growth malfunctions were observed.

The next step was to examine the behavior of mutants on xylose. At the beginning of the project both investigated kinases were annotated as xylose metabolism-related genes, and previous study suggested the *araK* (CAC1344) gene encodes also a xylulose kinase and is under the catabolic repression of glucose (Grimmler *et al.*, 2010). Only the *C. acetobutylicum xylB::int* mutant was incapable of growing on xylose and it was suspected that the presumed glucose catabolite repression in these conditions was too tight to allow the *araK* gene expression. Hence, no growth of *xylB::int* strain on xylose-supplemented medium was observed. However, several genes of the first xylose operon (CAC1344 – CAC1349) were annotated as genes involved in arabinose metabolism; therefore there was a possibility the CAC1344 gene could not be directly involved in xylose utilization, which would result in no complementation of the effects of the *xylB* disruption by *araK*. To investigate this theory, an analogical hungate test with arabinose as the sole carbon source was done. As presumed, on the arabinose the *araK::int* mutant showed no growth, while the *xylB::int* grew undisturbed.

The *C. acetobutylicum ΔaraKΔxylB* mutant, when tested in the hungate experiment, showed no growth on either arabinose or xylose.

## Batch culture tests

Batch cultures as an experiment are much more accurate than hungate tests, and can produce quite detailed information about the strain phenotype under provided conditions. The mutant and wild type behavior was investigated in batch cultures, using glucose, xylose and arabinose as the main or sole carbon sources, in both rich and minimal media. From the rich medium cultures, samples for sugar consumption and fermentation profile analysis were collected, while the synthetic media were used mainly to precisely establish the growth profile of mutant strains. In the minimal media supplemented monosaccharide was the only carbon and energy source, which provided trustworthy and accurate information on the growth profile. For all batch cultures multiple repetitions were done and representative experiments were chosen for graphic representation.

To examine the behavior of mutants under different nutritional conditions several batch culture tests in rich and minimal mediums with glucose, arabinose and xylose as the sole carbon sources were prepared. For all the batch cultures the precultures were prepared in the medium supplemented with glucose, unless otherwise stated. Low amounts of glucose were introduced into the medium of the main cultures with inocula, which is an explanation for the initial growth of arabinose and xylose kinase mutants on arabinose and xylose respectively. The inoculation of the CGM medium with no sugar supplementation with precultures grown on CGM+glucose also resulted in a short initial growth of *C. acetobutylicum* wild type strain, which was inhibited when glucose was fully consumed (data not shown).

Prior to investigation of the role of the *araK* gene in arabinose metabolism, the *C. acetobutylicum araK::int* mutant was investigated as a xylose kinase mutant. Therefore, the growth curves of WT, *araK::int* and *xylB::int* strains, cultivated in batch cultures with different concentrations of xylose in CGM medium, were compared in order to examine if various amounts of xylose added as a sole carbon source to the medium would reveal altered phenotypes of those strains. Five pentose concentrations: 60 mM, 100 mM, 150 mM, 200 mM and 265 mM were tested and juxtaposed on one growth curve graph. Regardless of the xylose concentration, the growth curves were mostly comparable with the growth curve on standard CGM medium (333 mM xylose) during the exponential growth phase. Results are shown in figure 3.3.7.

All subsequent experiments were conducted after the discovery of the role of the *araK* gene in arabinose metabolism, and therefore the *araK::int* mutant was investigated as an arabinose kinase disruptant.

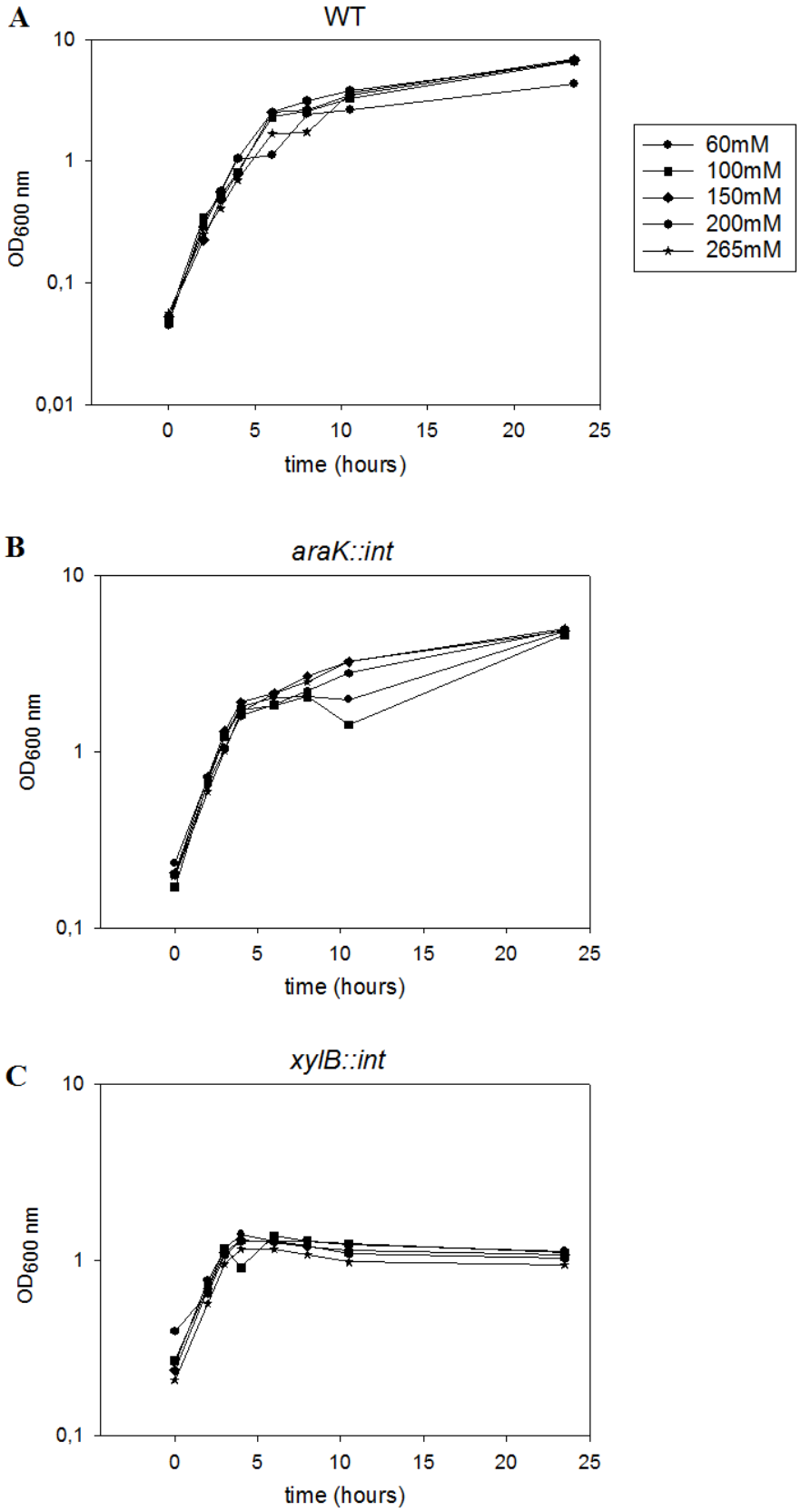



**Figure 3.3.7.** Growth curves of **(A)** *C. acetobutylicum* WT, **(B)** *araK::int* and **(C)** *xylB::int* during the first 25 hours of the exponential growth phase. The CGM medium was supplemented with different xylose concentrations: 60 mM, 100 mM, 150 mM, 200 mM and 265 mM, and the

optical density values are presented in a decimal logarithm scale. All precultures were prepared in CGM medium supplemented with 5% of glucose, hence the initial growth of the *xylB::int* mutant. The results shown come from the representative experiment. The experiments were repeated three times, giving similar results. For each optical density measurement three technical replicates were made, as described in section 2.3.

Next, the batch cultures in minimal medium (MM-MES), supplemented with 200 mM of three monosaccharides: glucose, xylose and arabinose, were prepared and the phenotypes of the wild type and mutant strains were compared. In these experiments all precultures were prepared in minimal medium supplemented with 6% glucose, hence the noticeable growth of the *araK* and *xylB* mutants at the beginning on arabinose and xylose respectively. The sugar concentrations of all the main cultures were analyzed and it was observed that when the glucose was fully consumed, the *araK::int* strain growth on arabinose was inhibited, and likewise the *xylB::int* growth on xylose.

All mutants were able to grow on glucose with only a moderate disadvantage compared to the wild type (Fig. 3.3.9 A). The *xylB::int* mutant was incapable of growing on xylose, which would confirm the gene's importance in xylose metabolism (Fig. 3.3.8 C), whereas the *araK::int* mutant was unable to grow on arabinose (Fig. 3.3.8 B).

The *C. acetobutylicum* wild type is known to grow slower and less efficiently on xylose compared to glucose or arabinose (Servinsky *et al.*, 2012; Ounine *et al.*, 1983), and similar behavior is shown by the *C. acetobutylicum araK::int* mutant (Fig. 3.3.8 B). Additionally, as on glucose, the *araK::int* mutant grew worse on xylose compared to the WT. No such relationship was observed for the *xylB::int* mutant when compared to the wild type on arabinose (Fig. 3.3.9 C). Furthermore, this mutant grew alike on glucose and arabinose. As in hungate tests, the *ΔaraKΔxylB* mutant did not grow on either arabinose or xylose (Fig. 3.3.8 D).

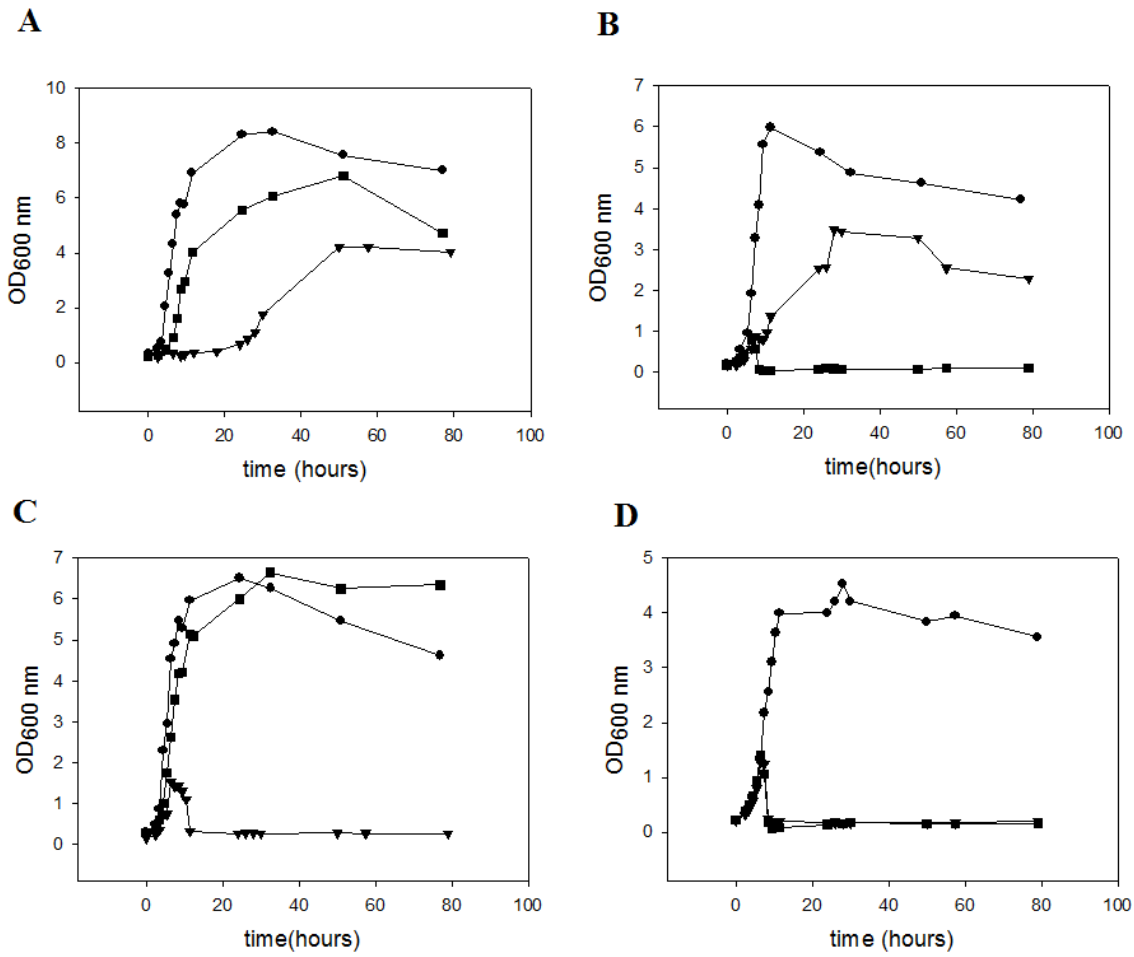



**Figure 3.3.8.** Growth curves of the *Clostridium acetobutylicum* **(A)** wild type, **(B)** *araK::int*, **(C)** *xylB::int* and **(D)** *ΔaraKΔxylB* in batch cultures of the MM-MES medium supplemented with three different monosaccharides: glucose (●), xylose (▼) or arabinose (■) as sole carbon sources. All precultures were prepared in a glucose-supplemented medium. The results shown come from the representative experiment. The experiments were repeated at least three times, giving similar results; for each optical density measurement three technical replicates were made, as described in section 2.3.

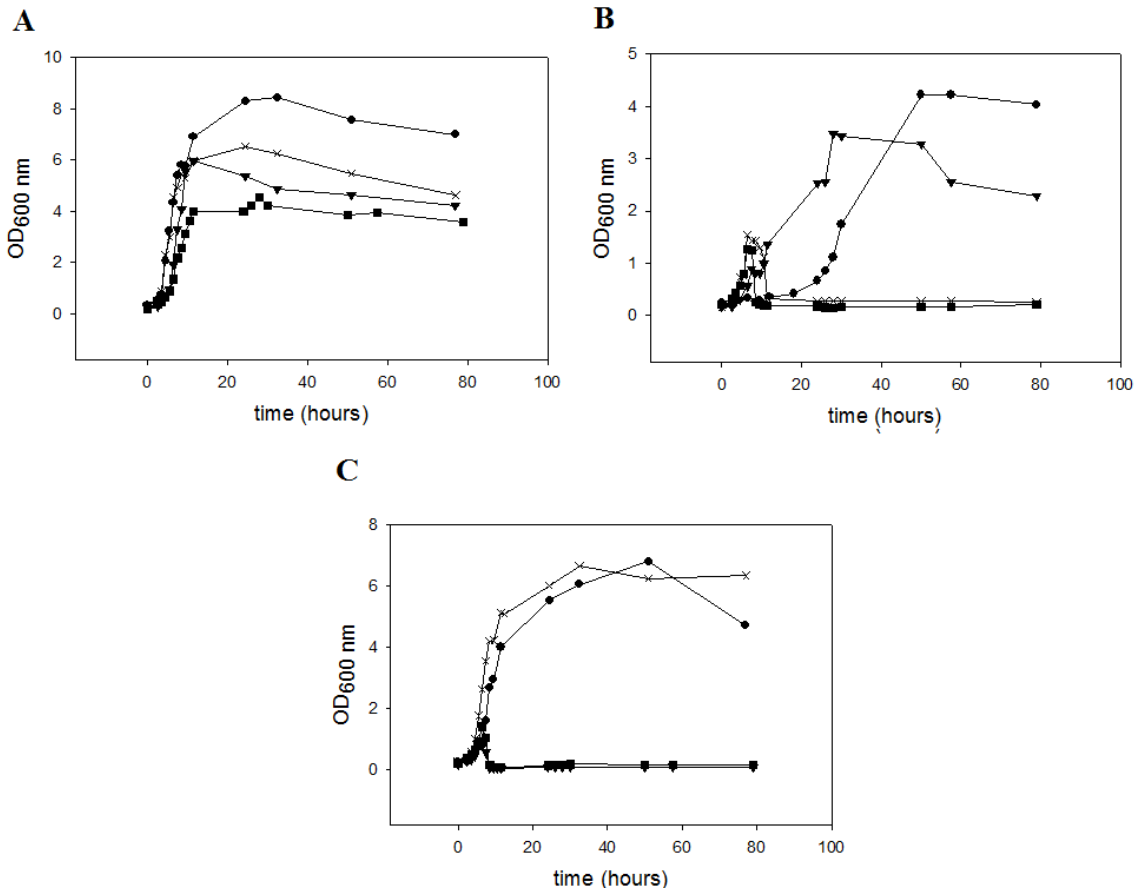



**Figure 3.3.9.** Growth curves of the *Clostridium acetobutylicum* wild type (●), *araK::int* (▼), *xylB::int* (✶) and *ΔaraKΔxylB* (■) in batch cultures of the MM-MES medium supplemented with three different monosaccharides: **(A)** glucose, **(B)** xylose or **(C)** arabinose as sole carbon sources. All precultures were prepared in a glucose-supplemented medium. The results shown come from the representative experiment. The experiments were repeated at least three times, giving similar results; for each optical density measurement three technical replicates were made, as described in section 2.3.

### 3.3.2.5.2. *Clostridium saccharobutylicum araK* and *xylB* ClosTron mutants

Hungate tests

The *C. saccharobutylicum* wild type and ClosTron mutant strains were tested in the same conditions as described in the section devoted to *C. acetobutylicum*. For all strains comparable growth curves in glucose, ribose and fructose-supplemented CGM medium were observed. As suspected, in hungates containing medium supplemented with xylose no growth of *C. saccharobutylicum xylB::int* was observed, and in the

arabinose-supplemented medium the *araK::int* mutant showed no growth, suggesting comparable gene functions in both the *C. acetobutylicum* and *C. saccharobutylicum* species.

Batch culture tests

The batch culture experiments were used to establish more detailed phenotypes of the *C. saccharobutylicum* mutant strains. In addition to the hungate tests, it was necessary to investigate all the strains more closely, and to describe similarities and differences between the *C. saccharobutylicum* and *C. acetobutylicum* mutants.

As in *C. acetobutylicum*, the wild type strain showed better growth on glucose in comparison to both ClosTron mutants (Fig. 3.3.11 A). The same observations were made for the two other sugars, xylose (Fig. 3.3.11 B) and arabinose (Fig. 3.3.11 C). In contrast to what is known for *C. acetobutylicum* (Fig. 3.3.8 A), no impaired growth of the *C. saccharobutylicum* wild type was observed on xylose when compared to arabinose (Fig. 3.3.10 A). The *C. saccharobutylicum xylB::int* mutant growth was inhibited in batch culture when xylose was the only carbon source (Fig. 3.3.10 B) as soon as the inoculum-derived glucose was fully consumed (data not shown), and analogical behavior was observed for the *araK::int* in arabinose-supplemented medium (Fig. 3.3.10 C). Both mutant strains grew as long as they had glucose in the medium and stopped growing when the hexose was completely consumed, similar to what has been observed for the *C. acetobutylicum* mutants.

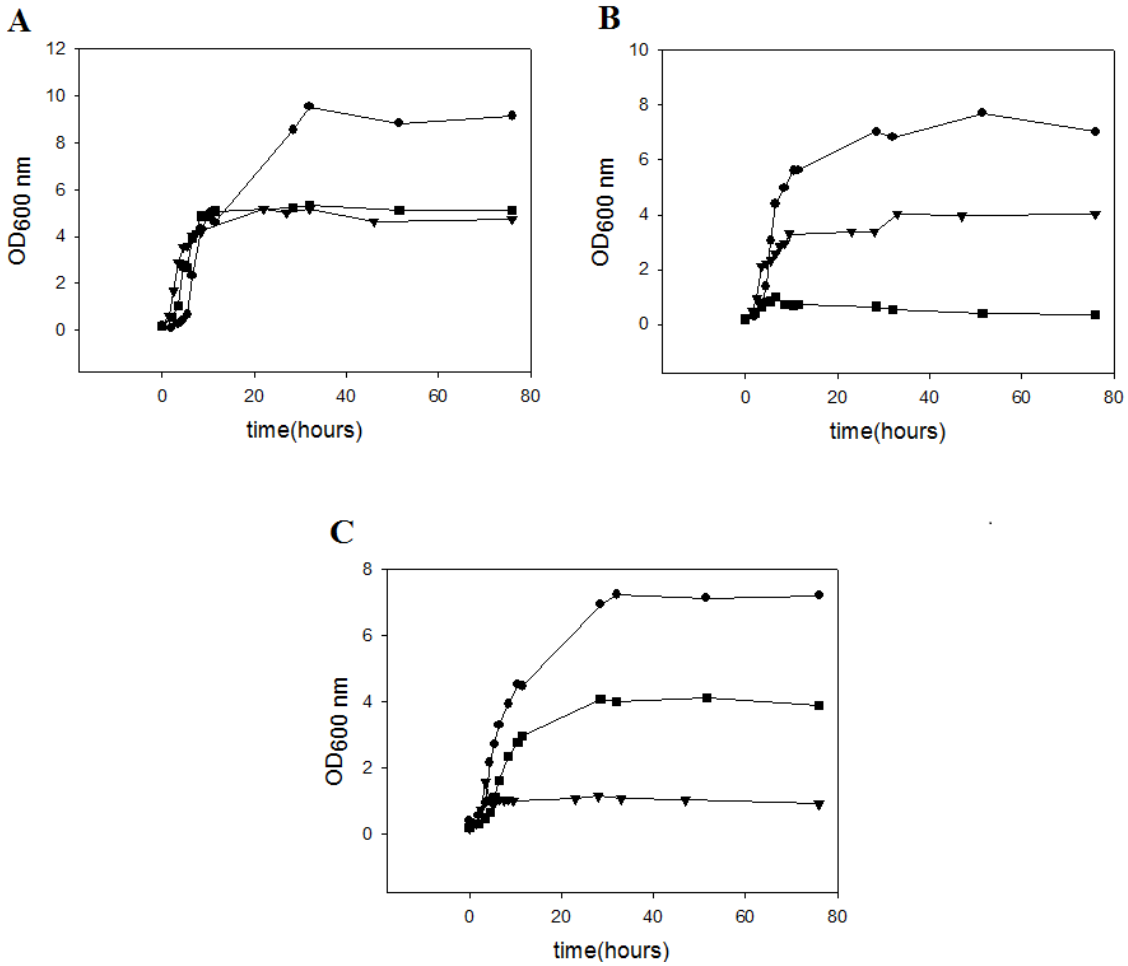



**Figure 3.3.10.** Growth curves of the *Clostridium saccharobutylicum* **(A)** wild type, **(B)** *araK::int* and **(C)** *xylB::int* in batch cultures of the MM-MES medium supplemented with three different monosaccharides: glucose (●), xylose (▼) or arabinose (■) as sole carbon sources. All precultures were prepared in glucose-supplemented medium. The results shown come from the representative experiment. The experiments were repeated at least three times, giving similar results; for each optical density measurement three technical replicates were made, as described in section 2.3.

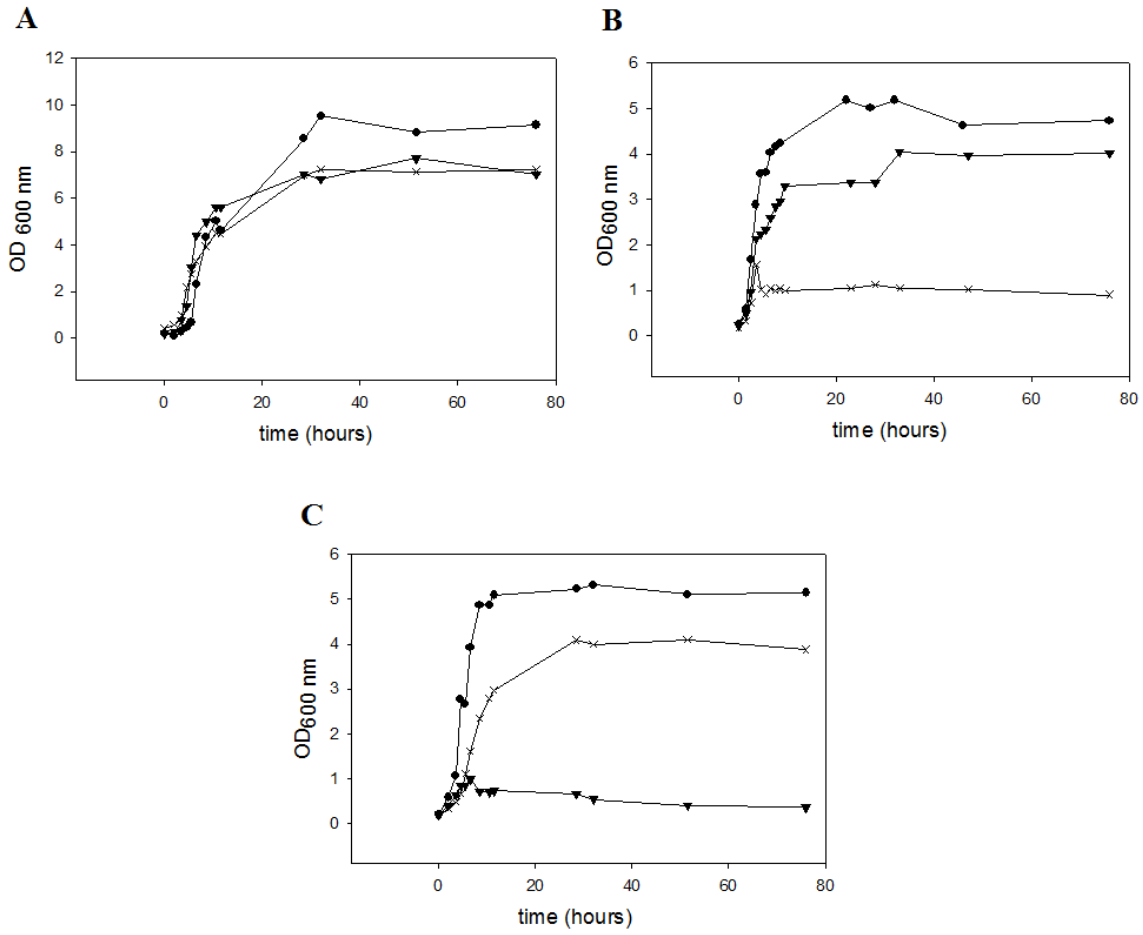



**Figure 3.3.11.** Growth curves of the *Clostridium saccharobutylicum* wild type (●), *araK::int* (▼) and *xylB::int* (✶) in batch cultures of the MM-MES medium supplemented with three different monosaccharides: **(A)** glucose, **(B)** xylose or **(C)** arabinose as sole carbon sources. All precultures were prepared in glucose-supplemented medium. The results shown come from the representative experiment. The experiments were repeated at least three times, giving similar results; for each optical density measurement three technical replicates were made, as described in section 2.3.

### 3.3.2.5.3. Fermentation profile of the *C. acetobutylicum xylB::int* and *araK::int* strains

To investigate the fermentation profile of the *C. acetobutylicum* mutant and wild type strains on the three different monosaccharides: glucose, xylose and arabinose, 200 ml batch cultures in standard rich medium (CGM), supplemented with one of the sugars as a sole carbon source, were prepared. In the conventional recipe one liter of the

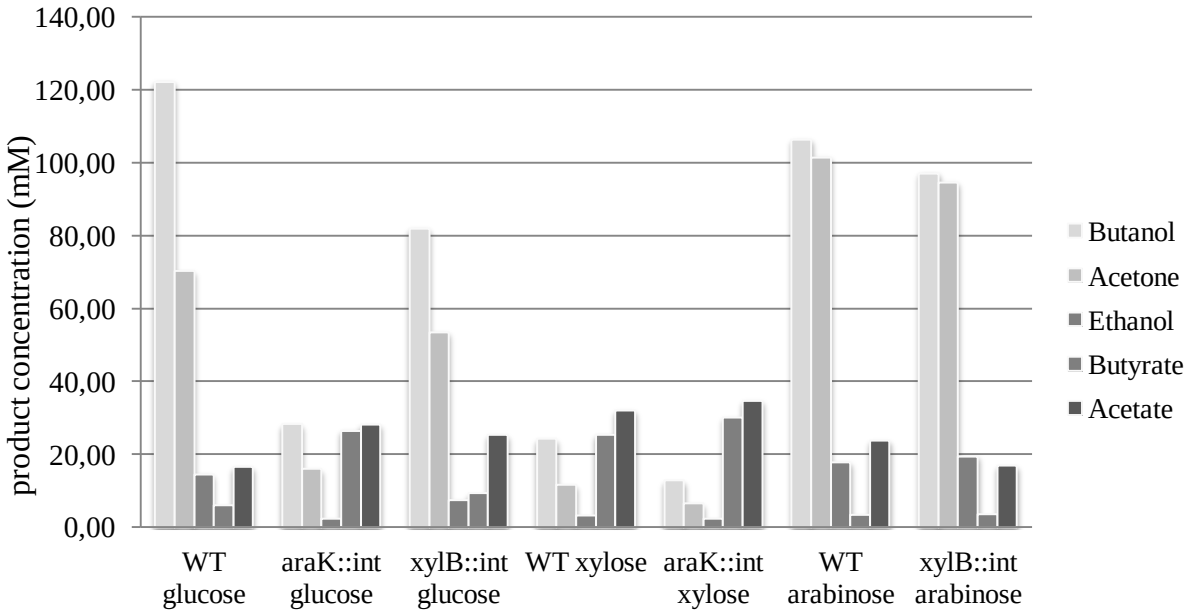
medium is supplemented with 50 g of a monosaccharide, giving 333 mM of xylose or arabinose and 278 mM of glucose.

The amount of butanol produced by the *C. acetobutylicum xylB::int* on glucose (81.98 mM) was lower than that produced by the wild type (122.21 mM) and was comparable to the butanol production of WT on arabinose (106.35 mM). The *AraK::int* strain on glucose produced a significantly lower amount of butanol (28.33 mM), like the wild type strain on xylose (24.31 mM).The *AraK::int* strain showed a less efficient butanol production profile, regardless of the sugar used as a carbon source (12.89 mM butanol on xylose), whereas the *xylB::int* mutant was capable of fermenting both glucose and arabinose nearly as effective as the wild type strain (Fig. 3.3.12 A and B). In xylose-supplemented media low solvent and high acid concentrations at the end of the experiment were detected, which would suggest a less efficient acid uptake and lower solvent production rate. Among all the strains investigated during this experiment the *araK::int* mutant produced the lowest amounts of solvents, while the *xylB::int* mutant was able to ferment glucose and arabinose quite effectively, compared to the WT strain. The acetone and ethanol production of the wild type strain on arabinose was higher (101.50 mM and 17.79 mM, respectively) than on glucose (70.31 mM and 14.50 mM). The same fermentation profile was shown by the *xylB::int* mutant, with 53.58 mM acetone on glucose and 94.60 mM on arabinose, and 7.47 mM of ethanol on glucose and 19.45 mM on arabinose.

*C. acetobutylicum* shows altered sugar metabolism patterns on glucose or xylose and arabinose, which could be indicated by the analysis of acetate and butyrate concentrations at the end of the fermentation experiment and calculation of the acetate:butyrate ratios (Servinsky *et al.*, 2012). The *C. acetobutylicum* wild type produced 16.62 mM, 32.11 mM and 23.90 mM of acetate with 6.08 mM, 25.41 mM and 3.44 mM of butyrate on glucose, xylose and arabinose, respectively. In samples of the *C. saccharobutylicum xylB::int* mutant, 9.33 mM of butyrate and 25.36 mM of acetate on glucose, with 3.55 mM of butyrate and 16.99 mM of acetate on arabinose, were detected. Measurements of acid production by the ClosTron mutant in the a*raK* gene showed 26.51 mM butyrate and 28.26 mM acetate on glucose, and 30.09 mM of butyrate and 34.71 mM of acetate on xylose, in the samples collected at the end of the fermentation experiment. These results show different a acetate:butyrate ratio on arabinose and glucose or xylose (Fig. 3.3.12 B), which will be further discussed in the

Discussion section. A table with the exact values for acid and solvent production is attached to the Supplementary Data section (Tab. 9.8).

**A**




**B**

| | Glucose | | | Xylose | | Arabinose | |
|---|---|---|---|---|---|---|---|
| | WT | *araK::int* | *xylB::int* | WT | *araK::int* | WT | *xylB::int* |
| Total sugar consumed (S), mM | 231.64 | 192.64 | 223.19 | 104.25 | 97.41 | 310.57 | 234.61 |
| Total solvent and acid production (SA), mM | 229.72 | 101.42 | 177.71 | 96.78 | 86.73 | 252.98 | 231.77 |
| S:SA ratio (%) | 99.17 | 52.65 | 79.62 | 92.83 | 89.03 | 81.45 | 98.79 |
| Butanol (%) | 53.20 | 27.94 | 46.13 | 25.11 | 14.86 | 42.02 | 41.93 |
| Acetone (%) | 30.61 | 15.81 | 30.15 | 12.13 | 7.66 | 40.12 | 40.82 |
| Ethanol (%) | 6.31 | 2.26 | 4.2 | 3.33 | 2.75 | 7.03 | 8.39 |
| Ac:Bt ratio | 2.73 | 1.07 | 2.71 | 1.26 | 1.15 | 6.95 | 4.79 |

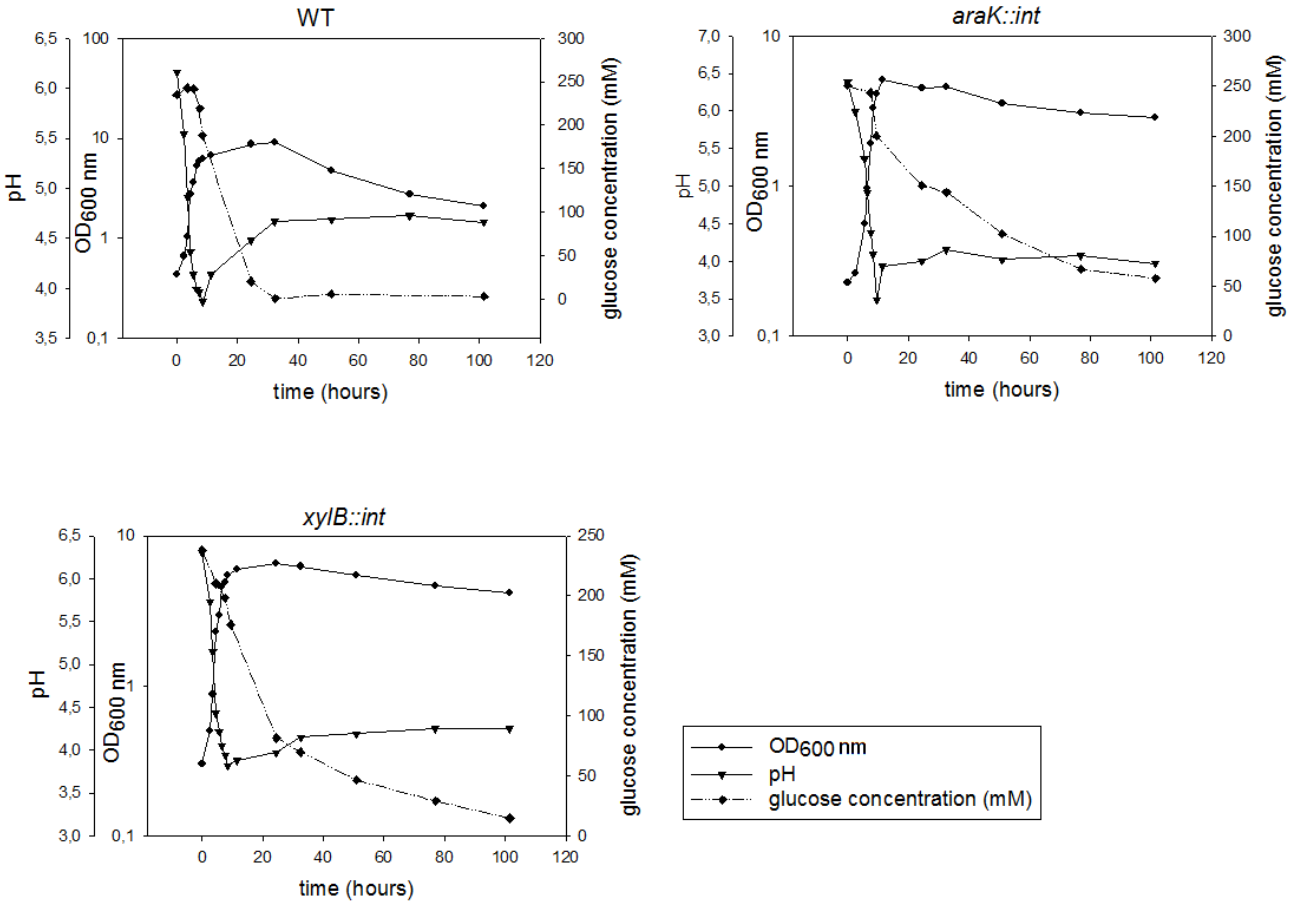
**Figure 3.3.12. (A)** Fermentation profile of the *C. acetobutylicum* wild type, *araK::int* and *xylB::int* mutant strains in CGM medium supplemented with 5% of monosaccharide; glucose (278 mM), xylose or arabinose (333 mM), measured by GC, as described in section 2.8.1. Samples were taken during solventogenic growth, 101 hours after inoculation. All precultures were prepared in

rich medium supplemented with the same sugar as the main culture, and results of the representative experiment are showed. **(B)** Total sugar consumption (S) compared to total solvent and acid production (SA) of the *C. acetobutylicum* WT, *araK::int* and *xylB::int* on glucose, xylose and arabinose. Amounts of butanol, acetone and ethanol produced are given as a percentage of the total solvent and acid production values (SA). Additionally, the acetate:butyrate ratios (Ac:Bt) were calculated for all strains on glucose, xylose and arabinose. The results shown come from the representative experiment. The experiments were repeated three times, giving similar results.

