# Improvement of *Corynebacterium glutamicum* for production of lysine and ectoine from industrial raw materials.

Dissertation

zur Erlangung des Grades

des Doktors der Ingenieurwissenschaften

der Naturwissenschaftlichen-Technischen Fakultät III

Chemie, Pharmazie, Bio- und Werkstoffwissenschaften

der Universität des Saarlandes

von

Rudolf Anton Schäfer

Saarbrücken

2016

Tag des Kolloquiums:	26.07.2016
Dekan:	Prof. DrIng. Dirk Bähre
Berichterstatter:	Prof. Dr. Christoph Wittmann
	Prof. Dr. Elmar Heinzle
Vorsitz:	Prof. Dr. Gert-Wieland Kohring
Akad. Mitarbeiter:	Dr. Björn Beckers

Für meinen Vater Otmar Schäfer.

### **PUBLICATIONS**

Partial results of this work have been published in advance authorized by the Institute of Biochemical Engineering (Technische Universität Braunschweig) and the Institute of Systems Biotechnology (Universität des Saarlandes) represented by Prof. Dr. Christoph Wittmann.

#### Peer reviewed articles:

Buschke N., Becker J., <u>Schäfer R.</u>, Kiefer P., Biedendieck R., Wittmann C.; 2013: Systems metabolic engineering of xylose utilizing *Corynebacterium glutamicum* for production of 1,5-diaminopentane. Biotechnology Journal, 8:557–570

Buschke N., <u>Schäfer R.</u>, Becker J., Wittman C.; 2013: Metabolic engineering of industrial platform microorganisms for biorefinery applications - optimization of substrate spectrum and process robustness by rational and evolutive strategies. Bioresource Technology 135:544–554.

Becker J.\*, <u>Schäfer R.</u>\*, Kohlstedt M., Harder B.J., Borchert N.S., Stöveken N., Bremer E., Wittmann C.; 2013: Systems metabolic engineering of *Corynebacterium glutamicum* for production of the chemical chaperone ectoine. Microbial Cell Factories 12:110.

#### **Conference contributions:**

<u>Schäfer R.</u>, Buschke N., Becker J., Wittmann C.: Systembiotechnologie von *Corynebacterium glutamicum* – Nachhaltige Produktion von Lysin aus Xylose und Hemicellulose, ProcessNet-Jahrestagung, Karlsruhe, 12.09.2012

<u>Schäfer R.</u>, Becker J., Buschke N., Wittmann C.: Systems metabolic engineering of *Corynebacterium glutamicum* –Sustainable production of chemicals from the hemicellulose sugar xylose. BioMicroWorld, Madrid, Spain, 02.10.2013

<sup>\*</sup>Authors contributed equally to the publication

# Contents

Ał	obrev	iations	i	iv
Ał	ostrad	ct		viii
Ζι	ısamı	menfas	sung	ix
1	Intro	oductio	n	1
	1.1	Gener	al introduction	1
	1.2	Main	objectives	5
2	The	oretica	I background	7
	2.1	C. glu	tamicum as industrial workhorse	7
		2.1.1	Core metabolism	7
		2.1.2	Lysine biosynthetic pathway	10
		2.1.3	Product portfolio	11
		2.1.4	Strain development and metabolic engineering	15
		2.1.5	Systems Biology	17
		2.1.6	Systems metabolic engineering	18
	2.2	Altern	ative renewable feedstocks for bio-based production	20
	2.3	Biopro	ocess technology of industrial lysine production	22
3	Mate	erials a	Ind Methods	26
	3.1	Bacter	rial strains and molecular work	26
	3.2	Chem	icals	31
	3.3	Media	for shake flask cultivations	32
	3.4	Media	for bioreactor cultivations	33
	3.5	Cultiv	ation conditions	35
		3.5.1	Shake flask cultivation	35

#### Contents

		3.5.2	Miniaturized cultivations	35
		3.5.3	Fed-batch bioprocesses	36
	3.6	Analyt	ical methods	37
		3.6.1	Sample preparation	37
		3.6.2	Quantification of cell concentration	38
		3.6.3	Quantification of metabolites	38
		3.6.4	Enzymatic assays	40
		3.6.5	Mass spectrometry	40
		3.6.6	Metabolic flux analysis	41
	3.7	Calcul	ation of production parameters	42
		3.7.1	Yields and rates	42
		3.7.2	NADPH supply and demand	44
		3.7.3	Economic parameters	44
4	Res	ults and	d Discussion	45
	4.1	Lysine	production of <i>C. glutamicum</i> at elevated temperature	45
		4.1.1	Growth and product formation	45
		4.1.2	Intracellular metabolic fluxes	48
		4.1.3	Lysine production by <i>C. glutamicum</i> Lys12 at elevated temper-	
			ature	62
		4.1.4	The metabolic response of <i>C. glutamicum</i> to elevated temperat-	
			ure is of industrial value	66
	4.2	Lysine	production from xylose	69
		4.2.1	Growth and lysine production of xylose-utilizing C. glutamicum	
			Xyl1	69
		4.2.2	Production from xylose in a fed-batch environment	70
		4.2.3	The engineered metabolism of C. glutamicum Xyl1 is tailored	
			for lysine production from xylose	76
		4.2.4	Slower kinetics for lysine production from xylose as compared	
			to glucose leaves space for further strain engineering	76
	4.3	Ectoin	e production by metabolically engineered <i>C. glutamicum</i>	79
		4.3.1	Re-design of <i>C. glutamicum</i> Lys1 for ectoine bio-synthesis	79
		4.3.2	C. glutamicum Ect1 efficiently produces ectoine from glucose .	80
		4.3.3	Increased temperature positively affects production	82

#### Contents

Appendix 96		96	
5	Conclusion	and Outlook	93
		production	91
	4.3.10	Metabolic changes point at potential bottlenecks in ectoine	
		performance	90
	4.3.9	High-cell density fermentation reveals excellent production	89
	4.3.8	Combined process and strain engineering increases ectoine	00
		from high salinity in a marker-free host system	86
	4.3.7	C. glutamicum Ect2 enables decoupling of ectoine production	
		Ect2 under fed-batch conditions	86
	4.3.6	Performance of the ectoine producing strain C. glutamicum	
		of the aspartate family	84
	4.3.5	Ectoine synthesis affects the intracellular pools of amino acids	
		ectoine producer	83
	4.3.4	Elimination of lysine secretion generates a second generation	

## **ABBREVIATIONS**

### Acronyms & abbreviations

BHI	Brain heart infusion
BHIS	Brain heart infusion with sorbitol
BM	Biomass
bp	Base pairs
cdm	Chemically defined medium
cdw	Cell dry weight
cww	Cell wet weight
ddH <sub>2</sub> O	Ultra pure water
DO	Dissolved Oxygen
EMP	Embden Meyerhof Parnas
Kan	Kanamycin
LB	Lysogeny broth
MCS	Multiple cloning site
n/a	Not applicable
OD	Optical density
ORI	Origin of replication
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription polymerase chain reaction
SOC	Super optimal/Catabolite repression medium
TCA	Tricarboxylic acid
Tet	Tetracycline

### Symbols

F	Farad
g	gram
g	Gravitational acceleration (9.81 m s <sup>-2</sup> )
h	hour
L	Liter
М	Molar (moles L <sup>-1</sup> )
min	minute
mol	moles; Avogadro constant (6.022 $\cdot 10^{23}$ molecules)
Ω	Ohm
t	tons
V	Volt
V	Volume

### **Metabolites**

AcCoA	Acetyl CoA
AKG	α-Ketoglutarate
ASA	Aspartate semialdehyde
ATP	Adenosine triphosphate
DAP	Diaminopimelate
DHAP	Dihydroxyacetone phosphate
E4P	Erythrose 4-phosphate
F6P	Fructose 6-phosphate
GAP	Glyceraldehyde 3-phosphate
Gly	Glycine
G6P	Glucose 6-phosphate
ICI	Isocitrate
Lys	Lysine
Mal	Malate
NAD	Nicotinamide adenine dinucleotide (oxidized form)
NADH	Reduced form of NAD
NADP	Nicotinamide adenine dinucleotide phosphate (oxidized form)

NADPH	Reduced form of NADP
OAA	Oxaloacetate
P5P	Pentose 5-phosphate
PEP	Phosphoenolpyruvate
PGA	3-Phosphoglycerate
Pyr	Pyruvate
SUC	Succinate
S7P	Sedoheptulose 7-phosphate
Tre	Trehalose

### Genes/Enzymes

asd	Aspartate semialdehyde dehydrogenase
ast	Aspartate transaminase
dapA	Dihydrodipicolinate synthase
dapB	Dihydrodipicolinate reductase
dapC	Succinyl-amino-ketopimelate transaminase
dapD	Tetrahydrodipicolinate succinylase
dapE	Succinyl-diaminopimelate desuccinylase
dapF	Diaminopimelate epimerase
ddh	Diaminopimelate dehydrogenase
ectA	2,4-Diaminobutyrate acetyltransferase
ectB	2,4-Diaminobutyrate transaminase
ectC	Ectoine synthase
ectD	Ectoine hydroxylase
fbp	Fructose 1,6-bisphosphatase
gapA	Glyceraldehyde 3-phosphate dehydrogenase
gdh	Glutamate dehydrogenase
gltA	Citrate synthase
gnd	6-Phosphoglucolactone dehydrogenase
hom	Homoserine dehydrogenase
icd	Isocitrate dehydrogenase
iolT1	myo-Inositol transporter

iolT2	myo-Inositol transporter
lysA	Diaminopimelate decarboxylase
lysC	Aspartokinase
lysE	Lysine permease
malE	Malic enzyme
pck	Phosphoenolpyruvate carboxykinase
pgl	6-Phosphogluconolactonase
ррс	Phosphoenolpyruvate carboxylase
ptsF	Phosphotransferase system (fructose)
ptsG	Phosphotransferase system (glucose)
ptsS	Phosphotransferase system (sucrose)
рус	Pyruvate carboxylase
sdhB	Succinate dehydrogenase
tal	Transaldolase
tkt	Transketolase
xylA	Xylose isomerase
xylB	Xylulokinase
zwf	Glucose 6-phosphate dehydrogenase

### ABSTRACT

*Corynebacterium glutamicum* is of huge importance in the production of amino acids and has an impressive portfolio of substrates and products.

In this thesis it was shown that elevated growth temperature led to decreased biomass yield and specific growth rate and to increased lysine yield. At 38 °C, the lysine yield was ca. 40-55 % higher than at 30 °C. Metabolic flux analysis of *C. glutamicum* Lys12 revealed 101 % flux through the pentose phosphate pathway at 38 °C, as compared to 86 % at 30 °C. In a fed-batch bioreactor setup, the cultivation at 38 °C led to a 10 % higher yield, as compared to 30 °C.

Also, *C. glutamicum* Lys12 was modified to allow for growth and lysine production from xylose via expression of the *Escherichia coli* genes *xylA* and *xylB*. The resulting strain *C. glutamicum* Xyl1 was able to grow at 0.17 h<sup>-1</sup> and converted xylose into lysine at a yield of 0.25 mol mol<sup>-1</sup>. In a more industry-like fed-batch bioreactor setup 116 g L<sup>-1</sup> lysine were produced with a xylose based medium.

Finally, the *Pseudomonas stutzeri ectABCD* operon (with the osmosensitive promoter exchanged for a constitutive one and the codon usage of *C. glutamicum* considered) was expressed in *C. glutamicum* lys $C^{T311I}$  to enable ectoine production. This strain produced ectoine (19 mol mol<sup>-1</sup>) from glucose, but also lysine, which was stopped by deletion of *lysE*. Elevated temperatures led to improved ectoine secretion and 4.5 g L<sup>-1</sup> ectoine could be produced in a fed-batch process at 35 °C.

### ZUSAMMENFASSUNG

*Corynebacterium glutamicum* ist von großer Bedeutung für die Produktion von Aminosäuren und weist ein beeindruckendes Spektrum an Substraten und Produkten auf.

In dieser Arbeit wurde gezeigt, dass erhöhte Wachstumstemperaturen zu reduzierter Biomasseausbeute und spezifischen Wachstumsrate sowie erhöhter Lysinausbeute führten. Bei 38 °C lag die Ausbeute ca. 40-55 % höher als bei 30 °C. Metabolische Flussanalyse zeigte 101 % Kohlenstofffluss durch den Pentosephosphatweg bei 38 °C, während bei 30 °C nur 86 % erreicht wurden. In einem Fedbatch Prozesses im Bioreaktor war die Lysinausbeute bei 38 °C 10 % höher als bei 30 °C.

Das Einbringen der *Escherichia coli* Gene *xylA* und *xylB* in *C. glutamicum* Lys12 ermöglichte außerdem Wachstum mit 0,17 h<sup>-1</sup> sowie Lysinproduktion mit einer Ausbeute von 0,25 mol mol<sup>-1</sup> auf Xylose. In einem Fedbatch Bioreaktorexperiment wurden in einem auf Xylose basierten industrienahen Medium 116 g L<sup>-1</sup> Lysin produziert.

Schließlich wurde noch das *Pseudomonas stutzeri ectABCD* Operon (mit ausgetauschtem Promotor und berücksichtigter Codonverwendung) in *C. glutamicum* Lys1 eingebracht, was die Ectoinproduktion von Glukose ermöglichte (19 mmol mol<sup>-1</sup>). Die Sekretion des Nebenprodukts Lysin wurde durch Deletion von *lysE* unterbunden. Wie bei der Produktion von Lysin führten erhöhte Temperaturen zu einer verbesserten Ectoinproduktion und in einem Fedbatch Bioreaktor konnten bei 35 °C 4,5 g L<sup>-1</sup> Ectoin produziert werden.

### INTRODUCTION

#### **1.1 General introduction**

Biotechnology, i.e. 'any technological application that uses biological systems, living organisms or derivatives thereof, to make or modify products or processes for specific use' [209], has been employed by humans for centuries. Many people are, however, not aware of the importance of this scientific field for their daily lives, even though biotechnology comprises a wide range of applications e.g. in the pharmaceutical or the food industry. The important flavor enhancer glutamic acid falls into the latter category and is produced on an industrial scale by the microorganism *Corynebacterium glutamicum*. This organism is closely linked to its most famous product, glutamate, that it was even named after it.

*C. glutamicum*, a Gram-positive rod shaped microorganism, was isolated in Japan in the mid 1950s [108] during a program screening for microorganisms able to produce glutamic acid. It was soon discovered that *C. glutamicum* is capable of producing other amino acids besides glutamate, first and foremost lysine [107]. Today, *C. glutamicum* is one of the most important microorganisms for fermentative production of amino acids and a workhorse of industrial biotechnology [15]. *C. glu-*



**Fig. 1:** Phase contrast microscopic image of *C. glutamicum* [22].

*tamicum* has earned the status as a industrial workhorse and a platform organism because of its robustness in large scale production processes, proven over many decades [59].

From early on, the increase of lysine yields and titers were major goals of strain optimization to enable more efficient production. Initially, this was realized by random mutagenesis and subsequent selection [98]. More recently, the genome sequence of *C. glutamicum* was unraveled [93] and by now systems metabolic engineering provides the possibility for systems-wide understanding of the intracellular mechanisms of the organism and optimization for improved production [18].

Continuously, this was driven by an increasing demand for lysine throughout the decades, illustrated in Fig. 2. Lysine, i.e. the biologically active L-enantiomer L-lysine, is an essential amino acid for most vertebrates and an important feed additive for the cereal-based fodder of the pig and poultry industry [230]. In such feedstuffs, lysine is one of the limiting amino acids [246]. Its concentration does not satisfy the demand of animal mast. Even though nutritional requirements in animals may vary with age and gender, and despite regional variations in fodder composition, L-lysine is the first limiting amino acid when breeding pigs and the second limiting in breeding poultry [94]. This can be visualized using the concept of LIEBIG'S barrel (Fig. 3) showing that a limitation in one compound can hinder the performance of the entire



**Fig. 2:** Worldwide production of L-lysine and poultry in the past decades. Lysine data until 2005 are taken from [163], 2009 data from [46] and later data/predictions from [142]. Poultry data are taken from [210].

system. Compensation for this limitation by addition of L-lysine leads to higher meat yield and lower ammonia burden in the manure [46]. The still increasing worldwide demand for meat predicts an increased lysine market for the future [142].

*C. glutamicum* is able to naturally utilize a wide range of substrates [125]. This includes many sugars such as glucose, fructose or sucrose, organic acids like succinate and lactate [35]. Traditionally, lysine has been produced by fermentation from sugar-based resources like starch hydrolysates or molasses [102, 238]. Clearly, for the production of a *'high-volume low-value'* product such as lysine, the raw material costs are a decisive factor for the profitability of a production plant [46]. An additional downfall of the traditional substrates is a high price volatility [142, 243]. Coupled to the criticism of using glucose-based, i.e. human consumable resources, to produce animal feed additives, these points provide a strong incentive for the search for alternative raw materials.

A promising candidate is the hemicellulose sugar xylose, a major building block of lignocellulose and widely abundant as a by-product in agriculture and in the pulp and paper industry [130], i.e. lignocellulosic waste products. The pentose



**Fig. 3:** LIEBIG'S barrel as a graphical representation of the nutritional value of lysine as a animal feed additive.

is, however, not a natural substrate for most strains of *C. glutamicum* [59], but an interesting raw material nevertheless. Turning a waste product into a valuable commodity chemical is not only of economic interest, but could also help to create a more ecologically sustainable approach to industrial lysine production [35]. Systems metabolic engineering has been shown to enable *C. glutamicum* to utilize xylose for growth and production by expressing xylulokinase and xylose isomerase from *E. coli* and from other microorganisms [36, 138].

With regard to the product portfolio, *C. glutamicum* is capable of producing a wide range of chemicals beyond L-amino acids: bulk products such as D-amino acids[196], diamines [103, 180], organic acids [155, 156] and biofuels like ethanol [82] or isobutanol [23], among other products [15]. In addition, also high-value chemicals become accessible through *C. glutamicum*.

Ectoine is a compatible solute and interesting agent for the pharmaceutical and cosmetic industry [61, 124, 160]. Ectoine is currently used in products like sunscreen and nasal spray due to its moisturizing, protective and anti-inflammatory qualities [204]. The current production relies on a process called 'bacterial milking', which is technically demanding and corrosive for the equipment [120]. Ectoine is not natively synthesized by *C. glutamicum* [14]. The biosynthetic route to ectoine starts from L-aspartate semialdehyde, a precursor of the biosynthetic route towards lysine. This makes *C. glutamicum* an excellent starting point for development [14].

Besides the molecular set up to handle the demanded bio-conversion efficiently, any microbial strain also has to perform under industrial process conditions in order to be an apt production strain. Large-scale production conditions are inherently more challenging for microorganisms as they comprise e.g. higher levels of compound inhomogeneity [30, 95] and huge variations in the local hydrostatic pressure.

Parameters influence growth and production capabilities of the used microorganisms and include, amongst others, pH value [187], initial substrate concentration [168] and temperature [14, 149, 151]. The set-point of controllable parameters in an industrial context should generally be profit-driven and focus on productivity (producing fast) and yield (producing efficiently). Other economic factors of biochemical engineering might also play a role, e.g. reduced amount of cooling water required.



**Fig. 4:** Microbial conversion of raw materials into value added chemicals. The depicted molecules reflect the substrates and products used in this work. One of each (glucose and lysine) is natural to *C. glutamicum* while the others are not naturally utilized/synthesized.

### 1.2 Main objectives

The aim of this work was to extend the biotechnological performance of *C. glutamicum* by addressing the three segments of the conversion of raw materials into value added chemicals (Fig. 4):

- Production of L-lysine from D-xylose. Implementation of a bioreactor fed-batch process emulating industrial production conditions.
- Characterization the correlation between lysine production and cultivation temperature for different strains of *C. glutamicum*. Unraveling of the underlying metabolic responses to changes in cultivation temperature.
- Broadening of the product portfolio of *C. glutamicum* towards the compatible solute ectoine, a promising substance for a diverse set of possible and existing applications in the cosmetic and pharmaceutical industry.

### **THEORETICAL BACKGROUND**

#### 2.1 C. glutamicum as industrial workhorse

#### 2.1.1 Core metabolism

The soil bacterium *C. glutamicum* is one of the most important microorganisms for amino acid production today [188]. The two amino acids with the largest worldwide production volume by far, L-glutamate and L-lysine, are mainly produced by this microorganism [16]. Contemporary lysine production uses mixtures of sugars (glucose, fructose, sucrose) from sugar cane, starchy biomass or molasses as substrate [102, 238].

The uptake of these sugars is mediated by three phosphoenolpyruvate dependent phosphotransferase systems (PTS, see Fig. 5). A fourth PTS with unknown function is reported in *C. glutamicum* [141]. Additionally, glucose is taken up via non-PTS routes [79, 114, 128]. At least one of these is uninhibited by high osmolarity, which can occur during fed-batch fermentations [60]. It was shown that overexpression of non-PTS glucose uptake in a PTS deficient strain can be beneficial for lysine production [79, 128]. Unlike other bacteria, the PTS uptake in *C. glutamicum* is only marginally affected by catabolite repression [242]. In fact, the uptake systems are rather expressed constitutively, which enables simultaneous consumption of mixtures of sugars [45]. This is a clear advantage for industrial fermentation, where media often contain a variety of such carbon sources.

The central carbon metabolism of *C. glutamicum* comprises the Embden-Meyerhof-Parnas (EMP) pathway, the pentose phosphate (PP) pathway, the tricarboxylic acid (TCA) cycle and a range of anaplerotic reactions around the pyruvate node (Fig.



**Fig. 5:** Sugar uptake in *C. glutamicum* ATCC 13032. PTS = phosphotransferase systems, Fru = fructose, Suc = sucrose, Glc = glucose, PEP = phosphoenolpyruvate, Pyr = pyruvate, G6P = glucose 6-phosphate, Suc6P = sucrose 6-phosphate, F1P = fructose 1-phosphate, FBP = fructose 1,6-bisphosphate, ATP = adenosine triphosphate, ADP = adenosine diphosphate, PolyP<sub>n</sub> = polyphosphate, IoIT = *myo*-inositole transporter, EII = enzyme II permease. GLK, PPGK and GGL2647 are glucokinases. Gene names are shown in brackets. The molecules representing the entry point into the central carbon metabolism are depicted with a dashed frame. Figure adapted from [242] and [79].

6). The EMP pathway catalyzes the conversion of phosphorylated sugars into phosphoenolpyruvate and pyruvate, two molecules that represent important metabolic intermediates and precursers for many metabolites [148]. Adenosine triphosphate (ATP) is generated along the way. The PP pathway comprises an oxidative part, generating NADPH, and a non-oxidative part, consisting of reversible interconversion reactions with erythrose 4-phosphate, ribose 5-phosphate and sedoheptulose 7-phosphate as pathway intermediates. Erythrose 4-phosphate is a precursor for aromatic amino acids [242] and ribose 5-phosphate is a building block for nucleotides [148]. Additionally, the PP pathway is a key route for lysine synthesis [136, 231, 234]. The quantification of metabolic fluxes through <sup>13</sup>C labeling studies revealed that this pathway supplies major amounts of NADPH during growth of C. glutamicum on typical industrial sugars such as glucose [136], sucrose [235] and xylose [36]. Fructose on the other hand is a suboptimal carbon source for lysine bio-synthesis [100]. The different entry point of fructose into the central carbon metabolism, omitting the PP pathway (Fig. 5 and 6) leads to reduced NADPH supply and lysine production [100]. Due to the high NADPH demand during lysine production, the PP pathway has been a focus of strain optimization efforts [10, 13, 149].

The TCA cyle is the core pathway of energy metabolism in many aerobic microorganisms, including *C. glutamicum*. Its major function is the generation of reducing



**Fig. 6:** Central carbon metabolism of *C. glutamicum*. PEP = phosphoenolpyruvate, Pyr = pyruvate, ATP = adenosine triphosphate, ADP = adenosine diphosphate, GTP = guanosine triphosphate, GDP = guanosine diphosphate, FADH = flavin adenine dinucleotide (*reduced*), NAD(P)H = nicotinamide adenine dinucleotide (phosphate) (*reduced*), NAD(P)<sup>+</sup> = nicotinamide adenine dinucleotide (phosphate) (*oxidized*), P<sub>i</sub> = phosphate, CoA(-SH) = coenzyme A (*unbound*), Q(H<sub>2</sub>) = coenzyme Q (*reduced*).

equivalents (NADH, NADPH, FADH) and biochemical energy as ATP. Additionally, the TCA cycle supplies anabolism with precursor molecules. In order to replenish the TCA cycle, *C. glutamicum* possesses reactions that supply TCA cycle intermediates – the so called anaplerotic reactions. For lysine production, the anaplerotic reactions involved in the formation of oxaloacetate/malate (OAA/Mal) from phosphoenolpyr-uvate/pyruvate (PEP/Pyr) are of immense importance as both oxaloacetate and pyruvate are product precursors. *C. glutamicum* has a remarkable array of enzymes at this metabolic node: PEP carboxylase (PEP $\rightarrow$ OAA) [49], Pyr carboxylase (Pyr $\rightarrow$ OAA) [162], PEP carboxykinase (OAA $\rightarrow$ PEP) [89], OAA decarboxylase (OAA $\rightarrow$ Pyr) [88] and malic enzyme (Mal $\rightarrow$ Pyr) [211, 242].

#### 2.1.2 Lysine biosynthetic pathway

The lysine synthesis (Fig. 7 A) starts from the TCA cycle intermediate oxaloacetate as well as from pyruvate. It branches off into two parallel routes at the level of  $L-\Delta^1$ -tetra-hydropicolinate, both of which converge at *meso*-diaminopimelate. One of the two routes, the multi-step succinylase pathway, is active even when only little ammonium is available [48, 182] (Fig. 7 B). In contrast, the one-step ammonia-dependent dehydrogenase pathway, is mainly active at a high concentration of ammonium [194] (Fig. 7 B). In total, the generation of one molecule of lysine requires one molecule of each, pyruvate and oxaloacetate, two molecules of ammonium and the oxidation of four molecules of NADPH to NADP<sup>+</sup>.

Lysine biosynthesis in *C. glutamicum* is mainly controlled by the enzyme aspartokinase, which is feedback inhibited by lysine and threonine. The release of this inhibition is a key to lysine producing strains [98]. Multiple point mutations have been reported in the literature that remove the feedback inhibition by the two amino acids [38, 150, 192, 244].

Since *C. glutamicum* natively does not possess pathways for lysine degeneration, lysine would accumulate inside the cell, if produced in large quantity. Therefore, *C. glutamicum* excretes lysine via a permease, encoded by *lysE* [217]. The low affinity of this permease towards its substrate ensures activity only in case of intracellular lysine accumulation [218].



**Fig. 7:** Lysine biosynthetic pathways in *C. glutamicum*. The precurser molecules are oxaloacetate and pyruvate (inside dashed frames). A: ATP = adenosine triphosphate, ADP = adenosine diphosphate, NADPH = nicotinamide adenine dinucleotide phosphate (*reduced*), NADP<sup>+</sup> = nicotinamide adenine dinucleotide phosphate (*oxidized*), CoA(-SH) = coenzyme A (*unbound*), Glu = glutamate, 2OG =  $\alpha$ -Ketoglutarate. Gene names (red italics) are listed on page vi. B: Carbon flux through the two parallel biosynthetic routes towards lysine at different levels of ammonium concentration [48, 182, 194].

### 2.1.3 Product portfolio

**Production of amino acids.** *C. glutamicum* is the dominant and classical producer of amino acids, originally discovered in a screening program for glutamate producing organisms [108]. Soon after that, L-lysine was also established as a product, secreted by a *C. glutamicum* mutant [107]. Today, these two amino acids are still the main products (by volume) of *C. glutamicum* by far [16], but over time *C. glutamicum* has developed into an industrial workhorse with a large range of products. It is not surprising that *C. glutamicum* has been used to produce other amino acids as L-methionine [25], L-arginine [77], L-tryptophan [74], L-alanine [90] and L-serine [161], reaching impressive titers of over 50 g L<sup>-1</sup> for some of these products [16]. Researchers also made use of the stereospecificity of biological reactions and used

C. glutamicum to produce D-amino acids [196].

**Other products.** Apart from amino acid production, *C. glutamicum* has been used for the generation of biofuels like ethanol [82] or isobutanol [23]. Ethanol is not only a biofuel, but also member of a revised list of promising chemicals that can serve as building blocks in a biorefinery based economy [28]. Glutamic acid, the major product of *C. glutamicum*, had been included in the initial list [227] and organic acids like succinate and lactate are included in both publications. The two organic acids are natural products of *C. glutamicum* during anaerobic cultivation [83]. Using metabolic engineering, titers of 146 g L<sup>-1</sup> (succinic acid [156]) and 120 g L<sup>-1</sup> (lactic acid [155]) could be reached. Other chemicals successfully produced in *C. glutamicum* are diamines like putrescine [139] and cadaverine [34], two important chemicals necessary for the production of bio-based polyamides.

Economical and ecological considerations. As researchers, politicians and the general public become increasingly aware of the scarcity of fossil resources in the recent decades, alternative production routes for fuels and commodity chemicals are sought after, basically as a way to short-cut the global carbon cycle [134]. A biotechnological approach to contribute to the solution of this pressing challenge is the concept of biorefineries, converting biomass (often waste biomass) into valuable products [35, 122]. Likewise, more traditional bioproduction, which relies on e.g. sugars as substrate, generates value added chemicals, i.e. chemicals of higher economical value than the educt, from renewable sources. It has many advantages over petrochemical processes, like stereospecificity, low process temperatures and independence of fossil raw materials, hence no risk of depletion of the substrate. The added monetary value varies widely for the different products. Bulk chemicals with high production volumes are usually less expensive due to the 'volume effect' or 'economy of scale' and because the high demand triggers competition. Fig. 8 illustrates this for the amino acid market, where the substances with the highest annual production (lysine and glutamate) have the lowest market value per unit while lower annual production amounts correspond to higher prices, as is the case for histidine and serine.

The market value of any given substance is a decisive factor when considering the production. In order to maximize profit, the three key performance indicators –



**Fig. 8:** Worldwide annual production and costs of amino acids [46]. All amino acids shown graph can be produced by *C. glutamicum*, although so far not all with economically feasible titers [15, 73, 214]. Arg = arginine, Asp = aspartic acid, Cys = cysteine, Glu = glutamic acid, Gly = glycine, His = Histidine, Ile = isoleucine, Leu = leucine, Lys = lysine, Phe = phenylalanine, Ser = serine, Thr = threonine, Val = valine.

titer, yield and productivity – have to be considered. For any production process it is desirable to attain high product titers in order to decrease the cost of downstream processing. Yet for 'high value' products, this factor might be of lesser importance, as even with increased formulation cost still a profit can be made due to high market price. This applies also to the yield. With valuable products, the relative low influence of inefficient (i.e. low-yield) substrate utilization to the overall profit margin leads to less emphasis on this key performance indicator. The opposite holds true for cheap bulk chemicals like lysine. Here, high titers and yields are crucial, as the substrate cost often makes up a significant portion of the overall cost [41, 98].

**Compatible solutes and ectoine.** Chemical chaperones are a class of substances that comprises small, polar and highly water-soluble organic osmolytes like trehalose, glycine betaine, proline and the tetrahydropyrimidines ectoine and 5-hydroxyectoine [29, 99, 184, 237]. They protect proteins against aggregation, promote their proper folding under otherwise denaturing conditions, and are fully compliant with cellular physiology, biochemistry and protein function [71, 199, 200, 241]. Due to these attractive properties, industry advertises ectoines as protective compounds for health



**Fig. 9:** Metabolic pathway for ectoine and 5-hydroxyectoine production from the building block L-aspartate-β-semialdehyde, synthesized from the TCA cycle intermediate oxaloacetate by the enzymes aspartokinase (Ask; EC: 2.7.2.4) and aspartate-semialdehyde-dehydrogenase (Asd; EC: 1.2.1.11) (see Fig. 7) and then converted into ectoine and 5-hydroxyectoine by L-2,4-diaminobutyrate transaminase (EctB; EC: 2.6.1.76) to form L-2,4-diaminobutyrate, a metabolite that is then acetylated by 2,4-diaminobutyrate acetyltransferase (EctA; EC: 2.3.1.178) to produce *N*-γ-acetyl-2,4-diaminobytuyrate, which is subsequently transformed to ectoine via a water elimination reaction by ectoine synthase (EctC; EC: 4.2.1.108). Ectoine can then serve as substrate for the formation of 5-hydroxyectoine through the activity of ectoine hydroxylase (EctD; EC: 1.14.11.). Gene names (red italics) are listed on page vi.

care products and cosmetics [61, 124, 160]. Ectoines possess excellent stabilizing effects on biological molecules; e.g. proteins, cell membranes, DNA, and even entire cells. To some extent they also find application as *in vivo* folding catalyst for the recombinant production of proteins [8] and as enhancers for polymerase chain reactions [181]. The anti-inflammatory effect of ectoine even suggests a medical oriented application in the future for treating lung inflammation [204] and colitis [1], and for tissue protection in ischemia [223]. These properties coupled with the current industrial bio-process for ectoine production being elaborate and detrimental to the equipment [33, 120] render the heterologous production of ectoines in *C. glutamicum* an interesting aim.

Although chemical synthesis of ectoines is certainly possible [181], their large-scale production with a high degree of purity and stereo-specificity is complicated and costly. Chemical synthesis was consequently out-competed by a biotechnological production route using the halophilic bacterium *Halomonas elongata* [175]. A major problem with this production is the requirement of high-salinity growth medium, which leads to slow production, corrosion of the equipment and an overall difficult to handle process [147].

In vivo, ectoine is synthesized from the precursor L-aspartate- $\beta$ -semialdehyde

(ASA), a central hub in microbial amino acid production [131], by three successive enzymatic steps that are catalyzed by the L-2,4-diaminobutyrate transaminase (EctB), the 2,4-diaminobutyrate acetyltransferase (EctA) and the ectoine synthase (EctC) [132, 157] (Fig. 9). A substantial subgroup of the ectoine producers can convert ectoine into 5-hydroxyectoine through the activity of the ectoine hydroxylase (EctD) [33, 170] (Fig. 9). The ectoine biosynthetic genes are normally organized in an operon (*ectABC*) [118, 119, 132, 169] that might also comprise the *ectD* gene [33, 57, 165]. Expression of the genes involved in the ectoine synthesis is typically induced in response to increased osmolarity [37, 118, 119, 184] or extremes in growth temperature [37, 119]. Transcriptional profiling indicated that the cellular levels of ASA could be a potential bottleneck in the synthesis of ectoines [21].

All in all, through the combination of its natural set-up as a robust and versatile soil microorganism and of metabolic optimization and systems metabolic engineering (see chapters 2.1.4 - 2.1.6), *C. glutamicum* is capable to produce a variety of value-added chemicals.

#### 2.1.4 Strain development and metabolic engineering

The improvement of *C. glutamicum* for industrial fermentation is a major goal since many decades. Due to a lack of insight into microbial biochemistry and genetics, at first, initial production strains of C. glutamicum could only be generated via selection, often coupled to active random mutagenesis. For example, glutamic acid producing strains were obtained by exposure of cell suspensions to UV radiation and selection involving coverage of mutated colonies with glutamate auxotrophic bacteria [127]. A second efficient approach for selection of desirable traits was the use of antimetabolites or analogs. A prominent example is the use of S-(2-aminoethyl) cysteine, an analog of lysine, and L-threonine, which led to the discovery of strains with a mutated aspartokinase, uncoupled from feed-back regulation [98]. These approaches led to the development of powerful production strains able to accumulate 100 g  $L^{-1}$  of L-lysine with yields of up to 50 % [72, 126]. On the other hand, random mutagenesis inherently caused the integration of hundreds of unspecific mutations into the genome. These can lead to detrimental changes and unwanted traits like weak stress response, poor growth capability or extensive nutritional requirement [9].

With a deeper understanding of microbial genetics, a rapidly increasing set of tools for their manipulation and the sequencing of the complete genome of *C. glutamicum* [72, 93, 207], genetic, i.e. metabolic engineering became feasible to rationally generate strains with improved production characteristics. One important tool for example was the levansucrase based system for selection of mutants, which allows simple identification of positive clones during integrative modification of *C. glutamicum* genes [84, 176].

A metabolic engineering project where this tool was successfully applied, was the production of ethanol in *C. glutamicum* [82]. Disruption of the native genes *ldhA* and *ppc* and the heterologous expression of the *Zymomonas mobilis* genes *adhB* and *pdc* led to an substantial increase in the ethanol produced. Metabolic engineering was also applied to *C. glutamicum* to broaden the substrate spectrum. Episomal expression of the *E. coli xylA* gene coding for xylose isomerase enabled xylose utilization [97]. The likewise employment of the genes *glpF*, *glpK* and *glpD*, also from *E. coli*, led to growth and lysine production from pure and crude glycerol [138, 171]. Another successful approach of metabolic engineering was the replacement of the native D-lactate dehydrogenase promoter of *C. glutamicum* by the strong promoter of the superoxide dismutase of the same organism [146]. This modification led to an improved utilization of lactate as a substrate and ultimately enhanced lysine production from that substrate.

Despite these successful examples, the inherent limitation, faced by metabolic engineering approaches, boils down to a lack of system wide knowledge. This is also the reason, why many bacterial modifications that can be identified as pure metabolic engineering, aim at more or less obvious targets like the usage of heterologous genes to enable novel substrates or products. Due to the complexity of the biochemical pathways and their regulation, metabolic engineering at a more intricate level requires a better understanding of the system to be modified, as presented in chapters 2.1.5 and 2.1.6.

#### 2.1.5 Systems Biology

The systems-wide, holistic approach toward understanding complex biological systems is referred to as systems biology. In this discipline, various omics-techniques as well as *in silico* modeling play a crucial role [159].

*Genomics* provides the genome sequence of an organism, which sets the frame for the capabilities of the bacterium [153]. The *C. glutamicum* ATCC 13032 genome has been sequenced by various research groups [72, 93]. It is over 3 Mbp in size and contains roughly 3000 genes. Genome scale stoichiometric models allowed the response simulation the *C. glutamicum* metabolism to environmental changes [109] and prediction of target for improved production of organic acids [189].

*Transcriptomics* reveals the set of genes, translated into mRNA, which can give context to the function of genes whose purpose could not be predicted by genomics [192]. Using techniques such as RT-PCR (reverse transcription polymerase chain reaction), a limited set of genes can be quantified with high precision [54]. DNA Microassays and whole transcriptome shotgun sequencing (RNA-seq) enable transcriptional analysis on a genomic scale [245]. The massive amount of data generated in such experimental set ups requires advanced computational power and techniques, for example normalization of the data for noise error reduction or clustering according to the co-expression of the genes [245].

*Proteomics* provides accurate information on the intracellular protein amount. Techniques like 2D-PAGE allow for the separation and quantification of many proteins in a complex matrix and e.g. MALDI-TOF MS (Matrix-assisted laser desorption/ionization– time of flight–mass spectrometry) and LC-ESI-MS/MS (liquid chromatography– electrospray ionization–mass spectrometry) analysis can be used to identify proteins [177].

*Metabolomics* quantifies intracellular metabolite levels and is useful to identify enzymatic bottlenecks via metabolite accumulation [66]. The method is technically challenging due to high turnover rates and huge chemical diversity of the analytes [26, 111].

*Fluxomics* is the most 'phenotypical' of all omics techniques. It most closely resembles the actual effect of certain conditions on the microbial central metabolism and relevant biosynthetic routes by providing information about the carbon fluxes through the various biochemical routes in an organism. This can be done by employing stoichiometric models in combination with data of biomass generated with isotopic tracers like <sup>13</sup>C-labeled sugars [112, 229]. With this technique it was possible to show the metabolic differences during growth and lysine production of *C. glutamicum* on glucose and fructose, like e.g. the drastically reduced carbon flux through the pentose phosphate pathway on the latter substrate [100]. Similarly, iso-

topic tracer studies have been used to unravel preferred use of the Entner-Doudoroff pathway in marine microorganisms [110], characterize the differences in the carbon metabolism between a wild type and a riboflavin producing strain of *Ashbya gossypii* [87] or even to investigate uropathogenic microbes [20]. These examples show the diversity and power of the tool of isotopic labeling.

For a true system-wide understanding, ideally information from different omics techniques is combined. The integration of fluxome, transcriptome and metabolome data of *C. glutamicum* ATCC 13287 e.g. revealed a strong correlation between the glucose uptake flux and the expression of glucose phosphotransferase genes, while the lysine biosynthetic genes were constantly expressed despite substantial changes in carbon flux [117].

#### 2.1.6 Systems metabolic engineering

As the name suggests, systems metabolic engineering combines knowledge and holistic viewpoints from systems biology with tools from metabolic engineering for a 'design-based creation of tailor-made overproducers that are optimized at the global level' [236].

A successful example for systems metabolic engineering is genome breeding [150]. Of the lysine over-producer *C. glutamicum* B-6 derived by classical mutagenesis, 16 relevant genes were compared to the ancestor strain ATCC 13032. Three mutations (V59A *hom*, T3111 *lysC* and P458S *pyc*) were then implemented in the genome of *C. glutamicum* ATCC 13032, which led to an over-producing strain with excellent properties and only minor and specific changes in the genetic background compared to the wild type.

A lysine hyper producing strain with even better production properties was generated by integration of 12 rational modifications into the *C. glutamicum* genome [18]. In addition to the three amino acid exchanged named above, other modifications were aimed at the attenuation of the TCA cycle (ATG $\rightarrow$ GTG start codon in *icd* [11]), the increase of carbon flux through the PP pathway (strong promoters P<sub>tuf</sub> and P<sub>sod</sub> in front of *fbp* and *tkt*, respectively [13]), the anaplerotic node ( $\Delta pck$ , P<sub>sod</sub>*pyc*) and an increased carbon flux towards lysine (2×*ddh* and *lysA*, P<sub>sod</sub>*dapB*, P<sub>sod</sub>*lysC*<sup>G1A</sup>). All these modification made use of the technique for stable integration of changes into the genome mentioned in chapter 2.1.4 [84]. This truly system wide approach led to a strain capable of producing up to 120 g  $L^{-1}$  lysine with a yield of 55%.

In an excellent example for the integration of multiple omics techniques for the generation of advanced production-strains, systems metabolic engineering was used to generate and improve a basic C. glutamicum strain for cadaverine production. Conversion of lysine to cadaverine was enabled by integration of the E. coli ldcC gene coding for the lysine decarboxylase into the C. glutamicum genome [103]. Hereby, the gene was codon-optimized for the expression in C. glutamicum and put under the control of the strong  $P_{tuf}$ -promoter and many modifications that had proven beneficial for lysine production (see above) were also implemented in this *C. glutamicum* DAP-3c strain. By-product formation by acetylation of the product was then eliminated by the use of deletion mutants which made more efficient production possible [104]. C. glutamicum DAP-3c was further improved by genomewide transcriptome analysis which unraveled a gene responsible for the export of cadaverine which also could subsequently be augmented by putting it under the control of the P<sub>tuf</sub>-promoter [105]. Finally, fluxome analysis allowed a detailed evaluation of the differences in the carbon flux between the strains [106]. Taken together, the consequent employment of systems metabolic engineering enabled the creation of a superior production strain able to secrete 90 g L<sup>-1</sup> cadaverine [106]

# 2.2 Alternative renewable feedstocks for bio-based production

*C. glutamicum* is able to utilize a wide variety of molecules as a substrate [125]. In addition to glucose, fructose, sucrose, ribose and mannose, also organic acids like acetate, succinate and lactate are used as a substrate [35]. The relatively low level of catabolite repression [242] and the ability to co-utilize multiple carbon sources are highly desirable for industrial fermentation [45].

Traditionally, large-scale fermentation with *Corynebacteria* use substrates like starch hydrolysate (North America) or sugar cane or beet molasses (Asia, Europe) [102, 238]. Hence it comes as no surprise that strain engineering in the past intensely focused on glucose [78, 113], fructose [100] and sucrose [235], the major constituents of these raw materials. Direct utilization of starch by *C. glutamicum* was achieved via heterologous expression and secretion of an amylase from *Streptomyces* 

griseus [185] and via surface display of a *Streptococcus bovis*  $\alpha$ -amylase [205, 206]. As high-starch materials like corn, potatoes or cassava are available worldwide, further efforts are important to enable more integrated biorefinery processes for *C. glutamicum*. Due to its high availability as a major feedstock in agriculture, silage is another substrate of interest [146]. Two major carbon sources of silage juice, lactate and glucose, can be readily utilized by *C. glutamicum* [145]. Lysine production on silage juice was successful with decent yields, but the final lysine titers of about 3 g L<sup>-1</sup> are far too low to be of industrial interest [145].

Coupled to increasing demands for bio-based chemicals, the use of traditional feedstocks, mentioned above, creates conflicts as they are also suitable for human nutrition or at least compete for arable land with food plants [179]. Hence, alternative, widely available, cheap, renewable feedstocks, not competing with human nutrition, are ethically, ecologically and economically desirable [34].

Whey is a waste product of the diary industry and hence abundantly at hand [7]. Yet, even though *C. glutamicum* has been genetically modified to utilize its main constituents lactose and galactose [9, 35], the production of L-lysine from this raw material is not beyond the proof of concept stage. A bit more advanced, the use of the non natural substrate glycerol has been established in *C. glutamicum* [171] and production of amino acids and putrescine was successfully implemented [139]. Glycerol used to be an attractive product itself at some point, but now is a promising substrate as large quantities are produced as waste during biodiesel production [247].

One extremely promising feedstock candidate, fulfilling the desirable criteria mentioned above is lignocellulose, a major component of plants. It can be derived from a multitude of different sources like forestry or agricultural (crop straw), industrial (pulp and paper industry) or municipal waste products [3, 6, 202]. Lignocellulose consists of cellulose (40-50%), lignin (10-25%) and hemicellulose (20-30%) [40]. Cellulose, the most abundant polymer on earth [130], is a polymer built from  $\beta$ -1,4-linked glucose units [5]. The easily digestible monomer is an advantage for the application in fermentation, but cellulose has to be treated intensely to break it down into smaller units [40]. Metabolic engineering of *C. glutamicum* allowed direct utilization of cellulose for glutamate production [35]. Cellulose, however, has an important technical use for production of paper and board. The waste streams from the pulp and paper industry consequently consist of mainly lignin and hemicellulose.

Name	Properties
Xylan	4-O-methyl-D-glucuronosyl & $\alpha$ -arabinofuranosyl residues on
	a β-1,4 xylopyranose backbone
Glucuronoxylan	2,3- or 2,3-linked glucuronate residues on a $\beta$ -1,4 xy-
	lopyranose backbone
Arabinoxylan	2,3- or 2,3-linked arabinose on a $\beta$ -1,4 xylopyranose back-
	bone
Glucomannan	$\beta$ -1,6 linked residues on a $\beta$ -1,4 mannose and glucose back-
	bone
Xyloglucan	$\beta$ -1,6 xylose residues on a $\beta$ -1,4 glucose backbone

**Table 1:** Different types of hemicelluloses. Like many plant derived polysaccharides, hemicelluloses exhibit a high molecular diversity [201].

Lignin is a diverse group of aromatic polymers of phenylpropane units [193] and a class of molecules not well suitable for microbial conversion [35]. For the pulp and paper industry as well as for the biofuel production, lignocellulosic raw material with low lignin content are desirable, but for burning, lignin-rich woods are preferable due to their high energy density [44].

Hemicellulose is a group of heteropolysaccharides (see Tab. 1) comprising different sugars like glucose, xylose, mannose, galactose, rhamnose, and arabinose with xylose usually being the most abundant one [36]. The sugar monomers are recovered from hemicellulose by acidic hydrolysis with relatively little effort and even at higher yield than glucose is obtained from cellulose [86]. As a beneficial trait, C. glutamicum tolerates growth inhibitors like furfurals, hydroxyl methyl furfurals and organic acids [173] typically present in hydrolysates of lignocellulose and hemicellulose [59]. Naturally, C. glutamicum does not use xylose as a substrate, but harbors xylose importers [174] and also xylB, encoding xylulokinase. This enzyme phosphorylates D-xylulose to D-xylulose 5-phosphate, the last step of xylose catabolism [24]. The first step, the conversion of D-xylose to D-xylulose is catalyzed by the enzyme xylose isomerase, coded for by xylA. Indeed, heterologous expression of xylA from E. coli, alone or in combination with xylB, enables C. glutamicum to use xylose as carbon and energy source [97]. This has been exploited to engineer C. glutamicum for the production of amino acids, putrescine [138], cadaverine [36] and succinic acid [220] from xylose-based raw materials.

