



GHENT UNIVERSITY

FACULTY OF BIOSCIENCE ENGINEERING

Project: academic year 2012-2013

The iGEM competition:

Synthetic biology based on standard parts

A new model for stabilised gene duplication

**Eline DE SCHUTTER**

**Renate DE SMEDT**

**Eva D'HAENE**

**Tine GOOVAERTS**

**Jules HAEZEBROUCK**

**Tim HEBBRECHT**

**Marlies VERSCHUUREN**

Promoter: Prof. dr. ir. M. DE MEY

Tutor: C. VERNIERS

## Copyright

*The authors and promoter give the permission to use this paper for consultation and to copy parts of it for personal use. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this paper.*

Ghent, June 2013

Promoter:

Prof. dr. ir. Marjan De Mey

Tutor:

Charlotte Verniers

Authors:

Eline De Schutter

Renate De Smedt

Eva D'haene

Tine Goovaerts

Jules Haezebrouck

Tim Hebbrecht

Marlies Verschuuren

## Preface

October 2012 – A small green poster appeals to third bachelor bioengineering students to participate in the iGEM competition within the scope of the upcoming bachelor paper. “Are you interested in synthetic biology?” Synthetic biology ... what on earth?

Now, several months later, we are finally beginning to grasp the vast possibilities of this new and rapidly expanding discipline. Our promoter, prof. Marjan De Mey, and our tutor, Charlotte Verniers, opened our eyes with their numerous ideas and invaluable feedback. We would like to thank them profusely for embarking with us on this adventure and for assisting us throughout. We are looking forward to learning even more from them as we begin to work out the practical side of our project.

Speaking of practical work. The interest of several sponsors has already given us the certainty that we will be able to test our idea in the lab. We are very grateful to our head sponsor ‘Bio Base Europe Pilot Plant’ for their belief in our project and their indispensable financial support. A big thank you also goes to Bioké, MRP Ghent Bio-Economy and Novolab for helping us to make our iGEM adventure possible.

We already had the chance to present our project to a broad public at the Open House Day of our faculty. We would like to thank the Faculty of Bioscience Engineering for giving us this opportunity and Katrien Van Impe of the faculty’s communications office for helping with the logistics.

Last but not least, we would like to thank our parents for their full support in every aspect of our education.

## Table of contents

1	The iGEM competition.....	1
2	Problem posing and setting our goals.....	2
2.1	Introduction: high gene expression in industrial biotechnology.....	2
2.2	Plasmid-based high gene expression .....	2
2.2.1	Plasmid maintenance imposes a metabolic burden on cells .....	2
2.2.2	Plasmids suffer from genetic instability .....	2
2.3	Chemically inducible chromosomal evolution, a system for stable gene overexpression .....	3
2.4	Antibiotic resistance, a major drawback of CICE.....	4
2.5	Our goal .....	5
3	Alternatives to antibiotic resistance.....	6
3.1	Triclosan and <i>fabI</i> overexpression.....	6
3.1.1	Mechanism of triclosan inhibition of bacterial fatty acid synthesis.....	6
3.1.2	Triclosan induced chromosomal evolution .....	7
3.1.3	Advantages and disadvantages .....	8
3.2	Operator-repressor titration .....	10
3.2.1	Repressor titration for plasmid selection.....	10
3.2.2	Repressor titration and CICE .....	11
3.3	Solvents .....	12
3.3.1	Accumulation of solvents in the membrane .....	12
3.3.2	Efflux pumps.....	13
3.3.3	Chaperon proteins.....	13
3.3.4	Solvents and chromosomal evolution .....	14
3.4	Cationic antimicrobial peptides.....	15
3.5	Toxin-Antitoxin systems .....	16
3.5.1	Types of TA-systems.....	16
3.5.2	TA systems on plasmids: postsegregational killing .....	18
3.5.3	Chromosomal TA systems: biological roles .....	19
3.5.4	The use of TA systems in chromosomal evolution.....	20
4	Developing a new model for chromosomal evolution without antibiotic marker.....	21
4.1	Our alternative to antibiotic resistance: the F plasmid <i>ccd</i> TA module .....	21
4.1.1	Function and mechanism of the F plasmid <i>ccd</i> TA module.....	21

4.1.2	Advantages and disadvantages of the F plasmid <i>ccd</i> TA module in CICHÉ.....	22
4.1.3	Our use of the F plasmid TA system .....	23
4.1.4	Choosing an inducible promoter for <i>ccdB</i> .....	23
4.2	Selecting a gene of interest .....	26
4.3	The role of RecA in homologous recombination.....	27
4.4	Selecting the appropriate homology regions for recombination.....	28
4.5	Our model for CICHÉ: an overview .....	28
References.....		29

## List of abbreviations

ACP	acyl carrier protein
ATP	adenosine-5'-triphosphate
CAMP	cationic antimicrobial peptides
cAMP	cyclic AMP
CAP	catabolite activator protein
<i>cat</i>	chloramphenicol acetyl transferase gene
ccd	control of cell death or coupled cell division
CIChE	chemically inducible chromosomal evolution
DAP	diaminopimelate
DCW	dry cell weight
dsDNA	double-stranded DNA
FAS	fatty acid synthase
GFP	green fluorescent protein
HSP	heat shock proteins
iGEM	International Genetically Engineered Machine
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<i>kan</i>	kanamycin resistance gene
kb	kilobase
kDa	kilodalton
<i>lacO</i>	lactose operator
<i>lacI</i>	lactose repressor
LPS	lipopolysaccharide
MIC	minimum inhibitory concentration
NAD <sup>+</sup>	nicotinamide adenine dinucleotide
ORT	operator-repressor titration
PBAD	arabinose-inducible araBAD promoter
PCD	programmed cell death
PHB	poly-3-hydroxybutyrate
PSK	postsegregational killing
RNAP	RNA polymerase
RND	resistance-nodulation-cell division
ssDNA	single-stranded DNA
TA	toxin-antitoxin

## List of figures

Figure 1: Allele segregation. ....	3
Figure 2: Chemically inducible chromosomal evolution. ....	4
Figure 3: Structure of triclosan, a member of the bisphenol family <sup>10</sup> .....	6
Figure 4: Fatty acid elongation in the type II FAS system. ....	7
Figure 5: Overproduction of lycopene using triclosan induced chromosomal evolution. ....	8
Figure 6: Differences in activity between FabI and FabI[G93V]. ....	9
Figure 7: Operator-repressor titration. ....	10
Figure 8: Repressor titration with <i>dapD</i> . ....	11
Figure 9: Repressor titration and ClChE. ....	12
Figure 10: An efflux pump of the resistance-nodulation-cell division (RND) family. ....	13
Figure 11: ClChE construct with <i>acrAB</i> and <i>tolC</i> as a selection marker. ....	14
Figure 12: ClChE construct with HSP genes <i>groES</i> and <i>groEL</i> as a selection marker. ....	14
Figure 13: Cationic antimicrobial peptides (CAMP) and the resistance mechanisms of bacteria. ....	15
Figure 14: Construct for chromosomal evolution by using cationic antimicrobial peptides. ....	16
Figure 15: TA systems type I mechanism. ....	17
Figure 16: TA systems type II mechanism. ....	17
Figure 17: TA systems type III mechanism. ....	18
Figure 18: Postsegregational killing. ....	19
Figure 19: Interaction between CcdB and GyrA. ....	22
Figure 20: The lactose operon. ....	24
Figure 21: The conformational change of AraC. ....	25
Figure 22: Protein expression system using T7. ....	26
Figure 23: Function of RecA. ....	27
Figure 24: Our model for ClChE. ....	28

# 1 The iGEM competition

The International Genetically Engineered Machine (iGEM) competition is a worldwide competition in the field of synthetic biology. This contest initially aimed at university students, but has in 2011 expanded to include a High School Division and an Entrepreneurship Division. iGEM began in January 2003 as a month-long course at MIT during the Independent Activities Period. This design course grew to a summer competition with 5 teams in 2004, 13 teams in 2005 (the year it became an international competition), 32 teams in 2006 and up to 190 teams from all over the world in 2012. Because of this increasing size, the competition was split into three regions in 2011: Europe, America, and Asia. Regional jamborees take place each year in October and the most impressive teams attending those events are selected to advance to the World Championship at MIT in November. Projects range from a rainbow of pigmented bacteria, to banana and wintergreen smelling bacteria, an arsenic biosensor and buoyant bacteria.

One of the aims of the competition is to build simple biological systems from standard, interchangeable parts and operate them in living cells. The iGEM competition helps the students by providing a library of standardized parts (called BioBrick standard biological parts) from the Registry of Standard Biological Parts. Over the summer, these parts and new parts designed by the students are used to build biological systems and they are operated in living cells. These systems are the so called 'genetic machines'. Successful projects produce cells that possess new properties by engineering sets of multiple genes together with mechanisms to regulate their expression.

Beyond just building biological systems, the broader goals of iGEM include the enabling of the systematic engineering of biology, the promotion of the development of tools for engineering biology and to help build a society that can productively and safely apply biological technology. iGEM's dual aspects of self-organization and imaginative manipulation of genetic material have demonstrated a new way to arouse interest in modern biology and to develop independent learning skills<sup>1</sup>.

Ghent University has not yet participated in this world-renowned competition in synthetic biology. That, however, is about to change. In this paper the first UGent team presents its project which it will participate with in 2013's iGEM competition: a new model for stabilised gene duplication.



## **2 Problem posing and setting our goals**

### **2.1 Introduction: high gene expression in industrial biotechnology**

The main goal of the industrial biotechnology is to increase the yield of the synthesis of biochemical products using microorganisms as production host. In general, this includes engineering large synthetic pathways and improving their expression<sup>2</sup>. This latter aspect is not always easy and depends on the complexity of the pathway and the role of the product in the metabolism of the cell<sup>3</sup>. Other facts that need to be considered are feedback inhibition, product stability, export, intracellular consumption, toxicity, etc. Until today, this host engineering is mainly based on deletion of substrate competing pathways as well as (over)expression of endogenous and/or heterologous genes using plasmids. In addition, genome based expression has also been reported, although only limited gene copy numbers could be attained.

In 2009, Tyo *et al.* developed a new, plasmid-free system for high gene copy expression in *E. coli*: chemically inducible chromosomal evolution (CICHE)<sup>2</sup>. This technique improves the genetic stability of the expressed pathway.

### **2.2 Plasmid-based high gene expression**

In industrial biotechnology, a common technique to express new synthetic products and pathways is the use of plasmids as vectors. Plasmids are easy to insert into cells and replicate independently from the genome, allowing strong gene expression. Overexpression is easily achieved by using plasmids with a medium or high copy number, different promoter systems, ribosome binding sites (RBS), etc. Thanks to plasmids, the industrial biotechnology has grown substantially over the past years. However, the use of plasmids entails some important disadvantages, which will be discussed below.

#### **2.2.1 Plasmid maintenance imposes a metabolic burden on cells**

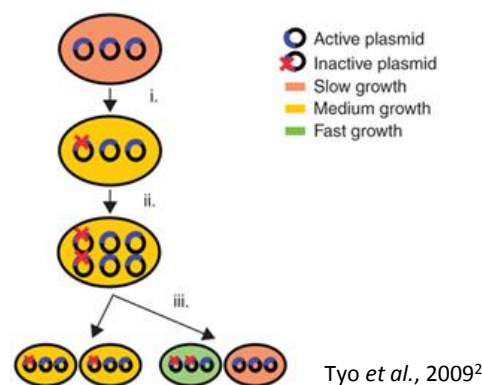
When plasmids are present in cells and replicate, they create a metabolic burden. This is defined by Bentley *et al.* (1990) as “*the amount of resources that are taken from the host cell metabolism for foreign DNA maintenance and replication*”<sup>4</sup>. As a result, metabolic load causes many alterations to the physiology and metabolism of the cell and reduces the cellular fitness<sup>5</sup>. The most common change is a delayed growth. This can be caused by the fact that new pathways for energy generation are activated due to the competition between cell propagation and plasmid replication<sup>6</sup>. This growth retardation leads to a lower yield of the desired product.

#### **2.2.2 Plasmids suffer from genetic instability**

Plasmids are genetically instable due to three processes: segregational instability, structural instability and allele segregation<sup>2</sup>. Segregational instability is caused by unequal distribution of plasmids to daughter cells, which leads to cells devoid of any plasmids. This problem has been solved by using selectable markers (e.g. antibiotic resistance) or post-segregational killing to remove plasmid-lacking cells. Structural instability leads to an incorrect expression of proteins. This problem

finds its cause in plasmids with a changed DNA sequence. However, these mutations occur at a low frequency<sup>2</sup>.

A third issue related to plasmids is allele segregation, which up to now remains mainly unaddressed. When a mutation occurs in the gene of interest, but not in the selectable marker, cells can emerge that are resistant to the selection, but not productive. As a result they cannot be removed by the use of selectable markers or post-segregational killing (**Figure 1**). After the mutation occurs, the plasmids are replicated and divided over daughters cells. The mutated plasmids can be divided in two different ways: either each daughter cell receives one mutated plasmid, or only one daughter cell receives both. In the latter case, the cell receiving both mutated plasmids produces less of the desired product and grows faster than the other cell. This growth advantage is due to the fact that the new synthetic pathway that has been inserted places a heavy metabolic burden on cells and reduces the cellular fitness. Therefore, cells with mutated plasmids accumulate and lead to a great productivity loss.