Additionally to the batch culture and fermentation experiments, the sugar consumption profiles of the wild type and mutant strains were analyzed. To this end samples from batch cultures prepared for the fermentation experiments were used, and sugar concentrations were calculated using appropriate kits for D-glucose, D-xylose or L-arabinose detection (Megazyme, Ireland). Glucose and arabinose were relatively well consumed, and only traces remained in the medium at the end of the experiment (Fig. 3.3.13 A and C). Xylose, on the contrary, was poorly consumed (Fig. 3.3.13 B). In all mutant strains the uptake of sugars was worse than in the wild type strain. In order to determine the effectiveness of the wild type and mutant strains in fermentation of each of the sugars during the experiment, the ratio of acid and solvent production (SA) to total sugar consumption (S) was calculated (Fig. 3.3.12 B) and shown as percentage values. For all the strains the ratio was comparable (around 80 – 99%), except for the *araK::int* mutant on xylose (52.65%). Butanol, acetone and ethanol amounts were given as a percentage of the total solvent and acid production (SA) value.

Differences in butanol, ethanol and acetone production between the cultures on glucose or xylose and arabinose were most probably the result of different metabolic pathways and will be discussed further.

**A**


WT
araK::int
xylB::int
pH
OD600 nm
glucose concentration (mM)
time (hours)
OD600 nm
pH
glucose concentration (mM)

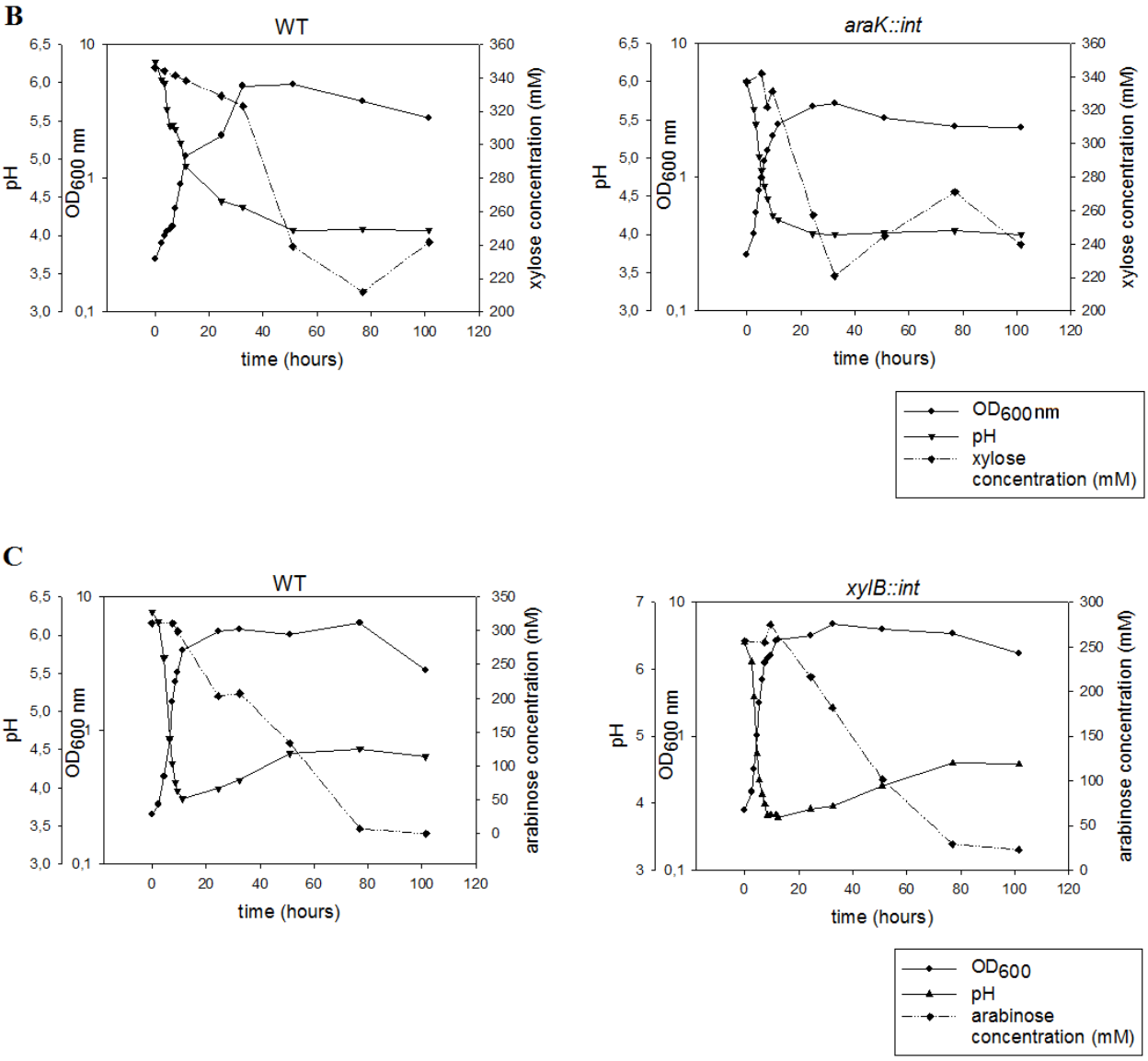



**Figure 3.3.13.** Graphic representation of the growth rate ($OD_{600}$ ●), pH values (▼) and sugar consumption (◆) of the *C. acetobutylicum* WT, *araK::int* and *xylB::int* strains on rich media supplemented with **(A)** glucose, **(B)** xylose and **(C)** arabinose. $OD_{600}$ values are presented on a logarithmic scale, and detailed information is provided in the legends. The results shown come from the representative experiment. The experiments were repeated at least three times, giving similar results. Three technical replicates were made for each optical density and pH measurement, as described in section 2.3, and for each sugar concentration measurements, as described in section 2.8.2.

### 3.3.2.5.4. Promoter mapping of the *araK* and *xylB* genes of *C. acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262

During the microarray analysis certain genes were observed to be expressed simultaneously under specific conditions (Grimmler *et al.*, 2010). Three out of the four investigated xylulose genes were believed to be the first genes of their operons. Promoters were searched for using the RLM-RACE method and their regions were detected upstream of the *C. acetobutylicum xylB*, *C. acetobutylicum araK* and *C. saccharobutylicum xylB* genes (Fig. 3.3.14). Prior to the RLM-RACE experiment, the BPROM from Soft Berry (www.softberry.com), an algorithm designed to find bacterial promoters, was used to map the presumed promoter regions *in silico*.

Expected sequences of -35 and -10 box, along with the number of bases between the promoter sequence and the START codon of the gene are given in figure 3.314. In *C. acetobutylicum* both kinases are the first genes of xylose and arabinose operons and in *C. saccharobutylicum* the kinase gene opens the xylose operon. The primers used to map the promoters in the RLM-RACE experiment are listed in table 9.7.

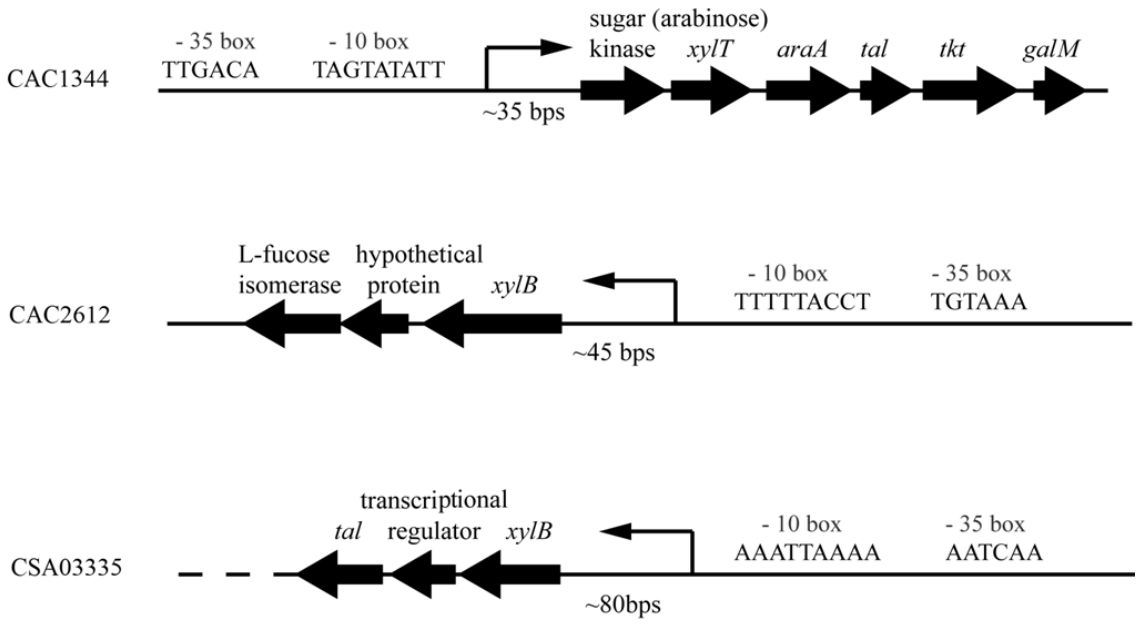



**Figure 3.3.14.** Graphic representation of the promoter regions of the *C. acetobutylicum* arabinose and xylose operons and the *C. saccharobutylicum* xylose operon. Predicted -10box

and -35box sequences are shown along with the distance from the promoter to the START codon of the first gene of the operon. Gene numbers of the first gene of the operons are given on the left (NCBI for *C. acetobutylicum* and JGI for *C. saccharobutylicum*). A dotted line indicates more genes belonging to the operon and a solid line indicates the whole operon is shown.

### 3.3.2.5.5. Complementation of *araK* and *xylB* gene disruption in mutants of *C. acetobutylicum* and *C. saccharobutylicum*

To test whether it is possible to reverse the effect of the intron disruption of *araK* and *xylB* kinases, complementation plasmids bearing functional copies of disrupted genes were proposed.

Two sets of complementation plasmids were designed – with wild type copies of the genes under clostridial strong and constitutive ferredoxin promoter ($p_{fdx}$) or with an upstream region of around 300 bps, containing their own promoter sequences (Fig. 3.3.15), mapped previously (Fig. 3.3.14). Since the *C. saccharobutylicum araK* gene is located in the middle of the operon, the complementation plasmid for the *C. saccharobutylicum araK::int* mutant was designed only with ferredoxin promoter. In total, seven plasmids were designed to complement the *C. acetobutylicum araK::int*, *C. acetobutylicum xylB::int*, *C. saccharobutylicum araK::int* and *C. saccharobutylicum xylB::int* mutants. The complementation experiment would reveal whether it is possible to reverse the phenotype of the *araK::int* and *xylB::int* mutants, assuming no polar effects were caused by the intron integration events.

This part of the work was done by Philipp Högel as a project for his B.Sc. thesis.

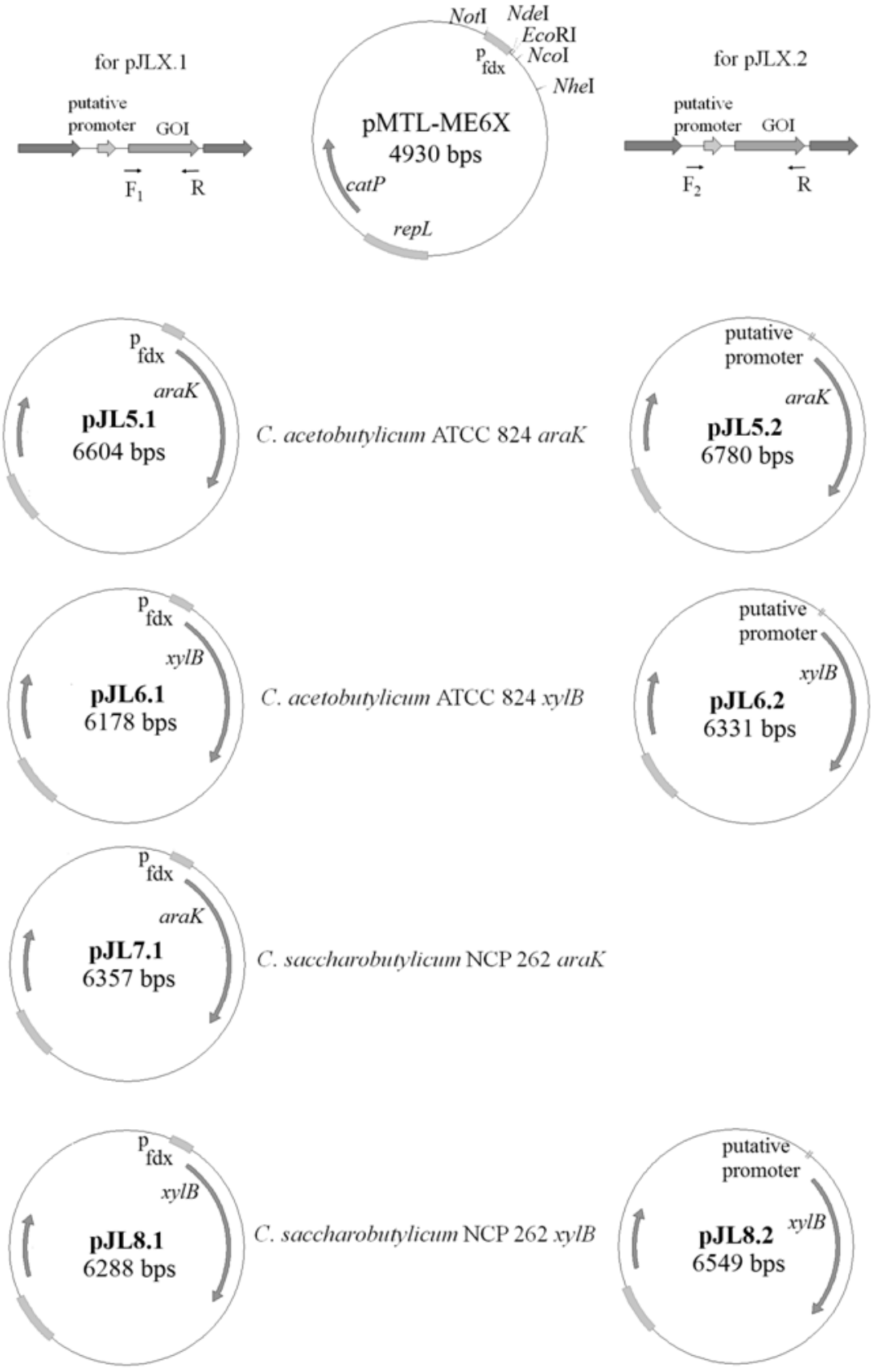



**Figure 3.3.15.** Schematic representation of the construction of pJL5, pJL6, pJL7 and pJL8 plasmids, designed to complement mutations in *C. acetobutylicum araK*, *xylB* and

*C. saccharobutylicum araK* and *xylB* genes, respectively. Plasmids with $p_{fdx}$ promoter are presented on the left and plasmids with native promoters on the right, along with graphic presentation of the cassette construction and the plasmid used as a backbone, pMTL-ME6X, at the top. The creation of exemplary complementation cassettes is shown above, for the pJLX.1 plasmids on the left, and pJLX.2 plasmids on the right. *RepL*, origin of replication for Gram-positive bacteria.

Plasmids with $p_{fdx}$ promoter: pJL5.1, pJL6.1, pJL7.1 and pJL8.1

The CAC1344 and CAC2612 genes were amplified using pJL5.1 fwd with pJL5 rev and pJL6.1 fwd with pJL6 rev primers, respectively, to obtain cassettes containing the gene regions only (Fig. 3.3.15). The CSA00774 and CSA03335 genes were amplified using pJL7.1 fwd with pJL7 rev and pJL8.1 fwd with pJL8 revprimers, respectively. The amplified cassettes were cloned into the pMTL-ME6X plasmid backbone using an *NdeI, NcoI* enzyme pair to create a pJL5.1 and *EcoRI, NheI* enzyme pair for pJL6.1. For the creation of both pJL7.1 and pJL8.1 *NdeI* and *NheI* enzymes were used. PMTL-ME6X (Tab. 2.2) bears a strong, constitutive clostridial promoter ($p_{fdx}$,) upstream of the polylinker (multiple cloning site; Fig. 3.3.15). Correct clones were isolated and sequenced using ME6X check fwd and rev primers. All the primers used in this study are listed in table 9.6.

Plasmids with native promoters – pJL5.2, pJL6.2 and pJL8.2

CAC1344 and CAC2612 genes were amplified using pJL5.2 fwd with pJL5 rev and pJL6.2 fwd with pJL6 rev primer pairs, respectively, to obtain cassettes containing the gene sequences with their predicted promoter regions. The CSA03335 gene with its promoter region was amplified using the pJL8.2 fwd and pJL8 rev pair. Cleaned PCR products were cloned into the pMTL-ME6X plasmid backbone using the *NotI, NcoI* enzyme pair for pJL5.2 and the *NotI, NheI* enzyme pair for pJL6.2 and pJL8.2 plasmids. All the primers used in this study are listed in table 9.6.

Correct complementation plasmids were introduced into the kinase mutants of either species, either by electroporation (*C. acetobutylicum*) or conjugation (*C. saccharobutylicum*). Colonies of the *C. acetobutylicum araK::int* – pJL5.1,

*C. acetobutylicum xylB::int* – pJL6.1, *C. saccharobutylicum araK::int* – pJL7.1 and *C. saccharobutylicum xylB::int* – pJL8.1 strains were obtained and their phenotypes were tested in the batch culture experiments. Since the *araK* and *xylB* genes were located under strong ferredoxin promoter, their expression should be uninterrupted and constitutive, sufficient to reverse the kinase disruption effect. To test this, 100 ml batch cultures in CGM medium were prepared, and the medium was supplemented with arabinose for the *C. acetobutylicum araK::int* and *C. saccharobutylicum araK::int* mutant strain or xylose for *C. acetobutylicum xylB::int* and *C. saccharobutylicum xylB::int* mutant strains.

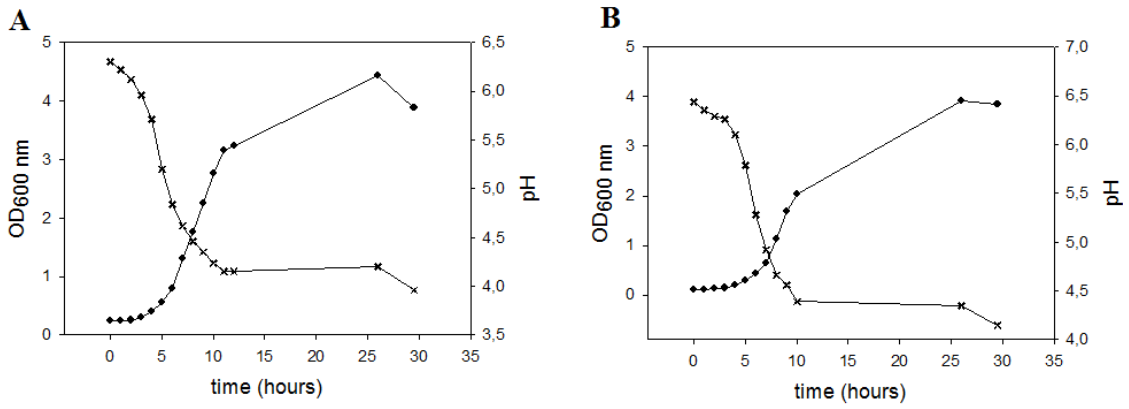



**Figure 3.3.16.** Growth curves of ClosTron mutants of *C. acetobutylicum*, the **(A)** *araK::int* and **(B)** *xylB::int* strains bearing the complementation plasmids. *C. acetobutylicum araK::int* was complemented with pJL5.1 and *xylB::int* with pJL6.1 plasmids. Both mutants showed active growth on sugar substrates. They were not able to ferment prior to the complementation due to the disruption of the kinases. The results shown come from the representative experiment. The experiments were repeated two times, giving similar results; for each optical density and pH measurement three technical replicates were made, as described in section 2.3.

The phenotypes of the *C. acetobutylicum araK::int* and of *xylB::int* mutants were reversed by functional *araK* and *xylB* genes introduced on plasmids, respectively, and in batch cultures the growth of *araK::int* on arabinose and of *xylB::int* on xylose-supplemented medium was observed. Both mutants were able to use either of the two pentoses as a sole carbon and energy source again (Fig. 3.3.16). However, the reversal of the mutant phenotypes was not observed for the *C. saccharobutylicum* ClosTron

mutants, and the reason behind this is not yet clear. Plasmids with putative promoters (pJLX.2) were not created during the work in the laboratory due to cloning problems.

### 3.3.3. Diauxic growth of *C. acetobutylicum* ATCC 824 on D-glucose and L-arabinose mixture

The diauxic growth profile on a mixture of sugars is caused by a phenomenon called carbon catabolite repression, and most often the glucose presence inhibits the expression of genes involved in the metabolism of other sugars (Servinsky *et al.*, 2010; Tangney *et al.*, 2003; Brückner *et al.*, 2002; Saier *et al.*, 1996). This phenomenon is present in *C. acetobutylicum* ATCC 824 on a xylose-glucose mixture (Grimmler *et al.*, 2010) and there was still some doubt whether similar dependence exists for the arabinose-glucose combination. Therefore, a batch culture experiment in a minimal MES medium with a mixture of 50 mM glucose and 150 mM arabinose as carbon sources was prepared and the growth and concentrations of both sugars were monitored every hour. At the beginning of the growth, in the early exponential growth phase, no changes in the arabinose concentrations were observed, while the glucose concentration in the medium quickly decreased.

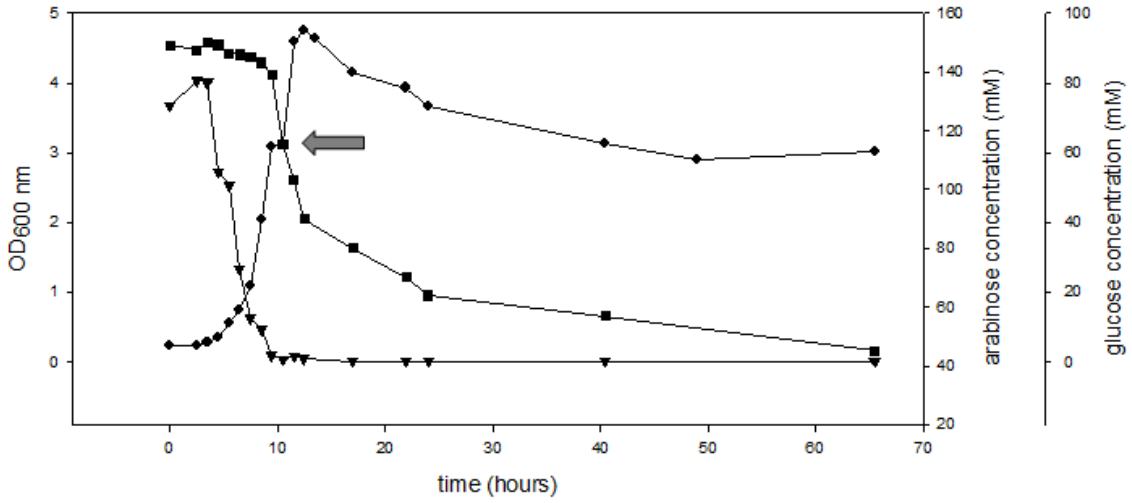



**Figure 3.3.17.** Diauxic growth curve of the *C. acetobutylicum* wild type strain in minimal MES medium, supplemented with 50 mM glucose and 150 mM arabinose. The black line with

diamonds indicates the growth curve ($OD_{600\ nm}$, right vertical axis). Sugar concentrations are shown as the line with inverted triangles for glucose and the line with squares for arabinose, both in mM (left vertical axis). The arrow indicates the presumable point of the switch between glucose and arabinose metabolism, with a short plateau phase. The results shown come from the representative experiment. The experiments were repeated three times, giving similar results; for each optical density and sugar concentration measurement three technical replicates were made, as described in section 2.3 and 2.8.2, respectively.

A short growth inhibition was observed when the glucose was nearly completely consumed, after which changes in the arabinose concentrations were detected (Fig. 3.3.17). Arabinose concentrations decreased through the late exponential and the stationary phase of clostridial growth until the end of the experiment, 65 hours after the inoculation time. Precultures were prepared in MM-MES medium with 6% glucose.

### 3.3.4. Investigation of the phosphoketolase role in the arabinose metabolism of *Clostridium acetobutylicum* ATCC 824

#### 3.3.4.1. Creation of the ClosTron mutant in phosphoketolase gene

The phosphoketolase pathway is associated mainly with *Bifidobacteria* and heterofermentative lactic acid bacteria, but recent studies have shown that it is also used in some *Clostridium* species as a part of the arabinose metabolism pathway (Servinsky *et al.*, 2012). Phosphoketolase (Pkt) was described as a bifunctional enzyme, capable of converting xylulose-5-P (X5P) and inorganic phosphate into glyceraldehyde-3-P (G3P) and acetyl-P or fructose-6-P (F6P) and inorganic phosphate into erythrose-4-P and acetyl-P (Servinsky *et al.*, 2012). Out of the three solventogenic strains investigated in this study only *C. acetobutylicum* possesses the phosphoketolase gene. The *pkt* homolog has not been found in *C. beijerinckii* or *C. saccharobutylicum*, based on their genome analysis and the results of the BLAST and DELTA-BLAST experiments, with the genome of *C. acetobutylicum* used as a reference sequence.

To test the impact and importance of phosphoketolase for arabinose metabolism two ClosTron mutants with the intron insertion in the front (*C. acetobutylicum pkt::2int*) and

rear part of the gene (*C. acetobutylicum pkt::1int*) were created, using the primers listed in table 9.3. First, the *pkt::1int* mutant was created, but since no major differences in its growth on arabinose were observed compared to the wild type, the second mutant, *pkt::2int*, was constructed (Fig. 3.3.18). Both mutants were checked with the PCR method using the *pkt* check fwd and rev primers (Tab. 9.5) and sequenced, and a single integration event in the *pkt::2int* strain was confirmed with the Southern Blot (Fig. 3.3.5). The *Pkt::2int* genomic DNA was digested with *HindIII* restrictase and the visible DNA fragment size is of 6.5 kb.

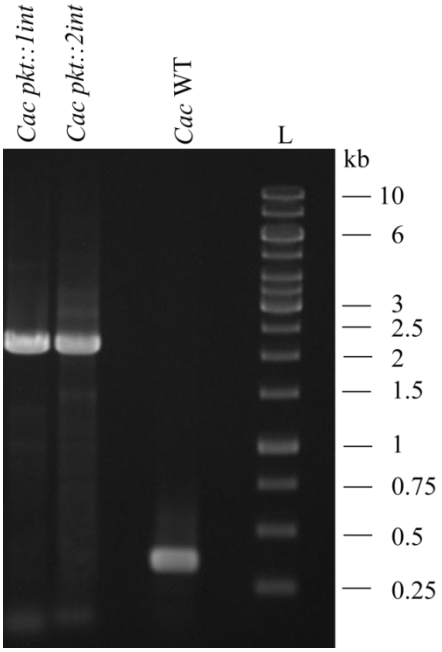



**Figure 3.3.18.** Integration event of the *pkt*-targeting intron into the *C. acetobutylicum* genome. For *pkt::1int* the *pkt* check 1 primer pair was used, and for the *pkt::2int* mutant the *pkt* check 2 primer pair was used, both producing a band of a size of around 2.45 kb. The PCR on the wild type genomic DNA was done using a *pkt* check 1 primer pair and produced a band of 352 bps. L, 1kb DNA ladder; WT, wild type DNA template.

### 3.3.4.2. Analysis of the arabinose fermentation profile of the phosphoketolase mutants

Phosphoketolase is known to play a great part in the pentose fermentation in several organisms, by altering the carbon flow in pentose metabolism pathways (Servinsky *et al.*, 2012; Papini *et al.*, 2012 Xiao *et al.*, 2011; Ohara *et al.*, 2006; Bustos *et al.*, 2005). To test whether and how disruption of the phosphoketolase gene would affect the arabinose fermentation in *C. acetobutylicum*, the growth profile of both mutants in a medium supplemented with 5% glucose or arabinose was investigated and compared with the wild type. Mutants grew similarly in the glucose-supplemented medium, yet not as efficiently as the wild type strain (Fig. 3.3.19 A). However, *C. acetobutylicum pkt::1int* and *pkt::2int* showed significant differences during their exponential growth phase on arabinose. The duplication time of both mutant strains was much higher than the duplication time of the wild type and the *pkt::2int* mutant showed impaired growth when compared to the *C. acetobutylicum* WT and *pkt::1int* strains, even though at the end of the experiment all cultures reached similar $OD_{600}$ values (Fig. 3.3.19 A). This would suggest the disruption of the *pkt* gene resulted in an alteration of the monosaccharide metabolism, most probably from the phosphoketolase pathway to the pentose phosphate pathway. It was also proved that the *pkt* gene is not crucial for arabinose metabolism, and *C. acetobutylicum* is capable of fermenting this pentose in the absence of functional phosphoketolase. To investigate whether the WT on xylose and *pkt::2int* on arabinose would present similar growth profiles, an experiment in MM-MES medium supplemented with 5% xylose for the WT and 5% arabinose for *pkt::2int* and *pkt::1int*, additionally, was prepared. Rather significant differences between the growth curves of the mutant and wild type strains were observed in favor of those strains grown on arabinose (Fig. 3.3.19 B).

*The Pkt* gene was believed to be expressed with two other genes – ribulose-5-phosphate -4-epimerase (*araD*) and L-arabinose isomerase (*araA*). In this study an approach to find the promoter sequence using the RLM-RACE method was made and a promoter region upstream of the *araD* was detected (Fig. 3.3.19 D). Both -35 and -10 box sequences were predicted using the BPROM algorithm (SoftBerry) *in silico* and are shown along with the number of nucleotides between the promoter region and the

START sequence in the graphic (Fig. 3.3.19 D). The primers used to map the promoter in the RLM-RACE experiment are listed in table 9.7.

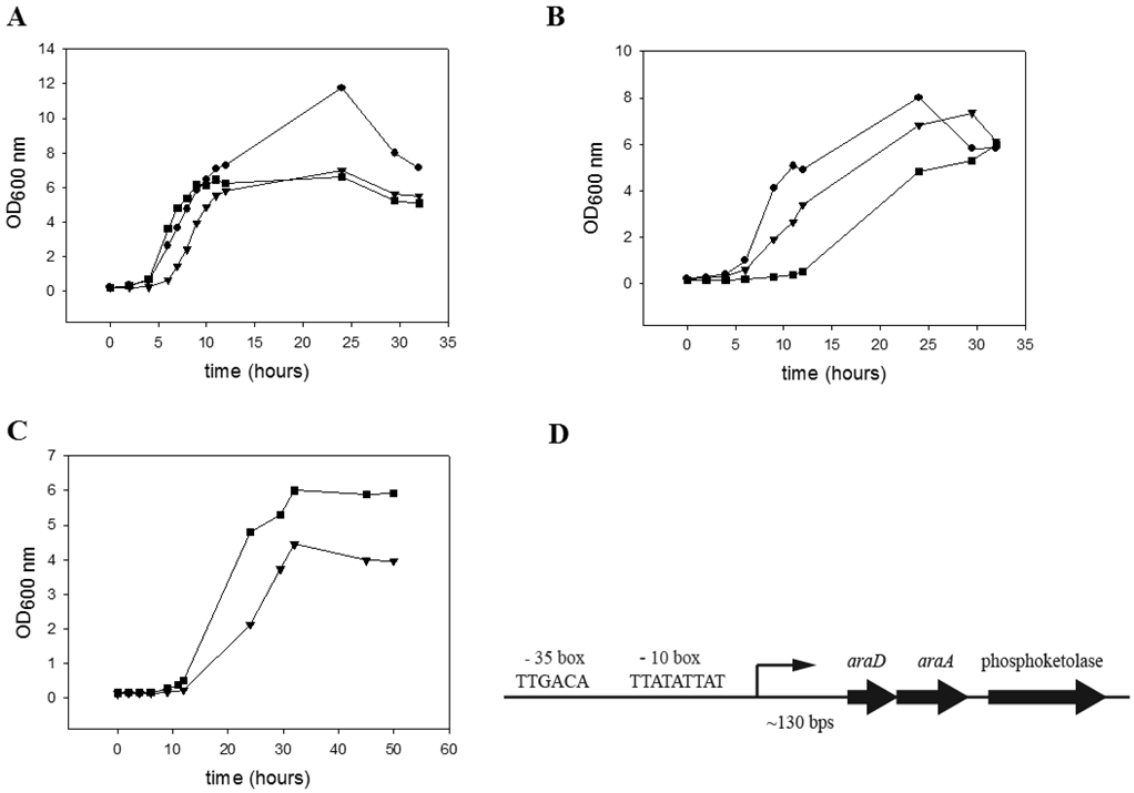



**Figure 3.3.19.** Growth curves showing the exponential growth phase of the *C. acetobutylicum* wild type ( ●), *pkt::1int* (▼) and *pkt::2int* (■) on **(A)** MM-MES medium supplemented with glucose and **(B)** and MM-MES medium supplemented with arabinose only. **(C)** Comparison of the growth curves of the *pkt::2int* mutant strain on MM-MES medium supplemented with 200 mM arabinose ( ■) and 200 mM xylose (▼) during the exponential phase of growth. **(D)** Prediction of the promoter sequence of the phosphoketolase-containing operon, located upstream of the *araD* gene (CAC1341). Predicted -10box and -35box sequences are shown along with the distance from the promoter to the START codon of the first gene of the operon. The growth curve results shown in this figure come from the representative experiment. The experiments were repeated at least three times, giving similar results; for each optical density measurement three technical replicates were made, as described in section 2.3.