### 2.3 Bioprocess technology of industrial lysine production

Beyond the biosynthetic properties of microorganisms, used for a desired bioconversion, their robustness to cultivation in large volumes is crucial for industrial efficiency. The cultivation vessel provides a defined environment and serves as a barrier to the outside. Modern bioreactors at laboratory scale are equipped with advanced measurement and control technology, allowing fine-tuned settings of environmental parameters [52]. Important parameters controlled by the operator include the temperature, the pH value, the supply with oxygen, substrate or anti-foaming agents and the power input required for mass transfer (homogenization, dispersion) and heat transfer. Within a certain operational window, the variation of these parameters controls the microbial behavior and hence the resulting bioprocess. Bioreactors are operated in different modes, i.e. batch, fed-batch and continuous mode, depending on the purpose of the cultivation.

In a batch cultivation, all nutrients (with the exception of oxygen and pH correction and antifoam agents) are present in the reactor at the time of inoculation. A batch fermentation requires less sophisticated equipment then the other operational modes, but allows no stringent control over certain cultivation parameters.

In a fed-batch cultivation, nutrients are supplied to the vessel. This feeding phase follows an initial batch phase. There are different possible feed regimes, e.g. permanent, exponentially increasing feed and feed pulses. The biggest advantage of this operational mode is the higher total amount of substrate, that can be transferred into the vessel in a controlled manner without causing (e.g. oxygen-) limitation, substrate inhibition or by-product formation. This also leads to high final product concentrations in the vessel, which is advantageous for the downstream processing. During production phase, auxotrophic strains require low levels of certain supplement at all times, which can also be achieved by the fed-batch operation. Early homoserine-auxotrophic *C. glutamicum* strains could be utilized for lysine production, if homoserine or threonine was present in limiting concentration [143, 144]. Compared to the batch mode, fed-batch processes need to be constructed with a larger amount of and also more sophisticated peripheral equipment. In addition to the storage tank and the pump(s) necessary for the feeding, sensors and software

apt for the task also have to be in place. A special case of a fed-batch fermentation is the repeated fed-batch (also called semi-continuous mode), in which a fraction of the culture broth of the previous fermentation remains in the vessel, and is used as inoculum for the next fed-batch.

The continuous mode is characterized by a steady (and in most cases constant and balanced) flow in and out of the bioreactor. This leads to steady state conditions, which can be easily controlled via the feed rate by the operator. The resulting high productivity is gained at the cost of product concentration, which is not as high as in fed-batch cultivations [69]. Again, advanced measurement and control technology is required to ensure proper operation, but the continuous mode allows for smaller reactor sizes, which reduces investment costs.

Industrial production of L-lysine is usually conducted in up to 700 m<sup>3</sup> large stirredtank bioreactors, operated in fed-batch, repeated fed-batch or continuous mode [98]. The latter two have the advantage of shorter down-times and higher productivity, but exhibit an increased risk of contamination or, in some cases, genetic instability [72]. Air lift reactors exhibit a more energy-efficient oxygen transfer, but their overall oxygen transfer rate is too low for the high cell density cultivation of modern lysine producing *C. glutamicum* strains [98]. The immense production volumes make multiple, successively larger seed reactors necessary. This is also due to the tendency of *C. glutamicum* to show a prolonged lag phase if the initial cell concentration at inoculation is too low. In this so called seed train the maintenance of monoseptic conditions is absolutely essential, as a contaminant like e.g. *Bacillus sp.* would outgrow *C. glutamicum* [98].

Raw materials, currently used in large scale lysine production, include sugar cane sucrose, starch hydrolysates or molasses as carbon sources, depending on the location of the production site, with corn starch hydrolysis e.g. dominating in North America [68]. Ammonia in dissolved, gaseous or saline form is used to provide nitrogen. Alternatively, complex sources like corn steep liquor are also used. The use of well-defined raw materials is beneficial for process reproducibility and has increased in the last two decades [98]. On the other hand, raw materials are a major cost factor in the production of a bulk chemical with a low price like lysine (see Fig. 8) [41, 163].

Among all parameters, temperature is an important, yet relatively easily controllable parameter of a bioprocess, as no advanced measurement and control technology is required. Its measurement is usually conducted with simple thermocouple or resistance thermometers and regulated via cooling jackets and coils. In contrast to chemical reactions, bioconversions do not produce heat very rapidly. The bigger problem lies in the fact that biological reactions take place at comparatively low temperature, only 5 to 20 °C above the temperature of the cooling water, which leads to poor heat removal capacity [213]. The impact of temperature on the process, however, is substantial. Microorganisms exhibit optimal growth at specific temperatures, depending on the species. For different strains of C. glutamicum, optimal temperatures around 25 to 37 °C are reported [2] and cultivation in the past decades is typically performed at 30 °C [98]. Below and above this optimum, specific growth rate is reduced. Too high temperatures eventually lead to cell death. The optimal temperature for growth also might not be the best temperature for the bioprocess as in some cases slower growth is desirable, e.g. to avoid oxygen limitation. In order to maintain a constant temperature despite the heat generated by the metabolic activity of the microorganisms and the mechanical power input of the stirrer, bioreactors are equipped with cooling jackets and coils [58]. With higher volumes this becomes a more and more demanding task as the specific volume (volume per surface area) gets smaller. This means that the metabolic heat, generated by the microorganisms increases cubically with the volume, but the cooling surface of the vessel wall increases only quadratically, often making expensive cooling coils inside of the reactor necessary, when scaling up a bioprocess [91].

Amino acid bulk production is often located in close proximity to substrate manufacturers [47]. In case of sugar cane this means tropical regions [102], where the temperature difference between process and cooling water might be even lower, which further increases the problem to cool large scale production. From a bioprocess point of view this illustrates the advantages of running fermentations at higher temperature. Clearly, the limitation for this are microorganisms, which are sensitive towards the temperature of their environment. Yet, the clear economic benefits make it worthwhile to investigate the possible improvement of lysine production bioprocesses by increased temperature.
# MATERIALS AND METHODS

#### 3.1 Bacterial strains and molecular work

**Pre-existing strains.** All strains of *Corynebacterium glutamicum* used in this work were based on the wild type strain *C. glutamicum* ATCC 13032 and were obtained from the American Type Culture Collection (Manassas, VA, USA) and from previous work [2, 18] (Tab. 2, Fig. 10). Strains were kept as cryostocks made by mixing equal volumes of 60 % glycerol and culture broth with exponentially growing cells and freezing them in liquid nitrogen before storage at -80 °C.

*C. glutamicum* Lys1 has a T311I amino acid exchange in the *lysC* gene, which results in an aspartokinase, resistant to feedback inhibition by lysine and threonine. The rationally designed strain *C. glutamicum* Lys12 comprises 12 genomic modifications [18]. *E. coli* strains DH5 $\alpha$  and NM522 (Invitrogen GmbH, Darmstadt, Germany) were utilized for amplification and methylation of plasmids, respectively. The plasmid pTc15AcglM [164], which contained an origin of replication for *E. coli* and tetracycline resistance as selection marker, served as expression vector for DNA methyl transferase, which added the *C. glutamicum* specific methylation pattern to plasmid DNA, upon co-expression in *E. coli* NM522.

**Generation of** *C. glutamicum* Lys12K. The episomal vector pClik 5a MCS [115], which contains a multiple cloning site, an origin of replication for *E. coli* and for *C. glutamicum* as well as a kanamycin resistance selection marker was transformed into *E. coli* DH5 $\alpha$  cells via a heat shock. For that, 50 µL of *E. coli* DH5 $\alpha$  cells, growing exponentially on LB medium, and 3 µL (1-50 ng) plasmid DNA were incubated on ice for 30 minutes, shifted to 45 °C for 45 seconds and immediately planted back

Name	Relevant feature	Ref./Source
E. coli DH5α	Heat shock competent cells for plasmid transforma- tion	Invitrogen
E. coli NM522 <sup>a</sup>	Heat shock competent cells for plasmid transforma- tion and <i>C. glutamicum</i> methylation pattern via plas- mid pTc15AcglM	Invitrogen
C. glutamicum ATCC 13032	Wild type strain	[2, 108]
C. glutamicum Lys1	<i>C. glutamicum</i> ATCC 13032 with a feedback resistant aspartokinase	[13]
C. glutamicum Ect1	<i>C. glutamicum</i> Lys1 heterologously expressing the synthetic <i>ectABCD</i> gene cluster	[14, 27]
C. glutamicum Ect2	C. glutamicum Ect1 with deletion of the lysE gene	[14, 65]
C. glutamicum Lys12	Lysine hyper producer	[18]
C. glutamicum Lys12K <sup>b</sup>	<i>C. glutamicum</i> Lys12 containing the episomal pClik plasmid	this work
C. glutamicum Xyl1 <sup>b</sup>	<i>C. glutamicum</i> Lys12 containing the episomal pClik <i>xylAB</i> plasmid	this work
pTc15AcglM	Contains an <i>E. coli</i> ORI and genes for tetracycline resistance as well as for a <i>C. glutamicum</i> specific DNA methyl transferase	[164]
pClik 5a MCS	Episomal replicating plasmid containing an MCS, an <i>E. coli</i> and <i>C. glutamicum</i> ORI and a kanamycin resistance gene	[115]
pClik 5a MCS P <sub>gro</sub> xylAB	pClik 5a MCS with <i>E. coli xylA</i> and <i>xylB</i> genes	[36]
pClik int <i>sacB</i>	Integrative plasmid containing an MCS, an <i>E. coli</i> ORI and a levansucrase as well as a kanamycin resistance gene	[84]
pClik int <i>sacB</i> P <sub>tuf</sub> ectABCD	pClik int <i>sacB</i> for the replacement of <i>C. glutamicum ddh</i> with an codon optimized ectoine synthesis cluster under the control of the <i>C. glutamicum tuf</i> promoter	[14]
pClik int sacB ∆lysE	pClik int <i>sacB</i> for the disruption of the <i>C. glutamicum lysE</i> gene	[105]

**Table 2:** Bacterial strains and plasmids used in this work. The genealogy of the strains is shown in Fig. 10.

<sup>*a*</sup>Tetracycline resistant

<sup>b</sup>Kanamycine resistant

on ice for 2 minutes. Subsequently, 900  $\mu$ L SOC medium was added, cells were incubated at 37 °C for one hour and were then plated on LB-agar plates with an appropriate antibiotic. To harvest amplified plasmids, the Gene Jet Plasmid Miniprep K 0503 or the Gene Jet Plasmid Midiprep K0481 (both Thermo Scientific, Schwerte, Germany) were used. To achieve the methylation of the DNA apt for utilization by *C. glutamicum*, the plasmid amplification step was repeated, using *E. coli* NM522.



**Fig. 10:** Genealogical tree of the *C. glutamicum* strains used in this work. *C. glutamicum* ATCC 13032 was obtained from the American Type Culture Collection (Manassas, VA, USA) Lys1 [2] and Lys12 [18] are described in the literature. The generation of *C. glutamicum* Ect1 and Ect2 was described previously [14, 27, 65]. A red background color represents ectoine production, a yellow one xylose utilization and strains resistant to kanamycin are labeled Kan<sup>R</sup>.

Plasmids, isolated from *E. coli* NM522, were used to transform *C. glutamicum* Lys12 cells using electroporation. *C. glutamicum* cells were grown in BHI medium up to an optical density of 1.5, harvested via centrifugation  $(13000 \times g, 5 \text{ min})$ , washed 4 times with and resuspended in ice cold 10% glycerol to a concentration of 0.125-0.25 g<sub>cww</sub> mL<sup>-1</sup>. In chilled electroporation cuvettes (2 mm gap, Bio-Rad, Hercules, CA, USA), 100 µL of the resuspended cells and 200-500 ng plasmid DNA were mixed. Electroporation was performed at 2.5 kV, 25 F and 400  $\Omega$  with the Gene-Pulser electroporator (Bio-Rad, Hercules, CA, USA). After the electric pulse, 1 mL of BHIS medium was added and the cells were then incubated for 1.5 hours at 30°C, then were spread on BHIS agar plates, containing kanamycin as a selection marker, and finally were incubated at 30°C for two days. Cells selected for on these plates were the desired clones and were named *C. glutamicum* Lys12K.

**Generation of** *C. glutamicum* Xyl1. The episomal vector pClik 5a MCS P<sub>gro</sub>xylAB was used for expression of the xylose operon [36]. The plasmid contained genes for xylose isomerase (*xylA*, ECK3554) and xylulokinase (*xylB*, ECK3553) from *E. coli* K-12 MG1655 under control of the strong constitutive *C. glutamicum* promoter P<sub>gro</sub>, a multiple cloning site, an origin of replication for *E. coli* and for *C. glutamicum* as

Table 3:	Primers	used	in	this	work.

Sequence (5'-3')	Function
CACCCTCAACAGTTGAGTGCTGGCAC ATTGTCTGTTGTGCCCAGTCATAG	Verifying integration of the <i>xylAB</i> gene in pClik 5a MCS
AGAGTACCTGGGACGCAGCGTCG CGTCGCGTGCGATCAGATCGGT	Verifying integration of the $P_{tuf}$ ectABCD gene cluster & disruption of the ddh gene
CTCGTTGAACACGGGAGGAA CACGGACAAATATCTCGATG	Verifying disruption of the <i>lysE</i> gene

well as a kanamycin resistance selection marker. The process of transformation of *C. glutamicum* Lys12 was analogous to the one described for *C. glutamicum* Lys12K above. For the verification of of the successful transformation, clones from the LB<sup>Kan</sup> agar plates were transferred to agar plates with a minimal medium containing xylose as the sole carbon. One of the colonies able to grow in presence of kanamycin and on xylose was chosen and named *C. glutamicum* Xyl1.

For strain validation, polymerase chain reaction was performed in addition to physiological confirmation, like antibiotic resistance and growth on xylose. PCR was done using a mastercycler EP gradient PCR machine (Eppendorf, Hamburg, Germany) and either the PWO master, the PCR master (both Roche, Basel, Switzerland) or the KAPAHiFi Kit (Peqlab, Erlangen, Germany). Primers were obtained from Invitrogen (Darmstadt, Germany) and are listed in Tab. 3. The annealing temperature  $T_a$  of primers was calculated according to equation 1, were C, G, A and T stand for the nucleobases cytosine, guanine, adenine and thymine, respectively.

$$T_{a} = \begin{cases} \frac{64 + (n_{G} + n_{C} - 16.4)}{n_{A} + n_{T} + n_{C} + n_{G}} - 5, & \text{if } \sum n_{i} > 13\\ (n_{A} + n_{T}) \cdot 2 + (n_{G} + n_{C}) \cdot 4 - 5, & \text{if } \sum n_{i} \le 13 \end{cases}$$
(1)

**Generation of** *C. glutamicum* **Ect1 and Ect2.** Heterologous production of ectoine in *C. glutamicum* was based on the *ectABCD* biosynthetic gene cluster from *Pseudo-monas stutzeri* A1501 [198]. The codon usage was adjusted to that preferred by *C. glutamicum* using the proprietary GeneOptimizer<sup>®</sup> software (Geneart, Regensburg, Germany). The DNA sequence of the codon-optimized *ectABCD* variant gene cluster, along with the original sequence of *P. stutzeri* A1501, can be seen in the



**Fig. 11:** Construct for the integration of the ectoine synthesis cluster into *C. glutamicum*. It contained the codon-optimized *ectABCD* gene cluster based on that present in *P. stutzeri* A1501 under the control of the promoter for the *tuf* gene from *C. glutamicum*. For genome-based integration via a double-recombination, the construct was additionally equipped with flanking regions of about 560 bp DNA sequences, derived from the upstream and downstream regions of the *ddh* gene. Recognition sites for the restriction enzyme SpeI were added to facilitate cloning of this DNA fragment into the vector pClik int sacB [13]. The *ddh* gene, encoding diaminopimelate dehydrogenase, was chosen as integration site to minimize competing carbon flux towards lysine.

appendix (Fig. 36). The complete 4577 bp-sized construct (Fig. 11) was provided by Geneart (Regensburg, Germany) and included a nucleotide sequence consisting of a 200-bp DNA segment that carried the strong and constitutively active promoter for the expression of the structural *tuf* gene (NCgl0480) of the elongation factor Tu of *C. glutamicum* and the aforementioned optimized *ectABCD* gene cluster. Additionally, the construct was flanked by 560 bp-sized homologous recombination sites for genome-based integration of the construct into the structural *ddh* gene (NCgl2528), encoding diaminopimelate dehydrogenase. Artificial SpeI digestion sites were added at the 5' and 3' end.

Ligation of this construct into the pClik int *sacB* plasmid was conducted with the Rapid Dephos & Ligation Kit (Roche, Basel, Switzerland). The resulting plasmid pClik int *sacB*  $P_{tuf}ectABCD$  was used for the integration of the synthetic gene cluster into *C. glutamicum* Lys1. It contained the ectoine-production construct (Fig. 11), an origin of replication for *E. coli* as well as a kanamycin resistance and a *sacB* gene as selection markers.

The further treatment of the plasmid (heat shock into *E. coli*, electroporation into *C. glutamicum* Lys1) was done as described above, but more plasmid DNA (1-5  $\mu$ g) was used in the electroporation step. Single *C. glutamicum* colonies, able to grow in the presence of kanamycin, were further cultivated without selective

pressure to allow for the second recombination event and were subsequently plated on  $CM_{Sac}$  agar plates [84]. A colony sensitive to kanamycin and insensitive to sucrose, *C. glutamicum* Lys1  $\Delta ddh P_{tuf}ectABCD$ , was named *C. glutamicum* Ect1.

The plasmid pClik int *sacB*  $\Delta lysE$  [105] contained homologous recombination sites for genome-based disruption of the *lysE* gene (NCgl1214), an origin of replication for *E. coli* as well as a kanamycin resistance and a *sacB* gene as selection markers. Transformation of this plasmid into *C. glutamicum* Ect1 was done as described above. The resulting strain, *C. glutamicum* Ect1  $\Delta lysE$ , was named *C. glutamicum* Ect2. The construction of *C. glutamicum* Ect1 was performed by Nicole Borchert [27] and strain *C. glutamicum* Ect2 was constructed by Björn Johannes Harder [65].

For strain validation, polymerase chain reactions were performed in addition to physiological confirmation, like ectoine production. PCR was done as described above with primers listed in Tab. 3.

### 3.2 Chemicals

Complex media ingredients like BHI (brain heart infusion) powder, yeast extract, peptone and tryptone as well as agar were obtained from Becton & Dickinson (Franklin Lakes, NJ, USA). Isotopically labeled substrates, i.e. 1-<sup>13</sup>C and U-<sup>13</sup>C glucose were obtained from Euriso-Top (Saarbrücken, Germany). If not specified otherwise, all other chemicals were obtained from Sigma-Aldrich (St. Lois, MO, USA) or Merck (Darmstadt, Germany) at analytical grade quality. Water was purified by a Milli-Q Integral water purification system (Merck, Darmstadt, Germany).

#### 3.3 Media for shake flask cultivations

Complex cultivation media used in this work are listed in Tab. 4. Chemically defined media are given in Tab. 5. If a solid version of a medium was required, 18 g L<sup>-1</sup> agar were added before sterilization. The sterilization was performed by autoclaving at 121 °C and 2 bar for 20 min. If required, kanamycin and tetracycline were used as antibiotics at concentrations of 50 and 12.5 mg L<sup>-1</sup>, respectively.

Chemically defined media cdm1 and cdm2 were prepared by mixing sterile stock solutions (Tab. 5) and adding autoclaved deionized water. Solutions (A)-(G) and (K)

Medium	Amount (g)	Component	Water (mL)	Comment
BHI	37	BHI	1000	
BHIS	37	BHI	750	250 mL 2 M sorbitol solution added after autoclaving
СМ	2.5 5 10 5	NaCl Yeast extract Peptone Meat extract	930	50 mL 40 g L <sup>-1</sup> urea solution and 20 mL 500 g L <sup>-1</sup> glucose solution added after autoclaving
CM <sup>Sac</sup>	2.5 5 10 5 100	NaCl Yeast extract Peptone Meat extract Sucrose	930	50 mL 40 g L <sup>-1</sup> urea solution and 20 mL 500 g L <sup>-1</sup> glucose solution added after autoclaving
LB	5 5 10	NaCl Yeast extract Tryptone	1000	
SOC	0.5 5 0.186 20	NaCl Yeast extract KCl Tryptone	975	5 mL 2 M MgCl <sub>2</sub> solution and 20 mL 1 M glucose solution added after autoclaving

 Table 4: Composition of complex media.

were sterilized by autoclaving, (H)-(J) by filtration (0.2  $\mu$ m Minisart filter; Sartorius, Göttingen, Germany). The medium cdm1 was used for shake flask cultivations with xylose as the carbon source and the glucose-based reference cultivations thereof. The cdm2 medium was used for all other shake flask cultivations on chemically defined medium.

### 3.4 Media for bioreactor cultivations

**Lysine production from xylose.** The batch medium for lysine production from xylose contained (per liter medium): 72.4 g root beet molasses (Paik Kwang Industrial, Gunsan, South Korea), 35 mL corn steep liquor (BASF, Ludwigshafen, Germany), 85 g xylose, 250  $\mu$ L H<sub>3</sub>PO<sub>4</sub> (85%), 40 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 250 mg KH<sub>2</sub>PO<sub>4</sub>, 100 mg MgSO<sub>4</sub>•

Stock	Concentra-	Component	Comment	<u>mL in 1 L</u>		
solution	tion (g L <sup>-1</sup> )			cdm1	cdm2	
А	40 2.2 8	NaCl CaCl <sub>2</sub> MgSO <sub>4</sub>		25	25	
В	150	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	рН 7.0	100	100	
С	100	Glucose			100	
D	100	Xylose or glucose		100		
E	100	<i>N</i> -(2-Acetamido)-2- aminoethanesulfonic acid (ACES)	pH 7.4	100		
F		K <sub>X</sub> H <sub>X</sub> PO <sub>4</sub> <sup>a</sup>	рН 7.8		100	
G	2	FeSO <sub>4</sub> •7H <sub>2</sub> O	рН 1.0	10	10	
Н	30	3,4-dihydroxybenzo- ic acid	in 0.3 M NaOH solution	1	1	
Ι	0.2 0.2 0.05 0.02 0.02 0.01	$\begin{array}{l} FeCl_{3} \cdot 6H_{2}O \\ MnSO_{4} \cdot H_{2}O \\ ZnSO_{4} \cdot H_{2}O \\ CuCl_{2} \cdot 2H_{2}O \\ Na_{2} B_{4}O_{7} \cdot 10H_{2}O \\ (NH_{4})_{6}Mo_{7}O_{24} \cdot 4H_{2}O \end{array}$	рН 1.0	10	10	
J	0.025 0.05 0.05	Biotin Thiamin Ca-pantothenate		20	20	
K	23	K <sub>2</sub> HPO <sub>4</sub>		10		

 Table 5: Composition of chemically defined media.

 $^{a}2$  M KH<sub>2</sub>PO<sub>4</sub> and 2 M K<sub>2</sub>HPO<sub>4</sub> mixed to achieve a pH of 7.8

 $7H_2O$ , 11 mg FeSO<sub>4</sub>• $7H_2O$ , 10 mg citric acid, 9 mg biotin, 15 mg thiamin-HCl, 60 mg Ca-pantothenate and 18 mg nicotinamide. The feed medium, added during the fed-batch phase, contained 162.5 g molasses, 585 g xylose and 40 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> per liter.

**Ectoine production from glucose.** One liter of batch medium contained the following stock solutions: (A) 25 mL salt solution (40 g  $L^{-1}$  NaCl, 2.2 g  $L^{-1}$  CaCl<sub>2</sub>, 8 g  $L^{-1}$ 

MgSO<sub>4</sub>), (B) 50 mL ammonium solution (400 g L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, pH 7.0), (C) 100 mL sugar solution (500 g L<sup>-1</sup> glucose), (D) 100 mL buffer solution (2 M KH<sub>2</sub>PO<sub>4</sub> and 2 M K<sub>2</sub>HPO<sub>4</sub> mixed to achieve a pH of 7.8), (E) 10 mL yeast extract (50 g L<sup>-1</sup>), (F) 10 mL iron solution (2 g L<sup>-1</sup> FeSO<sub>4</sub>·7H<sub>2</sub>O pH 1.0), (G) 1 mL complexing agent (30 g L<sup>-1</sup> 3,4-dihydroxybenzoic acid in 0.3 M NaOH solution), (H) 10 mL trace element solution (200 mg FeCl<sub>3</sub>·6H<sub>2</sub>O, 200 mg MnSO<sub>4</sub>· H<sub>2</sub>O, 50 mg ZnSO<sub>4</sub>· H<sub>2</sub>O, 20 mg CuCl<sub>2</sub>·2H<sub>2</sub>O, 20 mg Na<sub>2</sub> B<sub>4</sub>O<sub>7</sub>·10H<sub>2</sub>O and 10 mg (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O per liter, pH 1.0) and (I) 20 mL vitamin solution (25 mg biotin, 50 mg thiamin and 50 mg Ca-pantothenate per liter). The feed medium comprised the same stock solutions in the following composition per liter: (A) 5 mL, (B) 125 mL, (C) 850 mL, (E) 10 mL, (I) 50 mL. Solutions (D), (F), (G) and (H) were omitted. Solutions (A)-(E) were sterilized by autoclaving, solutions (F)-(I) were sterilized by filtration (0.2 µm Minisart filter; Sartorius, Göttingen, Germany).

**Lysine production from glucose at elevated temperature.** One liter of medium contained the following stock solutions: (A) 25 mL salt solution (40 g L<sup>-1</sup> NaCl, 2.2 g L<sup>-1</sup> CaCl<sub>2</sub>, 8 g L<sup>-1</sup> MgSO<sub>4</sub>), (B) 176.5 mL ammonium/sugar solution (248 g L<sup>-1</sup> (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 766 g L<sup>-1</sup> glucose, pH 7.0), (D) 20 mL buffer solution (2 M KH<sub>2</sub>PO<sub>4</sub> and 2 M K<sub>2</sub>HPO<sub>4</sub> mixed to achieve a pH of 7.8), (E) 10 mL iron solution (2 g L<sup>-1</sup> FeSO<sub>4</sub>·7H<sub>2</sub>O pH 1.0), (F) 1 mL complexing agent (30 g L<sup>-1</sup> 3,4-dihydroxybenzoic acid in 0.3 M NaOH solution), (G) 10 mL trace element solution (200 mg FeCl<sub>3</sub>·6 H<sub>2</sub>O, 200 mg MnSO<sub>4</sub>·H<sub>2</sub>O, 50 mg ZnSO<sub>4</sub>·H<sub>2</sub>O, 20 mg CuCl<sub>2</sub>·2H<sub>2</sub>O, 20 mg Na<sub>2</sub> B<sub>4</sub>O<sub>7</sub>·10H<sub>2</sub>O and 10 mg (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O per liter, pH 1.0) and (H) 20 mL vitamin solution (25 mg biotin, 50 mg thiamin and 50 mg Ca-pantothenate per liter).

The feed medium comprised the same stock solutions in the following composition per liter: (A) 5 mL, (B) 969.5 mL, (E) 5 mL, (F) 0.5 mL, (G) 5 mL and (H) 10 mL. Solutions (A)-(D) were sterilized by autoclaving and (E)-(H) by filtration (0.2  $\mu$ m Minisart filter; Sartorius, Göttingen, Germany). The sterility of solution (B) was ensured by first autoclaving the components in their dry form separately and then dissolve them in sterile water.

# 3.5 Cultivation conditions

#### 3.5.1 Shake flask cultivation

Dispersing a small aliquot of cryostocks on BHI agar and incubation for two days at 30 °C gave rise to enough cell material to be used in suspension cultures. Single colonies from those agar plates were used as inoculum for a first pre-culture in complex BHI medium. All shake flask cultures were incubated in baffled shake flasks filled with 10 % medium at 230 rpm on an orbital shaker with a shaking diameter of 5 cm (Multitron, Infors AG, Bottmingen, Switzerland). After 12 hours of incubation, cells were harvested (5 min,  $9800 \times g$ , 4 °C), including a washing step with sterile NaCl (9 g L<sup>-1</sup>), and were then used as inoculum for a second pre-culture in minimal medium. Cells were grown to mid exponential phase, were harvested as described above, and were then used to inoculate main cultures in minimal medium. On-line measurement of dissolved oxygen and pH in shake flasks was conducted with immobilized sensor spots and a flask reader system (SFR, Presens, Regensburg, Germany), installed in the rotary shaker [180].

#### 3.5.2 Miniaturized cultivations

Small-scale cultivations of *C. glutamicum* strains were carried out in 1 mL cdm2 in 48well flower plates (m2p-labs, Baesweiler, Germany) at 700 rpm using the Biolector system (DASGIP, Jülich, Germany). The temperature was varied between 27 °C and 42 °C. To avoid evaporation of the medium, the plates were sealed with a gaspermeable membrane (Aera Seal, Sigma Aldrich, Steinheim, Germany). Inoculum preparation involved shake flask cultivation as described above, and these cells were then subsequently used to inoculate the multi-well plates to an initial optical density of 0.3. Cell growth was monitored online via measurement of the optical density at 620 nm. After 10 h (27 °C, 30 °C, 35 °C) and 20 h (42 °C), extracellular and intracellular accumulation of ectoines was determined. Process monitoring, data collection, and data processing were carried out with the software suit BioLection (m2p-labs, Baesweiler, Germany). Miniaturized cultivations were carried out by Nicole Borchert [27].

#### 3.5.3 Fed-batch bioprocesses

Bioreactors used were stirred tank reactors with 1 L working volume (SR07000-DLS/CWD4 Bioblock, DASGIP AG, Jülich, Germany), equipped with two 6-blade rushton impellers. The set-up also included temperature, pH (Mettler Toledo, Gießen, Germany) and dissolved oxygen (DO) probes (VisiFerm, Hamilton, Bonaduz, GR, Switzerland), aeration (MX4/4 aeration system, DASGIP AG, Jülich, Germany) and pumps for the addition of base and feeding solution (MP8 pump system, DASGIP AG, Jülich, Germany). In all cultivations, the initial volume was 300 mL medium. The pH was controlled at  $6.9 \pm 0.2$  by addition of 25 % NH<sub>3</sub> (aq) using the software DASGIP control (DASGIP AG, Jülich, Germany).

**Lysine production from xylose.** For preparation of inoculum, cells from glycerol stocks of *C. glutamicum* Xyl1 were streaked out on BHI agar and incubated for two days at 30 °C. Cells were then transferred to BHI medium with 10 g L<sup>-1</sup> glucose for rapid biomass generations and incubated for 12 hours at 30 °C at 230 rpm on a rotary shaker with a shaking diameter of 2.5 cm (Certomat IS, Sartorius, Göttingen, Germany). Afterwards, cells were harvested by centrifugation (10 min, 9800×g, 4 °C), resuspended in 20 mL of sterile 9 g L<sup>-1</sup> NaCl solution and then were transferred into the bioreactor with a syringe. During the process, temperature was kept at 30  $\pm$  0.3 °C. The stirrer speed (400-1200 rpm) and aeration rate (6-12 L h<sup>-1</sup>) were adjusted manually to keep the dissolved oxygen concentration above 20% relative saturation.

**Ectoine production from glucose.** Ectoine production of *C. glutamicum* Ect2 was also investigated by fed-batch cultivation. Cells pre-grown on agar plates as described above were transferred in BHI medium with 20 g L<sup>-1</sup> glucose and were incubated for 10 hours at 35 °C at 230 rpm on a rotary shaker with a shaking diameter of 2.5 cm (Certomat IS, Sartorius, Göttingen, Germany). Cells were then harvest using a Heraeus Multifuge 4KR centrifuge (Thermo Fisher Scientific, Rockford, MA, USA,  $5350 \times g$ , 10 min), re-suspended in 20 mL fermentation medium and then transferred into the bioreactor with a syringe as an inoculum. Dissolved oxygen was maintained above 30 % of relative saturation by adjustment of stirrer speed and aeration rate. These values were initially set to 18 L h<sup>-1</sup> and 400 rpm, respectively. The temperature

was kept constant at  $35 \pm 0.3$  °C. The feeding was automatically controlled via the dissolved oxygen (DO) level by triggering feed pulses upon pronounced increases in the DO, which indicated complete sugar utilization by the cells. The fermentation for ectoine production was done in collaboration with Michael Kohlstedt [111] and Björn Johannes Harder [65].

**Lysine production from glucose at elevated temperature.** For production at elevated temperature, cells were grown in a series of pre-cultures (see 3.5.1), then were harvested, re-suspended in a small volume of fermentation medium and used as inoculum. Aeration rate (8 to 16 L h<sup>-1</sup> air) and stirrer speed (800 to 1200 rpm) were varied to keep the dissolved oxygen level above 20% saturation. Temperature was kept constant at  $38 \pm 0.3$  °C or  $30 \pm 0.3$  °C as specified below. The feed was added pulse-wise, coupled to the dissolved oxygen signal as described above.

## 3.6 Analytical methods

#### 3.6.1 Sample preparation

For the quantification of substrates and of secreted products in the culture supernatant, the biomass was separated from the culture broth by centrifugation  $(13000 \times g, 5 \text{ min}, 4 \degree \text{C} \text{ and } 10000 \times g, 15 \text{ min}, 4 \degree \text{C}$ , centrifuge 5415R Eppendorf, Hamburg, Germany).

For the quantification of intracellular concentrations of ectoine and hydroxyectoine, cells from 1 mL culture were harvested by centrifugation (13000×g, 5 min, 4 °C). The exact sample volume was determined gravimetrically on an analytical balance (CP255D, Sartorius, Göttingen, Germany). In parallel, the optical density was measured to quantify the harvested biomass amount. The supernatant was decanted, and the cells were subsequently dried for 12 h in a speedvac apparatus (Concentrator 5301, Eppendorf, Hamburg, Germany). Cells were re-suspended in 500 µL Bligh & Dyer solution (MeOH:CHCl<sub>3</sub>:H<sub>2</sub>0, 10:5:2) and disrupted mechanically (FastPrep<sup>®</sup>-24, 3 min, 4 m s<sup>-1</sup>, MP Biomedicals, Santa Ana, CA, USA) using glass beads with a diameter of 0.04 mm. The disruption step was repeated, after adding 130 µL H<sub>2</sub>O and 130 µL chloroform to the slurry. Phases were then separated by centrifugation (10000×g, 15 min centrifuge 5415R, Eppendorf, Hamburg, Germany). The aqueous

phase was transferred into a novel reaction tube and subsequently evaporated to dryness in a speedvac apparatus. The remaining solids were dissolved in 80-100  $\mu$ L H<sub>2</sub>O and centrifuged (15 min, 10000×g, 4 °C); metabolite concentrations were then quantified by HPLC analysis (see section 3.6.3). Sampling for quantification of the intracellular amino acid concentration was performed by fast filtration as described in the literature [26].

#### 3.6.2 Quantification of cell concentration

The concentration of cells was determined by measurement of the optical density of the culture broth at 660 nm in plastic cuvettes (Sarstedt, Nümbrecht, Germany) and water as a zero-value. For *E. coli* cells, a wave length of 600 nm was used. For optical densities, determined with the Libra S11 instrument (Biochrome, Cambridge, UK), the correlation between optical density and cell dry weight was 0.255 g<sub>cdw</sub> L<sup>-1</sup> OD<sup>-1</sup>, as reported in earlier work [11]. For determining the corresponding correlation for a UV-1600PC spectrophotometer (VWR, Radnor, PA, USA), cells were cultivated in 200 mL chemically defined medium in 2 L baffeled shake flasks after two precultures on BHI and defined medium as described before. At different time points, 10-14 mL of broth were withdrawn into pre-weighed reaction tubes, centrifuged (10 min, 9800×g) and washed three times with deionized water. The cell pellets were then dried until weight constancy (about 72 h), cooled to room temperature in an desiccator and then weighed. The resulting correlation factor was 0.39 g<sub>cdw</sub> L<sup>-1</sup> OD<sup>-1</sup> (30 °C) and 0.34 g<sub>cdw</sub> L<sup>-1</sup> OD<sup>-1</sup> (38-40 °C).

#### 3.6.3 Quantification of metabolites

Glucose, trehalose and organic acids were quantified by HPLC (Agilent 1260 Infinity Series, Agilent Technologies, Waldbronn, Germany), using a Microgard pre-column (Cation+ H+ 30x4.6, Bio-Rad, Hercules, CA, USA) and an Aminex HPX-87H main column (Bio-Rad) as solid phase, 5 mM  $H_2SO_4$  (55 °C, 0.7 mL h<sup>-1</sup>) as mobile phase and quantification via external standards and refraction index detection (Agilent Technologies, Waldbronn, Germany). Alternetively, xylose, glucose, fructose and sucrose were quantified in diluted cultivation supernatant with HPLC (LaChromElite, Hitachi, Chiyoda, Japan) on a Meta-Carb 87C carbohydrate column (300×7.8, Varian Inc., Palo Alto, CA, USA) at 85 °C with de-ionized water as mobile phase (0.6 mL min<sup>-1</sup>) and detection via refraction index. At-line analysis of glucose during shake-flask experiments for ectoine production and of xylose during fed-batch cultivations was performed by enzymatic quantification of the sugars in 1:10 diluted supernatant samples using the biochemical analyzer YSI 2700 Select (YSI, Yellow Springs, OH, USA).

Amino acids were quantified by HPLC (Agilent 1200 Series, Agilent Technologies, Waldbronn, Germany) on a reverse phase column (Gemini5u, Phenomenex, Aschaffenburg, Germany) with fluorescence detection, after pre-column derivatization with *o*-phthalaldehyde [116]. A gradient of eluent A (40 mM NaH<sub>2</sub>HPO<sub>4</sub>, pH 7.8) and B (45% acetonitrile, 45% methanol, 10% water) served as mobile phase. Quantification involved  $\alpha$ -aminobutyric acid as internal standard [103]. In samples from fed-batch production on xylose, lysine was quantified by HPLC (LaChromElite, Hitachi, Chiyoda, Japan) on an Ionospher 5C column (100×3, Varian Inc., Palo Alto, CA, USA) at 40°C, a flow rate of 1.5 mL min<sup>-1</sup> (20 mM tartaric acid, 4.9 mM ethylenediamine, 1.8 mM citric acid in 95% H<sub>2</sub>O and 5% methanol) and detection via refraction index. For lysine, mass concentrations given refer to lysine·HCl.

Ammonium was quantified in diluted supernatants by HPLC (ICS-2000, IonPac CS16 column, IonPac CG16 pre-column, Dionex, Sunnycale, CA, USA) at a flow of 0.5 mL min<sup>-1</sup> at 40 °C with methansulfonic acid as mobile phase. Detection was performed by a conductivity detector (L7470, LaChromElite, Hitachi, Chiyoda, Japan).

Quantification of ectoine and hydroxyectoine was performed by HPLC analysis (LaChrome, Merck-Hitachi, Darmstadt, Germany) using a ProntoSil C18 AQ + column ( $125 \times 4$  mm, Knauer, Berlin, Germany) with a Nucleosil C18 AQ + pre-column ( $120 \times 5$  mm, Knauer, Berlin, Germany). As mobile phase, a phosphate buffer was used (0.8 mM K<sub>2</sub>HPO<sub>4</sub>; 6.0 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 7.6) at a flow rate of 1 mL min<sup>-1</sup> and at 40 °C. The injection volume was 2 µL. Detection was carried out with a diode array detector (L7450, LaChrome, Merck-Hitachi, Darmstadt, Germany) at 220 nm.

#### 3.6.4 Enzymatic assays

For preparation of a cell extract, cells, cultivated on minimal salt medium, were harvested by centrifugation (5 min, 9800×g, 4 °C), washed with disruption buffer

[97] and re-suspended in the same buffer to a final concentration of 0.25  $g_{cww}$  mL<sup>-1</sup>. In a reaction cup, filled with 0.5 mL of glass beads (0.15-0.25 mm  $\emptyset$ ), 1 mL of the suspension was then disrupted using a homogenizer (4 m s<sup>-1</sup>, 2×30 s with a 5 min break in between, FastPrep-24, MP Biomedicals, Santa Ana, CA, USA). Cell debris was removed by centrifugation (5 min, 14000×g, 4 °C). The protein content of the obtained extract was measured with a bicinchoninic acid (BCA) protein assay (Thermo Fisher Scientific, Rockford, MA, USA).

Xylose isomerase activity was determined in a reaction mixture that contained 50 mM Tris-HCl (pH 7.4), 10 mM MgCl<sub>2</sub>, 0.15 mM NADH, 0.5 U mL<sup>-1</sup> sorbitol dehydrogenase (Roche, Basel, Switzerland), 50 mM xylose and 20  $\mu$ L crude cell extract [56, 215]. Xylulokinase activity was measured in a reaction mixture that contained 50 mM Tris-HCl (pH 7.5), 2 mM MgCl<sub>2</sub>, 2 mM ATP, 0.2 mM phosphoenolpyruvate, 0.2 mM NADH, 10 U mL<sup>-1</sup> pyruvate kinase (Sigma-Aldrich, St. Lois, MO, USA), 10 U mL<sup>-1</sup> lactate dehydrogenase (Sigma-Aldrich, St. Lois, MO, USA), 13 mM xylulose and 0.5  $\mu$ L crude cell extract [51]. Diaminopimelate dehydrogenase activity was determined in a reaction mixture that contained 200 mM glycine (pH 10.5), 10 mM MgCl<sub>2</sub>, 2 mM NADP, 4 mM *meso*-diaminopimelate and 50  $\mu$ L crude cell extract [39].

Quantification of the activity was realized by measuring the change of NADH absorption at 340 nm (Sunrise, Tecan, Männedorf, Switzerland). It was converted to enzymatic units, i.e. the amount of enzyme that catalyzes the reaction at a rate of 1  $\mu$ mol min<sup>-1</sup>.

#### 3.6.5 Mass spectrometry

Labeling patterns of proteinogenic amino acids and of secreted trehalose, required for calculation of metabolic fluxes, were determined from the proteinogenic amino acids and from the secreted trehalose by GC-MS [100]. Sample preparation and measurements were carried out as described previously [26], using the GC system 7890A gas chromatograph and a quadrupole MS detector (inert MSD 5979C, Agilent Technologies, Waldbronn, Germany). The <sup>13</sup>C labeling patterns were determined from three samples, which were taken at different time points during cultivation, to verify the isotopic steady-state. The mass spectra of alanine, valine, threonine, aspartate, glutamate, serine, phenylalanine, glycine and tyrosine from the cell proteins and of the trehalose from the culture supernatants, which were corrected for the

natural abundance of all stable isotopes then served as input for the flux calculation.

#### 3.6.6 Metabolic flux analysis

Metabolic flux calculation was performed by Judith Becker at the Institute of Systems Biology at the Saarland University. The metabolic network for flux estimation comprises the central metabolism with glycolysis, the pentose phosphate pathway, the tricarboxylic acid cycle and the glyoxylate shunt. All reactions for the interconversion of the three-carbon metabolites of glycolysis and the four-carbon metabolites of the TCA cycle present in *C. glutamicum* Lys12 as well as typical by-product-forming pathways were included.

For flux estimation, the metabolic network of *C. glutamicum* was implemented in the open source software OpenFlux [34, 167]. The software was extended to handle <sup>13</sup>C labeling data from the two parallel experiments by changing the subroutines 'leastSQ' and 'mdvGenerator' (Fig. 37 in the appendix). In addition, the specification file, reflecting the parallel labeling experiments ('inputSubEMU.mat'), was adapted accordingly, using the guided substrate input specifications [166].

Directly measured fluxes from three biological replicates (glucose uptake, formation of biomass, lysine and by-products) were used to calculate the free fluxes in the metabolic network of *C. glutamicum* Lys12 [34, 167]. The set of fluxes that gave the minimal deviation between the experimental and the simulated mass isotopomer fractions considered to be the best calculation of the intracellular flux distribution. For error considerations, a weighted sum of least-squares was applied [229]. The labeling data were corrected for the natural labeling of the inoculum and for natural isotopes [228]. Statistical analysis providing 90 % confidence intervals of the estimated fluxes was carried out by Monte-Carlo approach [167, 233].

#### 3.7 Calculation of production parameters

#### 3.7.1 Yields and rates

For calculation of yields during cultivations, the amount of product  $(P_{t_n})$  formed and substrate  $(S_{t_n})$  utilized for each sample time  $t_n$  was calculated. The term product applies to biomass as well as to any metabolite formed. For simple shake flask cultivations at low temperatures, i.e. cultivations in chapters 4.3 and 4.2, equations 2 and 3 were used. Hereby,  $c_{P,t_n}$  denotes the concentration of the product at time  $t_n$ ; the denomination of the substrate concentration followed suit.

$$P_{t_n} = c_{P,t_n} - c_{P,t_0} \tag{2}$$

$$S_{t_n} = c_{S,t_0} - c_{S,t_n} \tag{3}$$

For cultivations where the broth volume ( $V_t$ ) significantly changed due to factors other than sampling, i.e. all fed-batch cultivations and shake flask cultivations in chapter 4.1, where evaporation was considered due to higher temperatures, equations 4 and 5 were used

$$P_{t_n} = V_{t_n} \cdot c_{P,t_n} - V_{t_0} \cdot c_{P,t_0} + \sum_{n=t_0}^{t_n} (V_{sample,n} \cdot c_{P,n})$$
(4)

$$S_{t_n} = V_{t_0} \cdot c_{S,t_0} - V_{t_n} \cdot c_{S,t_n} - \sum_{n=t_0}^{t_n} (V_{sample,n} \cdot c_{S,n})$$
(5)

In all cases, the yield was determined by linear regression according to equation 6.

$$Y_{\frac{P}{S},t_{n}} = \frac{\sum_{n=t_{0}}^{t_{n}} P_{n} \cdot S_{n}}{\sum_{n=t_{0}}^{t_{n}} S_{n}^{2}}$$
(6)

Product per biomass yield  $Y_{\frac{p}{X},t}$  for the calculation of specific product formation rate (see equation 9) were calculated using equation 7.

$$Y_{\frac{P}{X},t_{n}} = \frac{\sum_{n=t_{0}}^{t_{n}} P_{n} \cdot X_{n}}{\sum_{n=t_{0}}^{t_{n}} X_{n}^{2}}$$
(7)

For the calculation of the maximum specific growth rate  $\mu_{max}$ , equation 8 was used. Here, the amount of biomass  $X_t$  was calculated according to equation 2 or 4 and  $n_t$  represents the amount of sample points used for the calculation. Like equation 6, equation 8 also a linear regression, but for the calculation of  $\mu_{max}$  only samples considered to be within the period of exponential growth were included, not all samples.

$$\mu_{max,t_n} = \frac{n \cdot \sum_{n=t_0}^{t_n} (\ln(X_n) \cdot n) - \sum_{n=t_0}^{t_n} n \cdot \sum_{n=t_0}^{t_n} (\ln(X_n))}{n \cdot \sum_{n=t_0}^{t_n} n^2 - \left(\sum_{n=t_0}^{t_n} n\right)^2}$$
(8)

Specific rates of product formation  $(q_P)$  and substrate uptake  $(q_S)$  were calculated according to equation 9 and 10, respectively.

$$q_{P,t_n} = \mu_{max,t_n} \cdot \frac{P,t_n}{X,t_n} \tag{9}$$

$$q_{S,t_n} = \frac{\mu_{max,t_n}}{Y_{\frac{X}{S},t_n}} \tag{10}$$

#### 3.7.2 NADPH supply and demand

Following the stoichiometry of lysine production (chapter 2.1.2, Fig. 7) and the metabolic routes in *C. glutamicum* (Fig. 6), equations 11 and 12 were used to calculate the NADPH supply and demand respectively.

$$Supply_{NADPH} = v_{MalE} + v_{Icd} + v_{Zwf} + v_{Gnd}$$
(11)

$$Demand_{NADPH} = 4 \times v_{LysA} + v_{Biomass} \times 16.429 \frac{mmol}{g}$$
(12)

The corresponding genes and names to the enzymes, abbreviated here, can be found in the list on page vi. The NADPH demand for biomass production is taken from [231].