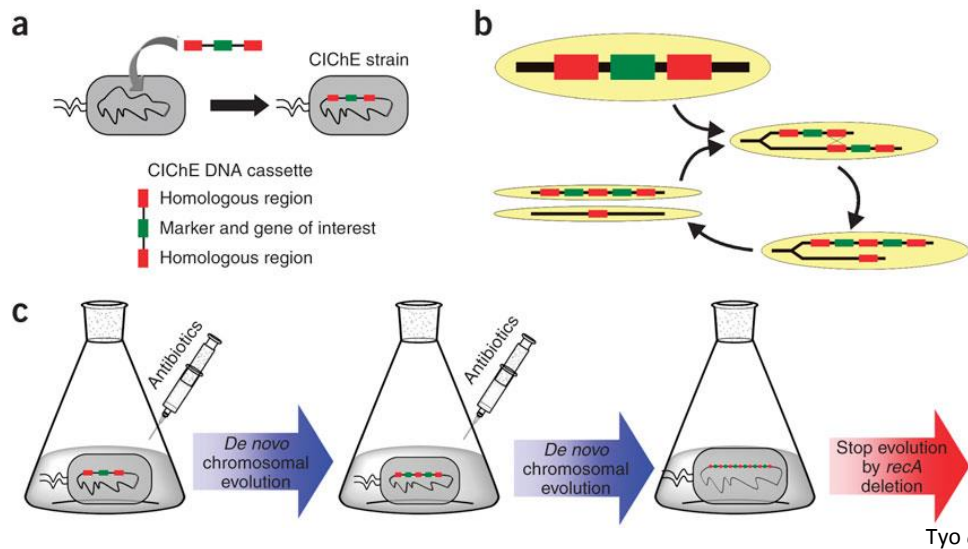
**Figure 1: Allele segregation.** Allele segregation results in a rapid loss of productivity in plasmids (i) A mutation occurs in the gene of interest leading to a non-productive plasmid. (ii) The plasmids are replicated. (iii) The plasmids are divided over daughters cells. The mutated plasmids can be separated over the two daughter cells or one daughter cell can receive both mutated plasmids. This last, non-productive cell will grow faster than the other cells, leading to a great productivity loss<sup>2</sup>.

## 2.3 Chemically inducible chromosomal evolution, a system for stable gene overexpression

In 2009, Tyo *et al.* developed a technique for the stable, high copy expression of a gene of interest in *E. coli* without the use of high copy number plasmids, thus avoiding their previously stated negative characteristics<sup>2</sup>. They called this plasmid-free, high gene copy expression system 'chemically inducible chromosomal evolution'. In this method, the gene of interest is integrated in the microbial genome and then amplified to achieve multiple copies and reach the desired expression level. Genomic integration guarantees ordered inheritance, resolving the problem of allele segregation.

CiChE works as follows: First a construct, containing the gene(s) of interest and the antibiotic marker *chloramphenicol acetyl transferase (cat)* flanked by homologous regions, is delivered to and subsequently integrated into the *E. coli* genome. The construct can be amplified in the genome

through tandem gene duplication by *recA* homologous recombination. Then the strain is cultured in increasing concentrations of chloramphenicol, providing a growth advantage for cells with increased repeats of the construct and thereby selecting for bacteria with a higher gene copy number (**Figure 2**). This process is called chromosomal evolution. When the desired gene copy number is reached, *recA* is deleted, thereby fixing the copy number.



**Figure 2: Chemically inducible chromosomal evolution.** (a) The CICHe construct containing the gene of interest, a selectable marker (green) and homologous regions (red) is integrated into the *E. coli* chromosome. (b) Tandem gene duplication through *recA* homologous recombination. (c) Chromosomal evolution: by adding antibiotics to the medium the selection pressure increases and only cells with a higher gene copy number are able to survive. When the desired copy number is reached, *recA* is deleted<sup>2</sup>.

It has been shown that approximately 40 and even up to 50 gene copies can be attained, using chromosomal evolution<sup>2, 7</sup>. It has also been demonstrated that, while plasmid-bearing strains lose their productivity after 40 generations due to allele segregation, gene copy number and productivity of CICHe-strains remain stable even after 70 generations<sup>2, 7</sup>. This genetic stability is considered to be the most important asset of CICHe.

CICHe has already been applied for the recombinant expression of poly-3-hydroxybutyrate (PHB)<sup>2</sup> and lycopene pathways<sup>2, 7</sup>. The lycopene yield in CICHe-strains was found to be 60% higher than in plasmid-bearing strains<sup>2</sup>. Using CICHe-strains for the recombinant production of PHB resulted in equivalent yields compared to plasmid-bearing strains<sup>2</sup>. The use of CICHe-strains poses nonetheless several advantages. They require no selection markers and can be cultured, after deletion of *recA*, without the addition of antibiotics to the medium. This is in contrast to strains containing plasmids, which still need antibiotic selection. Also, in CICHe-strains yields can be tuned by varying the antibiotic concentration during chromosomal evolution.

## 2.4 Antibiotic resistance, a major drawback of CICHe

In the CICHe technique described by Tyo *et al.* (2009), a gene for antibiotic resistance is used as a selection marker<sup>2</sup>. During chromosomal evolution, the construct delivered to the chromosome is duplicated. So is the antibiotic resistance gene, as this is the driving pressure for the tandem

duplication. The antibiotic resistance gene used by Tyo *et al.* (2009) is the type I *cat* gene, coding for chloramphenicol resistance<sup>2, 8</sup>. A disadvantage of this gene, is that after multiple duplication, it causes a significant enlargement of the genome: the gene consists of 1102 bp<sup>9</sup>. This causes a heavy burden on the genome, rising as the gene copy number increases. Therefore, a CICHÉ model with a smaller marker gene would be favourable.

Our biggest concern, however, is the fact that the final result of this CICHÉ technique is a strain carrying multiple antibiotic resistance genes. Once the gene is introduced into the genome, it will be passed on to next generations of the strain via vertical gene transfer. And even though the DNA cassette's homologous regions have low affinity to the genome of *E. coli*<sup>2</sup> and the genome of other, possibly pathogenic, strains, horizontal gene transfer cannot be entirely ruled out.

The fact that the strain eventually contains multiple antibiotic resistance genes, withholds this CICHÉ technique from being applicable in for example food industry, given several regulations as well as consumer's mind-set.

## **2.5 Our goal**

Our goal is to improve the current model for CICHÉ by eliminating the antibiotic resistance gene. By doing so, CICHÉ would become more widely applicable in industrial biotechnology. The whole industry would be able to benefit from the advantages of stabilized gene duplication without having to worry about hyper-resistant bacteria.

We will take a closer look into several mechanisms that may be able to replace the current system. After weighing their pros and cons we will develop a new model in detail, using the mechanism of our choice.

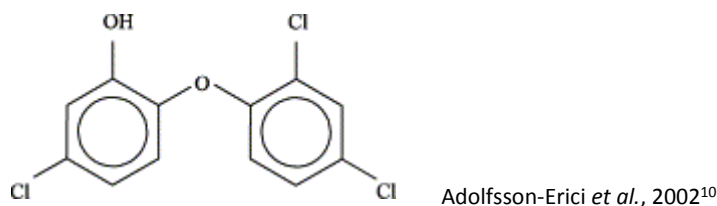
### 3 Alternatives to antibiotic resistance

In this section several alternative mechanisms, which may be able to replace the system with antibiotics and an antibiotic resistance gene, will be examined. What we are looking for is a mechanism to put pressure on cells and to create an advantage for cells with increased construct copy numbers.

#### 3.1 Triclosan and *fabI* overexpression

##### 3.1.1 Mechanism of triclosan inhibition of bacterial fatty acid synthesis

An alternative to antibiotic markers for selection could be the overexpression of a host essential gene to enable selection in the presence of a chemical inhibitor of the gene product. Recently, the compatibility of the biocide triclosan (**Figure 3**) as a selective agent with CICH<sub>E</sub> has been demonstrated by Chen *et al.* (2013)<sup>7</sup>. As plasmid-borne marker the growth essential target gene *fabI* was used. *fabI*-triclosan is a well-studied gene-inhibitor pair. FabI, also known as enoyl acyl carrier protein (ACP) reductase, catalyses fatty acid elongation. Triclosan prevents a proper functioning of FabI.

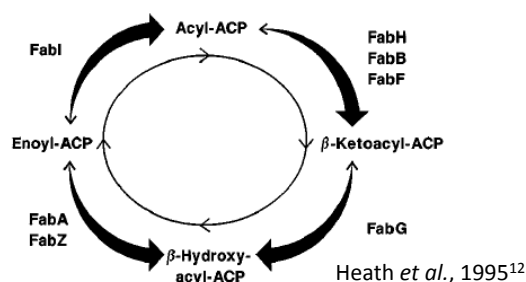


**Figure 3: Structure of triclosan, a member of the bisphenol family<sup>10</sup>.**

The first stage in membrane lipid biogenesis is the fatty acid biosynthesis. In most bacteria, this pathway is catalysed by a series of small, soluble proteins that are each encoded by a discrete gene. This type of synthesis is called the type II fatty acid synthase (FAS) system. The type I FAS of eukaryotes, in contrast, is a dimer of a single large, multifunctional polypeptide. Therefore, the bacterial pathway possesses several unique sites ideal for selective inhibition by chemotherapeutic agents<sup>11</sup>.

There are four basic reactions in each cycle of fatty acid elongation in the type II FAS system (**Figure 4**). The first step is the condensation of acetyl-CoA with malonyl-ACP, forming  $\beta$ -ketoacyl-ACP. This reaction is catalysed by  $\beta$ -ketoacyl-ACP synthase (FabH), the product of the *fabH* gene. The keto ester is then reduced by the NADPH-dependent  $\beta$ -ketoacyl-ACP reductase (FabG) with the formation of  $\beta$ -hydroxyacyl-ACP. The third step is a dehydration in which enoyl-ACP is formed, catalysed by  $\beta$ -hydroxyacyl-ACP dehydrase (either FabA or FabZ). The final step in each cycle is the conversion of *trans*-2-enoyl-ACP to acyl-ACP and is catalysed by the NADH-dependent enoyl-ACP reductase (FabI)<sup>12</sup>. *E. coli* cells possess only a single NADH-dependent enoyl-ACP reductase encoded by the *fabI* gene. Because of the importance of fatty acid biosynthesis for cell growth and functioning of the cell, this pathway is an attractive target for the development of antibacterial agents, such as triclosan<sup>13</sup>. Triclosan is a broad-spectrum antibacterial agent that inhibits FabI through binding at the ACP-enoyl

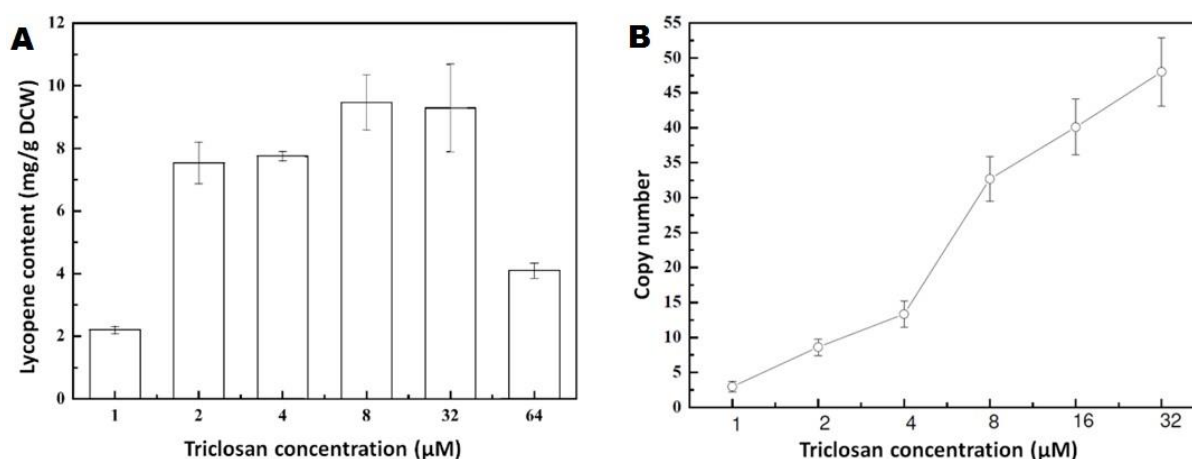
substrate site. This binding increases the enzyme's affinity for nicotinamide adenine dinucleotide (NAD<sup>+</sup>) and results in the formation of a stable FabI-triclosan-NAD<sup>+</sup> ternary complex, which is unable to participate in fatty acid synthesis<sup>14</sup>.



**Figure 4: Fatty acid elongation in the type II FAS system.** Cycles of fatty acid elongation are pulled by enoyl-ACP reductase (FabI). There are four reactions in each cycle of fatty acid elongation. The first step is a condensation reaction catalysed by  $\beta$ -ketoacyl-ACP synthase (FabH) in which acyl-ACP is transformed into  $\beta$ -ketoacyl-ACP. Subsequent cycles of elongation are catalysed by condensing enzyme I (FabB) or II (FabF). This step is followed by a reduction catalysed by FabG, the NADPH-dependent  $\beta$ -ketoacyl-ACP reductase.  $\beta$ -hydroxyacyl-ACP is then dehydrated by either FabA or FabZ, both  $\beta$ -hydroxyacyl-ACP dehydrases, with the formation of enoyl-ACP. The final step is again a reduction step, catalysed by FabI, the NADH-dependent enoyl-ACP reductase. In this reaction, *trans*-2-enoyl-ACP is transformed into acyl-ACP. The inner arrows indicate the overall direction of the cycle. The outer thick arrows indicate the equilibrium positions for the enzymatic reactions<sup>12</sup>.

### 3.1.2 Triclosan induced chromosomal evolution

Chen *et al.* (2013) implemented CICH<sub>E</sub> for the overproduction of lycopene, a carotenoid, using triclosan instead of an antibiotic as the selective agent and the essential growth gene *fabI* of *E. coli* instead of the antibiotic resistance gene as the selective marker<sup>7</sup>. By using a series of integration expression vectors, genes of interest can be site-specifically inserted into *E. coli* by transformation. The desired gene copy number can then be obtained by triclosan induction. Triclosan binds to FabI to stabilize a FabI-triclosan-NAD<sup>+</sup> complex and therefore puts an end to the fatty acid biosynthesis. The inhibition can be reversed through *fabI* overexpression. So the more the *fabI* gene is duplicated, the more triclosan can be tolerated by the cell in order to survive. Lycopene production in CICH<sub>E</sub> strains increases with increasing triclosan concentration during chromosomal evolution. In the research performed by Chen *et al.* (2013), the lycopene production reached a maximum at a triclosan concentration of 8  $\mu$ M. The copy number reached about 30 in the CICH<sub>E</sub> strains, which is the equivalent copy number of a medium copy plasmid. Adding more triclosan caused an increase in gene copy number, but the lycopene production remained the same (**Figure 5**). This indicates that there is an optimal copy number of the *crt* genes (these are the genes responsible for the biosynthesis of carotenoids) for efficient production of lycopene<sup>7</sup>. It may be concluded that the *fabI*-triclosan gene-inhibitor pair is compatible with CICH<sub>E</sub> and thus may be considered as a potential alternative to antibiotic markers for selection.