## 3.4. Continuous fermentation of *C. acetobutylicum* ATCC 824 in xylose-supplemented, phosphate-limited medium

A genome-wide transcriptional analysis of the phosphate-limited culture provides information on the regulation of multiple genes during two tightly controlled growth phases – acidogenesis and solventogenesis, and the transition between both, separating solvent production from sporulation. The advantage of this method, compared to a pH-controlled batch fermentation, is the possibility of maintaining a constant growth rate with the undisrupted inflow of nutrients and outflow of metabolic products secreting to the medium. The gene expression profile of the wild type *C. acetobutylicum* ATCC 824 strain was examined during acidogenic and solventogenic growth, as well as during the switch between acidogenesis and solventogenesis in the continuous culture on a xylose-supplemented, phosphate-limited medium, and the genes regulated during this experiment were analyzed and described as in the previous work on glucose (Grimmler *et al.*, 2011). Two independent experiments were conducted and the results obtained in both were compared. For closer analysis one of them was chosen as a representative experiment and the other was considered a control. In this part their fermentation, the growth and xylose consumption profiles are shown, and furthermore the genes expressed or inhibited during the entire continuous culture experiment are divided into four groups and analyzed.

### 3.4.1. The phenotype of *C. acetobutylicum* wild type strain in phosphate-limited, xylose-supplemented continuous culture

Two continuous culture experiments were conducted for the purpose of this work. The first one was treated as a representative experiment, and the second as a control.

In the first, representative experiment all spores were washed out after 72 h and the culture entered steady acidogenic growth 96 h after the inoculation of the fermentor, exhibiting typical butyric acid fermentation, with 10.68 mM of acetate and 77.78 mM of butyrate detected. A stable pH of 5.7 was maintained by the constant addition of 2 M KOH, which after 144 h was discontinued; the pH values decreased to 4.5 in 19 h and

45 min, and the culture entered steady-state solventogenic growth 216 h after the inoculation time. During solventogenesis the culture produced 30.15 mM of butanol and 22.32 mM of acetone (Fig. 3.4.1 I B), and these results are comparable with what was observed in the culture grown on glucose (Grimmler *et al.*, 2010). The concentrations of ethanol were nearly unchanged during the entire continuous culture experiment: 4.13 mM during acidogenesis (pH 5.7) and 4.48 mM during solventogenesis (pH 4.5), as observed previously (Grimmler *et al.*, 2010).

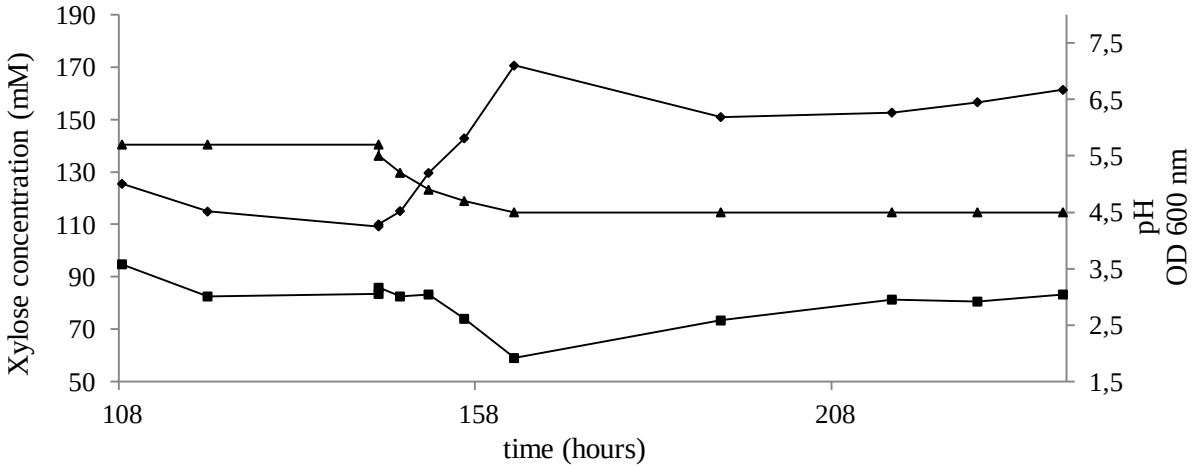
In the second, control experiment, spores were washed out after 72 h and steady-state acidogenic growth was also observed starting 72 h after the inoculation time. During the butyric acid fermentation 42.28 mM of acetate and 61.26 mM of butyrate were detected in the medium. The transition from acidogenesis to solventogenesis was initiated after 137 h and lasted for 23 h and 15 min. Steady-state solventogenic growth was reached after 212 h. In the second culture higher butanol production (44.51 mM) and generally higher amounts of other solvents, with 25.94 mM of acetone and 7.69 mM of ethanol were observed during solventogenesis (Fig. 3.4.1 II B). The concentrations of ethanol, as observed in the first experiment, remained nearly unchanged during the experiment (5.19 mM at pH 5.6 and 7.69 mM at pH 4.5). In both cultures the highest concentrations of solvents were observed during the last shift (pH 4.5) of the transition from acidogenesis to solventogenesis (Fig. 3.4.1. I and II B).

The xylose concentration increased in both cultures during the transition phase between acidogenesis and solventogenesis, and the changes correlated with the decrease in the optical density values in each culture (Fig. 3.4.1 I and II A). Changes in the xylose consumption were also correlated with rising amounts of butanol, which is known to inhibit xylose uptake (Ounine *et al.*, 1985). Detailed information on the growth, xylose consumption and fermentation profile of the representative experiment is given in the Supplementary Data section (Tab. 9.9).

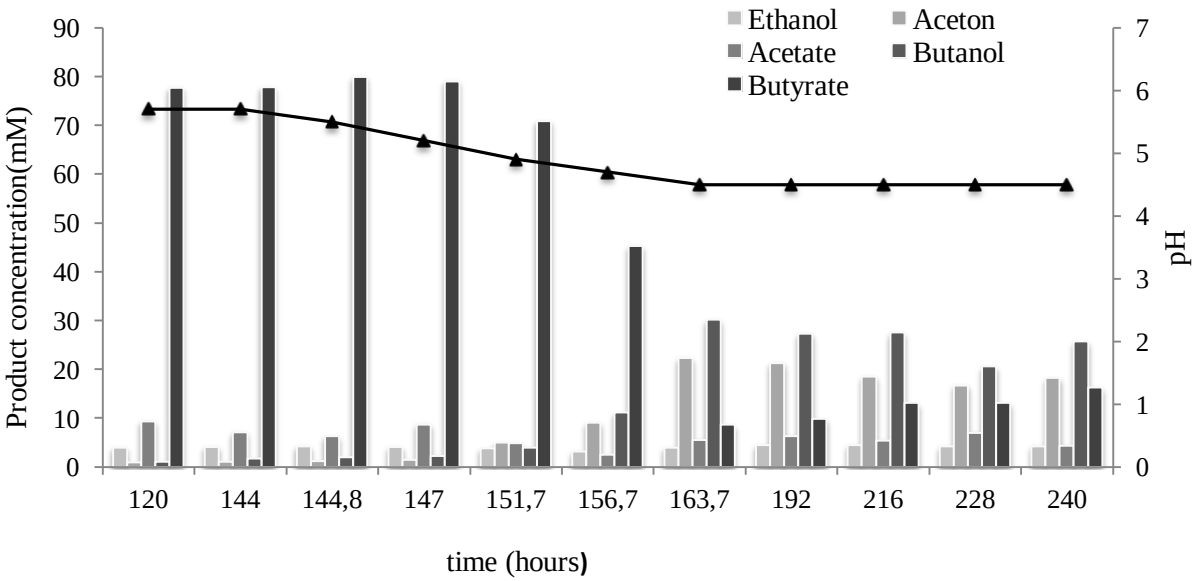
Furthermore, the expression profile of the genes regulated during acidogenesis and solventogenesis and the transition between both states were examined and compared with the results obtained for the glucose-supplemented culture (Grimmler *et al.*, 2010) and other published microarray experiments.

**I.**

**A**

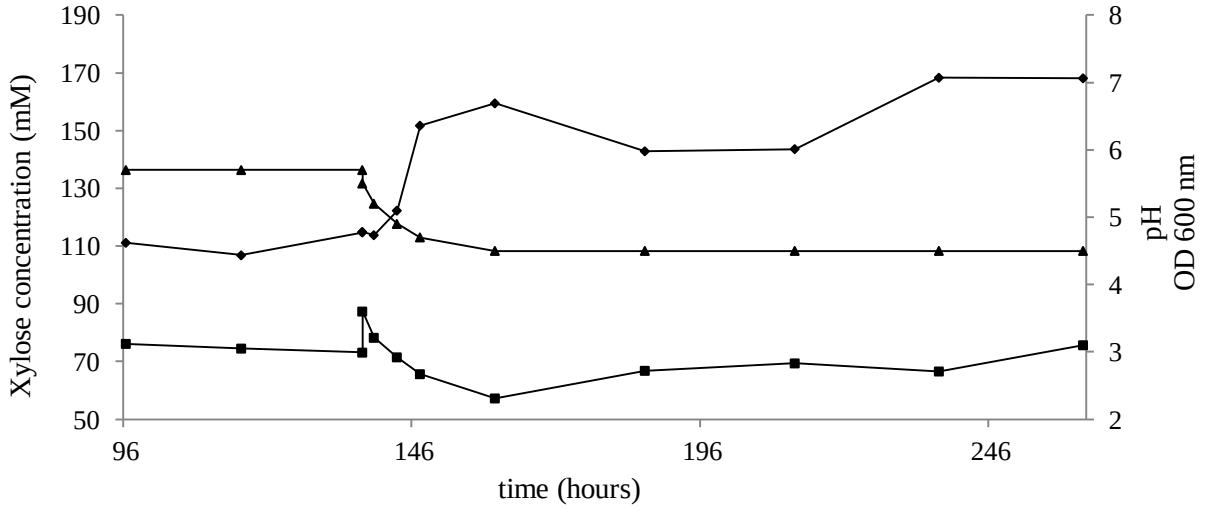

Xylose concentration (mM)
190
170
150
130
110
90
70
50
7,5
6,5
5,5
4,5
3,5
2,5
1,5
pH
OD 600 nm
108
158
208
time (hours)

**B**


Ethanol
Aceton
Acetate
Butanol
Butyrate
Product concentration(mM)
90
80
70
60
50
40
30
20
10
0
7
6
5
4
3
2
1
0
pH
120
144
144,8
147
151,7
156,7
163,7
192
216
228
240
time (hours)

**II.**

**A**




**B**

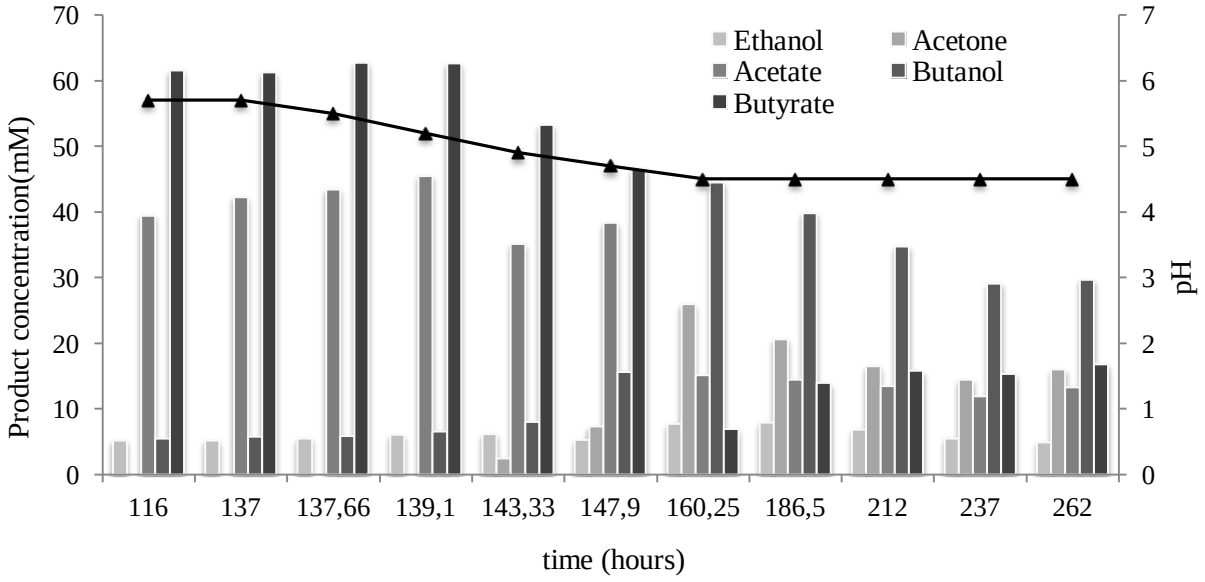



**Figure 3.4.1.** Growth, xylose consumption and fermentation profile of the first (I) and second (II) phosphate-limited continuous culture of the wild type *C. acetobutylicum* strain. **(A)** Xylose concentrations in the supernatant (●), $OD_{600}$ (■) and pH (▲) values of steady-state acidogenic growth, the transition from acidogenesis to solventogenesis and the steady-state solventogenic growth are shown. **(B)** The concentrations of the fermentation products, namely ethanol, butanol, acetone, acetate and butyrate measured during the entire continuous culture experiment are shown as a bar chart, with a solid line indicating the pH changes (▲).Three

technical replicates were made for each optical density and pH measurement, as described in section 2.3, and for each xylose concentration measurements, as described in section 2.8.2.

### 3.4.2. Transcriptional analysis of the xylose-supplemented continuous culture

Transcriptional analysis of the genes regulated during the continuous culture experiment give detailed information on the metabolism of the cell during acidogenesis, solventogenesis and the shift from acid to solvent production. In this work two independent experiments were conducted, and the data from both were analyzed. Based on the quality and the amount of data obtained from both microarray experiments, the first experiment was treated as representative and further described, while the second one was treated as a control.

Samples from the acidogenesis (A; pH 5.7) and solventogenesis (S; pH 4.5), as well as from each pH shift (pH 5.5; 5.2; 4.9; 4.7 and 4.5) of the switch between both steady-state growth phases (Fig. 3.4.2) were collected. A Sample from steady solventogenic growth (S) was used as a reference, and the results of this analysis are shown below.

All the data were divided into four groups – the first one would collect all the genes upregulated during the acidogenesis and the second all the genes upregulated during solventogenesis. However, some of the genes remained non-regulated during steady growth, but their expression differed through the metabolic shift. Therefore the third group, collecting the genes upregulated during the transition phase, and the fourth group, which gathers genes inhibited through the switch between acidogenesis and solventogenesis, were created. During the analysis of the data some of the groups of genes were observed to be highly upregulated, while the others were strongly downregulated. For the discussion certain groups of genes were chosen based on the expression pattern and the data that have been previously published.

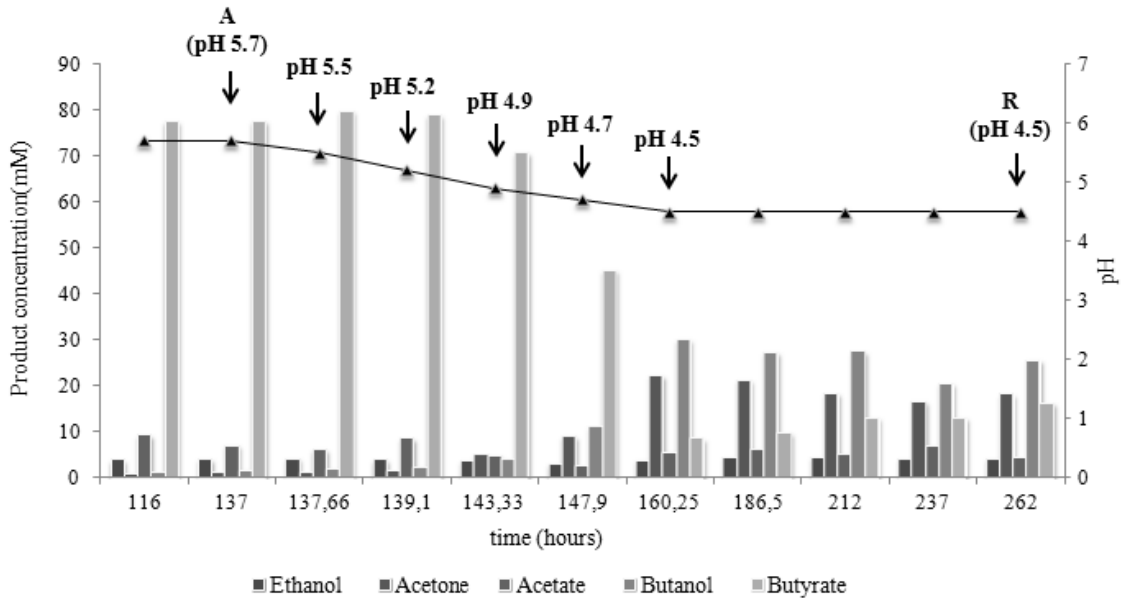



**Figure 3.4.2.** During the entire continuous culture experiment seven time points were chosen to collect samples for microarray experiments and are marked on the graph by arrows. The first sample was collected during the acidogenesis (A; pH 5.7), the following samples during the shift from the acid to solvent production (pH 5.5 – pH 4.5) and the last, used as a reference sample, was taken during solventogenesis (R; pH 4.5), as indicated on the graph.

### 3.4.2.1. Genes significantly upregulated during steady-state acidogenic growth (group 1)

The first group includes 141 genes induced during the acidogenesis. Among them several gene clusters were observed to be strongly upregulated. Two uncharacterized proteins were annotated as orthologs of the *ygaT* (CAP0036) and *ygaS* (CAP0037) genes of *Bacillus subtilis*. Both genes were upregulated during the acidogenesis and the metabolic switch, but downregulated during solventogenesis. Genes coding for the periplasmic hydrogenase small subunit *mbhs* (CAP0141) and the large subunit *mbhl* (CAP0142), and the hydrogenase maturation protease delta subunit, HyaD-like (CAP0143) were also found to be highly upregulated during steady-state acidogenic growth and the switch between acidogenesis and solventogenesis. Products of the *mbhs* and *mbhl* genes are involved in hydrogen uptake, which may contribute to a more balanced NADP(H) metabolism, needed for butanol production (Hu *et al.*, 2011).

Furthermore, the genes involved in carbohydrate transport and metabolism (CAC0662 – CAC0668), including the ABC-type sugar transporters and sugar permeases, were highly upregulated during acidogenesis and the beginning of the transition phase between both steady-state growths (at pH 5.5). When the pH decreased the expression of the genes became more inhibited. ABC transporters are present among multiple species of bacteria and are used for the ATP-related transport of different substrates such as ions, sugars, lipids, proteins, etc. The uptake of pentoses is believed to be directed through the ATP-binding cassettes (Servinsky *et al.*, 2010) and the expression pattern of the CAC0662 – CAC0668 genes correlated with xylose consumption rates during the entire continuous culture. Along with the genes involved in carbohydrate transport, three genes involved in sucrose and glycogen metabolism, the *glgC* glucose-1-phosphate adenylyltransferase (CAC2237), the *glgC* ADP-glucose pyrophosphorylase (CAC2238) and *glgA* glycogen synthase (CAC2239) were upregulated during the steady-state acidogenic growth. A similar regulation pattern was characteristic for the *glgP* gene (CAC1664), coding for glycogen phosphorylase. Additionally, genes encoding the glycosyltransferases were found to be upregulated during acidogenesis and the transition to solventogenesis. CAC2345, CAC2346, CAC2350 and CAC2351, annotated as the genes of polysaccharide metabolism, involved in the cell wall biogenesis, were also upregulated during steady-state acidogenic growth. Apart from the genes of carbohydrate metabolism, two genes involved in amino acid transport and metabolism, CAC0014 and CAC0015, were found to be strongly upregulated during steady-state acidogenic growth and during most pH shifts of the transition from acidogenesis to solventogenesis.

In the first experiment multiple sporulation genes were observed to be upregulated during acidogenesis and the shift from acid to solvent production. Among these genes several groups were determined, including CAC1336 – CAC1338 (hypothetical protein, spore coat protein COTJB and spore coat protein COTJC); CAC1694 – CAC1697 (sigma factor E processing enzyme SpoIIGA, SigE sporulation sigma factor, SigG sporulation sigma factor and uncharacterized conserved protein – YMXH *B. subtilis* homolog) and CAC2086 – CAC2093, consisting of stage III sporulation proteins (SpoIIIAH, SpoIIIAG, SpoIIIAF, SpoIII AE, SpoIII AD, SpoIIIAC, SpoAB, SpoIIIAA). The SpoIIID-coding gene (CAC2859) was significantly upregulated during acidogenesis and the transition to solventogenesis. A similar pattern was presented by a

large cluster of genes annotated as possibly involved in cell envelope and outer membrane biogenesis (CAC2345 – CAC2350).

Among the genes belonging to the first group, a big cluster of genes (CAC2576 – CAC2581) was positively regulated during acidogenesis and the transition to solventogenesis. It is composed of the 6-pyruvoyl-tetrahydropterin synthase-related protein, GGDEF-domain-containing protein (inactivated), glycosyltransferase, hypothetical protein, hypothetical protein (CF-41 family) and conserved 6-pyruvoyl-tetrahydropterin synthase-related domain, and all these genes encode key molecules, involved in essential processes of the cell (Zhang *et al.*, 2012; Haft *et al.*, 2012; Reader *et al.*, 2004).

### 3.4.2.2. Genes significantly upregulated during steady-state solventogenic growth (group 2)

The second group consists of 95 genes upregulated during steady-state solventogenic growth. Here, several genes located on the pSOL1 plasmid were identified. A part of a *sol*-operon, namely aldehyde dehydrogenase (*aad*, CAP0162), the butyrate-acetoacetate CoA-transferase subunit A (*ctfA*, CAP0163) and butyrate-acetoacetate CoA-transferase subunit B (*ctfB*, CAP0164) was upregulated during solventogenesis only. The genes were strongly inhibited during acidogenesis and became induced when the pH of the continuous culture decreased to pH 5.2. The acetoacetate decarboxylase (*adc*, CAP0165) gene was observed to be non-regulated during steady-state acidogenic or solventogenic growth, but strongly induced during the transition phase between both growth states.

Furthermore, the genes involved in carbohydrate transport and metabolism were positively regulated during solventogenesis. The xylanase, glycosyl hydrolase family (*xynb*, CAP0053), xylanase/chitin deacetylase family enzyme (CAP0054) and secreted pectate lyase from the polysaccharide lyase family (*pell*, CAP0056) were upregulated during steady solventogenic growth, but inhibited at the beginning of the transition from acidogenesis to solventogenesis. A similar regulation pattern was observed for single predicted xylanase/chitin deacetylase-encoding genes (CAC2383, CAC2396 and

CAC3017) and pectate lyase genes (CAC0574, CAC1968 and CAC3387). Two other genes were observed to be upregulated only during solventogenesis, and inhibited during steady-state acidogenic growth and metabolic switch: the glycerol-3-phosphate dehydrogenase (*glpA*, CC1322) is involved in glycerophospholipid metabolism and the galactokinase (*galK*, CAC2959) takes part in carbohydrate metabolic pathways.

Additionally, a large group of genes supposed to be involved in cellulose degradation was observed to be strongly inhibited during acidogenesis; the inhibition decreased during the pH shift and the genes were upregulated in steady-state solventogenic growth. The CAC0910 – CAC0920 cluster codes mostly for *cel* genes, and is composed of a cellulosomal scaffolding protein precursor, possible processive endoglucanase family 48 (CelF ortholog), possible non-processive endoglucanase family 5 (CelA homolog), possible non-processive endoglucanase family 9 (CelG ortholog), a cellulosome integrating cohesin-containing protein, cellulase A, CelG ortholog, cellulose-binding endoglucanase family 9 (CelL), possible non-processive endoglucanase family 5 (mannose A ortholog), secreted sialidase and a protein related to the MIFH/DOPD protein family of unknown function. The CelE cellulase ortholog, dockerin domain (CAC0561) and two endoglucanase genes (CAC0826 and CAC3469) were regulated in a similar manner to the cellulosome operon. These genes of *C. acetobutylicum*, supposedly involved in cellulose degradation process were described and analyzed previously (Grimmler *et al.*, 2011; López-Contreras, 2003 and 2004).

Interestingly, three genes coding for uncharacterized, small conserved proteins (CAC0410, CAC0411 and CAC0413) were detected to be downregulated during steady-state acidogenic growth and nearly the entire transition from acidogenesis to solventogenesis. They are annotated as homologs of YUKE/YFJA proteins, and belong to the WXG100 family, present among Gram-positive species. Studies suggest they are related to a secretion system (Garufi *et al.*, 2008). A similar regulation pattern was observed for the genes of the CAC3710 – CAC3713 cluster, bringing together an uncharacterized protein – a YUKC *Bacillus subtilis* homolog, two hypothetical proteins and another uncharacterized conserved protein of the YUKE/YFJA *B. subtilis* family.

Furthermore, the second group includes also a small operon, CAC2392 – CAC2393, encoding two uncharacterized ATPase components of an ABC transporter, involved in ATP catabolic process. Both genes were strongly downregulated during acidogenesis

and the majority of the transition to solventogenesis, but they became induced during steady-state solventogenic growth.

### 3.4.2.3. Genes significantly upregulated during the transition from acidogenesis to solventogenesis (group 3)

In this group 227 genes, upregulated during the transition from acidogenesis to solventogenesis, but non-regulated during steady growth states, were collected.

Two genes, pyruvate decarboxylase (*pdc*, CAP0025) and a hypothetical protein (CAP0026) were upregulated during the transition from acidogenesis to solventogenesis at pH values of between 5.2 and 4.9. Furthermore, positive regulation of thiolase B, *thlB* (CAP0078), and the transcriptional regulator *thlR* (CAP0079), was observed. They were non-regulated during acidogenesis and their expression was induced during the metabolic switch at pH values of 5.2 – 4.7. An identical regulation pattern was observed for acetoacetate decarboxylase (*adc*, CAP00165) and a predicted acetyltransferase (CAC2468).

During the shift from acidogenesis to solventogenesis, the genes involved in aromatic amino acid biosynthesis were noticed to be positively regulated, among them 3-deoxy-7-phosphoheptulonate synthase, prephenate dehydrogenase and *aroBACFK* genes (CAC0892 – CAC0898). The *aro* genes are involved in the shikimate pathway synthesis of aromatic amino acids, namely phenylalanine, tyrosine and tryptophan. The *TrpABFCD* (CAC3157 – CAC3161) operon was also found to be strongly upregulated during the transition from acidogenesis to solventogenesis. Similar regulation was shown for the *pflBA* genes (CAC0980 – CAC0981), coding for pyruvate-formate lyase. Pyruvate-formate lyase from *C. acetobutylicum* was recently identified as a biosynthetic enzyme for purine biosynthesis in our laboratory (Hönicke, personal).

Furthermore, group 3 includes genes involved in purine (*purECFMNHD*, CAC1390 – CAC1396) and pyrimidine (*pyrDZFIB*, CAC2650 – CAC2654) metabolism. In addition, thioredoxin reductase, *trxB* (CAC1548), was strongly upregulated during the

transition from acidogenesis to solventogenesis, along with two neighboring genes, thioredoxin (*trxA*, CAC1547) and glutathione peroxidase (*bsaA*, CAC1549).

Genes involved in hyperosmosis and the heat shock response of *C. acetobutylicum* (CAC1280 – CAC1283) were found to be upregulated during the metabolic switch, and downregulated at the end of the transition to solventogenesis. The *GroEL/ES* (CAC2703 – CAC2704) operon and molecular chaperone (CAC3714) gene were similarly regulated, while the *htpG* heat shock protein gene was found to be non-regulated during the majority of the switch from acidogenesis to solventogenesis, but downregulated at its last pH shift (pH 4.5).

### 3.4.2.4. Genes significantly downregulated during the transition from acidogenesis to solventogenesis (group 4)

In this group 24 genes significantly downregulated during the switch between acidogenesis and solventogenesis, but non-regulated during both of the steady-state growth phases, were collected.

Among these genes a large cluster of genes involved in xylan degradation (CAP0114 – CAP0120) composed of a possible beta-xylosidase, XynD (endo-1,4-beta-xylanase), xylanase, possible beta-xylosidase and three possible xylan degradation enzymes, was identified. Although the CAP0114 – CAP0120 genes became upregulated at the last pH shift (pH 4.5), they were downregulated during nearly the entire switch from acidogenesis to solventogenesis, and therefore described as belonging to the fourth group. An uncharacterized conserved membrane protein (CAC0193) and glycosyltransferase (CAC0194) were strongly downregulated during the entire transition from acidogenesis to solventogenesis. Both genes are annotated as involved in cell wall biosynthesis. Moreover, the *ugpAEB* genes (CAC0427 – CAC0429) coding for sn-glycerol-3-phosphate ABC transporters were strongly downregulated during the majority of the metabolic switch.

Additionally, the *argCJ* genes (CAC2390 – CAC2391), involved in arginine biosynthesis, were downregulated during the entire metabolic shift from acid to solvent production.

Aconitase (*citB*, CAC0971) and isocitrate dehydrogenase (*citC*, CAC0972) were strongly inhibited during the transition from acidogenesis to solventogenesis (at pH 5.2 – 4.7). Like other strictly anaerobic organisms, *C. acetobutylicum* contains a bifurcated TCA cycle (Crown *et al.*, 2011; Amador-Noguez *et al.*, 2010). A similar regulation pattern is shown by isopropylmalate synthase (CAC0970). The protein contains conserved domains described as involved in isopropylmalate/homocitrate/citramalate synthesis (Marchler-Bauer *et al.*, 2013).

Furthermore, a strong inhibition of the operon composed of the *arsE*-family transcriptional regulator, the probable arsenical resistance operon repressor, and the cation transport P-type ATPase (CAC2242 – CAC2241) was observed during nearly the entire transition from acidogenesis to solventogenesis. The CAC2241 gene contains a heavy-metal associated domain and the whole operon is believed to take part in the heavy metal ion resistance mechanism.

Transcriptional regulations of the chosen genes and gene clusters responsible for acid and solvent production, cellulosome formation, sporulation cascade, sugar, amino acis and fatty acid metabolism from all four groups were further discussed and compared with the data from previous transcriptional experiments in the Discussion, section.4.4

## 3.5. The analysis of the Rnf system in *Clostridium beijerinckii* NCIMB 8052

### 3.5.1. The Clostron *rnfC* mutant creation

Although most of this work focused on the solvent production related to pentose metabolism, the investigation of the Rnf complex of *C. beijerinckii* could give very interesting insights into the energy metabolism of the microorganism. Energy metabolism is strongly linked to sugar uptake and utilization pathways. For example, the glycolysis and gluconeogenesis pathways of *C. beijerinckii* bring together genes responsible not only for direct sugar molecule conversion, but also for ATP and $NAD^+$-NADH production and the maintenance of the transmembrane ion gradient. *C. beijerinckii*, like *C. saccharobutylicum*, possesses the Rnf complex (Poehlein *et al.*,

2013) responsible for the ferrodoxin-dependant reduction of $NAD^+$, which generates a transmembrane proton gradient. $NAD^+$ was regarded as a universally used electron donor in classic bioenergetics, but, in anaerobes, ferredoxin is widely used as a more electronegative electron carrier. The Rnf complex is present among many Gram-positive bacteria (Tremblay *et al.*, 2012; Wang Y *et al.*, 2013 a; Biegel *et al.*, 2009; Müller *et al.*, 2008), and was described as playing an important role in nitrogen fixation and energy metabolism. It was first discovered in *Rhodobacter capsulatus* (Schmehl *et al.*, 1993; Jouanneau *et al.*, 1998) and has been described and investigated in multiple organisms since then (Tremblay *et al.*, 2012; Sarkar *et al.*, 2012; Biegel *et al.*, 2011; Biegel & Müller, 2010; Dixon & Kahn, 2004). In *C. beijerinckii* NCIMB 8052, the RnfABCDGE complex is located in one operon, CBEI2449 – CBEI2454 (Fig. 3.5.1), therefore to disrupt the whole complex, a ClosTron mutant was created in the first gene of the operon, the *rnfC* (CBEI2449). To this end the pMTL007S-E2::CBEI2449-783|784s plasmid was designed and created as described in section 2.5.2 of Materials and Methods. Correct mutants were isolated based on the PCR check on the DNA template isolated from the fastest growing colonies using the *rnfC* check fwd and rev primers (Tab. 9.5) and one integration event was confirmed with Southern Blot (Fig. 3.5.2). For the blotting procedure mutant genomic DNA was digested with *Hind*III and *Eco*RV restrictases overnight and should produce a band of around 6450 bps size. In this particular Southern Blot experiment the standard NEB 1kb ladder (New England Biolabs) was used instead of the biotinylated 2-Log DNA Ladder (New England Biolabs).

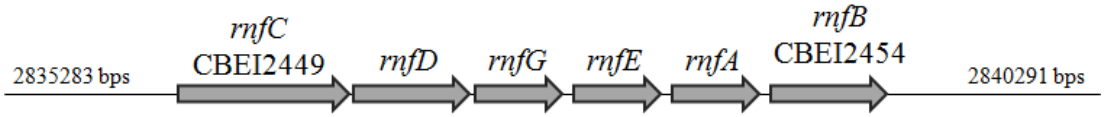



**Figure 3.5.1.** Schematic structure of the Rnf complex gene cluster. The *RnfC* gene (RnfABCDGE type electron transport complex subunit C) is the first of the operon, followed by the *rnfD* (subunit D), *rnfG* (subunit G), *rnfE* (subunit E), *rnfA* (subunit A) and *rnfB* (subunit B) subunits. Gene numbers are shown according to the NCBI database.