## 3.7.3 Economic parameters

For the assessment of the economic effects of different cultivation temperatures (chapter 4.1), equations 13 and 14 were used to calculate revenues. Equation 13 hereby represents the fundamental concept and equation 14 the more mathematical notation, showing that in addition to the market prices of the substrate and product also the product yield, the operational cost of the production plant and the production rate play a role.

$$Revenue = \{Market \ price_{Product}\} - \{Cost_{Substrate}\} - \{Operational \ cost\}$$
(13)

$$\text{Revenue} = \left[\frac{\textcircled{e}}{mol_{Lys}}\right] - \left[\frac{\textcircled{e}}{mol_{Gluc}} \cdot Y_{\frac{S}{p}}\right] - \left[\frac{\underbrace{\textcircled{e}}{L\cdot h}}{\frac{mol_{Lys}}{L\cdot h}}\right]$$
(14)

# **RESULTS AND DISCUSSION**

# 4.1 Lysine production of *C. glutamicum* at elevated temperature

#### 4.1.1 Growth and product formation

Temperature is one of the most fundamental process parameters in microbiology and biotechnology. While it appears straightforward in research to keep cultivation conditions constant in order to enable the comparison of different experiments, traditionally used parameters might not be desirable for production performance. Almost all cultivations with *C. glutamicum* have been performed at a temperature around 30 °C [98]. However, it has been reported that certain strains may show increased production performance at higher temperature [149]. In this regard, first studies should investigate underlying metabolic responses to temperature in order to eventually identify mechanisms for an enhancement of production and further improvement of industrial applications of *C. glutamicum*.

The *C. glutamicum* strains under investigation comprised the entire range with regard to lysine production, including the wild type *C. glutamicum* ATCC 13032, the basic lysine producer *C. glutamicum* Lys1 and the hyper-producing strains *C. glutamicum* Lys12 and Lys12K (Tab. 2, Fig. 10). *C. glutamicum* ATCC 13032 hereby is the predecessor of all the strains, strain Lys1 displays lysine secretion due to the decoupling of the enzyme aspartokinase from metabolic feedback inhibition. The genome of the hyper-producers *C. glutamicum* Lys12 and Lys12K have been engineered on a systems-wide level to allow for optimized lysine production [18]. The latter strain additionally possesses the empty plasmid pClik 5a MCS as further



**Fig. 12:** Screening of *C. glutamicum* Lys12K growth and lysine production properties over a range of temperatures in shake flasks. The data were corrected for evaporation and sampling and the error bars represent the standard deviation of two flasks (for 30 and 40 °C four and three flasks, respectively).

metabolic burden (Tab. 2).

An initial screening study investigated the impact of temperature on the growth physiology of a highly advanced lysine producer in a range from the conventional value 30 to 40 °C. *C. glutamicum* Lys12K, the strain with the highest metabolic burden, was chosen for the first experiments as it was expected to be most sensitive to environmental changes. Growth experiments in shake flasks revealed a considerable temperature robustness of the strain (Fig. 12). Fastest growth of Lys12K occurred at 30 °C, but the strain was capable to grow up to 40 °C, the highest temperature tested. In contrast to a decreased growth efficiency, lysine production was enhanced at higher temperature. The lysine yield at 40 °C (0.44 mol mol<sup>-1</sup>) was 55 % higher than that at 30 °C (0.28 mol mol<sup>-1</sup>). The process productivity showed a maximum at 35.5 °C.

For three temperatures, 30, 38 and 40  $^{\circ}$ C, shake flask cultivations were performed of the four strains mentioned above and the robust performance at higher temperature was also observed for them (Tab. 6). Generally, higher temperature positively influenced lysine production at the expense of growth, i.e. biomass yield and specific growth rate. To some extent, the studied strains differed in their quantitative response to the imposed temperature. All lysine producing strains showed an increase in the lysine yield of about 40-50 % with a temperature increase from 30 to 38  $^{\circ}$ C. However, the highest temperature tested (40  $^{\circ}$ C) was only beneficial for lysine in

**Table 6:** Impact of cultivation temperature on growth kinetics and stoichiometry in the wild type *C. glutamicum* ATCC 13032 and lysine producing strains *C. glutamicum* Lys1, Lys12 and Lys12K, all grown on glucose minimal medium. The data comprise yields for the main products lysine ( $Y_{Lys/Glc}$ ) and biomass ( $Y_{X/Glc}$ ), maximum specific growth rate ( $\mu_{max}$ ) and yields for by-products trehalose ( $Y_{Tre/Glc}$ ), succinate ( $Y_{Suc/Glc}$ ) and glycine ( $Y_{Gly/Glc}$ ) and reflect mean values and deviations from three biological replicates (for Lys12K 30 and 38 °C four and two replicates, respectively). The underlying concentration measurement data are corrected for evaporation and sampling.

Strain	Temp. [°C]	μ <sub>max</sub> [h <sup>-1</sup> ]	Y <sub>X/Glc</sub> [g/mol]	Y <sub>Lys/Glc</sub> [mol/mol]	Y <sub>Tre/Glc</sub> [mmol/mol]	Y <sub>Suc/Glc</sub> [mmol/mol]	Y <sub>Gly/Glc</sub> [mmol/mol]
ATCC 13032	30 38 40	$\begin{array}{c} 0.41{\pm}0.01\\ 0.23{\pm}0.01\\ 0.12{\pm}0.00\end{array}$	$\begin{array}{c} 94.4{\pm}4.7\\ 57.7{\pm}5.2\\ 63.8{\pm}1.4\end{array}$	n/a n/a n/a	$\begin{array}{c} 3.7{\pm}0.2 \\ 10.4{\pm}0.6 \\ 12.3{\pm}0.5 \end{array}$	$\begin{array}{c} 10.9{\pm}0.0\\ 16.8{\pm}0.9\\ 15.5{\pm}1.3\end{array}$	$9.0{\pm}0.3$ $7.4{\pm}0.6$ $2.5{\pm}0.4$
Lys1	30 38 40	$\begin{array}{c} 0.39{\pm}0.00\\ 0.31{\pm}0.00\\ 0.08{\pm}0.01 \end{array}$	$\begin{array}{c} 86.9{\pm}3.6\\ 52.7{\pm}5.8\\ 37.8{\pm}1.2\end{array}$	$\begin{array}{c} 0.08{\pm}0.01\\ 0.12{\pm}0.00\\ 0.22{\pm}0.06\end{array}$	$\begin{array}{c} 4.3{\pm}0.0\\ 10.5{\pm}0.7\\ 15.4{\pm}1.2\end{array}$	$\begin{array}{c} 11.0{\pm}0.0\\ 16.6{\pm}0.5\\ 12.2{\pm}0.6\end{array}$	$5.2{\pm}0.2\\4.2{\pm}0.2\\1.6{\pm}1.1$
Lys12	30 38 40	$\begin{array}{c} 0.24{\pm}0.01\\ 0.11{\pm}0.00\\ 0.03{\pm}0.00 \end{array}$	$57.9{\pm}2.6$ $33.2{\pm}1.4$ $16.1{\pm}0.2$	$\begin{array}{c} 0.27{\pm}0.02\\ 0.43{\pm}0.01\\ 0.41{\pm}0.03\end{array}$	$\begin{array}{c} 5.7{\pm}0.9 \\ 14.4{\pm}1.1 \\ 12.6{\pm}0.4 \end{array}$	$\begin{array}{c} 6.0{\pm}2.0\\ 4.5{\pm}0.0\\ 6.0{\pm}0.1 \end{array}$	$\begin{array}{c} 4.0{\pm}0.7\\ 1.7{\pm}0.1\\ 0.5{\pm}0.1\end{array}$
Lys12K	30 38 40	$\begin{array}{c} 0.25{\pm}0.01\\ 0.10{\pm}0.02\\ 0.03{\pm}0.01 \end{array}$	$58.9{\pm}9.5\\25.8{\pm}4.9\\20.2{\pm}1.2$	$\begin{array}{c} 0.28{\pm}0.02\\ 0.43{\pm}0.01\\ 0.44{\pm}0.01\end{array}$	$\begin{array}{c} 3.1{\pm}1.3\\ 14.5{\pm}0.1\\ 12.7{\pm}0.3\end{array}$	$\begin{array}{c} 6.2{\pm}0.1 \\ 9.0{\pm}0.0 \\ 5.9{\pm}0.1 \end{array}$	$\begin{array}{c} 3.8{\pm}0.5\\ 0.1{\pm}0.0\\ 0.4{\pm}0.0\end{array}$

*C. glutamicum* Lys1, not for the highly engineered producers. With regard to growth, the highly advanced lysine producers appeared more sensitive than the wild type *C. glutamicum* ATCC 13032 and the basic producer *C. glutamicum* Lys1. The highest lysine yield achieved in these shake flask studies amounted to almost 60% of the theoretical maximum of 0.75 mol mol<sup>-1</sup> [197].

By-products formed at significant levels were trehalose, glycine and succinate (Tab. 6). The secretion of the disaccharide trehalose increased with higher temperature. In contrast, glycine formation was reduced as temperatures increased, whereby the relative changes differed among the strains to a varying degree. Succinate secretion did not follow a temperature-specific pattern but remained within a similar range for each strain. Leucine was secreted only by *C. glutamicum* ATCC 13032 at elevated temperatures to levels of 46 and 75  $\mu$ M at 38 and 40 °C respectively. Glutamate was formed at levels above 10  $\mu$ M only at the highest temperature of 40 °C to concentrations of 45, 30, 52 and 18  $\mu$ M for *C. glutamicum* ATCC 13032, Lys1, Lys12 and Lys12K, respectively. The increased levels of trehalose and glutamate at elevated



**Fig. 13:** Quantitative physiological characterization of *C. glutamicum* Lys12 grown at 30 and 38 °C. A: Linear correlation of produced biomass and lysine relative to the amount of glucose consumed by the cells. B: Exponential biomass production over the relevant time window of the cultivation. The correlations show the metabolic steady state during exponential growth of the cells.

temperatures are easily understood as those two compounds are known compatible solutes and synthesized by *C. glutamicum* under temperature stress conditions [43, 55, 85]. The observed change in growth and product formation was a clear indication for an altered physiology in the studied strains.

#### 4.1.2 Intracellular metabolic fluxes

**Flux calculation and NADPH balance.** The clear effect on growth and production suggested a temperature dependent metabolic shift towards a production-favorable carbon utilization. To assess the intracellular flux distribution, *C. glutamicum* Lys12 was grown on [1-<sup>13</sup>C] glucose as well as on an equimolar mixture of [U-<sup>13</sup>C] glucose and naturally labeled glucose in parallel. Subsequent GC/MS analysis of secreted trehalose and of amino acids (alanine, valine, threonine, aspartate, glutamate, serine, phenylalanine, glycine and tyrosine) from hydrolyzed cell protein delivered precise <sup>13</sup>C labeling data (Tab. 7).

The obtained experimental data were used to calculate metabolic flux distributions for each temperature (Fig. 15 and 16). The set of intracellular fluxes that gave the minimum deviation between experimental and simulated labeling patterns was taken as best estimate for the intracellular flux distribution. For both studied temperatures, identical flux distributions were obtained with multiple starting values for the flux **Table 7:** Relative mass isotopomer fractions of amino acids from hydrolyzed cell protein and the secreted trehalose from *C. glutamicum* Lys12, grown at 30 and 38 °C in 99% [1-<sup>13</sup>C] glucose and in an equimolar mixture of naturally labeled glucose and 99% [U-<sup>13</sup>C] glucose. Data denote experimental GC-MS data (exp) and calculated labeling patterns (calc), corresponding to the optimal flux fit. The mass isotopomer *m* represents the relative amount of non-labeled isotopomer, *m*+1 represents the amount of singly labeled mass isotopomer fraction and so on. Amino acids and trehalose were analyzed by GC/MS as trimethyldimethylsilyl derivatives and as trimethylsilyl derivate, respectively. The considered ion cluster for each compound is denoted by the *m*/*z* value of its mono-isotopic mass. (*Table continued on page 50*)

			1-1	<sup>13</sup> C			U-	<sup>13</sup> C	
		30	°C	38	38 °C		30 °C		°C
		exp	calc	exp	calc	exp	calc	exp	calc
Ala m∕z 260 (Alanine)	m m+1 m+2	0.582 0.300 0.094	0.588 0.296 0.093	0.605 0.281 0.092	0.610 0.278 0.090	0.358 0.149 0.123	0.355 0.145 0.121	0.353 0.157 0.133	0.351 0.155 0.132
Val m⁄z 288 (Valine)	m m+1 m+2	0.455 0.356 0.140	0.457 0.357 0.138	0.493 0.336 0.127	0.494 0.338 0.126	0.166 0.092 0.220	0.156 0.088 0.219	0.169 0.101 0.217	0.154 0.098 0.219
Thr <i>m/z</i> 404 (Threonine)	m m+1 m+2	0.392 0.357 0.172	0.386 0.360 0.174	0.421 0.346 0.162	0.421 0.346 0.162	0.165 0.202 0.199	0.166 0.195 0.195	0.193 0.201 0.183	0.189 0.196 0.181
Asp <i>m/z</i> 418 (Aspartate)	m m+1 m+2	0.390 0.356 0.174	0.386 0.359 0.175	0.418 0.345 0.164	0.420 0.345 0.162	0.164 0.201 0.204	0.166 0.194 0.195	0.192 0.201 0.184	0.189 0.195 0.181
Glu <i>m/z</i> 432 (Glutamate)	m m+1 m+2	0.316 0.358 0.209	0.317 0.363 0.206	0.354 0.353 0.192	0.356 0.356 0.189	0.088 0.130 0.231	0.084 0.120 0.235	0.108 0.128 0.242	0.096 0.122 0.253
Ser <i>m∕z</i> 390 (Serine)	m m+1 m+2	0.514 0.315 0.131	0.510 0.317 0.131	0.533 0.301 0.127	0.525 0.306 0.129	0.318 0.177 0.149	0.319 0.174 0.148	0.295 0.196 0.180	0.300 0.191 0.184
Phe <i>m∕z</i> 336 (Phenylalanine)	m m+1 m+2	0.390 0.366 0.170	0.385 0.367 0.173	0.424 0.352 0.156	0.420 0.355 0.159	0.056 0.053 0.091	0.054 0.049 0.092	0.067 0.058 0.090	0.054 0.054 0.095

parameters, suggesting that global minima were identified in the examined cases. Excellent agreement between experimentally determined and calculated mass isotopomer ratios was achieved for both examined conditions (Tab. 7) and metabolic (Fig. 13) and isotopic (Fig. 14) steady state could be validated.

The NADPH supply calculations (equations 11 and 12) in the  $^{13}$ C cultivations from the fluxes and the Monte-Carlo approach showed an NADPH supply of 2.28±0.17

			1- <sup>13</sup> C				U-	<sup>13</sup> C	
		30	°C	38 °C		30	°C	38	°C
		exp	calc	exp	calc	exp	calc	exp	calc
Gly <i>m/z</i> 246 (Glycine)	m m+1	0.752 0.176	0.751 0.177	0.749 0.179	0.743 0.185	0.384 0.187	0.378 0.190	0.387 0.196	0.381 0.196
Tyr <i>m∕z</i> 466 (Tyrosine)	m m+1 m+2	0.336 0.351 0.195	0.332 0.355 0.200	0.363 0.342 0.183	0.362 0.349 0.187	0.050 0.053 0.087	0.047 0.049 0.087	0.061 0.059 0.087	0.047 0.052 0.090
Tre <i>m/z</i> 361 (Trehalose)	m m+1 m+2	0.155 0.550 0.187	0.149 0.554 0.188	0.180 0.528 0.186	0.179 0.531 0.185	0.295 0.122 0.091	0.287 0.120 0.096	0.271 0.123 0.100	0.269 0.122 0.104
Ala <i>m/z</i> 232 (Alanine)	m m+1 m+2	0.613 0.297 0.090	0.620 0.292 0.088	0.639 0.275 0.086	0.643 0.272 0.085	0.411 0.148 0.441	0.407 0.141 0.452	0.408 0.161 0.431	0.404 0.158 0.437
Val <i>m∕z</i> 260 (Valine)	m m+1 m+2	0.463 0.354 0.135	0.472 0.354 0.131	0.503 0.334 0.122	0.511 0.333 0.118	0.186 0.096 0.377	0.178 0.089 0.384	0.188 0.108 0.364	0.176 0.104 0.373
Thr <i>m∕z</i> 376 (Threonine)	m m+1 m+2	0.422 0.356 0.165	0.415 0.360 0.167	0.451 0.342 0.154	0.447 0.345 0.155	0.196 0.257 0.296	0.192 0.246 0.308	0.227 0.244 0.303	0.219 0.237 0.321
Asp <i>m/z</i> 390 (Aspartate)	m m+1 m+2	0.422 0.354 0.166	0.415 0.359 0.168	0.445 0.340 0.158	0.446 0.344 0.156	0.195 0.253 0.295	0.192 0.246 0.308	0.226 0.241 0.301	0.218 0.236 0.322
Glu <i>m∕z</i> 330 (Glutamate)	m m+1 m+2	0.451 0.361 0.148	0.448 0.359 0.143	0.478 0.333 0.135	0.487 0.340 0.129	0.160 0.144 0.333	0.149 0.127 0.345	0.176 0.137 0.338	0.156 0.128 0.348
Ser <i>m/z</i> 362 (Serine)	m m+1 m+2	0.551 0.320 0.129	0.545 0.324 0.131	0.571 0.304 0.125	0.563 0.310 0.127	0.372 0.194 0.435	0.371 0.187 0.442	0.349 0.244 0.407	0.351 0.242 0.406
Phe <i>m/z</i> 234 (Phenylalanine)	m m+1 m+2	0.440 0.370 0.140	0.442 0.375 0.141	0.480 0.352 0.125	0.483 0.357 0.125	0.070 0.057 0.155	0.067 0.052 0.158	0.079 0.063 0.149	0.067 0.059 0.158
Gly <i>m/z</i> 218 (Glycine)	m m+1	0.827 0.173	0.826 0.174	0.825 0.175	0.822 0.178	0.465 0.535	0.456 0.544	0.473 0.527	0.463 0.537
Ser <i>m/z</i> 288 (Serine)	m m+1 m+2	0.597 0.305 0.098	0.596 0.308 0.097	0.620 0.286 0.094	0.615 0.291 0.094	0.392 0.177 0.431	0.393 0.168 0.441	0.368 0.231 0.401	0.373 0.229 0.400
Phe <i>m∕z</i> 302 (Phenylalanine)	m m+1	0.731 0.194	0.716 0.206	0.728 0.196	0.713 0.209	0.386 0.200	0.368 0.201	0.389 0.209	0.372 0.208
Asp <i>m/z</i> 316 (Aspartate)	m m+1 m+2	0.461 0.355 0.139	0.457 0.362 0.141	0.490 0.337 0.130	0.492 0.343 0.129	0.209 0.259 0.290	0.206 0.249 0.305	0.242 0.243 0.297	0.235 0.237 0.320

**Table 7:** *(Continued from page 49)* Relative mass isotopomer fractions of amino acids from hydrolyzed cell protein and the secreted trehalose.



**Fig. 14:** Mass-charge ratio of different amino acids at different biomass concentrations during the cultivation of *C. glutamicum* Lys12 on <sup>13</sup>C glucose showing the isotopic steady-state of the cultivation. A: 30 °C with [1-<sup>13</sup>C] glucose, B: 30 °C with an equimolar [<sup>13</sup>C6] and naturally labeled glucose mixture, C: 38 °C with [1-<sup>13</sup>C] glucose, D: 38 °C with an equimolar [<sup>13</sup>C6] and naturally labeled glucose mixture. ● alanine, ○ valine, ■ threonine, □ glutamate, ▲ serine, △ phenylalanine, ▶ glycine,

and  $2.97\pm0.08 \text{ mol}_{\text{NADPH}} \text{ mol}_{\text{Glucose}}^{-1}$  at 30 and 38 °C, respectively. On the demand side, the values were  $2.03\pm0.12 \text{ mol}_{\text{NADPH}} \text{ mol}_{\text{Glucose}}^{-1}$  at 30 °C and  $2.23\pm0.07$ mol\_{NADPH} mol\_{\text{Glucose}}^{-1} at 38 °C. This means a supply to demand ratio of 1.12 at 30 °C of 1.33 at 38 °C. This ratio not being closed, but above the value of 1, indicates that not all biochemical reactions have been taken into account. On the demand side, this could be caused by the factor for the NADPH demand of biomass production in our set up possibly being lower than the one taken from the literature and used in equation 12. As NADPH is involved in the generation of nucleic and fatty acids [195], higher molecular turnover and changes in the cell-membrane composition [32] might explain the even higher NADPH-excess at elevated temperatures. On the supply side the only of the major NADPH-generating enzymes [195] not considered in the calculation of the balance (equation 11) is the polyphosphate/ATP-dependent NAD kinase (EC:2.7.1.23) [129]. Due to the already existing apparent over supply, a strong generation of NADPH by this enzyme is unlikely. Fundamentally, the supply to demand ratio exceeding a value of 1 shows that carbon fluxes as calculated can well account for the extensive demand in redox power put forward by lysine yields as high as the ones seen in *C. glutamicum* Lys12.

**Metabolic fluxes through glycolysis and pentose phosphate pathway.** At the reference temperature of 30 °C, *C. glutamicum* Lys12 channeled 86 % of the consumed glucose into the PP pathway (Fig. 15), which is in good accordance with previous estimates [18]. The PP pathway flux could be determined with high precision, which becomes evident from the narrow 90 % confidence interval. The metabolically important branching point between glycolysis and PP pathway hence was reflected sensitively in the labeling data. At elevated temperature, cells metabolized glucose via the PP pathway even more than at 30 °C. The relative flux was as high as 101 % (Fig. 16). Thus, cells almost exclusively channel the substrate into the PP pathway, despite the EMP route being fully functional.

The strong connection between the oxidative part of the PP pathway and lysine production has been pointed out in the introduction (chapter 2.1.1) and is based on the high NADPH demand during the bio-synthesis of the amino acid, as the synthesis of one molecule of lysine requires 4 molecules of NADPH. The integrated inspection of <sup>13</sup>C metabolic flux data from this work with that of other lysine-producing strains even suggests a direct correlation of flux through the PP pathway and lysine biosynthesis (Fig. 17). Consequently, metabolic engineering efforts largely have been aimed at increased carbon flux through the PP pathway. For example, replacement of the PP pathway enzymes glucose 6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH) by mutated variants with superior kinetic properties successfully pulled extra carbon into the PP pathway and increased lysine production [10, 152]. In addition, the pushing of carbon into the PP pathway through over expression of gluconeogenic fructose 1,6-bisphosphatase (FBP) improved lysine production in *C. glutamicum* [13].

An interesting approach in this direction reported on a disrupted glycolysis through elimination of phosphoglucoisomerase, which forced *C. glutamicum* to completely convert glucose via the PP pathway [135]. Unfortunately, such PGI null mutants exhibit imbalanced growth and production. It is clear from the non-uniformity of the trehalose labeling and the overall existing labeling information from [1-<sup>13</sup>C] glucose presented in this work, that a significant inter-conversion of glucose 6-phosphate and fructose 6-phosphate was present. The reason behind this is, that the



**Fig. 15:** *In vivo* activity of the central metabolic pathways of *C. glutamicum* LYS12 at 30 °C. *(Caption continued on next side.)* 

**Fig. 15:** (*Caption continued from previous side*). *In vivo* activity of the central metabolic pathways of *C. glutamicum* LYS12 at 30 °C during growth on glucose. All fluxes are given as molar percentage of the specific glucose uptake rate of  $4.1\pm0.15$ , which was set to 100 %. An arrow indicates the direction of the net flux of reversible reactions. Statistic evaluation was carried out by Monte-Carlo analysis. The errors reflect the corresponding 90 % confidence intervals for the different fluxes. AcCoA = acetyl CoA, AKG =  $\alpha$ -ketoglutarate, BM = biomass, DAP = diaminopimelate, DHAP = dihydroxyacetone phosphate, E4P = erythrose 4-phosphate, F6P = fructose 6-phosphate, GAP = glyceraldehyde 3-phosphate, Gly<sub>ex</sub> = extracellular glycine, G6P = glucose 6-phosphate, ICI = isocitrate, Lys = lysine, Lys<sub>ex</sub> = extracellular lysine, MAL = malate, OAA = oxaloacetic acid, PEP = phosphoenolpyruvate, PGA = 3-phosphoglycerate, Pyr = pyruvate, P5P = pentose 5-phosphate, SUC = succinate, SUC<sub>ex</sub> = extracellular succinate, S7P = sedoheptulose 7-phosphate, Tre<sub>ex</sub> = extracellular trehalose.

6-phosphoglucoisomerase dehydrogenase decarboxylates 6-phosphoglucoisomerase at the isotopically labeled C1 atom when  $[1^{-13}C]$  is used, thus eliminating any labeling information, unless said reversibility of the phosphoglucoisomerase is active. This also underlines the high flexibility of the metabolic routes to adapt to environmental changes. In addition to this obvious importance of phosphoglucoisomerase to fine-tune metabolism, pathway simulations provide evidence that deletion of this enzyme should not be the first choice to reach high PP pathway fluxes, because optimal lysine production needs this enzyme to re-cycle carbon from lower parts of metabolism back into the PP pathway [109, 140]. This re-directing of carbon is exactly what happened in the 38 °C cultivations of *C. glutamicum* Lys12 (Fig. 16), a strain metabolically designed for optimized lysine production. The combination of bioprocess and metabolic engineering hence led to the highest flux through the PP

**Fig. 16:** (*Figure on next side*). *In vivo* activity of the central metabolic pathways of *C. glutamicum* LYS12 at 38 °C during growth on glucose. The color of the arrows indicate the deviation of the *in vivo* fluxes in positive (green, solid) or negative (red, hatched) direction from the theoretical optimum for lysine production [34, 140], when compared to the fluxes at 30 °C. For the flux from the PEP/Pyr to the OAA/Mal pool marked with an asterisk, the flux is closer to the theoretical optimum only when the net-flux between those metabolite pools is considered. All fluxes are given as molar percentage of the specific glucose uptake rate of  $3.9\pm0.17$  mol g<sup>-1</sup> h<sup>-1</sup> which was set to 100%. An arrow indicates the direction of the net flux of reversible reactions. Statistic evaluation was carried out by Monte-Carlo analysis. The errors reflect the corresponding 90% confidence intervals for the different fluxes. The same abbreviations as in Fig. 15 were used.



**Fig. 16:** (*Caption on previous side.*) *In vivo* activity of the central metabolic pathways of *C. glutamicum* LYS12 at 38 °C.



**Fig. 17:** Correlation between the carbon fluxes through the pentose phosphate pathway and towards lysine production determined with <sup>13</sup>C-Flux analysis in different strains of *C. glutamicum.* ● This work,  $\bigcirc$  [13],  $\blacksquare$  [10],  $\square$  [12],  $\blacklozenge$  [18],  $\diamondsuit$  [232],  $\blacktriangle$  [100],  $\triangle$  [117],  $\blacktriangledown$  [234],  $\bigtriangledown$  [101],  $\triangleright$  [136]  $\triangleleft$  [158].

pathway ever reported for *C. glutamicum* and as a consequence to the highest lysine yield ever achieved. Reaching lysine yields even closer to the theoretical maximum would require even higher PP pathway fluxes [109, 140].

**Fluxes at the pyruvate node.** The pyruvate node in *C. glutamicum* comprises a number of reactions (chapter 2.1.1), but the outline of the labeling experiments in this work allowed only for the resolution of net-fluxes between the oxaloacetate/malate (OAA/Mal) and the phosphoenolpyruvate/pyruvate (PEP/Pyr) pool. The carbon re-cycling, which occurs at both temperatures, is much stronger at 38 °C. In both experimental set-ups, the net-flux led from the PEP/Pyr pool to the OAA/Mal pool, namely 43.9 % and 51.5 % for 30 and 38 °C, respectively. The flux at the higher temperature was hence closer to the optimal theoretical value of 75 % [109, 140]. At 38 °C, a lower flux from pyruvate to acetyl-CoA occurs. The incorporation or release of  $CO_2$  taking place in the anaplerotic reactions with regard to biomass formation will be discussed later in this chapter.

**Metabolic fluxes through the TCA cycle.** The tricarboxylic acid (TCA) cycle exhibited fully cyclic operation with little difference at both temperatures as shown e.g. by the entry reaction into the TCA cycle, citrate synthase (34.6 % at 30 °C and 30.4 % at 38 °C). This carbon flux was rather low compared to other *C. glutamicum* strains, including the predecessors of Lys12 [18]. Of those predecessor strains, the last 5 modifications implemented (in order: C1372T *pyc*, P<sub>sod</sub>*pyc*, A1G *icd*, P<sub>tuf</sub>*fbp*, P<sub>sod</sub>*tkt*) all led to reduced flux through the TCA cycle as well as to higher fluxes through the PP pathway and toward lysine formation [18].

Obviously, rather low amounts of carbon were available for the formation of energy in the form of ATP at both temperatures, which is likely due to the enormous withdrawal of the TCA cycle-derived building block oxaloacetate for lysine production and to a lower degree due to biomass formation, competing with energy forming reactions. The energy shortage thus was one off the reason of the reduced growth rate at high temperature, which was particularly pronounced in these more advanced producers (Tab. 6). At 38 °C the hyper-producers *C. glutamicum* Lys12 and Lys12K reach only 46 and 40 % of their specific growth rate at 30 °C. This temperature-dependent reduction is less drastic for the wild type strain *C. glutamicum* ATCC 13032 (56 %) and the basic lysine producer *C. glutamicum* Lys1 (79 %). With increasing temperature, the energy demand for cellular maintenance increases [4, 64, 133]. This has a more severe effect on strains like the advanced producers, which exhibit a lower energy metabolism already at low temperatures and need to commit a bigger percentage of it to cellular maintenance, which led to a stronger decrease of the specific growth rate with rising temperatures for those strains.

**By-product and biomass formation.** By-product formation of *C. glutamicum* Lys12 was low for either of the tested temperatures with only trehalose, glycine and succinate being formed under the conditions of the flux experiments. Only the disaccharide trehalose was secreted at higher levels at the elevated temperature, as can be expected due to its role as a compatible solute, as mentioned earlier. Glycine and succinate both showed higher production at 30 °C, yet even that amounted only to low concentrations.

The biomass production of *C. glutamicum* Lys12 was reduced by over 40 % at 38  $^{\circ}$ C when compared to the values at 30  $^{\circ}$ C (Tab. 6), which obviously also is reflected in the carbon fluxes. One of the contributing factors to the lower biomass yield

at higher temperatures was the stronger  $CO_2$  production. There are two carbon dioxide creating reactions of the TCA cycle, catalyzed by the isocitrate dehydrogenase (Icd) and oxoglutarate dehydrogenase. While the carbon flux through the latter was equal for both temperatures, the  $CO_2$  generation by Icd was higher at 30 °C. The other  $CO_2$ -generating reaction where this applies, is the conversion of pyruvate to acetyl-CoA by the pyruvate dehydrogenase. All other reactions of the central metabolism involving  $CO_2$ , namely the ones in the PP pathway (6-phosphogluconate dehydrogenase), the lysine biosynthetic pathway (diaminopimelate decarboxylase) and the reactions of the pyruvate node (see chapter 2.1.1) had higher (net-)fluxes at 38 °C. Summed up, the carbon loss due to  $CO_2$  generation is 9 % higher at 38 °C than at 30 °C.

All in all, the carbon loss through by-product formation in *C. glutamicum* Lys12 was low at both temperatures. The higher loss due to  $CO_2$  generation and the massively increased lysine production explain the significant reduction in biomass yield at elevated temperature.

**Correlation of fluxome to transcriptome.** Fluxome analysis of the central carbon metabolism and the lysine biosynthesis pathway of *C. glutamicum* Lys12 at 30 and 38 °C revealed differences between the two cultivation conditions. PP pathway flux and net flux from the PEP/Pyr to the OAA/Mal pool both increased by about 20% and flux toward lysine by about 50% during cultivations at the elevated temperature (Tab. 8). On the other hand, carbon flux from pyruvate to acetyl-CoA and the subsequent flux to  $\alpha$ -ketoglutarate via isocitrate was reduced to 80 and 90% of the 30 °C-value at 38 °C, respectively. Carbon flux toward biomass was decreased by about 40% at the higher temperature. The by-products formed showed a mixed response to higher temperatures. While the carbon flux toward compatible solute trehalose showed the highest relative increase of about 140%, fluxes toward glycine and succinate were significantly reduced at 38 °C (Tab. 8).

In the literature, transcriptome analysis at 30 and 40 °C of the *C. glutamicum* ATCC 13032 based strain AHP3 is reported [149]. Like *C. glutamicum* Lys12, AHP3 also possesses the genomic modifications V59A *hom*, T3111 *lysC* and P458S *pyc*, but lacks the additional ones present in strain Lys12 (A1G *icd*,  $P_{tuf}fbp$ ,  $P_{sod}tkt$ ,  $\Delta pck$ ,  $P_{sod}pyc$ ,  $2 \times ddh$ ,  $2 \times lysA$ ,  $P_{sod}dapB$ , G1A  $P_{sod}lysC$ ). These differences in the genomic background of the two strains have to be kept in mind as it makes a direct comparison

of the data a bit difficult. However, one could still use the combined inspection on the effect of elevated temperature on transcriptome and fluxome as a first qualitative glance to possibly underlying metabolic and regulatory mechanisms. Tab. 8 connects calculated flux-ratios to the genes, whose transcription level is given in the literature.

The increase in PP pathway and lysine production fluxes are reflected by enhanced gene expression only for a few exceptions (*asd, rpi, pgl*). Even though those pathways underwent heavy genomic optimization in *C. glutamicum* Lys12, the increased lysine production at elevated temperatures in strain AHP3 allow the fairly safe assumption, that fluxes through said metabolic routes were also higher. Higher fluxes, despite lower transcription (i.e. lower enzyme amounts) indicate a higher enzymatic activity which could have been caused by the increased temperature.

This is different at the anaplerotic node. Flux from pyruvate and phosphoenolpyruvate to oxaloacetate is catalyzed by pyruvate carboxylase (*pyc*) and phosphoenolpyruvate carboxylase (*ppc*). The flux from the OAA/Mal to the PEP/Pyr pool can be catalyzed by malic enzyme (*malE*), oxaloacetate decarboxylase (*odX*) and by phosphoenolpyruvate carboxykinase (*pck*) in *C. glutamicum* AHP3, while the latter enzyme is eliminated in *C. glutamicum* Lys12. Here, both the fluxes from the PEP/Pyr to the OAA/Mal pool and the ones in the reverse direction were considerably increased at high temperatures. The transcriptome analysis on the other hand revealed, that genes involved in the catabolic direction (*pyc, ppc*) were expressed at lower levels at the higher temperature while the anabolic genes (*malE, pck*) show an opposite response. Transcriptome data for *odX* were not available. These findings indicate that both genes, *malE* and *odX*, are potential targets for further metabolic engineering. Even though this would mean a lower drain on the lysine precursor oxaloacetate, in the case of *malE*, this would also be coupled to a lower NADPH production, which might be negative for the production of the amino acid.

It seems, with regard to carbon core metabolism, that most changes in gene transcription and in flux are below a factor of two, with only a few exceptions, highlighted in Fig. 18. Besides the mentioned strong increase in flux and transcription level of malic enzyme, it is striking that the transcription of *gapA* and *gltA* was considerably reduced without the analogous fluxes acting accordingly. Especially for *gltA*, coding for citrate synthase and hence for a reaction competing with lysine formation, this might point to attenuation of the gene being another possible target

**Table 8:** Ratios of fluxes at 38 °C to fluxes at 30 °C ( $R_{Flux}$ ), the corresponding genes and the ratios of transcriptional levels (40 over 30 °C values,  $R_{Transc.}$ ) from the literature [149]. The comments point out the difference between the strains used for fluxome analysis (*C. glutamicum* Lys12) and for transcriptone data (*C. glutamicum* AHP3). Gene names can be found on page vi and in the literature [67]. Metabolite abbreviations are identical to Fig. 15. *Table continued on next side* 

Flux	R <sub>Flux</sub>	Gene	R <sub>Transc.</sub>	Comment
G6P→P5P	1 18	zwf opcA pgl	0.65 1.00 1.80	P <sub>sod</sub> (tkt,tal,zwf,opcA,pgl) in Lys12
	1.10	gnd rpe rpi	0.55 0.70 1.25	
G6P→Tre	2.42	n/a	n/a	
G6P→F6P	n/a	pgi	0.40	flux reversed, hence no R <sub>Flux</sub>
F6P→GAP	0.94	glpX pfp fda	0.90 1.10 0.80	P <sub>tuf</sub> fbp in Lys12, fbp named glpX in [149]
GAP,S7P→E4P,F6P	1.20	tal	0.60	$P_{1}(tkt tal xwf oncA ngl) in$
P5P,E4P $\rightarrow$ F6P,GAP 2 $\times$ P5P $\rightarrow$ GAP,S7P	1.24 1.20	tkt	0.67	Lys12
DHAP→GAP	0.94	tpiA	0.60	
GAP→PGA	0.99	gapA gapB pgk	0.37 0.70 0.55	
PGA→PEP	1.02	gpmB gpm eno	0.60 1.30 0.60	
Pyr→ACoA	0.78	pdhA pdhB lpd	0.80 0.60 0.70	
PEP/Pyr→OAA/Mal	1.78	ррс рус	0.55 0.70	P <sub>sod</sub> pyc in Lys12
OAA/Mal→PEP/Pyr	3.03	malE odX pck	2.50 n/a 2.00	$\Delta pck$ in Lys12
ACoA→ICI	0.88	gltA acn	0.31 1.35	

Flux	R <sub>Flux</sub>	Gene	R <sub>Transc.</sub>	Comment
ICI→AKG	0.88	icd	0.70	A1G icd in Lys12
OAA/Pyr→Lys	1.48	dapA dapB dapC dapE dapF ddh lysC aspC asd lysA	0.80 0.55 0.70 n/a 0.80 0.65 0.65 0.90 0.75 1.30 0.65	2×ddh, 2×lysA, P <sub>sod</sub> dapB, G1A P <sub>sod</sub> lysC in Lys12
Lys→Lys <sub>ex</sub>	1.55	lysE	1.00	
Suc→OAA	1.01	sdhB sdhA fumH	2.43 1.10 0.70	
AKG→Suc	1.00	odhA lpd sucC	1.00 0.75 1.30	
$Suc{\rightarrow}Suc_{ex}$	0.71	n/a	n/a	
$Gly \rightarrow Gly_{ex}$	0.50	n/a	n/a	
$\rightarrow$ BM	0.58	n/a	n/a	

 Table 8: Table continued from previous side

for further engineering. However, care has to be taken as two different strains are compared here.

# 4.1.3 Lysine production by *C. glutamicum* Lys12 at elevated temperature

Obviously, temperature has a strong impact on the metabolism of *C. glutamicum* and, in particular, on lysine production. Based on the findings, it appeared promising to explore the increased performance of higher temperature in the relevant production set-up of a fed-batch fermentation. To this end, *C. glutamicum* Lys12-cultivations


log<sub>2</sub> Gene expression change

**Fig. 18:** Flux changes and transcriptional changes of reactions and genes in the central metabolism of *C. glutamicum* occurring with the switch from 30 °C to higher cultivation temperatures (38 and 40 °C for flux and transcription data, respectively). Transcriptome data are taken (and partly estimated) from the literature [149]. Values larger than 1 indicate an at least twofold change. The flux assigned to malic enzyme in this graph might be partly due to oxaloacetate decarboxylase, but no transcriptome data are available. Gene names are listed on page vi.

were conducted at 30 and 38  $^\circ C.$ 

*C. glutamicum* Lys12 was able to grow and produce L-lysine under both conditions. For each temperature, two bioreactors were run in parallel as replicates, which showed good agreement. At 38 °C, the final lysine titer was significantly increased. After a process time of 90 hours, 119 and 127 g L<sup>-1</sup> L-lysine were produced, at 30 and at 38 °C, respectively (Fig. 19). Particularly, the absolute amount of lysine increased up to the end at 38 °C (Fig. 20 B), while it leveled out at 30 °C (Fig. 20 A). Obviously, cells retained their production capability to a higher degree at the elevated temperature. This can also be seen from the frequency of the feed pulses (vertical gray bars in Fig. 19), which at 38 °C was less reduced to the dissolved oxygen signal in the bioreactor. Hence active glucose utilization by the cells led to faster depletion of the substrate and subsequent rapid increase of the dissolved oxygen signal which again triggered the next feeding pulse.

At both temperatures, the initial glucose was depleted after approximately 20 hours. The pulse-wise addition of feeding solution caused the glucose concentration never to exceed 8 g  $L^{-1}$  at the sampling time points. The volume of a feed-pulse was



**Fig. 19:** Lysine production *C. glutamicum* Lys12 at 30 °C (A) and 38 °C (B). The data represent mean values and standard deviation of 2 runs. The light gray vertical peaks show the operation of the feed pump. The pump and DO data are exemplary taken from one of the runs. The dashed vertical line marks the transition from batch to feed phase. Wall growth of the cells at 38 °C towards the end of the cultivation made accurate determination of the cell dry weight difficult. This can be seen from the large error bars and the sudden change of the biomass concentration right at the end of the process.

set to prevent any given pulse to add glucose to a concentration higher than 30 g L<sup>-1</sup>. The volume increase and hence the total amount of glucose added to the bioreactor was higher for 30 °C (159 g) than for 38 °C (146 g) as can be seen from the broth volume (Fig. 19) and the total glucose uptake over time (Fig. 20). The reduced glucose consumption coupled to higher lysine production led to the overall increased yield at 38 °C, as compared to 30 °C (Fig. 21).



**Fig. 20:** Lysine production *C. glutamicum* Lys12 at 30 °C (A) and 38 °C (B). The absolute values of the concentration values in Fig. 19 are shown.



**Fig. 21:** Lysine yield (A: 30 °C, B: 38 °C) during fed-batch cultivation of *C. glutamicum* LYS12. The dashed vertical line marks the transition from batch to feed phase. Different symbols represent values for the two replicates.

As already reflected by the shake flask studies, less biomass was formed in the fedbatch process at 38  $^{\circ}$ C. Additionally, the biomass concentration in the broth at 38  $^{\circ}$ C



**Fig. 22:** Lysine productivity during fed-batch cultivation of *C. glutamicum* LYS12. The dashed vertical line marks the transition from batch to feed phase.

declined towards the end of the process, whereas it remained constant at 30 °C. This effect was likely not due to actual cell death, but rather resulted from increased wall growth during the last period of the cultivation. At both temperatures, the specific growth rate during the exponential growth phase was lower than expected from the shake flask experiments (0.1 h<sup>-1</sup> and 0.08 h<sup>-1</sup> for 30 °C and 38 °C, respectively). This implies that during the bioreactor cultivation, the higher temperature led to a decrease of the specific growth rate by 20 % as compared to 54 % in shake flasks. One possible explanation is the high initial sugar concentration of 115 g L<sup>-1</sup> glucose which might have led to substrate inhibition during the batch phase. In accordance, the bioreactor cultivation at the reference temperature showed a higher lysine yield (0.38 g g<sup>-1</sup>) than expected from the shake flask cultivations (Fig. 21 A).

The final concentration of 127 g L<sup>-1</sup> L-lysine, produced at 38 °C, is among the highest titers reported in the literature [18, 239]. The overall yield was boosted by about 10 % to 0.42 g g<sup>-1</sup> through the temperature increase (Fig. 21 B). The productivity (Fig. 22) showed a similar time course for both temperatures with a steep increase during the batch phase, a slightly reduced increase during the first 10 hours of the feeding phase followed by a constant decrease until the end of the fermentation. While the peak productivity reached was higher at 30 °C (2.9 g L<sup>-1</sup> h<sup>-1</sup> compared to 2.7 g L<sup>-1</sup> h<sup>-1</sup> at 38 °C), the productivity during the batch phase and more notably by the end of the feeding phase was higher for the elevated temperature.

Potential measures for further optimization of the bioprocess include the adjust-

ment of the feed pulses towards slower feed-rates and higher volumes. This would effectively lead to fewer pulses and hence fewer situations of glucose limitation for the cells. Yet this approach would need to be executed with care as addition of inhibiting amounts of the substrate also must be avoided. Another possible lever is the initial glucose concentration, which could be optimized to allow for a more efficient bioprocess. Lastly, a temperature gradient throughout the course of the cultivation might lead to further improved performance. For this, further investigations are needed to characterize the cellular response to such an incremental or gradual change of the environment.

## 4.1.4 The metabolic response of *C. glutamicum* to elevated temperature is of industrial value

Temperature is a major factor of industrial fermentation as it affects the behavior of the microorganisms that catalyze the bioconversion. In addition, higher temperature largely reduces the risk of contamination. Furthermore, it reduces the amount of cooling water necessary to maintain constant temperature in industry scale fermentation vessels. This is especially true in tropical regions, where typical carbon sources for lysine production are grown [149] and many lysine production plants are located [47]. It appears thus more economical to perform fermentations at elevated temperature. Thereby, impact of the temperature on microbial growth and production properties must be considered.

It has been observed that lysine producing *C. glutamicum* and related species show varying growth and production performances during cultivation at 40 °C, as compared to the commonly used 30 °C [151]. As presented in this work, increased product yields at elevated temperature were observed for *C. glutamicum* Lys1 and also for *C. glutamicum* Lys12 and a newly derived strain (Tab. 6). Furthermore, high temperature was also found beneficial for fed-batch production of lysine (Fig. 19), resulting in a high titer of 127 g L<sup>-1</sup> and an efficient conversion of the substrate (42 g g<sup>-1</sup>). Fed-batch processes are most commonly used for biotechnological production of lysine. For lysine, a '*low value, high volume*' commodity, product yield is of crucial importance, because the substrate amounts to about 50 % of the variable costs of production [41, 98] and a higher product yield therefore is a major economic advantage.



**Fig. 23:** Cost efficiency of identical lysine fermentations, considering the ratio of product to raw material (sugar) cost. The solid lines represent the temperature dependence of the profit at a given lysine-to-glucose price ratio. The bold green lines represent the upper and lower boundaries (UB and LB, respectively) of ratios from the years 2008 to 2014. The data for the calculation of the boundaries were taken from online resources and databases [221, 222]. The data for the profit are based on interpolated values for the yield and productivity shown in Fig. 12 and on equation 14. Red markers are placed at the maximum value of each profit curve. It becomes clear that at a lower ratio of value creation (i.e. a smaller profit margin), the temperature with the greatest profitability should be rather high. In such a scenario, high yields are preferable over fast production. The data should be taken as a qualitative trend, as fixed costs and additional variable costs other than the profit margin are not considered.

Temperature variation enables an easy fine-tuning of key performance indicators: shifting either to higher productivity or to higher yield. This provides manufacturers the opportunity to quickly react to price developments on the market with an existing production plant set-up merely by changing this one process variable, even for the same strain. Fig. 23 visualizes this effect and shows how optimal temperature, i.e. the most profitable temperature, depends on the lysine-to-glucose price ratio. In case the glucose price is comparably low, the optimal temperature approaches the the one with the highest productivity, 35.5 °C. If, on the other hand, the profit margin

of glucose to lysine is low (i.e. a high glucose or/and a low lysine price), it is more profitable to produce lysine with a higher yield. In this scenario, higher cultivation temperatures are preferable.

The basic idea (equation 13) and the mathematical formulation (equation 14) behind Fig. 23 are admittedly simplified. For a more reliable calculation of the optimal temperatures, more than information about market prices of substrate and product and the temperature dependence of the microbial production behavior would be necessary. Additionally, detailed knowledge of engineering aspects of the different temperatures (e.g. cooling costs) as well as of business numbers (e.g. market volume of product) are also required. It is, however, noteworthy, that none of the profitmaxima shown in Fig. 23 suggest a temperature close to  $30 \,^{\circ}$ C, which is often used in scientific studies of *C. glutamicum*, as the optimal production temperature.