Chen *et al.*, 2013<sup>7</sup>

**Figure 5: Overproduction of lycopene using triclosan induced chromosomal evolution.** (A) Results of lycopene production in CICH strains resistant to different triclosan concentrations. (B) The *crtI* gene copy number in CICH strains resistant to different triclosan concentrations<sup>7</sup>. The copy number keeps increasing with higher triclosan concentrations, whereas the lycopene production reaches a maximum at 8 μM triclosan<sup>7</sup>.

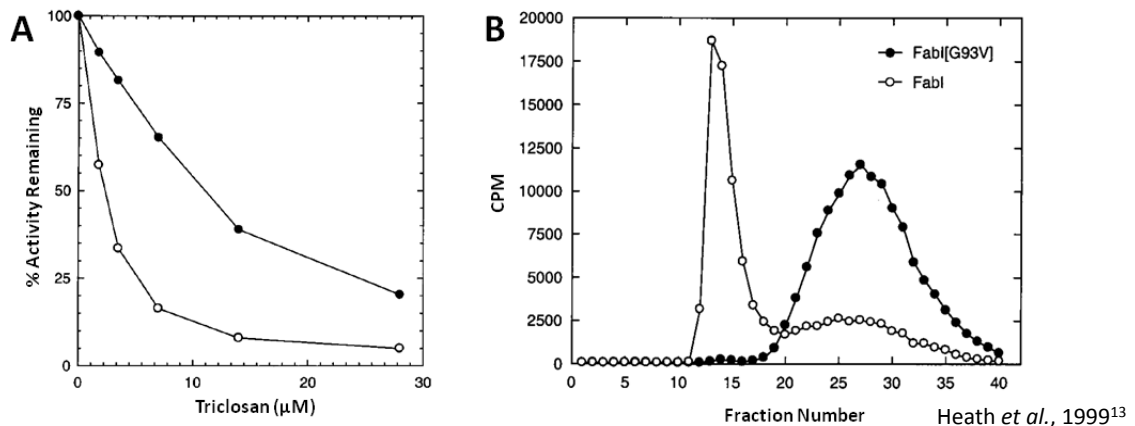
### 3.1.3 Advantages and disadvantages

Triclosan is a biocide that fulfills the criteria of a non-antibiotic. It is a stable, easy to handle and inexpensive chemical and is approved for use in many hygiene, household and industrial applications<sup>14</sup>. It is a chlorinated bisphenol and was originally patented as a herbicide, but was soon realized to possess a broad spectrum antibacterial action<sup>11</sup>. Over the last 30 years, triclosan has become the most potent and widely used bisphenol. It is used in many consumer and health care products, such as hand soaps, surgical scrubs, shower gels, toothpastes, mouthwashes and many more<sup>15</sup>.

A disadvantage is that cells overexpressing the *fabI* gene display toxic effects in the absence of triclosan, suggesting an “addictive” effect<sup>14</sup>. By adding triclosan, the toxicity can be suppressed. Therefore, when overexpression of FabI and triclosan is used separately, it is toxic for the cell, but both combined they enhance growth and production through their gene-inhibitor interaction<sup>14</sup>. A solution for this problem might be the use of an inducible promoter. The expression of *fabI* can be switched on and off by placing the gene under the control of for example a  $P_{BAD}$  promoter (arabinose-inducible *araBAD* promoter). In the absence of the inducer arabinose, no FabI will be produced and thus no toxic effect will occur in the absence of triclosan.

Another disadvantage is the concern that triclosan use may contribute to antibiotic resistance. Resistant mutants were indeed produced in laboratory conditions, but studies of bacteria in non-laboratory conditions exposed to triclosan did not find a correlation between antibiotic resistance and reduced triclosan susceptibility<sup>14</sup>. In the laboratory, three point mutations in *fabI* were observed. They increased triclosan Minimum Inhibitory Concentration (MIC) up to 95-fold<sup>14, 16</sup>. For example, the missense mutation in the *fabI* gene that leads to the expression of FabI[G93V] causes resistance to triclosan. Triclosan does not increase the binding of  $NAD^+$  to FabI[G93V], in contrast to the binding of  $NAD^+$  to the regular FabI (Figure 6). Cells expressing the FabI[G93V] protein have a MIC for triclosan that is 64-fold higher than the wild type strain. The difference in effectiveness of triclosan against

wild-type strains compared with *fabI*<sup>G93V</sup> seems to be the ability of triclosan to irreversibly inhibit FabI, but not FabI[G93V], as a function of the time<sup>13</sup>.



**Figure 6: Differences in activity between FabI and FabI[G93V]. (A)** Triclosan inhibition of FabI and FabI[G93V]: By adding the same concentration triclosan, the mutation in FabI still shows a greater percentage of activity, in contrast to a quick decrease in activity of the normal FabI. **(B)** Triclosan-induced binding of [3H]NAD<sup>+</sup> to FabI and FabI[G93V]: Triclosan induced the high affinity binding of NAD<sup>+</sup> to FabI but not to FabI[G93V]<sup>13</sup>.

Herbert P. Schweizer (2001) suggested a link between triclosan and antibiotics<sup>15</sup>. A cell can become resistant against triclosan by following mechanisms: target mutations, increased target expression, active efflux from the cell and enzymatic inactivation/degradation. These are the same types of mechanisms as those involved in antibiotic resistance; therefore, a link between triclosan and antibiotics can be concluded. The extensive use of triclosan may indeed aid in the development of antimicrobial resistance, in particular cross-resistance to antibiotics. Cross-resistance is the tolerance to a usually toxic substance as a result of exposure to a similarly acting substance. The link between triclosan and antibiotics can be illustrated by two key findings. First, triclosan and antibiotics share multidrug efflux systems as a common mechanism of resistance and they both cause expression of these pumps by selecting similar mutations in the respective regulatory loci. Second, in *M. tuberculosis* the mutation leading to isoniazid resistance was also obtained by selecting for triclosan resistance.

The connection between triclosan and antibiotics has thus clearly been recognised, but there is still not enough knowledge of how this relates to the real world<sup>15</sup>. In a natural environment, triclosan resistance has been slow to develop compared to antibiotic resistance. Possible explanations are the poor solubility and rapid degradation of triclosan, the low competitive fitness of FabI mutants, and the tripartite nature of the FabI-NAD<sup>+</sup>-triclosan complex. However, when the *fabI* gene is inserted in a vector, it is possible that the vector could transfer horizontally and thus induce resistance to triclosan in wild type bacteria. Standard precautions in the handling of genetically modified microorganisms should therefore be maintained<sup>14</sup>.

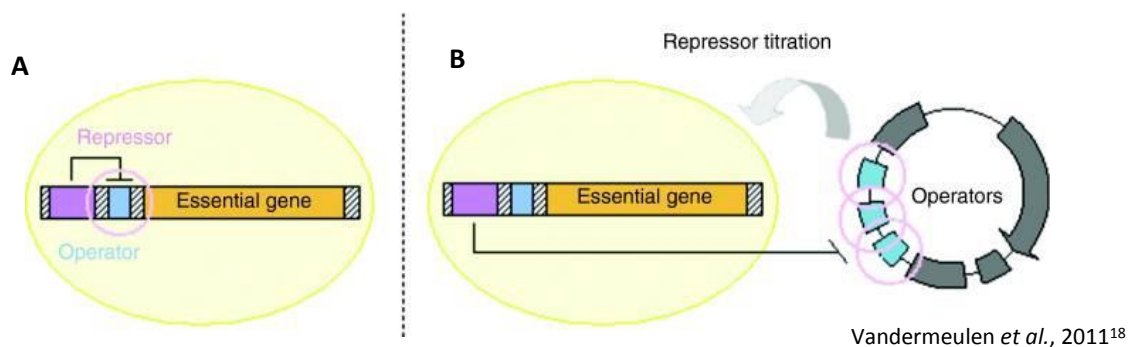


## 3.2 Operator-repressor titration

### 3.2.1 Repressor titration for plasmid selection

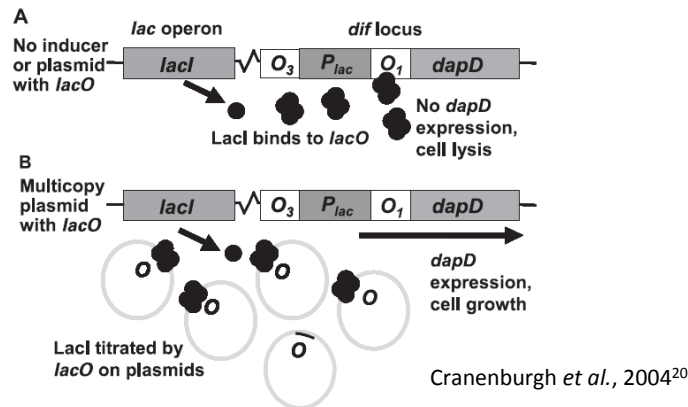
In the search for a plasmid selection mechanism without the use of antibiotic resistance markers, Williams *et al.* (1998) developed a novel system called operator-repressor titration (ORT)<sup>17</sup>.

In ORT, an essential gene is placed under the control of a negatively regulated operator and promoter (e.g. binding of the LacI repressor on *lacO*). As long as the repressor is bound to the operator, the cell is not able to survive. By introducing extra operator sequences in the cell, on a plasmid, the repressor is titrated from the operator that controls the expression of the essential gene (**Figure 7**). As a result, only cells that maintain plasmids containing the operator will be able to survive, in this way allowing selection of a plasmid without the use of an antibiotic resistance marker.



**Figure 7: Operator-repressor titration.** An essential gene is placed under the control of a regulated operator and promoter. **(A)** In the absence of extra operator sequences the gene is repressed and the cell will not be able to grow. **(B)** By adding additional operator sequences to the cell the repressor is titrated from the operator, allowing cell growth<sup>18</sup>.

Williams *et al.* (1998) developed a model system using the kanamycin resistance gene (*kan*) under the control of the lactose operator and promoter<sup>17</sup>. In a medium containing kanamycin, only cells with plasmids containing *lacO* were able to grow, as *lacO* titrated the repressor from the kanamycin resistance gene. However, in this system antibiotics must still be supplemented to the medium for plasmid selection and an antibiotic resistance gene is inserted into the chromosome. To circumvent these drawbacks, a system was designed in which a naturally occurring essential chromosomal gene was placed under the control of *lacO*/P<sup>19</sup>. The essential chromosomal gene *dapD* was used. *dapD* encodes tetrahydrodipicolinate N-succinyltransferase, an enzyme that catalyses an important step in the DAP-synthesis pathway. DAP (diaminopimelate) crosslinks peptidoglycan in the cell wall of bacteria and is a precursor for lysine. Hence, when the *dapD* gene is not expressed, cell lysis occurs. Only in cells containing plasmids with *lacO*, the repressor is titrated from the *lacO* which controls the *dapD* gene and the *dapD* gene is expressed (**Figure 8**).

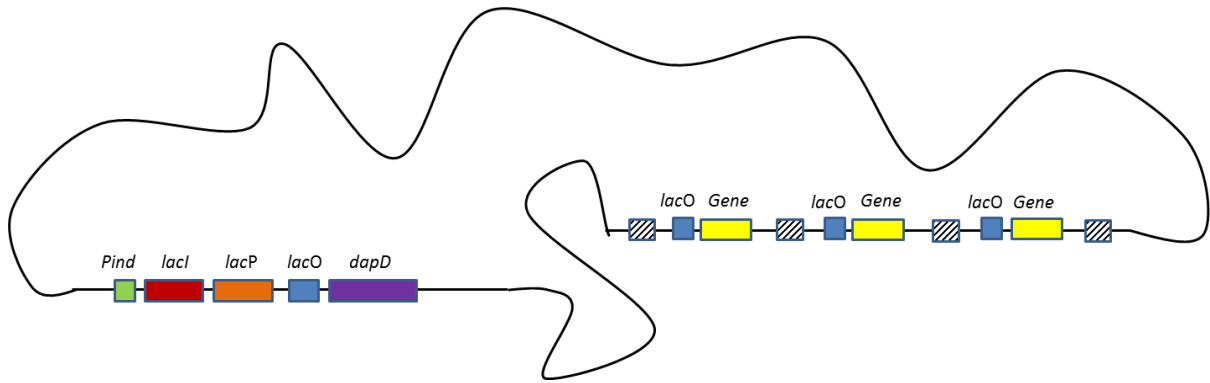


**Figure 8: Repressor titration with *dapD*.** Construct for ORT in (A) the absence and (B) presence of *lacO* containing multicopy plasmids<sup>20</sup>.

The use of repressor titration for plasmid selection and maintenance poses several advantages over the commonly used techniques. There is no need for an antibiotic resistance gene or supplementation of the medium with antibiotics. A number of antibiotic-free selection systems are already available<sup>18</sup>, but operator-repressor titration has an additional advantage over all of these systems. In contrast to the other systems, ORT requires no plasmid borne gene transcription, avoiding an additional metabolic burden for the host organism.