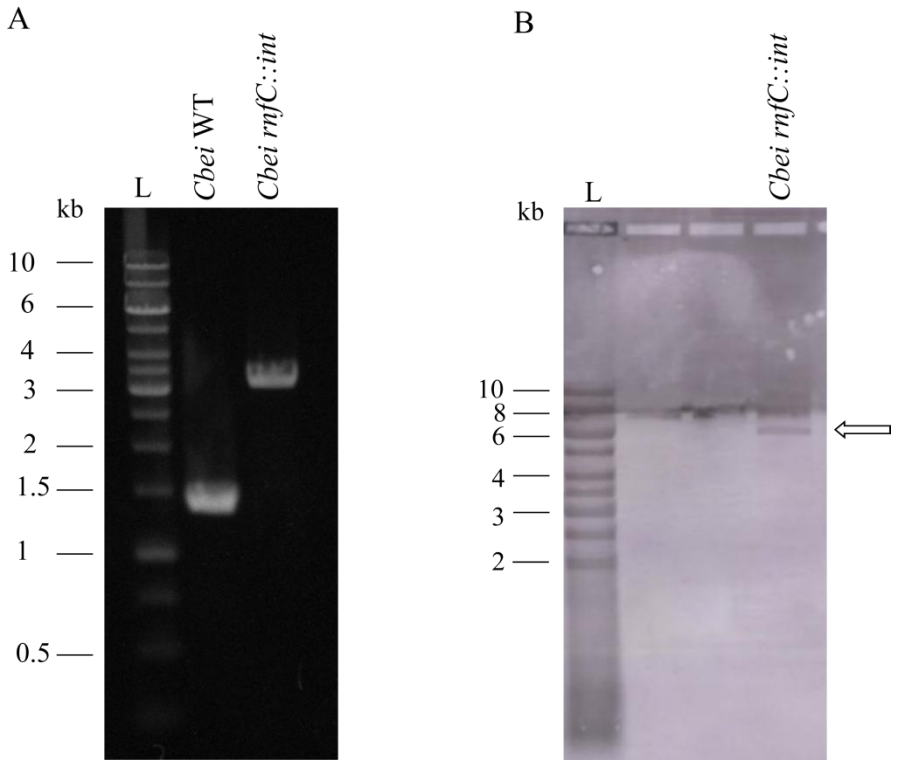



**Figure 3.5.2. (A)** Integration of the *rnfC*-targeting intron into the *C. beijerinckii* genome. Correct integration of the intron was determined by the PCR reaction. The wild type strain produced a band of 1379 bps, while *the rnfC::int* mutant produced a band of around 3479 bps. **(B)** A single integration event of the *rnfC*-targeting intron was proved with the Southern Blot on the *HindIII, EcoRV*-digested *rnfC::int* mutant DNA. L, DNA 1kb Ladder. The band produced by mutant DNA is indicated with an arrow; the band above is an unspecific artifact and is present also in both empty lanes, between the marker and *rnfC::int* lane.

### 3.5.2. Determination of the phenotype of the *C. beijerinckii rnfC::int* mutant in standard conditions

The aim of this part of the study was to examine whether disruption of the first gene of the Rnf complex would affect the phenotype of *C. beijerinckii* in standard growth conditions. The whole Rnf complex is bound to the membrane, and previous studies have shown that mutants in various parts of the *rnfABCDGE* operon result in a non-functional Rnf system (Schmehl *et al.*, 1993; Jouanneau *et al.*, 1998). In *C. beijerinckii* NCIMB 8052, the *rnfC* gene opens the *rnfABCDGE* operon, and therefore its disruption should affect the whole system, as it is believed its stability is

strongly dependent on the co-synthesis of all subunits (Kumagai *et al.*, 1997). The RnfC subunit is thought to be a soluble, 4[Fe-S]–containing protein and is characterized as associated with the membrane (Schmehl *et al.*, 1993). In mutants with insertions in genes downstream of *rnfC*, lower amounts of the RnfC protein were detectable, and in *rnfA, rnfB* or *rnfC* mutants no RnfC peptide was detected, probably due to a polar effect. Mutations resulted in the disruption of the whole complex, which often affects the general metabolism of the strain (Jouanneau *et al.*, 1998). It is already known that the Rnf complex of *Clostridium ljunghdalii* is active under both heterotrophic and autotrophic growth conditions, and that its disruption abolishes autotrophic growth of the strain and that the *C. ljunghdalii rnfAB* mutant shows greatly disproportionate growth in rich medium during heterotrophic growth, compared to the wild type (Tremblay *et al.*, 2012).

Growth profile tests prepared in the batch cultures based on the rich RCM medium revealed no detectable differences in the efficiency of growth or in the pH changes between the *C. beijerinckii* wild type and the *rnfC::int* mutant. Additionally, the rates of glucose consumption in both strains were examined (Fig 3.5.3 A) in the RCM with 10 g/l glucose supplementation. The results of one, representative experiment are shown below (Fig. 3.5.3 B).

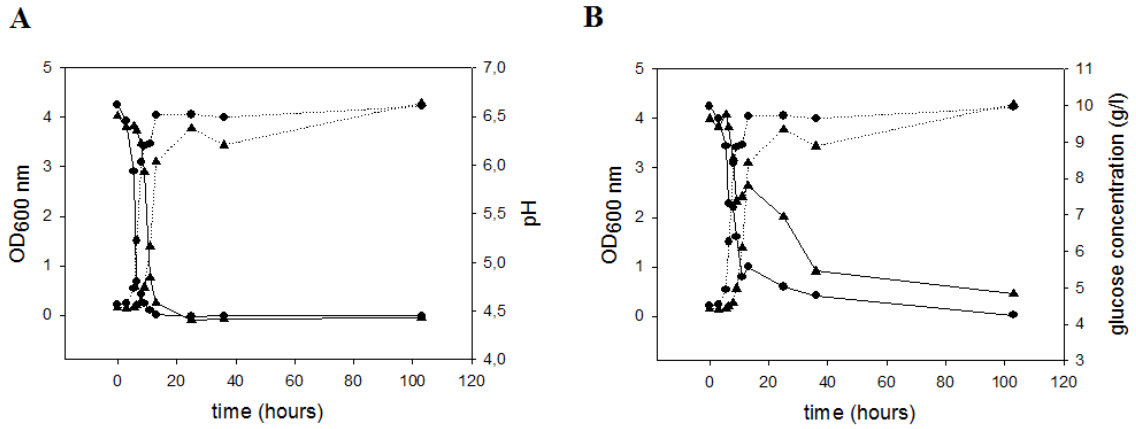



**Figure 3.5.3.** Representation of growth, pH changes and glucose consumption curves in batch culture experiments on the *C. beijerinckii* wild type (●) and *rnfC::int* (▲) mutant strain. **(A)** A comparison of the growth curve ($OD_{600\ nm}$, dotted line) and pH values (solid line) is shown with **(B)** a comparison of the growth curve ($OD_{600\ nm}$, dotted line) and glucose consumption profile (solid line) in both strains. The growth curve results shown in this figure come from the representative experiment. The experiments were repeated three times, giving similar results; three technical replicates were made for each optical density and pH measurement, as described in section 2.3, and for each glucose concentration measurements, as described in section 2.8.2.

Both wild type and *rnfC::int* mutant strains were able to grow to similar OD values and both reached the same pH values after 103 hours of cultivation. The amounts of glucose that remained in the medium were also comparable. It is therefore possible the Rnf complex is not constitutively active in *C. beijerinckii*, in contrast to what has been observed for *C. ljunghdalii.*

# 4. Discussion

## 4.1. The *upp*-based clean deletion system

Genetic manipulations of clostridia give a broad spectrum of possibilities for the investigation and improvement of solventogenic strains. Although intron mutagenesis systems (ClosTron, TargeTron) have been created and proved to be very efficient in many *Clostridium* species (Wang Y *et al.*, 2013 b; Heap *et al.*, 2010; Heap *et al.*, 2009; Yao & Lambowitz, 2007), there is still a great need for a clean deletion and insertion system.

Markerless mutagenesis systems based on the activity of various phosphoribosyltransferases (PRTases) have been established for many prokaryotic organisms. They were proved to be applicable in various Gram-negative and Gram-positive bacteria as well as in archeons (Kristich *et al.*, 2005; Bitan-Banin *et al.*, 2003; Fabret *et al.*, 2002; Peck *et al.*, 2000), and they all are based on a common PRTases feature, which is the conversion of modified purines and pyrimidines into their toxic base analogues (Results, 3.1). An *upp*-based approach for *C. acetobutylicum* was adapted from *B. subtilis* (Fabret *et al.*, 2002) and patented (Soucaille, Figge & Croux, 2008), while similar *pyrE*-based systems have been successfully applied in many other clostridia (Ng *et al.*, 2013; Tripathi *et al.*, 2010; Ehsaan, PhD thesis, 2013).

To broaden the spectrum of genetic tools fully available for *C. acetobutylicum* the *upp*-based system for allelic exchange mutagenesis (ACE) was designed. Several plasmids were constructed to remove the *upp* gene (CAC2879), using different replicative plasmid backbones and multiple variants of the deletion cassettes and various conditions were tested to increase the odds of a recombination event. Non-replicative (suicidal) plasmids were used previously and resulted in *pta, buk* and *aad* mutant creation (Green *et al.*, 1996; Green & Bennett, 1996), however they proved to be inefficient in *C. acetobutylicum* for the deletion of certain genes, the *upp* among them (Kraußе, PhD thesis, 2012). Allelic exchange mutagenesis is based on a phenomenon common to all living organisms, called homologous recombination.

Homologous recombination plays the role of a major DNA repair process in bacteria, but can also serve as a way to increase genetic variation during the process of horizontal

gene transfer. The whole process, although well described in model organisms, still remains poorly known in many species (Wigley, 2013; Blackwood *et al.*, 2013; Karpenshif *et al.*, 2012). In general, when two DNA molecules sharing a common DNA sequence (the 'homologous region') are present in the same cell, they may recombine *via* a Campbell-like mechanism to form a single molecule. When this principle is applied to the genomic DNA and plasmid DNA, the plasmid would integrate into the genome in a process called the first recombination. This DNA molecule is very unstable under certain conditions and a second recombination event normally occurs. A second recombination would result in the excision of the plasmid and either mutant creation through the allelic exchange, or reversion of the cell to the wild type. This feature was exploited to integrate the cassette composed of the regions of homology located up- and downstream of the *upp* gene in the *C. acetobutylicum* wild type genome, to create clean deletion mutants. This method can be also used to introduce genes or regulatory elements into the genomic DNA of the host. What remains very important is the double recombination procedure, enabling the application of different conditions and tight control over the whole process at every step. Moreover, the *upp*-based system allows the construction of multiple mutants in one strain without the need to remove the selection marker after each mutagenesis step, and without causing polar effects on genes located downstream of the gene of interest.

In *C. acetobutylicum* TCC 824 the *upp* gene overlaps the *rpiB* (CAC2880) with its START codon (Nöling *et al.*, 2001), and therefore previously used deletion cassettes, designed to delete the whole *upp* gene along with the START and STOP codons, could affect the formation of functional ribose 5-phosphate isomerase (*rpiB*). Therefore in this work each cassette was designed to leave the *upp* START and STOP codons, to assure no influence of the gene alteration on its neighboring genes. The same approach was used to create clean deletion plasmids for the *araK* (pJL3) and *xylB* (pJL4) genes in *C. acetobutylicum*.

During the attempts to create an in-frame *upp* deletion several first integrants were observed, although they always showed extremely poor growth and it was not possible to obtain any second recombinants. Even though different conditions were applied to improve the first integrants' survival rate such as a rich and buffered media, uracil supplementation, or proceeding to the second recombination event as quickly as possible, it was not possible to maintain these strains for long enough. It must be

underlined here that the DNA molecules created during the process of first recombination are very unstable, and that could have been the reason for the failure of the attempts to maintain the first recombinants for a longer time. It is possible that the loss of the first recombinants was due to the second recombination event, resulting in the wild type strain. In theory, during the second recombination step mutants and revertants should be observed in 1:1 ratio, however for some genes, especially genes crucial for the basic metabolism of the cell, the ratio is shifted in favor of the revertants. The difficulties observed during the *upp*-based system creation could also originate from the fact that certain regions of bacterial DNA are less likely to recombine, probably due to regulatory elements or simply the high importance of the genes. Therefore, another method to construct the *upp*-deficient strain was used, and although the ClosTron mutagenesis is regarded as extremely efficient, it was not possible to obtain the *upp*-deficient strain during multiple approaches. The reason behind this remains unclear as an *upp*-deficient *C. acetobutylicum* strain was described (Soucaille, Figge & Croux, 2008).

Regardless of good transformation rates, procuring the first integrants was relatively difficult even though, in further studies with allelic exchange experiments on the *ΔpyrE* strain of *C. acetobutylicum*, the first and second integrants within the arabinose and xylulose kinase genes were obtained without much effort. In the natural environment the genetic changes in a genome are not promoted in the presence of favorable conditions, while strong stress often results in increased mutation or gene transfer rates. Both reactions to stress are considered to be an evolutionary response to environmental changes, which require new features to be gained for faster adaptation, therefore an efficient genetic variation machine could be a key factor for the survival of an organism. It is already known that different kinds of stress induce the expression of certain recombination genes (Carr & Lambert, 2013; Winn *et al.*, 2003; Zhong & Priest, 2011), and applying stress conditions results in a higher expression of DNA repair genes, with the homologous recombination genes among them. In *C. acetobutylicum* the expression of sporulation cascade genes was believed to be linked with the induction of DNA repair genes, and was observed to be activated by the presence of carboxylic acids or solvents, low pH and the limitation of nutritious elements: especially sugar, phosphate and nitrogen salts (Alsaker *et al.*, 2009 and 2005; Bahl *et al.*, 1995). To improve the recombination rate different conditions were applied, including the limited nutrient, low

pH and 5-FU-derived stress conditions. Supplementation with 5-fluorouracil was used exclusively for pJL10 and pJL11 plasmids, as they were designed to integrate within the *upp* gene, and hence the disruption of the phosphoribosyltransferase function would occur during the first recombination event.

Additionally to the work done on *C. acetobutylicum*, the impact of 5-FU on the *Clostridium saccharobutylicum* NCP 262 wild type in liquid and agar cultures was investigated and the strain proved to be sensitive to the chemical. The sensitivity of some bacteria towards the 5-FU depends on the medium used for the cultivation of a strain, i.e. the phosphoribosyltransferase mutants of *E. coli, B. subtilis* or *E. faecalis*, require supplementation with purine ribonucleosides to be sensitized to 5-FU, because it can be converted to UMP in a sequential action of uridine phosphorylase and uridine kinase (Kristich *et al.*, 2005; Neuhard, 1983). However, no changes in the sensitivity to 5-FU of the *C. acetobutylicum* or *C. saccharobutylicum* wild types were observed, regardless of the media used.

For future work it is suggested that the influence of butanol, acetate- or butyrate-derived stress (Alsaker *et al.*, 2009; Tomas *et al.*, 2004), which is also considered a strong trigger for the upregulation of expression of many genes related to DNA repair, should be investigated. Furthermore, creating a pJL10- or pJL11-derived plasmid bearing a functional copy of the *upp* gene of another *Clostridium* species could be a solution for problems arising from the disruption of the *upp* gene of *C. acetobutylicum* during the first recombination event, and could possibly increase integration rate.
The first integration event achieved during this work should be regarded as the sign of an adequate approach, as in previous work on *upp*-deficient strain construction no first integrants were observed, regardless of the many different approaches used (Krauβe, PhD thesis, 2012).

## 4.2. Development of a methylation system for *Clostridium saccharobutylicum* NCP 262 and the analysis of two endonuclease mutants

The development of an efficient transformation system is a prerequisite for the genetic engineering of the biotechnologically important strain *C. saccharobutylicum* NCP 262. A characterization of the restriction-modification operons was made during this work and a triparental conjugation protocol for a methylation system was established to overcome the restriction defense of the host and increase the conjugation rate.

### 4.2.1. Characterization of restriction-modification operons

*C. saccharobutylicum* NCP 262 has two restriction-modification systems (RM1 and RM2) annotated in the genome sequence (Poehlein *et al.*, 2013). The first RM system (RM1) consists of three genes: the restriction subunit (*hsdR1*, CSA00451), the methylation subunit (*hsdM1*, CSA00552) and the specificity subunit (*hsdS1*, CSA00453). A second RM system (RM2) contains three subunits: *hsdR2* (CSA02977), *hdsM2* (CSA02981) and *hdsS2* (CSA02979) and two hypothetical genes, CSA02978 and CSA2980.

Further investigation, based on the peptide BLAST experiments and analysis of the structure of both operons revealed that the HsdR1 peptide most likely belongs to the IA family, and the HsdR2 to the IC family, of restriction enzymes (Kulik and Bickle, 1996). For this analysis *E. coli* restriction enzymes were chosen, as most well-known and comprehensively described.

### 4.2.2. Methylation system for *C. saccharobutylicum* NCP 262

Despite the biotechnological importance of this solventogenic organism, no successful transformation of *C. saccharobutylicum* has yet been described. Among clostridia various RM systems have been reported for pathogenic and non-pathogenic species (Guss *et al.* 2012; Purdy *et al.* 2002; Pyne *et al.* 2013), and since nothing is yet known about the specificity of the restriction systems of *C. saccharobutylicum*, the *in vivo* methylation by the methyltransferases (MTases) from the organism was used to overcome the restrictase activity. Since no positive results were obtained with an *in*

*vitro* methylation method, a tri-parental conjugation approach based on the *in vivo* DNA modification was developed.

Conjugation is a common cell-to-cell, horizontal genetic transfer mechanism among bacteria (Griffiths *et al.*, 2000). The donor cell bears the F-plasmid ($F^+$ phenotype), carrying genes responsible for pili biosynthesis and the formation of proteins responsible for attaching to the recipient cell. A tri-parental conjugation with two donor *E. coli* strains was developed and proved to work efficiently, although the details of the plasmid transfer between the donor and recipient cells in this approach remain unclear. Previous work on *C. difficile* showed all plasmids from one donor cell are transferred to recipient cells (Purdy *et al.*, 2002), and therefore it is possible that the DNA transfer occurs first between two *E. coli* donors, and then between *E. coli* and *Clostridium* strains.

### 4.2.3. Efficiency of conjugation

The restriction-modification systems of microorganisms serve as simple defense mechanisms against extraneous DNA (Murray, 2002 and 2000; Krüger & Bickle, 1983). When the unmethylated plasmid DNA was used for conjugation, no transconjugants of the wild type *C. saccharobutylicum* were observed. However, when the DNA was methylated with one of the native MTases of *C. saccharobutylicum*, multiple plasmid-bearing colonies were noticed. Experiments with wild type recipients with and without *in vivo* methylation of the donor DNA suggested the RM2 system of *C. saccharobutylicum* has a higher significance for restriction as well as for methylation. The tri-parental conjugation method described in this work resulted in an efficiency rate higher by one order of magnitude compared to what has been reported previously on similar conjugation approaches for *C. difficile* (Mullany *et al.*, 1991; Purdy *et al.*, 2002).

### 4.2.4. Different origins of replication for *C. saccharobutylicum*

Each replicative plasmid possesses a region called the origin of replication (*ori*), and this region is very often host-specific (Mott & Berger, 2007; Kelman & Kelman, 2004; Baker & Wickner, 1992). Several types of *ori* applicable for clostridia were

described before (Heap *et al.*, 2009), yet no information on their function in *C. saccharobutylicum* has been obtained so far. Four different functional origins of replication in *C. saccharobutylicum* were established: the pIM13, pBP1, pCB102 and pCD6 (Heap *et al.*, 2007; Purdy *et al.*, 2002; Mermelstein & Papoutsakis, 1993; Davis, 1998). Furthermore, the pIM13 and pBP1 were observed to remain in cells under no selective pressure for multiple restreak events, while the pCB102 and pCD6, were lost relatively quickly (Tab. 3.2.1). These results gave an insight into information important for different mutagenesis procedures. Stable plasmids are of great value for plasmid-derived gene silencing or expression, while unstable plasmids are used especially for genome manipulations, such as knock-outs and knock-ins, where there is a need to remove the exogenous DNA as quickly as possible.

### 4.2.5. Phenotype analysis of the *hsdR1::int* and *hsdR2::int* mutants of *C. saccharobutylicum*

One of the characteristic traits of type I RM-systems is that the HsdM and HsdS subunits are transcribed from a different promoter than the HsdR subunit. Therefore, any disruption of the *hsdR* gene should not affect the activity of the MTase (Wilson & Murray 1991) and would give reliable information on the activity and specificity of restrictases. Disruption of either of the restrictase genes reduces the restriction activity of *C. saccharobutylicum* to the point where conjugation with unmethylated plasmid DNA is possible. It also increased the efficiency of conjugation when either an MTase1 or MTase2-methylated plasmid was used, which accords with previous observations. During this work the flippase system was used in an attempt to remove the *ermB* marker from either of the *hsdR::int* mutant strains, which would result in a markerless restrictase-deficient strain, therefore making it accessible for further ClosTron experiments,. However this system did not prove efficient in this experiment and it was noticed that the *ermB* cassette removal is highly troublesome (Heap *et al.*, 2014).

In summary, this work shows it is possible to transconjugate DNA into the *C. saccharobutylicum* NCP 262 using *in vivo* methylation by its own methyltransferases. Experiments with wild type, *hsdR1::int* and *hsdR2::int* recipients conjugated with methylated and non-methylated plasmid DNA suggested that the RM2 system of *C. saccharobutylicum* might play a more significant role in restriction and

methylation. The conjugation experiment with unmethylated plasmid showed higher rates for the *hsdR2::int* recipient, while methylation with pJL2 containing the MTase from RM2 was more effective for conjugation with WT recipients. Disruption of either of the two restrictases was another possibility to increase conjugation rates and does not influence the growth rate of *C. saccharobutylicum*.

## 4.3. Pentose metabolism in solventogenic clostridia

### 4.3.1. Comparison of the xylose and arabinose operons of *C. saccharobutylicum* NCP 262 and *C. acetobutylicum* ATCC 824

Although the genes related to the pentose metabolism of *C. acetobutylicum* were annotated and described (Grimmler *et al.*, 2010; Servinsky *et al.*, 2009; Rodionov *et al.*, 2001), knowledge of the xylose and arabinose metabolism of another important solventogenic microorganism, *C. saccharobutylicum*, was still very limited.

Two gene clusters presumably involved in xylose metabolism have been annotated in the genomes of *Clostridium acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262 (Poehlein *et al.*, 2013; Nölling *et al.*, 2011). In *C. acetobutylicum* they have been described and investigated (Grimmler *et al.*, 2010), therefore this microorganism was chosen as a model with which to compare and study the genes of *C. saccharobutylicum*. However, further studies on two pentose kinase genes, *araK* and *xylB*, have shown that the first of the two investigated clusters of *C. acetobutylicum* is involved in arabinose rather than xylose metabolism.

During analysis of the arabinose and xylose metabolism genes several differences in the operon structure and the gene order between both strains were found (Fig. 4.1). In *C. acetobutylicum*, the transketolase and transaldolase genes are only present in the first operon, involved mainly in arabinose metabolism, while in the *C. saccharobutylicum* genome they are present in both operons, like in *C. beijerinckii* (Wang *et al.*, 2011). Two clusters of *C. saccharobutylicum* genes involved in arabinose and xylose metabolism comprise genes divided into several operons of *C. acetobutylicum* (Fig. 4.1). This would suggest the first cluster of *C. saccharobutylicum* consists of one operon, CSA00700 – CSA00775, containing a transcriptional regulator, L-ribulose-5-

phosphate-4-epimerase, a transaldolase (*tal*), transketolase (*tkt*), sugar kinase and hypothetical protein, and a single L-arabinose isomerase gene, located upstream (CSA00769), while the second cluster forms one operon, CSA03335 - CSA03329, consisting of xylulokinase, a transcriptional regulator or sugar kinase, *tal*, *tkt*, a hypothetical protein, L-fucose isomerase and another hypothetical protein, with a promoter region which we mapped upstream of the CSA03335 gene.

As a part of the analysis of the arabinose and xylose operons of *C. acetobutylicum* and *C. saccharobutylicum*, four promoter regions, located upstream of the L-ribulose-5-phosphate 4-epimerase (*araD*), arabinose kinase (*araK*), and xylulose kinase (*xylB*) of *C. acetobutylicum* and the xylulose kinase (*xylB*) of *C. saccharobutylicum* were recognized and described.

### 4.3.2. Analysis of *C. acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262 *xylB* and *araK* mutants

Xylose and arabinose are the predominant pentoses in plant-derived material, and detailed information on their input in ABE fermentation is of great value for the biotechnological industry nowadays. Unfortunately, knowledge of the pentose metabolism, especially in *C. saccharobutylicum*, is still very limited.

Although two pentose-related operons of *Clostridium acetobutylicum*, CAC1344 – CAC1349 and CAC2612 – CAC2610, have been annotated and described as xylose operons (Grimmler *et al.*, 2010), subsequent experiments suggested the first operon is involved rather in arabinose than in xylose metabolism (Servinsky *et al.*, 2010; Servinsky *et al.*, 2012; Zhang *et al.*, 2012).

In both operons sugar kinases are present and homologous genes in the operons of *C. saccharobutylicum*, named *araK* and *xylB*, were found. In the NCIMB and KEGG databases the *C. acetobutylicum araK* gene is still annotated as *sugar kinase, possible xylulose kinase* (Karp *et al.*, 2005; Kanehisa *et al.*, 2014), and transcriptional data on *C. acetobutylicum* cultivated on xylose revealed that both the *xylB* and *araK* genes are significantly upregulated on xylose (Grimmler *et al.*, 2010).

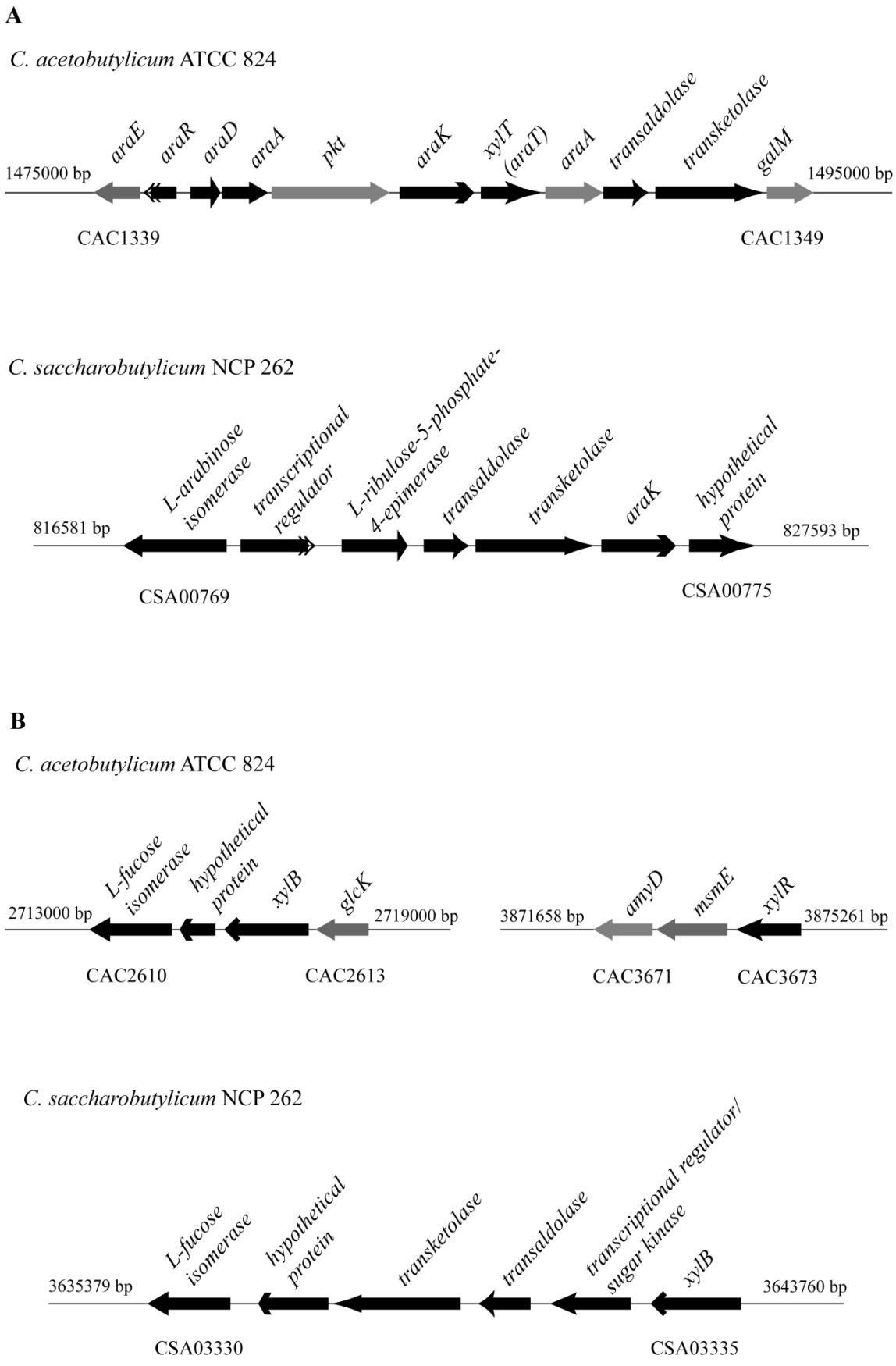



**Figure 4.1.** Schematic comparison of the genes of *Clostridium acetobutylicum* and *Clostridium saccharobutylicum* involved in **(A)** arabinose and **(B)** xylose metabolism.

Homologous genes are black-coloured and indicated with the same arrow-shaped coding. No significant homology was found for grey-coloured genes within the operons of *C. saccharobutylicum*. Gene products are described in Results section 3.3 and in the text; *amyD*, sugar ABC transporter permease; *msmE*, sugar ABC transporter substrate-binding protein. *C. acetobutylicum* gene numbers are given according to the NCBI database, and *C. saccharobutylicum* – to the JGI database.

In both operons sugar kinases are present and homologous genes in the operons of *C. saccharobutylicum*, named *araK* and *xylB,* were found. In the NCIMB and KEGG databases the *C. acetobutylicum araK* gene is still annotated as *sugar kinase, possible xylulose kinase* (Karp *et al.*, 2005; Kanehisa *et al.*, 2014), and transcriptional data on *C. acetobutylicum* cultivated on xylose revealed both the *xylB* and *araK* genes are significantly upregulated on xylose (Grimmler *et al.*, 2010).

In this work it is demonstrated that the *araK* gene in both solventogenic clostridia is involved in arabinose metabolism by showing that the *araK* disruptants in both organisms are not capable of using this pentose as a sole carbon and energy source, and yet they can still grow on xylose. Therefore it is possible, that the upregulation of the CAC1344 – CAC1349 operon in the presence of xylose could be caused by similarities in the pathways that both pentoses are metabolized through (Servinsky *et al.*, 2012; Xiao *et al.* 2011). Mutants in the xylulose kinases (*xylB*) of both strains were unable to grow on xylose exclusively, and showed no growth impairment in an arabinose-supplemented medium. A congruent role of *xylB* homologs in xylose metabolism was suggested for all Firmicutes (Gu *et al.*, 2010).Moreover, neither of the mutations in arabinose or xylose kinases of both strains affected their ability to ferment two other monosaccharides largely present in the plant biomass – fructose and ribose.

Although glucose kinases are known to catalyze the phosphorylation of different hexoses under certain conditions, no such activity of the enzyme towards pentoses was observed (Xiao *et al.*, 2011; Kawai *et al.*, 2005), and regarding the results obtained in this study it is most probable that neither arabinose nor xylose kinase can initiate a phosphorylation reaction of the other pentose.

The intron mutagenesis might cause a polar effect on the expression of the genes located downstream. However, it was possible to restore the growth of *araK::int* on

arabinose and *xylB::int* on xylose in *C. acetobutylicum*, by introducing a functional copy of the native *araK* or *xylB* genes, respectively, under a strong, constitutive clostridial promoter. However, little is known about the regulatory elements applicable for *C. saccharobutylicum*, therefore no reversion of the effects of *araK* and *xylB* disruptions could possibly be caused by the inefficient activity of the chosen promoter.

In this work no growth of the *araK::int* mutant in *C. acetobutylicum* was observed in the medium supplemented with arabinose only.However, results of a recent study on a similar mutant in *C. acetobutylicum* were published suggesting the *araK* disruptant is still capable of growing in a minimal medium with arabinose as the sole carbon source (Zhang *et al.*, 2012). Although the growth profile of *C. acetobutylicum araK::int* in rich (CGM) and minimal medium (MES) was repeatedly examined during this study, comparable results were never observed. It is difficult to discuss these differences, as the information provided by Zhang and colleagues would suggest that for both experiments similar conditions were applied, and yet the discrepancy between the results obtained cannot be negligible. Intron mutations can influence the genes downstream of the integration site (the polar effect) although it is rather unlikely this was the reason for the *araK::int*‘s inability to grow on arabinose since it was possible to reverse the *araK* disruption effect with a functional copy of the gene.

### 4.3.3. Fermentation profile of *xylB::int* and *araK::int* mutants in *C. acetobutylicum*

In general, xylose and arabinose were shown to be a less efficient source of energy, when compared to glucose (Saddler *et al.*, 1983), and it was noted that *C. acetobutylicum* produces higher amounts of solvents while fermenting monosaccharides arising from the *erythro* configuration, i.e. glucose, mannose or arabinose, than from the *threo* configuration, like xylose or galactose (Moo-Young, 1989). Differences between the solvent production by the *C. acetobutylicum* wild type and mutant strains grown on glucose or arabinose were relatively small, while the solvent production on glucose and xylose, or arabinose and xylose, differed significantly. These differences might result from different pathways being used for the glucose (the Embden-Meyerhoff-Parnas pathway), xylose (the pentose phosphate

pathway) and arabinose (the phosphoketolase pathway) metabolism, as suggested before (Servinsky *et al.*, 2012; Xiao *et al.*, 2011).

In *Lactobacillus plantarum* the disruption of both phosphoketolase genes and the introduction of a transketolase resulted in a full switch of xylose metabolism from the phosphoketolase pathway (PKP) to the pentose phosphate pathways (PPP), which resulted in higher fermentation rates (Okano *et al.*, 2009). During this study, however, better solvent production was detected in samples grown on arabinose, and therefore presumably metabolized through the PKP, than on xylose, which is directed to the PPP.

In the context of the efficiency of ABE fermentation it was observed that the functional *araK* gene is nearly as important for high solvent production, as the sugar used as a carbon source. Analysis of samples taken from cultures grown on different sugars revealed that the *araK::int* mutant produced the lowest amounts of solvents when compared to the other strains, regardless of the monosaccharide additive. Significant differences in acetone, butanol and ethanol production and sugar utilization between the wild type strain on glucose or arabinose and xylose proved that xylose is the less preferable carbon and energy source among the three monosaccharides (Ounine *et al.*, 1983; Saddler *et al.*, 1983). Moreover, the results obtained during the fermentation experiment underlined the differences in the behavior of *C. acetobutylicum* in batch and continuous cultures. The wild type strain produced comparable amounts of acids and solvents on xylose and on glucose during the continuous fermentation culture, while in the batch culture the fermentation profile on glucose and on xylose differed significantly.