### 4.2 Lysine production from xylose

## 4.2.1 Growth and lysine production of xylose-utilizing *C. glutamicum* Xyl1

Naturally, *C. glutamicum* ATCC 13032 strains cannot use xylose as carbon source [24]. In order to implement xylose utilization, *C. glutamicum* Lys12 was transformed with the episomal plasmid pClik5a P<sub>gro</sub>xylAB, which harbored the xylose catabolic genes *xylA* and *xylB* from *E. coli* (Tab. 2). The gene *xylA* codes for xylose isomerase, while *xylB* codes for xylulokinase. The two enzymes catalyze the conversion from xylose to xylulose and further on to xylose 5-phosphate, an intermediate of the pentose phosphate pathway, respectively (Fig. 28). *C. glutamicum* natively possesses a *xylB* gene, so the heterologous expression of *E. coli xylA* allowed *C. glutamicum* to grow on xylose [97]. The heterologous expression of both genes from *E. coli* also enabled the bio-synthesis of cadaverine from hydrolyzed hemicellulose in *C. glutamicum* [36].

The mechanisms of xylose transport into *C. glutamicum* cells is not clear at this stage [97], but there are indications for the involvement of hexose phosphotransferase systems [219]. Additionally, an arabinose transporter coded by the *C. glutamicum* 31831 gene *araE* has been shown to enhance growth of a *C. glutamicum* R based strain at low xylose concentrations [174].

The process of transferring the plasmid pClik5a  $P_{gro}xylAB$  into *C. glutamicum* Lys12 is described in chapter 3.1. The resulting strain was designated *C. glutamicum* Xyl1. In contrast to its parent *C. glutamicum* Lys12, the novel mutant formed colonies within two-day incubation on agar plates with xylose as sole carbon source. Additionally, *C. glutamicum* Xyl1 exhibited significant activity of xylose isomerase  $(119 \pm 24 \text{ U g}_{\text{protein}^{-1}})$  and xylulokinase  $(153 \pm 11 \text{ U g}_{\text{protein}^{-1}})$  respectively, while the parent strain showed no activity for xylose isomerase and only a basal level for xylulokinase  $(6 \pm 5 \text{ U g}_{\text{protein}^{-1}})$ .

During growth on xylose minimal medium, xylose was consumed from early on (Fig. 24 A). Cells reached a specific growth rate of  $0.17 \text{ h}^{-1}$  in chemically defined medium containing xylose (Tab. 9). With decreasing levels of xylose, growth slowed down slightly, but the carbon source was depleted completely. Additionally, *C. glutamicum* Xyl1 accumulated substantial amounts of lysine. The lysine yield (0.30 C-mol C-mol<sup>-1</sup>) was constant over time (Fig. 25 A), which indicated that about



**Fig. 24:** Growth and lysine production of *C. glutamicum* Xyl1 on minimal medium with xylose (A) and glucose (B) as sole carbon and energy source. The data reflect means and deviations from three independent replicates.

30% of substrate carbon was directed to product formation (Tab. 9). During the cultivation, no by-products were formed to significant concentration. The strain *C. glutamicum* Xyl1 produced lysine also from glucose and exhibited a molar lysine yield of 0.35 C-mol C-mol<sup>-1</sup> on this substrate (Fig. 24 B, Fig. 25 B). The maximum specific growth rate on the hexose sugar was slightly higher (0.20 h<sup>-1</sup>), as compared to growth on xylose (Tab. 9). On glucose, the disaccharide trehalose was the only by-product detected, and reached about 50 mg L<sup>-1</sup> at the end of the cultivation. The maximum specific rates of substrate uptake and lysine production were similar for both substrates. In all cases, the pH remained constant at  $7.1\pm0.2$  and the level of the dissolved oxygen remained above 65% of saturation, which indicated fully aerobic growth.

To investigate whether the effects of elevated temperatures observed for glucose



**Fig. 25:** Lysine yield of *C. glutamicum* Xyl1 on minimal medium with xylose (A) and glucose (B) as sole carbon and energy source. The different symbols reflect data from three independent replicates.

also apply on xylose, additional shake flask experiments were conducted at 38 °C (Tab. 9). In contrast to glucose (see chapter 4.1), higher temperature reduced the specific growth rate without increasing the lysine yield. This behavior was surprising as the reduction in the biomass yield was similar to cultivation on glucose, but the carbon seemed not to be re-directed towards lysine biosynthesis. A possible explanation could be the entry point of xylose derived xylose 5-phosphate into the central metabolism, which which is located downstream of the oxidative part of the PP pathway. This might limit the flux increase into the PP pathway and the concomitant increase in NADPH production (Fig. 15 and 16), which appears to play an important role for lysine production [136, 231, 234] (Fig. 17). Additional <sup>13</sup>C flux analysis on xylose could shed more light into the underlying cellular mechanisms. Hereby, the response of the increased temperature on the carbon flux through the TCA cycle would be of particular interest. Growth on xylose under more stressful conditions might require an increased maintenance and hence energy demand.

#### 4.2.2 Production from xylose in a fed-batch environment

The production performance of the novel *C. glutamicum* Xyl1 strain on xylose-based medium was next investigated in a fed-batch process (Fig. 26). The medium contained a mixture of xylose, molasses and corn steep liquor, in order to reflect a

**Table 9:** Growth and production physiology of *C. glutamicum* Xyl1 on chemically defined medium with glucose or xylose as sole carbon source at 30 and 38 °C. The data are means and standard deviations from three biological replicates and comprise maximal specific rates of growth ( $\mu_{max}$ ), substrate uptake ( $q_S$ ), lysine production ( $q_P$ ), and yields for biomass ( $Y_{X/S}$ ) and lysine ( $Y_{P/S}$ ,molar,  $Y_{P/S}$ ,C-molar).

Temperature	Parameter	Glucose	Xylose
30°C	$\mu [h^{-1}] q_{S} [mmol g^{-1} h^{-1}] q_{P} [mmol g^{-1} h^{-1}] Y_{P/S},molar [mol mol^{-1}] Y_{P/S},C-molar [C-mol C-mol^{-1}] Y_{P/S} q^{-1}]$	$\begin{array}{c} 0.20 \pm 0.01 \\ 5.3 \pm 0.3 \\ 1.5 \pm 0.1 \\ 0.35 \pm 0.01 \\ 0.35 \pm 0.01 \\ 0.22 \pm 0.00 \end{array}$	$\begin{array}{c} 0.17 \pm 0.01 \\ 5.7 \pm 0.7 \\ 1.4 \pm 0.1 \\ 0.25 \pm 0.01 \\ 0.30 \pm 0.01 \\ 0.10 \pm 0.00 \end{array}$
38 °C	$ \begin{array}{c} \mu \ [h^{-1} \ ] \\ Y_{P/S}, molar \ [mol \ mol^{-1} \ ] \\ Y_{X/S} \ [g \ g^{-1} \ ] \end{array} $	n/a n/a n/a n/a	$\begin{array}{c} 0.07 \pm 0.00 \\ 0.07 \pm 0.00 \\ 0.25 \pm 0.01 \\ 0.07 \pm 0.01 \end{array}$

typical medium for industrial production. The process started with a batch phase to achieve a fast increase in biomass concentration already in early process phases. The initial mixture of 115 g L<sup>-1</sup> total sugar (Fig. 26 A) contained mainly xylose (84 g L<sup>-1</sup>), plus smaller amounts of sucrose (13 g L<sup>-1</sup>), fructose (9 g L<sup>-1</sup>), and glucose (7 g L<sup>-1</sup>). During the first 20 h, cell and lysine concentration increased steeply, while the sugar was rapidly utilized (Fig. 26 A). The maximum specific growth rate during this phase was 0.09 h<sup>-1</sup>. Interestingly, *C. glutamicum* co-consumed all available sugars, with preferential metabolization of the hexoses. As soon as the total sugar level remained around 30 g L<sup>-1</sup>. In the following 20 hours, the cell concentration increased to a value that just compensated dilution of the medium by the addition of feed and base. However, lysine accumulated further during the entire process. After about 120 hours of cultivation, the feed was stopped to allow metabolization of the residual sugar. The final titer of lysine was 116 g L<sup>-1</sup>.

Dissolved oxygen (DO) was available in sufficient amounts throughout the whole process (Fig. 26 B). Following a steep decline of the oxygen saturation in the beginning of the cultivation, the lowest value of about 20 % coincided with the near depletion of the hexose sugars. After that, the DO increased almost steadily,



**Fig. 26:** Lysine production with *C. glutamicum* Xyl1 in a fed-batch process on xylose-based medium containing industry-grade molasses. The dashed vertical line marks the start and the end of the feed phase. The data points are means and deviations from two independent experiments.

indicating a reduced activity in the oxygen-requiring biochemical reactions like the energy metabolism. Ammonia concentration also remained at non-limiting levels at least until the end of the feeding phase (Fig. 26 B). The  $K_m$  value of the ammonia transporter proteins in *C. glutamicum* is lower than 50  $\mu$ M [137, 191]. Hence, the concentrations present during the cultivation allowed for sufficient uptake.

Both, specific substrate uptake rate and product formation rate peaked during the



**Fig. 27:** Lysine yield of *C. glutamicum* Xyl1 from fed-batch production on xylose-based medium. The consumed substrate represents total sugar. The data are means and deviations from two independent experiments.

initial batch phase and reached values of 0.17 g g<sup>-1</sup> h<sup>-1</sup> and 0.05 g g<sup>-1</sup> h<sup>-1</sup> respectively (Fig. 26 C). The space-time yield reached a maximum of 1.5 g L<sup>-1</sup> h<sup>-1</sup> after 30 hours. During the first 90 hours, the lysine yield was constant at 0.38 g g<sup>-1</sup> (0.39 C-mol C-mol<sup>-1</sup> based on all sugars) and then dropped slightly, while the overall yield was 0.36 g g<sup>-1</sup> (Fig. 27).

For almost 40 years, lysine has been produced through fermentation from hexose sugars. Throughout the years, the highly competitive lysine market has been a major driver for continuous breeding of the major producer *C. glutamicum* towards enhanced key performance indicators titer, yield and productivity [16]. Global pathway engineering (systems metabolic engineering) has delivered optimized strains for lysine production from traditional hexose-based substrates like starch, raw sugar and molasses. Such globally engineered producers reach lysine titers of 120 g L<sup>-1</sup> within 30 h of cultivation and carbon conversion yields of 0.55 g g<sup>-1</sup> [18].

In the developing era of bio-based economy, strain engineering for lysine production now expands to the utilization of alternative raw materials, particularly non-food biomass such as lignocellulose and hemicellulose, most abundant in forestry and agricultural residues [3] or in waste streams of the pulp and paper industry [130]. The application of systems metabolic engineering strategies to optimize *C. glutami*-

Table 10: Performance of C. glutamicum strains (i	in [190] the closely related <i>Brevibacterium</i>
lactofermentum) for industrial lysine production.	. All studies used fed-batch processes for
production.	

Major carbon source	Complex ingredient	Titer [g L <sup>-1</sup> ]	Ref.
Glucose	Molasses	120	[18]
Glucose	Molasses	105	[190]
Glucose	-	90	[80]
Glucose	-	80	[150]
Xylose	Molasses	116	this work

*cum* for lysine production from a xylose-based medium proved very successful as the novel strain *C. glutamicum* Xyl1 was capable of reaching a lysine titer of 116 g L<sup>-1</sup> in a fed-batch process. Hereby, medium composition reflected that of large-scale industrial cultivations, which often include cheap components such as molasses and corn steep liquor [68]. The main carbon source was xylose, and the supplemented molasses provided additional nutrients. Nitrogen for growth and lysine production originated from corn steep liquor and ammonium.

So far, lysine titers above 100 g L<sup>-1</sup> have been only reported for production from glucose based media (Tab. 10), the preferred carbon source for *C. glutamicum* and also most other microorganisms [80], so that the titer reached here keeps up with reported benchmarks. Similarly, product yields of up to 0.38 g g<sup>-1</sup> (Fig. 27) lie well in the range of attractive production yields reported in the literature [72]. Together with recent studies, which demonstrate robustness of *Corynebacteria* against inhibitors from hemicellulose hydrolysates [59] and access to an interesting product portfolio from such raw materials [36, 138], this development sets an important benchmark for hemicellulose biotechnology of *C. glutamicum*.

## 4.2.3 The engineered metabolism of *C. glutamicum* Xyl1 is tailored for lysine production from xylose

Previous metabolic engineering of *C. glutamicum* for xylose assimilation focused on local implementation of the xylose degradation pathway [97]. This probably explains the lower performance of the resulting producers [138, 220], which did not achieve the level required for industrial production. For lysine, none of the previously developed strains achieved economically attractive yields on xylose [138]. However, yield is a major criterion for the economic success of bulk products such as lysine, for which the raw material is a major cost driver [41]. *C. glutamicum* strain Xyl1 exhibits an overall yield of 0.39 C-mol C-mol<sup>-1</sup>. This value is 2.5 fold higher than that obtained previously (Fig. 29). It is 52% of the theoretical optimum for the production of lysine from xylose (0.75 C-mol C-mol<sup>-1</sup>, 0.625 mol mol<sup>-1</sup>), predicted by elementary flux mode analysis [34] and approaches the performance of top lysine producers on conventional glucose-based raw materials (Fig. 29). The key to this high performance is previous systems wide metabolic pathway engineering (Fig. 28) with a set of 14 modifications. Hereby, the performance benefits from the engineered core metabolism, specifically tailored for xylose-based through fluxome and transcriptome based strain engineering [34]. Likewise, this combination of targets enables high-yield production of *C. glutamicum* Xyl1 is well suited for hyper-production of different chemicals from xylose-based feedstocks.

### 4.2.4 Slower kinetics for lysine production from xylose as compared to glucose leaves space for further strain engineering

Xylose is less efficient than glucose with regard to space-time yield, i.e. productivity. The maximum value of 1.5  $g_{lysine} L^{-1} h^{-1}$  achieved (Fig. 26 C) is less than half of the value of 4.0 g L<sup>-1</sup> h<sup>-1</sup> for production from glucose [18]. Without doubt, xylose seems still attractive for production, considering the low price and high availability of residues and waste materials containing xylose. However, this leaves space for further improvement. *C. glutamicum* Xyl1 metabolizes xylose and glucose with similar specific rates (Tab. 9), which indicates successful introduction of the xylose operon. However, the influx of xylose in terms of carbon is only 5/6 of the value for glucose, due to the smaller number of carbon atoms in the pentose sugar. Additionally, growth on xylose results in lower growth rate and biomass yield, as compared to glucose (Fig. 24, Fig. 25, Tab. 9).f Therefore, *C. glutamicum* Xyl1 could benefit from further engineering of xylose utilization. Notably, the overexpression of the *Xanthomonas campestris* xylose isomerase and of the *C. glutamicum* xylulokinase



**Fig. 28:** Tailoring *C. glutamicum* Xyl1 for efficient lysine production from the hemicellulose sugar xylose by systems-wide pathway engineering. Bold green arrows indicate reactions with enhanced gene expression, bold dashed red arrows mark the reactions designed towards reduced carbon flux. The modifications at the sites numbered 1 to 10 have been implemented in previous work [11, 13, 18, 150] as mentioned in the chapter 2.1.6. The modifications at the site 11 allow for xylose utilization in *C. glutamicum* Xyl1 presented in this work. 1:  $P_{sod}lysC$ , G1A, C932T; 2:  $2 \times ddh$ ; 3:  $\Delta pck$ ; 4:  $P_{sod}dapB$ ; 5:  $2 \times lysA$ ; 6: hom, T176C; 7:  $P_{sod}pyc$ , C1372T; 8: *icd*, A1G; 9:  $P_{tuf}fbp$ ; 10:  $P_{sod}tkt$ , *tal*, *zwf*, *opcA*, *pgl*; 11: pClik 5a MCS  $P_{gro}xylAB$ . Gene names (red italics) are listed on page vi.



**Fig. 29:** Performance of lysine and diaminopentane producing *C. glutamicum* strains using xylose (blue, solid bars) and glucose (green, hatched bar) as the main carbon source. The yields given in darker colors represent experimentally achieved values. The yields given in light colors represent the theoretical maxima, as deduced from *in silico* modeling by elementary flux mode analysis. a: this work; b: estimated from reference [138]; c and d: taken from reference [18] and [34], respectively.

provides improved growth rate and biomass yield in recombinant *C. glutamicum* strains [138] and thus appears to be an interesting additional target for *C. glutamicum* Xyl1. Indeed, the created producer possesses endogenous activity of xylulokinase which could be enhanced by up-stream insertion of a strong promoter routinely available for *C. glutamicum* [10].

Taken together, this work demonstrates high-titer production of lysine production from a xylose-based industrial medium with *C. glutamicum* Xyl1, rationally engineered for this purpose. The excellent performance of the recombinant *C. glutamicum* strain on xylose, naturally not accessible by the bacterium, further opens the door to hemicellulose raw materials and an even wider portfolio of processes of this industrial workhorse beyond classical hexose-based biotechnology.

# 4.3 Ectoine production by metabolically engineered *C. glutamicum*

#### 4.3.1 Re-design of *C. glutamicum* Lys1 for ectoine bio-synthesis

The metabolic production routes for lysine and ectoine both share the metabolite aspartate- $\beta$ -semialdehyde (ASA, Fig. 7 A and 9). The enzyme aspartokinase, generating its precursor, aspararatyl-phosphate, is under tight feedback control by threonine and lysine in *C. glutamicum* ATCC 13032 [92]. Hence, for heterologous synthesis of ectoine in *C. glutamicum*, the basic lysine-producer *C. glutamicum* Lys1 was chosen as a suitable genetic background as it possesses a feedback-resistant aspartokinase [18] and thereby circumvents the native biochemical pathway regulation for the synthesis of ASA. Thus, an adequate supply of the precursor molecule for ectoine biosynthesis (Fig. 9) was aimed for with the chosen starter strain *C. glutamicum* Lys1.

For the recombinant DNA experiments an ectABCD gene cluster from an isolate of Pseudomonas stutzeri [240], a well-known ectoine and hydroxyectoine producer [186, 198], was chosen. To adjust the codon usage of the ectoine gene cluster employed by the natural host strain P. stutzeri A1501 to that preferred by C. glutamicum, the ectABCD genes were chemically synthesized and codon-optimized (sequence see Fig. 36 in the appendix). The expression of the naturally osmotically inducible operon [198] was uncoupled from osmotic-stress-derived signal transduction processes by positioning the synthetic ectABCD gene cluster under control of a strong and constitutive C. glutamicum promoter that is driving the expression of the structural gene (tuf) for the elongation factor Tu of C. glutamicum [13]. The resulting P<sub>tuf</sub>ectABCD construct was additionally provided at its 5'- and 3'-ends with regions flanking the non-essential diaminopimelate dehydrogenase gene (ddh) for targeted integration into the chromosome of *C. glutamicum* by a double-homologous recombination event (Fig. 11). The *ddh* gene was chosen as integration site as it represents the more active one of the two metabolic branches for the synthesis of lysine (Fig. 7) during high availability of ammonium, which is usually the case during production processes at industrial scale [48, 182]. The intention was to lower the carbon flux via this pathway towards lysine to a priori diminish lysine formation as a major competitor for the building block aspartate semialdehyde (Fig. 7 A and 9).

Successful integration and inactivation was verified by PCR analysis (Fig. 30, left





**Fig. 30:** Agarose electrophoresis gels. L = DNA base pair ladder, 1 = Synthetic *ectABCD* gene cluster amplified from genomic DNA from *C. glutamicum* Ect1 [27], 2 = negative control, 3 = lysE in *C. glutamicum* Ect1, 4 = lysE in *C. glutamicum* Ect2 [65].

**Fig. 31:** Specific diaminopimelate dehydrogenase activity in the strains *C. glutamicum* Lys1 and Ect1.

gel) and determination of the Ddh activity (Fig. 31). The first showed the integration of the 3448 bp long gene fragment (see Fig. 11) in the genome of *C. glutamicum* Ect1 and the latter the absence of diaminopimelate dehydrogenase activity in the strain due to the interruption of the *ddh* gene by said DNA fragment. In this way, a recombinant strain *C. glutamicum* Ect1 was constructed that carried a single-copy of the  $P_{tuf}ectABCD$  gene cluster stably integrated into a well-defined site in the chromosome and that should be able to produce ectoine/hydroxyectoine in the absence of osmotic stress.

## 4.3.2 *C. glutamicum* Ect1 efficiently produces ectoine from glucose

Evaluation of the pattern and levels of ectoine/hydroxyectoine production in the newly constructed *C. glutamicum* Ect1 strain was conducted by cultivation in a chemically defined minimal medium with glucose as the carbon source. The experiment revealed that the  $P_{tuf}ectABCD$  gene cluster was functionally expressed and that the Ect1 strain produced and secreted ectoine in shake-flask cultures already from early on (Fig. 32). Ectoine accumulated in the culture supernatant up to a final titer of



**Fig. 32:** Cultivation profile of the heterologous ectoine producer strain *C. glutamicum* Ect1. The *C. glutamicum* strain Ect1 was cultivated in shake flasks at 30 °C in a chemically defined medium. At the indicated time intervals, consumption of glucose and the extracellular accumulation of L-lysine, ectoine, and 5-hydroxyectoine were monitored. The data shown represent mean values and corresponding standard deviations from three biological replicates.

1.3 mM, after the cells had completely depleted carbon source. About 2 % of the consumed glucose was converted into ectoine as reflected by the molar yield of 19.4 mmol mol<sup>-1</sup> (Tab. 11).

Hydroxyectoine was also detected in the culture supernatant, but at a rather low concentration (< 0.1 mM). In addition to both compatible solutes, *C. glutamicum* Ect1 secreted substantial amounts of L-lysine into the medium, 2.3 mM in total. L-lysine production thus exceeded that of the desired products ectoine and hydroxy-ectoine. Quantification of the intracellular levels of ectoine and hydroxyectoine in *C. glutamicum* Ect1 revealed that ectoine accumulated up to 130  $\mu$ mol g<sub>cdm</sub><sup>-1</sup> in the cytosol. Hydroxyectoine was detected in significantly lower amounts (ca. 5  $\mu$ mol g<sub>cdm</sub><sup>-1</sup>). Hence, relative amounts of these two compounds inside the cells corresponded roughly to the production titers (Fig. 32).

#### 4.3.3 Increased temperature positively affects production

The optimal cultivation temperature for the recombinant production of ectoine and hydroxyectoine by *C. glutamicum* Ect1 was assessed by miniaturized cultivations in a temperature range between  $27 \,^{\circ}$ C and  $42 \,^{\circ}$ C. Interestingly, this revealed that



**Fig. 33:** Influence of cultivation temperature on the growth and ectoine production performance of *C. glutamicum* Ect1. Strain Ect1 was grown in chemically defined medium with glucose on a miniaturized scale at the indicated growth temperatures. The specific growth rate  $\mu_{max}$ , ectoine secretion, and intracellular accumulation of ectoine and hydroxyectoine were determined. Ectoines were quantified after 10 hours (27 °C– 35 °C) and 20 hours (42 °C) of cultivation. The data shown represent mean values and corresponding standard deviations from three biological replicates.

ectoine production was improved by increased temperature (Fig. 33). As compared to the reference cultivation conditions for *C. glutamicum* Ect1 at 30 °C, secretion was more than doubled when the temperature was set to 35 °C. The enhanced production performance was also reflected by a slight increase of the intracellular ectoine level. The higher cultivation temperature also positively influenced the intracellular amounts of hydroxyectoine, and the growth performance of strain Ect1 as reflected by a 28 % increase of the specific growth rate (Fig. 33). The higher ectoine concentration in the supernatant was taken as positive indication for a better production performance at 35 °C.

**Table 11:** Growth and production performance of *C. glutamicum* strains Lys1, Ect1 and Ect2 during batch cultivation on a mineral salt medium with glucose as carbon source at 30 °C (Lys1, Ect1 & Ect2) and 35 °C (Ect2). The data represent mean values and standard deviations from three biological replicates and denote specific rates for growth ( $\mu_{max}$ ), substrate uptake (q<sub>S</sub>) and product formation (q<sub>Ect</sub> and q<sub>Lys</sub>). Additionally, yields for biomass (Y<sub>X/S</sub>), ectoine (Y<sub>Ect/S</sub>), lysine (Y<sub>Lys/S</sub>), trehalose (Y<sub>Tre/S</sub>),  $\alpha$ -ketoglutarate (Y<sub>AKG/S</sub>) and hydroxyectoine (Y<sub>EctOH/S</sub>) are given. Data for *C. glutamicum* Lys1 are taken from previous work [18].

C. glutamicum strain	Lys1	Ect1	Ect2	Ect2
Temperature	30 ° C	30 ° C	30 °C	35 °C
$\mu_{max}$ [h <sup>-1</sup> ]	$0.38 \pm 0.01$	$0.38 \pm 0.01$	$0.38\pm 0.01$	$0.38\pm 0.01$
$q_{S}$ [mmol g <sup>-1</sup> h <sup>-1</sup> ]	$\textbf{4.86} \pm 0.10$	$3.82 \pm 0.08$	$3.51 \pm 0.05$	$3.74 \pm 0.10$
q <sub>Ect</sub> [mmol g <sup>-1</sup> h <sup>-1</sup> ]	n/a	$0.07 \pm 0.01$	$0.09\pm0.01$	$0.12\pm 0.01$
q <sub>Lys</sub> [mmol g <sup>-1</sup> h <sup>-1</sup> ]	$0.39 \pm 0.02$	$0.20\pm0.01$	$0.00\pm0.00$	$0.00\pm0.00$
$Y_{X/S}$ [g mol <sup>-1</sup> ]	$82.1 \pm 1.3$	$87.2 \pm 2.8$	$101.5\pm0.6$	$90.9 \pm 6.3$
Y <sub>Ect/S</sub> [mmol mol <sup>-1</sup> ]	n/a	$19.4 \pm 1.5$	$24.6 \pm 0.6$	$32.0\pm 0.8$
Y <sub>Lys/S</sub> [mmol mol <sup>-1</sup> ]	$81.2 \pm 3.2$	$53.8 \pm 2.5$	$0.0\pm0.0$	$0.0\pm0.0$
Y <sub>Tre/S</sub> [mmol mol <sup>-1</sup> ]	$9.4 \pm 0.4$	$5.6 \pm 0.1$	$7.2 \pm 0.9$	$6.7 \pm 0.2$
Y <sub>AKG/S</sub> [mmol mol <sup>-1</sup> ]	$0.0\pm0.0$	$2.6\pm 0.4$	$1.0\ \pm\ 0.1$	$1.9\pm0.1$
Y <sub>EctOH/S</sub> [mmol mol <sup>-1</sup> ]	n/a	$0.6\pm0.0$	$0.6\pm0.0$	$0.9\pm0.1$

### 4.3.4 Elimination of lysine secretion generates a second generation ectoine producer

Ectoine production of the first-generation producer Ect1 was limited by substantial carbon loss related to L-lysine synthesis and subsequent secretion of this amino acid into the growth medium. To avoid excretion of L-lysine, the gene (*lysE*) for the L-lysine exporter LysE [216] was inactivated in the genetic background of strain Ect1 to yield the second-generation ectoine producer *C. glutamicum* Ect2. Validation was carried out by PCR (Fig. 30, right gel), confirming that a 575 bp segment of the *lysE* gene was missing in *C. glutamicum* Ect2 compared to Ect1. First, the novel *C. glutamicum* Ect2 strain was cultivated in shake flasks in glucose minimal medium at 30 °C to allow a direct comparison of the performance with the parent Ect1 strain. The deletion of the *lysE* gene had a beneficial influence on ectoine production. As compared to the parent strain *C. glutamicum* Ect1, the molar ectoine yield was increased by 27 % (Tab. 11). At the same time, L-lysine secretion was completely eliminated. However, the additional carbon available for strain Ect2 was not completely channeled towards ectoine production; it was instead recruited for

biomass formation. To take benefit from the improved ectoine production at elevated growth temperature observed for strain Ect1 (Fig. 33), the performance of strain Ect2 at 35 °C was also investigated in greater detail.

Similarly to the *C. glutamicum* Ect1 strain, the higher growth temperature positively influenced the production performance of the strain Ect2. The yield was increased from 25 mmol mol<sup>-1</sup> at 30 °C to 32 mmol mol<sup>-1</sup> at 35 °C. When compared to the basic proof-of-concept approach, the allover ectoine yield was improved in strain Ect2 by 65 %. In addition, strain Ect2 did not suffer from high fluxes to the by-product L-lysine (Tab. 11). The specific growth rate of strain *C. glutamicum* Ect2 was hardly affected by the elevated cultivation temperature (Tab. 11); however, the yield in biomass was reduced, likely as a stress response of *C. glutamicum* to the increase in growth temperature. This finding is different from the results presented in chapter 4.1 where it was shown that in addition to reduced biomass yields, C. glutamicum strains Lys1, Lys12 and Lys12K also show lower specific growth rates. It is likely that ectoine synthesized by C. glutamicum Ect2 acts as a compatible solute and enables stable growth rates despite increased temperature. Taken together, the substantially improved ectoine yield and the unaffected glucose uptake rate of strain Ect2 resulted in an about 70% increased specific productivity of the second-generation ectoine producer C. glutamicum Ect2 (Tab. 11).

## 4.3.5 Ectoine synthesis affects the intracellular pools of amino acids of the aspartate family

Aspartate semialdehyde is an important metabolite and critical branch point with regard to biosynthesis of the aspartate family of amino acids [131]. Since ectoine biosynthesis is dependent on a good supply of ASA as well (Fig. 9) [21, 198], the intracellular pools of amino acids belonging to the aspartate family were examined. Most desirable, integration of the synthetic ectoine cluster into the Ddh lysine branch resulted in substantial decrease of the intracellular accumulation of lysine as major by-product in the Ect1 strain (Tab. 12). This was taken as positive indication of a successfully lowered carbon flux towards lysine biosynthesis, thus increasing the ASA availability for the novel product ectoine. Ectoine indeed accumulated in substantially higher levels in the cytosol than lysine did (Tab. 12). Upon elimination of the lysine exporter, however, the metabolite pattern completely changed. The novel

**Table 12:** Concentration of free intracellular amino acids of the aspartate family and of intracellular ectoine of the lysine producing strain *C. glutamicum* Lys1 and its ectoine producing derivatives *C. glutamicum* Ect1 and Ect2. Cells were grown at 30 °C and 35 °C in mineral salt medium. The data represent mean values with standard deviations from two biological replicates, each sampled at three different optical densities (OD 2, OD 4 and OD 8).

Temperature	Compound		C. glutamicur	n
	$[\mu mol g_{cdw}^{-1}]$	Lys1	Ect1	Ect2
30 °C	Aspartate Asparagine Threonine Lysine Ectoine	$\begin{array}{c} 11.8 \pm 3.3 \\ 0.7 \pm 0.3 \\ 5.1 \pm 0.9 \\ 23.4 \pm 3.1 \\ n/a \end{array}$	$\begin{array}{c} 8.1 \pm 2.9 \\ 24.6 \pm 6.0 \\ 3.9 \pm 0.6 \\ 14.5 \pm 4.1 \\ 126.8 \pm 25.5 \end{array}$	$\begin{array}{c} 7.0 \pm 0.3 \\ 20.1 \pm 4.9 \\ 3.0 \pm 0.9 \\ 76.7 \pm 11.3 \\ 34.1 \pm 14.2 \end{array}$
35 °C	Aspartate Asparagine Threonine Lysine Ectoine	$\begin{array}{c} 9.1 \pm 0.9 \\ 1.4 \pm 0.2 \\ 6.8 \pm 0.5 \\ 24.4 \pm 6.3 \\ n/a \end{array}$	$\begin{array}{c} 7.6 \pm 0.6 \\ 28.1 \pm 4.0 \\ 6.5 \pm 0.4 \\ 16.3 \pm 2.0 \\ 158.5 \pm 20.7 \end{array}$	$\begin{array}{c} 7.9 \pm 0.6 \\ 17.0 \pm 2.4 \\ 5.2 \pm 0.4 \\ 52.8 \pm 12.6 \\ 36.1 \pm 7.6 \end{array}$

strain *C. glutamicum* Ect2 accumulated up to  $77\mu$ mol  $g_{cdm}^{-1}$  of lysine even exceeding the value of the basic lysine producer *C. glutamicum* Lys1 more than three fold. Simultaneously, the intracellular ectoine level dropped substantially (Tab. 12). Obviously, elimination of lysine secretion did not circumvent high carbon fluxes towards this amino acid but resulted in increased intracellular accumulation. In addition to the most obvious changes regarding the intracellular lysine and ectoine level, also the formation of aspartate and threonine was slightly affected. As compared to the parent lysine producer *C. glutamicum* Lys1, the cytosolic concentrations of these two amino acids were slightly reduced in the synthetic ectoine cell factories. The intracellular asparagine level, however, was strongly increased. Whereas only marginal amounts were found in *C. glutamicum* Lys1, both Ect1 and Ect2 exhibited comparably high intracellular asparagine levels, also competing with ectoine biosynthesis and secretion. The cultivation temperature only had a marginal effect on the intracellular amino acid accumulation. For ectoine, a slight trend towards higher accumulation was observed at 35 °C as indicated by the miniaturized temperature screening.



**Fig. 34:** Production performance of the advanced ectoine-producer *C. glutamicum* Ect2 during fed-batch fermentation. During the batch phase, the oxygen saturation in the fermenter was kept constant at 30 % by variation of the stirrer velocity and the aeration rate. Automated feeding was initiated by a  $pO_2$ -based signal [18]. Glucose concentration was thereby kept below 5 g L<sup>-1</sup>. The data shown represent mean values from two independent fermentation experiments.

### 4.3.6 Performance of the ectoine producing strain *C. glutamicum* Ect2 under fed-batch conditions

To assess the overall production performance of *C. glutamicum* Ect2 under conditions more relevant for an industrial process, the ectoine production by this strain was benchmarked in a fed-batch process (Fig. 34). Ectoine was secreted from early on during growth and accumulated in the growth medium to a final concentration of 4.5 g L<sup>-1</sup> after 16 hours of fermentation. The production efficiency differed significantly between the batch and the feeding phase of the process (Fig. 35). The initial 45 g L<sup>-1</sup> glucose was already consumed after 8 hours and mainly served for the production of biomass (Fig. 34). This was reflected by the rapid increase of the cell dry weight to 25 g L<sup>-1</sup>. As soon as the feeding was started, a shift in the production pattern towards the formation of ectoine was observed. Overall, the feeding phase contributed to more than 80 % of the total ectoine production (Fig. 34). The yield obtained during the batch phase (28 mmol mol<sup>-1</sup>) was similar to the yield obtained during shake flask cultivation of strain Ect2 (Tab. 11). In the feeding phase, however, it increased 10-fold up to about 300 mmol mol<sup>-1</sup> (Fig. 35). This substantial increase provided an overall space time yield of 6.7 g L<sup>-1</sup> d<sup>-1</sup> ectoine.



**Fig. 35:** Ectoine yield achieved in the different cultivation phases during fed-batch fermentation (see Fig. 34) of the advanced ectoine-producer *C. glutamicum* Ect2. The data shown represent mean values from two independent fermentation experiments.

## 4.3.7 *C. glutamicum* Ect2 enables decoupling of ectoine production from high salinity in a marker-free host system

The stabilizing and function-preserving effects of ectoines have led to considerable interests in these compounds and hence to the development of a variety of practical applications [61, 124, 160]. As a consequence, ectoine and its derivative hydroxyectoine are considered as valuable natural microbial products and they have gained significant market value in recent years. Their biotechnological production has reached the scale of tons on an annual basis. The industrial production of ectoines is currently achieved by bacterial milking of *H. elongata* [175, 184], a process that has recently been improved by the inclusion of mutants of this bacterium that cannot catabolize ectoines and those that accumulate these compounds in the high salinity medium as a result of a defect in the ectoine/hydroxyectoine-specific TeaABC uptake system [62].

The construction and characterization of a synthetic microbial cell factory for the production of ectoine described here relies on the robust metabolism of *C. glutamicum* [35, 226]. This bacterium incorporates features desirable for large scale production

processes which are responsible for its tremendous rise and success as industrial production host [15]. One of those beneficial features is the considerable knowledge base for the genetic manipulation and large-scale fermentation processes of the industrial workhorse C. glutamicum [15]. Most advantageous, the C. glutamicum genome sequence lacks ectoine catabolic genes [93] eliminating product loss due to the reuse of ectoine and hydroxyectoine as carbon or nitrogen sources – a drawback of natural producer H. elongata [62]. In developing the synthetic C. glutamicum ectoine cell factory several strategies were combined simultaneously to optimize its production. Adaptation of the codon usage appeared promising as this strategy has proven beneficial for the heterologous production of diaminopentane recruiting lysine decarboxylase from E. coli [103]. Expression control via the constitutively active tuf promoter of C. glutamicum [13, 103] not only decoupled ectABCD expression from its natural osmo-stress responsive regulation [198], but also provided a simple and robust promoter system with no specialized needs for the control of its activity. Beyond this, the careful design of the cellular chassis guaranteed good supply of the building block ASA which is normally tightly controlled in many microorganisms including C. glutamicum [92, 131]. In addition, and in an effort to diminish carbon flux towards L-lysine, the dehydrogenase branch of L-lysine biosynthesis [39, 194] was inactivated.

In comparison to plasmid-based synthesis systems [172, 183, 186, 198] the strategy presented here benefits from a stable genome-based integration of the synthetic  $P_{tuf}ectABCD$  gene cluster (Fig. 11), an approach that has also been employed for the recombinant production of ectoines by *H. polymorpha* [50]. Constitutive gene expression in the synthetic *C. glutamicum* cell factory obviated the need for the use of high salinity growth media, conditions that are typically required to trigger enhanced expression of these genes in natural producers [33, 118, 119] such as *H. elongata* [184]. The novel synthetic cell factory *C. glutamicum* Ect2 thus avoids the considerable drawbacks of high salinity growth media during fermentation processes that invoke corrosion problems on the reactor systems [172] and thereby drive up the costs for their appropriate design and maintenance. Related to the economic relevance, also previous studies have addressed the issue of a salt decoupled ectoine production with the natural producers *P. stutzeri* and *Chromohalobacter salexigens*, and the heterologous hosts *H. polymorpha* and *E. coli* [50, 172, 183, 186].

## 4.3.8 Combined process and strain engineering increases ectoine production in *C. glutamicum*

As successful proof of concept, the first generation *C. glutamicum* Ect1 strain revealed salinity decoupled ectoine and hydroxyectoine production and secretion. Interestingly, the summed molar yields of L-lysine and ectoine added up to 73.2 mmol mol<sup>-1</sup>, a value that perfectly matches with the L-lysine yield of the parent strain *C. glutamicum* Lys1 [18]. Hence, the carbon demand for ectoine synthesis seems to be completely satisfied by the cells on the expense of L-lysine. Still, L-lysine production dominated in the basic ectoine producer. This issue was addressed in a second round of metabolic engineering through inactivation of the gene coding for the L-lysine exporter LysE. This strategy has been applied for optimized diaminopentane production by *C. glutamicum* [105]. Beyond the 27 % gain in ectoine yield (Tab. 11), the inactivation of the lysine exporter avoided a contamination of the excreted ectoine, a feature that will clearly facilitate the downstream processing for the recovery of ectoine from the culture broth.

Further improved ectoine production was then achieved by an increase in the growth temperature from 30 °C to 35 °C. In combination, genetic and process engineering proved to be highly beneficial as reflected in a 65% increased yield and a 70% increased specific productivity (Tab. 11). A rather unexpected finding of our study was the observation that the recombinant strain produced so little hydroxyectoine (Tab. 11). This is surprising because P. stutzeri A1501, on whose genetic blueprint the synthetic *ectABCD* gene cluster is based, produces hydroxyectoine very efficiently and in much greater quantities than ectoine [186, 198]. It is currently unclear why there is so little hydroxyectoine formed by the *C. glutamicum* P<sub>tuf</sub>ectABCD strains. Overall, both recombinant C. glutamicum ectoine producers already performed admirably in shake-flask experiments. They exhibited excellent specific production rates of 9.9 mg g<sup>-1</sup> h<sup>-1</sup> (*C. glutamicum* Ect1) and 17.1 mg g<sup>-1</sup> h<sup>-1</sup> (*C. glutamicum* Ect2), thus exceeding that of the native and the heterologous ectoine production hosts *H. elongata* (7.1 mg g<sup>-1</sup> h<sup>-1</sup>) [175] and *E. coli* (2 mg g<sup>-1</sup> h<sup>-1</sup>) [183]. However, a closer examination of the achieved final yield of 32 mmol mol<sup>-1</sup> revealed that it was so far not possible to harness the complete carbon used by C. glutamicum Lys1 for L-lysine synthesis for the recombinant production of ectoine.

## 4.3.9 High-cell density fermentation reveals excellent production performance

To take benefit from high cell densities [18, 154], which cannot readily be achieved in shake-flask experiments, the performance of the second generation ectoine producer *C. glutamicum* Ect2 was benchmarked under carefully controlled fed-batch conditions. The ectoine production efficiency differed significantly during the batch and the feeding phase of the fermentation process. That kind of desired shift in the product spectrum is often intentionally induced at industrial scale production by appropriate process operations. In general, this leads to an optimized channeling of the substrate to the desired product within the feeding-phase [18, 34, 73, 123]. Similarly, a significant 10-fold increase of the ectoine yield in the feeding phase of the *C. glutamicum* Ect2 fermentation (Fig. 35) could be observed. Though there was detectable growth-associated ectoine production during the batch phase, the larger fraction of ectoine seemed to be produced by cells during the feeding phase. The achieved final titer of 4.5 g L<sup>-1</sup> already approached to that of currently described industrial production systems [53, 121, 212].

The excellent performance of the *C. glutamicum* strain Ect2 under fed-batch conditions allowed the synthetic cell factory to achieve an overall space-time yield of 6.7 g L<sup>-1</sup> d<sup>-1</sup> ectoine which is among the highest productivities reported so far in the literature [53, 121, 212]. Better performance was, so far, only achieved with *H. boliviensis* [212], and *Chromohalobacter salexigens* [53]. These do, however, rely on high salinity and involve complex process operation strategies [53] thus driving up the production costs.

### 4.3.10 Metabolic changes point at potential bottlenecks in ectoine production

When introducing the synthetic ectoine gene cluster in lysine producing *C. gluta-micum* Lys1, carbon flux within the cell was partly re-routed from the common intermediate ASA to ectoine as indicated by the drop in extracellular and intracellular lysine ([18], Tab. 12), an effect equally observed at 30 °C and 35 °C. However, elimination of lysine secretion resulted in a strong increase of the intracellular pool of this amino acid. The additional carbon, now potentially elevating the ASA pool,

was not efficiently recruited for driving ectoine biosynthesis. Obviously, the native lysine route is still superior to the non-native route towards the novel product ectoine despite the elimination of the dehydrogenase branch. In subsequent rounds of strain engineering, this might be targeted by additional down-regulation of other lysine biosynthetic genes including dihydrodipicolinate synthase (*dapB*) and diaminopimelate decarboxylase (*lysA*) previously identified as key reactions for lysine production [18].

In addition, manipulating the strength of the ribosome-binding sites of the various ect genes might improve their expression. An additional issue involves the transport processes related to the ectoine metabolism. Most favorable for production processes, C. glutamicum excretes ectoine into the growth medium, a feature that was also observed for other recombinant ectoine production systems [50, 183]. So far the intracellular metabolite levels (Tab. 12) do not suggest limitations in the transport processes but experience has shown that engineering the excretion systems for amino acids and related compounds is beneficial for the performance of C. glutamicum [16, 19, 31, 81, 105, 216]. In addition, it appears even more important to address the uptake system of C. glutamicum comprising the three transporters EctP, LcoP, and ProP [224]. In native producers, these transport systems not only serve for the scavenging of stress protectants from natural resources [225], but they also function as salvage systems to retrieve compatible solutes that leak or are actively released from the producer cells [62, 63, 70]. As previously shown for the production of aromatic amino acids [75, 76], the production efficiency of C. glutamicum might significantly suffer from product re-uptake into the cell.

In addition, there is strong indication that the compatible solute transporter TeaABC in *H. elongata* is integrated in the regulation of ectoine biosynthesis. *De novo* biosynthesis of the cell is immediately decreased when externally supplied ectoine is taken up via TeaABC [62]. As the regulatory mechanism is unknown so far, similar regulation patterns involving the *C. glutamicum* ectoine transporters might also limit the biosynthetic efficiency in the recombinant production host.

In contrast to the currently used *H. elongata* strain for the commercial production of ectoines [124, 184] or other microbial species that have been suggested as alternative natural producers [121, 178, 186, 198, 212], *C. glutamicum* has a long history for the industry level production of valuable natural products by large-scale fermentation procedures [68, 81]. Knowledge gained during the development of *C. glutamicum* into an excellent performing microbial cell factory [17, 18, 34] can now be brought

to bear in a scale-up process for ectoine production by the recombinant *C. glutamicum* Ect2 strain. In addition, the ability to readily manipulate the genetic blueprint of *C. glutamicum* on a genome-wide scale [18, 203, 208] will allow the application of rational and systems-wide metabolic engineering approaches to further improve the performance of the second generation ectoine producer *C. glutamicum* Ect2. The success of these approaches has recently unequivocally been demonstrated in the development of highly efficient L-lysine [18] and diaminopentane [34, 103, 104] *C. glutamicum* cell factories.

#### CHAPTER 5

### **CONCLUSION AND OUTLOOK**

Biotechnological production of chemicals will play an increasingly important role in the global economy of the future to replace processes based on petrochemistry [42]. In order to reach this goal, scientists and engineers have to integrate the entire spectrum of biotechnological methods, ranging from molecular biology to biochemical engineering and including analytical tools to gain an insight into intraand extracellular processes to precisely engineer them for the production of value added chemicals.

C. glutamicum, a classical amino acid producer is a proven robust and versatile production strain for a wide range of chemicals [15]. The results presented in this work make use of the useful traits of C. glutamicum and show further improvements in the substrate and product portfolio. Heterologous expression of genes from Escherichia coli and Pseudomonas stutzeri allowed the utilization of xylose as a substrate and the secretion of the compatible solute ectoine as a product, respectively. The latter modification also included optimization of the genetic sequence for the codon usage in *C. glutamicum* and utilization of a constitutive promoter. This allowed the production of ectoine without the need for media of high salinity as the operon, comprising the genes for ectoin production, is naturally under control of a promoter triggered by osmotic stress. Thus the product portfolio of C. glutamicum could be widened by a substance of interest for the cosmetic industry, while at the same time making a simpler biotechnological production route for ectoine possible, which is less destructive to the equipment than the bacterial milking processes applied for natural producers [175, 184]. Further measures to increase the ectoin titer could focus on the attenuation of lysine production to avoid the intracellular accumulation of this amino acid. Additionally, genetic modifications that have proven useful for

increasing lysine production [18, 150] could be implemented in *C. glutamicum* Ect2. To focus purely on ectoine production, the gene cluster could be shortened by the *ectD* gene to avoid hydroxyectoine co-production.

The implementation of the *xylA* and *xylB* genes from *E. coli* enabled conversion of xylose to lysine, and integrated into an engineered core metabolism, enables lysine titers above 100 g L<sup>-1</sup> [138]. This is an important aspect for the industrial applicability, for which final titer, yield and productivity are of utmost importance [41]. The significance of xylose utilization for the production of valuable chemicals stems from the need for alternative substrates for microbial processes. Those alternatives must meet standards like non-competition with human nutrition. Xylose meets this requirement and, a major building block of hemicellulose, has the benefit of being globally available as a waste product in agriculture [3, 6, 202]. Further improvements could be made to equip *C. glutamicum* Xyl1 with the genetic tools to utilize arabinose, another pentose sugar present in hemicellulose, i.e. by transforming *E. coli araBAD* genes. This approach has been successfully applied in *C. glutamicum* before [96, 179]. *C. glutamicum* Xyl1 could thus use hemicellulose hydrolysates more efficiently for production of lysine [36].

In addition to the successful implementation on the molecular level, the effect of temperature, an important factor in biotechnological production was investigated. It was shown for a wide range of *C. glutamicum* strains, how cultivation at increased temperature increases lysine yield. Through <sup>13</sup>C flux analysis it was possible to characterize the effect of the higher temperature on the metabolism of *C. glutamicum*. This revealed a higher flux through the PP pathway and an increased net flux from the phosphoenolpyruvate/pyruvate pool to the oxaloacetate/malate pool.