### 3.2.2 Repressor titration and CICH<sub>E</sub>

It may be possible to use repressor titration as a selection tool in chromosomal evolution, replacing the current system of antibiotics and an antibiotic resistance gene. An essential chromosomal gene, such as *dapD*, should be placed under the control of for example *lacO*. This can be done by the same method that Cranenburgh *et al.* (2001) used: insertion of an ectopic *lac-dapD* locus into the *dif* region of the chromosome<sup>19</sup>. In CICH<sub>E</sub> one wants to be able to select those cells with a sufficient number of CICH<sub>E</sub>-construct copies in the chromosome instead of selecting for cells that contain the desired plasmids. The operators, used to titrate the repressor from the *lacO* controlling the essential gene, should be integrated into the construct containing the gene or pathway of interest. By regulating the amount of repressor present in the cell (selection pressure), one could select for those cells that contain a sufficient number of construct copies. Dosage of the repressor, LacI, can be achieved by replacing the constitutive promoter of the *lacI* gene with an inducible promoter (Figure 9).



**Figure 9: Repressor titration and CIChE.** *E. coli* chromosome with CIChE-construct (*lacO* and gene of interest flanked by homologous regions), *dapD* gene under control of *lacO*/P and *lacI* repressor under control of an inducible promoter (*Pind*).

To minimise the size of the CIChE-construct, an ideal version of *lacO* could be used instead of the originally described combination of *lacO*<sub>1</sub> and *lacO*<sub>3</sub><sup>20</sup>. This ideal *lacO* is only 20 bp, instead of the 113 bp for the combination of *lacO*<sub>1</sub> and *lacO*<sub>3</sub> with optimal spacing, and binds the repressor tenfold more tightly than *lacO*<sub>1</sub>. It has been demonstrated that repressor titration levels were higher in systems with the ideal *lacO*<sup>20</sup>.

The question arises if repressor titration levels in this system will be sufficient for chromosomal evolution, since ORT was developed using high copy number plasmids (500-700 copies per cell)<sup>17, 19</sup>. Although it has been shown that lower copy number pBR322 plasmids (39-55 copies per cell) could also reach sufficient repressor titration levels<sup>20</sup>, this does not guarantee the applicability to attain even lower copy numbers during CIChE (1-40). Hence, the sensitivity of the dose-response relationship will have to be investigated, to study whether lower repressor titration levels can be reached, yielding less copies of the gene of interest in the chromosome. In addition, the dosage of repressor protein will need to be very tightly regulated.

### 3.3 Solvents

Many solvents are toxic to microorganisms and are therefore used as disinfectants. By adopting another phenotype, microorganisms can become more tolerant to solvents. Solvent tolerance is a very useful characteristic in industrial biotechnology and may be used as a selection marker in CIChE.

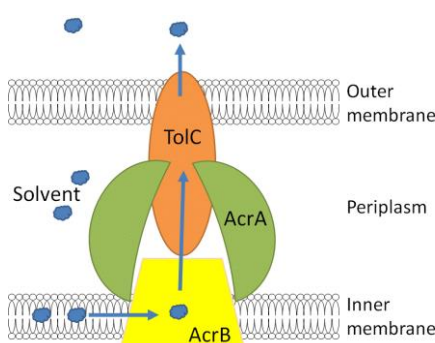
#### 3.3.1 Accumulation of solvents in the membrane

Solvents can accumulate in the cytoplasmic membrane, which increases the fluidity<sup>21, 22</sup>. Some bacteria can spread the solvent through the entire lipid bilayer, diminishing the effect of the solvent. Other microorganisms can modify the density of their membrane by changing their fatty acids<sup>21, 22</sup>. They produce the enzyme cis-trans isomerase to transform the cis unsaturated fatty acids to trans fatty acids, making their membrane less permeable and denser<sup>21, 22</sup>. However, these membrane modifications only cause temporary tolerance and are not sufficient to survive high concentrations of solvents for an extended period of time. Additional strategies are needed to create a more permanent tolerance. This is described in the sections below.

### 3.3.2 Efflux pumps

Some bacteria protect themselves against solvents by producing efflux pumps. In this way they can actively transport antibacterial products out of their cell and reduce the intracellular concentration to subtoxic levels. Research into this strategy is mostly focused on antibiotic resistance, as it is an important mechanism of the natural resistance to antibiotics<sup>22</sup>.

Efflux pumps are membrane transporter proteins, driven by the transmembrane proton gradient. They are energy-dependent. The most efficient efflux pumps, in presence of solvents, are those of the resistance-nodulation-cell division (RND) family<sup>21, 22</sup>. Typical for RND-efflux pumps is that the export of solvents happens in a single energy-coupled step. An efflux pump of this family forms a multicomponent complex of three proteins (**Figure 10**), encoded by the *acrAB-tolC* genes<sup>22</sup>. The first protein, called AcrB, is a transmembrane transporter. The second is a membrane fusion protein, called AcrA and the last protein is a channel protein in the outer membrane, called TolC.



**Figure 10: An efflux pump of the resistance-nodulation-cell division (RND) family.** Efflux pumps of this family consist of three different proteins: TolC, AcrA and AcrB. TolC is an outer membrane protein, AcrA a membrane fusion protein and AcrB is a transporter protein. Solvents enter the cell by porins and accumulate in the membrane. The efflux pumps excrete the solvents and result in a more tolerant phenotype.

### 3.3.3 Chaperon proteins

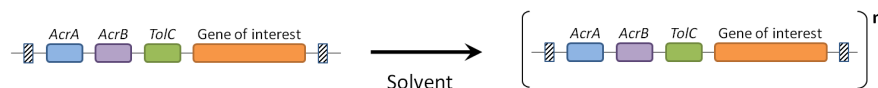
In solvent stress situations, genes involved in heat stress response are up-regulated<sup>23, 24</sup>. These heat stress genes encode chaperon proteins such as heat shock proteins (HSP)<sup>25</sup>. Some examples are GroES and GroEL. These help prevent damage or misfolding of proteins in the cell by refolding and transporting the damaged proteins. This refolding is ATP-dependent.

The overexpression of HSPs results in an improvement of the solvent tolerant phenotype<sup>25</sup>. Single gene integrations and overexpression cause only a minor increase in solvent tolerance, but the co-overexpression of a combination of different HSPs has additional benefits leading to a higher increase in tolerance. Overexpression does however not always lead to improvements<sup>25</sup>. So to conclude, heat shock proteins not always give a higher solvent tolerant phenotype.

### 3.3.4 Solvents and chromosomal evolution

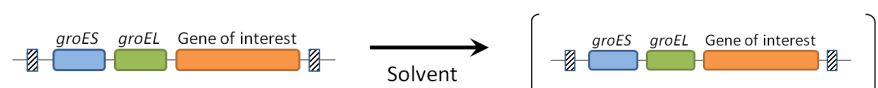
There is a possibility to use the solvent stress responses as selective marker instead of the antibiotic marker in CICH<sub>E</sub>. To succeed in our goal, the solvent may not cause too much damage, because cells would not survive long and this will make it impossible to get expression of the gene of interest. By placing the genes that increase solvent tolerance in the CICH<sub>E</sub> construct, they can be integrated into the chromosome. After transformation, the solvent can be added to the growth medium and chromosomal evolution could be achieved. As the degree of tolerance depends on the quantity of genes causing solvent tolerance, it should be possible to create a selection pressure by raising solvent concentrations in the growth medium and the chromosome should evolve to contain a higher number of construct copies.

The construct could consist of the genes coding for AcrAB-TolC efflux pumps and the gene of interest (**Figure 11**). However, some problems could arise using this construct, as efflux pumps are known to possess some disadvantages. Our biggest concern is that efflux pumps do not only excrete solvents, but also antibiotics and other antibacterial products<sup>22</sup>. Consequently, using efflux pumps would have a similar result in terms of antibiotic resistance as the use of an antibiotic resistance gene. So no progress towards our goal would be made. A second major disadvantage is the perforation of the membrane, caused by these efflux pumps. This could lead to many weak regions and an unstable membrane. Additionally, a high number of pumps would demand an enormous amount of energy if all present pumps are activated.



**Figure 11: CICH<sub>E</sub> construct with *acrAB* and *tolC* as a selection marker.**

Another possible construct can consist of the the *groESL* operon, to create a higher solvent tolerance by producing heat shock proteins (**Figure 12**). The effectiveness of HSPs depends on the inducible promoter and the copy number of the plasmid on which the HSP genes are present<sup>25</sup>. Integrating these genes into the genome bring an improvement by increasing the HSP efficiency. HSPs are only a part of the solvent stress response and try to diminish the effect of solvents. However, the solvent stays intracellular, so interference with other cell components can still occur.



**Figure 12: CICH<sub>E</sub> construct with HSP genes *groES* and *groEL* as a selection marker.**

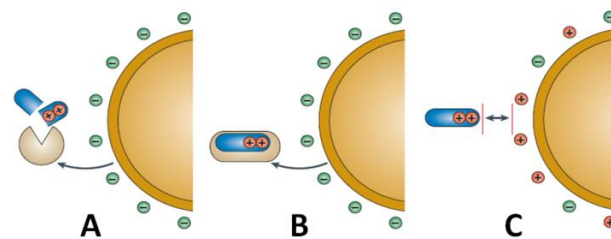
The genes that are responsible for solvent stress tolerance (see above: efflux pumps and HSP) each deliver only a small part of the solvent tolerant phenotype. So, to give bacteria a sufficient degree of tolerance, a combination of many different genes is needed. As a result the construct will be longer. This can cause several problems. On the one hand a long construct makes integration inefficient. On the other hand duplication of a long construct is more difficult and could be harmful to the bacterial cell. The length of the construct will not be the only problem. Solvent tolerance is a result of a

combination of the above described mechanisms. It is not clear if the duplication of only one mechanism is sufficient to tolerate the degree of solvent that will be used as selection pressure in chromosomal evolution. Because it would be very difficult to find an excellent and useful combination of solvent stress tolerance genes and the possible difficulties mentioned above (e.g. membrane instability), tolerance to solvents would not be the best solution to our problem.

### 3.4 Cationic antimicrobial peptides

Another possible alternative to the antibiotic marker may be the use of cationic antimicrobial peptides (CAMP)<sup>26</sup>. CAMPs are linear, microbicidal host-defence molecules that occur in nature. They occur in almost all kinds of life as a protection against microbial organisms. The antibacterial activity can be found in the binding properties to the lipopolysaccharides (LPS) of the bacterial outer membrane<sup>26, 27</sup>, which leads to damage of the membrane.

Bacteria have tried to defend themselves against CAMPs. A number of defence mechanisms were already discovered. Some bacteria can degrade CAMPs by producing peptidases and proteases (**Figure 13.A**)<sup>26</sup>. If proteolysis is too limited, additional strategies are needed, such as the inactivation of CAMPs after binding with inhibitors (**Figure 13.B**). The capturing and elimination of CAMPs depends on direct recognition of certain sequences or structural motifs in the CAMP. The inhibitors prevent the binding to LPS and in this way diminish the membrane damage<sup>26</sup>. Another defence mechanism is altering the net charge of the membrane, as the antimicrobial activity of CAMPs is very dependent on the ionic environment (**Figure 13.C**). This is achieved by a higher production of peptidoglycan, teichoic acids, lipid A and phospholipids<sup>26</sup>. As a result, the normally negatively charged membrane is partially neutralised and this decreases the affinity for CAMPs and repulses them. Microorganisms combine these strategies to heighten their resistance to CAMPs.

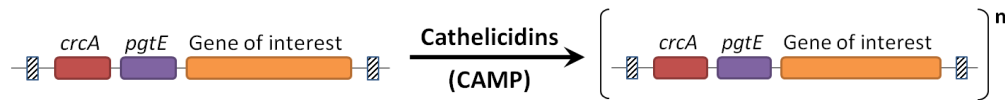


Peschel et al., 2006<sup>26</sup>

**Figure 13: Cationic antimicrobial peptides (CAMP) and the resistance mechanisms of bacteria. (A)** Proteolysis. **(B)** Producing inhibitors: CAMPs are captured and inactivated. **(C)** Modifying the net charge of the membrane to repulse CAMPs<sup>26</sup>.

CAMP tolerance could be used as a selection marker instead of the antibiotic resistance gene in chromosomal evolution. The *phoPQ* operon may be used to achieve this<sup>28, 29</sup>. This operon contains the genes *pgtE* and *pagP*. The former encodes an extracellular membrane protein that cleaves the CAMP cathelicidin by proteolysis (**Figure 13.A**)<sup>29</sup>. The second gene encodes a PagP transferase, which adds palmitate to lipid A<sup>29, 30</sup>. This alters the membrane net charge (**Figure 13.C**). When used in *E. coli*, the gene *crcA* would be better to use than *pagP*, as this gene is a homologue in *E. coli*<sup>28</sup>. The ClChE construct (**Figure 14**) should consist of two genes, namely *crcA* and *pgtE* and cathelicidins (a type of CAMP) could be used to create a selection pressure. By adding different concentrations of

cathelicidins, it might be possible to conclude how many times the gene of interest is present in the cell, because the higher the degree of tolerance, the higher the copy number of the integrated genes.



**Figure 14: Construct for chromosomal evolution by using cationic antimicrobial peptides.** Cathelicidin is used as CAMP in this figure. The gene *crcA* encodes a proteolysis enzyme to cleave CAMPs. The second gene, *pgtE*, encodes a transferase for palmitate in lipid A. This changes the membrane charge and repulses CAMPs.

CAMPs are very similar to antibiotics, but bacteria have not developed highly effective resistance mechanisms to CAMPs yet. In most organisms, there is a huge collection of CAMPs with only minor differences in sequence<sup>26</sup>. This means there is a big cocktail of defence agents, which makes it difficult for microorganisms to become resistant to the entire cocktail of CAMPs at once. Thus even if the bacteria would acquire resistance to one CAMP, it would not have such a dramatic effect as in the case of antibiotic resistance. A disadvantage of CAMPs could be the need for higher concentrations than the antibiotics equivalent. The genes that are responsible for CAMP tolerance each deliver only a small part of the tolerant phenotype. It is not clear if the duplication of only one mechanism is sufficient to tolerate the concentration of CAMPs that will be used as pressure in chromosomal evolution. To give bacteria a sufficient degree of tolerance, a combination of multiple mechanisms will be necessary, which makes the construct longer. This can cause problems similar to those encountered in the solvent tolerance construct. The integration of a long construct is inefficient and duplication of a long fragment is more difficult for the cell.