It is possible that the higher growth rates and better fermentation profile observed on arabinose compared to xylose were caused by the pentose-derived carbon flux through the phosphoketolase pathway, which is regarded as energetically more effective than the pentose phosphate pathway (Servinsky *et al.*, 2012). Pentose metabolism through the phosphoketolase pathway also results in a different acetate:butyrate production ratio, directing the cell metabolism to higher acetone production. It was proposed that the phosphoketolase in *Clostridium acetobutylicum* allows the Embden-Meyerhof-Parnas pathway to be skipped and more carbons to be turned into acetate rather than butyrate, which increases the acetate:butyrate (Ac:Bt) ratio. Indeed, higher Ac:Bt ratios were observed in samples grown on arabinose compared to the samples from cultures grown

on glucose or xylose (Fig. 3.3.12), as was shown before (Servinsky *et al.*, 2012) indicating a different carbon flow.

### 4.3.4. Carbon catabolite repression and diauxic growth on a mixture of D-glucose and L-arabinose of *C. acetobutylicum* ATCC 825

Carbon catabolite repression is a common phenomenon among many bacteria and plays the role of an evolutionary mechanism that enables the microbes to use the most efficient energy source first (Servinsky *et al.*, 2010; Tangney *et al.*, 2003; Brückner *et al.*, 2002; Saier *et al.*, 1996). Glucose is regarded as a preferable carbon source to arabinose or xylose, and arabinose is preferred to xylose (Desai & Rao, 2010) However in *Clostridium saccharoperbutylacetonicum* N1-4 the arabinose metabolism was found to be superior to glucose (Yoshida *et al.*, 2012). Interestingly, no carbon catabolite repression was observed for *C. beijerinckii* in a medium with mixtures of glucose, mannose, arabinose and xylose (Ezeji *et al.*, 2008). The *C. acetobutylicum* wild type shows diauxic growth on a mixture of glucose and xylose (Gu *et al.*, 2010; Grimmler *et al.*, 2010). However the over-expression of three genes of the pentose phosphate pathway: predicted xylose proton symporter (CAC1345), xylose isomerase (CAC2610) and xylose isomerase (CAC2612) in the *ΔglcG* strain, lacking enzyme II of the D-glucose phosphoenolpyruvate-dependent phosphotransferase system, resulted in better utilization of both xylose and arabinose in the presence of glucose in the medium (Xiao *et al.*, 2011; Aristidou & Penttilä, 2000). In this work the glucose catabolite repression in the wild type strain on the mixture of glucose and arabinose was observed with a characteristic short plateau phase when the organism switched from hexose to pentose metabolism, which remains in accordance with previous observations (Xiao *et al.*, 2011).

### 4.3.5. Fermentation of arabinose by *C. acetobutylicum* phosphoketolase (*pkt*) mutants

In the majority of clostridia, arabinose and xylose were traditionally thought to be metabolized through the pentose phosphate pathway (Xiao *et al.*, 2011; Jurgens *et al.*, 2012; Gheshlaghi *et al.*, 2009; Jones & Wood, 1986). *C. acetobutylicum*, however, is one of the few *Clostridium* species possessing in their genomes the phosphoketolase gene (*pkt*), a close homolog of D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase (*xfp*) from *Bifidobacterium lactis* (Meile *et al.*, 2001), and recent findings state that arabinose enters the EMP not through the PPP, but through the PKP, which would explain the advantage in the growth on this pentose compared to xylose (Servinsky *et al.*, 2012). It is still debated whether in *C. acetobutylicum* the *pkt* gene is highly induced only by arabinose (Liu *et al.*, 2012; Servinsky *et al.*, 2010), while in some organisms, mainly in Lactobacilli, the phosphoketolase pathway is used for the xylose heterolactic acid fermentation (Ohara *et al.*, 2006; Bustos *et al.*, 2005; Tanaka *et al.*, 2002; Chaillou *et al.*, 1998). Also, among different fungi species, the PKP is commonly used for xylose metabolism (Papini *et al.*, 2012; Panagiotou *et al.*, 2008; Thykaer & Nielsen, 2007; Sonderedder *et al.*, 2004). In *C. acetobutylicum* the phosphoketolase gene was found to be induced by xylose (Grimmler *et al.*, 2010) and it was suggested that it was a part of the xylose metabolism pathway by contributing up to 40% of the xylose catabolic flux (Liu *et al.*, 2012). High upregulation of the *pkt* gene expression was observed in a diauxic growth experiment on a glucose and xylose mixture, after the cells switched to pentose metabolism (Grimmler *et al.*, 2010). However, during the batch culture fermentation experiment the acetate:butyrate ratios of cultures grown on glucose and on xylose were similar. The carbon flux through the PKP results in higher acetone production, which was not observed in the xylose-supplemented medium.

A traditional view of *C. acetobutylicum* metabolism indicates that the organism is able to metabolize arabinose through the pentose phosphate pathway (Jurgens *et al.*, 2012; Jones & Woods, 1986; Ounine *et al.*, 1983); hence disruption of the phosphoketolase gene should not result in the inability to use this monosaccharide as a sole carbon and energy source. Both ClosTron mutants in the phosphoketolase gene of *C. acetobutylicum* were indeed capable of growing in a medium supplemented only

with arabinose, but they grew much less during the exponential growth phase when compared to the wild type strain, and the growth of the *pkt::2int* mutant was impaired compared to the growth of the *pkt::1int* mutant. Genes in regions located downstream of the *pkt* are expressed from their own promoter, therefore they should not by affected by the possible polar effect and the differences in the growth curves must come exclusively from the partial or complete inactivation of the *pkt* gene. Disparities between both mutant strains could be explained by the fact that *pkt::1int* has the intron inserted into a rear part of the gene, while *pkt::2int* was disrupted close to the middle of the gene, which could more likely affect the active centre of the phosphoketolase.

In the absence of a functional phosphoketolase, the arabinose metabolism of *C. acetobutylicum* should be re-directed through the PPP. Studies on mRNA have shown transaldolase and transketolase, two proteins common to all pentoses metabolism, were found to be expressed in a similar way on both xylose and arabinose (Servinsky *et al.*, 2012). However, the growth of both *pkt* mutants on arabinose was still more effective than the growth of the wild type strain on xylose. The differences between the rates of xylose and arabinose metabolism could therefore be a result not only of a functional phosphoketolase gene, but also of a more complex regulation.

## 4.4. Continuous fermentation of *C. acetobutylicum* ATCC 824 on xylose-supplemented, phosphate-limited medium

Experiments on the small scale continuous fermentation of *C. acetobutylicum* were so far conducted on mixtures of different sugars or glucose only, but never exclusively on xylose (Millat *et al.*, 2013; Schwarz *et al.*, 2012; Grimmler *et al.*, 2011; Napoli *et al.*, 2010).

PH-controlled phosphate-limited continuous fermentation gives information on the metabolism of the *C. acetobutylicum* wild type culture grown on xylose and allows separation of the three phases of growth – acidogenesis, solventogenesis and the transition between those steady states. During acidogenesis mainly acetate and butyrate were measured, while ethanol or butanol levels were relatively low. Concentrations of the solvents increased at the end of the metabolic switch, and concentrations of acids

decreased, but were still detectable during steady-state solventogenic growth (Fischer *et al.*, 2006; Bahl *et al.*, 1982).

Although the dilution rate was identical to a comparable culture grown on glucose (Grimmler *et al.*, 2011) and the profile of growth was very similar, the OD values of the xylose-supplemented culture were 0.5 $OD_{600}$ units lower. A pH-induced metabolic switch took significantly longer than on glucose (Grimmler *et al.*, 2011), and decreased cell density was observed, which reached the lowest values at the last point of the shift, and could have been caused by butanol-induced stress (Tomas *et al.*, 2004; Vollherbst-Schneck *et al.*, 1984, Lin & Blaschek, 1983). The fermentation profiles of butanol and ethanol were comparable to the culture on glucose, while the butyrate and acetate levels were significantly higher, and the acetone concentration was lower. The first xylose-supplemented culture returned more quickly to its OD values from before the switch, after 77.75 h, while the second one needed more time, and reached the starting OD values after 102.25 h. However, both xylose-supplemented cultures needed less time to reach steady solventogenic state, compared to the glucose-supplemented cultures (Grimmler *et al.*, 2011). Although it is commonly known in batch cultures that *C. acetobutylicum* grows much more slowly on xylose compared to glucose, we observed no significant differences in the growth rate during continuous culture on xylose, compared to what has been shown in a study on glucose (Grimmler *et al.*, 2011). Optical density values were similar during acidogenesis and solventogenesis, as in the work of Grimmler, but differed from previous work on glucose (Grupe & Gottschalk, 1992). However, the time required to reaching steady-state solventogenic growth and the transition took significantly longer, which can be explained by less carbon molecules being accessible per gram of sugar compared to glucose. Previous work on *C. acetobutylicum* wild type batch cultures on xylose suggested that the strain grows more slowly and shows an impaired solvent production profile, compared to the same experiment conducted on glucose, and the differences were very significant. In the continuous culture experiment, however, both growth and fermentation profiles of the wild type strains grown on glucose or xylose were comparable, and differed only insignificantly.

Xylose concentration increased during the two last pH shifts of the transition from acidogenic to solventogenic growth and it was strongly related to higher butanol production, which did not occur on glucose (Grimmler *et al.*, 2011). Butanol has an

inhibitory effect on sugar analog uptake (Bowles & Ellefson, 1985; Moreira *et al.*, 1981) and was shown to have a greater effect on xylose than on glucose (Ounine *et al.*, 1985). Higher xylose concentrations correlated with a higher acid uptake, which could also be the cause of lower pentose utilization. After the metabolic shift from acidogenesis to solventogenesis the concentrations of acids rose again and reached stable values during solventogenesis, yet the xylose concentration in the medium did not decrease. The highest solvent concentration was observed at the last pH shift (pH 4.5) of the metabolic switch, whereas on glucose it was reached during steady-state solventogenic growth (Grimmler *et al.*, 2011). It was suggested previously there are several main triggers for the switch from acid to solvent production in *C. acetobutylicum*: high carboxylic acid concentrations (over 28 mM), a decrease in the pH of the medium and limitations in iron, nitrogen or phosphate (Dürre *et al.*, 2002; Girbal & Soucaille, 1998; Bahl *et al.*, 1995; Terraciano & Kashket, 1986). In continuous culture the shift occurs exclusively due to the pH change (Grupe & Gottschalk, 1992), and in this experiment high concentrations of acids (over 60 mM of butyrate and 40 mM of acetate) present during the steady-state acidogenic growth did not induce solvent production.

Along with the analysis of the phenotype the microarray data from six different time points (Fig. 3.4.2) were investigated. All the genes significantly regulated during the representative continuous culture experiment were divided into four groups. In the first group 141 genes upregulated during steady-state acidogenic growth were described and the second group was composed of 95 genes upregulated only during steady-state solventogenic growth. The third and fourth groups gather together genes non-regulated during these metabolic states. Exactly 226 genes found to be induced during the transition from acidogenesis to solventogenesis created the third group and 24 genes downregulated during the transition – the fourth group. The groups were analyzed by searching for the transcriptional pattern of the genes involved in the general- and butanol-induced stress response, sporulation, solvent and acid production, cellulosome synthesis, and the metabolism of sugars, and amino and fatty acids. Chosen genes were described in more detail and discussed. The genes and the binary logarithm values of their expression are listed in the tables.

### 4.4.1. Dehydrogenases and thiolases

The dehydrogenases and thiolases of *C. acetobutylicum* play an important role in acid and solvent production. An important group of these genes is located on the pSOL1 megaplasmid, forming the *sol*-operon, and is induced at the onset of solventogenesis (Jones *et al.*, 2008; Zhao *et al.*, 2005; Alsaker *et al.*, 2005).

Surprisingly, throughout the whole continuous culture experiment on xylose, no significant regulation of 3-hydroxybutyryl-CoA dehydrogenases (CAC2708 and CAC2009) or enoyl-CoA hydratases (CAC2712, CAC2012 and CAC2016) was observed (Tab. 4.4.1). The *Hbd* (CAC2708) and *crt* (CAC2712) genes play an important role in butyrate and butanol production (Lehmann & Lütke-Eversloh, 2011), and were observed to be highly induced during steady-state acidogenic growth in the continuous culture experiment on glucose (Grimmler *et al.*, 2011) and yet remained non-regulated on xylose. However, the butanol amounts produced on xylose did not differ significantly from the concentration obtained on glucose. This might suggest the genes were probably constitutively induced through the whole fermentation process.

Under butanol stress induced in an acidogenic chemostat culture of *C. acetobutylicum* no transcription of genes for solvent formation was observed (Schwarz *et al.*, 2012), in contrast to previous reports, suggesting that the expression of solvent genes is dose-dependent (Alsaker et. al., 2004; Tomas *et al.*, 2004; Alsaker *et al.*, 2009). No induction caused by the presence of butanol in the medium was observed, and it was assumed that their expression was caused by the change in the pH of the culture and the influence of other genes.

*Clostridium acetobutylicum* has two paralogs of alcohol/aldehyde dehydrogenases, *aad* (CAP0162) and *adhE* (CAP0035) and two paralogs of thiolases, *thlA* (CAC2873) and *thlB* (CAP0078). The *aad* gene is located in the *sol*-operon (CAP0162 – CAP0164; Fischer *et al.*, 1993) and was strongly repressed during steady-state acidogenic growth; its expression rose through the metabolic switch and the genes were upregulated during steady-state solventogenic growth. A comparable regulation pattern was observed on glucose (Grimmler *et al.*, 2011).

| | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAP0025 | *pdc* | 0.76 | 0.68 | 1.28 | 1.59 | 2.131 | -1.44 | -0.76 |
| CAP0161 | *solR* | -0.22 | 0.63 | 1.18 | 0.18 | 0.97 | | 0.22 |
| CAP0162 | *aad* | -3.32 | -1.42 | 1.67 | 1.88 | 4.01 | -0.34 | 3.32 |
| CAP0163 | *ctfA* | -4.07 | -1.65 | 1.20 | 1.60 | 3.38 | -0.85 | 4.07 |
| CAP0164 | *ctfB* | -3.50 | -1.36 | 1.01 | 1.33 | 2.56 | -1.40 | 3.50 |
| CAP0165 | *adc* | -0.16 | 0.44 | 3.22 | 3.05 | 2.69 | 0.34 | 0.16 |
| CAP0078 | *thlB* | 0.31 | 0.58 | 1.65 | 2.13 | 1.27 | -1.77 | -0.31 |
| CAP0079 | *thlR* | -0.15 | 1.56 | 1.5 | 2.01 | 1.98 | -0.52 | 0.15 |

**Table 4.4.1.** Comparison of the transcription profile of the genes involved in solvent formation: thiolases and dehydrogenases in a continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to a transcription level in steady state solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate the transcription levels of the genes were below the detection limit.

A significant reduction in the *aad* expression at the last pH shift of the transition from acidogenesis to solventogenesis (pH 4.5) correlates with a strong decrease in the acetate and butyrate concentrations and an increase in the solvent concentrations detectable in the culture. A similar expression pattern can be observed for the whole *sol* operon (CAP0162 – CAP0164), also encoding two butyrate-acetoacetate CoA-transferase subunits, *ctfA* and *ctfB*. Acetoacetate decarboxylase (*adc*, CAP0165) was significantly upregulated during the metabolic switch, but not during steady-state acidogenic growth (Tab. 4.4.1). The *aad-ctfA-ctfB-adc* locus is believed to be repressed by SolR (encoded by CAP0161; Nair *et al.*, 1999). However, no correlation was observed between the expression of *sol*-locus and *solR*gene (Tab. 4 4.1). More recent work has been published on this particular regulation, emphasising that the expression of genes responsible for solvent production is activated by a multivalent transcription factor Spo0A through a region upstream of the *sol*-operon promoter, containing three imperfect repeats and a

putative Spo0A-binding motif (Thormann *et al.*, 2002), which stands in agreement with the results shown in this work.

*ThlB* expression increased only during the switch, and during steady-state acidogenic and solventogenic growth it remained non-regulated (Tab. 4.4.1). The transcription levels of *thlA* and *adhE* did not change throughout the entire continuous fermentation experiment. In *C. acetobutylicum* ATCC 824 grown in a phosphate-limited, glucose-supplemented continuous culture, the *adhE* and *thlA* genes were upregulated during acidogenic growth, but repressed and non-regulated, respectively, during steady solventogenic growth (Grimmler *et al.*, 2011). Although in the batch culture of the *buk* (butyrate kinase) mutant, the *adhE* and *sol*-operon expression patterns were found to be similar (Zhao *et al.*, 2005) and in the experiments done by Grimmler and colleagues (2011) to be antagonistic, here no correlation between them was found.

In *Clostridium acetobutylicum* DSM 792 the *thlA* gene is identical with the one of *C. acetobutylicum* ATCC 824 (Stim-Herndon *et al.*, 1995), and the second thiolase (*thlB*) forms an operon with a presumed transcriptional regulator, *thlR*, and a hypothetical protein, *thlC*. Contrary to what has been observed for *thlA*, the thiolase B was poorly expressed in a batch culture during acid and solvent production by the DSM 792 strain (Winzer *et al.*, 2000). In the continuous culture however, the *thlB* was non-regulated during acidogenesis and solventogenesis, and its expression increased significantly during the transition between these metabolic states. While it was suggested that the *thlR* in *C. acetobutylicum* DSM 792 might repress the transcription of the other two genes in the operon (*thlB* and *thlC*), no such dependence was observed in a continuous culture on xylose.

The pyruvate decarboxylase gene (*pdc*, CAP0025) was induced only during the transition from acidogenic to solventogenic growth and repressed at its end (Tab. 4.4.1). A similar regulation pattern was shown in a glucose-supplemented continuous culture (Grimmler *et al.* 2011) and the association of *pdc* with solvent production was suggested before in a batch culture experiment (Zhao *et al.*, 2005). Pyruvate decarboxylase is responsible for the decarboxylation of pyruvic acid to acetaldehyde and carbon dioxide, the first then presumably being reduced to ethanol by alcohol dehydrogenase, *aad* (CAP0162), and hence their simultaneous activation during the metabolic switch to solventogenesis.

At the onset of solventogenesis, larger amounts of acetyl-CoA are required for the acid activation. Among the genes involved in acetyl coenzyme A metabolism, the *pflBA* (CAC0980 – CAC0981) operon was significantly upregulated during the transition from acidogenic to solventogenic growth, while the *citBC* operon (CAC0971 – CAC0972) was strongly inhibited, which accords with results of previous experiment on glucose (Grimmler *et al.*, 2011).

### 4.4.2. Redox sensors and NAD-biosynthesis

In prokaryotic metabolism, $NAD^+$ is involved in the redox mechanisms, and carries electrons from one reaction to another. Its hydrogenated form, NADH, is used as a reducing agent to donate electrons, while $NAD^+$ accepts electrons from other molecules, and therefore acts as an oxidizing agent.

In *C. Acetobutylicum* the solvent production requires high amounts of NADH, and NAD-biosynthesis genes, *nadABC*, are highly induced in solventogenesis (Jones *et al.*, 2008). Here, the upregulation of *nadABC* was observed exclusively during the transition from acidogenic to solventogenic growth; subsequently the genes were repressed at the last pH shift and remained non-regulated during solventogenesis.

Changes in cellular NADH/$NAD^+$ levels and the redox state are controlled by the regulators belonging to the Rex family, and in *C. acetobutylicum* the CAC2713 gene codes for Rex, a redox-sensing transcriptional repressor (Wietzke & Bahl, 2012). Of the many genes regulated by Rex (Wang Q*et al.*, 2013) almost none of them were significantly regulated in a xylose continuous culture, except for two clusters, CAP0162 – CAP0164 and CAC0014 – CAC0015. Regulation of the CAC0014 – CAC0015 genes was not observed in a glucose continuous culture (Grimmler *et al.*, 2011). On xylose however it was strongly induced in the presence of high butyrate concentrations, during acidogenesis and the metabolic shift to solventogenesis, but not regulated when the butanol level increased (Tab. 4.4.2), which would correlate with the results obtained in a pH-controlled batch culture (Wang Q *et al.*, 2013). The *aad-ctfA-ctfB* locus expression was induced briefly during the shift and showed no association with butyrate levels (Tab. 4.4.1).

| | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAC0014 | Aminotransferase | 4.93 | 7.48 | 2.41 | 7.80 | 3.52 | 0.95 | -4.93 |
| CAC0015 | *serA* | 3.83 | 7.54 | 2.01 | 7.48 | 3.29 | 0.60 | -3.83 |
| CAC1023 | *nadC* | 0.69 | 0.72 | 1.79 | 2.01 | 2.116 | -2.13 | -0.69 |
| CAC1024 | *nadB* | 0.74 | 0.70 | 1.75 | 1.85 | 1.781 | -2.06 | -0.74 |
| CAC1025 | *nadA* | 1.08 | 0.70 | 1.35 | 1.42 | 1.851 | -1.77 | -1.08 |

**Table 4.4.2.** Comparison of the transcription profiles of genes involved in $NAD^+$-NADH biosynthesis in a continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to the transcription level in steady state solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate the transcription levels of the genes were below the detection limit.*SerA*, D-3-phosphoglycerate dehydrogenase; *nadC*, Nicotinate-nucleotide pyrophosphorylase; *nadB*, L-aspartate oxidase; *nadA*, quinolinate synthetase.

### 4.4.3. Stress response

Continuous fermentation puts the culture under high acid- or solvent-mediated stress. GroES/EL is expressed as a result of butanol stress and is believed to be a key component of the butanol-stress response, along with the ClpC and DnaKJ proteins (Schwarz *et al.*, 2012; Alsaker *et al.*, 2005; Tomas *et al.*, 2004) However some studies show their importance also in the butyrate stress-response at low pH values (Alsaker & Papoutsakis, 2005; Tomas *et al.*, 2003 b). The stress response genes of *B. subtilis* were described and divided into IV groups (Schumann *et al.*, 2002). Based on this classification, in *C. acetobutylicum* the HrcA-GrpE-DnaK-DnaJ, GroES-GroEL and HtpG were classified as class I heat-shock proteins (*hsp*) genes, HrcA-dependent, and *clp* family genes as class III *hps* genes, CtsR-dependent (Alsaker *et al.*, 2009; Tomas *et al.*, 2004). Genes of both classes were identified in *C. acetobutylicum*, with a CtsR-

binding region upstream of ClpC (Tomas *et al.*, 2004). The expression of stress response genes was shown to be dose-dependent (Tomas *et al.*, 2004).

The increased expression of GroES and GroEL proteins may serve to stabilize biosynthetic machinery (Tomas *et al.*, 2004), and overexpression of the heat shock proteins increases solvent tolerance not only in *C. acetobutylicum* (Mann *et al.*, 2012; Tomas *et al.*, 2003 a), but also in *Lactococcus lactis*, *L. paracasei* (Desmond *et al.*, 2004) and *Saccharomyces cerevisiae* (Vianna *et al.*, 2008). The *C. acetobutylicum* strain overexpressing butanol-derived stress proteins GroESL was able to metabolize glucose even under high solvent concentrations in batch cultures (Tomas *et al.*, 2003 a).

In batch cultures it was also observed that high butyrate stress (over 50 mM) leads to the activation of purine metabolism and ATP synthesis coupled to proton transport genes, while riboflavin, cobalamin biosynthesis and ribosome-related genes seem to be silenced (Alsaker *et al.*, 2009). Butyrate (over 40 mM) or butanol (over 60 mM) stress in the pH-controlled batch culture was observed to lead to the activation of genes targeted by HrcA and CtsR, grouped in clusters CAC1280 – CAC1283 (*hrcA-grpE-dnaK-dnaJ*), CAC2703 – CAC2704 (*groEL/ES*), CAC3189 – CAC3192 (*clpC*, ATP:guanido-phosphotransferase, hypothetical protein, *ctsR* transcriptional regulator) and CAC3315 (*htpG*) (Wang Q*et al.*, 2013). However, during the analysis of the transcription profile of the continuous culture it was observed that, regardless of high butyrate concentrations during acidogenesis (over 60 mM) and at the beginning of the switch to solventogenesis (pH 5.5, 60mM butyrate), neither of these genes was regulated, except for the *groEL/ES* operon (Tab. 4.4.3). They were upregulated significantly at lower pH values (pH 5.2 – 4.9), until the butanol concentrations rose, which is consistent with the results obtained previously on glucose (Grimmler *et al.*, 2011). The *htpG* (CAC3315) transcription level did not change throughout the entire continuous culture experiment, and was only downregulated at the last pH shift (pH 4.5) of the transition from acidogenesis to solventogenesis (Tab. 4.4.3). It is possible therefore that the HrcA and CtsR regulation might be connected to the metabolic switch between the two physiological states of the cell, acidogenesis and solventogenesis. However, regarding the constant inflow of nutrients and outflow of fermentation products the stress in a continuous culture might be much lower compared to any batch culture. Moreover, it was suggested previously that the exposure of *C. acetobutylicum* to butanol-derived stress results in the upregulation of major stress response genes,

including *clpC, htpG,groEL/ES, dnaKJ*and *hrcA* (Alsaker *et al.*, 2005; Tomas *et al.*, 2004; Alsaker *et al.*, 2004). However, studies done on a glucose continuous culture with no additional butanol supplementation clearly showed these genes to have been expressed even before the butanol was produced (Grimmler *et al.*, 2011). The stress response genes were upregulated during the transition from acidogenesis to solventogenesis, when the pH decreased to below 5.8, which is consistent with what is known about their role in the adaptation of a cell to changed conditions. This kind of stress response to low pH values is common among lactic acid bacteria (Kang *et al.*, 2010; Frees *et al.*, 2003). The results presented in this thesis confirm these data, and the stress response genes, especially CAC3189 – CAC3192 and the *groEL/ES* operons were highly induced when the pH values decreased below 5.5, even though no significant amount of butanol was yet measured (2 g/l).

| | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAC1280 | *hrcA* | 0.24 | 1.01 | 2.15 | 1.46 | 1.14 | -1.79 | -0.24 |
| CAC1281 | *grpE* | -0.06 | 0.95 | 1.93 | 1.53 | 1.056 | -2.04 | 0.06 |
| CAC1282 | *dnaK* | -0.24 | 0.72 | 1.87 | 1.50 | 0.90 | -2.37 | 0.24 |
| CAC1283 | *dnaJ* | 0.23 | 0.70 | 2.14 | 1.77 | 1.348 | -1.11 | -0.23 |
| CAC3189 | *clpC* | 0.54 | 1.56 | 3.29 | 2.83 | 1.954 | -1.37 | -0.54 |
| CAC3190 | *yacI* | 0.52 | 1.47 | 3.04 | 2.71 | 1.835 | -1.13 | -0.52 |
| CAC3191 | *yacH* | 0.59 | 1.37 | 3.16 | 2.84 | 1.573 | -1.36 | -0.59 |
| CAC3192 | *ctsR* | -0.60 | 0.75 | 2.19 | 2.28 | 0.844 | -1.54 | 0.60 |
| CAC3714 | *hsp18* | -0.98 | 0.78 | 3.34 | 2.96 | 1.601 | -2.54 | 0.98 |
| | | | | | | | | |
| CAC2703 | *groEL* | -0.05 | 1.15 | 2.76 | 2.01 | 0.97 | -2.45 | 0.05 |
| CAC2704 | *groES* | -1.29 | 1.75 | 2.52 | 2.37 | 1.05 | -2.03 | 1.29 |
| CAC3315 | *htpG* | 0.66 | 0.89 | 1.68 | 0.44 | 0.09 | -2.06 | -0.66 |

**Table 4.4.3.** Comparison of the transcription profiles of genes involved in a stress-response of a *C. acetobutylicum*in continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to the transcription level in steady state

solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate that the transcription levels of the genes were below the detection limit.

Most of the genes became repressed at the end of the metabolic shift from acidogenesis to solventogenesis, when butanol production reached its highest level (Tab. 4.4.3), and were non-regulated during steady-state solventogenic growth (Sauer & Dürre, 1995), like during acidogenesis. Previous experiments on the glucose-supplemented continuous culture showed the heat shock proteins were observed to be induced shortly before upregulation of the solvent genes (Pich *et al.*, 1990), and a similar relation between both gene groups was observed for a culture grown on xylose.

As observed previously, glycerol-3-phosphate dehydrogenase (*glpA*, CAC1322) was strongly upregulated during steady-state solventogenic growth, but not during acidogenesis (Alsaker *et al.*, 2009).

### 4.4.4. Purine biosynthesis

The genes involved in purine metabolism, grouped in several operons, *purE* – *purC* (CAC1390 – CAC1391), *purF* – *purM* – *purN* – *purH* (CAC1392 – CAC1395) and *purD* (CAC1396), are presumably involved in resistance to acid stress, strongly induced in the presence of high butyrate concentrations and generally upregulated during the exponential growth phase in batch cultures (Wang Q*et al.*, 2013; Jones *et al.*, 2008). In this study the activation of *pur* genes was observed not in the presence of butyrate, but at the onset of solventogenesis. No strong connection to any acid stress response was noticed, as the transcription level of *pur* genes increased only during the last pH shifts of the switch between acidogenesis and solventogenesis (pH 4.9 – 4.5). The *purQL* (CAC1655) was not regulated throughout the entire continuous culture experiment (Tab. 4.4.4). Although it was reported previously that prolonged acetate and butyrate stress increases the expression of purine-biosynthesis genes in batch cultures (Alsaker *et al.*, 2009), no regulation of the *pur* operons in the presence of high butyrate concentrations during steady acidogenic growth was observed in the continuous culture.

Although the *pur* genes were upregulated during the transition from acidogenic to solventogenic growth at pH 4.9 – 4.5, when the butyrate levels decreased and butanol concentration was increasing, they were still non-regulated during steady-state solventogenic growth. It is therefore possible their expression is rather pH-dependent.

In batch cultures some of the pyrimidine biosynthesis genes are known to be downregulated under acetate, butyrate and butanol-derived stresses (Alsaker *et al.*, 2009). Here, in the continuous culture experiment the *pyr* genes, *pyrDFIB* (CAC2650, CAC2652 – CAC2654) were significantly upregulated at pH 4.9 and 4.5, when butyrate and acetate levels had already decreased.

| | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAC1390 | *purE* | -0.04 | 0.89 | -0.059 | 2.90 | 2.71 | 1.56 | 0.04 |
| CAC1391 | *purC* | -0.44 | 0.82 | -0.26 | 2.70 | 2.52 | 1.50 | 0.44 |
| CAC1392 | *purF* | -0.42 | 0.81 | -0.28 | 2.48 | 2.24 | 1.34 | 0.42 |
| CAC1393 | *purM* | -0.81 | 0.71 | -0.44 | 2.72 | 2.27 | 1.54 | 0.81 |
| CAC1394 | *purN* | -0.70 | 0.67 | -0.41 | 2.76 | 2.62 | 1.70 | 0.70 |
| CAC1395 | *purH* | -0.24 | 0.70 | -0.14 | 2.60 | 2.93 | 1.84 | 0.24 |
| CAC1396 | *purD* | -0.93 | 0.40 | -0.67 | 1.86 | 1.76 | 1.40 | 0.93 |

**Table 4.4.4.** Comparison of the transcription profiles of genes involved in purine metabolism (*pur* genes) in the continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to the transcription level in steady state solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate that the transcription levels of the genes were below the detection limit.

### 4.4.5. Sporulation genes and transcriptional regulators

In the continuous culture experiment on xylose, several clusters of sporulation genes were observed to be induced during steady acidogenic growth and during the switch between acidogenesis and solventogenesis. Two of the clusters, CAC1695 – CAC1696 *(sigE and sigG*) and CAC2086 – CAC2093 (stage III sporulation proteins, *SpoIII*) were significantly upregulated during steady-state acidogenic growth and at the beginning of the transition from acidogenesis to solventogenesis.

The so-called 'master regulator', the *spo0A* gene (CAC2071) was non-regulated during steady-state acidogenic or solventogenic growth and was upregulated only during the metabolic shift between these steady states. Interestingly, in the glucose-supplemented continuous culture, no significant regulation of *spo0A* was noticed (Grimmler *et al.*, 2011). However, Grimmler and colleagues described two AbrB transcriptional regulators (CAC0310 and CAC3647), also playing an important role in the initiation of sporulation and solventogenesis (Scotcher *et al.*, 2005; Strauch *et al.*, 1989), as being upregulated during acidogenic growth, but silenced during the transition to solventogenesis. Although their data would suggest the AbrB regulator could be involved in the regulation of the metabolic switch from acid to solvent production, no expression of those genes was observed in this experiment.

Spo0A is believed to induce genes involved in solvent production by binding upstream of the promoter region of *sol*-operon (Thormann *et al.*, 2002). In the xylose-supplemented continuous culture no positive correlation between the expression of the *spo0A* and *sol*-operon or the sporulation genes was observed, showing their independence from Spo0A regulation (Tab. 4.4.5). However, many sporulation genes: namely stage III sporulation proteins (CAC2086 – CAC2093), sigma sporulation factors (CAC1689, CAC1694 – CAC1697) and spore coat proteins (CAC1337 – CAC1338) were upregulated during acidogenesis and through the transition from acidogenic to solventogenic growth, yet downregulated during solventogenesis. In *B. subtilis*, *spo0A* positively regulates the expression of sporulation genes, including the *sigE* operon (*C. acetobutylicum* ortholog CAC1694 – CAC1695), *sigF* operon (CAC2306 – CAC2308) and *spoIIE* (CAC3205), and similar regulation has been observed *in C. acetobutylicum* pH-controlled batch fermentations (Jones *et al.*, 2008; Alsaker & Papoutsakis, 2005) and in this study. In the experiment on a glucose-supplemented

continuous culture, the positive regulation of the *sigF*-operon, involved in the early sporulation process, was observed during the transition from acidogenesis to solventogenesis (Grimmler *et al.*, 2011).

A *spo0A*-deficient strain shows no expression of *sol*-operon genes, but high expression of the genes related to the motility and chemotaxis of the cell (Tomas *et al.*, 2003 b). In *Bacillus subtilis* the Spo0A protein negatively regulates flagellar (*fli*-operons), chemotaxis and DNA replication genes (Molle *et al.*, 2003), and similar behavior was found in the *C. acetobutylicum buk* mutant (Zhao *et al.*, 2005). In the xylose-supplemented continuous culture experiment no significant changes in the regulation of *fli*or chemotaxis genes were observed. Experiments on continuous cultures showed that. regardless of the lack of *spo0A* regulation, only a few, if any, flagellar genes were described as repressed or induced (Schwarz *et al.*, 2012; Grimmler *et al.*, 2011), while in batch cultures these genes were negatively regulated by *spo0A* expression (Alsaker *et al.*, 2009; Zhao *et al.*, 2005; Alsaker *et al.*, 2004, Tomas *et al.*, 2003 b). Another batch culture study showed chemotaxis and motility genes are mostly affected by acetate stress (but not by butyrate or butanol), even though under all three stress conditions the *spo0A* expression was comparable (Alsaker *et al.*, 2009). These results would suggest that the conditions applied in continuous cultures are not triggering somehow any changes in cell motility or chemotaxis gene expression, possibly due to the constant inflow of nutrients, which does not require an active carbon and energy source search by the organism.