Interestingly, the effect of temperature, described above, was also observed for *C. glutamicum* strains producing ectoine. It did, on the other hand, not occur during lysine production on xylose. This underlines the importance of systems wide knowledge about the production host, as such phenomena cannot be easily understood. The increase in PP pathway carbon flux, observed during growth on glucose at elevated temperature, hence might not take place on xylose.

To reach the aim of an economically feasible biotechnological production of chemicals, continuous improvement of existing systems and innovation for new routes is necessary. The results presented in this work give clues on how to proceed. For an improved lysine production from xylose, the carbon flux through the oxidative part of the PP pathway should be increased via molecular modification or maybe co-feeding of a hexose sugar. The findings of the metabolic flux analysis at elevated temperature suggest the PP pathway and the PEP/Pyr-OAA/Mal interconversion as promising targets to further advance lysine production. Finally it becomes apparent from the economic considerations presented in chapter 4.1.4 how important the definition of a goal is for any attempt to optimization. In this example, the term 'improved lysine production' could mean either higher yields, higher titers or higher productivity, but either of those possible goals requires a different approach.

### **Appendix**

### Ectoine synthese cluster nucleotide sequence

	(1)	1		10		20			30				47
EctSC_codon_optimized	(1)	AGCGI	TATTC	TCTG	CAAC	TAGT	GCAT	GTTG	GATG	CAAT	GGT	TGCAG	GCG
EctSC_native	(1)	AGCGI	TATTC	TCTG	CAAC	TAGT	GCAT	GTTG	GATG	CAAT	GGT	TGCAG	GCG
												<ul> <li>Sectio</li> </ul>	n 2
	(48)	48		60	)		70		8	)			94
EctSC codon optimized	(48)	CCACT	GAGC	ATCT	rgggi	AACC	TCAT	GCAT	GAGC	CGCA.	ACA	CCATO	TG
EctSC_native	(48)	CCACI	GAGC	ATCT	rgggi	AACC	TCAT	GCAT	GAGC	CGCA	ACA	CCATO	CTG
	. ,											- Sectio	n 3
	(95)	95	100		,110			120		130			141
EctSC codon optimized	(95)	CCCAC	CCGTT	GGAA	TAGC	CAAC	AATA.	AAGA	TCCT	CTTG.	ATG	CCAT	ACG
EctSC_native	(95)	CCCAC	CGTT	GGAA	TAGC	CAAC	AATA.	AAGA	TCCT	CTTG.	ATG	CCAT	ACG
	. ,											- Sectio	n 4
	(142)	142	,19	50		160		1	70				188
EctSC codon optimized	(142)	TGTTO	GCCCA	AGTG	CGTG	GCGA	GTTT	TACA	AAGA	ACCC	CAC	ATCAT	CA
EctSC_native	(142)	TGTTO	GCCCA	AGTG	CGTG	GCGA	GTTT	TACA	AAGA	ACCC	CAC	ATCAT	CA
	. ,											- Sectio	n 5
	(189)	189		200		2	210		220	)			235
EctSC_codon_optimized	(189)	ATGCO	СТААА	TGGC	GGGT	ATTT	TCAT	CCAA	ACCC	AACC	GCG	CATCA	\TT
EctSC_native	(189)	ATGCO	TAAA	TGGC	GGGT	ATTT	TCAT	CCAA	ACCC	AACC	GCG	CATCA	ΔTT
	. ,											- Sectio	n 6
	(236)	236			250		26	50		270			282
EctSC_codon_optimized	(236)	CCAAT	GCTG	ATCC	ACCC	CATC	CGGA	ТААА	CCAC	CATG	AAC	GGCA	ACG
EctSC_native	(236)	CCAAT	GCTG	ATCC	ACCC	CATC	CGGA	ТААА	CCAC	CATG	AAC	GGCA	ACG
												- Sectio	n 7
	(283)	283	,290	)	ŝ	300		310	)				329
EctSC_codon_optimized	(283)	GATCA	AAAG	TCCT	GTTG	GTGA	AGCT	GCGC	CCCA	CAGA	TCC	TGACI	GC
EctSC_native	(283)	GATCA	AAAG	TCCT	GTTG	GTGA	AGCT	GCGC	CCCA	CAGA	TCC	TGACI	GC
												- Sectio	n 8
	(330)	330		340		,35	0		360				376
EctSC codon optimized	(330)	TGGGA	AGCCA	TGAA	AATA	GATC	AGCG	CATC	CGTG	GTGG.	AAC	CAAAA	٩GG
EctSC_native	(330)	TGGG	AGCCA	TGAA	AATA	GATC	AGCG	CATC	CGTG	GTGG.	AAC	CAAAA	٩GG
_	、 ,											- Sectio	n 9
	(377)	377		2	390		400			410			423
EctSC_codon_optimized	(377)	CTCAR	CAAT	ACGA	AACG	TTCG	CTTT	CGGT	CCTG	ATGA	AAG	AGATO	TC
EctSC_native	(377)	CTCAF	CAAT	ACGA	AACG	TTCG	CTTT	CGGT	CCTG	ATGA	AAG	AGATO	Этс

**Fig. 36:** Nucleotide sequence of the ectoine gene cluster. The upper rows represent the sequence optimized for gene expression in *C. glutamicum*. The lower rows show the native *Pseudomonas stutzeri* sequence. (*The sequence is continued on pages 97 to 101*)

#### Appendix

	(424)	424 430		440	450	460	470
EctSC_codon_optimized	(424)	CCTGAATC	ATCATCT	AAGTATGO	CATCTCGGT	TAAGCTCGAC	CAGGACA
EctSC_native	(424)	CCTGAATC	ATCATCT	AAGTATGC	CATCTCGGI	TAAGCTCGAC	CAGGACA
							- Section 11
	(471)	471	480	490	50	0	517
EctSC_codon_optimized	(4/1)	GTGCCACC	ACAATTTT.	I'GGAGGA'I	TACAAGAA	ACATGACCAA	CATCCGC
EctSC_native	(4/1)	GIGCCACC	ACAATTT	IGGAGGAI	TACAAGAA	ACATGACCAA	Section 12
	(518)	518	530	5	40	550	564 Section
EctSC codon ontimized	(518)	GTAGCTAT	CGTGGGC		GTACCTG	- 500 GACGCAGCG	TCGAAAA
EctSC native	(518)	GTAGCTAT	CGTGGGC	FATCTAGA	GTACCTGO	GACGCAGCG	TCGAAAA
	. ,						- Section 13
	(565)	565 570	1	580	590	600	611
EctSC_codon_optimized	(565)	GTGGCCGT	TACCCTG	CGAATGTC	CACAGGG	TAGCTGGTAG	GTTTGAAA
EctSC_native	(565)	GTGGCCGT	TACCCTG	CGAATGTC	CACAGGG	TAGCTGGTAG	STTTGAAA
	(6.1.0)	010			0.10		- Section 14
FatCC and an antimized	(612)	612	620 aammaaaa	630	640		658
EctSC_codon_optimized	(612)	ATCAACGC	CGTTGCCC	TTAGGA1	TCAGTAAC	TGGCACATI	TTGTAAT
	(012)	AICAACGC	69119666	JIAGGAI	ICAGIAAC	LIGGCACAII	- Section 15
	(659)	659	670	680	0	690	705
EctSC codon optimized	(659)	GCGCTAGA	TCTGTGT	GCTCAGTC	TTCCAGG	TGCTTATCA	CAGTGAA
EctSC_native	(659)	GCGCTAGA	TCTGTGT	GCTCAGTC	CTTCCAGGO	CTGCTTATCA	CAGTGAA
							- Section 16
	(706)	706	,72	0	730	740	752
EctSC_codon_optimized	(706)	AGCAAAAC	CAATTCG	IGGCTGCC	GAAAGTCGI	TAGCCACCAC	GAAGTCC
EctSC_native	(706)	AGCAAAAC	CAATTCG	IGGCTGCG	GAAAGTCGI	TAGCCACCAC	GAAGTCC
	(752)	752 70	20	770	790		- Section 17
EctSC codon ontimized	(753)	ACCACCAC	00 Δ ΤΤ Δ Δ Τ Δ Δ Τ (				199
EctSC_codon_optimized	(753)	AGGAGGAC	ATACAAT	GCCTACCC	TAAAAAG	GAATTCAATC	CAACAACC
	( )						- Section 18
	(800)	800	810	820	3,	30	846
EctSC_codon_optimized	(800)	CAAAGGGC	ATCGTGC	IGTCCTTC	CCCAACCGT	GATGCTGCG	GTCGCCCA
EctSC_native	(800)	CCAAAGGC	ATTGTTT	FGAGTTTC	CCCCACCG	TAATGCTCCG	STCGCCCA
	(847)	847	860		870	880	893
EctSC_codon_optimized	(847) (847)	847 ACCGATGG	860 CGACGGC	FACAACCI	870 GCACCAGO	880 CTCGTGGCAC	893 GCTGCCA
EctSC_codon_optimized EctSC_native	(847) (847) (847)	847 ACCGATGG ACCGACGG	860 CGACGGC CGACGGT	FACAACCI FACAACCI	870 GCACCAGO TCATCAGO	880 CTCGTGGCAC CTGGTGGCGC	893 GCTGCCA GCTGCCA
EctSC_codon_optimized EctSC_native	(847) (847) (847)	847 ACCGATGG ACCGACGG	860 CGACGGC CGACGGT	FACAACCI FACAACCI	870 GCACCAGO TCATCAGO	880 CTCGTGGCAC CTGGTGGCGC	893 CGCTGCCA CGCTGCCA - Section 20
EctSC_codon_optimized EctSC_native	(847) (847) (847) (847) (894)	847 ACCGATGG ACCGACGG 894 900	860 CGACGGC CGACGGT	FACAACCI FACAACCI 910	870 GCACCAGO TCATCAGO 920	880 CTCGTGGCAC CTGGTGGCGC 930	893 GCTGCCA GCTGCCA - Section 20 940
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(847) (847) (847) (894) (894) (894)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG	860 CGACGGC CGACGGT ATACCAAC	FACAACCI FACAACCI _910 CTCCGTGT FTCGGTCI	870 GCACCAGO TCATCAGO <u>920</u> CACTGCAAO	880 CTCGTGGCAC CTGGTGGCGC 930 CCTGCTGCAC	893 GCTGCCA GCTGCCA - Section 20 940 GTGCTCCG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG	860 CGACGGC CGACGGT ATACCAA ATACCAA	PACAACCT PACAACCT 910 CTCCGTGT FTCGGTCT	870 GCACCAGO TCATCAGO 920 CACTGCAAO	880 CTCGTGGCAC CTGGTGGCGC 930 CCTGCTGCAG CCTGCTGCAG	893 CGCTGCCA CGCTGCCA - Section 20 940 TGCTCCG TGTTCCG - Section 21
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (894)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941	860 CGACGGC CGACGGT ATACCAA ATACCAA	FACAACCI FACAACCI 910 CTCCGTGI FTCGGTCI 960	870 GCACCAGO TCATCAGO 920 CACTGCAAO CACTGCAAO	880 CTCGTGGCAC CTGGTGGCGC 930 CCTGCTGCAG CCTGCTGCAG	893 GCTGCCA GCTGCCA - Section 20 940 GTGCTCCG GTGTTCCG - Section 21 987
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(847) (847) (847) (894) (894) (894) (894) (941) (941)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCA	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG	FACAACCT FACAACCT 910 CTCCGTGT FTCGGTCT 960 CAATCGCP	870 TCATCAGC 920 TACTGCAAC CACTGCAAC 97 GCAGAAAA	880 CTCGTGGCAC CTGGTGGCGC 930 CCTGCTGCAG CCTGCTGCAG 0 ACGCACAGGG	893 CGCTGCCA CGCTGCCA - Section 20 940 GTGCTCCG GTGTTCCG - Section 21 987 GCGAACTG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (894) (941) (941) (941)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCA	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG	JACAACCT JACAACCT 210 CTCCGTGT TTCGGTCT 960 CAATCGCA CCATCGCC	870 TCATCAGC 920 TACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAGAA	880 CTCGTGGCAC CTGGTGGCGC 200 CCTGCTGCAC CCTGCTGCAC 0 ACGCACAGGG ACGCCCAAGG	893 CGCTGCCA CGCTGCCA - Section 20 940 GTGCTCCG GTGTTCCG - Section 21 987 GCGAACTG GCGAGCTG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (894) (941) (941) (941)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG	FACAACCT FACAACCT 910 TTCGGTGT 960 CAATCGCA CCATCGCC	870 GCACCAGO TCATCAGO 920 CACTGCAAC GCAGAAAA GCAGAAAA GCAGAAAA	880 ETCGTGGCAC ETGGTGGCGC 930 ECTGCTGCAC CCTGCTGCAC 0 ACGCACAGGG ACGCCCAAGG	893 CGCTGCCA CGCTGCCA - Section 20 940 CTGCTCCG CGCTCCG - Section 21 987 CGCAACTG CGCACTG - Section 22
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (988)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT 988	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (1000	PACAACCT PACAACCT 910 TTCCGTGT 960 CAATCGCA CCATCGCC	870 GCACCAGO P20 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA 010	880 ETCGTGGCAC ETGGTGGCGC 930 ECTGCTGCAC CCTGCTGCAC 0 ACGCACAGGG ACGCCCAAGG 1020	893 CGCTGCCA CGCTGCCA - Section 20 940 TGCTCCG TGTTCCG - Section 21 987 CGCAACTG CGCAACTG CGCAACTG - Section 22 1034
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (988) (988) (988)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT 988 GTGGGCTT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (ATACCG CATCTCC	FACAACCT FACAACCT 910 TTCGGTGT TTCGGTGT 960 CAATCGCA CCATCGCC 1 GGCTACCG	870 GCACCAGO P20 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA 010 CTCCACCAT	880 CTCGTGGCAC CTGGTGGCGC 930 CCTGCTGCAC CCTGCTGCAC 0 ACGCACACAGGC ACGCCCAAGGC 1020 CCCCGTCCAG	893 CGCTGCCA CGCTGCCA - Section 20 940 TGCTCCG TGCTCCG - Section 21 987 CGCAACTG CGCACTG - Section 22 1034 CGACCG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (988) (988) (988)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT 988 GTGGGCTT GTCGGTTT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (1000 CATCTCC CATCTCG	PACAACCT PACAACCT 910 CTCCGTGT TTCGGTCT 960 CAATCGCA CCATCGCC ,1 GGCTACCC GGTTACCC	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA 010 GTCCACCAT	880 ETCGTGGCAC ETGGTGGCGC 930 CCTGCTGCAC CCTGCTGCAC 0 ACGCACACAGGC ACGCCCAAGGC 1020 FCCCGTCCAC	893 CGCTGCCA CGCTGCCA - Section 20 940 FTGCTCCG FTGCTCCG FGCTCCG CGACTG CGAACTG CGAACTG CGAACTG CGAACTG - Section 22 1034 CATACCCT - Section 23
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (988) (988) (988) (988)	847 ACCGATGG ACCGACGG 894 900 GCCACTGG GCCCCTCG 941 ATTTCGCT 988 GTGGGCTT GTCGGTTT 1035 1040	860 CGACGGC CGACGGT ATACCAA 950 GATACCGG GACACCGG ,1000 CATCTCCG	PACAACCT PACAACCT 910 CTCCGTGT TTCGGTCT 960 CAATCGCA CCATCGCC GGCTACCC GGCTACCC 1050	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA 010 TCCACCAT GCCCCCTT 1060	880 ETCGTGGCAC ETGGTGGCGC 930 CCTGCTGCAC CCTGCTGCAC 0 ACGCACACGGC ACGCCCAAGG 1020 FCCCGTCCAC CCCGTCCAC 1070	893 CGCTGCCA CGCTGCCA - Section 20 940 FTGCTCCG FTGCTCCG FGCTCCG SCGAACTG CGAACTG CGAACTG CGAACTG CGAACTG CGAACTG - Section 22 1034 FATACCCT SACACGCT - Section 23 1081
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (1000 CATCTCC CATCTCC GGCAGGT	PACAACCT PACAACCT 910 CTCCGTGT TTCGGTCT 960 CAATCGCA CCATCGCC 1050 CGCAGTGC	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA 010 TCCACCAT 6CCCCCTT 1060 GATTCCTCC	880 2TCGTGGCAC 2TGGTGGCGC 930 2CTGCTGCAC 2CTGCTGCAC 0 ACGCACACGGC ACGCCCAAGG 1020 TCCCGTCCAG CCCCGTCCAG 1020 TCCCGTCCAG 1070 2ATGCGTGCG	893 GCTGCCA GCTGCCA - Section 20 940 FTGCTCCG GTGTTCCG 987 GCGAACTG - Section 21 987 GCGAACTG - Section 22 1034 GATACCCT - Section 23 ATACCCT - Section 23 1081 CAGGGCC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT GTTCGTCT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (ACACCGC CATCTCC GGCAGGT GGCAGGT	PACAACCT PACAACCT 910 TTCGGTGT TTCGGTGT 960 CAATCGCA CCATCGCC GGCTACCG GGCTACCG 1050 CGCAGTGC CGCAGTGC	870 GCACCAGO TCATCAGO P20 CACTGCAAC CACTGCAAC 97 GCCAGAAAA GCAGAAAA 010 TCCCACCAT CCCCCCT 1060 GATTCCTCC GACAGTTCC	880 2TCGTGGCAC 2TGGTGGCGC 930 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 TCCCGTCCAG CCGCGGCCGG ,1070 2ATGCGTGGC 3ATGCGCGGT	893 CGCTGCCA CGCTGCCA - Section 20 940 FTGCTCCG FTGCTCCG 967 GCGACTCC 987 GCGACTG - Section 22 1034 GACGCC - Section 23 1034 CACGCCT - Section 23 1081 CAGGGCC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG (ACACCG CATCTCC CATCTCC GGCAGGT GGCAGGT	PACAACCT PACAACCT P10 TTCGGTGT TTCGGTCT 960 CAATCGCA CCATCGCC GGCTACCG GGCTACCG GGTTACCG	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA 100 TCCACCAT 1060 CATTCCTCC CACCACCAT	880 2TCGTGGCAC 2TGGTGGCGC 930 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 CCCCGTCCAG CCCCCGTCCAG CCCCCGTCCAG CCCCCGTCCAG CCCCCGTCCAG CCCCCGTCCAG CCCCCCGTCCAG CCCCCCCCGTCCAG CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	893 CGCTGCCA CGCTGCCA - Section 20 940 FTGCTCCG FTGCTCCG GTGTCCG 987 CGAACTG CGAACTG CGAACTG CGAACTG CGAACTG CGACCCT - Section 23 1081 CAGGGCC - Section 24
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1082)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTCT GTTCGTCT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG 1000 CATCTCC CATCTCG GGCAGGT GGCAGGT	PACAACCT PACAACCT P10 TTCGGTGT TTCGGTCT 960 CAATCGCA CCATCGCC GGTACCG GGTACCG GGTACCG CGCAGTGC CGCCGTCC 1100	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA 100 TCCACCAT 1060 GATTCCTCC GAGTTCC 11110	880 CTGGTGGCAC 27GGTGGCGC 2300 CCTGCTGCAG CCTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 CCCCGTCCAG CCCCGTCCAG CCGCGGCCGG 1070 CATGCGTGGC ATGCGCGGC	893 GCTGCCA GCTGCCA - Section 20 940 FTGCTCCG GTGTCCG GTGTCCG 987 GCGAACTG - Section 22 1034 GCGAGCTG - Section 23 1081 CAGGGCC - Section 24 1128
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1082) (1082)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035           1040           GTTCGTGT           1082           TGGCACTG	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG ATACCG CATCTCC CATCTCC GGCAGGT GGCAGGT GGCAGGT	PACAACCT PACAACCT P10 TTCGGTGT TTCGGTCT 960 CAATCGCA CCATCGCC GGTACCG GGTACCG GGTACCG CGCCGTCC CGCCGTCC 1000 TGCTGGCA	870 GCACCAGO TCATCAGO 920 CACTGCAAC GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA GCAGAGAA 100 TTCCACCAT 1060 GATTCCTCC JACAGTTCC	880 2TCGTGGCAC 2TGGTGGCGC 230 2CTGCTGCAG 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 TCCCGTCCAG CCCGGCCGG 1070 2ATGCGTGGC 3ATGCGCGGG 2ACGCGTGGC	893 CGCTGCCA CGCTGCCA - Section 20 940 FTGCTCCG FTGCTCCG FTGCTCCG CGACTG CGAACTG CGAACTG CGAACTG CGAACTG CGAACTG CGAACTG CGACCT CGACCT CCACGCC - Section 24 1128 CACGCGAA
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1082) (1082) (1082)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT GTCCGTCT 1082 TGGCACTG TGGCACTG	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC GGCAGGT GGCAGGT 1090 CGCATGC	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CCATCGCC J100 CGCAGTGCA TGCTGGCA	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA GCAGACAC 100 GACTCCCCCT 1060 GATTCCCCC 11110 CCCGACCGC	880 2TCGTGGCAC 2TGGTGGCGC 930 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 TCCCGTCCAG TCCCGTCCAG CGCGGCCGG 1070 ATGCGTGGC ATGCGCGGGCGG 2ACGCGTGGC 2CCGGGTCGC	893 CGCTGCCA CGCTGCCA - Section 20 940 TGCTCCG TGCTCCG TGCTCCG GCGACTG - Section 21 987 CGCACTG CGAGCTG - Section 22 1034 TATACCCT - Section 23 CAGGGCC - Section 24 1128 CAGGGCA - Section 27 - Section 24 - Section 24 - Section 24 - Section 25 - Section 24 - Section 25 - Section 25
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1035) (1082) (1082) (1082)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035         1040           GTTCGTGT           1082           TGGCACTG           TGGCACTG           TGGCCCTG	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC GGCAGGT GGCAGGT GGCAGGT CGCATGC CGCATGC	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CCATCGCC 1050 CGCAGTGC CGCCGTCG 1100 IGCTGGCA IGCTGGCA	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 010 TCCACCAT 1060 GATTCCTCC 11100 GATTCCTCCC 11100 GATTCCTCCCCT 11100 GATTCCTCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCT 11100 GATTCCTCCCCCCCCT 11100 GATTCCTCCCCCCCT 11000 GATTCCTCCCCCCT 11000 GATTCCTCCCCCCT 11000 GATTCCTCCCCCCCCT 11000 GATTCCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	880 2TCGTGGCAC 2TGGTGGCGC 930 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 TCCCGTCCAG TCGCGGCCGG 1070 ATGCGTGGC ATGCGTGGC ATGCGCGGGCGG 1460	893 GCTGCCA GCTGCCA - Section 20 940 GTGCTCCG GTGTCCG GCGACTG - Section 21 987 GCGAACTG GCGAGCTG - Section 22 1034 GATACCCT - Section 23 1081 CAGGGCC - Section 24 1128 CAGGCGAA - Section 25 - Section 24 1128 - Section 24 1128 - Section 24 - Section 25 - Section 24 - Section 24 - Section 25 - Section 24 - Section 25 -
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1035) (1032) (1082) (1082) (1082) (1129)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035         1040           GTTCGTGT           1082           TGGCACTG           TGGCACTG           TGGCACTG           TGGCCCTG	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC GGCAGGT GGCAGGT GGCAGGT CGCATGC CGCATGC 1140	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CAATCGCC 1050 CGCCGTCG 1050 CGCCGTCG 1100 IGCTGGCA IGCTGGCA 1110	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 010 TCCACCAT 1060 SATTCCTCA CACTGACCGC CTGACCGCC 50	880 CTGGTGGCAG 930 CTGCTGCAG CTGCTGCAG CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG ,1020 CCCGGCCGG ,1070 CATGCGGGCGG ATGCGCGGGGG CCCGGGTGGC ,1160	893 GCTGCCA GCTGCCA Section 20 940 GTGCTCCG GTGTCCG GCGACTG GCGACTG GCGACTG GCGACTG GCGACTG GCGACCG Section 22 1034 GACACGCT ACACGGCC Section 24 1128 CAGGGCA CAGGGCA CAGGCAA CCGCGAA CCGCGAG Section 25 1175
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1035) (1032) (1082) (1129) (1129) (1129)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCA 988 GTGGGCTT GTCGGGTTT 1035 1040 GTTCGTGT GTTCGTCT 1082 TGGCACTG TGGCCCTG 1129 TACGGCGT	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC GGCAGGT GGCAGGT CGCATGC CGCATGC CGCATGC (1140 GCGCTAC	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CCATCGCC 1050 CGCCGTCGCA I050 CGCCGTCGCA I050 CGCCGTCGCA I100 IGCTGGCA IGCTGGCA IGCTGGCA I112 IGCTGGCA I112 IGCTGGCA	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 010 TCCACCAT 1060 SATTCCTCC SACCGCCTT 1060 SATTCCTCC SACCGCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCCTT 1060 SATTCCTCCCCCTT 1060 SATTCCTCCCCCTT 1060 SATTCCTCCCCTT 1060 SATTCCTCCCCTTCCCCCTT 1060 SATTCCTCCCCCTTCCCCCTT 1060 SATTCCTCCCCCTTCCCCCTTCCCCCCTTCCCCCCCTTCCCC	880 2TCGTGCCAC 2TGGTGCAC 230 2CTGCTGCAC 2CTGCTGCAC 0 ACGCACACAGGC ACGCCCAAGC ,1020 TCCCGTCCAC CCCCGTCCAC ATGCGCGCCGC 2ATGCGCGGCGC 2ACGCCGGTCGC ,1160 TCCCCAGATA	893 CGCTGCCA CGCTGCCA Section 20 940 TGCTCCG TGCTCCG SCGACTG CGAACTG CGAACTG CGAACTG CGAACTG CGAACTG CGACCT Section 22 1034 CAGGGCC CAGGGCC CAGGGCC CAGGGCC CAGGGCC Section 24 1128 CAGGCGAA TCGCGAA CCGCGAG Section 25 1175 CACGCGCGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (988) (1035) (1035) (1035) (1035) (1032) (1082) (1082) (1082) (1082) (1082) (1082) (1082)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT GTCGTCT 1082 TGGCACTG TGGCACTG TGGCACTG TGGCACTG TGGCACTG	860 CGACGGC CGACGGT ATACCAA 950 GATACCG GACACCG CATCTCC GGCAGGT GGCAGGT CGCATGC CGCATGC CGCATGC CGCATGC (1090 CGCATGC CGCATGC	JACAACCT ACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CCATCGCC 1050 CGCAGTGCC CGCCGTCG 1100 IGCTGGCA IGCTGGCA IGCTGGCA III0	870 GCACCAGO TCATCAGO 920 CACTGCAAC CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAGAA 010 TCCACCAT 1060 SATTCCTCC 1110 CTGACCGC CCGCCATCT CCACCATCT CCACCATCT	880 2TCGTGCCAC 2TGGTGCCAC 230 2CTGCTGCAG 2CTGCTGCAG 2CTGCTGCAG 0 ACGCACAGGG ACGCCCAAGG 1020 TCCCGTCCAG 1070 ATGCGCGGCCGG 2ATGCGCGGCGG 2ACGCCGGTCGC 1160 TCCCCAGATA CCCCCAGATA	893 GCTGCCA GCTGCCA Section 20 940 GTGTCCG GTGTCCG GCGACTG GCGACTG GCGACTG GCGACTG GCGACTG GCGACTG CAGGGCC CAGGGCC CAGGGCC CAGGGCC CAGGGCC Section 24 1128 CAGGCGAA CTCGCGAG Section 25 1175 ACGCGCC Section 26 CACGGCC Section 26 CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCAA CACGCCCAA CACGCCCAA CACGCCCAA CACGCCCAA CACCCCCAA CACGCCCAA CACGCCCAA CACGCCCAA CACGCCCAA CACGCCCCAA CACGCCCAA CACCCCCCAA CACGCCCCAA CACGCCCAA CACCCCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCAA CACCCCCCAA CACCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCCAA CACCCCCCCC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (988) (1035) (1035) (1035) (1035) (1035) (1082) (1082) (1082) (1129) (1129) (1176)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035         1040           GTTCGTGT           1082           TGGCACTG           TGGCACTG           TGGCACTG           TACGGCGT           1129           TACGGCGT           TACGGCGT           1176	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC GGCAGGT GGCAGGT CGCATGC CGCATGC CGCATGC 1090 CGCATGC CGCATGC (1140 GCGCTAC) 1140	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CAATCGCA CATCGCA CATCGCA CATCGCA CATCGCA J100 IGCTGGCA IGCTGGCA IGCTGGCA ATGGAAAC J050 90	870 GCACCAGO TCATCAGO 920 CACTGCAAO CACTGCAAO 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 100 TCCACCAT 1060 SATTCCTCC CACCATCT CCACCATCT 1200	880 2TCGTGCCAC 930 2CTGCTGCAC CCTGCTGCAC CCTGCTGCAC 0 ACGCACAGGC ACGCCCAAGC ,1020 TCCCGTCCAC TCGCGGCCGC ATGCGCGGCCGC ,1070 CATGCGTGGC CCCGGGTCGC ,1160 TCCCCAGATA TCGCCGGACA ,1210	893 CGCTGCCA CGCTGCCA Section 20 940 TGCTCCG TGCTCCG SCGACTCG CGACTG CGACTG CGACTG CGACTG CGACCTG Section 22 1034 CACGCCT CCCGACCT CACGGCC CACGGCC CACGGCCA CCCGCGAG Section 25 ACGCGCGC CCCCCCCCCCCCCCCCCCCCCCCCCCCCC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (988) (1035) (1035) (1035) (1035) (1035) (1035) (1032) (1082) (1082) (1082) (1129) (1129) (1176) (1176) (1176)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035         1040           GTTCGTGT           1082           TGGCACTG           TGGCACTG           TACGGCGT           1129           TACGGCGT           TACGGCGT           1176           ATCCCAGG	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG ACACCG CATCTCC CATCTCC GGCAGGT GGCAGGT CGCATGC CGCATGC (1140 GCGCTACA GCGTTACA 11140	JACAACCT JACAACCT JACAACCT JCCGTGT TCCGTCT J60 CAATCGCA CATCGCC J100 TGCTGGCA CGCCGTCG J100 TGCTGGCA TGCTGGCA ATGGAAACC 90 CAAACGCG	870 GCACCAGO TCATCAGO 920 CACTGCAAO CACTGCAAO 97 GCAGAAAAA GCAGAAAAA GCAGAAAAA GCAGAAAAA 010 TCCACCAT 1060 GATCGCCCT 1100 CCACCATCT 1200 GCATTCGA	880 CTGGTGGCAG 930 CTGCTGCAG CTGCTGCAG CTGCTGCAG ACGCCCAAGG 1020 CCCGGCCAG 1020 CCCGGCCAG 1070 CATGCGCGGCGG 0 CCCGGGTCGC 1160 CCCCAGATA CCCCCAGATA CCCCCAGATA	893 CGCTGCCA CGCTGCCA Section 20 940 TGCTCCG TGCTCCG TGCTCCG CGACTG CGACTG CGACTG CGACCG Section 22 1034 CAGGGCC CAGGGCC CAGGGCC CAGGGCCA CAGGGCCA CAGGGCA CCAGGGCA Section 25 1175 CACGGGAG Section 26 CCCCGAGGCC Section 26 Section 2
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (988) (988) (988) (988) (1035) (1035) (1035) (1035) (1035) (1035) (1032) (1082) (1082) (1082) (1129) (1129) (1176) (1176) (1176)	847           ACCGATGG           ACCGACGG           894         900           GCCACTGG           GCCACTGG           941           ATTTCGCA           ATTTCGCA           988           GTGGGCTT           1035         1040           GTTCGTGT           1082           TGGCACTG           TGGCACTG           TACGGCGT           1129           TACGGCGT           TACGGCGT           1176           ATCCCAGG           GTCACAGG	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG (ACACCG CATCTCC CGCAGGT (1000 CGCAGGT CGCAGGT (1140 GCGCTACZ (1140 GCGCTACZ (1140 CCGCTGTT (CGCTGTT)	JACAACCT JACAACCT J10 TTCGGTGT TTCGGTGT J60 CAATCGCA CATCGCC J100 TGCTGGCA CGCCGTCG J100 TGCTGGCA TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGCA J100 TGCTGCA J100 TGCTGCCA J100 TGCTGCA J100 TGCTGCA J100 TGCTGCA J100 TGCTGCA J100 TGCTGCCA J100 TGCTGC	870 GCACCAGO TCATCAGO 920 CACTGCAAO CACTGCAAO GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 1000 GTCCACCAT 11000 GCACCATCI 1200 GCATTCGA	880 2TCGTGCGCAC 230 2CTGCTGCAC 2CTGCTGCAC 2CTGCTGCAC 0 ACGCACAGGC ACGCCCAAGC ,1020 TCCCGTCCAC TCGCGGCCGG ,1070 ATGCGCGGCGG 2ACGCGGGCGG ,1160 TCCCCAGATA TCGCCGGACA ,1210 TCGCCTGGAT	893 CGCTGCCA CGCTGCCA Section 20 940 TGCTCCG TGCTCCG CGACTCG CGACTG CGACTG CGACTG CGACTG CGACCG Section 22 1034 CAGGGCC CAGGGCC CAGGGCC CAGGGCCA CCGCGAG CSCCION 25 1175 CCGCGACT CCCACT CCCACT CCCACT CCCACT
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1032) (1129) (1126) (1176) (	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT 988 GTGGGCTT GTCGGTTT 1035 1040 GTTCGTGT 1035 1040 GTTCGTGT 1082 TGGCACTG TGGCACTG TGGCCCTG 1129 TACGGCGT ACCGCGT	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG (ACACCG CATCTCC CGCAGGT (1000 CGCAGGT CGCAGGT (1140 GCGCTACA GCGTTACA 1140 CCCTGTT CCCTGTT	JACAACCT JACAACCT J10 CTCCGTGT TTCGGTCT J60 CAATCGCA CATCGCC J100 CGCCGTCGCA J100 IGCTGGCA IGCTGGCA J100 IGCTGGCA J110 IGCTGGCA J110 IGCTGGCA J110 IGCTGGCA J110 IGCTGGCA J110 IGCTGGCA J110 IGCTGGCA J100 IGCTGGCA J100 IGCTGGCA J100 IGCTGGCA J100 IGCTGGCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA J100 IGCTGCCA IGCCA	870 GCACCAGO TCATCAGO 920 CACTGCAAO CACTGCAAO 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 100 TCCACCAT 1060 GATTCCACCAT 1110 CCTGACCGC CCACCATCT 1200 GCATTCGA	880 2TCGTGCCAC 930 2CTGCTGCAC CTGCTGCAC CCTGCTGCAC 0 ACGCACAGGC ACGCCCAAGC ,1020 TCCCGTCCAC TCGCGGCCGG ,1070 ATGCGCGGCGG CCCGGGTCGC ,1160 TCCCCAGATA TCGCCGGACA ,1210 TCGCCTGGAT	893 CGCTGCCA CGCTGCCA Section 20 940 TGCTCCG TGCTCCG TGCTCCG CGACTG CGACTG CGACTG CGACTG CGACTG CGACCG Section 22 1034 CACGGCC Section 24 1128 CAGGGCC CAGGGCC CAGGGCCA CCGCGAG Section 25 1175 CACGGGCA Section 26 20 CCACGGCAC Section 26 20 20 20 20 20 20 20 20 20 20 20 20 20
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (1035) (1129) (1129) (1176) (1176) (1176) (1176) (1122) (112) (112) (112) (112) (112) (112) (112) (112) (11	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCA 988 GTGGGCTT GTCGGTTT 1035 _1040 GTTCGTGT 1035 _1040 GTTCGTGT 1082 TGGCACTG TGGCACTG TGGCCCTG 1129 TACGGCGT 1176 ATCCCAGG GTCACAGG	860 CGACGGC CGACGGT ATACCAA ATACCAA 950 GATACCG GACACCG (ACACCG CATCTCC CGCAGGT (1000 CGCAGGT CGCAGGT (1140 GCGCTACC GCGTTACC 1140 CGCGTTACC 1140 CGCGTTACC 230	JACAACCT JACAACCT JACAACCT JO CACCGTGT JGGCTACCG GGTTACCG GGTTACCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGTCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCG IO50 CGCCGCGCGC IO50 CGCCGCGC IO50 IO5	870 GCACCAGO 920 CACTGCAAC 920 CACTGCAAC 97 GCAGAAAA GCAGAAAA GCAGAAAA GCAGAAAA 010 TCCACCAT 1060 GATTCCACCAT 1110 CCTGACCGC CCACCATCT 1200 GCATTCGAC GCATTCGAC GCATTCGAC 1250	880 CTGGTGGCAG 930 CTGCTGCAG CTGCTGCAG CTGCTGCAG ACGCCCAAGG 1020 CCCCGTCCAG CCCGGCCGG 1070 CATGCGGGCGG CCCGGGTCGC 1160 CCCCCAGATA CCCCCAGATA CCCCCGGACA	893 GCTGCCA GCTGCCA Section 20 940 GTGCTCCG GTGTTCCG GCGACTG GCGACTG GCGACTG GCGACTG GCGACTG GCGACTG GCGACCG Section 22 1034 GACGCGCA GCACGCG Section 24 1128 CAGGCGAG Section 25 1175 ACGCGAC Section 25 1175 CCCACGGCG Section 26 1222 GCAACT GCCAACT GCCAACT CCCACC Section 27 1289 CCCACC Section 27 CCCACC Section 27 CCCACCC Section 27 CCCACCCC Section 27 CCCACCCC Section 27 CCCACCCC Section 27 CCCACCCC Section 27 CCCACCCCCCCC Section 27 CCCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(847) (847) (847) (894) (894) (894) (941) (941) (941) (941) (941) (941) (988) (988) (988) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1035) (1032) (1129) (1129) (1176) (1176) (1176) (11223) (1223)	847 ACCGATGG ACCGACGG GCCACTGG GCCACTGG GCCCCTCG 941 ATTTCGCA ATTTCGCT 988 GTGGGCTT GTCGGTTT 1035 _1040 GTTCGTGT 1035 _1040 GTTCGTGT 1082 TGGCACTG TGGCCCTG 1129 TACGGCGT 1176 ATCCCAGG GTCACAGG GCACCACC GCACCACC	860 CGACGGC CGACGGC CGACGGC CGACGGC ATACCAA 950 GATACCGG GATACCGG ATACCGC GGCAGGT GGCAGGT GGCAGGT GGCAGGT (1090 CGCATGCC CGCATGCC 1140 GCGCTACC GCCTGTT CGCCTGTT CGCCGCTGTT CGCCACCCC	JACAACCT JACAACCT JACAACCT JCCGTGT TCCGTGT TCCGTCT J60 CAATCGCA CAATCGCA CGCCGTCG J100 TGCTGGCA J100 TGCTGGCA J100 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA J110 TGCTGGCA TGCTGGCA TGCTGGCA TGCTGGCA TGCTCGCA	870 GCACCAGO 920 CACTGCAAC 920 CACTGCAAC 97 GCAGAAAAA GCAGAAAAA GCAGAAAAA 010 TCCACCAT 1060 GATTCCACCAT 1110 CCGCCACCATCT 1200 GCACTCGAC GCATTCGAC CCACCATCT 1200 GCATTCGAC CCACCATCA	880 CTGGTGGCAG 930 CTGCTGCAG CTGCTGCAG CTGCTGCAG ACGCCCAAGG 1020 CCCGCGCCAG 1020 CCCGGCCGG 1070 CATGCGCGGCGG 2ACGCGGCGG 1070 CACGCGGGCGG 1070 CACGCGGGCGG 1070 CCCGGGCGGCGG 1070 CCCGGGCGGCGG 1070 CCCGGGCGGCGG 1070 CCCGGGCGGGCGG 1070 CCCGGGCGGGCGG 1070 CCCGGGCGGGCGG 1070 CCCGGGCGGGCGG 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGCGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCGGGCGGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCGGGCGGCGGC 1070 CCCCCGGGCGGCGGC 1070 CCCCCGGGCGGCGGC 1070 CCCCCGGGCGGC 1070 CCCCCGGGCGGC 1070 CCCCCGGGCGGC 1070 CCCCCGGGCCGGC 1070 CCCCCGGGCCGGC 1070 CCCCCGGCCGGGC 1070 CCCCCGGGCCGGC 1070 CCCCCGGCCGGC 1070 CCCCCGGCCGGC 1070 CCCCCCGGCCGGC 1070 CCCCCCGGCCGGC 1070 CCCCCCGGCCGGC 1070 CCCCCCGGCCGGC 1070 CCCCCCGGC 1070 CCCCCCGGC 1070 CCCCCCGGC 1070 CCCCCCCGGC 1070 CCCCCCCGGCGC 1070 CCCCCCCGGC 1070 CCCCCCCCGGC 1070 CCCCCCCCCCCGGC 1070 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	893 GCTGCCA GCTGCCA - Section 20 940 GTGTCCGG GTGTCCG - Section 21 987 GCGAACTG GCGAACTG GCGACTG - Section 23 1031 CAGGGCC - Section 23 1081 CAGGGCC - Section 23 1081 CAGGGCC - Section 23 1081 CAGGGCC - Section 25 1175 ACGCGAA TCGCGAG - Section 25 1175 ACGGGGC - Section 26 CAGGGCCA - Section 27 1229 CGCAACT - Section 27 1269 CGCAACT - Section 27 1269 CGCACCAC