### 3.5 Toxin-Antitoxin systems

The need for antibiotics may also be eliminated by using toxin-antitoxin systems (TA). These systems are widely distributed among bacteria and archaea. The toxins produced by these systems can slow down or stop cell growth by interfering with certain molecules that are essential in cellular processes like DNA replication, cell wall synthesis, ATP synthesis etc. (see section 3.5.3)<sup>31</sup>. Under normal conditions, the toxin is inhibited by the antitoxin, which is encoded in the same operon as the toxin. Research revealed different types of TA systems, which will be discussed below.

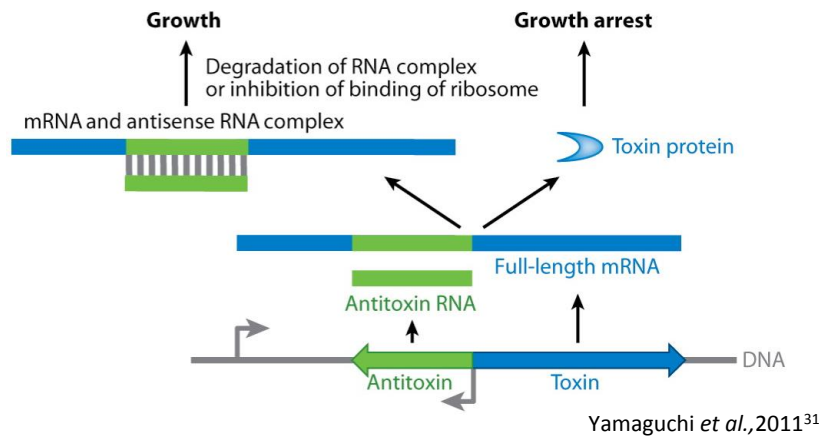
#### 3.5.1 Types of TA-systems

##### 3.5.1.1 Type I

The first category of TA systems is characterised by an antisense mechanism that regulates the expression of the toxin gene<sup>31</sup>. This antisense RNA is transcribed from the same toxin region, but in reversed orientation, and encodes the antitoxin (**Figure 15**). This antitoxin anneals to the toxin mRNA and forms double-stranded RNA across the ribosome binding site<sup>32</sup>. This induces the degradation of the toxin mRNA or the blocking of the ribosome binding site. As a result, the amount of toxin in the cell is reduced.

Type I systems probably evolved by gene duplication and are therefore rare. To date, type I systems are not well represented in chromosomes and their function is still unknown<sup>32</sup>.

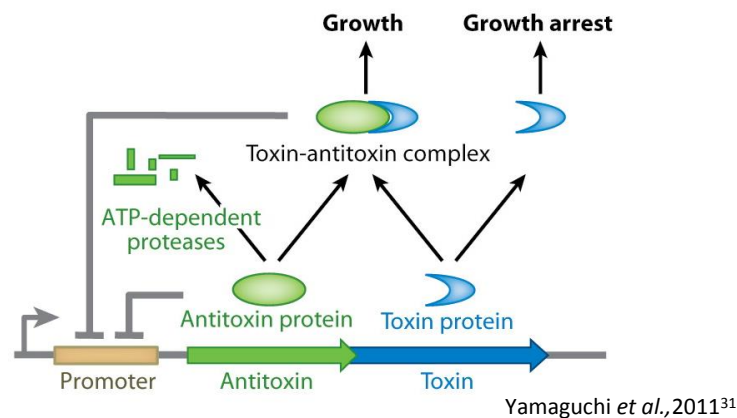




**Figure 15: TA systems type I mechanism.** The antitoxin is localised on the same operon, but is transcribed in reversed orientation. By forming a double-stranded RNA complex, the toxin mRNA is degraded<sup>31</sup>.

### 3.5.1.2 Type II

Antitoxins in type II TA systems are proteins and inhibit the corresponding toxin by forming a stable TA protein complex (**Figure 16**). These systems are encoded by an operon that contains the two genes encoding the toxin and antitoxin. The antitoxin has to be continuously produced, as it is less stable than the toxin. Because of this feature, the cell has a feedback mechanism to prevent growth arrest. The TA complex functions as a repressor that binds the promoter and thereby inhibits transcription. Thus, when the antitoxin is degraded and the concentration of the TA complex decreases, the TA operon is derepressed and produces more toxin and antitoxin. Finally, the concentration of the antitoxin is reestablished. The rank of the two genes differs, which can contribute to different toxin-antitoxin ratios<sup>31</sup>.



**Figure 16: TA systems type II mechanism.** The antitoxin protein inhibits the toxin protein by forming a stable complex with the toxin. This complex (as well as the antitoxin itself) is part of a feed-back inhibition of the promoter<sup>31</sup>.

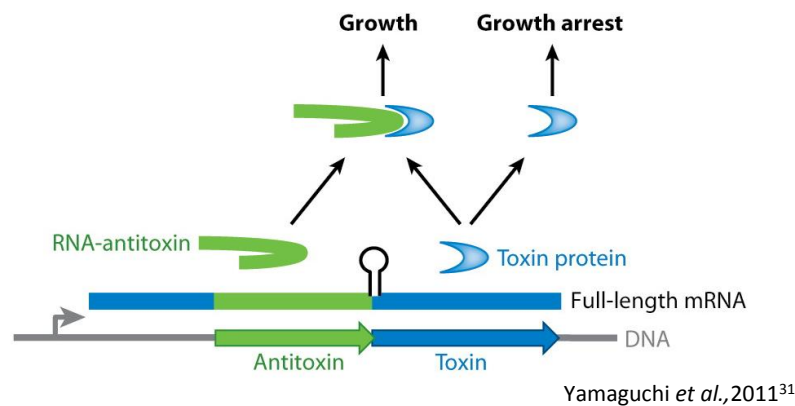
Type II systems are common among the genomes of bacteria and archaea, as they move from one genome to another by horizontal gene transfer. The toxin and antitoxin are dependent, as they are both necessary for survival. Without the antidote the cell dies, but an efficient toxin might be vital to maintain a functional antitoxin<sup>32</sup>.

### 3.5.1.3 Type III

In the last category of TA systems, an RNA antitoxin directly inhibits the toxin<sup>31</sup> (**Figure 17**). These RNA antitoxins are pseudoknots, which contain internal stemloops<sup>33,34</sup>. The toxin and its antidote are



once more encoded by the same operon. A transcriptional terminator between the two genes, regulates the ratio of toxin and antitoxin in the cell<sup>34</sup>.

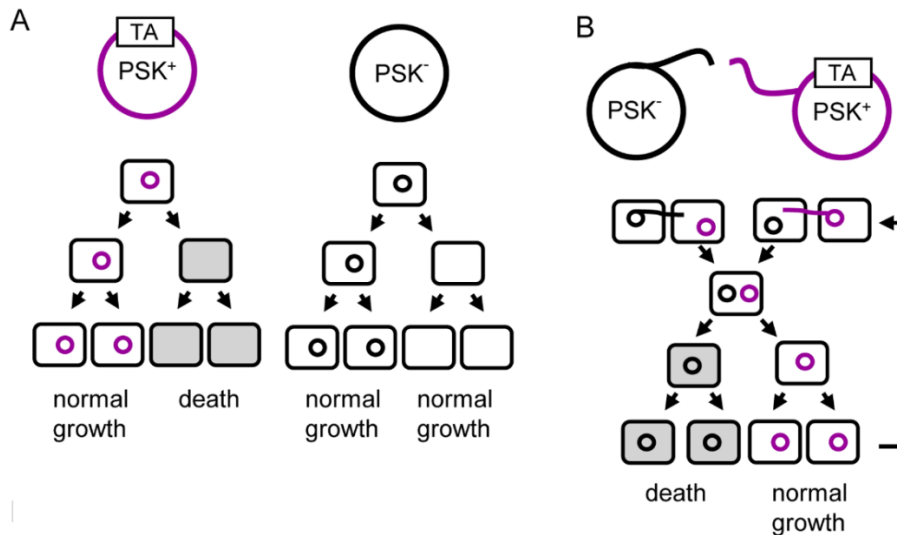


**Figure 17: TA systems type III mechanism.** The antitoxin RNA directly inhibits the toxin protein. A transcriptional terminator regulates the relative amounts of toxin and antitoxin in the cell<sup>31</sup>.

Recently, three new families were identified, suggesting that more remain to be discovered<sup>33</sup>.

### 3.5.2 TA systems on plasmids: postsegregational killing

TA systems were first discovered on plasmids. These TA-systems are also referred to as postsegregational killing (PSK) systems. There are two hypotheses explaining the principle of PSK and they are both based on the occurrence of a PSK<sup>+</sup> (containing a TA system) plasmid and a random PSK<sup>-</sup> plasmid in the cell population. The first hypothesis is based on the instability of the antitoxin after the PSK<sup>+</sup> plasmid loss<sup>35</sup> (**Figure 18.A**). As a result, the more stable toxin is no longer inhibited and kills the PSK<sup>+</sup>-free host cell. This results in a population where the majority of the cells contains the plasmid. Cells that contain a PSK<sup>-</sup> plasmid without a TA system do not suffer from plasmid loss. A second hypothesis is called the selfish gene hypothesis and is based on the will to obtain a competitor-free environment<sup>35, 36</sup> (**Figure 18.B**). When the two plasmids belong to the same incompatibility group, they cannot appear in the same cell. When horizontal transmission takes place, the two plasmids cannot be maintained in the same bacteria and one of them has to be degraded. When the PSK<sup>+</sup> plasmid is degraded, the host is killed, following the first hypothesis. On the other hand, when the PSK<sup>-</sup> plasmid is degraded, the bacteria survives.



Van Melder et al., 2009<sup>35</sup>

**Figure 18: Postsegregational killing. (A)** Vertical transmission. Loss of a PSK<sup>+</sup> (with TA system) plasmid leads to cell death, because the antitoxin is less stable than the toxin. Therefore, the toxin is no longer inhibited. **(B)** Horizontal transmission. The incompatibility of the two plasmids leads to the degradation of one of the plasmids. Bacteria where the PSK<sup>-</sup> plasmid remains in the cell are killed due to the instability of the antitoxin (similar to A)<sup>35</sup>.

### 3.5.3 Chromosomal TA systems: biological roles

Later on, chromosomal TA systems were also identified. The abundance of these TA systems in the genome depends on the type of TA systems. To date, there are many hypotheses and models which try to explain the mechanisms and the roles of TA systems in bacteria. All of them are based on the fact that TA systems are involved in stress management<sup>31, 32</sup>. Under stress conditions, the TA systems kill the cell (programmed cell death, PCD) or put them in a dormant stage (stasis)<sup>32, 35</sup>. Both hypotheses are based on the release of stress-induced proteases and peptides that trigger the TA systems to liberate the toxin by shutting down the antitoxin. The released toxins can interfere with different processes within their hosts<sup>31</sup>.

Many toxins inhibit DNA gyrase, which results in the blocking of DNA replication. Others can cleave mRNA or can interfere with the ribosomes themselves. As a result, translation is inhibited which stops cell growth. Other types of toxins directly inhibit the cell division by binding to the cytoskeletal proteins (like FtsZ, a tubulin-like protein). At last, all type I systems interfere with ATP synthesis by affecting the membrane and the proton gradient. This also results in inhibition of cell growth and a reduced viability.

TA systems are also linked to persistence<sup>31, 35</sup>. Persister cells can survive the presence of antibiotics by entering a dormant state, even though they are genetically sensitive.

Another aspect of the TA systems is their protection against DNA loss. There are TA systems associated with the maintenance of mobile genetic elements. These TA systems are active when the genetic elements are circularized and are not integrated in the chromosome. The PSK system is probably responsible for the maintenance of these genetic elements<sup>32</sup>. TA systems not only interfere with the protection of the host's DNA, but also protect the host against invading DNA. The mechanisms which are responsible for this, are still unknown<sup>32</sup>.

### **3.5.4 The use of TA systems in chromosomal evolution**

In our opinion, a TA system could eliminate the need for antibiotics as a selection tool in ClChE. It may be possible to achieve chromosomal evolution by integrating an antitoxin in the ClChE-construct and adding varying concentrations of the toxin as selection pressure.

Considering all the alternatives discussed above, the use of TA systems seems to be the best option. Some important disadvantages of the other systems made us decide against them: the toxicity of *fabI* and the link between triclosan and antibiotic resistance, the uncertainty of sufficient repressor titration levels in ORT, destabilization of the membrane using efflux pumps and the similarity of CAMPs to antibiotics. We developed a new model for ClChE using a TA system. This is discussed in section 4.

## 4 Developing a new model for chromosomal evolution without antibiotic marker

The goal of our iGEM project is to develop a CICH system independent of antibiotics and antibiotic resistance genes. Therefore, we will investigate the use of TA systems in CICH. Following choices need to be addressed:

- How using a TA system as pressure for tandem replication?
- Which gene of interest will be put into the CICH construct?
- How controlling tandem replication by using *recA* as an on/off-switch?
- Which homology regions will be used?

Below each of these choices are described in detail.

### 4.1 Our alternative to antibiotic resistance: the F plasmid *ccd* TA module

#### 4.1.1 Function and mechanism of the F plasmid *ccd* TA module

The F plasmid *ccd* (control of cell death or coupled cell division) TA module encodes the toxin CcdB (11.7 kDa) and its antidote CcdA (8.7 kDa)<sup>37</sup>. This type II TA system is the first identified and best studied of all TA systems, which is why we have chosen to use it in our model. The target of the CcdB protein is the A subunit of DNA gyrase<sup>38, 39</sup>. This gyrase is an essential type II topoisomerase, occurring in all bacteria but not in eukaryotes. This makes it a popular target for many antibacterial agents, such as the quinolones and coumarins<sup>40</sup>. Bacterial gyrases have the property to introduce negative supercoils into DNA, which makes them unique among topoisomerases<sup>41</sup>.