It seems the Spo0A upregulation may be caused by the changes of pH in the culture and not the butyrate, acetone or butanol accumulation, as the *spoA* is non-regulated in both steady-state gworths, acidogenic and solventogenic, which tends to confirm the suggestion that the expression of sporulation genes remains mostly unaffected by metabolite stress (Alsaker *et al.*, 2009).

Although several groups of sporulation genes were upregulated in this experiment, in multiple continuous culture experiments, conducted by our research group on glucose, no regulation of sporulation genes was ever observed.

### 4.4.6. Cellulosome formation and polysaccharide metabolism

*C. acetobutylicum* possesses a cluster of genes coding for a putative cellulosome (CAC0910 – CAC0919) and although the organism is not known to metabolize cellulose, it produces higher extracellular cellulolytic activities when cultured on lichenan or xylose, compared to cellobiose and glucose (López-Contreras, 2003 and 2004). The cellulosome cluster represents a high homology to the operon described in *C. cellulolyticum* (Bélaich*et al.*, 1997; Gal *et al.*, 1997) and *C. cellulovorans* (Tamaru *et al.*, 2000). In the continuous culture it was strongly repressed during the steady acidogenic phase and through the metabolic shift to solventogenesis (Tab. 4.4.5 B). As in the continuous culture experiment on glucose, the cellulosome genes were upregulated during solventogenic growth (Grimmler *et al.*, 2011).

Since the cellulosome-coding genes were strongly downregulated during acidogenesis and upregulated during solventogenesis and there was a constant inflow of the carbon source and nutrients, it was suggested that low pH values or a high butanol concentration could be the trigger for these genes' expression. These results agree with what has been observed previously (Grimmler *et al.*, 2011), although it seemed that, in a xylose continuous culture, higher upregulation of putative cellulosome genes during steady-state solventogenic growth was noticed. These observations are consistent with the results of López-Contreras (2004 and 2003), where a significantly larger amount of Cel-related proteins was detected on xylose than on glucose.

Additionally, other genes related to the metabolism of polysaccharides were found to be upregulated during steady-state solventogenic growth or at the end of the transition from acidogenesis to solventogenesis (Tab. 4.4.5 A).

In the continuous culture on glucose, genes involved in alternative substrate utilization were upregulated during solventogenesis (Grimmler *et al.*, 20101), and a similar pattern was observed for the xylose-supplemented culture. Pectate lyases (CAP0056, CAC0574, CAC1968, and CAC3387) are involved in the decomposition of pectate to oligosaccharides and play an important role in maceration of plant tissues (Ouattara *et al.*, 2011). Xylanases (CAP0053 – CAP0054, CAC2396 and CAC3017) are responsible for the degradation of xylan components in hemicelluloses to xylose (Beg *et al.*, 2001).

| **A** | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAP0053 | Xylanase *(xynb)* | -1.66 | -2.47 | -1.56 | -1.03 | -0.46 | 1.73 | 1.66 |
| CAP0054 | Xylanase/chitin deacetylase family enzyme | -1.65 | -2.72 | -1.48 | -1.40 | -0.56 | 1.70 | 1.65 |
| CAC2396 | Predicted xylanase/chitin deacetylase | -1.76 | -1.34 | -1.09 | -1.52 | -2.31 | -0.37 | 1.76 |
| CAC3017 | Predicted xylanase/chitin deacetylase | -2.15 | -2.19 | -1.15 | -1.10 | -1.68 | -1.19 | 2.15 |
| CAP0056 | Pectate lyase *(pell)* | -2.00 | -0.75 | -1.25 | -1.11 | 0.35 | 0.76 | 2.00 |
| CAC0574 | Pectate lyase | -2.78 | -2.91 | -2.77 | -2.29 | -0.57 | 0.36 | 2.78 |
| CAC1968 | Pectate lyase related enzyme | -2.22 | -2.11 | -2.21 | | -1.17 | 0.52 | 2.22 |
| CAC3387 | Pectate lyase | -2.73 | -2.82 | -1.79 | -1.68 | -0.90 | -0.29 | 2.73 |

| **B** | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAC0561 | *celE* | -3.29 | -3.26 | -2.61 | -2.79 | -0.7 | -0.61 | 3.29 |
| CAC0826 | *celE* | -1.77 | -2.22 | -2.16 | -3.08 | -1.62 | -0.25 | 1.77 |
| CAC0910 | *cbp* | -6.72 | -6.57 | -5.38 | -5.11 | -2.21 | -0.41 | 6.72 |
| CAC9011 | *celF* | -6.51 | -6.51 | -5.24 | -5.21 | -2.15 | -0.52 | 6.51 |
| CAC0912 | *celA* | -7.38 | -6.32 | -5.41 | -5.32 | -2.00 | -0.68 | 7.38 |
| CAC0913 | *celH* | -6.72 | -5.97 | -5.16 | -4.92 | -1.62 | -0.65 | 6.72 |
| CAC0914 | *ccp* | -6.80 | -6.57 | -5.27 | -5.18 | -1.68 | -0.58 | 6.80 |
| CAC0915 | *engA* | -5.47 | -5.88 | -4.47 | -4.74 | -1.58 | -0.66 | 5.47 |
| CAC0916 | *celG* | -6.38 | -6.64 | -5.35 | -5.18 | -1.85 | -0.81 | 6.38 |
| CAC0917 | *celL* | -6.38 | | -5.51 | | -2.29 | -0.78 | 6.38 |
| CAC0918 | *manA* | -6.27 | | -5.72 | -5.54 | -2.53 | -0.89 | 6.27 |
| CAC0919 | sialidase | -5.80 | -6.01 | -4.92 | -5.04 | -2.86 | -0.92 | 5.80 |
| CAC3469 | endoglucanase | -3.99 | -4.21 | -4.12 | -4.15 | -2.19 | 0.97 | 3.99 |

**Table 4.4.5.** Comparison of the transcription profiles of **(A)** genes involved in complex sugar metabolism and **(B)** genes coding for putative cellulosome (*cel* genes) in the continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the

metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to the transcription level in steady state solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate the transcription levels of the genes were below the detection limit.

Pectate lyases and xylanases remained downregulated during steady-state acidogenic growth and at the beginning of the metabolic switch, and became upregulated during the solventogenesis.

The α-amylase (CAP0168) gene, located on the pSOL1 megaplasmid, was also upregulated at the end of the transition from acidogenic to solventogenic growth, but remained non-regulated in both acidogenesis and solventogenesis. As in the previous experiment (Grimmler *et al.* 2011), upregulation of glycerol-3-phosphate dehydrogenase (*glpA*, CAC1322) and galactokinase (*galK*, CAC2959) was observed during solventogenic growth.

Since a simple carbon source (xylose) is not a limiting factor in our experiment, it is possible that the triggers of the expression of cellulosome and polysaccharide metabolism genes are similar – a low pH or solvent production.

### 4.4.7. Fatty acid and amino acid biosynthesis

In *B. subtilis* fatty acid synthesis is controlled by *spo0A* and precedes the endospore formation (Pedrido *et al.*, 2013). In batch cultures several operons containing genes of fatty acid metabolism were regulated (CAC2007 – CAC2019, CAC3568 – CAC3580 or *fab* genes) (Schwarz *et al.*, 2012; Tomas *et al.*, 2004; Zhao *et al.*, 2005). In the continuous culture no regulation of the fatty acid metabolism was observed. The 4'-phosphopantetheinyl transferase (*acpS*, CAC0489) is involved in the transfer of acyl fatty acid intermediates as a part of the fatty acid and lipid biosynthesis, and was upregulated during steady-state acidogenic growth (Tab. 4.4.6). The experiment on a glucose-supplemented continuous culture showed upregulation of the genes of fatty acid

biosynthesis (CAC1988 – CAC2019) during solventogenesis (Grimmler *et al.*, 2011). Although the transcription of the whole cluster was detectable during the entire xylose-supplemented continuous culture, no regulation of either of the genes was observed.

Different genes involved in amino acid biosynthesis were regulated during the entire xylose-supplemented continuous culture experiment. Serine is required for cysteine biosynthesis (Schaffer *et al.*, 2002), and the genes involved in serine biosynthesis, CAC0014 and CAC0015, were strongly upregulated during steady-state acidogenic growth and during the switch between acidogenesis and solventogenesis, until the pH of the medium decreased to a value of 4.5. This regulation pattern correlated with a marked decrease in butyrate concentrations in the medium and an increase in butanol and acetone levels. In batch cultures, however, cysteine was observed to be synthesized as a response to acetate and butyrate stress (Alsaker *et al.*, 2009).

The *aro* locus (CAC0892 – CAC0899) is coding for genes of chorismate biosynthesis, an intermediate for aromatic amino acid biosynthesis, namely phenylalanine, tyrosine and tryptophan. The *trp* locus (CAC3157 – CAC3163) encodes genes responsible for converting chorismate to tryptophan. The regulation pattern of the *aro* genes (CAC0892 – CAC0898) resembled the regulation of *pur* genes (Tab. 4.4.4).

Genes remained non-regulated during steady-state acidogenic growth and solventogenic growth and become upregulated only during part of the transition between these metabolic states, which would contrast with previous observations made in batch cultures, demonstrating the induction of the *aro* locus on butyrate and butanol stress, while *trp* genes were previously found to be silenced by all solvent and acid stresses (Alsaker *et al.*, 2009).

Similarly, the two *arg* genes (CAC2388 – CAC2389) described as being induced on butyrate and inhibited on butanol (Alsaker *et al.*, 2009), remained downregulated during nearly the entire continuous culture on xylose, except for steady-state solventogenic growth.

| | | pH 5.7 (A) | pH 5.5 | pH 5.2 | pH 4.9 | pH 4.7 | pH 4.5 | pH 4.5 (S) |
|---|---|---|---|---|---|---|---|---|
| CAC0014 | aminotransferase | 4.93 | 7.48 | 2.41 | 7.8 | 3.5 | 1 | -4.93 |
| CAC0015 | *serA* | 3.83 | 7.54 | 2.01 | 7.5 | 3.3 | 0.6 | -3.83 |
| CAC0892 | 3-deoxy-7-phosphoheptulonate synthase | 0.78 | 1.85 | 0.38 | 2.87 | 3.735 | 1.218 | -0.78 |
| CAC0893 | Prephenate dehydrogenase | 0.58 | 2.10 | 0.57 | 2.66 | 3.521 | 1.036 | -0.58 |
| CAC0894 | *aroB* | 0.26 | 1.93 | 0.43 | 2.7 | 3.589 | 1.232 | -0.26 |
| CAC0895 | *aroA* | 0.67 | 1.45 | 0.32 | 2.24 | 3.459 | 0.888 | -0.67 |
| CAC0896 | *aroC* | 0.27 | 0.96 | 0.07 | 2.01 | 3.134 | 0.901 | -0.27 |
| CAC0897 | *aroF* | -0.06 | 1.44 | 0.26 | 2.23 | 3.383 | 1.403 | 0.06 |
| CAC0898 | *aroK* | -0.53 | 0.99 | -0.27 | 1.81 | 2.646 | 1.029 | 0.53 |
| CAC2388 | *argD* | -1.99 | -0.724 | -2.77 | -4.158 | -0.898 | -1.89 | 1.99 |
| CAC2389 | *argB* | -2.06 | -2.149 | -3.22 | | -1.86 | -1.87 | 2.06 |
| CAC2390 | *argC* | -0.41 | -0.56 | -2.22 | -4.00 | -0.40 | -2.00 | 0.41 |
| CAC2391 | *argJ* | -0.34 | -0.22 | -1.86 | -3.60 | -0.08 | -1.80 | 0.34 |
| CAC0489 | *acpS* | 2.47 | 1.32 | 2.28 | 1.5 | 0.5 | 1.9 | -2.47 |

**Table 4.4.6.** Comparison of the transcription profiles of genes involved in amino and fatty acid biosynthesis and metabolism in the continuous culture in a xylose-supplemented, phosphate-limited medium. RNA samples for the transcription analysis were taken during acidogenesis (pH 5.7; A), solventogenesis (pH 4.5; S) and the metabolic shift (pH 5.5 – 4.5), and the expression ratios are shown as logarithms to the basis of 2. The RNA collected at the end of solventogenesis was used as a reference. Positive values indicate the upregulation of a gene and negative values indicate the downregulation of a gene, compared to the transcription level in steady state solventogenic growth (pH 4.5; S). Genes were regarded as significantly upregulated or repressed when their $\log_2$ expression ratios were ≤ 1.6 or -1.6 ≥, respectively. Blank cells indicate that the transcription levels of the genes were below the detection limit.

## 5. Summary

Solventogenic clostridia are able to use various pentoses to produce energy in the process called Acetone-Butanol-Ethanol fermentation. *Clostridium acetobutylicum* ATCC 824 can ferment two common plant-derived pentoses, xylose and arabinose. However, their utilization is blocked by the presence of glucose, due to the carbon catabolite repression. Arabinose, compared to xylose, was more efficient as a carbon and energy source and its fermentation resulted in much higher solvent yields. The acetone-to-butanol ratio of cultures grown on arabinose was much higher compared to cultures grown on glucose or xylose, and it was suggested that the carbon flux in *C. acetobutylicum* grown on arabinose is directed through the phosphoketolase pathway, which would also explain the differences between the growth and fermentation profiles on both pentoses. Disruption of the phosphoketolase gene resulted in an impaired growth of the mutant strain on arabinose during the entire exponential phase of growth, and most probably it directed the carbon flux through the pentose phosphate pathway.

In the metabolism of xylose and arabinose in *C. acetobutylicum* ATCC 824 and *C. saccharobutylicum* NCP 262, two important genes, encoding xylulose kinase (*xylB*) and arabinose kinase (*araK)* are involved. In this work intron mutants of the kinase genes were created in both solventogenic clostridia, and additionally the *araK* and *xylB* genes of *C. acetobutylicum* were deleted through the allelic exchange (ACE). *Clostridium* strains without functional arabinose kinase were not able to utilize arabinose as a sole carbon and energy source, but it was possible to complement the intron-derived mutation effect by introducing a functional copy of the *araK* gene under the promoter of the ferredoxin gene in *C. acetobutylicum*. Similarly, strains with disrupted xylulose kinase were not able to grow on xylose, but introducing a functional copy of the *xylB* gene under the ferredoxin promoter in the *C. acetobutylicum* intron mutant reversed this phenotype. A *C. acetobutylicum ΔaraK ΔxylB* strain was incapable of fermenting either xylose or arabinose.

Although in batch cultures *C. acetobutylicum* showed an impaired growth and fermentation profile on xylose compared to glucose, no such observation was made in the continuous culture experiment. DNA microarray-based transcriptional analysis of

the genes regulated during the two metabolic states of the cell: acidogenesis and solventogenesis, and during the transition between both, revealed many similarities in the regulation of important cellular processes between glucose- (previous work from this group) and xylose-supplemented continuous cultures. Among genes regulated in a similar manner on xylose and glucose, the genes involved in solvent production (*sol*-operon, *thlB* and *aad*), cellulosome formation (*cel* genes), stress response (*groEL/ES*, *dnaKJ*, *grpE* or *hrcA*) and the tricarboxylic acid cycle (*citBC*) were described. However, the transcriptional pattern of some genes differed between both cultures. Significant regulation of genes involved in amino acid biosynthesis (*aro* and *trp* genes), redox state (*nadABC*) and sporulation (*spoIII* genes, *sigF*genes, *sigG*, *sigE* or *spo0A*) was observed on xylose, but not on glucose.

Solvent production depends not only on productive sugar fermentation, but also on the effective energy conversion directed by, *inter alia*, the Rnf complex. This complex is common among anaerobic Gram-positive species and is used for energy conservation by converting the transfer of electrons from reduced ferredoxin to NAD into a sodium or proton gradient. In *C. beijerinckii*, however, disruption of the *rnfC* gene, coding for the first gene of the *rnf* cluster and a subunit of the complex, resulted in no visible changes in the growth curve or glucose consumption rates when compared to the wild type strain.

Additionally, this study focused on the need for an accessible clean deletion system for solventogenic clostridia and the creation of an *upp*-based system for *C. acetobutylicum*. Despite the construction of several variants of deletion vectors and the application of different conditions to improve the integration rate, no clean deletion of the *upp* gene was achieved, although in several attempts the plasmid integrants, representing the first recombination event in the clean deletion method, were observed.

Furthermore, for the genetic manipulations of *C. saccharobutylicum* NCP262, a methylation system and tri-parental conjugation approach were created. The pJL1 and pJL2 methylation plasmids, based on the native methyltransferase genes from the restriction-modification systems RM1 and RM2 of the bacterium, were constructed and it was shown that it is possible to transconjugate DNA to *C. saccharobutylicum* using *in vivo* methylation by its own methyltransferases. Experiments with three clostridial recipients: the wild type strain and mutants in the restriction subunits (*hsdR*) of the

restriction-modification systems RM1 and RM2, with and without *in vivo* methylation of donor DNA suggested the RM2 of *C. saccharobutylicum* has a higher significance for restriction, as well as for methylation. Conjugation with an unmethylated plasmid showed higher rates for the *hsdR2*-disrupted recipient, while methylation with pJL2 containing the methyltransferase from RM2 was more effective for conjugation with wild type recipients.

## 6. Zusammenfassung

Solventogene Clostridien sind in der Lage, verschiedene Pentosen für ihre Energiegewinnung mittels Aceton-Butanol-Ethanol Fermentation zu verwenden.

*Clostridium acetobutylicum* ATCC 824 fermentiert zwei pflanzliche Pentosen, D-Xylose und L-Arabinose, jedoch wird deren Verwertung in Anwesenheit von Glukose aufgrund der Kohlenstoff-Katabolitrepression unterdrückt. Arabinose scheint im Vergleich zu Xylose die effizientere Kohlenstoff- und Energiequelle zu sein, deren Fermentation in einer signifikant erhöhten Lösungsmittelproduktion resultierte. *C. acetobutylicum* ATCC 824 zeigte zudem ein deutlich erhöhtes Aceton/Butanol Verhältnis bei Wachstum auf Arabinose verglichen zu Kulturen mit Glukose oder Xylose. Der Katabolismus von Arabinose scheint über eine Phosphoketolase zu erfolgen. Die unterschiedlichen Abbauwege der beiden Pentosen erklären die Unterschiede im Wachstums- und Fermentationsverhalten auf diesen Substraten. Die Inaktivierung des Phosphoketolase-Gens führte während der gesamten exponentiellen Wachstumsphase zu einem verminderten Wachstum der Mutante auf Arabinose, was die Bedeutung dieses Gens für den Arabinose-Abbau belegte.

An der Verwertung von Xylose und Arabinose sind sowohl in *C. acetobutylicum* ATCC 824, als auch in *C. saccharobutylicum* NCP 262 zwei entscheidende Gene beteiligt, die für Xylulose-Kinase (*xylB*) und Arabinose-Kinase (*araK*) kodieren. Im Rahmen dieser Arbeit konnten Intron-Mutanten dieser Kinasegene in beiden solventogenen Clostridien erstellt werden. Darüber hinaus wurden die Gene *araK* und *xylB* in *C. acetobutylicum* ebenfalls durch *allelic exchange* (ACE) deletiert. *Clostridium*-Stämme können ohne eine funktionsfähige Arabinose-Kinase Arabinose nicht mehr als einzige Kohlenstoff- und Energiequelle nutzen. Dieser Phänotyp konnte durch eine funktionsfähige Kopie des *araK*-Gens unter Kontrolle des Ferredoxin-Gen-Promotors in *C. acetobutylicum* komplementiert werden. Analoges gilt für Stämme mit einer inaktivierten Xylulose-Kinase. Mit der Inaktivierung des *xyl*B-Gens war ein Wachstum auf Xylose ausgeschlossen. Auch dieser Phänotyp konnte durch eine funktionsfähige Kopie des *xylB*-Gens unter Kontrolle des Ferredoxin-Gen-Promotors in einer der entsprechenden Intron-Mutante komplementiert werden. Der Stamm *C. acetobutylicum ΔaraKΔxylB* konnte beide Pentosen, Xylose und Arabinose, nicht fermentieren. Obwohl

*C. acetobutylicum* in Batch-Kulturen mit Xylose ein vermindertes Wachstums- und Fermentationsverhalten zeigte, konnte dieser Effekt in kontinuierlichen Kulturen nicht bestätigt werden.

Transkriptionsanalysen mit DNA-Microarrays zeigten bei kontinuierlichen Kulturen auf Glucose (frühere Arbeiten der Arbeitsgruppe) und Xylose ein ähnliches Expressionsverhalten von Genen, die während der Acido- und Solventogenese sowie während der Transitionsphase reguliert waren. Diese ähnliche Genexpression konnte für Gene der Lösungsmittelproduktion (*sol*-operon, *thlB* and *aad*), der Cellulosomen-Bildung (*cel* genes), der Stress-Antwort (*groEL/ES, dnaKJ, grpE* or *hrcA*) und der Citrat-Bildung (*citBC*) nachgewiesen werden. Andererseits unterschied sich das Transkriptionsprofil einiger Gene in beiden Kulturen deutlich voneinander. So zeigten Gene, die an der Aminosäurebiosynthese (*aro* und *trp*), des Redoxzustands (*nadABC*) oder der Sporulation (*spoIII* genes, *sigF* genes, *sigG, sigE* or *spo0A*) beteiligt sind, eine signifikante Regulation auf Xylose, nicht aber auf Glukose.

Die Lösungsmittelproduktion ist nicht allein von einer produktiven Zucker-Fermentation, sondern ebenfalls von einer effektiven Energie-Umwandlung etwa durch den Rnf Komplex abhängig. Unter anaeroben Gram-positiven Bakterien ist der Rnf Komplex weit verbreitet. Er produziert einen Natrium- bzw. Protonengradienten bei der Übertragung von Elektronen von reduziertem Ferredoxin auf $NAD^+$. In *C. beijerinckii* zeigte jedoch die Inaktivierung des *rnfC* Gens, welches „das erste Gen des *rnf*-Genclusters dieses Bakteriums darstellt, keine ersichtlichen Veränderungen in der Wachstumskurve oder der Glukose-Verbrauchsrate im Vergleich zum Wildtyp-Stamm.

Die vorliegende Arbeit beschäftigte sich außerdem mit der Entwicklung eines einfachen *clean deletion* Systems für solventogene Clostridien durch Gegenselektion mit einem auf *upp* basierenden System in *C. acetobutylicum*. Trotz der Konstruktion etlicher Varianten von Deletionsvektoren und der Verwendung unterschiedlicher Bedingungen, um die Integrationsrate zu verbessern, konnte keine Deletion des *upp* Gens nachgewiesen werden. In einigen Versuchen konnte jedoch die Vektor-Integration, der erste Schritt in der Deletionsmethodik, beobachtet werden.

Ferner wurde für die genetische Manipulation von *C. saccharobutylicum* NCP 262 ein Methylierungssystem und ein tri-parentales Konjugationsverfahren entwickelt. Es wurden die Methylierungsvektoren pJL1 und pJL2 mit den Genen für die nativen

Methyltransferasen der Restriktions/Modifikations-Systeme RM1 und RM2 des Bakteriums konstruiert. Es konnte gezeigt werden, dass infolge der *in vivo* Methylierung durch die eigenen Methyltransferasen DNA in *C. saccharobutylicum* konjugiert werden kann. Drei *Clostridium* Rezipienten-Stämme, der Wildtyp-Stamm und Mutanten in den Restriktions-Untereinheiten (*hsdR*) der Restriktions/ModifikationsSysteme RM1 bzw. RM2, wurden mit und ohne *in vivo* Methylierung der Donor-DNA zum Einsatz gebracht. Die Experimente zeigten, dass RM2 von *C. saccharobutylicum* offenbar eine größere Bedeutung sowohl für Restriktion als auch Methylierung besitzt: Die Konjugation mit unmethyliertem Plasmid zeigte eine höhere Rate mit dem *hsdR2 knock-out*-Rezipienten und die Methylierung mit der auf pJL2 kodierten Methyltransferase von RM2 war deutlich effektiver bei der Konjugation mit dem Wildtyp-Rezipienten.

## 7. Publications derived from this work

Published:

Development of an *in vivo* methylation system for the *solventogen Clostridium saccharobutylicum* NCP 262 and analysis of two endonuclease mutants.

Lesiak, Justyna Maria; Liebl, Wolfgang and Ehrenreich, Armin

J Biotechnol. 2014 Aug 1;188C:97-99. doi: 10.1016/j.jbiotec.2014.07.005.

Corresponding authors:

Dr. Armin Ehrenreich; Lehrstuhl für Mikrobiologie TU-München; Emil-Ramann Str. 4; D-85350 Freising. Email: aehrenr@mikro.biologie.tu-muenchen.de

Prof. Dr. Wolfgang Liebl; Lehrstuhl für Mikrobiologie TU-München; Emil-Ramann Str. 4; D-85350 Freising. Email: wliebl@wzw.tum.de

### Abstract

Restriction-Modification (RM) systems are a common defence mechanism against foreign DNA among bacteria. For this reason they often represent a limiting factor for the development of transformation and conjugation protocols. The genome of the biotechnologically important solventogen *Clostridium saccharobutylicum* NCP 262 contains two operons coding for genes of presumed type I RM systems belonging to the families A and C. We developed an efficient triparental mating system to transfer DNA to *C. saccharobutylicum* by conjugation, which includes an *in vivo* methylation of the donor DNA. Furthermore we describe increased rates of conjugation in knock-out mutants of the restrictase subunits of both RM systems.

In preparation:

Investigation of xylose and arabinose metabolism in solventogenic clostridia.

Lesiak, Justyna Maria; Liebl, Wolfgang and Ehrenreich, Armin

## 8. Literature


1. Alegre MT, Rodriguez MC, Mesas JM (2004). Transformation of *Lactobacillus plantarum* by electroporation with *in vitro* modified plasmid DNA. FEMS Microbiology Letters 241 (2004), 73-77.
2. Alsaker KV, Papoutsakis ET (2005). Transcriptional program of early sporulation and stationary-phase events in *Clostridium acetobutylicum*. J Bacteriol. 2005 Oct;187(20):7103-18.
3. Alsaker KV, Paredes C, Papoutsakis ET (2009). Metabolite stress and tolerance in the production of biofuels and chemicals: gene-expression-based systems analysis of butanol, butyrate, and acetate stresses in the anaerobe *Clostridium acetobutylicum*. Biotechnol Bioeng. 2010 Apr 15;105(6):1131-47.
4. Alsaker KV, Spitzer TR, Papoutsakis ET (2004). Transcriptional analysis of *spo0A* overexpression in *Clostridium acetobutylicum* and its effect on the cell's response to butanol stress. J Bacteriol. 2004 Apr;186(7):1959-71.
5. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402.
6. Amador-Noguez D, Feng XJ, Fan J, Roquet N, Rabitz H, Rabinowitz JD (2010). Systems-level metabolic flux profiling elucidates a complete, bifurcated tricarboxylic acid cycle in *Clostridium acetobutylicum*. J Bacteriol. 2010 Sep;192(17):4452-61.
7. Aristidou A, Penttilä M (2000). Metabolic engineering applications to renewable resource utilization. Curr Opin Biotechnol. 2000 Apr;11(2):187-98.
8. Bahl H, Andersch W, Gottschalk G (1982). Continuous production of acetone and butanol by *Clostridium acetobutylicum* in a two-stage phosphate-limited chemostat. Eur J Appl Microbiol Biotechnol; 15: 201–205.
9. Bahl H, Müller H, Behrens S, Joseph H, Narberhaus F (1995). Expression of heat shock genes in *Clostridium acetobutylicum*. FEMS Microbiol Rev. 1995 Oct;17(3):341-8.
10. Baker TA, Wickner SH (1992). Genetics and enzymology of DNA replication in *Escherichia coli*. Annu Rev Genet. 1992;26:447-77.

11. Beg QK, Kapoor M, Mahajan L, Hoondal GS (2001). Microbial xylanases and their industrial applications: a review. Appl Microbiol Biotechnol. 2001 Aug;56(3-4):326-38.
12. Bélaich JP, Tardif C, Bélaich A, Gaudin C (1997). The cellulolytic system of *Clostridium cellulolyticum*. J Biotechnol. 1997 Sep 16;57(1-3):3-14.
13. Berezina OV, Brandt A, Yarotsky S, Schwarz WH, Zverlov VV (2009). Isolation of a new butanol-producing *Clostridium* strain: high level of hemicellulosic activity and structure of solventogenesis genes of a new *Clostridium saccharobutylicum* isolate. Syst Appl Microbiol. 2009 Oct;32(7):449-59.
14. Biegel E, Müller V (2010). Bacterial Na+-translocating ferredoxin: NAD+ oxidoreductase. Proc Natl Acad Sci U S A. 2010 Oct 19;107(42):18138-42.
15. Biegel E, Schmidt S, González JM, Müller V (2011). Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. Cell Mol Life Sci. 2011 Feb;68(4):613-34.
16. Biegel E, Schmidt S, Müller V (2009). Genetic, immunological and biochemical evidence for a Rnf complex in the acetogen *Acetobacterium woodii*. Environ Microbiol. 2009 Jun;11(6):1438-43.
17. Bitan-Banin G, Ortenberg R, Mevarech M (2003). Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. J Bacteriol. 2003 Feb;185(3):772-8.
18. Blackwood JK, Rzechorzek NJ, Bray SM, Maman JD, Pellegrini L, Robinson NP. (2013). End-resection at DNA double-strand breaks in the three domains of life. Biochem Soc Trans. 2013 Feb 1;41(1):314-20.
19. Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL (2012). Domain enhanced lookup time accelerated BLAST. Biology Direct 7:12.
20. Bowles LK, Ellefson WL (1985). Effects of butanol on *Clostridium acetobutylicum*. Appl Environ Microbiol. 1985 Nov;50(5):1165-70.
21. Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976 May 7;72:248-54.
22. Branduardi P, Longo V, Berterame NM, Rossi G, Porro D (2013). A novel pathway to produce butanol and isobutanol in *Saccharomyces cerevisiae*. Biotechnol Biofuels. 2013 May 4;6(1):68.

23. Brückner R, Titgemeyer F (2002). Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. FEMS Microbiol Lett. 209:141-8.

24. Bunker RD, Bulloch EM, Dickson JM, Loomes KM, Baker EN (2013). Structure and function of human xylulokinase, an enzyme with important roles in carbohydrate metabolism. J Biol Chem. 2013 Jan 18;288(3):1643-52.

25. Bustos G, Moldes AB, Cruz JM, Domínguez JM (2005). Influence of the metabolism pathway on lactic acid production from hemicellulosic trimming vine shoot hydrolyzates using *Lactobacillus pentosus*. Biotechnol Prog. 2005 May-Jun;21(3):793-8.

26. Cao Y, Xian M, Zou H, Zhang H (2013). Metabolic engineering of *Escherichia coli* for the production of xylonate. PLoS One. 2013 Jul 5;8(7).

27. Carr AM, Lambert S (2013). Replication stress-induced genome instability: the dark side of replication maintenance by homologous recombination. J Mol Biol. 2013 Nov 29;425(23):4733-44.

28. Cartman ST, Heap JT, Kuehne SA, Cockayne A, Minton NP (2010). The emergence of 'hypervirulence' in *Clostridium difficile*. Int J Med Microbiol. 2010 Aug;300(6):387-95.

29. Chaillou S, Bor YC, Batt CA, Postma PW, Pouwels PH (1998). Molecular cloning and functional expression in *Lactobacillus plantarum* 80 of *xylT*, encoding the D-xylose-H+ symporter of *Lactobacillus brevis*. Appl Environ Microbiol. 1998 Dec;64(12):4720-8.

30. Cornillot E, Croux C, Soucaille P (1997). Physical and genetic map of the *Clostridium acetobutylicum* ATCC 824 chromosome. J Bacteriol. 179:7426-34.

31. Crown SB, Indurthi DC, Ahn WS, Choi J, Papoutsakis ET, Antoniewicz MR (2011). Resolving the TCA cycle and pentose-phosphate pathway of *Clostridium acetobutylicum* ATCC 824: Isotopomer analysis, in vitro activities and expression analysis. Biotechnol J. 2011 Mar;6(3):300-5.

32. Curatti L, Brown CS, Ludden PW, Rubio LM (2005). Genes required for rapid expression of nitrogenase activity in *Azotobacter vinelandii*. Proc Natl Acad Sci U S A. 2005 May 3;102(18):6291-6.

33. Dahms AS (1974). 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochem Biophys Res Commun. 1974 Oct 23;60(4):1433-9.

34. Davis TO 1998. Regulation of botulinum toxin complex formation in *Clostridium botulinum* type A NCTC 2916. Ph.D. thesis, Open University.

35. Desai TA, Rao CV (2010), Regulation of arabinose and xylose metabolism in *Escherichia coli*. Appl Environ Microbiol. 76:1524-32.

36. Desmond C, Fitzgerald GF, Stanton C, Ross RP (2004). Improved stress tolerance of GroESL-overproducing *Lactococcus lactis* and probiotic *Lactobacillus paracasei* NFBC 338. Appl Environ Microbiol. 2004 Oct;70(10):5929-36.

37. Dixon R, Kahn D (2004). Genetic regulation of biological nitrogen fixation. Nat Rev Microbiol. 2004 Aug;2(8):621-31.

38. Dredge R, Radloff SE, van Dyk JS, Pletschke BI (2011). Lime pretreatment of sugar beet pulp and evaluation of synergy between ArfA, ManA and XynA from *Clostridium cellulovorans* on the pretreated substrate. 3 Biotech. 2011 Oct;1(3):151-159.

39. Dürre P (2007). Biobutanol: an attractive biofuel. Biotechnol J. 2007 Dec;2(12):1525-34.

40. Dürre P, Bahl H (1996). Microbial Production of Acetone/Butanol/Isopropanol, in: Biotechnology 2nd Edn. Vol. 6 (M. Roehr, Ed.), VCH, Weinheim, 1996, pp. 229-268.

41. Dürre P, Böhringer M, Nakotte S, Schaffer S, Thormann K, Zickner B (2002). Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*. J Mol Microbiol Biotechnol. 2002 May;4(3):295-300.

42. Ehsaan M, (2013). Development & Exploitation of Gene Tools for Metabolic Engineering in Saccharolytic Clostridia. PhD thesis, The University of Nottingham.

43. Ezeji TC, Qureshi N, Blaschek HP (2007). Production of acetone butanol (AB) from liquefied corn starch, a commercial substrate, using *Clostridium beijerinckii* coupled with product recovery by gas stripping. J Ind Microbiol Biotechnol. 34:771-7.

44. Fabret C, Ehrlich SD, Noirot P (2002). A new mutation delivery system for genome-scale approaches in *Bacillus subtilis*. Mol Microbiol. 2002 Oct;46(1):25-36.

45. Feng X, Zhao H (2013). Investigating xylose metabolism in recombinant *Saccharomyces cerevisiae* via $^{13}C$ metabolic flux analysis. Microb Cell Fact. 2013 Nov 18;12:114.