Fig. 36: Ectoine nucleotide sequence. (Continued from page 96).
(1270)	1270	1280	1290	1300	1316
EctSC_codon_optimized (1270)	CACGAA	GATGAAGT	GCTGTACCG	IGCAGGCCCATTO	CACCGTGTCCCA
EctSC_native (1270)	CACGAG	GACGAGGT	GCTCTACCG	CGCCGGCCCGTTC	CACCGTTTCCCA
					——— Section 29
(1317)	1317	13	30 ,1	1340 1350	) 1363
EctSC_codon_optimized (1317)	CCTGGA	AGAAGAAC	TGAAAGAAC <i>i</i>	ACGCATAAGAAGI	CCAGGAGGACA
EctSC_native (1317)	TCTAGA	AGAAGAGC'	I'CAAGGAGCA	ACGCATGAGAAGI	CCAGGAGGACA
(1264)	1264	1270	1290	1200	Section 30
(1364) EctSC codon ontimized (1264)	1304 TACAAT				
EctSC_codoII_optiIII2ed (1364)	TACAAI	GAAAACCI	TTGAACTGA	ACGAAICCCGIGI	TCGCAGCTACT
	111011111		i i dimic i dim	ii diiii eenooo	Section 31
(1411)	1411	.1420	.1430	.1440	1457
EctSC codon optimized (1411)	GCCGCT	CCTTCCCA	GTGGTGTTC <i>I</i>	AAGCAGGCACAG	GTGCAGAACTG
EctSC_native (1411)	GCCGTT	CCTTCCCC	GTGGTCTTC <i>I</i>	AAGCAGGCCCAG	GCGCCGAACTG
					Section 32
(1458)	1458	147	0 14	80 1490	1504
EctSC_codon_optimized (1458)	GTGACC	CAGGATGG	CAAGCGCTAC	CATCGATTTCTTC	GCAGGCGCAGG
EctSC_native (1458)	GTCACT	CAGGACGG	CAAGCGCTAC	CATCGACTTCCTC	CGCTGGTGCCGG
	4505 4	- 10	4500	4500	Section 33
(1505) EstCC and an antimized (1505)	1505 1		1520	1530 <sup>1</sup>	540 1551
EctSC_codon_optimized (1505)	CACCET	GAACTACG	GCCACAACCA	ACCCAGTGCTGA	AGCAGGCACTGC
	CACGCI	CAACIACO	JUCACAACCA	ACCEGGIGCICAP	Section 34
(1552)	1552	1560	1570	1580	1598
EctSC codon optimized (1552)	TGGAAT	ACATCGAA	TCCGATGGC	ATCACCCACGGC	CTGGATATGTAC
EctSC_native (1552)	TCGAGT	ACATCGAG	AGCGACGGC	ATCACCCACGGC	CTGGACATGTAC
					——— Section 35
(1599)	1599	1610	1620	1630	1645
EctSC_codon_optimized (1599)	ACCGAA	GCAAAAGA	ACGCTTCCT	GGAAACCTTCAAC	CCGCCTGATCCT
EctSC_native (1599)	ACCGAA	GCCAAGGA	GCGTTTCCT	CGAAACCTTCAAC	CCGGCTGATCCT
					——— Section 36
(1646)	1646		1660	_1670 _16	80 1692
EctSC_codon_optimized (1646)	GGAACC	ACGCGGTA:	IGGGCGATT <i>I</i>	ACCGCATGCAGT	CACCGGTCCAA
ECTSC_native (1646)	CGAGCC	GCGCGGCA:	IGGGCGACTA	ACCGCATGCAGT	CACCGGCCCGA
(1693)	1693	1700	1710	1720	1739
(1693) EctSC_codon_optimized (1693)	1693 CCGGCA	1700 CCAACGCA	1710 GTGGAAGCAG	<mark>,1720</mark> GCAATGAAGCTGC	1739 GCACGCAAGGTG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693)	1693 CCGGCA CCGGCA	_1700 CCAACGCAC CCAACGCGC	1710 GTGGAAGCAC GTCGAGGCGC	, <mark>1720</mark> GCAATGAAGCTGC GCGATGAAGCTGC	1739 GCACGCAAGGTG GCGCGCAAGGTC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693)	1693 CCGGCA CCGGCA	_1700 CCAACGCAC CCAACGCGC	1710 GTGGAAGCAC GTCGAGGCGC	,1720 GCAATGAAGCTGC GCGATGAAGCTGC 1770	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740)	1693 CCGGCA CCGGCA	_1700 CCAACGCAC CCAACGCGC _1750	1710 GTGGAAGCAC GTCGAGGCGC 1760	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1740)	1693 CCGGCA CCGGCA 1740 ACCGGT	_1700 CCAACGCAC CCAACGCG 	,1710 GTGGAAGCAC GTCGAGGCGC ,1760 CATCATTTCC CATCATCAG	<u>1720</u> GCAATGAAGCTGG GCGATGAAGCTGG <u>1770</u> CTTCACCAACGGG	1739 CACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCTG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1740)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG	_1700 CCAACGCA( CCAACGCG( _1750 CGCAACAA( CGCAACAA(	,1710 GTGGAAGCAG GTCGAGGCG ,1760 CATCATTTCC CATCATCAGT	<u>,</u> 1720 GCAATGAAGCTGG GCGATGAAGCTGG <u>,</u> 1770 CTTCACCAACGGG TTTCACCAACGGG	1739 CACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG Section 39 Section 39
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1740) (1787)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG		1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG		1739 CACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG Section 39 1833
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1747) EctSC_codon_optimized (1787)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT	_1700 CCAACGCAC CCAACGCGC 	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAGT 300 11 TGGCAGCAAG	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770 TTCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA	1739 CACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG Section 39 1833 ACCACCGTGGCG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) EctSC_native (1787)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT		1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAGT 000 1 TGGCAGCAAG	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770 TTCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCZ	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ATCACCGCGGCGCG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1740) (1787) EctSC_codon_optimized (1787) EctSC_native (1787)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 18 CGGTGCACC TGGCGCGCC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAGT 000 1 TGGCAGCAAG	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770 TTCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ATCACCGCGGCG Section 40
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1740) EctSC_native (1787) EctSC_codon_optimized (1787) EctSC_native (1787) (1834)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 18 CGGTGCACC TGGCGCGCC 1840	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 1000 1 GGCAGCAAG 1GGCCGCCAG	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770 TTCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ATCACCGCGGGCG CTCACCGCGGGCG Section 40 1870 1880
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_native (1740) EctSC_native (1787) EctSC_codon_optimized (1787) EctSC_native (1787) (1834) EctSC_codon_optimized (1834)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834 GTTCCG	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 18 CGGTGCAC TGGCGCGCC 1840 GCATCGGCC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 1000 11 GGCCGCCAG 1850 CTGACCGATC	,1720 GCAATGAAGCTGC GCGATGAAGCTGC ,1770 TTCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGC	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ACCACCGTGGCG CGCGCGCGCG Section 40 1870 1880 CCATACGCAAAC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) EctSC_native (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834)	1693 CCGGCA CCGGCA ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834 GTTCCG GCTCCG	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 18 CGGTGCAC TGGCGCGCC 1840 GCATCGGCC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 100 110 100 110 110 110 110 11	,1720 GCAATGAAGCTGC GCGATGAAGCTGC TTCACCAACGGC TTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGC GTCAGCCGCATGC	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ACCACCGTGGCG CGTACGCCAAC CGTACGCCAAC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) EctSC_native (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) (1901)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834 GTTCCG GCTCCG	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 18 CGGTGCAC TGGCGCGCC 1840 GCATCGGCC GCATCGGCC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAGT 1000 1 IGGCAGCAAC 1850 CTGACCGATC CTCACCGATC	,1720 GCAATGAAGCTGC GCGATGAAGCTGC 17TCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGC GTCAGCCGCATGC	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ACCACCGTGGCG TCCACCGCGGCG Section 40 1870 1880 CCATACGCCAAC CCGTACGCCAAC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1881)	1693 CCGGCA CCGGCA 1740 ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834 GTTCCG GCTCCG 1881 TAGTTC	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 188 CGGTGCAC 1840 GCATCGGCC (1890	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAGT 1000 1 GGCAGCAAC 1850 CTGACCGATC CTCACCGATC 1900	,1720 GCAATGAAGCTGC GCGATGAAGCTGC 17TCACCAACGGC TTTCACCAACGGC 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGC GTCAGCCGCATGC ,1910	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ATCACCGCGGGGG CATCCGCGCGCG CGTACGCCAAC CGTACGCCAAC CGTACGCCAAC Section 41 1927
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_native (1834) (1881) EctSC_codon_optimized (1881) EctSC_codon_optimized (1881)	1693           CCGGCA           CCGGCA           1740           ACCGGT           ACCGGT           1787           CTCCAT           CAGCAT           1834           GTTCCG           1881           TACTTCC	,1700 CCAACGCAC CCAACGCGC ,1750 CGCAACAAC CGCAACAAC 188 CGGTGCACC 1840 GCATCGGCC ,1890 GGCGATAAAC GCCGATAAAC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 1000 (1) TGGCAGCAACA 1850 CTGACCGATC CTCACCGATC 1900 GACCAACACC	,1720 GCAATGAAGCTGC GCGATGAAGCTGC 1770 TTCACCAACGGC TTCACCAACGGC 1810 1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGC 1910 CATCGGCATGATC	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCGG CTTCCACGGCGG Section 39 1833 ACCACCGTGGCG ATCACCGCGGGGG CACCGCGCAAC CGTACGCCAAC CGTACGCCAAC CGTACGCCAAC Section 41 1927 GGATAAGCTGCT
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_native (1834) EctSC_native (1834) EctSC_native (1834) EctSC_native (1831)	1693 CCGGCA CCGGCA ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1834 GTTCCG GCTCCG 1881 TACTTC TATTTC	,1700 CCAACGCA CCAACGCG CCAACGCG CGCAACAAC (GCAACAAC 188 CGGTGCAC 1840 GCATCGGCC (1890 GGCGATAAC GGCGATAAC	1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 1000 (1) TGGCAGCAACA 1850 CTGACCGATC CTCACCGATC 1900 GACCAACACC	,1720 GCAATGAAGCTGO GCGATGAAGCTGO TTTCACCAACGGO TTTCACCAACGGO ITTCACCAACGGO CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGO GTCAGCCGCATGO ,1910 CATCGGCATGATO	1739 GCACGCAAGGTG GCGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGCGCGG ACCACCGTGGCG ACCACCGTGGCG ACCACCGTGGCG CGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCACC Section 41 1927 GGATAAGCTGCT Section 42 Section 42 Section 41 927 SGACAAGCTGCT Section 41 Section 41 Se
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834) (1881) EctSC_codon_optimized (1881) EctSC_codon_optimized (1881) (1928)	1693         CCGGCA         1740         ACCGGT         ACCGGT         1787         CTCCAT         CAGCAT         1834         GTTCCG         1881         TACTTC         1881         TACTTC         1928		1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCACAG 1000 (1) TGGCAGCAACA 1000 (2) 1000 (2) 10	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA ,1860 GTGTCCCGCATGG GTCAGCCGCATGG TCAGCCGCATGATC CATCGGCATGATC 50 1960	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 TTCCACGGCTG TTCCACGGCTG CTTCCACGGCTG CCACGCGTGGCG ACCACCGTGGCG ACCACCGTGGCG 1870 1880 CCATACGCAAAC CCGTACGCAAAC CCGTACGCCAC CCGTACGCCAC CCGTACGCCAAC CCGTACGCCAC C
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1831) EctSC_codon_optimized (1831) EctSC_codon_optimized (1831) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928)	1693         CCGGCA         CCGGCA         1740         ACCGGT         ACCGGT         1787         CTCCAT         CAGCAT         1834         GTTCCG         1881         TACTTC         1928         GTCCGA		1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCAG 1000 (1) TGGCAGCAACA 1000 (2) TGGCCGCCAC 1850 CTGACCGATCG 1900 GACCAACACC 0 (19) CCGGCATCG2	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA ,1860 GTGTCCCGCATGG GTCAGCCGCATGG ,1910 CATCGGCATGATG 50 ,1960 ATAAGCCAGCAG	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 TTCCACGGCTG TTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG ACCACCGTGGCG CCATACGCAAC CCGTACGCAAC CCGTACGCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCA
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1881) EctSC_codon_optimized (1881) EctSC_codon_optimized (1881) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928)	1693 CCGGCA CCGGCA ACCGGT ACCGGT ACCGGT ACCGGT 1787 CTCCAT CAGCAT 1834 GTTCCG GCTCCG 1881 TACTTC TATTTC 1928 GTCCGA CTCCGA		1710 GTGGAAGCAG GTCGAGGCGG 1760 CATCATTTCC CATCATCACAT 1000 (1) TGGCAGCAACA 1000 (2) TGGCCGCCAC 1850 CTGACCGATCG GACCAACACC 0 (19) CCGGCATCG2 GCGGCATCG2	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCCGCATGG ,1910 CATCGGCATGATG CATCGGCATGATG 50 ,1960 ATAAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG ACCACCGTGGCG CCATACGCAAC CCGTACGCAAC CGTACGCCAAC CGTACGCCACT GGACAAGCTGCT GGCAAGCTGCT Section 42 1974 CAGTGATCGTCGCG CGTGATCGTCGC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834) (1881) EctSC_codon_optimized (1881) EctSC_native (1881) EctSC_native (1881) EctSC_native (1882) (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_native (1928)	1693           CCGGCA           CCGGCA           1740           ACCGGT           ACCGGT           ACCGGT           1787           CTCCAT           CAGCAT           1834           GTTCCG           1881           TACTTC           1928           GTCCGA           CTCCGA	_,1700 CCAACGCAC CCAACGCGC .,1750 CGCAACAAC CGCAACAAC .18 CGGTGCACC TGGCGCGCC .1840 GCATCGGCC .,1890 GCCATCGGCC .,1890 GGCGATAAC GGCGACAAC .,1944 TCCATCCTC CCCGTCCAC	1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATTCATCATCATCATCAG 1760	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGG ,1910 CATCGGCATGATG 50 ,1960 ATAAGCCAGCAGCA	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG ACCACCGTGGCG CGTACGCAAC CGTACGCAAC CGTACGCCAAC CGTACGCCACT GGACAAGCTGCT GGACAAGCTGCT 1974 CAGTGATCGTGG CGTGATCGTCG GGTGATCGTCG CGTGATCGTC CGTCG CGTGATCGTC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834) (1881) EctSC_native (1881) EctSC_native (1881) EctSC_native (1881) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_native (1928) EctSC_native (1928)	1693           CCGGCA           CCGGCA           1740           ACCGGT           ACCGGT           ACCGGT           1787           CTCCAT           CAGCAT           1834           GTTCCG           1881           TACTTC           1928           GTCCGA           CTCCGA           1928           GTCCGA           1975		1710 GTGGAAGCAG GTCGAGGCGG CATCATTTCC CATCATTCCC CATCATCAG 1760 17GGCAGCAAC 1GGCCGCCAC 1850 CTGACCGATCG 1900 GACCAACACC GACCAACACC 0 19 CCGGCATCGZ 3CGGGATCGZ 1990	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGG ,1910 CATCGGCATGATG 50 ,1960 ATAAGCCAGCAGCA CACAGCCGCCGC 2000 2	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG CCCTACGCAACC CCGTACGCAACC CCGTACGCAACC CGTACGCCAACC CGTACGCCACCT GGACAAGCTGCT GGACAAGCTGCT GGACAGCTGCTG CGTGATCGTGG CGTGATCGTGG CGTGATCGTCG CGTGA
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1881) EctSC_native (1881) EctSC_native (1881) EctSC_native (1882) (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975)	1693           CCGGCA           CCGGCA           1740           ACCGGT           ACCGGT           1787           CTCCAT           CAGCAT           1834           GTTCCG           1881           TACTTC           1928           GTCCGA           1975           1975		1710 GTGGAAGCAG GTCGAGGCGG CATCATTTCC CATCATTCAC GOO 11 TGGCAGCAAC TGGCCGCCAC 1850 CTGACCGATCC GACCAACACC GACCAACACC CGGCATCG2 CCGGCATCG2 1990 GAAGGCGGTC	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA GTGCCCGCATGG ,1910 CATCGGCATGATGA 50 ,1960 ATAAGCCAGCAGCA CAAGCCCGCCGG 2000 2 CTGAACACCGCATG	1739 CACGCAAGGTG CGCGCAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG CCATACGCAAC CGTACGCAAC CGTACGCAAC CGTACGCCAAC CGTACGCCACT GGACAAGCTGCT GGACAAGCTGCT GGACAGCTGCT GGACAGCTGCT GGCAACGTCGCG CGTGATCGTCG CGTGATCGTCG CGGTGATCGTCG CGGTGATCGTCG CGGCAGAATGG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1881) EctSC_native (1881) EctSC_codon_optimized (1881) EctSC_native (1881) (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975)	1693 CCGGCA CCGGCA ACCGGT ACCGGT ACCGGT ACCGGG 1787 CTCCAT CAGCAT 1884 GTTCCG GCTCCGA 1881 TACTTC TATTTC 1928 GTCCGA CTCCGA 1975 _1 AAGTGG AGGTGG		1710 GTGGAAGCAG GTCGAGGCGG GTCGATCATTTCC CATCATTACATTACA GOO 11 TGGCAGCAACA TGGCCGCCAC 1850 CTGACCGATCG 1900 GACCAACACCG GACCAACACCG 1990 GAAGGCGGTCG	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGG ,1910 CATCGGCATGATG CATCGGCATGATG 50 ,1960 ATAAGCCAGCAGCA CCAAGCCCGCCGC 2000 2 CTGAACACCGCAT	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG ACCACCGTGGCG CCGTACGCAAC CCGTACGCAAC CCGTACGCAAC CCGTAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GACAAGCTGCT GGCGACGCGACGCG Section 43 2010 2021 CCGCCGAGAATGG CCGCCGAGTGGC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1831) EctSC_native (1881) EctSC_codon_optimized (1881) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_native (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975)	1693           CCGGCA           CCGGCA           1740           ACCGGT           ACCGGT           1787           CTCCAT           CAGCAT           1884           GTTCCG           1881           TACTTC           1928           GTCCGA           1975           AGTGG           AGTGG	,1700 CCAACGCAC CCAACGCGC (CCAACGCGC CGCAACAAC (CGCAACAAC 1880 GCATCGGCC ,1890 GCCATCGGCC ,1890 GCCGATCAGCC (CCCGTCCAC 980 TGCAGGGCC TCCAGGCCC	1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATTACAT 1760 CATCATTACAT 1760 CATCATCAG 1760 CATCATCAG 1760 CATCATCAG 1760 CTGACCGAACAC 1850 CTGACCGATCG 1900 GACCAACACC 0 19 CCGGCATCGA 1990 GAAGGCGGTC GAAGGCGGTC	,1720 GCAATGAAGCTGO GCGATGAAGCTGO TTCACCAACGGO TTCACCAACGGO 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA GTGTCCCGCATGO ,1860 GTGTCCCGCATGO ,1910 CATCGGCATGATO CATCGGCATGATO 50 ,1960 ATAAGCCAGCAGCA CCAAGCCCGCCGO 2000 2 CTGAACACCGCAT	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCAAC CCGTACGCCACC Section 41 1927 GGATAAGCTGCT GGACAAGCTGCT GGACAAGCTGCT Section 43 2010 2021 CCGCCGAGATGG CCGCCGAGTGG CCGCCGAGTGG CCGCCGAGTGG CCGCCGAGTGG CCGCCGAGTGG
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1831) EctSC_codon_optimized (1831) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975) EctSC_native (1975) EctSC_native (1975)	1693 CCGGCA CCGGCA ACCGGG ACCGGG ACCGGG 1787 CTCCAT CAGCAT 1884 GTCCG GTCCGA 1881 TACTTC TATTTC 1928 GTCCGA CTCCGA 1975 _1 AAGTGG AGGTGG		1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATCAGT 300 11 TGGCAGCAAC TGGCCGCCAC 1850 CTGACCGATCG 1900 GACCAACACC GACCAACACC 0 19 CCGGCATCGZ 3CCGGCATCGZ 1990 GAAGGCGGTC 2040 A ACCTCTTCC	,1720 GCAATGAAGCTGO GCGATGAAGCTGO 1770 TTCACCAACGGO TTCACCAACGGO 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGO ,1910 CATCGGCATGATO CATCGGCATGATO 50 ,1960 ATAAGCCAGCAGCAGO ACAAGCCCGCCGO 2000 2 CTGAACACCGCAT CTGAACACCGCAT CTGAACACCGCAT 2050	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCG CACCCGCGGCGAAAC CCGTACGCCAAC CCGTACGCCAAC Section 41 1974 CAGTGATCGTGG GGGCGACGCGAATGG CGGCCGAGATGG CGGCCGAGATGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGCCCAC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1831) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022)	1693 CCGGCA CCGGCA ACCGGG ACCGGG ACCGGG ACCGGG 1787 CTCCAT CAGCAT 1884 GTCCGA GTCCGA 1881 TACTTC TATTCC 1928 GTCCGA GTCCGA 1975 _1 AAGTGG AGGTGG		1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATCAGT 000 11 TGGCAGCAAC TGGCCGCCAC 1850 CTGACCGATCG GACCAACACC GACCAACACC 0 190 GACCAACACC GACGGATCGZ J990 GAAGGCGGTC 2040 AAAGCTGTCC	,1720 GCAATGAAGCTGO GCGATGAAGCTGO 1770 TTCACCAACGGO TTCACCAACGGO 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGO 1910 CATCGGCATGATO CATCGGCATGATO 50 ,1960 ATAAGCCAGCATGATO 2000 2 CTGAACACGCATO 2000 2 CTGAACACGCATO 2000 2 CTGAACACGCATO CATCGGCAAGCACGAA	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CACCGCGGCGGCG CACCGCGGCGGCG CCATACGCAAAC CCGTACGCCAAC CCGTACGCCAAC Section 41 1927 GGATAAGCTGCTG GGCGAGAATGG CGGCCGAGATGG CGGCCGAGATGG CGGCCGAGTGGAT Section 44 2068 ATGCTGCTGCTGAT
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834) EctSC_codon_optimized (1831) EctSC_codon_optimized (1881) EctSC_codon_optimized (1928) EctSC_native (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_native (2022)	1693 CCGGCA CCGGCA ACCGGG ACCGGG ACCGGG ACCGGG 1787 CTCCAT CAGCAT 1884 GTCCGA GTCCGA 1881 TACTTC TATTCC 1928 GTCCGA GTCCGA GTCCGA 1975 _11 AAGTGG AGGTGG ATGCGCC		1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATCAGT 000 11 TGGCAGCAAC TGGCCGCCAC 1850 CTGACCGATC 1900 GACCAACACC GACCAACACC 0 19 CCGGCATCGZ GAAGGCGGTC GAAGGCGGTC CAAGCTGTGG GAAGGCTGTGG	,1720 GCAATGAAGCTGO GCGATGAAGCTGO TTCACCAACGGO TTCACCAACGGO 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA ,1860 GTGTCCCGCATGO GTGTCCGGCATGO ,1910 CATCGGCATGATO CATCGGCATGATO CATCGGCATGATO CATCGGCATGATO CTGAACACGCAGCA 2000 2 CTGAACACGCAGCA CCGCCAGCACGAA CCGCAAGCACGAA CCGCCAGCACGAA	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CACCGCGGCGGCG CACCGCGGCGGCG CCATACGCAAAC CCGTACGCCAAC CCGTACGCCAAC Section 41 1927 GGATAAGCTGCTG GGCGAGATGG CGGCGAGATGG CGGCCGAGATGG CGGCCGAGATGG CGGCCGAGTGG CGGCCGGCGGAGTGG CGGCCGGCGGAGTGG CGGCCGGCGGGCGG CGGCCGGCGGG CGGCCGGC
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) (1740) EctSC_codon_optimized (1740) EctSC_codon_optimized (1747) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_native (1834) EctSC_codon_optimized (1831) EctSC_codon_optimized (1881) EctSC_native (1881) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_native (2022)	1693         CCGGCA         CCGGCA         ACCGGCA         ACCGGCA         ACCGGCA         1740         ACCGGCA         ACCGGTA         CTCCAT         CAGCAT         1834         GTCCGA         1881         TACTTCC         1928         GTCCGA         1928         GTCCGA         1975         AAGTGG         AGGTGCC         2022         ATGCGCC         2069		1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCAGG CATCATTTCC CATCATCAGT 000 11 TGGCAGCAAC TGGCCGCCAC ,1850 CTGACCGATC CTCACCGATC ,1900 GACCAACACC 0 19 CCGGCATCGZ GGGGATCGZ ,1990 GAAGGCGGTC GAAGGCGGTC GAAGCTGTGC GAAGCTGTGC	,1720 GCAATGAAGCTGO GCGATGAAGCTGO 1770 TTCACCAACGGO TTCACCAACGGO 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGO 1910 CATCGGCATGATO CATCGGCATGATO 50 ,1960 ATAAGCCAGCATGATO 50 ,1960 ATAAGCCAGCAGCA CCGCAAGCACGCATO 2000 22 CTGAACACCGCATO CTGAACACCGCATO 2000 22 CTGAACACCGCATO 2000 22 CTGAACACCGCATO CCGCAAGCACGAA	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 TTCCACGGCTG TTCCACGGCTG TTCCACGGCTG CACCGCGGCGG CACCGCGGCGG Section 40 1870 1880 CCATACGCAAAC CCGTACGCCAAC Section 41 1927 GATAAGCTGCTG GACAAGCTGCTG Section 42 1974 CGGCGAAGCTGCTG Section 43 2010 2026 Section 44 2068 ATGCTGCTGAT Section 45 2015 Section 45 2016 Section 45 Section 45 Sect
(1693) EctSC_codon_optimized (1693) EctSC_native (1693) EctSC_native (1693) EctSC_codon_optimized (1740) EctSC_codon_optimized (1787) EctSC_codon_optimized (1787) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1834) EctSC_codon_optimized (1881) EctSC_codon_optimized (1881) EctSC_native (1881) EctSC_codon_optimized (1928) EctSC_codon_optimized (1928) EctSC_codon_optimized (1975) EctSC_codon_optimized (1975) EctSC_native (1975) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022) EctSC_codon_optimized (2022)	1693 CCGGCA CCGGCA ACCGGCA ACCGGCA ACCGGCA IT740 ACCGGT CTCCAT CAGCAT I884 GTTCCG GTTCCGA I881 TACTTCC TATTTCC I928 GTCCGA GTCCGA I975 _1: AAGTGG AGGTGGC ATGCGCC 2069 CGTGGA	,1700 CCAACGCA CCAACGCG CCAACGCG CGCAACAAC CGCAACAAC 1880 GCATCGGCC 1840 GCATCGGCC 1890 GGCGATAAC GGCGACAAC 1944 TCCATCCTC CCCGTCCAC 980 TGCAGGGCC 2030 AAGCTGGA AAGCTCGAC	1710 GTGGAAGCAG GTCGAGGCGG GTCGAGGCGGG CATCATTTCC CATCATCAG IGGCCGCCAG IGGCCGCCAG 1850 CTGACCGATC CTCACCGATC 1900 GACCAACACC 0 19 CCGGCATCGA GAAGGCGGTC GAAGGCGGTC 2040 AAAGCTGTGG 2040 AGGCAGGCTC	,1720 GCAATGAAGCTGG GCGATGAAGCTGG TTCACCAACGGG TTCACCAACGGG 17TCACCAACGGG 1810 ,1820 CCGGCAACCAGCA CCGGCAACCAGCA 1860 GTGTCCCGCATGG GTGTCCCGCATGGC 1910 CATCGGCATGATG 50 ,1960 ATAAGCCAGCATGATG 50 ,1960 ATAAGCCAGCAGCA CCGCAAGCACGCAG 2000 22 CTGAACACCGCAT CTGAACACGCAGCAG 2000 22 CTGAACACCGCAT CCGCAAGCACGAA CCGCCAAGCACGAA 2050 CCGCCAAGCACGAA	1739 CACGCAAGGTG GCGCGAAGGTC Section 38 1786 CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCTG CTTCCACGGCGGCG ACCACCGTGGCGG CACCGCGGCGAAC CCGTACGCCAAC CCGTACGCCAAC Section 41 1927 GATAAGCTGCTG GGCGAACGTGCT GGCGAAGCTGCT Section 43 2010 2021 CCGCCGAGAATGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CGGCCGAGTGG CCGCCGAGATGG CGGCCGAGTGGT Section 45 2115 CACCTTCTTCT

Fig. 36: Ectoine nucleotide sequence. (Continued from page 97).

(2)	116)	2116	2130	2140	2150	2162
EctSC codon optimized (2)	116)	CCTTCGAAGAA	TGGGCAT	CCAGCCAGA	TATCGTGACCC	TGTCCAAG
EctSC_native (2)	116)	GCTTCGAAGAG	ATGGGCAT	CCAGCCGGA	FATCGTCACGC	TGTCCAAG
						<ul> <li>Section 47</li> </ul>
(2)	163)	2163 2170	2180	21	90	2209
EctSC_codon_optimized (2)	163)	TCCCTGTCCGGC	CTACGGCCI	IGCCATTCG	CAATGGTGCTG	CTGCGCCA
EctSC_native (2)	163)	TCGCTGTCCGGC	CTACGGCCI	FGCCGTTCG	CCATGGTGTTG	CTGCGCCA
						- Section 48
(2)	210)	2210 22	20	2230	2240	2256
EctSC_codon_optimized (2)	210)	AGAACTGGATCA	AGTGGAAGO			CTTCCGTG
ECISC_Halive (2.	210)	AGAGCIGGACCF	AGIGGAAGU	LCCGGCGAAG	LACAACGGCAC	Section 40
(2)	257)	2257	2270	2280	2200	- Section +9
EctSC codon ontimized (2)	257)	CCAACAACCACC			ACCACCACTCC	A A C A CTTC
EctSC_codon_optimized (22	257)	GCAACAACCATC	CATTCGT(	TACGGCGGC	CACGACCAGICG	AGCACTTC
						<ul> <li>Section 50</li> </ul>
(2)	304)	2304 2310	2320	2330	0 2340	2350
EctSC codon optimized (2)	304)	TGGCAGAACGAT	GCATTCG	CAAACTCCG	IGAAGGCAAAG	GGCAAGCG
EctSC_native (2	304)	TGGCAGAACGAC	CGCGTTCGC	CCAACAGCG	IGAAGGCCAAG	GGCAAGCG
						<ul> <li>Section 51</li> </ul>
(22	351)	2351 ,2360	) 2	2370	2380	2397
EctSC_codon_optimized (22	351)	CATTGCAGATGO	GCATGCAG	CGCATCATC	CGTCGCCACGG	TCCAGACT
EctSC_native (23	351)	CATCGCCGACG	GCATGCAG	CGCATCATC	CGTCGCCACGG	CCCGGATT
						<ul> <li>Section 52</li> </ul>
(23	398)	2398	2410	2420	2430	2444
EctSC_codon_optimized (2	398)	CCCTGTTCCTG	AGGGTCGC	CGGTATGAT	GATCGGCATCT	CCTGCCCA
EctSC_native (2.	398)	CGCTGTTCCTCA	AAGGGCCGG	CGGGATGAT	GATCGGCATCA	GCTGCCCC
		2445 2450	2460	0.470	2490	- Section 53
(24	445)		2460	2470		2491
EctSC_codon_optimized (24	445)	GATGGCGAGAT	GCAGCAGO		CCACGCATIC	GAAAACGG
	.,,,)	GAIGGEGAGAII			GCCACGCCTTC	- Section 54
(2)	4071	2492 2500	25	10 2	2520	2538
EctSC codon optimized (24	492)	CCTGGTGATCGA	AACCTCC	GTGCACAC	TCCGAAGTGGT	GAAGTGCC
EctSC native (2	492)	CCTGGTGATCGA	GACCAGC	GGCGCCCAC	AGCGAAGTGGT	CAAGTGCC
	- /					
(2)	E201	2520 2	550	2560	2570	2595
(2) EctSC codon ontimized	539)	2539 2	2550	2560	2570	2585
(2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539)	2539 2 TGTGCCCACTGA	550 ATCATCTCC	2560 CGATGAGCAG	2570 GATCGATCAGG	2585 CACTGTCC CACTTTCC
2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA	2550 ATCATCTCC ATCATCAG	2560 CGATGAGCAG CGATGAGCAG	2570 GATCGATCAGG GATCGACCAGG	2585 CACTGTCC CACTTTCC — Section 56
(2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539) 539)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA 2586	2550 ATCATCTCO ATCATCAGO 2600	2560 CGATGAGCAG CGATGAGCAG 2610	2570 GATCGATCAGG GATCGACCAGG 2620	2585 CACTGTCC CACTTTCC - Section 56 2632
(2) EctSC_codon_optimized (2) EctSC_native (2) 	539) 539) 539) 539) 586)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA 2586 ATCCTGGATAAG	2550 ATCATCTCO ATCATCAGO 2600 GCATTCGO	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGAG	2570 GATCGATCAGG GATCGACCAGG 2620 IGTCCGAACAG	2585 CACTGTCC CACTTTCC — Section 56 2632 ACCGAAAA
(2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539) 586) 586) 586)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA 2586 ATCCTGGATAAG ATCCTCGACAAG	550 ATCATCTCC ATCATCAG 2600 GGCATTCG GGCCTTTGC	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA	2570 GATCGATCAGG GATCGACCAGG 2620 IGTCCGAACAG IGAGCGAGCAG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA
(2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539) 539) 586) 586) 586)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA 2586 ATCCTGGATAAG ATCCTCGACAAG	2550 ATCATCTCC ATCATCAGO 2600 GGCATTCGO GGCCTTTGO	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA	2570 GATCGATCAGG GATCGACCAGG 2620 IGTCCGAACAG IGAGCGAGCAG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57
(2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_native (2) (2)	539) 539) 539) 586) 586) 586) 586)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAC       ATCCTCGACAAC       2633     2640	550 ATCATCTC ATCATCAG 2600 GGCATTCG GGCCTTTG 2650	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA D 26	2570 GATCGATCAGG GATCGACCAGG 2620 IGTCCGAACAG IGAGCGAGCAG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 2679
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2)	539) 539) 539) 586) 586) 586) 586) 633)	2539         2           TGTGCCCACTGA         2           TCTGCCCGCTGA         2           2586         2           ATCCTGGATAAC         2           ATCCTGGACAAC         2           2633         2640           CCAGGCATCCTA         2	550 ATCATCTCC TCATCAG 2600 GGCATTCGC GGCCTTTGC 2650 AGGAAGTCC	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA D 26 CAGGAGGAGC	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA – Section 57 2679 GTGCGCAC
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 539) 586) 586) 586) 586) 586) 633) 633)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTCGACAAA       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA	550 ATCATCAG 2600 GGCATTCGG GGCCTTTGC 2650 AAGAAGTCC	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA D 26 CAGGAGGAC CAGGAGGAC	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA Section 57 2679 GTGCGCAC GTCAGAAC
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2)	(539) (539) (539) (539) (539) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (586) (539) (539) (539)	2539 2 TGTGCCCACTGA TCTGCCCGCTGA 2586 ATCCTGGATAAG ATCCTCGACAAG 2633 2640 CCAGGCATCCTA CCAGGCATCCTA	2550 ATCATCTCC 2600 GGCATTCGC GGCCTTTGC 2650 AAGAAGTCC AAGAAGTCC	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGA CCGCCGTGA CCGCCGTGA CAGGAGGAC CAGGAGGAC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2740	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 2679 GTGCGCAC GTCAGAAC – Section 58
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_native (2) (2) EctSC_codon_optimized (2)	(539) (539) (539) (539) (539) (586) (586) (586) (633) (633) (633) (633) (633)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAG       ATCCTCGACAAG       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     263	2550 ATCATCTCC 2600 GGCATTCGG GGCCTTTGG 2650 AAGAAGTCC AAGAAGTCC 90	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC 0 26 CAGGAGGACA CAGGAGGACA 2700	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG GAGCGAGCAG ATACAATGATC ATACAATGATC 2710 AACCTCCACTC	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 2679 GTGCGCAC GTCAGAAC – Section 58 2726
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2)	539) 539) 539) 586) 586) 586) 586) 586) 586) 586) 586	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAG       ATCCTCGACAAG       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     266       CCTGGCAGAATG	550 ATCATCTCC 2600 GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 GCCGAAAAGA	2560 CGATGAGCAG CGATGAGCAG 2610 CAGCAGTGAC CCGCCGTGAC 0 26 CAGGAGGACAC CAGGAGGACAC 2700 ACCGATCGCA	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA – Section 57 2679 GTGCGCAC GTCAGAAC – Section 58 2726 CCAGACCG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_native (2)	539) 539) 539) 586) 586) 586) 586) 586) 586) 586) 586	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTCGACAAC       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     26       CCTGGCAGAATC       CCTCGCCGAGTAC	550 ATCATCTCC 2600 GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 GCGAAAAG2 GCGAAAAG2	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAG CCGCCGTGA D 260 CAGGAGGAGCAC CAGGAGGAGCAC 2700 ACCGATCGC2	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA ACCGAGAA - Section 57 2679 GTGCGCAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG - Section 59
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_native (2) (2) (2) (2) (2) (2) (2) (2) (2) (2)	539) 539) 5539) 5586) 5586) 5586) 6333) 6333) 6333) 6633) 6680) 6680) 6680) 6680) 7271)	2539         2           TGTGCCCACTGA         7           TCTGCCCGCTGA         7           2586         7           ATCCTGGATAAG         7           2633         2640           CCAGGCATCCTA         7           2680         26           CCTGGCAGAATG         7           2680         26           CCTGGCAGAATG         7           2727         7	2550 ATCATCTCC ATCATCAG 2600 GGCATTCGC GGCCTTTGC 2650 AAGAAGTCC AAGAAGTCC 90 GCGAAAAG7 2740	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC D 260 CAGGAGGAC CAGGAGGAC 2700 ACCGATCGC 2750	2570 GATCGATCAGG GATCGACCAGG CGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA ACCGAGAAA - Section 57 2679 GTGCGCAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG CCAGACCG - Section 59 2773
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) (2)	539) 539) 558) 5586) 5586) 5586) 6333) 6333) 6633) 6680) 6680) 6680) 6680) 7277) 727)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAG       ATCCTCGACAAG       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     266       CCTGGCAGAATG       CCTGGCCGAGATG       CCTCGCCGAGATG       2727       GCACCTGGGATT	550 ATCATCTCC ATCATCAG 2600 GGCATTCGC GGCCTTTGC 2650 AAGAAGTCC 90 GCGAAAAG2 GCGAAAAG2 2740 CCCCCCGC	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC D 260 CAGGAGGACZ 2700 ACCGATCGCZ ACCGATCGCZ 2750 CATGCTGCT	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA - Section 57 2679 GTGCGCAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGC
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2)	539) 539) 5586) 5586) 5586) 6586) 6633) 6633) 6633) 6633) 6680) 6680) 6680) 727) 727) 727)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAG       ATCCTCGACAAG       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     268       CCTGGCAGAATG       CCTGGCCGAGATG       2727       GCACCTGGGAT       GCACCTGGGACA	2550 ATCATCAG 2600 GGCATTCGC GGCCTTTGC 2650 AGGAAGTCC AGGAAGTCC 90 GCGAAAAG 2740 CCCACCCGC GCCACCCGC	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC D 260 CAGGAGGAGC CAGGAGGAGC 2700 ACCGATCGCZ ACCGATCGCZ 2750 CATGCTGCTC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA ACCGAGAAA - Section 57 2679 GTGCGCAC GTCAGAAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGC AGGTGGGC
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 559) 5586) 5586) 6333) 6333) 6333) 6680) 6880) 6880) 727) 727) 727)	2539     2       TGTGCCCACTGA       TCTGCCCGCTGA       2586       ATCCTGGATAAG       ATCCTCGACAAG       2633     2640       CCAGGCATCCTA       CCAGGCATCCTA       2680     268       CCTGGCAGAATG       CCTGGCGAGATG       2727       GCACCTGGGAT       GCACCTGGGAT	550 ATCATCTCC ATCATCAG GGCATTCG GGCCTTTGC 2650 AAGAAGTCC AAGAAGTCC 90 GCGAAAAG GCGAAAAG 2740 CCCACCGCG	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGAC 2700 ACCGATCGCC 2750 CATGCTGCTC	2570 GATCGACCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA ACCGAGAAA - Section 57 2679 GTGCGCAC GTCAGAAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGC AGGTGGGC AGGTGGGA - Section 60
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) (2) EctSC_codon_optimized (2) (2) EctSC_codon_optimized (2) (2) (2) (2) (2) (2) (2) (2) (2) (2)	539) 539) 559) 5586) 5586) 6333) 6333) 6333) 6333) 6680) 6880) 6880) 727) 727) 727) 727)	25392TGTGCCCACTGATCTGCCCGCTGA2586ATCCTGGATAAGATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680268CCTGGCAGAATGCCTCGCCGAGTG2727GCACCTGGGACA2000 CCTGGCAGATG27742780	2550 ATCATCAG ATCATCAG 2600 GGCATTGG GGCCTTTGG 2650 AGGAAGTCC AGGAAGTCC 90 GCGAAAAGJ CGAAAAGJ CGAACCCGC AGCACCCGC 2790	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGC CAGGAGGAC 2700 ACCGATCGCC 2750 CATGCTGCTC CATGCTGCTC 2800	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGGCGAGCAG GO ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA - Section 57 2679 GTGCGCAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGCA AGGTGGGCA - Section 60 2820
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2)	539) 539) 5586) 5586) 6333) 6333) 6680) 6680) 6680) 7277) 7277) 7277) 7277) 7774)	25392TGTGCCCACTGATCTGCCCGCTGA2586ATCCTGGATAAGATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680268CCTGGCAGAATGCCTCGCCGAGATG2727GCACCTGGGACA27742780TTCTCCTTCCAC	550 ATCATCAG 2600 GGCATTCG GGCCTTTGC 2650 AAGAAGTCC 00 GCGAAAAGTCC 00 CCCACCCGC AGCACCCGC 2790 CATCACCAC	2560 CGATGAGCAG CGATGAGCAG CGACGAGCAG CCGCCGTGA CCGCCGTGA CCGCCGTGA CAGGAGGACA CAGGAGGACA 2700 ACCGATCGCC CATGCTGCTC CATGCTGCTC CATGCTGCTC 2800 CCATCTACG	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA	2585 CACTGTCC CACTTTCC - Section 56 2632 ACCGAAAA - Section 57 2679 GTGCGCAC GTCAGAAC - Section 58 2726 CCAGACCG CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGA AGGTGGGA - Section 60 2820 ACCCACAT
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_native (2) EctSC_native (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_native (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 5586) 5586) 6333) 6333) 6333) 6680) 6680) 6680) 7277) 7277) 7277) 7277) 7277) 7277) 7277)	25392TGTGCCCACTGATCTGCCCGCTGATCTGCCCGCTGA2586ATCCTGGATAAGATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTACCAGGCAGCATCCTA2680268CCTGGCAGAATGCCTCGCCGAGATG2727GCACCTGGGATG27742780TTCTCCTTCCAC	2550 ATCATCAG 2600 GGCATTCG GGCCTTTGC 2650 AAGAAGTC AAGAAGTC AGGAAGTC 2740 CCCACCCGC AGCACCCGC 2790 CATCACCAC	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGACA CAGGAGGACA 2700 ACCGATCGCC CATGCTGCTC CATGCTGCTC CATGCTGCTC CATGCTGCTC CATCTACGC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA – Section 57 2679 GTGCGCAC GTCAGAAC – Section 58 2726 CCAGACCG CCAGACCG CCAGACCG – Section 59 2773 AGGTGGGA AGGTGGGA – Section 60 2820 ACCCACAT
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2)	539) 539) 539) 5586) 5586) 633) 633) 633) 633) 6680) 6680) 6680) 7227) 7227) 7727) 7727) 7727) 7727)	25392TGTGCCCACTGATCTGCCCGCTGATCTGCCCGCCGACAA2586ATCCTCGACAAA26332640CCAGGCATCCTACCAGGCATCCTA2680268CCTGGCAGAATCCCTCGCCGAGTC2727GCACCTGGGATAGCACCTGGGACA27742780TTCTCCTTCCACTTCTCCTTCCAC	2550 ATCATCAG 2600 GGCATTCG GGCCTTTGC 2650 AAGAAGTCC AAGAAGTCC 90 CCCAACCCG AGCACCCGC 2740 CCCACCCGC 2740 CCCACCCGC 2790 CATCACCACCCCC	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGGAC 2700 ACCGATCGCC 2750 CATGCTGCTC CATGCTGCTC CATGCTGCTC CATGCTGCTC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA – Section 57 2679 GTGCGCAC GTCAGAAC GTCAGAAC CCAGACCG CCAGACCG CCAGACCG - Section 59 2773 AGGTGGGC AGGTGGGC AGGTGGGA – Section 61
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (	539) 539) 539) 5586) 5586) 5586) 633) 633) 633) 633) 6680) 680) 680) 680) 727) 727) 7727) 7727) 7727) 7727) 7727) 7727) 7727)	25392TGTGCCCACTGATCTGCCCGCTGATCTGCCCGCGCTGA2586ATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTACCAGGCAGCATCCTACCTGGCAGAATG2727GCACCTGGGACA27742780TTCTCCTTCCACTTCTCCTTCCAC28212830	2550 ATCATCTCC ATCATCAG 2600 GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 90 90 90 90 90 90 90 90 90	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGGAC 2700 ACCGATCGCC ACCGACCGCC 2750 CATGCTGCTC CATGCTGCTC CATGCTGCTC CATCTACGC	2570 GATCGATCAGG GATCGACCAGG IGAGCGAACAG IGAGCGAGCAG ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA CCGGCAGCGAG 2850	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA – Section 57 2679 GTGCGCAC GTCAGAAC GTCAGAAC CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG Section 59 2773 AGGTGGGA – Section 60 2820 ACCCACAT – Section 61 2867
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2)	539) 539) 5586) 5586) 5586) 6333) 6333) 6680) 6680) 6680) 6680) 727) 77270 7727) 77270 7727) 77270	25392TGTGCCCACTGATCTGCCCGCTGATCTGCCCGCTGA2586ATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680263CCTGGCAGAATGCCTGGCAGAATG2727GCACCTGGGACT27742780TTCTCCTTCCACTTCTCCTTCCAC28212830CCACTACCAGAA	2550 ATCATCTCC ATCATCAG 2600 GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 90 90 90 90 90 90 90 90 90	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGGAC 2700 ACCGATCGCC ACCGACCGCC 2750 CATGCTGCTC CATGCTGCTC CATGCTGCTC CATGCTGCTC CATCTACGC CCATCTACGC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG ATACAATGATC 2710 AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA CGGCAGCGAG 2850 TACTGCATCTC	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA – Section 57 2679 GTGCGCAC GTCAGAAC GTCAGAAC GTCAGAACG CCAGACCG CCAGACCG CCAGACCG CCAGACCG Section 59 2773 AGGTGGGA – Section 60 2820 ACCCACAT – Section 61 2867 CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2)	539) 539) 5586) 5586) 5586) 6333) 6333) 6680) 6680) 6680) 6680) 6680) 7277) 77270 7727) 77270 77200000000	25392TGTGCCCACTGATCTGCCCGCTGA2586ATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680263CCTCGCCGAGATG2727GCACCTGGCAGATG27742780TTCTCCTTCCAC28212830CCACCAGAACCACCAGAACCACCAGAA	2550 ATCATCTCC ATCATCAG 2600 GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC AAGAAGTCC AGGAAAAG2 CCACCCCC AGCACCCCCC AGCACCCCCC CATCACCACC CCACTTCC	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGCAC CAGGAGGAGCAC 2700 ACCGACCGCC 2750 CATGCTGCTGC CATGCTGCTGC CATGCTGCTGC CATGCTGCTGC CATGCTGCTGC CATGCTGCTGC CATGCTGCTGC 2840 GAATCCGTC	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG GAGCGAGCAG ATACAATGATC 2710 AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA CGGCAGCGAG 2850 IACTGCATCAG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA – Section 57 2679 GTGCGCAC GTCAGAAC – Section 58 2726 CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA – Section 60 2820 ACCCACAT ACGCACAT – Section 61 2867 CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (	539) 539) 5586) 5586) 5586) 6333) 6333) 6630) 6680) 6680) 6680) 6680) 7277) 7777) 7777) 7774) 7774) 7774) 7774) 7774) 7774) 7774) 7774) 7774)	2539         2           TGTGCCCACTGA         TGTGCCCGCTGA           TCTGCCCGCTGA         TCTGCCGCGACAAG           2586         ATCCTGGATAAG           ATCCTGGATAAG         ATCCTGGATAAG           2633         2640           CCAGGCATCCTA         CCAGGCATCCTA           2680         268           CCTGGCAGAATG         CCTGGCAGATG           2727         GCACCTGGGACA           GCACCTGGGACA         TTCTCCTTCCAC           2774         2780           TTCTCCTTCCAC         Z821           2821         2830           CCACTACCAGAA           2868	2550 ATCATCTCC ATCATCAG 2600 GGCATTCGC GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 GCGAAAAG7 CCCACCCGC AGCACCGCC 2790 CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACCACCACCACCACCACCACCACCACCACCACC	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGCAC CAGGAGGAGCAC 2700 ACCGATCGCC 2750 CATGCTGCTGCT CATGCTGCTGCT CATGCTGCTGCT CATGCTGCTGCT 2800	2570 GATCGATCAGG GATCGACCAGG IGTCCGAACAG IGAGCGAGCAG 60 ATACAATGATC ATACAATGATC 2710 AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACAG 0 2810 CAGGCTCCGAA CCGGCAGCGAG 2850 IACTGCATCTC IACTGCATCAG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 2773 GTGCGCAC GTCAGAACG GTCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA AGGTGGGA AGGTGGGA AGGTGGGA AGCCACAT ACGCACAT ACGCACAT Section 61 2867 CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (	539) 539) 5586) 5586) 5586) 6333) 6333) 6630) 6680) 6680) 6680) 6680) 7277) 7777) 7777) 7777) 77740) 77740000000000	2539         2           TGTGCCCACTGA         TCTGCCCGCTGA           TCTGCCCGCTGA         ATCCTGGATAAG           ATCCTCGACAAG         ATCCTCGACAAG           2633         2640           CCAGGCATCCTA         CCAGGCATCCTA           2680         263           CCTGGCAGAATG         CCTGGCAGAATG           2727         GCACCTGGGACA           GCACCTGGGACA         TTCTCCTTCCAC           2774         2780           TTCTCCTTCCAC         TTCTCCTTCCAC           2821         2830           CCACTACCAGAA         CCACTACCAGAA           2868         GCGAAATCGAAA	550 ATCATCTCC ATCATCAG 2600 GGCATTCGC GGCATTCGC 2650 AAGAAGTCC AAGAAGTCC 90 GCGAAAAG7 CCCACCCGC AGCACCCGC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACC 2790 CATCACCACCACCACC 2790 CATCACCACCACCACCACCACCACCACCACCACCACCACC	2560 CGATGAGCAG CGATGAGCAG CAGCAGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGCAC CAGGAGGAGCAC 2700 CACCGATCGCC 2750 CATGCTGCTG CATGCTGCTG CCATCTACGG CCATCTACGG CCATCTACGG CCATCTACGG CCATCTACGG CCATCTACGG CCATCGTGCTC 2880 CGATGCCAC	2570 GATCGATCAGG GATCGACCAGG GAGCGAGCAG GAGCGAGCAG GAGCGAGCAG CATACAATGATC AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACA 0 2810 CAGGCTCCGAA CGGCCAGCGAG 2850 FACTGCATCTC FACTGCATCAG 2900 CATCTACAAGA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 2779 GTGCGCAC GTCAGAAC GTCAGAAC CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA AGGTGGGA AGGCGCACAT ACGCACAT ACGCACAT CGGCAACG CGGCAACG CGGCAACG CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (	539) 539) 5586) 5586) 6333) 6333) 6333) 6630) 6680) 6680) 6680) 6680) 7277) 7777) 7777) 7777) 7777) 7774) 77774) 77740) 77740) 77740) 77740) 77740) 777740000000000	25392TGTGCCCACTGATCTGCCCGCTGA2586ATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680263CCTCGCCGAGATG2777GCACCTGGGAATG27742780TTCTCCTTCCAC28212830CCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAA	550           ATCATCTCC           ATCATCTCC           ATCATCAGO           2600           GGCATTCGC           2650           AGAAGTCC           2000           GCCATTGC           2650           AGAAGTCC           90           GCGAAAAG2           GCGAAAAG2           2740           CCCACCCGC           AGCACCCCCC           ACCACCCCCC           ACCACTCCC           ACCATTGCC           ACCATTGCC	2560 CGATGAGCAG CGATGAGCAG CGCCGTGAC CCGCCGTGAC CCGCCGTGAC CAGGAGGAGGAC CAGGAGGAGGAC 2700 CACCGATCGCC 2750 CATGCTGCTG CATGCTGCTG CCATCTACGG CCATCTACGG CCATCTACGG CCATCTACGG CCATCGCGTG CAGGGCAAG CGATGGCAAG CGATGGCAAG	2570 GATCGATCAGG GATCGACCAGG GGCGAGCAG GGO ATACAATGATC ATACAATGATC 2710 AAGGTCCACAG 2760 CAAGGACGACAG 2760 CAAGGATGATA CAAGGACGACAGA 0 2810 CAGGCTCCGAA CGGCAGCGAG 2850 TACTGCATCTC TACTGCATCTC ACTGCATCAGA 2900 GATCTACAAGA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA – Section 57 2726 GTGCGCAC GTCAGAACG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA AGGTGGGA ACCCACAT ACGCACAT Section 61 2867 CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2) EctSC_native (2)	539) 539) 5586) 5586) 633) 633) 633) 633) 633) 633) 633) 6	2539         2           TGTGCCCACTGA         TCTGCCCGCTGA           TCTGCCCGCTGA         ATCCTGGATAAG           ATCCTCGACAAG         ATCCTCGACAAG           2633         2640           CCAGGCATCCTA         CCAGGCATCCTA           CCAGGCATCCTA         2680           2680         268           CCTCGCCGAGAATG         CCTCGCCGAGTG           2727         GCACCTGGGACA           GCACCTGGGACA         CCTCCCTCCAC           2774         2780           TTCTCCTTCCAC         CCACTACCAGAA           CCACTACCAGAA         CCACTACCAGAA           CCACTACCAGAA         CCACTACCAGAA           GCGAAATCGAAAA         GCGAGATCGAAA	550 ATCATCTCC ATCATCAG 2600 GGCATTCGC 2650 AGGAAGTCC AGGAAGTCC AGGAAGTCC 2740 CCCACCCGC AGCACCCCC 2790 CATCACCACC CATCACCACCCCC ACCATTCC ACCATTGCC	2560 CGATGAGCAG CGATGAGCAG CGACGAGTGAG CCGCCGTGA CCGCCGTGA CCGCCGTGA CAGGAGGAGAC CAGGAGGAGAC 2700 CATGCTGCTG CATGCTGCTG CATGCTGCTG CCATCTACGG CCATCTACGG CCATCTACGG CCATCGTGCTG CCATCGTGCTG CCATCGTGCTG CCATCGTGGCAAG CGATGGCAAG CGACGGCAAG	2570 GATCGATCAGG GATCGACCAGG GATCGACCAG GGCGAGCAG GGO ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACACA 0 2810 CAGGCTCCGAA CGGCAGCGAG 2850 TACTGCATCTC AACTGCATCAGA 2900 GATCTACAAGA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA – Section 57 272 GTGCGCAC GTCAGAAC GTCAGAAC CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCACAT ACGCCACAT ACGCCACAT – Section 61 2807 CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG Section 62 2914
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (	539) 539) 5580) 5586) 5586) 633) 633) 633) 633) 633) 633) 680) 680) 680) 680) 680) 680) 727) 777) 777) 777) 777) 777) 7774) 7774) 7774) 7774) 821) 821) 821) 821) 821) 821) 821) 821	2539         2           TGTGCCCACTGA         TCTGCCCGCTGA           TCTGCCCGCTGA         ATCCTGGATAAG           ATCCTCGACAAG         ATCCTCGACAAG           2633         2640           CCAGGCATCCTA         CCAGGCATCCTA           2680         268           CCTGGCAGAATG         CCTGGCAGATG           2727         GCACCTGGGACA           GCACCTGGGACA         CTCTCCTCCAC           2774         2780           TTCTCCTTCCAC         CCACTACCAGAA           CCACTACCAGAA         CCACTACCAGAA           2821         2830           CCACTACCAGAA         GCGAAATCGAAA           2868         GCGAAATCGAAA           GCACTACCAGAA         2915	550 ATCATCTCC ATCATCAG 2600 GGCATTCG 2650 AGGAAGTCC AGGAAGTCC AGGAAGTCC 200 CCACCCCGC AGCAACCCCC 2790 CCACCCCCC AGCACCCCCC 2790 CATCACCACCCCC AGCACTTCC ACCATTCCC 2880 ACCATTGCC 2930	2560 CGATGAGCAG CGATGAGCAG CGACGAGTGAC CGCCGTGAC CCGCCGTGAC CAGGAGGAGGAC CAGGAGGAGAC 2700 CACCGATCGCC 2750 CATGCTGCTG CATGCTGCTG CCATCTACGC CCATCTACGC CCATCTACGC CCATCTACGC CCATCGCGGTG CAGGGGCAAC CGACGGCAAC CGACGGCAAC	2570 GATCGATCAGG GATCGACCAGG GATCGACCAG GGCGAGCAG GGO ATACAATGATC ATACAATGATC 2710 AAGGTCCACAG 2760 CAAGGACGACAG 2760 CAAGGATGATA CAAGGACGACAGA 0 2810 CAGGCTCCGAA CGGCAGCGAG 2850 TACTGCATCTC FACTGCATCAGA 2900 GATCTACAAGA GATCTACAAGA	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAA – Section 57 272 GTGCGCAC GTCAGAAC GTCAGAAC CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA AGGTGGGA Section 60 2820 ACCCACAT ACGCACAT – Section 61 2867 CGGCAACG
(2) EctSC_codon_optimized (2) EctSC_native (2) EctSC_codon_optimized (2)	539) 539) 5580) 5586) 5586) 6333) 6333) 6333) 6333) 6333) 6680) 6880) 6880) 6880) 7277) 7774) 7774) 7774) 7774) 7774) 7774) 8211] 8211) 8211) 8211) 8211) 8211) 8211] 82	25392TGTGCCCACTGATCTGCCCGCTGA2586ATCCTCGACAAG26332640CCAGGCATCCTACCAGGCATCCTA2680268CCTCGCCGAGATG2777GCACCTGGCAGATTGCACCTGGGACACAT27742780TTCTCCTTCCAC28212830CCCCACACAGAACCACTACCAGAACCACTACCAGAACCACTACCAGAAACCACTACCAGAAA2868GCGAAATCGAAA29152920GGCACCCTGTACC	550 ATCATCTCC ATCATCAG 2600 GGCATTCG GGCATTCG 2650 AGGAAGTCC AGGAAGTCC 90 GCGAAAAGT 2740 CCCACCCGC AGCAACGCG 2790 CATCACCAC 290 CATCACCAC 2900 CATCACTTCC 290 CATCACTTCC 290 CATCACTTCC	2560 CGATGAGCAG CGATGAGCAG CGACGAGTGAG CCGCCGTGA CCGCCGTGA CAGGAGGAGAC 2700 ACCGATCGCC 2750 CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCTG CATGCTGCCG CATGCCGCAAG CGACGGCAAG CGACGGCAAG 2940 AAAAGCACG	2570 GATCGATCAGG GATCGACCAGG GATCGACCAG GGCGAGCAG GGO ATACAATGATC ATACAATGATC 2710 AAGGTGCACTC AAGGTCCACAG 2760 CAAGGATGATA CAAGGACGACAG 2760 CAAGGATGATA CAAGGACGACAGA 2760 CAAGGACGACAGA 2760 CAAGGACGACAGA 2760 CAAGGACGACAGA 2760 CAAGGACGACAGA 2760 CAAGGACGACAGA 2760 CAAGGACGACAGA 2850 FACTGCATCTC FACTGCATCTC ACGGCCACAGA 2900 GATCTACAAGA 2950 ATGAACACCTG	2585 CACTGTCC CACTTTCC – Section 56 2632 ACCGAAAA ACCGAGAAA – Section 57 2679 GTGCGCAC GTCAGAACG CAGACGC CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG CCAGACCG AGGTGGGA – Section 60 2820 ACCCACAT ACGCCACAT ACGCACAT – Section 61 2867 CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG CGGCAACG Section 62 2914 TCGAACCA – Section 63 2961

Fig. 36: Ectoine nucleotide sequence. (Continued from page 98).