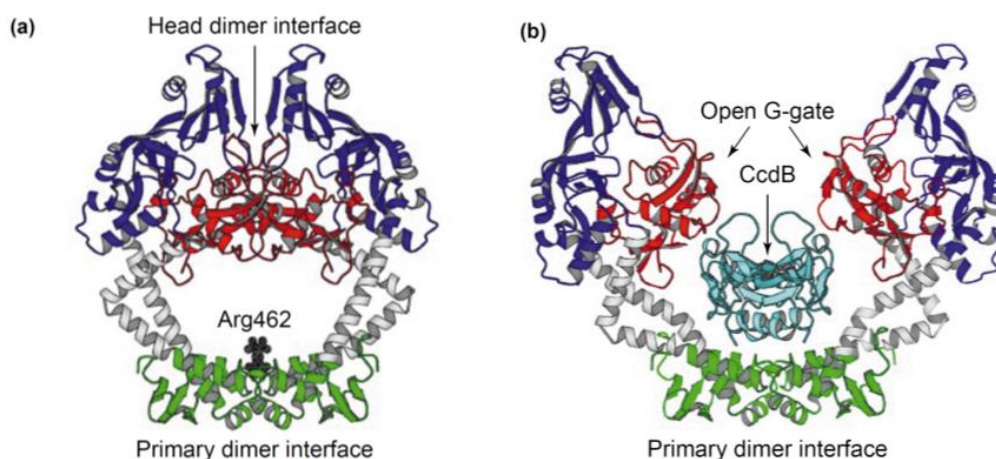
*E. coli* DNA gyrase is a heterotetramer, consisting of two GyrA and two GyrB subunits, forming an A<sub>2</sub>B<sub>2</sub> complex. The major role of the GyrA subunit (97 kDa) consists of the breakage and reunion of DNA, whereas the GyrB subunit (90 kDa) is responsible for ATP binding and hydrolysis<sup>40</sup>. The normal working mechanism of supercoiling of DNA by gyrase initiates with the wrapping of a DNA segment (called the DNA gate or G-segment) around the entire A<sub>2</sub>B<sub>2</sub> complex. This segment then undergoes a double-stranded cleavage, which facilitates the passage of another piece of DNA (called the DNA transfer or T-segment). When this latter piece has passed the break, the broken G-segment is resealed<sup>37, 42</sup>. The intermediate complex in which the enzyme is covalently attached to the 5' phosphate termini of the cleaved DNA G-segment, is called the cleavable complex<sup>39</sup>.

In the absence of its antidote CcdA, CcdB causes reduced DNA synthesis, activation of the SOS pathway, cell filamentation and eventually cell death<sup>37</sup>. When CcdB interacts with a gyrase:DNA complex, it stabilises the cleavable complex<sup>43</sup>. In this mode of action, CcdB acts as a poison, promoting DNA breakage mediated by gyrase, which is an ATP-dependent process<sup>39, 44</sup>. In an in vitro transcription assay in which T7 RNA polymerase and ribonucleotides were added to a CcdB:gyrase:DNA complex, truncated transcripts were formed. This indicates that the ternary CcdB:gyrase:DNA complex can block transcription by RNA polymerases<sup>42</sup>. CcdB can also act as an inhibitor, interacting with free gyrase and thus forming a CcdB:gyrase complex which inactivates the

gyrase<sup>43</sup>. A crucial amino acid of GyrA in the interaction between CcdB and GyrA seems to be Arg462, as a GyrA462 mutant is resistant to CcdB<sup>44</sup>. A model for the binding of CcdB with the 59 kDa of the aminoterminal region of GyrA (GyrA59) is given in **Figure 19**. In this figure it can be seen that when CcdB is bound, it forces the catalytic domain of gyrase into an open conformation<sup>45</sup>.

The antidote CcdA can inhibit CcdB by the formation of a tight non-covalent complex<sup>45</sup>. This prevents CcdB from making a covalent complex with gyrase. But even after CcdB has formed a complex with gyrase, CcdA can still reverse the blocking CcdB caused<sup>42</sup>. When CcdA and CcdB are present in a 1:1 ratio, they bind to the DNA operator/promoter region as a (CcdA<sub>2</sub>CcdB<sub>2</sub>)<sub>n</sub> complex, which indicates that the *ccd* module is autoregulated<sup>46</sup>.

As long as no F plasmid-loss occurs, CcdA and CcdB are both present and hence CcdA prevents CcdB from being toxic. When F plasmids do get lost, however, CcdA is fastly degraded by Lon protease<sup>47</sup> ( $t_{1/2} \approx 30$  min in the absence of CcdB and  $t_{1/2} \approx 60$  min in the presence of CcdB)<sup>48</sup>, whereas CcdB has a higher stability and thus remains in the cell without inhibition by CcdA. This makes it possible for CcdB to perform its toxic effect on DNA gyrase.



Couturier *et al.*, 1998<sup>37</sup>

**Figure 19: Interaction between CcdB and GyrA.** Red: head dimer interface; green: primary dimer interface; dark blue: tower domain; light grey: connecting helices; turquoise: CcdB dimer **(a)** Free GyrA59 (the 59 kDa amino-terminal breaking-rejoining domain of GyrA) in its closed conformation. Arg462 is essential in GyrA-CcdB complex formation. **(b)** Opening up of the GyrA dimer creates space for a CcdB dimer. When CcdB binds to this conformation, religation of the DNA G-segment is prevented<sup>37</sup>.

#### 4.1.2 Advantages and disadvantages of the F plasmid *ccd* TA module in CICH

An advantage of using the *ccdA* gene as selectable marker is its relatively small size. CcdA consists of 72 amino acids and has a molecular mass of 8.3 kDa, which makes it roughly three times smaller than the antibiotic marker used by Tyo *et al.* (2009)<sup>2, 8, 49</sup>. It might not even be necessary to use the complete *ccdA* gene, as experiments have shown that truncated CcdA with the first 36 amino acids missing (CcdA<sup>37-72</sup>) suffices to protect *E. coli* cells from the toxic effects of CcdB<sup>50</sup>. Hence, the integration construct would be significantly smaller than that used by Tyo *et al.* (2009)<sup>2</sup>. A possible disadvantage is that an altered use of the *ccd* TA module in our model might cause unexpected interactions in the cell.

#### 4.1.3 Our use of the F plasmid TA system

We will attempt to use the CcdA-CcdB TA system as a selection tool for CICHÉ. As CcdA has no toxic effects in the absence of CcdB, the *ccdA* gene will be used as selectable marker integrated in the CICHÉ construct, under the control of a constitutive promoter. By adding a plasmid bearing the *ccdB* gene under the control of a chemically inducible promoter, we suggest it is possible to regulate the amount of CcdB present in the cell titrating the chemical inducer. If this CICHÉ system works, cells with a higher gene copy number could be attained by titrating higher concentrations of the chemical inducer, as these have more CcdA to compensate for the toxic CcdB.

When selecting a suitable plasmid for genetic engineering, an important factor that needs to be considered is the plasmid copy number. The copy number determines the gene dosage available for expression and depends on the plasmid's mode of replication, which is controlled by its origin of replication<sup>51</sup>. Many high or low copy number plasmids are known.

In nature, the genes encoding the CcdB toxin and the CcdA antidote in *E. coli* are found together in the *ccd* operon on F plasmids<sup>52</sup>. These plasmids are large (95 kb), conjugative plasmids with a low copy number<sup>37</sup>. In order to ensure that the used *ccd* system will be effective in this experiment, it seems evident to apply a method that mirrors the natural mechanism by using a low copy number plasmid. Generally in overexpression studies, high copy number plasmids are used, because high gene expression is desired. In this experiment, however, the produced protein has toxic effects on the host cell. Therefore, low copy plasmids are more suitable<sup>51</sup>. On top of that, low copy number plasmids have a lot of advantages, such as segregational stability and low metabolic burden on the host strain<sup>53</sup>. Therefore, a low copy number plasmid will be used to provide *E. coli* with CcdB in our chromosomal evolution system.

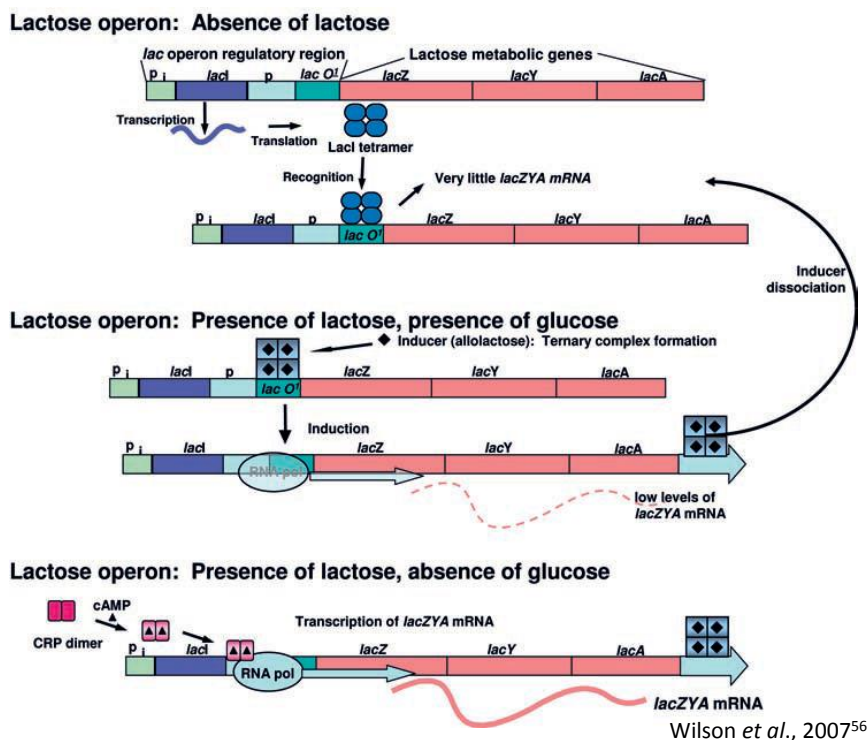
F plasmids contain both the *ccdA* and *ccdB* gene. In this experiment, however, only the *ccdB* gene is desired on the plasmid, hence normal F plasmids cannot be used. A first possibility to solve this problem would be to utilise an F plasmid without the *ccdA* gene. Another possibility would be to clone the *ccdB* gene into another low copy plasmid. This has already been done before in high copy number plasmids. For example the pKIL18/19 plasmid is now used as positive-selection vector<sup>54</sup>. Because the iGEM competition offers an enormous library of standardised parts (BioBricks), cloning the *ccdB* gene into a low copy plasmid, available in the Registry of Standard Biological Parts, seems to be the best option.

#### 4.1.4 Choosing an inducible promoter for *ccdB*

To increase the gene copy number in the chromosome, we will have to be able to regulate the amount of CcdB in the cell. To achieve this, an inducible promoter will be used. An inducible promoter is a combination of an operator and a promoter. The operator may be switched on or off. This, in turn, controls the transcription of the gene. Sometimes it is possible to regulate transcription according to the amount of stimulus the promoter receives. This technique is commonly used in the biotechnology industry for protein overexpression. It is needed because a continuous production of a

foreign protein could cause a metabolic burden on the cell, which may lead to the inhibition of cell growth. An inducible promoter can be used to prevent this. In this way, the production of the desired protein only starts when the ideal conditions are met<sup>55</sup>.

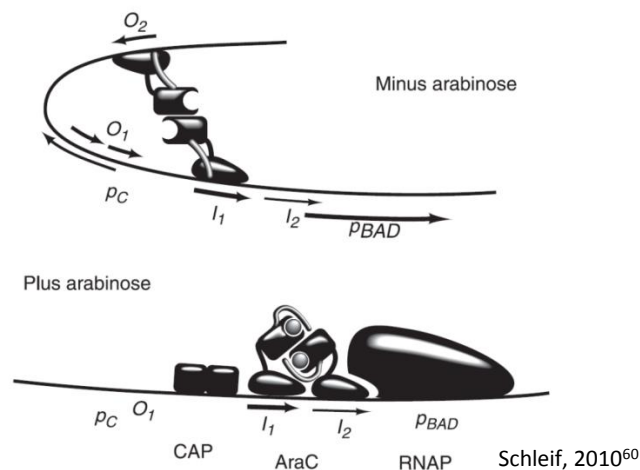
The mechanism of an inducible promoter was first discovered by Jacob and Monod in 1961, based on the metabolism of lactose in *E. coli*<sup>56</sup>. It was known that the bacteria prefer to utilise glucose over lactose as energy source. The organism started to metabolise the lactose only when all glucose was depleted. They discovered that a protein, the repressor (LacI), inhibits the transcription of the required genes to metabolise lactose by binding with a regulatory element, called the operator (*lacO*). When inducer molecules bind to the repressor, its affinity for the operator is reduced. This releases the repressor from the operator and enables transcription. The natural inducer for LacI is allolactose, an analogue of lactose. Yet, when *E. coli* is supplied with both glucose and lactose, it still metabolises glucose first. This is because the RNA polymerase has a low affinity to the promoter. When glucose concentrations are low, cyclic AMP (cAMP) is made. This binds onto the cyclic AMP-dependent catabolite activator/repressor protein (CAP), which in turn binds onto the promoter. This enhances the affinity of the RNA polymerase. Thus, the cell will metabolise lactose when there is a low concentration of glucose and a high concentration of lactose (**Figure 20**)<sup>56,57</sup>.



**Figure 20: The lactose operon.** In the absence of lactose, LacI is formed and binds onto the *lac* operator. This inhibits the transcription of the genes (*lacZ*, *lacY* and *lacA*). When lactose and glucose are present, allolactose binds to the LacI tetramer. The affinity of LacI changes and it releases the operator. This allows the RNA polymerase (RNAP) to bind to and transcribe the genes. Because the affinity of the RNAP to the promoter is low, there is only a low amount of transcription. However, when glucose is absent but lactose is present, cAMP is generated. This binds onto CAP (cyclic AMP-dependent catabolite activator/repressor protein) and together they increase the affinity of RNAP for its promoter. This increases the transcription<sup>56</sup>.