46. Fischer RJ, Helms J, Dürre P (1993). Cloning, sequencing, and molecular analysis of the sol operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. J Bacteriol. 1993 Nov;175(21):6959-69.

47. Fischer RJ, Oehmcke S, Meyer U, Mix M, Schwarz K, Fiedler T, Bahl H (2006). Transcription of the *pst* operon of *Clostridium acetobutylicum* is dependent on phosphate concentration and pH. J Bacteriol. 2006 Aug;188(15):5469-78.

48. Fitzpatrick JM, Johnston DA, Williams GW, Williams DJ, Freeman TC, Dunne DW, Hoffmann KF (2005). An oligonucleotide microarray for transcriptome analysis of *Schistosoma mansoni* and its application/use to investigate gender-associated gene expression. Mol Biochem Parasitol. 2005 May;141(1):1-13.

49. Frees D, Vogensen FK, Ingmer H (2003). Identification of proteins induced at low pH in *Lactococcus lactis*. Int J Food Microbiol 2003; 87: 293–300.

50. Freier D, Mothershed CP, Wiegel J (1988). Characterization of *Clostridium thermocellum* JW20. Appl Environ Microbiol. 1988 Jan;54(1):204-211.

51. Gal L, Pages S, Gaudin C, Belaich A, Reverbel-Leroy C, Tardif C, Belaich JP (1997). Characterization of the cellulolytic complex (cellulosome) produced by *Clostridium cellulolyticum*. Appl Environ Microbiol. 1997 Mar;63(3):903-9.

52. Garufi G, Butler E, Missiakas D (2008). ESAT-6-like protein secretion in *Bacillus anthracis*. J Bacteriol. 2008 Nov;190(21):7004-11.

53. Gheshlaghi R, Scharer JM, Moo-Young M, Chou CP (2009). Metabolic pathways of clostridia for producing butanol. Biotechnol Adv. 2009 Nov-Dec;27(6):764-81.

54. Girbal L, Soucaille P (1998). Regulation of solvent production in *Clostridium acetobutylicum*. Trends Biotechnol. 16:11 – 16.

55. Green EM, Bennett GN (1996). Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC 824. Appl Biochem Biotechnol. 1996 Spring;57-58:213-21.

56. Green EM, Boynton ZL, Harris LM, Rudolph FB, Papoutsakis ET, Bennett GN (1996). Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. 1996 Aug;142 ( Pt 8):2079-86.

57. Griffiths AJF, Miller JH, Suzuki DT, *et al.*(2000). An Introduction to Genetic Analysis. 7th edition. New York: W. H. Freeman; 2000.

58. Grimmler C, Held C, Liebl W, Ehrenreich A (2010). Transcriptional analysis of catabolite repression in *Clostridium acetobutylicum* growing on mixtures of D-glucose and D-xylose. J Biotechnol. 150:315-23.

59. Grimmler C, Holger J, Krauße D, Fischer R-J, Bahl H, Dürre P. Liebl W, Ehrenreich A (2011). Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. J Mol Microbiol Biotechnol. 20:1-15.

60. Grupe H, Gottschalk G (1992). Physiological events in *Clostridium acetobutylicum* during the shift from acidogenesis to solventogenesis in continuous culture and presentation of a model for shift induction. Appl Environ Microbiol. 1992 Dec;58(12):3896-902.

61. Gu Y, Ding Y, Ren C, Sun Z, Rodionov DA, Zhang W, Yang S, Yang C, Jiang W (2010). Reconstruction of xylose utilization pathway and regulons in Firmicutes. BMC Genomics. 2010 Apr 21;11:255

62. Guss AM, Olson DG, Caiazza NC, Lynd LR (2012). Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. Biotechnol. Biofuels 5(1):30.

63. Haft DH, Payne SH, Selengut JD (2012). Archaeosortases and exosortases are widely distributed systems linking membrane transit with posttranslational modification. J Bacteriol. 2012 Jan;194(1):36-48.

64. Hartmanis MG, Gatenbeck S (1984). Intermediary metabolism in *Clostridium acetobutylicum*: levels of enzymes involved in the formation of acetate and butyrate. Appl Environ Microbiol. 1984 Jun; 47(6):1277-83.

65. Heap JT, Ehsaan M, Cooksley CM, Ng YK, Cartman ST, Winzer K, Minton NP (2012). Integration of DNA into bacterial chromosomes from plasmids without a counter-selection marker. Nucleic Acids Res. 2012 Apr;40(8):e59.

66. Heap JT, Kuehne SA, Ehsaan M, Cartman ST, Cooksley CM, Scott JC, Minton NP (2010). The ClosTron: Mutagenesis in *Clostridium* refined and streamlined. J Microbiol Methods. 2010 Jan;80(1):49-55.

67. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007). The ClosTron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods. 2007 Sep;70(3):452-64.

68. Heap JT, Pennington OJ, Cartman ST, Minton NP (2009). A modular system for *Clostridium* shuttle plasmids. J Microbiol Methods. 2009 Jul;78(1):79-85.

69. Heap JT, Theys J, Ehsaan M, Kubiak AM, Ludwig D, Paesmans K, van Mellaert L, Knox R, Kuehne SA, Lambin P, Minton NP (2014). Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors *in vivo*. Oncotarget, Jan 12 2014

70. Held C, (2013). Funktionelle Genomanalyse von *Clostridium ljungdahlii*. PhD thesis, Technische Universität München.

71. http://img.jgi.doe.gov

72. http://www.clostron.com

73. http://www.ncbi.nlm.nih.gov

74. http://www.softberry.com

75. http://www.typei-rm.info

76. Hu S, Zheng H, Gu Y, Zhao J, Zhang W, Yang Y, Wang S, Zhao G, Yang S, Jiang W (2011). Comparative genomic and transcriptomic analysis revealed genetic characteristics related to solvent formation and xylose utilization in *Clostridium acetobutylicum* EA 2018. BMC Genomics. 2011 Feb 2;12:93.

77. Janscak P, Dryden DTF, Firman K (1998). Analysis of the subunit assembly of the type IC restriction-modification enzyme *EcoR*124I. Nucleic Acids Res. 26:4439-45.

78. Jin L, Zhang H, Chen L, Yang C, Yang S, Jiang W, Gu Y (2014). Combined overexpression of genes involved in phosphate pentose pathway enables enhanced d-xylose utilization by *Clostridium acetobutylicum*. J Biotechnol. 2014 Jan 9. pii: S0168-1656(14)00003-0.

79. Johnson JL, Toth J, Santiwatanakul S, Chen JS (1997). Cultures of "*Clostridium acetobutylicum*" from various collections comprise *Clostridium acetobutylicum, Clostridium beijerinckii*, and two other distinct types based on DNA-DNA reassociation. Int J Syst Bacteriol.

80. Jones DT, Woods DR (1986). Acetone-butanol fermentation revisited. Microbiol Rev. 1986 Dec;50(4):484-524.

81. Jones SW, Paredes CJ, Tracy B, Cheng N, Sillers R, Senger RS, Papoutsakis ET (2008). The transcriptional program underlying the physiology of clostridial sporulation. Genome Biol. 2008;9(7):R114.

82. Jouanneau Y, Jeong HS, Hugo N, Meyer C, Willison JC (1998). Overexpression in *Escherichia coli* of the *rnf* genes from *Rhodobacter capsulatus*-characterization

of two membrane-bound iron-sulfur proteins. Eur J Biochem. 1998 Jan 15;251(1-2):54-64.

83. Jurgens G, Survase S, Berezina O, Sklavounos E, Linnekoski J, Kurkijärvi A, Väkevä M, van Heiningen A, Granström T (2012). Butanol production from lignocellulosics. Biotechnol Lett. 2012 Aug;34(8):1415-34.

84. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014). Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014 Jan;42(Database issue):D199-205.

85. Kang KH, Lee JS, Yoo M, Jin I (2010). The influence of HtrA expression on the growth of *Streptococcus* mutans during acid stress. Mol Cells. 2010 Mar;29(3):297-304.

86. Kannan P, Cowan GM, Daniel AS, Gann AA, Murray NE (1989). Conservation of organization in the specificity polypeptides of two families of type I restriction enzymes. J Mol Biol. 209:335-44.

87. Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahrén D, Tsoka S, Darzentas N, Kunin V, López-Bigas N (2005). Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res. 2005 Oct 24;33(19):6083-9.

88. Karpenshif Y, Bernstein KA (2012). From yeast to mammals: recent advances in genetic control of homologous recombination. DNA Repair (Amst). 2012 Oct 1;11(10):781-8.

89. Kawai S, Mukai T, Mori S, Mikami B, Murata K (2005). Hypothesis: structures, evolution, and ancestor of glucose kinases in the hexokinase family. J Biosci Bioeng. 2005 Apr;99(4):320-30.

90. Keis S, Bennett CF, Ward VK, Jones DT (1995). Taxonomy and phylogeny of industrial solvent-producing clostridia. Int J Syst Bacteriol. 45:693-705.

91. Kelman LM, Kelman Z (2004). Multiple origins of replication in archaea. Trends Microbiol. 2004 Sep;12(9):399-401.

92. Kim BH, Gadd GM (2008). Bacterial physiology and metabolism. Cambridge University Press, Cambridge.

93. Kraußе D, (2012). Einfluss der Redox-Balance auf die Solventogenese in *Clostridium acetobutylicum* ATCC 824. PhD thesis, Technische Universität München.

94. Kristich CJ, Manias DA, Dunny GM (2005). Development of a method for markerless genetic exchange in *Enterococcus faecalis* and its use in construction of a *srtA* mutant. Appl Environ Microbiol. 2005 Oct;71(10):5837-49.

95. Kuehne SA, Minton NP. (2012) ClosTron-mediated engineering of *Clostridium*. Bioengineered 3: p. 247-254.

96. Kulik EM, Bickle TA (1996). Regulation of the activity of the type IC *Eco*R124I Restriction Enzyme. J.Mol.Biol. 264, 891-906.

97. Kumagai H, Fujiwara T, Matsubara H, Saeki K (1997) Membrane localization, topology, and mutual stabilization of the *rnfABC* gene products in *Rhodobacter capsulatus* and implications for a new family of energy-coupling NADH oxidoreductases. Biochemistry 36:5509–5521

98. Lee J, Seo E, Kweon DH, Park K, Jin YS (2009). Fermentation of rice bran and defatted rice bran for butanol-5 production using *Clostridium beijerinckii* NCIMB 8052. J Microbiol Biotechnol. 19:482-90.

99. Lehmann D, Lütke-Eversloh T (2011). Switching *Clostridium acetobutylicum* to an ethanol producer by disruption of the butyrate/butanol fermentative pathway. Metab Eng. 2011 Sep;13(5):464-73.

100. Levitz R, Chapman D, Amitsur M, Green R, Snyder L, Kaufmann G (1990). The optional *E. coli prr* locus encodes a latent form of phage T4-induced anticodon nuclease. EMBO J. 9:1383-9.

101. Li Q, Li L, Rejtar T, Lessner DJ, Karger BL, Ferry JG (2006). Electron transport in the pathway of acetate conversion to methane in the marine archaeon *Methanosarcina acetivorans*. J Bacteriol. 2006 Jan;188(2):702-10.

102. Li Z, Xiao H, Jiang W, Jiang Y, Yang S (2013). Improvement of solvent production from xylose mother liquor by engineering the xylose metabolic pathway in *Clostridium acetobutylicum* EA 2018. Appl Biochem Biotechnol. 2013 Oct;171(3):555-68.

103. Liew ST, Arbakariya A, Rosfarizan M, Raha AR (2006). Production of solvent (acetone-butanol-ethanol) in continuous fermentation by *Clostridium saccharobutylicum* DSM 13864 using gelatinized sago starch as a carbon source. Mal J Microbiol, 2:45-50.

104. Lin YL, Blaschek HP (1983). Butanol Production by a Butanol-Tolerant Strain of *Clostridium acetobutylicum* in Extruded Corn Broth. Appl Environ Microbiol. 1983 Mar;45(3):966-73.

105. Linville JL, Rodriguez M Jr, Land M, Syed MH, Engle NL, Tschaplinski TJ, Mielenz JR, Cox CD (2013). Industrial Robustness: Understanding the Mechanism of Tolerance for the Populus Hydrolysate-Tolerant Mutant Strain of *Clostridium thermocellum*. PLoS One. 2013 Oct 21;8(10):e78829.

106. Liu L, Zhang L, Tang W, Gu Y, Hua Q, Yang S, Jiang W, Yang C (2012). Phosphoketolase pathway for xylose catabolism in *Clostridium acetobutylicum* revealed by 13C metabolic flux analysis. J Bacteriol. 194:5413-22.

107. López-Contreras AM, Gabor K, Martens AA, Renckens BA, Claassen PA, Van Der Oost J, De Vos WM (2004).Substrate-induced production and secretion of cellulases by *Clostridium acetobutylicum*. Appl Environ Microbiol. 2004 Sep;70(9):5238-43.

108. López-Contreras AM, Martens AA, Szijarto N, Mooibroek H, Claassen PA, van der Oost J, de Vos WM (2003). Production by *Clostridium acetobutylicum* ATCC 824 of CelG, a cellulosomal glycoside hydrolase belonging to family 9. Appl Environ Microbiol. 2003 Feb;69(2):869-77.

109. Ma M, Li J, McClane BA (2012). Genotypic and phenotypic characterization of *Clostridium perfringens* isolates from Darmbrand cases in post-World War II Germany. Infect Immun. 2012 Dec;80(12):4354-63.

110. Mann MS, Dragovic Z, Schirrmacher G, Lütke-Eversloh T (2012). Over-expression of stress protein-encoding genes helps *Clostridium acetobutylicum* to rapidly adapt to butanol stress. Biotechnol Lett. 2012 Sep;34(9):1643-9.

111. Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH (2013). CDD: conserved domains and protein three-dimensional structure. Nucleic Acids Res. 2013 Jan;41:D348-52.

112. Mattam AJ, Yazdani S (2013). Engineering *E. coli* strain for conversion of short chain fatty acids to bioalcohols. Biotechnol Biofuels. 2013 Sep 10;6(1):128.

113. McKee AE, Rutherford BJ, Chivian DC, Baidoo EK, Juminaga D, Kuo D, Benke PI, Dietrich JA, Ma SM, Arkin AP, Petzold CJ, Adams PD, Keasling JD, Chhabra SR (2012). Manipulation of the carbon storage regulator system for metabolite remodeling and biofuel production in *Escherichia coli*. Microb Cell Fact. 2012 Jun 13;11:79.

114. Meile L, Rohr LM, Geissmann TA, Herensperger M, Teuber M (2001). Characterization of the D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase gene (*xfp*) from *Bifidobacterium lactis*. J Bacteriol. 183:2929-36.

115. Mermelstein LD, Papoutsakis ET (1993). In vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage *phi* 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 59 1077-81.

116. Millat T, Janssen H, Thorn GJ, King JR, Bahl H, Fischer RJ, Wolkenhauer O (2013). A shift in the dominant phenotype governs the pH-induced metabolic switch of *Clostridium acetobutylicum* in phosphate-limited continuous cultures. Appl Microbiol Biotechnol. 2013 Jul;97(14):6451-66.

117. Molle V, Fujita M, Jensen ST, Eichenberger P, González-Pastor JE, Liu JS, Losick R (2003). The Spo0A regulon of *Bacillus subtilis*. Mol Microbiol. 2003 Dec;50(5):1683-701.

118. Moo-Young M, Cooney CL, Humphrey AE. (1989). Comprehensive Biotechnology, 1st edition. Pergamon Press.

119. Moreira AR, Ulmer DC, Linden JC(1981). Butanol toxicity in the butylic fermentation. Biotechnol Bioeng Sym 11:567–579.

120. Mott ML, Berger JM (2007). DNA replication initiation: mechanisms and regulation in bacteria. Nat Rev Microbiol. 2007 May;5(5):343-54.

121. Mullany P, Wilks M, Tabaqchali S (1991). Transfer of Tn916 and Tn916 delta E into *Clostridium difficile*: demonstration of a hot-spot for these elements in the *C. difficile* genome. FEMS Microbiol Lett. 1991 Apr 15;63(2-3):191-4.

122. Müller V, Imkamp F, Biegel E, Schmidt S, Dilling S (2008). Discovery of a ferredoxin:NAD+-oxidoreductase (Rnf) in *Acetobacterium woodii*: a novel potential coupling site in acetogens. Ann N Y Acad Sci. 2008 Mar;1125:137-46.

123. Murray N (2002). Immigration control of DNA in bacteria: self versus non-self. 2001 Fred Griffith review lecture. Microbiology. 2002 Jan;148(Pt 1):3-20.

124. Murray NE (2000). Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). Microbiol Mol Biol Rev, 64 412–434.

125. Nair RV, Green EM, Watson DE, Bennett GN, Papoutsakis ET (1999). Regulation of the *sol* locus genes for butanol and acetone formation in *Clostridium acetobutylicum* ATCC 824 by a putative transcriptional repressor.J Bacteriol. 1999 Jan;181(1):319-30.

126. Neuhard J (1983). Utilization of preformed pyrimidine bases and nucleosides, p. 95–148. In A. Munch-Petersen (ed.), Metabolism of nucleotides, nucleosides, and nucleobases in microorganisms. Academic Press, Inc., New York, N.Y.

127. Ng YK, Ehsaan M, Philip S, Collery MM, Janoir C, Collignon A, Cartman ST, Minton NP (2013). Expanding the repertoire of gene tools for precise manipulation of the *Clostridium difficile* genome: allelic exchange using *pyrE* alleles. PLoS One. 2013;8(2):e56051.

128. Ni Y, Xia Z, Wang Y, Sun Z (2013). Continuous butanol fermentation from inexpensive sugar-based feedstocks by *Clostridium saccharobutylicum* DSM 13864. Bioresour. Technol. 129:680-5.

129. Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR (2001). Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J Bacteriol. 183:4823-38.

130. O'Brien RW, Morris JG (1971). Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. J Gen Microbiol. 1971 Nov;68(3):307-18.

131. Ohara H, Owaki M, Sonomoto K (2006). Xylooligosaccharide fermentation with *Leuconostoc lactis*. J Biosci Bioeng. 2006 May;101(5):415-20.

132. Okano K, Yoshida S, Yamada R, Tanaka T, Ogino C, Fukuda H, Kondo A (2009). Improved production of homo-D-lactic acid via xylose fermentation by introduction of xylose assimilation genes and redirection of the phosphoketolase pathway to the pentose phosphate pathway in L-Lactate dehydrogenase gene-deficient *Lactobacillus plantarum*. Appl Environ Microbiol. 2009 Dec;75(24):7858-61.

133. Ouattara HG, Reverchon S, Niamke SL, Nasser W (2011). Molecular identification and pectate lyase production by *Bacillus strains* involved in cocoa fermentation. Food Microbiol. 2011 Feb;28(1):1-8.

134. Ounine K, Petitdemange H, Raval G, Gay R (1983). Acetone-butanol production from pentoses by *Clostridium acetobutylicum*. Biotech Lettr 1983, 9 (5):605-610.

135. Ounine K, Petitdemange H, Raval G, Gay R (1985). Regulation and butanol inhibition of D-xylose and D-glucose uptake in *Clostridium acetobutylicum*.Appl Environ Microbiol. 1985 Apr;49(4):874-8.

136. Panagiotou G, Andersen MR, Grotkjaer T, Regueira TB, Hofmann G, Nielsen J, Olsson L (2008). Systems analysis unfolds the relationship between the phosphoketolase pathway and growth in *Aspergillus nidulans*. PLoS One. 2008;3(12):e3847.

137. Papagianni M (2012). Recent advances in engineering the central carbon metabolism of industrially important bacteria. Microb Cell Fact. 30;11:50.

138. Papini M, Nookaew I, Siewers V, Nielsen J (2012). Physiological characterization of recombinant *Saccharomyces cerevisiae* expressing *the Aspergillus nidulans* phosphoketolase pathway: validation of activity through $^{13}$C-based metabolic flux analysis. Appl Microbiol Biotechnol. 2012 Aug;95(4):1001-10.

139. Peck RF, DasSarma S, Krebs MP (2000). Homologous gene knockout in the archaeon *Halobacterium salinarum* with ura3 as a counterselectable marker. Mol Microbiol. 2000 Feb;35(3):667-76.

140. Pedrido ME, de Oña P, Ramirez W, Leñini C, Goñi A, Grau R (2013). Spo0A links de novo fatty acid synthesis to sporulation and biofilm development in *Bacillus subtilis*. Mol Microbiol. 2013 Jan;87(2):348-67.

141. Perutka J, Wang W, Goerlitz D, Lambowitz AM (2004). Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. J Mol Biol. 336: 421-39.

142. Pich A, Narberhaus F, Bahl H (1990). Induction of heat shock proteins during initiation of solvent formation in *Clostridium acetobutylicum*. Appl Microbiol Biotechnol (1990) 33:697 – 704.

143. Poehlein A, Hartwich K, Krabben P, Ehrenreich A, Liebl W, Dürre P, Gottschalk G, Daniel R (2013). Complete Genome Sequence of the Solvent Producer *Clostridium saccharobutylicum* NCP 262 (DSM 13864). Genome Announc. 2013 Nov 27;1(6).

144. Price C, Pripfl T, Bickle TA (1987). *Eco*R124 and *Eco*R124/3: the first members of a new family of type I restriction and modification systems. Eur J Biochem. 167:111-5.

145. Purdy D, O'Keeffe TA, Elmore M, Herbert M, McLeod A, Bokori-Brown M, Ostrowski A, Minton NP. (2002). Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol. 46 (2): 439-52.

146. Pyne ME, Moo-Young M, Chung DA, Chou CP (2013). Development of an electrotransformation protocol for genetic manipulation of *Clostridium pasteurianum*. Biotechnol. Biofuels 6(1):50.

147. Qureshi N, Saha BC, Cotta MA (2007). Butanol production from wheat straw hydrolysate using *Clostridium beijerinckii*. Bioprocess Biosyst Eng. 2007 Nov;30(6):419-27.

148. Reader JS, Metzgar D, Schimmel P, de Crécy-Lagard V (2004). Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine. J Biol Chem. 2004 Feb 20;279(8):6280-5.

149. Roberts GA, Chen K, Cooper LP, White JH, Blakely GW, Dryden DT (2012). Removal of a frameshift between the *hsdM* and *hsdS* genes of the *Eco*KI Type IA DNA restriction and modification system produces a new type of system and links the different families of Type I systems. Nucleic Acids Res. 40:10916-24.

150. Rodionov DA, Mironov AA, Gelfand MS (2001). Transcriptional regulation of pentose utilisation systems in the *Bacillus/Clostridium* group of bacteria. FEMS Microbiol Lett. 2001 Dec 18;205(2):305-14.

151. Rupnik M, Wilcox MH, Gerding DN (2009). *Clostridium difficile* infection: new developments in epidemiology and pathogenesis. Nat Rev Microbiol. 2009 Jul;7(7):526-36.

152. Sabathé F, Croux C, Cornillot E, Soucaille P (2002). AmyP, a reporter gene to study strain degeneration in *Clostridium acetobutylicum* ATCC 824. FEMS Microbiol Lett. 2002 Apr 23;210(1):93-8.

153. Rydzak T, McQueen PD, Krokhin OV, Spicer V, Ezzati P, Dwivedi RC, Shamshurin D, Levin DB, Wilkins JA, Sparling R (2012). Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression. BMC Microbiol. 2012 Sep 21;12:214.

154. Saddler JN, Yu EK, Mes-Hartree M, Levitin N, Brownell HH (1983). Utilization of enzymatically hydrolyzed wood hemicelluloses by microorganisms for production of liquid fuels. Appl Environ Microbiol. 1983 Jan;45(1):153-60.

155. Saier MH Jr, Chauvaux S, Cook GM, Deutscher J, Paulsen IT, Reizer J, Ye JJ (1996). Catabolite repression and inducer control in Gram-positive bacteria. Microbiology. 142:217-30.

156. Sambrook, J (2001) Molecular cloning. A laboratory manual. Cold Spring Harbor, New York.

157. Sambrook, J, Fritsch EF, Maniatis T. (1989). Molecular cloning: A laboratory manual. 2nd ed. Cold Spring Harbor Lab., Cold Spring Harbor, NY.

158. Sánchez B, Zúñiga M, González-Candelas F, de los Reyes-Gavilán CG, Margolles A (2010). Bacterial and eukaryotic phosphoketolases: phylogeny, distribution and evolution. J Mol Microbiol Biotechnol. 18:37-51.

159. Sá-Nogueira I, Mota LJ (1997). Negative regulation of L-arabinose metabolism in *Bacillus subtilis*: characterization of the *araR (araC)* gene. J Bacteriol. 179:1598-608.

160. Sarkar A, Köhler J, Hurek T, Reinhold-Hurek B (2012). A novel regulatory role of the Rnf complex of *Azoarcus* sp. strain BH72. Mol Microbiol. 2012 Jan;83(2):408-22.

161. Sauer U, Dürre P (1995). Differential induction of genes related to solvent formation during the shift from acidogenesis to solventogenesis in continuous culture of *Clostridium acetobutylicum*. FEMS Microbiol Lett 1995; 125: 115–120.

162. Schäffer AA, Aravind L, Madden TL, Shavirin S, SpougeJL, Wolf YI, Koonin EV, Altschul SF (2001). Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements. Nucleic Acids Res. 29:2994-3005.

163. Schaffer S, Isci N, Zickner B, Dürre P (2002). Changes in protein synthesis and identification of proteins specifically induced during solventogenesis in *Clostridium acetobutylicum*. Electrophoresis. 2002 Jan;23(1):110-21.

164. Schmehl M, Jahn A, Meyer zu Vilsendorf A, Hennecke S, Masepohl B, Schuppler M, Marxer M, Oelze J, Klipp W (1993). Identification of a new class of nitrogen fixation genes in *Rhodobacter capsulatus*: a putative membrane complex involved in electron transport to nitrogenase. Mol Gen Genet. 1993 Dec;241(5-6):602-15.

165. Schumann W, Hecker M, Msadek T (2002). Regulation and function of heat-inducible genes in *Bacillus subtilis*. In: Soneshein A, *et al.*(Eds.), *Bacillus subtilis* and its clostest relatives: from genes to cells. ASM Press, Washington, DC, pp. 359 – 368.

166. Schwarz KM, Kuit W, Grimmler C, Ehrenreich A, Kengen SW (2012). A transcriptional study of acidogenic chemostat cells of *Clostridium acetobutylicum*-

cellular behavior in adaptation to n-butanol. J Biotechnol. 2012 Oct 31;161(3):366-77.

167. Scotcher MC, Rudolph FB, Bennett GN (2005). Expression of *abr*B310 and SinR, and effects of decreased abrB310 expression on the transition from acidogenesis to solventogenesis, in *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol 71:1987–1995.

168. Servinsky MD, Germane KL, Liu S, Kiel JT, Clark AM, Shankar J, Sund CJ (2012). Arabinose is metabolized *via* a phosphoketolase pathway in *Clostridium acetobutylicum* ATCC 824. J Ind Microbiol Biotechnol. 39:1859-67.

169. Servinsky MD, Kiel JT, Dupuy NF, Sund CJ (2010). Transcriptional analysis of differential carbohydrate utilization by *Clostridium acetobutylicum*. Microbiology. 156:3478-91.

170. Skrzypek E, Piekarowicz A (1989). The *EcoD*XX1 restriction and modification system: cloning the genes and homology to type I restriction and modification systems. Plasmid. 21:195-204.

171. Sonderegger M, Schümperli M, Sauer U (2004). Metabolic engineering of a phosphoketolase pathway for pentose catabolism in *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2004 May;70(5):2892-7.

172. Soucaille P., Frigge R, Croux C (2008). Process for chromosomal integration and DNA sequence replacement in clostridia. United States Patent Application 20080085558.

173. Stim-Herndon KP, Petersen DJ and Bennett GN (1995). Molecular characterization of the acetyl coenzyme A acetyltransferase (thiolase) from *Clostridium acetobutylicum* ATCC 824. Gene 154: 81–85

174. Strauch MA, Spiegelman GB, Perego M, Johnson WC, Burbulys D, Hoch JA (1989). The transition state transcription regulator AbrB of *Bacillus subtilis* is a DNA binding protein. EMBO J. 1989;8: 1615–1621.

175. Tamaru Y, Karita S, Ibrahim A, Chan H, Doi RH (2000). A large gene cluster for the *Clostridium cellulovorans* cellulosome. J Bacteriol. 2000 Oct;182(20):5906-10.

176. Tanaka K, Komiyama A, Sonomoto K, Ishizaki A, Hall SJ, Stanbury PF (2002). Two different pathways for D-xylose metabolism and the effect of xylose concentration on the yield coefficient of L-lactate in mixed-acid fermentation by

the lactic acid bacterium *Lactococcus lactis* IO-1. Appl Microbiol Biotechnol. 2002 Oct;60(1-2):160-7.

177. Tangney M, Galinier A, Deutscher J, Mitchell WJ (2003). Analysis of the elements of catabolite repression in *Clostridium acetobutylicum* ATCC 824. J Mol Microbiol Biotechnol. 6:6-11.

178. Terracciano JS, Kashket ER (1986). Intracellular Conditions Required for Initiation of Solvent Production by *Clostridium acetobutylicum*. Appl Environ Microbiol. 1986 Jul;52(1):86-91.

179. Thormann K, Feustel L, Lorenz K, Nakotte S, Dürre P (2002). Control of butanol formation in *Clostridium acetobutylicum* by transcriptional activation. J Bacteriol. 2002 Apr;184(7):1966-73.

180. Thykaer J, Nielsen J (2007). Evidence, through $^{13}$C-labelling analysis, of phosphoketolase activity in fungi. Process Biochem 2007 Jan;42 1050-5.

181. Tomas CA, Alsaker KV, Bonarius HP, Hendriksen WT, Yang H, Beamish JA, Paredes CJ, Papoutsakis ET (2003 b). DNA array-based transcriptional analysis of asporogenous, nonsolventogenic *Clostridium acetobutylicum* strains SKO1 and M5. J Bacteriol. 2003 Aug;185(15):4539-47.

182. Tomas CA, Beamish J, Papoutsakis ET (2004). Transcriptional analysis of butanol stress and tolerance in *Clostridium acetobutylicum*. J Bacteriol. 2004 Apr;186(7):2006-18.

183. Tomas CA, Welker NE, Papoutsakis ET (2003 a). Overexpression of *groESL* in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. Appl Environ Microbiol. 2003 Aug;69(8):4951-65.

184. Tremblay PL, Zhang T, Dar SA, Leang C, Lovley DR (2012). The Rnf complex of *Clostridium ljungdahlii* is a proton-translocating ferredoxin:NAD+ oxidoreductase essential for autotrophic growth. MBio. 2012 Dec 26;4(1):e00406-12.

185. Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, McCool JD, Warner AK, Rajgarhia VB, Lynd LR, Hogsett DA, Caiazza NC (2010). Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. Appl Environ Microbiol. 2010 Oct;76(19):6591-9.

186. Vianna CR, Silva CL, Neves MJ, Rosa CA (2008). *Saccharomyces cerevisiae* strains from traditional fermentations of Brazilian cachaça: trehalose metabolism, heat and ethanol resistance. Antonie Van Leeuwenhoek. 2008 Jan-Feb;93(1-2):205-17.

187. Vieira J, Messing J (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19, 259-268.

188. Vollherbst-Schneck K, Sands JA, Montenecourt BS (1984). Effect of butanol on lipid composition and fluidity of *Clostridium acetobutylicum* ATCC 824. Appl Environ Microbiol. 1984 Jan;47(1):193-4.

189. Wang Q, Venkataramanan KP, Huang H, Papoutsakis ET, Wu CH (2013). Transcription factors and genetic circuits orchestrating the complex, multilayered response of *Clostridium acetobutylicum* to butanol and butyrate stress. BMC Syst Biol. 2013 Nov 6;7:120.

190. Wang R, Li L, Zhang B, Gao X, Wang D, Hong J. (2013). Improved xylose fermentation of *Kluyveromyces marxianus* at elevated temperature through construction of a xylose isomerase pathway. J Ind Microbiol Biotechnol. 2013 Aug;40(8):841-54.

191. Napoli F, Olivieri G, Russo ME, Marzocchella A, Salatino P (2010). Butanol production by *Clostridium acetobutylicum* in a continuous packed bed reactor.J Ind Microbiol Biotechnol. 2010 Jun;37(6):603-8.

192. Wang R, Zhang L, Wang D, Gao X, Hong J (2011). Identification of a xylulokinase catalyzing xylulose phosphorylation in the xylose metabolic pathway of *Kluyveromyces marxianus* NBRC1777. J Ind Microbiol Biotechnol. 2011 Oct;38(10):1739-46.

193. Wang Y, Li X, Mao Y, Blaschek HP (2013 a). Genome-wide dynamic transcriptional profiling in *Clostridium beijerinckii* NCIMB 8052 using single-nucleotide resolution RNA-Seq. BMC Genomics. 2012 Mar 20;13:102.

194. Wang Y, Li X, Milne CB, Janssen H, Lin W, Phan G, Hu H, Jin YS, Price ND, Blaschek HP (2013 b). Development of a gene knockout system using mobile group II introns (Targetron) and genetic disruption of acid production pathways in *Clostridium beijerinckii*. Appl Environ Microbiol. 2013 Oct;79(19):5853-63.

195. Weimberg R (1961). Pentose oxidation by *Pseudomonas fragi.* J Biol Chem. 1961 Mar;236:629-35.

196. Wells CL, Wilkins TD (1996). Clostridia: Sporeforming Anaerobic Bacilli. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 18.

197. Wietzke M, Bahl H (2012). The redox-sensing protein Rex, a transcriptional regulator of solventogenesis in *Clostridium acetobutylicum*. Appl Microbiol Biotechnol. 2012 Nov;96(3):749-61.

198. Wigley DB (2013). Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. Nat Rev Microbiol. 2013 Jan;11(1):9-13.

199. Wilson GG, Murray NE (1991). Restriction and modification systems. Annu. Rev. Genet. 25, 585-627.

200. Winn LM, Kim PM, Nickoloff JA (2003). Oxidative stress-induced homologous recombination as a novel mechanism for phenytoin-initiated toxicity. J Pharmacol Exp Ther. 2003 Aug;306(2):523-7.

201. Winzer K, Lorenz K, Zickner B, Dürre P (2000). Differential regulation of two thiolase genes from *Clostridium acetobutylicum* DSM 792. J Mol Microbiol Biotechnol. 2000 Oct;2(4):531-41.

202. Xiao H, Gu Y, Ning Y, Yang Y, Mitchell WJ, Jiang W, Yang S (2011). Confirmation and elimination of xylose metabolism bottlenecks in glucose phosphoenolpyruvate-dependent phosphotransferase system-deficient *Clostridium acetobutylicum* for simultaneous utilization of glucose, xylose, and arabinose. Appl Environ Microbiol. 77:7886-95.

203. Yao J, Lambowitz AM (2007). Gene targeting in gram-negative bacteria by use of a mobile group II intron ("Targetron") expressed from a broad-host-range vector. Appl Environ Microbiol. 2007 Apr;73(8):2735-43.