	(2962)	2962	2970	2	980	2990		3008
EctSC_codon_optimized	(2962)	TGGC	TCCGAAGA	TATGAAG	CTGGCC	TGCGTGTT	CAACCC	ACCACTGA
EctSC_native	(2962)	TGGC	AGCGAAGA	CATGAAG	CTGGCC	TGCGTCT	CAACCC	GCCGCTCA
								– Section 65
<b>F 100</b> 1 11 1	(3009)	3009	30	20	3030	30	40	3055
EctSC_codon_optimized	(3009)	ACGG	TCGCGAAG	IGCACGA	TGAATC	CGGTGTGTG	l'ACCCAC'	I'GGAAGCA
EctSC_native	(3009)	ACGG	GCGCGAAG	IGCATGA	CGAAAG	CGGCGTCT	rareere.	Section 66
	(2056)	3056		3070	3(	180	3000	- Section 00 3102
EctSC codon ontimized	(3050)	GAAA	CCGTGTAA	TACCGGA	GCTCGA			
EctSC_codon_optimized	(3056)	GAAA	CCGTCTGA	TACCGGA	GCTCGA	TCACGAA	TCCAGG/	AGGACATA
	(0000)							- Section 67
	(3103)	3103	3110	312	20	3130		3149
EctSC_codon_optimized	(3103)	CAAT	GCAGGCAG	ATCTGTA	CCCATC	CCGTCAAC	GAAGATCA	AGCCATCC
EctSC_native	(3103)	CAAT	GCAAGCCG	ACCTGTA	TCCCTC	GCGCCAG	GAAGACC	AGCCCAGC
								<ul> <li>Section 68</li> </ul>
	(3150)	3150	316	0	3170	3180	)	3196
EctSC_codon_optimized	(3150)	TGGC	AAGAACGC	CTGGATC	CAGTGG	TGTACCGO	CTCCGAT	CTGGAAAA
EctSC_native	(3150)	TGGC	AGGAACGC	CTGGATC	CGGTCG	TCTACCGO	CAGCGAC	CTGGAGAA
								- Section 69
<b>E 100</b> 1 11 1 1	(3197)	3197		3210	322	20	3230	3243
EctSC_codon_optimized	(3197)	CGCA	CCAATCGC	AGCCGAA	CTGGTG	GAACGCT	L'CGAACG(	CGACGGCT
EctSC_native	(3197)	TGCG	CCGATCGC	GCAGAG	CIGGIC	GAACGCI	rcgaacg	Section 70
	(2244)	2244	3250	3260		3270	2280	- Section 70
EctSC codon ontimized	(3244)	3244 ACCT		3200 277777	CTTCTC	SZIU CCCACAT(		3290
EctSC_couon_optimized	(3244)	ACCT	GGTCATCC	CAATCT	GTTCAG	CGCCGAC	JAAGIGG(	COTGTTT
	(3211)		001011100		0110110	000001100		– Section 71
	(3291)	3291	.3300		3310	3320		3337
EctSC codon optimized	(3291)	CGCG	CAGAACTG	GAACGCA	TGCGCC	AGGATCC	AGCAGTG	GCAGGCTC
EctSC_native	(3291)	CGCG	CCGAACTC	GAGCGCA	TGCGCC	AGGACCC	GCCGTC	GCCGGTTC
								- Section 72
	(3338)	3338		3350	3360	2	3370	3384
EctSC_codon_optimized	(3338)	CGGC	AAGACCAT	CAAAGAA	CCAGAT	TCCGGTG	CAATCCG	CTCCGTGT
EctSC_native	(3338)	CGGC	AAGACCAT	CAAGGAA	CCCGAC	AGCGGTGC	CGATCCG	CTCGGTGT
	(3385)	3385	3390	3400		3410	3420	3431
EctSC_codon_optimized	(3385) (3385)	3385 TCGC	3390 AATCCACA	3400 AGGATAA	CGAACT	3410 GTTCGCTC	3420 CGCGTGGG	3431 CAGCAGAT
EctSC_codon_optimized EctSC_native	(3385) (3385) (3385)	<u>3385</u> TCGC TCGC	3390 AATCCACA CATCCACA	3400 AGGATAA AGGACAA	CGAACT CGAGCT	3410 GTTCGCTC GTTCGCCC	3420 CGCGTGGC CGCGTCGC	3431 CAGCAGAT CAGCCGAC
EctSC_codon_optimized EctSC_native	(3385) (3385) (3385)	3385 TCGC TCGC	3390 AATCCACA CATCCACA	3400 AGGATAA AGGACAA	CGAACT CGAGCT	3410 GTTCGCTC GTTCGCCC	3420 CGCGTGGC CGCGTCGC	3431 CAGCAGAT CAGCCGAC – <b>Section 74</b>
EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432)	3385 TCGC TCGC 3432	3390 AATCCACA CATCCACA 3440	,3400 AGGATAA AGGACAA 3	CGAACT CGAGCT 450	3410 GTTCGCTC GTTCGCCC ,3460	3420 CGCGTGGG CGCGTCGG	3431 CAGCAGAT CAGCCGAC – Section 74 3478
EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432)	3385 TCGC TCGC 3432 GAAC	3390 AATCCACA CATCCACA 3440 GCACCGCA	3400 AGGATAA AGGACAA 3 GGTATCG	CGAACT CGAGCT 450 CACGCT	3410 GTTCGCTC GTTCGCCC ,3460 TCATCCTC	3420 CGCGTGGC CGCGTCGC	3431 CAGCAGAT CAGCCGAC – Section 74 3478 GATCTGTA
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432)	3385 TCGC TCGC 3432 GAAC GAGC	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG	CGAACT CGAGCT 450 CACGCT CCCGCT	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTC TCATCCTT	3420 CGCGTGGC CGCGTCGC GGGTGGCC GGCGGGCC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA GACCTGTA Corting 75
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432)	3385 TCGC TCGC 3432 GAAC GAGC	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG	CGAACT CGAGCT 450 CACGCT CCCGCT 2500	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTC TCATCCTC	3420 CGCGTGGG CGCGTCGG GGCGGCG 10	3431 CAGCAGAT CAGCCGAC - Section 74 GATCTGTA GACCTGTA - Section 75 2525
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479)	3385 TCGC TCGC 3432 GAAC GAGC 3479	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC 34	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG	CGAACT CGAGCT 450 CACGCT CCCGCT 3500	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTC TCATCCTC 35	3420 CGCGTGGC CGCGTCGC GGGTGGCC TGGCGGCC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA GACCTGTA - Section 75 3525
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC 34 CACCAGTC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 190 CCGCATG CCGCATG	CGAACT CGAGCT 450 CACGCT CCCGCT 3500 AACTTC	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTC TCATCCTT 35 AAGCCAGC	3420 CGCGTGGC CGCGTCGC GGGTGGCC TGGCGGCC 10 CCTTCACC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 3ATCTGTA 3ACCTGTA - Section 75 3525 CGGCAAAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC 34 CACCAGTC CATCAGTC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 190 CCGCATG 3CGAATG	CGAACT CGAGCT 450 CACGCT CCCGCT 3500 AACTTC AACTTC	3410 GTTCGCCC GTTCGCCC TCATCCTC TCATCCT 35 AAGCCAGC AAGCCCGC	3420 CGCGTGGC CGCGTCGC GGGTGGCC TGGCGGCC 10 GCTTCACC	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA 3ACCTGTA - Section 75 3525 CGGCAAAG - Section 76
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3479)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC 34 CACCAGTC CATCAGTC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG GGCATCG CCGCATG 3CGAATG 3540	CGAACT CGAGCT 450 CACGCT CCCGCT 3500 AACTTC AACTTC	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTC TCATCCTT 35 AAGCCAGC AAGCCCGC	3420 CGCGTGGC CGCGTCGC GGGTGGCC TGGCGGCC 10 GCTTCACC 3560	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA 3ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAGG - Section 76 3572
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3479) (3526) (3526)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT	3390 AATCCACA CATCCACA 3440 GCACCGCA GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG GGCATG 3CGCATG 3540 ACTCCGA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC 33 TTTCGA	3410 GTTCGCT GTTCGCC 3460 TCATCCTT 35 AAGCCAG( AAGCCCG( 550 AACCTGG(	3420 CGCGTGGC CGCGTGGC TGGCGGCC 10 GCTTCACC 3560 CACATCG2	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA 3ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3479) (3526) (3526) (3526)	3385 TCGC TCGC GAAC GAAC GAGC GAGC GAGC GAGC	3390 AATCCACA CATCCACA 3440 GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC.	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG GCGATG GCGATG GCGATG 3540 ACTCCGA ACTCCGA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC 330 TTTCGA	3410 GTTCGCT GTTCGCC 3460 TCATCCTT 35 AAGCCAGC AAGCCCGC 550 AACCTGGC GACCTGGC	3420 CGCGTGGC CGCGTGGC TGGCGGCG TGGCGGCG CGCTTCACC 3560 CACATCG2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA 3ACCTGTA SACCTGTA - Section 75 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3526) (3526) (3526)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT	3390 AATCCACA CATCCACA 3440 GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC.	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG GCGCATG GCGATG 3540 ACTCCGA ACTCCGA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC 38 TTTCGA	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTT 35 AAGCCAGC AAGCCGC 550 AACCTGGC GACCTGGC	3420 CGCGTGGC CGCGTGGC TGGCGGCC 10 GCTTCACC 3560 CACATCG2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA SATCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 AGGACGGC - Section 77
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (35273)	3385 TCGC TCGC GAGC GAGC GAGC GAGC GAGC GAGC	3390 AATCCACA CATCCACA 3440 GCACCGCC 34 CACCAGTC CACCAGTC CATCAGTC CTACTGGC 3580	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 3 CCGCATG 3 CCGCGCATG 3 CCGCGCA 3 CCGCGCA 3 CCGCGCA 3 CCGCGCA 3 CCGCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCGCA 3 CCCCGCGCGCCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCGCA 3 CCCCGCCGCA 3 CCCCGCGCA 3 CCCCGCGCGCCGCGC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC 38 TTTCGA TTTCGA	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTT 355 AAGCCAGC AAGCCGGC 550 AACCTGGC GACCTGGC	3420 CGCGTGGC CGCGTGGC GGCGGCC 10 GCTTCACC GCTTCACC 3560 CACATCG2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAGG CGGCAAGG - Section 76 3572 AAGATGGC - Section 77 3619
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3573) (3573)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG CGTG 3526 AATT AGTT 3573 ATGC	3390 AATCCACA CATCCACA 3440 GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC CTACTGGC 3580 CACGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG CCGCATG 3CGAATG 3540 ACTCCGA ACTCCGA 3550 CGCTGCC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC 38 TTTCGA TTTCGA 00 TGTCCT	3410 GTTCGCT GTTCGCC 3460 TCATCCTC 355 AAGCCAGC AACCTGGC GACCTGGC 3600 GCTCCATC	3420 CGCGTGGC CGCGTGGC IGGCGGCC 10 SCTTCACC CACATCGA CACATCGA	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC - Section 77 3619 ACCGATAA
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3573) (3573) (3573)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC	3390 AATCCACA CATCCACA 3440 GCACCGCC GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC CTACTGGC 3580 CACGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 3CGAATG 3CGAATG 3540 ACTCCGA ACTCCGA 3550 CGCTGCC CGCTGCC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 38 TTTCGA TTTCGA TTTCGA TGTCCT	3410 GTTCGCC GTTCGCCC 3460 TCATCCTC CATCCTC 355 AAGCCAGC AACCTGGC GACCTGGC 3600 GCTCCATC	3420 CGCGTGGC CGCGTCGC GGCGGCC 10 GCTTCACC CACATCG2 CACATCG2 CCTGCTG2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC - Section 77 3619 ACCGATAA
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3576) (3526) (3526) (3526) (3573) (3573) (3573)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC	3390 AATCCACA CATCCACA GCACCGCA GCACCGCCA CACCAGTC CACCAGTC CTACTGGC CTACTGGC 3580 CACGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 3CGAATG 3540 ACTCCGA ACTCCGA ACTCCGA 3550 CGCTGCC CGCTGCC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33 TTTCGA TTTCGA TTTCGA TTTCGA	3410 GTTCGCT GTTCGCC 3460 TCATCCTC TCATCCTC 355 AAGCCAGC AAGCCCGC 550 GACCTGGC 3600 GCTCCATC	3420 CGCGTGGC CGCGTGGC CGCGGCG CGGCGGCC 10 GCTTCACC CCTTCACC CACATCG2 CCTGCTG2 CCTGCTG2	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACGGC - Section 77 3619 ACCGATAA ACCGACAA - Section 78
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3576) (3526) (3526) (3526) (3573) (3573) (3573) (3573) (3573)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC ATGC 3620	3390 AATCCACA CATCCACA GCACCGCA GCACCGCCA CACCAGTC CACCAGTC CTACTGGC CTACTGGC CTACTGGC CTACTGGC CACGCATG CACGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG CCGCATG 3540 ACTCCGA ACTCCGA ACTCCGA 3550 CGCTGCC CGCTGCC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33 TTTCGA TTTCGA TTTCGA TGTCCT 3640	3410 GTTCGCT GTTCGCC 3460 TCATCCTC TCATCCTC 355 AAGCCAGC AAGCCCGC 550 AACCTGGC GACCTGGC GCTCCATC GCTCCATC	3420 CGCGTGGC CGCGTGGC CGCGGCGC IGGCGGCC 10 GCTTCACC GCTTCACC CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACGGC - Section 77 3619 ACCGATAA - Section 78 3666
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3576) (3526) (3526) (3526) (3573) (3573) (3573) (3573) (3620) (3620) (3620)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG CGTG 3526 AATT AGTT 3573 ATGC ATGC 3620 CGAA	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA CACCAGTC CACCAGTC CTACTGGC CTACTGGC CTACTGGC CTACTGGC CACGCATG CACGCATG CACGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 2CGCATG 3540 ACTCCGA ACTCCGA 3540 3540 CGCTGCC CGCTGCC CGCTGCC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 360 TGTCCT 3640 CTGATG	3410 GTTCGCT GTTCGCC 3460 TCATCCTC TCATCCTC 355 AAGCCAGC AAGCCCGC 550 AACCTGGC GACCTGGC GCTCCATC GCTCCATC CTGATGCC	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC 10 GCTTCACC GCTTCACC CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACGGC - Section 77 3619 ACCGATAA - Section 78 3666 CACAAGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3573) (3526) (3526) (3526) (3526) (3573) (3573) (3573) (3620) (3620) (3620)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC ATGC 3620 CGAA CGAG	3390 AATCCACA CATCCACA GCACCGCA GCACCGCA CACCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CACGCATG CGCGCATG 363 CCACACAA	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 2 GGCATG 3 CCGCATG 3 CGCATGC 3 CGCTGCC CGCTGCC 0 CGGTCCA CGGCCCG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 38 TTTCGA TTTCGA TGTCCT 3640 CTGATG CTGATG	3410 GTTCGCT GTTCGCC 3460 TCATCCTT 35 AAGCCAGC AAGCCAGC 550 AACCTGGC GCTCCATC GCTCCATC CTGATGCC	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC 10 GCTTCACC 3560 CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCCGCTCC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACAGGC - Section 77 3619 ACCGATAA - Section 78 3666 CACAAGC CACAAGC Saction 70
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3576) (3526) (3526) (3526) (3526) (3573) (3573) (3573) (3620) (3620) (3620)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC 3620 CGAA CGAG	3390 AATCCACA CATCCACA GCACCGCA GCACCGCA GCACCGCC CACCAGTC CACCAGTC CTACTGGC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCCACAA	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG CCGCATG CCGCATG CGCATCGA ACTCCGA ACTCCGA CGCTGCC CGCTGCC CGCTGCC CGCTCCA CGGTCCA CGGCCCG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33 TTTCGA TTTCGA TTTCGA TGTCCT 3640 CTGATG CTGATG	3410 GTTCGCC GTTCGCC 3460 TCATCCTC TCATCCTC 355 AAGCCAGC AAGCCCGC 550 AACCTGGC GCTCCATC GCTCCATC CTGATGCC CTGATGCC	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC 10 GCTTCACC 3560 CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCGGCTCC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACAAGC - Section 77 3619 ACCGATAA - Section 78 3666 CACAAGC CACAAGC - Section 79
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3573) (3573) (3573) (3620) (3620) (3667) (3667)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC ATGC 3620 CGAA CGAG 3667	3390 AATCCACA CATCCACA GCACCGCA GCACCGCA CACCAGTC CACCAGTC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCGCATG	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 90 CCGCATG 3540 ACTCCGA ACTCCGA ACTCCGA CGCTGCC CGCTGCC CGCTGCC CGCTGCC CGGTCCA CGGTCCA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33 TTTCGA TTTCGA TTTCGA GO CTGATG CTGATG CTGATG CTGATG	3410 GTTCGCT GTTCGCC 3460 TCATCCTC TCATCCT 35 AAGCCAGC AAGCCCGC 550 AACCTGGC GCTCCATC GCTCCATC CTGATGCC CTGATGCC 10 CCCCACAA	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC 10 GCTTCACC 3560 CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCGGCTCC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA ACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACAGC - Section 77 3619 ACCGATAA - Section 78 3666 CCACAAGC CACAAGC - Section 79
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3620) (3667) (3667)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC 3573 ATGC 3620 CGAA CGAG 3667 ACTA	3390 AATCCACA CATCCACA GCACCGCA GCACCGCA GCACCGCC CACCAGTC CATCAGTC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCGCACAA CGCACACAA	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG 90 CCGCATG 3540 ACTCCGA ACTCCGA ACTCCGA CGCTGCC CGCTGCC CGCTGCC CGCTGCC CGGTCCA 3680 GCGTTGG GCGTCGC	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33500 TTTCGA TTTCGA TTTCGA TTTCGA GO CGCAAC AGCCAC	3410 GTTCGCC GTTCGCC 3460 TCATCCTC TCATCCTC 35 AAGCCAGC AAGCCCGC 550 AACCTGGC GCTCCATC GCTCCATC CTGATGCC CTGATGCC CCCAGAAA ACCGGAAA	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC 10 GCTTCACC 3560 CACATCG2 CACATCG2 CACATCG2 CCTGCTG2 CCTGCTG2 CCGGCTCC 3700 AACCACT2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA GACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACAGGC - Section 77 3619 ACCGATAA - Section 78 3666 CCACAAGC CACAAGC - Section 77 3713 ACGAAAAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3620) (3667) (3667) (3667) (3667)	3385 TCGC TCGC GAGC GAAC GAAC GAGC GAGC GAGC	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA 34 CACCAGTC CATCAGTC CTACTGGC CTACTGGC CGCGCATG 363 CCACACAA CCGCACAAA CCGCACAAA	3400 AGGATAA AGGACAA 3 GGTATCG GGCATCG GGCATCG CGCATG GCGATG CGCTCCGA CGCTGCC CGCTGCC 0 CGGTCCA CGGCCCG 3680 GCGTCGG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 33 TTTCGA TTTCGA TTTCGA 00 TGTCCT 3640 CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC	3410 GTTCGCT GTTCGCC 3460 TCATCCTT 35 AAGCCAG( AAGCCCGC 550 AACCTGGC GACCTGGC GCTCCATC 3650 CCTGATGCC CTGATGCC 00 CCCAGAAA	3420 CGCGTGGC CGCGTGGC CGCGTGGC CGGCGGCC CGGCTCACC 3560 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCCGGCTCC 3700 AACCACT2	3431 CAGCAGAT CAGCCGAC - Section 74 3ATCTGTA GACCTGTA SACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG - Section 76 3572 AAGATGGC AGGACAGGC - Section 77 3619 ACCGACAAG CCACAAGC CACAAGC CACAAGC - Section 79 3713 ACGAAAAG - Section 80
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3620) (3667) (3667) (3667) (3667) (3667) (3667)	3385 TCGC TCGC GAGC GAAC GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3526 AATT AGTT 3527 ACGAA CGAA CGAA CGAA CGAA CGAA CGAA CG	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA 34 CACCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CACCACA CGCGCATG 3580 CACGCATG CGCGCATG CGCGCATG CGCGCACAA CGCGCACAA	3400 AGGATAA AGGACAA GGCATCG GGCATCG GGCATCG CGCATG GCGATG CGCATGGA ACTCCGA ACTCCGA ACTCCGA CGCTGCC CGCTGCC 0 CGGTCCA CGGCCCG 3680 GCGTCGG GCGTCGG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 38 TTTCGA TTTCGA TTTCGA TTTCGA CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTT 35 AAGCCAG( AAGCCCGC 550 AACCTGGC GACCTGGC GCTCCATC 3660 GCTCCATC CTGATGCC 00 CCCAGAAA ACCCGAAA	3420 CGCGTGGC CGCGTGGC CGCGTGGC CGGCGGCC CGGCTCACC 3560 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCCGGCTCC 3700 AACCACT2 AATCACT2	3431 CAGCAGAT CAGCCGAC - Section 74 347C SATCTGTA 362CTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAAG CGGCAAG - Section 77 3619 ACCGATAA - Section 78 3666 CACAAGC CACAAGC CACAAGC - Section 79 3713 ACGAAAAG ACGACAAAG - Section 80 3760
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3573) (3620) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (36714) (3714)	3385 TCGC TCGC GAGC GAGC GAGC GAGC GAGC GAGC	3390 AATCCACA CATCCACA GCACCGCC GCACCGCC 34 CACCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCGCATG CGTGCGCT 3720 TGCGCAAG	3400 AGGATAA AGGACAA GGCATCG GGCATCG GCCATCG CCGCATG CCGCATG CCGCATG CGCTCCGA CGCTGCC CGCTGCC CGCTGCC CGCTGCC CGGTCCG CGGTCCG CGGTCGG CGGTCGG CGGTCGG CGGTCGG CGGTCGG CGGTCGG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC 38 TTTCGA TTTCGA TTTCGA 0 CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC	3410 GTTCGCTC GTTCGCCC 3460 TCATCCTT 35 AAGCCAGC AAGCCCGC 550 AACCTGGC GACCTGGC GCTCCATC 3660 GCTCCATC 3660 CCTGATGCC 0 CCCAGAAA ACCCGGAAA 3740 TCCCTGAT	3420 CGCGTGGC CGCGTGGC CGCGTGGC CGGCGCC CGGCTCACC 3560 CACATCG2 CACGCTCACC CCTGCTG2 CACGCTCC CCGGCTCC 3700 AACCACT2 AATCACT2 3750 CCAGAAC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAGG - Section 77 3619 ACCGATAA - Section 78 3666 CACAAGC CACAAGC CACAAGC CACAAGC - Section 79 3713 ACGAAAAG ACGAAAAAG - Section 80 3760 CCCTGTC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3714) (3714) (3714)	3385 TCGC TCGC GAGC GAGC GAGC GAGC GAGC GAGC	3390 AATCCACA CATCCACA GCACCGCC GCACCGCC 34 CACCAGTC CATCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCGCATG CGTGCGCT CGTGCGCT CGTGCGCT CGTGCGCT	3400 AGGATAA AGGACAA GGTATCG GGCATCG CCGCATG CCGCATG CCGCATG CCGCATG CCGCTGCC CGCTGCC CGCTGCC CGCTGCC CGCTGCC CGGTCCG 3680 GCGTCGG CGGTCGG CGGTCGG CGGTCGG CGAAAA CAGAAA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC AACTTCGA TTTCGA TTTCGA TTTCGA CGCAC CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC TCGGCA	3410 GTTCGCT GTTCGCC 3460 TCATCCTT 35 AAGCCAGC AAGCCAGC 550 AACCTGGC GACCTGGC 3600 GCTCCATC GCTCCATC 3650 CTGATGCC 0 CCCAGAAA ACCCGGAAA ACCCGGAAA 3740 TCCCTGAT	3420 CGCGTGGC CGCGTGGC CGCGTGGC TGGCGGCC TGGCGGCC TO SCTTCACC 3560 CACATCG CACGCTCG CACGCTCC 3700 AACCACT 3750 CAGGACC CAGAAC	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA - Section 75 3525 CGGCAAAG CGGCAAGG CGGCAAGG CGGCAAGG - Section 77 3619 ACCGATAA CGGACAAGC CACAAGC CACAAGC CACAAGC CACAAGC CACAAGC CACAAGC CACAAGC Section 79 3713 ACGAAAAG ACGAAAAG - Section 80 3760 CCCTGTC AGCCTGAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3526) (3527) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3714) (3714)	3385 TCGC TCGC GAGC GAGC GAGC GAGC GAGC GAGC	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA CACCAGTC CATCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CTACTGGC CCACCATG CGCGCATG CGCGCATG CGCGCACA CGTGCGCT CGTGCGCT CGTGCGCT CGTGCGCACA	3400 AGGATAA AGGACAA GGTATCG GGCATCG CCGCATG CCGCATG CCGCATG ACTCCGA ACTCCGA ACTCCGA CGCTGCC CGCTGCC CGCTGCC CGGTCCA CGGCCCG 3680 GCGTCGG GCGTCGG GCGTCGG 3730 CAAGAAA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC AACTTC GACTTCGA TTTCGA TTTCGA TTTCGA CTGATG CTGATG CTGATG CTGATG CTGATG CTGATG CTGATG CTGATG CTGATG	3410 GTTCGCTG GTTCGCCC 3460 TCATCCTT 35 AAGCCAGC AAGCCGGC 550 AACCTGGC GACCTGGC 3600 GCTCCATCG CTGATGCC CTGATGCC 00 CCCAGAAA ACCCGAAA ACCCGAAA 3740	3420 CGCGTGGC CGCGTGGC CGCGTGGC CGGCGGCC 10 CCTTCACC 3560 CACATCG2 CCTGCTG2 CCTGCTG2 CCTGCTG2 CCGGCTCC 3700 AACCACT2 AATCACT2 3750 CCAGAAC2	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA ACTGTA - Section 75 3525 CGGCAAAG CGGCAAGG - Section 76 3619 ACCGATAA ACCGATAA CCGATAA CCGACAAGC CACAAGC CACAAGC CACAAGC CACAAGC CACAAGC Section 79 3713 ACGAAAAG - Section 81
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(3385) (3385) (3385) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3526) (3526) (3526) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (3667) (36714) (3714) (3714) (3761)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC AGGA CGAA CGAA CGAA CGAA CGAA CGAA	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA CACCAGTC CATCAGTC CATCAGTC CTACTGGC CTACTGGC CTACTGGC CTACTGGC CGCGCATG CGCGCATG CGCGCATG CGCGCACAA CCGCACAAA CCGCGCACAA CGTGCGCT CGTGCGCT CGTGCGCT CGTGCGCAAG TGCGCAAG	3400 AGGATAA AGGACAA GGCATCG GGCATCG CCGCATG CCGCATG CCGCATG CCGCTCCGA CCGCTGCC CCGCTGCC CCGCTGCC CCGGTCCA CGGCCCG CCGGTCCA CGGCCCG CCGGTCGG CCGGTCGG CCGGTCGG CCGGTCGG CCGGTCGG	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC AACTTC GAC TTTCGA TTTCGA TTTCGA CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC TCGGCA 3780	3410 GTTCGCT GTTCGCCC 3460 TCATCCTT 35 AAGCCAGC AAGCCAGC AAGCCAGC 550 AACCTGGC 3600 GCTCCATCG GCTCCATCG CTGATGCC 00 CCCAGAAA ACCGGAAA ACCGGAAA 3740 TCCCCGAC 3790	3420 CGCGTGGC CGCGTGGC CGCGTGGC CGGCGGCC 10 SCTTCACC CCTCCTCAC CACATCGA CCCGGCTCC CCGGCTCC 3700 AACCACTA ATCACTA 3750 CAGAACA	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA AGACCTGTA - Section 75 3525 CGGCAAAG CGGCAAGG - Section 76 3572 AGGACGGC - Section 77 3619 ACCGATAA AGGACGGC - Section 77 3619 ACCGATAA CGACAAGC CACAAGC CACAAGC CACAAGC - Section 79 3713 ACGAAAAG - Section 80 3760 FCCCTGAC - Section 81 3807
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(3385) (3385) (3385) (3432) (3432) (3432) (3432) (3479) (3479) (3479) (3479) (3526) (3526) (3526) (3526) (3526) (3526) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3573) (3620) (3667) (3667) (3667) (3667) (3667) (3667) (36714) (3714) (3761) (3761) (3761)	3385 TCGC TCGC 3432 GAAC GAGC 3479 CGTG CGTG 3526 AATT AGTT 3573 ATGC ATGC 3620 CGAA CGAG 3667 ACTA ACTA 3714 TCCC TCCC 3761 CGAA	3390 AATCCACA CATCCACA GCACCGCCA GCACCGCCA CACCAGTC CATCAGTC CATCAGTC CTACTGGCA CTACTGGCA CACCAGTA CGCGCATG CGCGCATG CGCGCATG CGTGCGCT CGTGCGCT CGTGCGCT CGTGCGCAG TGCGCAAG 3770 CTGGCATC	3400 AGGATAA AGGACAA GGTATCG GGCATCG CGCATG CGCATG CGCATGGA ACTCCGA CGCTGCC CGCTGCC CGCTGCC CGGTCCA CGGTCCA CGGTCGG CGGTCGG 3730 CAAGAAA CAGGAGA	CGAACT CGAGCT 450 CACGCT 3500 AACTTC AACTTC AACTTC AACTTC GAC TTTCGA TTTCGA TTTCGA CTGATG CTGATG CTGATG CTGATG CGCAAC AGCCAC TCGGCA TCGGCA 3780 GGCATC	3410 GTTCGCT GTTCGCC GTTCGCCC 3460 TCATCCTT 35 AAGCCAGC AAGCCAGC AACCTGGC 3600 GCTCCATC GCTCGATGCC CTGATGCC 00 CCCAGAAA ACCGGAAA 3740 TCCCTGAT CCCCGAC 3790 GATTGCGC	3420 CGCGTGGC CGCGTGGC GGCGGCC IGGCGGCC IO SCTTCACC CCTCTCAC CACATCG7 CACATCG7 CACGCTC CCGGCTCC 3700 AACCACT7 AATCACT7 3750 ICAGAAC7 CAGAAC7	3431 CAGCAGAT CAGCCGAC - Section 74 3478 GATCTGTA AGACCTGTA - Section 75 3525 CGGCAAAG CGGCAAAG CGGCAAGG - Section 77 3619 ACCGATAA AGGACGAC - Section 77 3619 ACCGATAA CCGACAAGC CACAAGC CACAAGC CACAAGC - Section 79 3713 ACGAAAAG CGACAAGC - Section 79 3713 ACGAAAAG CGACAAGC - Section 81 3807 ICCCAGCAG

Fig. 36: Ectoine nucleotide sequence. (Continued from page 99).