A commonly used inducible promoter is based upon the mechanism of the *lac* promoter, in combination with Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), which cannot be hydrolysed, as inducer<sup>58</sup>. IPTG is also capable of diffusing through the cell membrane without the help of LacY<sup>59</sup>.

Another inducible promoter is  $P_{BAD}$ , originating from the *araBAD* operon. It uses AraC as regulator, which is influenced by the presence of arabinose. Without arabinose, AraC binds to the  $I_1$  and  $O_2$  sites and causes the DNA to loop (**Figure 21**). Once arabinose binds to AraC its conformation changes so that it now binds to the  $I_1$  and  $I_2$  sites. This makes it possible for the polymerase to bind and to start transcription. This promoter is also stimulated by CAP<sup>60,61</sup>.

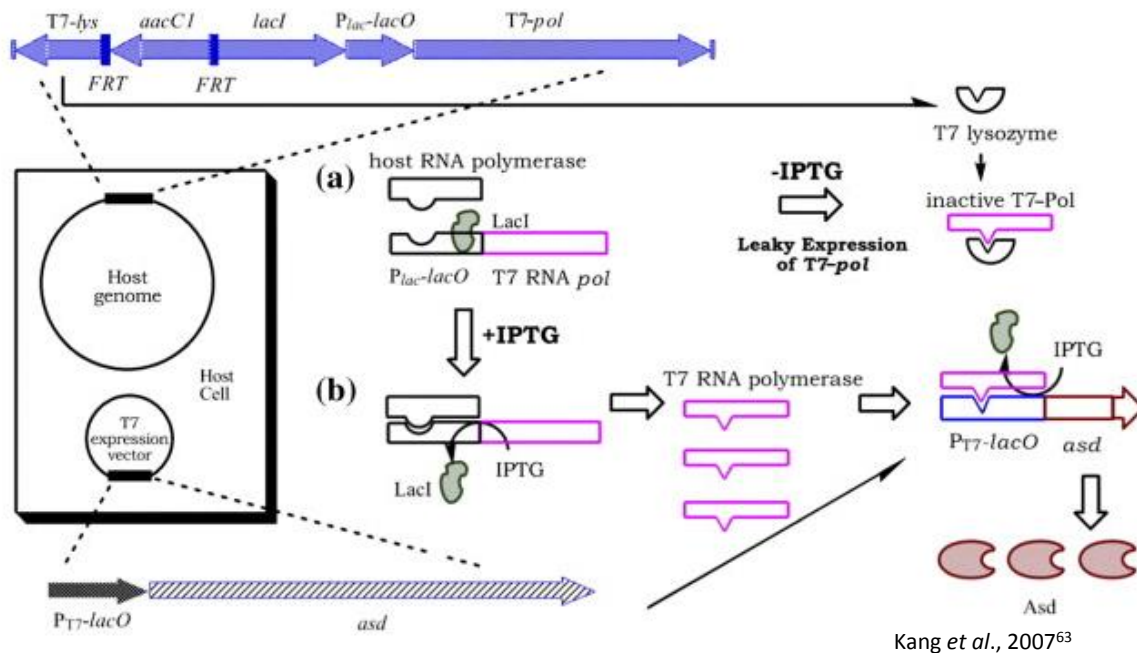


**Figure 21: The conformational change of AraC.** When there is no arabinose in the cell, AraC binds to the  $I_1$  and  $O_2$  site. This forms a loop in the DNA, which inhibits the transcription of the genes. When arabinose is present, AraC changes conformation and the loop disappears. The RNA polymerase (RNAP) is able to bind and transcribe the genes. The binding of RNAP is also enhanced by CAP<sup>60</sup>.

A somewhat different inducible promoter is T7. This promoter originated from the bacteriophage T7, along with the T7 RNA polymerase (T7 RNAP)<sup>62</sup>. This protein binds onto its specific promoter, and transcribes the associated genes. The T7 promoter is not recognised by the *E. coli* RNA polymerase<sup>63</sup>. This allows for good regulation of the transcription of genes, because there is no transcription as long as there is no T7 RNAP present<sup>63</sup>. By using an inducible promoter to regulate the production of T7 RNAP, it is possible to change the amount of transcription of the desired protein. This is usually done by an IPTG inducible promoter<sup>63,64</sup>.

Leaky expression, i.e. transcription that occurs when no stimulus is given to the promoter, can be reduced by using T7 lysozyme, an inhibitor of T7 RNAP. This inhibits the small amount of T7 RNAP that might be formed by leaky expression of the *lac* promoter used to transcribe T7 RNAP<sup>65</sup>. This is ideal for the transcription of a toxic protein, because even a small amount of transcription can be a burden to the cell and restrict cell growth<sup>66</sup>. A disadvantage of the use of T7 lysozyme is the weakening of the cell wall, as it cuts bonds in the cell wall of *E. coli*<sup>64,65</sup>. Another way to reduce the basal expression is adding glucose to the growth medium<sup>64</sup>. This limits the formation of cAMP, which plays a role in leaky expression<sup>67</sup>. A third way to reduce the amount of basal expression is by placing a *lac* operator close to the T7-promoter. This way, when LacI binds to the operator, T7 RNAP has a reduced chance to bind to the promoter and transcribe the desired gene<sup>66</sup> (**Figure 22**).





**Figure 22: Protein expression system using T7.** (a) Without IPTG, the LacI repressors bind to the *lac* operators, preventing the transcription of T7 RNA polymerase (T7 RNAP). There is still leaky expression, resulting in a low amount of T7 RNAP. This is intercepted by T7 lysozyme, which inhibits the polymerase. Further expression is inhibited by binding of LacI to the *lac* operator (*lacO*) of the gene of interest. (b) When IPTG is added, LacI is removed from the operators. This causes T7 RNAP to be produced in a greater amount than the T7 lysozyme. T7 RNAP binds to the T7 promoter, which is now free of LacI. This results in the transcription of the gene of interest, here represented by *Asd*<sup>63</sup>.

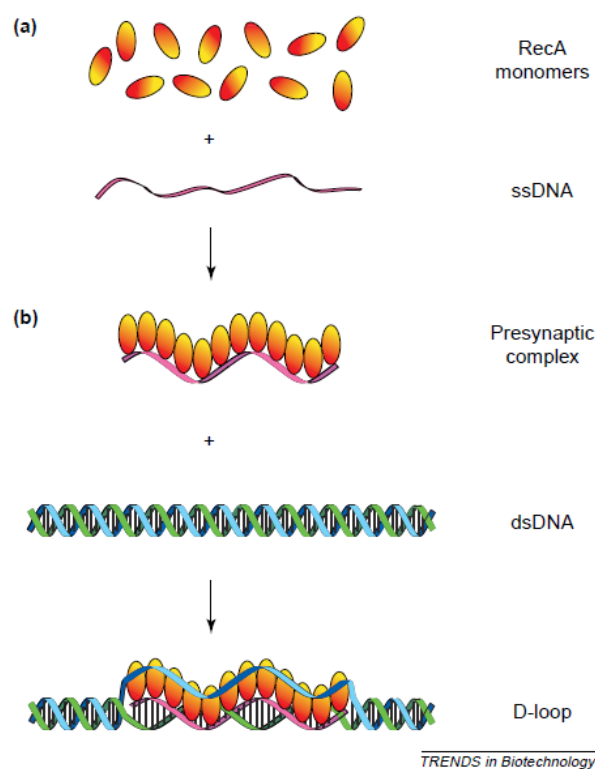
Because both the arabinose promoter and the lactose promoter show higher leaky expression<sup>64</sup>, we will be using the T7 promoter for the production of CcdB. It will be induced by IPTG, in a concentration ranging from 0,05 to 2,0 mM<sup>64</sup>. By using different concentrations we will try to find a correlation between the applied IPTG concentration and the amount of gene copy numbers in the genome.

## 4.2 Selecting a gene of interest

For our experiment, it would be useful to select a gene of interest that can easily be identified. Using a fluorescent reporter gene such as *gfp* is the obvious choice. GFP is a protein consisting of 238 amino acids, which absorbs blue light and emits green light<sup>68</sup>. This protein was first discovered in the jellyfish *Aequorea Victoria*, where it produces a stable green fluorescent light. When GFP is brought to expression in *E. coli* this also results in the production of green light when illuminated with an ultraviolet source<sup>68</sup>. A great advantage of GFP is the fact that no substrate is needed for its expression. Hence, GFP is ideal to measure gene expression, as it is not limited by substrate availability. The intensity of the fluorescence is proportional to the GFP presence in the cell, providing a first indication of the number of *gfp* genes present in the cell<sup>69</sup>. Consequently, fluorescence intensity could give us a first insight in the number of construct copies attained through chromosomal evolution.

### 4.3 The role of RecA in homologous recombination

CICHe uses homologous recombination as a method to achieve multiple copies of the gene of interest. This is achieved using RecA (coded by *recA*), a 38 kDa enzyme that is part of a family of ATPases. These enzymes are responsible for homologous recombination, a process that is essential to repair DNA<sup>70</sup>. RecA helps damaged DNA to find an intact homologous sequence and brings these two strands together<sup>71</sup>. It does so by polymerizing onto a single-stranded DNA molecule to form a helical nucleoprotein, called the presynaptic complex. This filament then searches a homologous region in a double-stranded DNA (dsDNA) molecule. When it binds with the dsDNA, it forms a D-loop, where one strand of the DNA is replaced by the single-stranded homologous DNA<sup>72</sup> (**Figure 23**). This is coupled with the hydrolysis of ATP<sup>73</sup>.



**Figure 23: Function of RecA.** (a) The RecA monomers polymerise onto the single-stranded DNA (ssDNA) molecule, and form the presynaptic complex. (b) The presynaptic complex searches for homologous regions in double-stranded DNA (dsDNA), where it forms a D-loop and integrates the ssDNA into the dsDNA<sup>72</sup>.

If both DNA molecules are double-stranded, linear and one has a single-stranded tail, it is possible that a cross-stranded structure is formed. This is called a Holliday junction<sup>74</sup>. It happens when a double-stranded break is repaired<sup>73</sup>.

When the desired number of copies of the gene of interest is reached, *recA* must be deleted. This stabilises the gene copy number, as there is no homologous recombination without RecA. An easy way in which this could be done is by adding the gene *recA* to the plasmid with *ccdB* and using an *E. coli* strain without *recA* in its genome. This way, if we take away *ccdB*, *recA* will also be deleted, fixating the copy number. A possible way to remove *ccdB* and *recA* in one step is by putting them on a heat sensitive plasmid. When *ccdB* and *recA* are no longer needed, they can be removed simply by

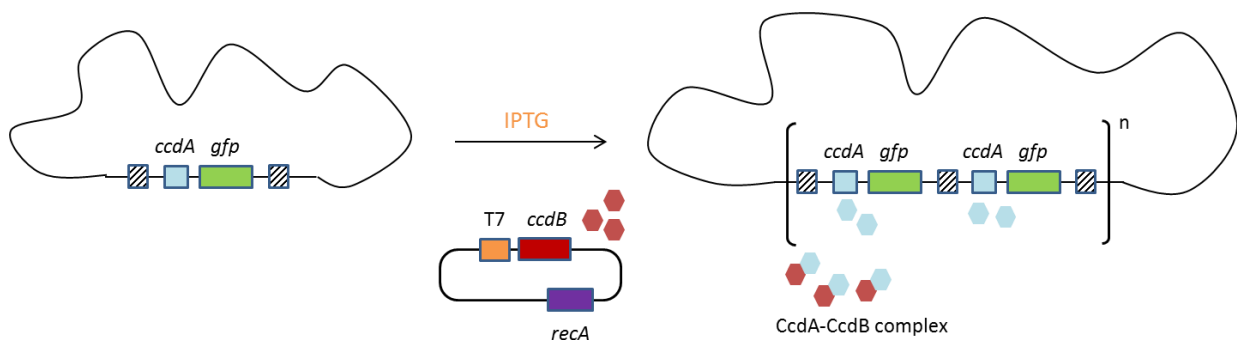
increasing the temperature. Another possibility is to use an *E. coli* strain that does have *recA* in its genome. When *recA* is no longer needed it is disabled using the method for single gene knock out developed by Datsenko and Wanner<sup>75</sup>. This method removes the gene through homologous recombination<sup>75</sup>.

#### 4.4 Selecting the appropriate homology regions for recombination

Homologous recombination is defined as the exchange of DNA between two molecules with an identical DNA sequence<sup>76</sup>. Such a sequence is called a homologous region. Thus, these regions need to be known in order for the crossover to take place. In the original CICE technique, a 1 kb region of the *chlB* gene from *Synechocystis* sp. PCC6803 is used as homologous flanking region at both ends of the gene of interest<sup>2</sup>. This region is selected because of its size and sequence. The efficiency of recombination increases with increasing homology<sup>77</sup>. However, the sequence cannot be too long, which is why a size of 1 kb is taken. *chlB* is a gene encoding subunits of light-independent protochlorophyllide reductase in *Synechocystis* sp. PCC6803<sup>78</sup>. The utilised sequence is noncoding and foreign in *Escherichia coli*. It is selected because it has low homology to the *E. coli* genome. This is required in order to ensure that the recombination is specific. If the sequence would have high homology to the genome, the gene of interest would be inserted into the genome too often. Because this region showed to be efficient in the experiment of Tyo *et al.* (2009)<sup>2</sup>, this region will also be used in our experiment.

#### 4.5 Our model for CICE: an overview

The figure below gives an overview of the model for CICE we will be testing, integrating all choices made in the sections above.



**Figure 24: Our model for CICE.** *E. coli* chromosome with construct containing the *ccdA* antitoxin gene and the *gfp* reporter gene, flanked by the homology regions of our choice (left). A heat sensitive plasmid containing *recA* and the *ccdB* toxin gene is transformed into the cell. Expression of *ccdB* is controlled by the inducible T7-promoter. By raising IPTG (inducer) concentrations, the selection pressure increases and chromosomal evolution is achieved.