204. Yoshida T, Tashiro Y, Sonomoto K (2012). Novel high butanol production from lactic acid and pentose by *Clostridium saccharoperbutylacetonicum*. J Biosci Bioeng. 2012 Nov;114(5):526-30.

205. Youell J, Firman K (2012). Mechanistic insight into Type I restriction endonucleases. Front Biosci (Landmark Ed). 17:2122-39.

206. Zhang L, Leyn SA, Gu Y, Jiang W, Rodionov DA, Yang C (2012). Ribulokinase and transcriptional regulation of arabinose metabolism in *Clostridium acetobutylicum*. J Bacteriol. 194:1055-64. Epub 2011 Dec 22.

207. Zhang S, Jiang W, Yu Y (2012). In silico analysis of GGDEF domain proteins functioning differently in *Xanthomonas campestris*. Mol Pathogen, 3 (1):1-5.

208. Zhao Y, Tomas CA, Rudolph FB, Papoutsakis ET, Bennett GN (2005). Intracellular butyryl phosphate and acetyl phosphate concentrations in *Clostridium acetobutylicum* and their implications for solvent formation. Appl Environ Microbiol. 2005 Jan;71(1):530-7.

209. Zhong W, Priest NK (2011).Stress-induced recombination and the mechanism of evolvability. Behavioral Ecology and Sociobiol. 65(3):493 – 502.

## 9. Supplementary Data

**Table 9.1.** List of primers used during creation of the upp-deficient strain. The restriction sites used for cloning are underlined and specific enzymes are indicated in the description.

| Name | Sequence (5' → 3') | Description |
|---|---|---|
| pCH1 check fwd | GCTCTCAGAGGCTCATAGAC | Check primers for all inserts in the pCH1 plasmid; Integration check for the pJL10 plasmid |
| pCH1 check rev | CGGGCCTCTTCGCTATTA | Check primers for all inserts in the pCH1 plasmid; Integration check for the pJL10 plasmid |
| *upp* up rev *Hind*III | ACTTGCCTCGGTAATG | Integration check for the pMAD*Kupp*- plasmid |
| check *upp* down fwd | CCATGGAGAAAGTGGCTCTTTGC | Integration check for the pMAD*Kupp*- plasmid |
| *upp* down fwd *Bam*HI | CACCTGCTGAAAAAGCT | Integration check for the pMAD*Kupp*- plasmid |
| check *upp* up rev | CCATGAATAGCAGATG | Integration check for the pMAD*Kupp*- plasmid |
| 520 LHA fwd | TTTTTTGGATCCTATTTTAACTGTTGGTAAAAG | Creation of the pJL9-520 left homology arm, *Bam*HI |
| 520 LHA rev | TGGAGGAATGAAATAATGTAAATTGCATAAATAAAAAGG | Creation of the pJL9-520 left homology arm |
| 520 RHA rev | TTTTTTTAAGCTTTATCTTCTTTTTTTGCGATTATG | Creation of the pJL9-520 right homology arm, *Hind*III |
| 520 RHA fwd | TATTTATGCAATTTACATTATT | Creation of the pJL9-520 right |

| | | |
|---|---|---|
| | TCATTCCTCCATTG | homology arm |
| *upp* check down fwd | CCATGGAGAAAGTGGCTCTTTGC | Integration check for the pJL9-520 plasmid |
| *upp* check down rev | CAGGAATAAGAGCAGCTGTATG | Integration check for the pJL9-520 plasmid |
| *upp* check up fwd | TCCAGGTGGAACAGTTGCCCTTAC | Integration check for the pJL9-520 plasmid |
| *upp* check up rev | CAGCAGGTGCATCTGCTATTCATGG | Integration check for the pJL9-520 plasmid |
| LHA *upp* in fwd | TTTTTTTCTAGAATTATTTTGTACCGAATAATCTATCTC | Creation of the pJL10 left homology arm, *Xba*I |
| LHA *upp* in rev | TTCTACAGGTATATGTCCAACCTTAGCAGCAGGTATTA | Creation of the pJL10 left homology arm |
| RHA *upp* in fwd | TAAGGTTGGACATATACCTGTAGAATACTTCTGTAAAC | Creation of the pJL10 right homology arm |
| RHA *upp* in rev | TTTTTTCTGCAGATGAGTAAAGTTACACAAATATCACATC | Creation of the pJL10 right homology arm, *Pst*I |
| LHA *upp* ins-out fwd | TTTTTTCTGCAGATTTGATGAGCAAGGTGCAAGGGTATC | Creation of the pJL11 left homology arm, *Pst*I |
| LHA *upp* ins-out rev | GTAAAATGTTAATAACACTTCTTAAGAAAAG | Creation of the pJL11 left homology arm |
| RHA *upp* ins-out fwd | GAAGTGTTATTAACATTTTACATTTAGTTATAC | Creation of the pJL11 right homology arm |
| RHA *upp* ins-out rev | TTTTTTTCTAGAACATTTAGAGAGTAAAAATATTGAGGTTAAAG | Creation of the pJL11 right homology arm, *Xba*I |
| *upp* genome check fwd | AAGTGGACGTGCAAGCGGAAAC | Integration check for the pJL10 / pJL11 plasmid |
| *upp* genome check | GGTGTAGAGGGTGACATTACT | Integration check for the pJL10 / |

| | | |
|---|---|---|
| rev | G | pJL11 plasmid |
| Clostron *upp* check rev | GTTGTAGGACCTGGATTAGC | Check primers for the ClosTron *C. acetobutylicum upp* mutant |
| RecA-pJL11 fwd | TTTTTTGGTACCTGATTTATTTTGCATAAAATGGATATAA | Creation of the pJL11-RecA plasmid, *Kpn*I |
| RecA-pJL11 rev | TTTTTTGAATTCTTATTTTGTATTATTTTCAGTTCCTTTTTCA | Creation of the pJL11-RecA plasmid, *Eco*RI |

**Table 9.2.** List of primers used during the creation of the methylation plasmids, pJL1 and pJL2, for *C. saccharobutylicum*. Restriction sites used for cloning are underlined and specific enzymes are indicated in the description. MT I and II check primers were used exclusively for sequence check of the MTase cassettes, and are designed to bind 0.7 – 1 kb from one another to produce significant sequencing results.

| **Name** | **Sequence (5' → 3')** | **Description** |
|---|---|---|
| 010 fwd | TTTTTTCTGCAGGATTTGATAGCACCATTTAAAGC | pJL010 plasmid creation, *Pst*I |
| 010 rev | TTTTTTCTAGAATATTGCCAATTACTTACCTTGTTATTC | pJL010 plasmid creation, *Xba*I |
| 020.1 fwd | TTTTTTCTGCAGGATATTAAATATATACTGCGAGGAG | pJL020 plasmid creation, *Pst*I |
| 020.1 rev | TTTTTTCTAGAATAATAAAGTATTTTTCCGCCTAAAC | pJL020 plasmid creation, *Xba*I |
| 020.2 fwd | TTTTTTCTAGATACTTTCCTTGTTTGGTTGGTAATG | pJL020 plasmid creation, *Xba*I |
| 020.2 rev | TTTTTTGGCGCCATTTCATCTACAAATGCAATAAAAG | pJL020 plasmid creation, *Ehe*I |

| | | |
|---|---|---|
| JetT7 fwd | TTTTTTACATGTGTCCCAATTAG TAGCATCACGCTGTGAG | T7 promoter amplification for pJL021 / pJL011 plasmid creation, *Pci*I |
| JetT7 rev | TTTTTTCTGCAGAAACTCGAGCC ATCCGGAAGATCTG | T7 promoter amplification for pJL021 / pJL011 plasmid creation, *Pst*I |
| 011/21 fwd | TTTTTTACTAGTCCAATTAGTAG CATCACGCTGTGAG | pJL022 / pJL012 plasmid creation, *Spe*I |
| 011 rev | TTTTTTCCCGGGGATCCTCTAGA ATATTGCCAATTA | pJL012 plasmid creation, *Sma*I |
| 021 rev | TTTTTTCCCGGGTCAGGCGCCAT TTCATCTACAAATGC | pJL022 plasmid creation, *Sma*I |
| PACYC_backbone fwd | TTTTTTCCCGGGATTATCACTTAT TCAGGCGTAGCAC | Amplification of the pACYC 184 backbone, *Sma*I |
| PACYC_backbone rev | TTTTTTACTAGTATTCTTGGAGT GGTGAATCCGTTAG | Amplification of the pACYC 184 backbone, *Spe*I |
| JL-Tet fwd | TTTTTTGGTACCAAATACGCCCG GTAGTGATCTTATTTC | pJL2 / pJL1 plasmid creation, *Kpn*I |
| JL-Tet rev | TTTTTTGGTACCGCTACGCCTGA ATAAGTGATAATCCC | pJL2 / pJL1 plasmid creation, *Kpn*I |
| pACYC check fwd | ATCCATCCGGCGTAATACGACTC AC | Check primer for sequencing of MTase cassettes on the pACYC 184 plasmid |
| pACYC check rev | AAGATCACTACCGGGCGTATTTG | Check primer for sequencing of MTase cassettes on the pACYC 184 plasmid |
| MT I check 1 | CCAGCTTGTGGTACAGCAGGATT TC | Forward check primer for sequencing of the MT I cassette |
| MT I check 2 | GATAAGTGGTTCTGGGTTGATAA | Forward check primer for |

| | | |
|---|---|---|
| | AG | sequencing of the MT I cassette |
| MT I check 3 | ATTTATCGAGATGTTTGGCGATCCG | Forward check primer for sequencing of the MT I cassette |
| MT II check 1 | ATGGTAGGTATACTAAAGCCACAAG | Forward check primer for sequencing of the MT II cassette |
| MT II check 2 | AGAGTGAAGAGCGAAGAAATAAGG | Forward check primer for sequencing of the MT II cassette |
| MT II check 3 | GGGAGATACCGAATAGTTGGAAGG | Forward check primer for sequencing of the MT II cassette |

**Table 9.3.** List of ClosTron primers used to create intron mutants in *C. acetobutylicum*, *C. saccharobutylicum* and *C. beijerinckii* strains. Integration sites are given in the names of primers with sense (s) or antisense (a) directions of integration. Additionally, all intron cassettes created using the PCR technique and cloned into pMTL007 plasmids were examined by sequencing, using the pMTL2 check primers.

| **Name** | **Sequence (5' → 3')** | **Description** |
|---|---|---|
| pMTL2 check fwd | TCTCAAGAAGTTGGAGGGTGAAG | Primers for the sequencing check of the ClosTron intron cassette |
| pMTL2 check rev | CCGTCAGGATGGCCTTCTGCTTAATTTG | Primers for the sequencing check of the ClosTron intron cassette |
| EBS universal | CGAAATTAGAAACTTGCGTTCAGT | ClosTron intron cassette |

| | | |
|---|---|---|
| primer | AAAC | creation, universal primer |
| *Csa hsdR1* - 2565\|2566s-IBS primer | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCGTTAGAGATAAC TTACCTTTCTTTGT | ClosTron intron cassette targeting the CSA00451 gene |
| *Csa hsdR1* - 2565\|2566s-EBS1d primer | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCGTTAGAGATAAC TTACCTTTCTTTGT | ClosTron intron cassette targeting the CSA00451 gene |
| *Csa hsdR1* - 2565\|2566s-EBS2 primer | TGAACGCAAGTTTCTAATTTCGAT TAATGCTCGATAGAGGAAAGTGT CT | ClosTron intron cassette targeting the CSA00451 gene |
| *Csa hsdR2* - 2220\|2221s-IBS | AAAAAAGCTTATAATTATCCTTAG ATTTCGGACAGGTGCGCCCAGAT AGGGTG | ClosTron intron cassette targeting the CSA02977 gene |
| *Csa hsdR2* - 2220\|2221s-EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCGGACAGCCTAAC TTACCTTTCTTTGT | ClosTron intron cassette targeting the CSA02977 gene |
| *Csa hsdR2* - 2220\|2221s-EBS2 | TGAACGCAAGTTTCTAATTTCGAT TAAATCTCGATAGAGGAAAGTGT CT | ClosTron intron cassette targeting the CSA02977 gene |
| *Cac araK* - 260\|261a-IBS | AAAAAAGCTTATAATTATCCTTAC CAATCGAACCAGTGCGCCCAGAT AGGGTG | ClosTron intron cassette targeting the CAC1344 gene |
| *Cac araK* - 260\|261a-EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCGAACCAATTAAC TTACCTTTCTTTGT | ClosTron intron cassette targeting the CAC1344 gene |
| *Cac araK* - 260\|261a-EBS2 | TGAACGCAAGTTTCTAATTTCGGT TATTGGTCGATAGAGGAAAGTGT CT | ClosTron intron cassette targeting the CAC1344 gene |
| *Cac xylB* - 843\|844s-IBS | AAAAAAGCTTATAATTATCCTTAG GAAGCGTTCATGTGCGCCCAGAT AGGGTG | ClosTron intron cassette targeting the CAC2612 gene |

| | | |
|---|---|---|
| *Cac xylB* - 843\|844s-EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCGTTCATACTAACT TACCTTTCTTTGT | ClosTron intron cassette targeting the CAC2612 gene |
| *Cac xylB* - 843\|844s-EBS2 | TGAACGCAAGTTTCTAATTTCGAT TCTTCCTCGATAGAGGAAAGTGTC T | ClosTron intron cassette targeting the CAC2612 gene |
| *Cac pkt1* - 1887\|1888s-IBS | AAAAAAGCTTATAATTATCCTTAC AAGACCCTGATGTGCGCCCAGAT AGGGTG | ClosTron intron cassette targeting the CAC1343 gene |
| *Cac pkt1* - 1887\|1888s-EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCCCTGATGTTAACT TACCTTTCTTTGT | ClosTron intron cassette targeting the CAC1343 gene |
| *Cac pkt1* - 1887\|1888s-EBS2 | TGAACGCAAGTTTCTAATTTCGGT TTCTTGTCGATAGAGGAAAGTGTC T | ClosTron intron cassette targeting the CAC1343 gene |
| *Cac pkt2* - 621\|622s –IBS | AAAAAAGCTTATAATTATCCTTAG CAGTCCTTCCTGTGCGCCCAGATA GGGTG | ClosTron intron cassette targeting the CAC1343 gene |
| *Cac pkt2* - 621\|622s -EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCCTTCCTATTAACT TACCTTTCTTTGT | ClosTron intron cassette targeting the CAC1343 gene |
| *Cac pkt2* - 621\|622s -EBS2 | TGAACGCAAGTTTCTAATTTCGAT TACTGCTCGATAGAGGAAAGTGT CT | ClosTron intron cassette targeting the CAC1343 gene |
| *Csa araK* - 1104\|1105s-IBS | AAAAAAGCTTATAATTATCCTTAG GAGTCCTTGCAGTGCGCCCAGATA GGGTG | ClosTron intron cassette targeting the CSA00774 gene |
| *Csa araK* - 1104\|1105s-EBS1d | CAGATTGTACAAATGTGGTGATA ACAGATAAGTCCTTGCATATAACT TACCTTTCTTTGT | ClosTron intron cassette targeting the CSA00774 gene |
| *Csa araK* - | TGAACGCAAGTTTCTAATTTCGAT | ClosTron intron cassette |

| | | |
|---|---|---|
| 1104\|1105s-EBS2 | TACTCCTCGATAGAGGAAAGTGTCT | targeting the CSA00774 gene |
| *Csa xylB* - 1143\|1144s-IBS | AAAAAAGCTTATAATTATCCTTAGATTTCGTTCGCGTGCGCCCAGATAGGGTG | ClosTron intron cassette targeting the CSA03335 gene |
| *Csa xylB* - 1143\|1144s-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCGTTCGCAGTAACTTACCTTTCTTTGT | ClosTron intron cassette targeting the CSA03335 gene |
| *Csa xylB* - 1143\|1144s-EBS2 | TGAACGCAAGTTTCTAATTTCGATTAAATCTCGATAGAGGAAAGTGTCT | ClosTron intron cassette targeting the CSA03335 gene |
| *Cbei rnfC* - 784s-IBS | AAAAAAGCTTATAATTATCCTTAGTTGTCGTACAAGTGCGCCCAGATAGGGTG | ClosTron intron cassette targeting the CBEI2449 gene |
| *Cbei rnfC* - 784s-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCGTACAAAATAACTTACCTTTCTTTGT | ClosTron intron cassette targeting the CBEI2449 gene |
| *Cbei rnfC* - 784s-EBS2 | TGAACGCAAGTTTCTAATTTCGATTACAACTCGATAGAGGAAAGTGTCT | ClosTron intron cassette targeting the CBEI2449 gene |
| *Cac upp* - 84\|85a-IBS primer | AAAAAAGCTTATAATTATCCTTAATATACCTCGCAGTGCGCCCAGATAGGGTG | ClosTron intron cassette targeting the CAC2879gene |
| *Cac upp* - 84\|85a-EBS1d | CAGATTGTACAAATGTGGTGATAACAGATAAGTCCTCGCATCTAACTTACCTTTCTTTGT | ClosTron intron cassette targeting the CAC2879gene |
| *Cac upp* - 84\|85a-EBS2 | TGAACGCAAGTTTCTAATTTCGGTTTATATCCGATAGAGGAAAGTGTCT | ClosTron intron cassette targeting the CAC2879gene |

**Table 9.4.** List of primers used for the molecular check of ClosTron and clean deletion mutants. Sequencing reactions of the regions with insertions or deletions were performed using the same primers as for the ClosTron mutations check. Restriction sites used for cloning are underlined and specific enzymes are indicated in the description.

| Name | Sequence (5' → 3') | Description |
|---|---|---|
| *pyrE* rev | GTGCCTTAACATCTAAGTTGAG | primer for the sequencing check of the *C. acetobutylicum ΔpyrE* strain |
| *pyrE* fwd | GCACAATTGTATTTGGACTTC | primer for the sequencing check of the *C. acetobutylicumΔpyrE* strain |
| pMTL-ME3check fwd | CGGCCGCTGTATCCATATGACC | Check primers for sequencing of the clean deletion cassettes on the pMTL-ME3 plasmid |
| pMTL-ME3check rev | GCTGCAAGGCGATTAAGTTGGG | Check primers for sequencing of the clean deletion cassettes on the pMTL-ME3 plasmid |
| LHA *araK* fwd | TTTTTTGAATTCACAACAGATAAG CCTGTAATATTTG | Creation of the pJL3 plasmid, left homology arm, *Eco*RI |
| LHA *araK* rev | AATTTTACTTTTACAATTTTAAAA CCCCTTCCC | Creation of the pJL3 plasmid, left homology arm |
| RHA *araK* fwd | GGGGTTTTAAAATTGTAAAAGTA AAATTTAGAAG | Creation of the pJL3 plasmid, right homology arm |
| RHA *araK* rev | TTTTTTTCTAGATTAGTACTAGTA GAAACGGCAGAAC | Creation of the pJL3 plasmid, right homology arm, *Xba*I |
| LHA *xylB* fwd | TTTTTTGAATTCTATGAGTAATCTT CTTCAATATC | Creation of the pJL4 plasmid, left homology arm, *Eco*RI |
| LHA *xylB* rev | TTTGATTGTGAGGTAAATAGTTAA | Creation of the pJL4 plasmid, |

| | | |
|---|---|---|
| | AAGTAG | left homology arm |
| RHA *xylB* fwd | TACTTTTAACTATTTACCTCACAA TCAAAC | Creation of the pJL4 plasmid, right homology arm |
| RHA *xylB* rev | TTTTTTCCATGGAAAGAAGGAGAT AAATTAGCTAATG | Creation of the pJL4 plasmid, right homology arm, *Nco*I |

**Table 9.5.** List of primers used for the molecular check of ClosTron and clean deletion mutants. Sequencing reactions of the regions with insertions or deletions were performed using the same primers as for the ClosTron mutations check. Additionally, all ClosTron mutants were examined for the presence of a correctly integrated RAM cassette with RAM primers, and all *C. acetobutylicum* samples were checked for the pSOL presence. An additional pair of check primers used to determine the species (161V and 630R) was used prior to the mutation process.

| **Name** | **Sequence (5' → 3')** | **Description** |
|---|---|---|
| RAM forward | ACGCGTTATATTGATAAAAATAA TAATAGTGGG | RAM cassette check for the correct integration site |
| RAM reverse | ACGCGTGCGACTCATAGAATTAT TTCCTCCCG | RAM cassette check for the correct integration site |
| *hsdR1* fwd | TGCGTCAATGCTTAAGATGCGTA TG | Check primers for the ClosTron*C. saccharobutylicum hsdR1::int* mutant |
| *hsdR1* rev | TCATCCGCCATTCTTCCAAATAT TC | Check primers for the ClosTron *C. saccharobutylicumhsdR1::int* mutant |
| *hsdR2* fwd | GGGAGAACGTGATTCTTTAAATG AG | Check primers for the ClosTron *C. saccharobutylicumhsdR2::int* mutant |

| | | |
|---|---|---|
| *hsdR2* rev | TAGTCCAATCTACGACATAGTTATC | Check primers for the ClosTron *C. saccharobutylicum hsdR2::int* mutant |
| *pkt* check 1 fwd | ATGTATGGCAGCAGGATCATAATGG | Check primers for the ClosTron *C. acetobutylicum pkt::1int* mutant |
| *pkt* check 1 rev | TGAAGGATTGTAACAGCTGCCAAAG | Check primers for the ClosTron *C. acetobutylicum pkt::1int* mutant |
| *pkt* check 2 fwd | GTTGGAGACGGAGAGGCAGAAACAG | Check primers for the ClosTron *C. acetobutylicum pkt::2int* mutant |
| *pkt* check 2 rev | TGGCCACTTTGGTCGTGAACAATCG | Check primers for the ClosTron *C. acetobutylicumpkt::2int* mutant |
| *Cac araK* check fwd | ATTTCGGGAAGGGGTTTTAAAATTG | Check primers for the ClosTron *C. acetobutylicumaraK::int* mutant |
| *Cac araK* check rev | GATTTAATCTGATAATCTTCTAAATTTTAC | Check primers for the ClosTron *C. acetobutylicumaraK::int* mutant |
| *Cac xylB* check fwd | ACTCTTTAAATAAAAGCGCTCCTAC | Check primers for the ClosTron *C. acetobutylicumxylB::int* mutant |
| *Cac xylB* check rev | AAAATTAAGGGGGTTTGATTGTG | Check primers for the ClosTron *C. acetobutylicumxylB::int* mutant |
| *Csa araK* check fwd | TGTCTTAGGGAAACCGTTTACTG | Check primers for the ClosTron *C. saccharobutylicumaraK::int* mutant |

| | | |
|---|---|---|
| *Csa araK* check rev | AAATTATCCACTGCTGCCTGTTC | Check primers for the ClosTron *C. saccharobutylicumaraK::int* mutant |
| *Csa xylB* check fwd | AATCCTGTTAAAGCTGGGTTTC | Check primers for the ClosTron *C. saccharobutylicum xylB::int* mutant |
| *Csa xylB* check rev | TGCGAGTCGAGCTTCCTATGATG | Check primers for the ClosTron *C. saccharobutylicum xylB::int* mutant |
| *Cbei rnfC* check fwd | GAGGTTTGTGAGGCGATTATTAATG | Check primers for the ClosTron *C. beijerinckii rnfC::int*mutant |
| *Cbei rnfC* check rev | TTCTTGTACTCCAATTTCGGAACTC | Check primers for the ClosTron *C. beijerinckii rnfC::int*mutant |
| *Cac araK* clean deletion check fwd | GATCCAGGTTTATTAGGACATATTG | Check primers for clean deletion of the *C. acetobutylicum araK* gene |
| *Cac araK* clean deletion check rev | TAGTACCATTGCAGGTACCATAG | Check primers for clean deletion of the *C. acetobutylicum araK* gene |
| *Cac xylB* clean deletion check fwd | TCCATTGTAAGCATCTACTACTG | Check primers for clean deletion of the *C. acetobutylicum xylB* gene |
| *Cac xylB* clean deletion check rev | GGGGGAACAACAATAAAAATGG | Check primers for clean deletion of the *C. acetobutylicum xylB* gene |
| pSOL check fwd | CCCAAGGCTGGTGATTGGTGGTTAC | Check primers for the presence of the pSOL1 megaplasmid |
| pSOL check rev | TTTGGCCAGGAACTGGAGATTGAGG | Check primers for the presence of thepSOL1 megaplasmid |
| 161V | AGAGTTTGATYMTGGCTC | Check primers for the species |

| | | |
|---|---|---|
| | | determination sequencing |
| 630R | CAKAAAGGAGGTGATCC | Check primers for the species determination sequencing |

**Table 9.6.** List of primers used to create complementation plasmids for ClosTron *araK* and *xylB* mutants in *C. acetobutylicum* and *C. saccharobutylicum.* Restriction sites used for cloning are underlined and specific enzymes are indicated in the description.

| **Name** | **Sequence (5' → 3')** | **Description** |
|---|---|---|
| ME6X +pfdx check fwd | AAAGATAGAATTAATGAGTAC TATAAAG | Check primers for sequencing of the complementation cassettes on the pMTL-ME6X plasmid, with $p_{fdx}$region |
| ME6X check fwd | TGTAGTAGCCTGTGAAATAAGT AAG | Check primers for sequencing of the complementation cassettes on the pMTL-ME6X plasmid, without $p_{fdx}$ region |
| ME6X check rev | AATGTGCCTTAACATCTAAGTT GAG | Check primers for sequencing of the complementation cassettes on the pMTL-ME6X plasmid |
| pJL5.1 fwd | TTTTTTCATATGCGGGAAGGGG TTTTAAAATTG | Creation of the complementation plasmid pJL5.1, *Nde*I |
| pJL5.2 fwd | TTTTTTGCGGCCGCGTTATATG GAAGAGGGAACTATTAC | Creation of the complementation plasmid pJL5.2, *Not*I |
| pJL5 rev | TTTTTTCCATGGATCAATCCTTT CGGCAATTACTTAG | Creation of the complementation plasmid pJL5, *Nco*I |
| pJL6.1 fwd | TTTTTTGAATTCGGTTTGATTGT GAGGTATTTATTAG | Creation of the complementation plasmid pJL6.1, *Eco*RI |

| | | |
|---|---|---|
| pJL6.2 fwd | TTTTTTGCGGCCGCTGGAGGAG TATCAAAGGCTGGAATG | Creation of the complementation plasmid pJL6.2, *Not*I |
| pJL6 rev | TTTTTTGCTAGCAAAGCGCTCC TACTTTTAAC | Creation of the complementation plasmid pJL6, *Nhe*I |
| pJL7.1 fwd | TTTTTTCATATGATTTTGAGAG GAGGTCTTTTTATTG | Creation of the complementation plasmid pJL7.1, *Nde*I |
| pJL7 rev | TTTTTTGCTAGCCATGCTGTAA TGGTCTATAATCAAC | Creation of the complementation plasmid pJL7.1, *Nhe*I |
| pJL8.1 fwd | TTTTTTCATATGGTTAAGGGAT AGATTTGAATTTAGG | Creation of the complementation plasmid pJL8.1, *Nde*I |
| pJL8.2 fwd | TTTTTTGCGGCCGCGGTTATAT GGTTAAGCACATGATAG | Creation of the complementation plasmid pJL8.2, *Not*I |
| pJL8 rev | TTTTTTGCTAGCACAGTTACAC TTTGCATGATACC | Creation of the complementation plasmid pJL8, *Nhe*I |

**Table 9.7.** List of primers used to perform the RLM-RACE experiments to map the promoter regions upstream of chosen genes. For each promoter region two primers were designed and named ‘in’ and ‘out’, according to the specifications of the manufacturer.

| **Name** | **Sequence (5’ → 3’)** | **Description** |
|---|---|---|
| RACE5'in CAC1344 | TCACCTGTTAGCTGCCA ATG | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum araK* gene |
| RACE5'out CAC1344 | CCTGCTAGTAACACCTT TGG | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum araK* gene |

| | | |
|---|---|---|
| RACE5' in CAC2612 | AGGGTGAAACCCGTTAATGC | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum xylB* gene |
| RACE5'out CAC2612 | CTGAGTACCACTAGCATCTG | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum xylB* gene |
| RACE5'in CAC1341 | TCAATACCGCTAACATTCCC | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum araD* gene |
| RACE5'out CAC1341 | GCCCAAGGTGAATGTGTATG | Primers for RLM-RACE promoter mapping upstream of the *C. acetobutylicum araD* gene |
| RACE5'in CSA31340 | AAGCCCATGCATCTGACCAC | Primers for RLM-RACE promoter mapping upstream of the *C. saccharobutylicum xylB* gene |
| RACE5'out CSA31340 | CTGCATACCACTAGCATCTG | Primers for RLM-RACE promoter mapping upstream of the *C. saccharobutylicum xylB* gene |

**Table 9.8.** Fermentation profile of the *C. acetobutylicum* wild type, *araK::int* and *xylB::int* mutant strains in CGM medium supplemented with 5% of monosaccharide: glucose (278 mM), xylose or arabinose (333 mM). Samples were taken in the solventogenic phase of growth, at 101 hours after the inoculation time. Results are shown with the computed acetate:butyrate (Ac:Bt) ratio.

| | | **Butanol** | **Acetone** | **Ethanol** | **Butyrate** | **Acetate** | **Ac:Bt ratio** |
|---|---|---|---|---|---|---|---|
| | | | | **(mM)** | | | |
| **Glucose** | WT | 122.21 | 70.31 | 14.50 | 6.08 | 16.62 | 2.73 |
| | *araK::int* | 28.33 | 16.03 | 2.29 | 26.51 | 28.26 | 1.07 |
| | *xylB::int* | 81.98 | 53.58 | 7.47 | 9.33 | 25.36 | 2.72 |
| **Xylose** | WT | 24.31 | 11.74 | 3.22 | 25.41 | 32.11 | 1.26 |
| | *araK::int* | 12.89 | 6.65 | 2.39 | 30.09 | 34.71 | 1.15 |
| **Arabinose** | WT | 106.35 | 101.50 | 17.79 | 3.44 | 23.90 | 6.95 |
| | *xylB::int* | 97.18 | 94.60 | 19.45 | 3.55 | 16.99 | 4.79 |

**Table 9.9.** Growth, xylose consumption and fermentation profile of the first, representative phosphate-limited continuous culture. Changes in pH, OD (nm=600), xylose (mM), acid and solvent concentrations (mM) during the transition from acidogenesis to solventogenesis are separated with dotted lines, and the changes in pH values are marked in bold font.

| pH | time | OD | Xylose | Butanol | Acetone | Ethanol | Butyrate | Acetate |
|---|---|---|---|---|---|---|---|---|
| | (h) | 600nm | | | (mM) | | | |
| 5.7 | 12 | 3.47 | 132.46 | 0.53 | 1.52 | 1.48 | 24.02 | 0.75 |
| 5.7 | 24 | 5.26 | 128.88 | 3.20 | 0.46 | 2.34 | 52.85 | 0.00 |
| 5.7 | 48 | 5.22 | 136.05 | 0.24 | 0.30 | 2.16 | 51.95 | 3.48 |
| 5.7 | 72 | 4.40 | 130.01 | 0.40 | 0.00 | 2.64 | 67.18 | 7.12 |
| 5.7 | 96 | 3.70 | 125.48 | 0.00 | 0.52 | 3.03 | 72.13 | 8.24 |
| 5.7 | 108 | 3.58 | 125.48 | 0.63 | 0.00 | 3.19 | 75.59 | 10.68 |
| 5.7 | 120 | 3.01 | 114.91 | 1.13 | 0.95 | 3.96 | 77.69 | 9.35 |
| 5.7 | 144 | 3.06 | 109.25 | 1.76 | 1.11 | 4.14 | 77.78 | 7.13 |
| 5.5 | 144.8 | 3.17 | 110.01 | 2.01 | 1.19 | 4.27 | 79.90 | 6.26 |
| 5.2 | 147 | 3.01 | 115.10 | 2.29 | 1.47 | 4.06 | 79.04 | 8.70 |
| 4.9 | 151.7 | 3.05 | 129.63 | 3.99 | 5.05 | 3.77 | 70.88 | 4.82 |
| 4.7 | 156.7 | 2.62 | 142.84 | 11.18 | 9.13 | 3.16 | 45.26 | 2.51 |
| 4.5 | 163.7 | 1.92 | 170.58 | 30.15 | 22.32 | 3.91 | 8.62 | 5.57 |
| 4.5 | 192 | 2.59 | 150.95 | 27.37 | 21.32 | 4.48 | 9.84 | 6.32 |
| 4.5 | 216 | 2.96 | 152.65 | 27.51 | 18.50 | 4.48 | 13.11 | 5.35 |
| 4.5 | 228 | 2.92 | 156.62 | 20.59 | 16.72 | 4.17 | 13.17 | 6.99 |
| 4.5 | 240 | 3.05 | 161.33 | 25.72 | 18.22 | 4.29 | 16.34 | 4.39 |

## Acknowledgements

First and foremost I want to thank Prof. Dr. Wolfgang Liebl for giving me the opportunity to do the doctoral studies in his Department of Microbiology and Dr. Armin Ehrenreich for his help and valuable discussions during the completion of my PhD thesis. I would also like to thank Prof. Dr. Nigel Minton for hosting me in his laboratory at the University of Nottingham, for his invaluable remarks on my work and help with the genetic systems for clostridia. Furthermore I thank Dr. Ying Zhang for her time and help.

Additionally, I wish to thank Rolf Daniel and Peter Dürre for sharing the genome sequence of *C. saccharobutylicum*, and the group of Dr. Wolfgang Schwarz for cooperation with the GC analysis.

My very special thanks go to the whole research group and all the people from the Department of Microbiology: Claudia, Daniel, Xu, Li, Ching-Ning, Markus, Michael, Björn, David, Liu, Hedwig, Benne and Johanna. Thank you all for your cooperation and for creating a great working atmosphere. I also want to thank Desiree for being my tutor and colleague, and I thank Chrissi and Nina for a warm welcome and the many nice times we spent together.
I thank Jennifer Wieland, Daniela Dyckhoff and Philipp Högel for their cooperation during my *Clostridium* project.

Moreover, I wish to thank all the members and leaders of the Marie Curie ClostNet network for an amazing scientific experience, many great discussions and wonderful non-official sessions. Additionally, I thank all the people who made my stay at the University of Nottingham a terrific, and also a fruitful time. I want to thank Muhammad Ehsaan for sharing his system for genetic manipulations in clostridia and Katrin Schwarz, Wouter Kuit and Steve Cartman for useful tips on working with clostridia.

I also want to thank Helga, Beate and Maria for maintaining order, for helping with the preparation of part of my experiments and nice lunch breaks.

My very special thanks go to Konny for being my friend and an amazing help in many moments – dziękuję Ci, Konny, za bycie moją przyjaciółką i niezastąpioną pomocą w wielu różnych momentach.

I wish to thank everyone who was there for me during all the four years of my PhD studies, my family and friends. Thank you very, very much.

This work was funded by funds from the EU within the Marie-Curie ITN project CLOSNET (Contract No. 237942).