## References

1. Randy, R. *et al.* (2013). Synthetic biology based on standard parts. <http://igem.org> (read April 16, 2013).
2. Tyo, K.E.J., Ajikumar, P.K. & Stephanopoulos, G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nature Biotechnology* **27**, 760-U115 (2009).
3. Dong, X.Y., Quinn, P.J. & Wang, X.Y. Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for the production of L-threonine. *Biotechnology Advances* **29**, 11-23 (2011).
4. Bentley, W.E., Mirjalili, N., Andersen, D.C., Davis, R.H. & Kompala, D.S. Plasmid-encoded protein - The principal factor in the metabolic burden associated with recombinant bacteria. *Biotechnology and Bioengineering* **35**, 668-681 (1990).
5. Glick, B.R. Metabolic load and heterologous gene-expression. *Biotechnology Advances* **13**, 247-261 (1995).
6. Silva, F., Queiroz, J.A. & Domingues, F.C. Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnology Advances* **30**, 691-708 (2012).
7. Chen, Y.-Y. *et al.* Chromosomal evolution of *Escherichia coli* for the efficient production of lycopene. *BMC biotechnology* **13**, 6 (2013).
8. Cunningham, F.X., Sun, Z.R., Chamovitz, D., Hirschberg, J. & Gantt, E. Molecular structure and enzymatic function of Lycopene Cyclase from the *Cyanobacterium synechococcus* sp strain PCC7942. *Plant Cell* **6**, 1107-1121 (1994).
9. Alton, N.K. & Vapnek, D. Nucleotide sequence analysis of the chloramphenicol resistance transposon Tn9. *Nature* **282**, 864-869 (1979).
10. Adolfsson-Erici, M., Pettersson, M., Parkkonen, J. & Sturve, J. Triclosan, a commonly used bactericide found in human milk and in the aquatic environment in Sweden. *Chemosphere* **46**, 1485-1489 (2002).
11. Heath, R.J., White, S.W. & Rock, C.O. Lipid biosynthesis as a target for antibacterial agents. *Progress in lipid research* **40**, 467-497 (2001).
12. Heath, R.J. & Rock, C.O. Enoyl-acyl carrier protein reductase (*fabI*) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*. *Journal of Biological Chemistry* **270**, 26538-26542 (1995).
13. Heath, R.J. *et al.* Mechanism of triclosan inhibition of bacterial fatty acid synthesis. *Journal of Biological Chemistry* **274**, 11110-11114 (1999).
14. Goh, S. & Good, L. Plasmid selection in *Escherichia coli* using an endogenous essential gene marker. *BMC biotechnology* **8**, 61 (2008).
15. Schweizer, H.P. Triclosan: a widely used biocide and its link to antibiotics. *FEMS Microbiology Letters* **202**, 1-7 (2001).
16. McMurry, L.M., Oethinger, M. & Levy, S.B. Triclosan targets lipid synthesis. *Nature* **394**, 531-532 (1998).
17. Williams, S.G. *et al.* Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. *Nucleic Acids Research* **26**, 2120-2124 (1998).
18. Vandermeulen, G., Marie, C., Scherman, D. & Preat, V. New Generation of Plasmid Backbones Devoid of Antibiotic Resistance Marker for Gene Therapy Trials. *Molecular Therapy* **19**, 1942-1949 (2011).
19. Cranenburgh, R.M., Hanak, J.A.J., Williams, S.G. & Sherratt, D.J. *Escherichia coli* strains that allow antibiotic-free plasmid selection and maintenance by repressor titration. *Nucleic Acids Research* **29** (2001).
20. Cranenburgh, R.M., Lewis, K.S. & Hanak, J.A.J. Effect of plasmid copy number and *lac* operator sequence on antibiotic-free plasmid selection by operator-repressor titration in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology* **7**, 197-203 (2004).

21. Segura, A. et al. Solvent tolerance in Gram-negative bacteria. *Current Opinion in Biotechnology* **23**, 415-421 (2012).
22. Fernandes, P., Ferreira, B.S. & Cabral, J.M.S. Solvent tolerance in bacteria: role of efflux pumps and cross-resistance with antibiotics. *International Journal of Antimicrobial Agents* **22**, 211-216 (2003).
23. Horinouchi, T. et al. Transcriptome analysis of parallel-evolved *Escherichia coli* strains under ethanol stress. *Bmc Genomics* **11**, 11 (2010).
24. Torres, S., Pandey, A. & Castro, G.R. Organic solvent adaptation of Gram positive bacteria: Applications and biotechnological potentials. *Biotechnology Advances* **29**, 442-452 (2011).
25. Zingaro, K.A. & Papoutsakis, E.T. Toward a Semisynthetic Stress Response System To Engineer Microbial Solvent Tolerance. *Mbio* **3**, 9 (2012).
26. Peschel, A. & Sahl, H.G. The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature Reviews Microbiology* **4**, 529-536 (2006).
27. Gunn, J.S. & Miller, S.I. PhoP-PhoQ activates transcription of *pmrAB*, encoding a two-component regulatory system involved in *Salmonella typhimurium* antimicrobial peptide resistance. *Journal of Bacteriology* **178**, 6857-6864 (1996).
28. Ernst, R.K., Guina, T. & Miller, S.I. *Salmonella typhimurium* outer membrane remodeling: role in resistance to host innate immunity. *Microbes and Infection* **3**, 1327-1334 (2001).
29. Pilonieta, M.C., Erickson, K.D., Ernst, R.K. & Detweiler, C.S. A Protein Important for Antimicrobial Peptide Resistance, YdeI/OmdA, Is in the Periplasm and Interacts with OmpD/NmpC. *Journal of Bacteriology* **191**, 7243-7252 (2009).
30. Miller, S.I., Ernst, R.K. & Bader, M.W. LPS, TLR4 and infectious disease diversity. *Nature Reviews Microbiology* **3**, 36-46 (2005).
31. Yamaguchi, Y., Park, J.H. & Inouye, M. in Annual Review Genetics, Vol 45, Vol. 45. (eds. B.L. Bassler, M. Lichten & G. Schupbach) 61-79 (Annual Reviews, Palo Alto; 2011).
32. Van Melderren, L. Toxin-antitoxin systems: why so many, what for? *Current Opinion in Microbiology* **13**, 781-785 (2010).
33. Blower, T.R. et al. Identification and classification of bacterial Type III toxin-antitoxin systems encoded in chromosomal and plasmid genomes. *Nucleic Acids Research* **40**, 6158-6173 (2012).
34. Short, F.L. et al. Selectivity and self-assembly in the control of a bacterial toxin by an antitoxic noncoding RNA pseudoknot. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E241-E249 (2013).
35. Van Melderren, L. & De Bast, M.S. Bacterial Toxin-Antitoxin Systems: More Than Selfish Entities? *Plos Genetics* **5**, 6 (2009).
36. Cooper, T.F. & Heinemann, J.A. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 12643-12648 (2000).
37. Couturier, M., Bahassi, E. & Van Melderren, L. Bacterial death by DNA gyrase poisoning. *Trends in Microbiology* **6**, 269-275 (1998).
38. Maki, S., Takiguchi, S., Miki, T. & Horiuchi, T. Modulation of DNA supercoiling activity of *Escherichia coli* DNA gyrase by F plasmid proteins. *Journal of Biological Chemistry* **267**, 12244-12251 (1992).
39. Bernard, P. et al. The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *Journal of Molecular Biology* **234**, 534-541 (1993).
40. Maxwell, A. DNA gyrase as a drug target. *Trends in Microbiology* **5**, 102-109 (1997).
41. Gellert, M., Mizuuchi, K., Odea, M.H. & Nash, H.A. DNA gyrase: An enzyme that introduces superhelical turns into DNA. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 3872-3876 (1976).
42. Critchlow, S.E. et al. The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: Induction of DNA cleavage and blocking of transcription. *Journal of Molecular Biology* **273**, 826-839 (1997).

43. Maki, S., Takiguchi, S., Horiuchi, T., Sekimizu, K. & Miki, T. Partner switching mechanisms in inactivation and rejuvenation of *Escherichia coli* DNA gyrase by F plasmid proteins LetD (CcdB) and LetA (CcdA). *Journal of Molecular Biology* **256**, 473-482 (1996).
44. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology* **226**, 735-745 (1992).
45. Dao-Thi, M.H. et al. Molecular basis of gyrase poisoning by the addiction toxin CcdB. *Journal of Molecular Biology* **348**, 1091-1102 (2005).
46. Dao-Thi, M.H. et al. Intricate interactions within the *ccd* plasmid addiction system. *Journal of Biological Chemistry* **277**, 3733-3742 (2002).
47. Vanmelderen, L., Bernard, P. & Couturier, N. Lon-dependent Proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Molecular Microbiology* **11**, 1151-1157 (1994).
48. VanMelderen, L. et al. ATP-dependent degradation of CcdA by Lon protease - Effects of secondary structure and heterologous subunit interactions. *Journal of Biological Chemistry* **271**, 27730-27738 (1996).
49. Murray, I.A. & Shaw, W.V. O-acetyltransferases for chloramphenicol and other natural products. *Antimicrobial Agents and Chemotherapy* **41**, 1-6 (1997).
50. De Jonge, N. et al. Rejuvenation of CcdB-Poisoned Gyrase by an Intrinsically Disordered Protein Domain. *Molecular Cell* **35**, 154-163 (2009).
51. Friehs, K. in *New Trends and Developments in Biochemical Engineering*, Vol. 86. (ed. T. Scheper) 47-82 (2004).
52. Miki, T., Park, J.A., Nagao, K., Murayama, N. & Horiuchi, T. Control of segregation of chromosomal DNA by sex factor F in *Escherichia Coli* - Mutants of DNA gyrase subunit A suppress *letD* (*ccdB*) product growth inhibition. *Journal of Molecular Biology* **225**, 39-52 (1992).
53. Carrier, T., Jones, K.L. & Keasling, J.D. mRNA stability and plasmid copy number effects on gene expression from an inducible promoter system. *Biotechnology and Bioengineering* **59**, 666-672 (1998).
54. Bernard, P., Gabant, P., Bahassi, E.M. & Couturier, M. Positive-selection vectors using the F-plasmid *ccdB* killer gene. *Gene* **148**, 71-74 (1994).
55. Donovan, R.S., Robinson, C.W. & Glick, B.R. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter. *Journal of Industrial Microbiology* **16**, 145-154 (1996).
56. Wilson, C.J., Zhan, H., Swint-Kruse, L. & Matthews, K.S. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cellular and Molecular Life Sciences* **64**, 3-16 (2007).
57. Lewis, M. The *lac* repressor. *Comptes Rendus Biologies* **328**, 521-548 (2005).
58. Fernandez-Castane, A., Caminal, G. & Lopez-Santin, J. Direct measurements of IPTG enable analysis of the induction behavior of *E. coli* in high cell density cultures. *Microbial Cell Factories* **11**, 9 (2012).
59. Keasling, J.D. Gene-expression tools for the metabolic engineering of bacteria. *Trends in Biotechnology* **17**, 452-460 (1999).
60. Schleif, R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *Fems Microbiology Reviews* **34**, 779-796 (2010).
61. Cole, S.D. & Schleif, R. A new and unexpected domain-domain interaction in the AraC protein. *Proteins-Structure Function and Bioinformatics* **80**, 1465-1475 (2012).
62. Kochetkov, S.N., Rusakova, E.E. & Tunitskaya, V.L. Recent studies of T7 RNA polymerase mechanism. *Febs Letters* **440**, 264-267 (1998).
63. Kang, Y., Son, M.S. & Hoang, T.T. One step engineering of T7-expression strains for protein production: Increasing the host-range of the T7-expression system. *Protein Expression and Purification* **55**, 325-333 (2007).

64. Terpe, K. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology* **72**, 211-222 (2006).
65. Studier, F.W. Use of bacteriophage-T7 lysozyme to improve an inducible T7 expression system. *Journal of Molecular Biology* **219**, 37-44 (1991).
66. Dubendorff, J.W. & Studier, F.W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. *Journal of Molecular Biology* **219**, 45-59 (1991).
67. Grossman, T.H., Kawasaki, E.S., Punreddy, S.R. & Osburne, M.S. Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene* **209**, 95-103 (1998).
68. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. Green Fluorescent Protein as a Marker for Gene-Expression. *Science* **263**, 802-805 (1994).
69. Soboleski, M.R., Oaks, J. & Halford, W.P. Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells. *Faseb Journal* **19**, 440-+ (2005).
70. Chen, Z.C., Yang, H.J. & Pavletich, N.P. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**, 489-U483 (2008).
71. Kuzminov, A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiology and Molecular Biology Reviews* **63**, 751-+ (1999).
72. Volodin, A.A., Voloshin, O.N. & Camerini-Otero, R.D. Homologous recombination and RecA protein: towards a new generation of tools for genome manipulations. *Trends in Biotechnology* **23**, 97-102 (2005).
73. Lusetti, S.L. & Cox, M.M. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annual Review of Biochemistry* **71**, 71-100 (2002).
74. Smith, G.R. Homologous recombination in prokaryotes. *Microbiological Reviews* **52**, 1-28 (1988).
75. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6640-6645 (2000).
76. Court, D.L., Sawitzke, J.A. & Thomason, L.C. Genetic engineering using homologous recombination. *Annual Review of Genetics* **36**, 361-388 (2002).
77. Kung, S.H., Retchless, A.C., Kwan, J.Y. & Almeida, R.P.P. Effects of DNA Size on Transformation and Recombination Efficiencies in *Xylella fastidiosa*. *Applied and Environmental Microbiology* **79**, 1712-1717 (2013).
78. Nakamura, K. & Hihara, Y. Photon flux density-dependent gene expression in *Synechocystis* sp PCC 6803 is regulated by a small, redox-responsive, LuxR-type regulator. *Journal of Biological Chemistry* **281**, 36758-36766 (2006).