	(3808)	3808		3820		3830		3840	3854
EctSC_codon_optimized	(3808)	GCTC	CGTGGT	GTTCT	TCGATT	GCAACA	CCATO	GCACGGCT	CCAACGGC
EctSC_native	(3808)	GCAG	CGTGGT	GTTCT	TCGACT	GCAACA	CCATO	GCACGGCT	CCAACGGC
									<ul> <li>Section 83</li> </ul>
	(3855)	3855	3860	í.	3870	388	30	3890	3901
EctSC_codon_optimized	(3855)	AACA	TCACCC	CATCCO	GCACGC'	TCCAAC	CTGTI	CTACGTG	TACAACCA
EctSC_native	(3855)	AACA	TCACGC	CCAGCO	GCGCGT.	AGCAAI	CTGTT	CTACGTC	TACAACCA
	(2002)	2002	201	<u>^</u>	2020		2020		- Section 64
EctCC codon ontimized	(3902)	3902	0.0 T A A C		3920		3930		3940
EctSC_couon_optimized	(3902)	CGIG	GATAAC	GCAGIC	GCAGGC.	TCCGTT	CIGCO	AACAGAA	ACCGCGCC
ECCSC_Hattive	(3302)		UNIANI		GCAUUC	recorr		INCENCIA	- Section 85
	(3949)	3949		3960	:	3970	3	980	3995
EctSC codon optimized	(3949)	CAGC	ATTCGT	GGCCG	AACGCG	AAAACT	TCAAG	CCTCTGG	ATATTCGC
EctSC_native	(3949)	CGGC	CTTTGT	CGCCG	AACGCG.	AGAATT	TCAAG	CCGCTGG	ACATTCGG
	. ,								<ul> <li>Section 86</li> </ul>
	(3996)	3996		40	)10	4020	)	4030	4042
EctSC_codon_optimized	(3996)	CCAC	AGCAGT	ACCTG	TAATCC	AGGATC	CATAC	CTGCTCT	CCCCAGAG
EctSC_native	(3996)	CCGC	AACAGT.	ATCTC	TGATCC.	AGGATC	CATAC	CTGCTCT	CCCCAGAG
									<ul> <li>Section 87</li> </ul>
	(4043)	4043	4050		4060		4070		4089
EctSC_codon_optimized	(4043)	AATC	TAGAGT.	ACCGA	TCTGAT	CGCACG	CGACO	TCTAATT	TAGCTCGA
EctSC_native	(4043)	AATC	TAGAGT.	ACCGA	TCTGAT	CGCACG	CGACO	TCTAATT	TAGCTCGA
	(	4000		4400		4.0		20	- Section 88
Fuce and a subscient	(4090)	4090		4100	41	10	412	20	4136
EctSC_codon_optimized	(4090)	GGGG	CAAGGA.	AACAG.	TGTGGT	TTCCTT	GCCTC		CTTTTTCAG
ECISC_Halive	(4090)	GGGG	CAAGGA.	AACAG	IGIGGI	IICCII	GCCIC	.IIIIAGC	_ Section 89
	(4127)	/137		115	0	4160		4170	- Section 09 /183
EctSC codon ontimized	(4137)	4137 ACCC	TGTCTT	CCCTC	CACCAA	4100 CACCAA	ACCAG		4103 TCACAAAA
EctSC_couon_optimized	(4137)	AGGG	TGTCTT	СССТС	GACCAA	GAGGAA	ACCAG	ACAGGCG	TGACAAAA
	(1207)								<ul> <li>Section 90</li> </ul>
	(4184)	4184	4190		4200	.4	210	4220	4230
E 100 1 11 1	(								
EctSC codon optimized	(4184)	ATCT	GGATTT	CCGCCI	AGGTTT'	TGGCAC	GCCTO	STCTGGTT	TAGGGGAT
EctSC_codon_optimized EctSC_native	(4184) (4184)	ATCT ATCT	GGATTT GGATTT	CCGCCI CCGCCI	AGGTTT' AGGTTT'	TGGCAC TGGCAC	GCCTO	STCTGGTT STCTGGTT	TAGGGGAT TAGGGGAT
EctSC_codon_optimized EctSC_native	(4184) (4184) (4231)	ATCT ATCT 4231	GGATTT GGATTT 42	CCGCCI CCGCCI 240	AGGTTT' AGGTTT' 425	IGGCAC IGGCAC 0	GCCTO GCCTO 4260	TCTGGTT TCTGGTT	TAGGGGAT TAGGGGAT 4277
EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231)	АТСТ АТСТ <u>4231</u> GAGA	GGATTT GGATTT 42 AACCGG	CCGCCA CCGCCA 240 ACACACA	AGGTTT AGGTTT 425	TGGCAC TGGCAC 0 AAAACT	GCCTG GCCTG 4260	TCTGGTT TCTGGTT ) TTTTTTCG	TAGGGGAT TAGGGGAT <u>4277</u> CCAATCTT
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231)	ATCT ATCT 4231 GAGA GAGA	GGATTT GGATTT 42 AACCGG AACCGG	CCGCCA CCGCCA 240 ACACACA ACACACA	AGGTTT' AGGTTT' 425 CGTGCC. CGTGCC.	TGGCAC TGGCAC 0 AAAACT AAAACT	GCCTG GCCTG 4260 TCGGC	TCTGGTT TCTGGTT ) TTTTTTCG TTTTTCG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231)	ATCT ATCT 4231 GAGA GAGA	GGATTT GGATTT 42 AACCGG AACCGG	CCGCCA CCGCCA 240 ACACACA ACACACA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC.	TGGCAC TGGCAC 0 AAAACT AAAACT	GCCTG GCCTG 4260 TCGGC	TCTGGTT TCTGGTT ) TTTTTTCG TTTTTTCG	TAGGGGAT TAGGGGAT <u>4277</u> CCAATCTT CCAATCTT – Section 92
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4231)	ATCT ATCT 4231 GAGA GAGA 4278	GGATTT GGATTT 42 AACCGG AACCGG	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290	AGGTTT AGGTTT 425 CGTGCC. CGTGCC.	IGGCAC IGGCAC 0 AAAACI AAAACI 4300	GCCTG GCCTG 4260 TCGGC TCGGC	TCTGGTT TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310	TAGGGGAT TAGGGGAT CCAATCTT CCAATCTT – Section 92 4324
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4278) (4278)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG	CCGCC2 CCGCC2 240 ACACACAC ACACAC 4290 TCTGG	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC	rggcac rggcac 0 AAAACT AAAACT 4300 rcggat	GCCTG GCCTG 4260 TCGGC TCGGC	TCTGGTT TCTGGTT TTTTTCG TTTTTTCG 4310	TAGGGGAT TAGGGGAT CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4231) (4278) (4278) (4278)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG	CCGCC2 CCGCC2 240 ACACACA ACACACA <u>4290</u> TCTGG TCTGG	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC	rggcac rggcac 0 AAAACT AAAACT 4300 rcggat rcggat	GCCTG GCCTG 4260 TCGGC TCGGC GAGGI	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 CGATTTCA	TAGGGGAT TAGGGGAT CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4231) (4278) (4278) (4278)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG	CCGCC2 CCGCC2 240 ACACACA ACACACA (4290 TCTGG2 TCTGG2	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC	rggCac rggCac 0 AAAACT AAAACT 4300 rcgGAT rcgGAT	GCCTG GCCTG 4260 TCGGC TCGGC GAGGI	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 GATTTCA	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT – Section 92 4324 TGGCCAAG TGGCCAAG – Section 93
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325)	ATCT ATCT 4231 GAGA GAGA GAGA GTCA GTCA 4325	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330	CCGCCA CCGCCA 240 ACACACA ACACACA 4290 TCTGG TCTGG	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340	IGGCAC IGGCAC 0 AAAACT 4300 ICGGAT ICGGAT ICGGAT	GCCTG GCCTG 4260 TCGGC TCGGC GAGGT GAGGT 50	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 CGATTTCA CGATTTCA 4360	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA	CCGCC/ CCGCC/ 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 GTTCG/	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG	IGGCAC IGGCAC 0 AAAAACT AAAACT 4300 ICGGAT ICGGAT ICGGAT 438 CAGGAT	GCCTG GCCTG A260 TCGGC TCGGC TCGGC GAGGI GAGGI 50 CGCTI	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 CGATTTCA CGATTTCA 4360 CTAAGGG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4278) (4325) (4325)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA	CCGCCZ CCGCCZ 440 ACACACA ACACACA 4290 TCTGG TCTGG GTTCGZ GTTCGZ	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG ACCTCG	IGGCAC IGGCAC 0 AAAAACT AAAACT ,4300 ICGGAT ICGGAT CAGGAT CAGGAT	GCCTG GCCTG TCGGC TCGGC TCGGCG GAGGT GAGGT CGCTT CGCTT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 CGATTTCA CGATTTCA 4360 CTAAGGG CTAAGGG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA 4325 ACTT ACTT	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA	CCGCCZ CCGCCZ 240 ACACACA ACACACA 4290 TCTGG TCTGG TCTGG GTTCGZ	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. ITTGCC ITTGCC 4340 ACCTCG ACCTCG	IGGCAC IGGCAC 0 AAAACT AAAACT 4300 ICGGAT ICGGAT 438 CAGGAT CAGGAT	GCCTG GCCTG TCGGCTG TCGGC TCGGCT GAGGT GAGGT GAGGT CGCTT CGCTT	TCTGGTT TCTGGTT ) TTTTTCG CTTTTTCG 4310 CGATTTCA CGATTTCA 4360 CCTAAGGG CCTAAGGG	TAGGGGAT TAGGGGAT 4277 CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC - Section 94
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4327)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA 4325 ACTT ACTT 4372	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA CTAAAA	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290 ICTGG7 ICTGG7 ICTGG7 4 GTTCG2 GTTCG2 GTTCG2	AGGTTT' AGGTTT' <u>425</u> CGTGCC. CGTGCC. TTTGCC' 4340 ACCTCG ACCTCG <u>4390</u>	IGGCAC IGGCAC 0 AAAACT 4300 ICGGAT ICGGAT 430 CAGGAT CAGGAT	2GCCTC 2GCCTC 7CGGC 7CGGGC 7GAGGT 7GAGGT 7GAGGT 7GAGGT 7GAGGT 7CGCTT 7CGCTT 7CGCTT	TCTGGTT TCTGGTT ) TTTTTCG TTTTTCG (310 CGATTTCA CGATTTCA 4360 CCTAAGGG CCTAAGGG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 TGGCCAAG TGGCCAAG TGGCCAAG TGGCCAAG CCATTAGC CCTTTAGC CCTTTAGC - Section 94 4418
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4325) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT ACTT 4372 GGAC	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA CAAACCT	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290 ICTGG7 ICTGG7 ICTGG7 ACGCCCCA	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG ACCTCG <u>4390</u> GATACC	rggcac rggcac 0 AAAACT 4300 rcggat rcggat caggat caggat caggat	ACCTC ACCTC ACCTC ACCTC ACCTC ACCTC ACCTT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG TTTTTCG 4310 CGATTTCA 4360 CCTAAGGG CCTAAGGG CCTAAGGG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG TGGCCAAG CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4327) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA A38 CAACCT. CAACCT.	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 GTTCG3 CTCG3 GTTCG3 CTCG3 GTTCG3 GTTCG3 CTCG3 GTTCG3 G	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG ACCTCG <u>4390</u> GATACC	rggcac rggcac 0 AAAACT 4300 rcggat rcggat caggat caggat caggat caggat	GCCTG GCCTG TCGGC TCGGC TCGGC GAGGT GAGGT GAGGT CGCTT CCCTT CCCTT 4400 GAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG 4310 CGATTTCA 4360 CCTAAGGG CCTAAGGG CCTCGACG CCTCGACG	TAGGGGAT TAGGGGAT 4277 CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG TGGCCAAG CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT CCTTAAAT
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4325) (4372) (4372) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC	GGATTT GGATTT 42 AACCGG AACCGG CGCCTG CGCCTG 4330 CTAAAA CTAAAA 438 CAACCT. CAACCT.	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 ACGCCCG AGGCCCC 4430	AGGTTT AGGTTT <u>425</u> CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG ACCTCG <u>4390</u> GATACC	IGGCAC IGGCAC 0 AAAACT AAAACT 4300 ICGGAT ICGGAT CAGGAT CAGGAT CAGGAT CAGGAT	GCCTC GCCTC GCCTC TCGGC TCGGC GAGGI GAGGI CCGCTI CCGCTI 4400 GAAAI GAAAI	TTTTTCG TTTTTCG TTTTTCG TTTTTCG TTTTTCG CTTTTTCG CTTTTTCG CTTTTCG CTTAAGGG CCTAAGGG CCTAAGGG CCTCGACG CCTCGACG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 94 4455
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4325) (4372) (4372) (4372) (4372) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC 4419 GCAC	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG 4330 CTAAAA CTAAAA 438 CAACCT. CAACCT.	2CGCC2 2CGCC2 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 TCTGG7 CTGG7 GTTCG2 GTTCG2 0 AGGCCCA AGGCCCA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG 4340 ACCTCG 4390 GATACC GATACC	TGGCAC TGGCAC 0 AAAAACT 	GCCTC GCCTC GCCTC TCGGC TCGGC GAGGI GAGGI CCGCTI CCGCTI 4400 GAAAI GAAAI	TTTTTTCG TTTTTTCG TTTTTTCG TTTTTTCG (TTTTTCG CTTTTTCG CTTTTTCG CTTAAGGG CTTAAGGG CTTCGACG CTTCGACG CTCCGACG CTCCGACG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 CACGATAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4372) (4372) (4372) (4372) (4372) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC 4419 GGAC GGAC	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG 4330 CTAAAAA CTAAAAA CAACCT. CAACCT. CAACCT. GATTGG.	2CGCC2 2CGCC2 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 TCTGG7 TCTGG7 AGGCCCA AGGCCA AGGCTAA AGCTAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG 4340 ACCTCG 4390 GATACCG	IGGCAC IGGCAC 0 AAAAACI 	GCCTC GCCTC GCCTC TCGGC TCGGC GAGGT GAGGT GAGGT CCGCTT GCGCTT GAAAT GAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (TTTTCA (TT	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4372) (4372) (4372) (4372) (4372) (4372) (4372) (4372) (4372)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC GGAC GGAC	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG 4330 CTAAAAA CTAAAAA 438 CAACCT. CAACCT. GATTGG.	2CGCC2 2CGCC2 240 ACACAA ACACAA 4290 TCTGG7 TCTGG7 TCTGG7 AGGTTCG2 0 AGGCCC AGGCCC AGGCCCA AGGCCAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG 4340 ACCTCG 4340 GATACCG GATACC GATACC	IGGCAC IGGCAC 0 AAAACI AAAACI A300 ICGGAT ICGGAT CAGGAT CAGGAT CAGGAT CATGTG CATGTG CATGTG CATGTG CATGTG CATGTG CATGTG CATGTG CATGTG	GCCTG GCCTG GCCTG TCGGC TCGGCG GAGGT GAGGT GAGGT CCGCTT CCGCTT 4400 GGAAAT GAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (GATTTCA CCTAAGGG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG - Section 96
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4	ATCT ATCT 4231 GAGA GAGA GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC GGAC GGAC 4419 GGAC GGAC	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (CTAAAAA CTAAAAA (CTAAAAA) (CTAAAAA) CAACCT. CAACCT. CAACCT. GATTGG.	2CGCC2 2CGCC2 240 ACACACA ACACACA 4290 TCTGG7 TCTGG7 TCTGG7 AGGTCG2 0 AGGCCCC AGGCCCC AGGCCCC AGGCCCA AGGCCAA AGCTAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG 4340 ACCTCG 4340 GATACCG GATACC GATACCA AAACCA	IGGCAC IGGCAC 0 AAAACI AAAACI A300 ICGGAI CCGGAI CAGGAI CAGGAI CATGTG CA	GCCTG GCCTG GCCTG TCGGC TCGGC GAGGT GAGGT GAGGT CCGCTT GCGCTT GAAAT A400 GGAAAT GAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (GATTTCA CCTAAGGG CCTCAAGGG CCTCGACG CC	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG - Section 96 4512
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT 4322 GGAC GGAC GGAC GGAC 4419 GGAC GGAC 4466 GATT	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (CTAAAAA CTAAAAA (CTAAAAA) (CTAAAAA) CTAAAAA CAACCT. CAACCT. CAACCT. GATTGG. GATTGG.	CCGCC2 CCGCC2 240 ACACACA ACACACA 4290 ICTGG7 TCTGG7 TCTGG7 AGGTCG2 0 AGGCCCC AGGCCCC AGGCCCC AGGCCCA AGGCCAA AGCTAA AGCTAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG 4340 ACCTCG 4340 GATACCG GATACC GATACCA AAACCA AAACCA AAACCA	IGGCAC IGGCAC 0 AAAACI AAAACI A300 ICGGAI CAGGAI CAGGAI CAGGAI CAGGAI CATGTG CATGTG CATGTG CATGTG CATGTG CATGTG CGAACA CGAACA CGAACA	GCCTG GCCTG GCCTG TCGGC TCGGC GAGGT GAGGT GAGGT GAGGT GAGGT GAGAAT GGAAAT GGAAAT AGCTGG GCTGG	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (GATTTCA (GATTTCA CCTAAGGG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG CCTCGACG (GATTTTC (GATTTCC (GATTTCC) (GATTTTC) (GATTTC) (GATTTC) (GATTT	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT - Section 95 4465 CACGATAG CACGATAG - Section 96 4512 TCGCGGTC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4372) (4372) (4419) (4419) (4419) (4466) (4466) (4466)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC GGAC 4419 GGAC GGAC 4466 GATT GATT	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (CTAAAA) CTAAAA (CTAAAA) (CTAAAA) CTAAAAA (CTAAAA) CAACCT. CAACCT. CAACCT. GATTGG. GATTGG.	CCGCC2 CCGCC2 240 ACACAA ACACAA 4290 TCTGG7 TCTGG7 TCTGG7 AGTTCG2 0 AGGCCCC AGGCCCC AGGCCCC AGGCCCC AGGCCCA AGGCTAA AGCTAA AGCTAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG AAACCCA AAACCA AAACCA AAACCA AAACCA AAACCA	IGGCAC IGGCAC 0 AAAACT AAAACT ,4300 ICGGAT ICGGAT CAGGAT CAGGAT CAGGAT CATGTG CATGTG CATGTG CATGTG CATGTG CGAACA CGAACA A490 GTTGGT	GCCTG GCCTG GCCTG TCGGC TCGGC GAGGI GAGGI GAGGI GAGGI GGAGTI GGAAAI GGAAAI GCTGG GCTGG	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (GATTTTCA (GATTTTCA (GATTTTCA (GATTTTCA (GATTTA (GATTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTA (GATTA (GATTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTTA (GATTA (GATTTA (G	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT - Section 95 4465 CACGATAG CACGATAG - Section 96 4512 TCGCGGTC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4322) (4372) (4372) (4479) (4419) (4419) (4419) (4466) (4466)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT 4372 GGAC GGAC GGAC 4419 GGAC GGAC 4466 GATT GATT	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (A330 CTAAAA CTAAAA (A38 CAACCT. CAACCT. GATTGG. GATTGG. GGGTCT GGGTCT	CCGCC2 CCGCC2 240 ACACACA ACACACA ACACACA ACACACA ACACACA ACACACA ACACACA ACACACA ACACACA ACACACA AC	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG AACCTCG AAACCA AAACCA AAACCA AAACCA B0 AGATTC	rggcac rggcac 0 AAAACT AAAACT ,4300 rCGGAT rCGGAT CAGGAT CAGGAT CAGGAT CATGTG CATGTG CATGTG CATGTG CATGTG CGAACA CGAACA ,4490 GTTGGT	ACCTC ACCTC	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (GATTTA (GATTTA	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG - Section 96 4512 TCGCGGTC TCGCGGTC - Section 97
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4322) (4372) (4372) (4372) (4419) (4419) (4419) (4466) (4466) (4466) (44513)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT 4372 GGAC GGAC GGAC GGAC 4419 GGAC GGAC GGAC 4466 GATT GATT 4513	GGATTT GGATTT 42 AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (A330 CTAAAA CTAAAA CTAAAA (A38 CAACCT. CAACCT. GATTGG. GATTGG. GGGTCT GGGTCT	CCGCC2 CCGCC2 240 ACACAAA ACACAAA ACACAAA ACACAAAA ACACAAA ACACAAAA ACACAAAA ACACAAAA ACACAAAA ACACAAA ACACAAAA ACACAAAA ACACAAAAA ACACAAAA ACACAAAA ACACAAAAA ACACAAAAAA	AGGTTT AGGTTT 425 CGTGCC. CGTGCC. CGTGCC. TTTGCC 4340 ACCTCG 4340 ACCTCG 4340 ACCTCG 4340 ACCTCG 4340 ACCTCG 4340 ACCTCG 4340 ACCTCG AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA	IGGCAC IGGCAC 0 AAAACT AAAACT ,4300 ICGGAT ICGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGACA CGAACA CGAACA CGAACA A490 GTTGGT	AGCTGG AGCTT AGAGGT GAGGT GAGGT GAGGT GAGGT GAGGT GAGGT AGAAAT AGCTGG GAAAT AGCTGG GAAAT AGCTGG GAAAT AGCTGG	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (4310 CGATTTCA CGATTTCA (4360 CCTCGACG CCTCACGACG CCTC	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG - Section 97 TCGCGGTC TCGCGGTC - Section 97
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4322) (4372) (4372) (4372) (4372) (4419) (4419) (4419) (4419) (4466) (4466) (4466) (4465) (4513) (4513)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT 4372 GGAC GGAC GGAC GGAC 4419 GGAC GGAC GGAC 4466 GATT GATT	GGATTT GGATTT AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (A330 CTAAAAA CTAAAAA CTAAAAA CTAAAAA CAACCT. GATTGG. GATTGG. GATTGG. GGGTCT GGGTCT GGGTCT	CCGCC2 CCGCC2 ACACACACA ACACACACACA ACACACACACACACA AC	AGGTTT AGGTTT AGGTTT (425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG (4340 ACCTCG (4340 ACCTCG (4390 GATACC) (4390 GATACC) (4390 GATACC) (4390 GATACC) (4390 GATACC) (4390 GATACC) (4390 GATACC) (4530 TAGTAA	IGGCAC IGGCAC 0 AAAACT AAAACT AAAACT CGGAT CGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGACA CGAACA CGAACA CGAACA CGAACA CGAACA CGACA	AS40 CTTGGAP CGAGGT CGAGGT CGAGGT CGAGGT CGAGGT CGAGGT CGCTT CGCTT CGCTT CGCTT CGCTT CGCTGG CGCTG CGCTG CGCTG CGCTG CGCTG CGCTG CGCGT CGCGT CGCGT CGCGT CGCGT CGCGT CGCGT CGCGGT CGCGGT CGCGT CGCGGT CGCGG CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CTCGGC CGCTG CGCGGT CGCGCT CGCGT CGCTG CGCC CGC CG	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (4310 CGATTTCA CGATTTCA (4360 CCTCGACG CCTCGA	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG Section 96 4512 TCGCGGTC TCGCGGTC Section 97 4559 GATGCCGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4322) (4372) (4372) (4372) (4372) (4419) (4419) (4419) (4419) (4466) (4466) (4466) (4466) (4513) (4513) (4513)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT 4372 GGAC GGAC GGAC GGAC 4419 GGAC GGAC GGAC 4466 GATT GATT 4513 GCGG GCGG	GGATTT GGATTT AACCGG. AACCGG. CGCCTG CGCCTG CGCCTG (CGCCTG (CGCCTG (CGCCTG (CGCCTG (CAACCT) CAACCT. CAACCT. GATTGG. GATTGG. GGGTCT GGGTCT GGGTCT AAGAAT AAGAAT	CCGCC2 CCGCC2 ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACA ACACACACACACA AC	AGGTTT AGGTTT AGGTTT (425 CGTGCC. CGTGCC. TTTGCC TTTGCC 4340 ACCTCG A340 ACCTCG (4390 GATACCC GATACC GATACC (3ATACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AGATTC (4530 TAGTAA	IGGCAC IGGCAC 0 AAAACT AAAACT AAAACT CGGAT CGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGACA CGAACA CGAACA GTTGGT GTTGGT	AS40 CTTGGAP CGAGGT CGAGGT CGAGGT CGAGGT CGAGGT CGCTT CGCTT CGCTT CGCTT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (4310 CGATTTCA CGATTTCA CGATTTCA (4360 CCTCGACG CCTCGA	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG CACGATAG CACGATAG - Section 97 4559 GATGCCGC
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4325) (4322) (4372) (4372) (4372) (4372) (4372) (4419) (4419) (4419) (4419) (4466) (4466) (4466) (4466) (4513) (4513) (4513)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT 4372 GGAC GGAC 4419 GGAC GGAC 4466 GATT GATT 4513 GCGG GCGG	GGATTT GGATTT ( AACCGG, AACCGG, CGCCTG CGCCTG ( CGCCTG CGCCTG ( A330 CTAAAAA CTAAAAA CTAAAAA CTAAAAA CTAAAAA CAACCT. GATTGG, GATTGG, GGGTCT GGGTCT GGGTCT AAGAAT AAGAAT	CCGCC2 CCGCC2 40 ACACAA ACACAA ACACAA ACACAA ACCCAA TCTGG7 TCTGG7 CTGG7 GTTCG3 O AGGCCCA AGGCCCA AGGCCCA AGGCCCA AGGCCA	AGGTTT AGGTTT AGGTTT (425 CGTGCC. CGTGCC. CGTGCC. TTTGCC 4340 ACCTCG AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA AAACCA	IGGCAC IGGCAC 0 AAAACT AAAACT AAAACT CGGAT CGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGACA CGAACA CGAACA GTTGGT GTTGGT	AS40 CTTGGAP CGAGGT CGAGGT CGAGGT CGAGGT CGAGGT CGCTT CGCTT CGCTT CGCTT CGCTG CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT CGAAAT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (4310 CGATTTCA CGATTTCA (4360 CCTAAGGG CCTCGACG CCTCGA	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG - Section 97 4559 GATGCCGC GATGCCGC - Section 98
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4184) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4325) (4325) (4372) (4372) (4372) (4372) (4419) (4419) (4419) (4419) (4466) (4466) (4466) (4513) (4513) (4513) (4560)	ATCT ATCT 4231 GAGA GAGA GTCA GTCA GTCA 4325 ACTT 4372 GGAC GGAC GGAC 4419 GGAC GGAC GGAC 4466 GATT GATT 4513 GCGG GCGG	GGATTT GGATTT (GATTT) (42 AACCGG, AACCGG, CGCCTG (CGCCTG (CGCCTG) (4330 CTAAAA CTAAAA (CTAAAA) (438 CAACCT, CAACCT, GATTGG, GATTGG, GGGTCT (GGGTCT) (4520 AAGAAT AAGAAT	CCGCC2 CCGCC2 ACACACACA ACACACA ACACACACA ACACACACACA ACACACACACACACA AC	AGGTTT' AGGTTT' <u>425</u> CGTGCC. CGTGCC. CGTGCC. TTTGCC' 4340 ACCTCG AAACCA AAACA AAACAA AAACAAAAAA	IGGCAC IGGCAC 0 AAAACT AAAACT AAAACT CGGAT ICGGAT ICGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGACA CGAACA GATGGT GTTGGT GTTGGT	AS40 AGCTGG AGCTGG AGAGGT CGAGGT CGAGGT CGAGGT CGAGGT CGCTT AGAAAT GGAAAT AGCTGG CGCTG CGCTG CGCTG CGCTG CGCGGT CGGGT CGGGT CGGGT CGGGT CGGGT CGGGT CGGGT CGGGT CGGT CGGT CGGGT CGGT	TCTGGTT TCTGGTT TTTTTCG TTTTTCG (A310 CGATTTCA CGATTTCA CGATTTCA (CGATTTCA (CGATTTCA CCTCGACG C	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC - Section 94 4418 TCTTAAAT TCTTAAAT - Section 95 4465 CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG Section 97 4559 GATGCCGC GATGCCGC - Section 98
EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_codon_optimized EctSC_native EctSC_codon_optimized EctSC_native	(4184) (4184) (4231) (4231) (4231) (4231) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4278) (4325) (4325) (4372) (4372) (4372) (4372) (4372) (4372) (4372) (4419) (4419) (4419) (4466) (4466) (4466) (4513) (4513) (4560) (4560)	ATCT ATCT 4231 GAGA GAGA 4278 GTCA GTCA 4325 ACTT ACTT 4372 GGAC GGAC 4419 GGAC GGAC 4466 GGATT GGATT GGATT 4513 GCGG GCGG GCGG 44560 AATCC	GGATTT GGATTT AACCGG AACCGG CGCCTG CGCCTG CGCCTG CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAA CTAAAAA CAACCT CAACCT GGGTCT GGGTCT AAGAAT AAGAAT	CCGCC2 CCGCC2 240 ACACACA ACACACA ACACACA ACCACACA ACCACACA ACCACACA ACCACACACACACACACACACACACACACACACACACACA	AGGTTT' AGGTTT' 425 CGTGCC. CGTGCC. CGTGCC. TTTGCC' 4340 ACCTCG 4340 ACCTCG 4340 GATACCG 4390 GATACCA 4390 GATACCA 4390 GATACCA 4390 AAACCA 4390 AAACCA 4390 AAACCA 4390 AAACCA 4390 AAACCA 4390 AAACCA 4390 AAACCA 4300 AAACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 AACCA 400 400 400 400 400 400 400 40	TGGCAC TGGCAC 0 AAAACT 4300 TCGGAT TCGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT CAGGAT GTTGGT GTTGGT GTTGGT TGCTGC	ASSACE AND A CONTRACTOR	TTTTTCG TTTTTCG TTTTTCG TTTTTCG CTTTTTCG CTTTTTCG CTTTTTCG CTTTTCCG CTTAAGGG CTTAAGGG CTCGACG	TAGGGGAT TAGGGGAT 4277 CCAATCTT CCAATCTT - Section 92 4324 TGGCCAAG TGGCCAAG - Section 93 4371 CCTTTAGC CCTTTAGC CCTTTAGC CCTTTAAC CCTTAAAT - Section 94 4418 TCTTAAAT - Section 95 CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG CACGATAG - Section 97 4559 GATGCCGC GATGCCGC - Section 98

Fig. 36: Ectoine nucleotide sequence. (Continued from page 100).

# Appendix

leastSQ	and all the state
original	modified
%%to load input substrate MDV	%%to load input substrate MDV
global guidedinSub	global guidedinsub
11 1sempty(guidedinsub)	11 isempty(guideainsub)    guideainsub
guidedinsub 0	U substrato EMII
load inputSubEMU	load inputSubEMU expl
end	load input SubEMU exp2
	end
%%load rates	%%load rates
<pre>v = - fluxCalc(u);</pre>	v = - fluxCalc(u);
v(find(abs(v)<1e-8))=0;	v(find(abs(v)<1e-8))=0;
<pre>v = appendSyn(v,u);</pre>	<pre>v = appendSyn(v,u);</pre>
loader_EMUModel	loader_EMUModel
	· · · · · · · · · · · · · · · · · · ·
	x_calc_big=[];
	TOad InputSubEMO_expl
EMIIModel	EMIModel
2110110401	Lindidadi
x sim	x sim
—	x calc big=[x calc];
	<pre>load inputSubEMU_exp2</pre>
	EMUModel
	x_sim
	x_calc_big=[x_calc_big; x_calc];
	u sele — u sele bie:
mdyGenerator	x_care = x_care_big;
original	modified
<pre>a = dir(fullfile(pwd, '*.mat'));</pre>	<pre>a = dir(fullfile(pwd, '*.mat'));</pre>
<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0;</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0;</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0;</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile 2 = 0;</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat')</pre>	<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat')</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat')</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU_expl.mat') AAVFile = 1; end if strcmp(a(i).name, 'inputSubEMU_exp2.mat') aAVFile_2 = 1;</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU_expl.mat') AAVFile = 1; end if strcmp(a(i).name, 'inputSubEMU_exp2.mat') AAVFile_2 = 1; end</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)     if strcmp(a(i).name,     'inputSubEMU.mat')         AAVFile = 1;     end end</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)     if strcmp(a(i).name,     'inputSubEMU.mat')         AAVFile = 1;     end end if AAVFile ==1;</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)     if strcmp(a(i).name,     'inputSubEMU.mat')         AAVFile = 1;     end end if AAVFile ==1;     load inputSubEMU.mat</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a)     if strcmp(a(i).name,     'inputSubEMU.mat')</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat')</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU;</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile 2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel EMUModel</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel EMUModel x_sim</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile 2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel EMUModel x_sim</pre>	<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; AAVFile 2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel EMUModel x_sim</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>
<pre>a = dir(fullfile(pwd,'*.mat')); AAVFile = 0; for i = 1:length(a) if strcmp(a(i).name, 'inputSubEMU.mat') AAVFile = 1; end end if AAVFile ==1; load inputSubEMU.mat else substrate_EMU; end loader_EMUModel EMUModel x_sim mdvOut = x_calc;</pre>	<pre>a = dir(fullfile(pwd, '*.mat')); AAVFile = 0; AAVFile_2 = 0; for i = 1:length(a)</pre>

**Fig. 37:** Reprogramming of the OpenFlux subroutine files 'leastSQ.m' and 'mdvGenerator' for comprehensive integration of two parallel <sup>13</sup>C-labeling experiments.

# BIBLIOGRAPHY

- Abdel-Aziz, H., Wadie, W., Abdallah, D. M., Lentzen, G. and Khayyal, M. T. (2013). 'Novel effects of ectoine, a bacteria-derived natural tetrahydropyrimidine, in experimental colitis'. *Phytomedicine* 20: 585–591.
- [2] Abe, S., Takayama, K. and Kinoshita, S. (1967). 'Taxonomical Studies on glutamic acid-producing bacteria.' *J. Gen. Appl. Microbiol.* 13: 279–301.
- [3] Adachi, N., Takahashi, C., Ono-Murota, N., Yamaguchi, R., Tanaka, T. and Kondo, A. (2013). 'Direct L-lysine production from cellobiose by *Corynebacterium glutamicum* displaying β-glucosidase on its cell surface'. *Appl. Microbiol. Biotechnol.* 7165–7172.
- [4] Allison, S. D., Wallenstein, M. D. and Bradford, M. A. (2010). 'Soil-carbon response to warming dependent on microbial physiology'. *Nature Geosci.* 3: 336–340.
- [5] Aristidou, A. and Penttilä, M. (2000). 'Metabolic engineering applications to renewable resource utilization'. *Curr. Opin. Biotechnol.* 11: 187–198.
- [6] Ballesteros, M., Sáez, F., Ballesteros, I., Manzanares, P., Negro, M., Martínez, J., Castañeda, R. and Oliva Dominguez, J. (2010). 'Ethanol Production from the Organic Fraction Obtained After Thermal Pretreatment of Municipal Solid Waste'. *Appl. Biochem. Biotechnol.* 161: 423–431.
- Barrett, E., Stanton, C., Zelder, O., Fitzgerald, G. and Ross, R. P. (2004).
   'Heterologous Expression of Lactose- and Galactose-Utilizing Pathways from Lactic Acid Bacteria in *Corynebacterium glutamicum* for Production of Lysine in Whey'. *Appl. Environ. Microbiol.* 70: 2861–2866.
- [8] Barth, S., Huhn, M., Matthey, B., Klimka, A., Galinski, E. A. and Engert, A. (2000). 'Compatible-Solute-Supported Periplasmic Expression of Functional Recombinant Proteins under Stress Conditions'. *Appl. Environ. Microbiol.* 66: 1572–1579.
- [9] Becker, J., Kind, S. and Wittmann, C. (2012). 'Systems Metabolic Engineering of *Corynebacterium glutamicum* for Biobased Production of Chemicals, Materials and Fuels'. In: *Systems Metabolic Engineering*. Ed. by Wittmann, C. and Lee, S. Y. Springer Netherlands, 151–191.
- [10] Becker, J., Klopprogge, C., Herold, A., Zelder, O., Bolten, C. J. and Wittmann, C. (2007). 'Metabolic flux engineering of L-lysine production in *Corynebacte*-

*rium glutamicum* – over expression and modification of G6P dehydrogenase'. *J. Biotechnol.* 132: 99–109.

- [11] Becker, J., Klopprogge, C., Schröder, H. and Wittmann, C. (2009). 'Metabolic Engineering of the Tricarboxylic Acid Cycle for Improved Lysine Production by *Corynebacterium glutamicum*'. *Appl. Environ. Microbiol.* 75: 7866–7869.
- [12] Becker, J., Klopprogge, C. and Wittmann, C. (2008). 'Metabolic responses to pyruvate kinase deletion in lysine producing *Corynebacterium glutamicum*'. *Microb. Cell Fact.* 7: 1–15.
- Becker, J., Klopprogge, C., Zelder, O., Heinzle, E. and Wittmann, C. (2005).
   'Amplified Expression of Fructose 1,6-Bisphosphatase in *Corynebacterium glutamicum* Increases In Vivo Flux through the Pentose Phosphate Pathway and Lysine Production on Different Carbon Sources'. *Appl. Environ. Microbiol.* 71: 8587–8596.
- [14] Becker, J., Schäfer, R., Kohlstedt, M., Harder, B. J., Borchert, N., Stöveken, N., Bremer, E. and Wittmann, C. (2013). 'Systems metabolic engineering of *Corynebacterium glutamicum* for production of the chemical chaperone ectoine'. *Microb. Cell Fact.* 12: 1–15.
- [15] Becker, J. and Wittmann, C. (2012). 'Bio-based production of chemicals, materials and fuels *Corynebacterium glutamicum* as versatile cell factory'. *Curr. Opin. Biotechnol.* 23: 631–640.
- [16] Becker, J. and Wittmann, C. (2012). 'Systems and synthetic metabolic engineering for amino acid production – the heartbeat of industrial strain development'. *Curr. Opin. Biotechnol.* 23: 718–726.
- [17] Becker, J. and Wittmann, C. (2013). 'Pathways at Work: Metabolic Flux Analysis of the Industrial Cell Factory *Corynebacterium glutamicum*'. In: *Corynebacterium glutamicum*. Ed. by Yukawa, H. and Inui, M. 23: Microbiology Monographs. Springer Berlin Heidelberg, 217–237.
- [18] Becker, J., Zelder, O., Häfner, S., Schröder, H. and Wittmann, C. (2011).
   'From zero to hero Design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production'. *Metab. Eng.* 13: 159–168.
- Becker, M., Börngen, K., Nomura, T., Battle, A. R., Marin, K., Martinac, B. and Krämer, R. (2013). 'Glutamate efflux mediated by *Corynebacterium glutamicum* MscCG, Escherichia coli MscS, and their derivatives'. *Biochim. Biophys. Acta* 1828: 1230–1240.
- Berger, A., Dohnt, K., Tielen, P., Jahn, D., Becker, J. and Wittmann, C. (2014).
   'Robustness and plasticity of metabolic pathway flux among uropathogenic isolates of *Pseudomonas aeroginosa*'. *PLoS ONE* 9: 1–14.
- [21] Bestvater, T., Louis, P., Galinski, E. et al. (2008). 'Heterologous ectoine production in *Escherichia coli*: by-passing the metabolic bottle-neck'. *Saline Systems* 4: 1–14.
- [22] Biopolako (2006). Corynebacterium.jpg. URL: https://commons.wikimedia.org/wiki/File:Corynebacterium.jpg.

- [23] Blombach, B., Riester, T., Wieschalka, S., Ziert, C., Youn, J.-W., Wendisch, V. and Eikmanns, B. (2011). *'Corynebacterium glutamicum* tailored for efficient isobutanol production.' *Appl. Environ. Microbiol.* 77: 3300–3310.
- [24] Blombach, B. and Seibold, G. (2010). 'Carbohydrate metabolism in *Corynebacterium glutamicum* and applications for the metabolic engineering of L-lysine production strains'. *Appl. Microbiol. Biotechnol.* 86: (5), 1313–1322.
- [25] Bolten, C. J. and Dickschat, J. (2010). 'Towards Methionine Overproduction in *Corynebacterium glutamicum* – Methanethiol and Dimethyldisulfide as Reduced Sulfur Sources'. *J. Microbiol. Biotechnol.* 20: 1196–1203.
- Bolten, C. J., Kiefer, P., Letisse, F., Portais, J.-C. and Wittmann, C. (2007).
   'Sampling for Metabolome Analysis of Microorganisms'. *Anal. Chem.* 79: 3843–3849.
- [27] Borchert, N. (2012). *Metabolic Engineering von Corynebacterium glutamicum zur Ectoinproduktion*. Master's thesis. Germany: Technische Universität Braunschweig.
- [28] Bozell, J. J. and Petersen, G. R. (2010). 'Technology development for the production of biobased products from biorefinery carbohydrates-the US Department of Energy's "Top 10" revisited'. *Green Chem.* 12: (4), 539–554.
- [29] Bremer, E. and Krämer, R. (2000). *Bacterial Stress Responses*. In: ed. by Storz G. andHengge-Aroni, R. ASM Press. Chap. Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes, 79–97.
- [30] Buchholz, J., Graf, M., Freund, A., Busche, T., Kalinowski, J., Blombach, B. and Takors, R. (2014). 'CO<sub>2</sub>/HCO<sub>3</sub><sup>-</sup> perturbations of simulated large scale gradients in a scale-down device cause fast transcriptional responses in *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 98: 8563–8572.
- [31] Burkovski, A. and Krämer, R. (2002). 'Bacterial amino acid transport proteins: occurrence, functions, and significance for biotechnological applications'. *Appl. Microbiol. Biotechnol.* 58: 265–274.
- [32] Burkovski, A. (2013). 'Cell Envelope of Corynebacteria: Structure and Influence on Pathogenicity'. *ISRN Microbiology*, 1–11.
- [33] Bursy, J., Kuhlmann, A. U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A. J. and Bremer, E. (2008). 'Synthesis and Uptake of the Compatible Solutes Ectoine and 5-Hydroxyectoine by *Streptomyces coelicolor* A3(2) in Response to Salt and Heat Stresses'. *Appl. Environ. Microbiol.* 74: 7286–7296.
- [34] Buschke, N., Becker, J., Schäfer, R., Kiefer, P., Biedendieck, R. and Wittmann,
   C. (2013). 'Systems metabolic engineering of xylose utilizing *Corynebacterium glutamicum* for production of 1,5-diaminopentane'. *Biotech. J.* 8: 557–570.
- [35] Buschke, N., Schäfer, R., Becker, J. and Wittman, C. (2013). 'Metabolic engineering of industrial platform microorganisms for biorefinery applications optimization of substrate spectrum and process robustness by rational and evolutive strategies'. *Bioresour. Technol.* 135: 544–554.

- [36] Buschke, N., Schröder, H. and Wittmann, C. (2011). 'Metabolic engineering of *Corynebacterium glutamicum* for production of 1,5-diaminopentane from hemicellulose'. *Biotech. J.* 6: 306–317.
- [37] Calderon, M. I. (2004). 'Complex regulation of the synthesis of the compatible solute ectoine in the halophilic bacterium *Chromohalobacter salexigens* DSM 3043T'. *Microbiology* 150: 3051–3063.
- [38] Chen, Z., Meyer, W., Rappert, S., Sun, J. and Zeng, A.-P. (2011). 'Coevolutionary Analysis Enabled Rational Deregulation of Allosteric Enzyme Inhibition in *Corynebacterium glutamicum* for Lysine Production'. *Appl. Environ. Microbiol.* 77: 4352–4360.
- [39] Cremer, J., Treptow, C., Eggeling, L. and Sahm, H. (1988). 'Regulation of enzymes of lysine biosynthesis in *Corynebacterium glutamicum*'. *J. Gen. Microbiol.* 134: 3221–3229.
- [40] Dale, B. E. (1987). 'Lignocellulose conversion and the future of fermentation biotechnology'. *Trends Biotechnol.* 5: 287–291.
- [41] Dale, B. and Kim, S. (2006). Biomass Refining Global Interest The Biobased Economy of the 21st Century. In: Biorefineries - Industrial Processes and Products: Status Quo and Future Directions. Ed. by Kamm, B., Gruber, P. and Kamm, M. Wiley VCH. Chap. 2, 41–66.
- [42] Dean, B., Dodge, T., Valle, F. and Chotani, G. (2006). Development of Biorefineries – Technical and Economic Considerations. In: Biorefineries - Industrial Processes and Products: Status Quo and Future Directions. Ed. by Kamm, B., Gruber, P. and Kamm, M. Wiley VCH. Chap. 3, 67–84.
- [43] Delaunay, S., Uy, D., Baucher, M., Engasser, J., Guyonvarch, A. and Goergen, J. (1999). 'Importance of Phosphoenolpyruvate Carboxylase of *Corynebacterium glutamicum* during the Temperature Triggered Glutamic Acid Fermentation'. *Metab. Eng.* 1: 334–343.
- [44] Demirel, Y. (2012). 'Energy and Energy Types'. In: *Energy*. Green Energy and Technology. Springer London, 27–70.
- [45] Dominguez, H., Cocaign-Bousquet, M. and Lindley, N. D. (1997). 'Simultaneous consumption of glucose and fructose from sugar mixtures during batch growth of *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 47: (5), 600–603.
- [46] Eggeling, L. (2013). 'Aminosäuren'. In: *Industrielle Mikrobiologie*. Ed. by Sahm, H., Antranikian, G., Stahmann, K.-P. and Takors, R. Springer Berlin Heidelberg, 109–126.
- [47] Eggeling, L. and Bott, M. (2015). 'A giant market and a powerful metabolism:
   L-lysine provided by *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 99: 3387–3394.
- [48] Eikmanns, B. J., Eggeling, L. and Sahm, H. (1993). 'Molecular aspects of lysine, threonine, and isoleucine biosynthesis in *Corynebacterium glutamicum*'. *Antonie van Leeuwenhoek* 64: 145–163.
- [49] Eikmanns, B. J., Follettie, M. T., Griot, M. U. and Sinskey, A. J. (1989). 'The phosphoenolpyruvate carboxylase gene of *Corynebacterium glutamicum*:

Molecular cloning, nucleotide sequence, and expression'. *Mol. Gen. Genet.* 218: 330–339.

- [50] Eilert, E., Kranz, A., Hollenberg, C. P., Piontek, M. and Suckow, M. (2013).
   'Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*'. J. Biotechnol. 167: 85–93.
- [51] Eliasson, A., Boles, E., Johansson, B., Österberg, M., Thevelein, J. M., Spencer-Martins, I., Juhnke, H. and Hahn-Hägerdal, B. (2000). 'Xylulose fermentation by mutant and wild-type strains of *Zygosaccharomyces* and *Saccharomyces cerevisiae*'. *Appl. Microbiol. Biotechnol.* 53: (4), 376–382.
- [52] El-Mansi, E. M. T. (2006). Fermentation Microbiology and Biotechnology. Ed. by El-Mansi, E. M. T., Bryce, C. F. A., Demain, A. L. and Allman, A. 2nd ed. CRC Press.
- [53] Fallet, C., Rohe, P. and Franco-Lara, E. (2010). 'Process optimization of the integrated synthesis and secretion of ectoine and hydroxyectoine under hyper/hy-po-osmotic stress'. *Biotechnol. Bioeng.* 107: 124–133.
- [54] Freeman, W. M., Walker, S. J. and Vrana, K. E. (1999). 'Quantitative RT-PCR: pitfalls and potential'. *BioTechniques* 26: 112–125.
- [55] Frings, E., Kunte, H. and Galinski, E. A. (1993). 'Compatible solutes in representatives of the genera *Brevibacterium* and *Corynebacterium*: Occurrence of tetrahydropyrimidines and glutamine'. *FEMS Microbiol. Lett.* 109: 25–32.
- [56] Gao, Q., Zhang, M., McMillan, J. and Kompala, D. (2002). 'Characterization of heterologous and native enzyme activity profiles in metabolically engineered *Zymomonas mobilis* strains during batch fermentation of glucose and xylose mixtures'. *Appl. Biochem. Biotechnol.* 98-100: (1), 341–355.
- [57] García-Estepa, R., Argandoña, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J. J. and Vargas, C. (2006). 'The *ectD* Gene, Which Is Involved in the Synthesis of the Compatible Solute Hydroxyectoine, Is Essential for Thermoprotection of the Halophilic Bacterium Chromohalobacter salexigens'. J. Bacteriol. 188: 3774–3784.
- [58] Gimbun, J., Radiah, A. D. and Chuah, T. (2004). 'Bioreactor design via spreadsheet—a study on the monosodium glutamate (MSG) process'. *J. Food Eng.* 64: 277–283.
- [59] Gopinath, V., Murali, A., Dhar, K. and Nampoothiri, K. (2012). '*Corynebacterium glutamicum* as a potent biocatalyst for the bioconversion of pentose sugars to value-added products'. *Appl. Microbiol. Biotechnol.* 93: (1), 95–106.
- [60] Gourdon, P., Raherimandimby, M., Dominguez, H., Cocaign-Bousquet, M. and Lindley, N. D. (2003). 'Osmotic stress, glucose transport capacity and consequences for glutamate overproduction in *Corynebacterium glutamicum*'. *J. Biotechnol.* 104: 77–85.
- [61] Graf, R., Anzali, S., Buenger, J., Pfluecker, F. and Driller, H. (2008). 'The multifunctional role of ectoine as a natural cell protectant'. *Clin. Dermatol.* 26: 326–333.
- [62] Grammann, K., Volke, A. and Kunte, H. J. (2002). 'New Type of Osmoregulated Solute Transporter Identified in Halophilic Members of the Bacteria

Domain: TRAP Transporter TeaABC Mediates Uptake of Ectoine and Hydroxyectoine in *Halomonas elongata* DSM 2581T'. *J. Bacteriol.* 184: 3078–3085.

- [63] Hagemann, M., Richter, S. and Mikkat, S. (1997). 'The *ggtA* gene encodes a subunit of the transport system for the osmoprotective compound glucosylglycerol in *Synechocystis* sp. strain PCC 6803.' *J. Bacteriol.* 179: 714– 720.
- [64] Hagerty, S. B., Groenigen, K. J. van, Allison, S. D., Hungate, B. A., Schwartz, E., Koch, G. W., Kolka, R. K. and Dijkstra, P. (2014). 'Accelerated microbial turnover but constant growth efficiency with warming in soil'. *Nature Clim. Change* 4: 903–906.
- [65] Harder, B. J. (2013). *Generierung und Charakterisierung von Corynebacterium glutamicum Stämmen für die Ectoinproduktion*. Master's thesis. Germany: Technische Universität Braunschweig.
- [66] Hasunuma, T., Sanda, T., Yamada, R., Yoshimura, K., Ishii, J. and Kondo, A. (2011). 'Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of *Saccharomyces cerevisiae*'. *Microb. Cell Fact.* 10: 1–13.
- [67] Hayashi, M., Mizoguchi Hiroshi ad Shiraishi, N., Obayashi, M., Nakagawa, S., Imai, J., Watanabe, S., Ota, T. and Ikeda, M. (2002). 'Transcriptome Analysis of Acetate Metabolism in Corynebacterium glutamicum Using a Newly Developed Metabolic Array'. *Biosci., Biotechnol., Biochem.* 66: 1337– 1344.
- [68] Hermann, T. (2003). 'Industrial production of amino acids by coryneform bacteria'. *J. Biotechnol.* 104: 155–172.
- [69] Hirao, T., Nakano, T., Azuma, T., Sugimoto, M. and Nakanishi, T. (1989).
   'L-Lysine production in continuous culture of an L-lysine hyperproducing mutant of *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 32: (3), 269–273.
- [70] Hoffmann, T., Blohn, C. von, Stanek, A., Moses, S., Barzantny, H. and Bremer,
   E. (2012). 'Synthesis, Release, and Recapture of Compatible Solute Proline
   by Osmotically Stressed *Bacillus subtilis* Cells'. *Appl. Environ. Microbiol.* 78: 5753–5762.
- [71] Ignatova, Z. and Gierasch, L. M. (2006). 'Inhibition of protein aggregation in vitro and in vivo by a natural osmoprotectant'. *P. Natl. Acad. Sci.* 103: 13357–13361.
- [72] Ikeda, M. and Nakagawa, S. (2003). 'The Corynebacterium glutamicum genome: features and impacts on biotechnological processes'. Appl. Microbiol. Biotechnol. 62: (2), 99–109.
- [73] Ikeda, M. (2003). 'Amino Acid Production Processes'. In: *Microbial Production of L-Amino Acids*. Ed. by Faurie, R., Thommel, J., Bathe, B., Debabov, V., Huebner, S., Ikeda, M., Kimura, E., Marx, A., Möckel, B., Mueller, U. and Pfefferle, W. 79: Advances in Biochemical Engineering/Biotechnology. Springer Berlin / Heidelberg, 1–35.

- [74] Ikeda, M. (2006). 'Towards bacterial strains overproducing L-tryptophan and other aromatics by metabolic engineering'. *Appl. Microbiol. Biotechnol.* 69: 615–626.
- [75] Ikeda, M. and Katsumata, R. (1994). 'Transport of aromatic amino acids and its influence on overproduction of the amino acids in *Corynebacterium glutamicum*'. J. Ferment. Bioeng. 78: 420–425.
- [76] Ikeda, M. and Katsumata, R. (1995). 'Tryptophan Production by Transport Mutants of Corynebacterium glutamicum'. Biosci., Biotechnol., Biochem. 59: 1600–1602.
- [77] Ikeda, M., Mitsuhashi, S., Tanaka, K. and Hayashi, M. (2009). 'Reengineering of a *Corynebacterium glutamicum* L-Arginine and L-Citrulline Producer'. *Appl. Environ. Microbiol.* 75: 1635–1641.
- [78] Ikeda, M., Mizuno, Y., Awane, S.-i., Hayashi, M., Mitsuhashi, S. and Takeno, S. (2011). 'Identification and application of a different glucose uptake system that functions as an alternative to the phosphotransferase system in *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 90: 1443–1451.
- [79] Ikeda, M., Noguchi, N., Ohshita, M., Senoo, A., Mitsuhashi, S. and Takeno, S. (2015). 'A third glucose uptake bypass in *Corynebacterium glutamicum* ATCC 31833'. *Appl. Microbiol. Biotechnol.* 99: 2741–2750.
- [80] Ikeda, M., Ohnishi, J., Hayashi, M. and Mitsuhashi, S. (2006). 'A genomebased approach to create a minimally mutated *Corynebacterium glutamicum* strain for efficient L-lysine production'. *J. Ind. Microbiol. Biot.* 33: (7), 610– 615.
- [81] Ikeda, M. and Takeno, S. (2013). 'Amino Acid Production by *Corynebacterium glutamicum*'. In: *Corynebacterium glutamicum*. Ed. by Yukawa, H. and Inui, M. 23: Microbiology Monographs. Springer Berlin Heidelberg, 107–147.
- [82] Inui, M., Kawaguchi, H., Murakami, S., Vertès, A. and Yukawa, H. (2004).
   'Metabolic Engineering of *Corynebacterium glutamicum* for Fuel Ethanol Production under Oxygen-Deprivation Conditions'. J. Mol. Microbiol. Biotechnol. 8: 243–254.
- [83] Inui, M., Murakami, S., Okino, S., Kawaguchi, H., Vertès, A. and Yukawa, H. (2004). 'Metabolic Analysis of *Corynebacterium glutamicum* during Lactate and Succinate Productions under Oxygen Deprivation Conditions'. *J. Mol. Microbiol. Biotechnol.* 7: 182–196.
- [84] Jäger, W., Schäfer, A., Pühler, A., Labes, G. and Wohlleben, W. (1992).
   'Expression of the *Bacillus subtilis* sacB gene leads to sucrose sensitivity in the gram-positive bacterium *Corynebacterium glutamicum* but not in *Streptomyces lividans*.' *J. Bacteriol.* 174: 5462–5465.
- [85] Jain, N. K. and Roy, I. (2008). 'Effect of trehalose on protein structure'. *Protein Sci.* 18: 24–36.
- [86] Jeffries, T. W. (1983). 'Utilization of xylose by bacteria, yeasts, and fungi'. In: *Pentoses and Lignin*. 27: Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, 1–32.

- [87] Jeong, B.-Y., Wittmann, C., Kato, T. and Park, E. Y. (2015). 'Comparative metabolic flux analysis of an *Ashbya gossypii* wild type strain and a high riboflavin-producing mutant strain'. *J. Biosci. Bioeng.* 119: 101–106.
- [88] Jetten, M. S. M. and Sinskey, A. J. (1995). 'Purification and properties of oxaloacetate decarboxylase from *Corynebacterium glutamicum*'. *Antonie van Leeuwenhoek* 67: 221–227.
- [89] Jetten, M. S. and Sinskey, A. J. (1993). 'Characterization of phosphoenolpyruvate carboxykinase from *Corynebacterium glutamicum*'. *FEMS Microbiol. Lett.* 111: 183–188.
- [90] Jojima, T., Fujii, M., Mori, E., Inui, M. and Yukawa, H. (2010). 'Engineering of sugar metabolism of *Corynebacterium glutamicum* for production of amino acid L-alanine under oxygen deprivation'. *Appl. Microbiol. Biotechnol.* 87: 159–165.
- [91] Junker, B. H. (2004). 'Scale-up methodologies for Escherichia coli and yeast fermentation processes'. *J. Biosci. Bioeng.* 97: 347–364.
- [92] Kalinowski, J., Cremer., J., Bachmann, B., Eggeling, L., Sahm, H. and Pühler, A. (1991). 'Genetic and biochemical analysis of the aspartokinase from *Corynebacterium glutamicum*'. *Mol. Microbiol.* 5: 1197–1204.
- [93] Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Krämer, R., Linke, B., McHardy, A. C., Meyer, F., Möckel, B., Pfefferle, W., Pühler, A., Rey, D. A., Rückert, C., Rupp, O., Sahm, H., Wendisch, V. F., Wiegräbe, I. and Tauch, A. (2003). 'The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins'. *J. Biotechnol.* 104: 5–25.
- [94] Karau, A. and Grayson, I. (2014). 'Amino Acids in Human and Animal Nutrition'. In: *Biotechnology of Food and Feed Additives*. Ed. by Zorn, H. and Czermak, P. 143: Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, 189–228.
- [95] Käß, F., Hariskos, I., Michel, A., Brandt, H.-J., Spann, R., Junne, S., Wiechert, W., Neubauer, P. and Oldiges, M. (2013). 'Assessment of robustness against dissolved oxygen/substrate oscillations for *C. glutamicum* DM1933 in twocompartment bioreactor'. *Bioproc. Biosyst. Eng.* 37: 1151–1162.
- [96] Kawaguchi, H., Sasaki, M., Vertès, A. A., Inui, M. and Yukawa, H. (2007).
   'Engineering of an l-arabinose metabolic pathway in Corynebacterium glutamicum'. *Appl. Microbiol. Biotechnol.* 77: 1053–1062.
- [97] Kawaguchi, H., Vertès, A. A., Okino, S., Inui, M. and Yukawa, H. (2006).
   'Engineering of a Xylose Metabolic Pathway in *Corynebacterium glutamicum*'. *Appl. Environ. Microbiol.* 72: 3418–3428.
- [98] Kelle, R., Hermann, T. and Bathe, B. (2005). 'L-Lysine Production'. In: *Handbook of Corynebacterium glutamicum*. CRC Press, 465–488.

- [99] Kempf, B. and Bremer, E. (1998). 'Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments'. *Arch. Microbiol.* 170: (5), 319–330.
- [100] Kiefer, P., Heinzle, E., Zelder, O. and Wittmann, C. (2004). 'Comparative metabolic flux analysis of lysine-producing *Corynebacterium glutamicum* cultured on glucose or fructose'. *Appl. Environ. Microbiol.* 70: 229–239.
- [101] Kim, M. H., Heinzle, E. and Wittmann, C. (2006). 'Deregulation of aspartokinase by single nucleotide exchange leads to global flux rearrangement in the central metabolism of *Corynebacterium glutamicum*'. J. Microbiol. Biotechnol. 16: 1174–1179.
- [102] Kimura, E. (2005). 'L-Glutamate Production'. In: *Handbook of Corynebacterium glutamicum*. CRC Press, 456–463.
- [103] Kind, S., Jeong, W. K., Schröder, H. and Wittmann, C. (2010). 'Systems-wide metabolic pathway engineering in *Corynebacterium glutamicum* for bio-based production of diaminopentane'. *Metab. Eng.* 12: 341–351.
- [104] Kind, S., Jeong, W. K., Schröder, H., Zelder, O. and Wittmann, C. (2010).
   'Identification and Elimination of the Competing *N*-Acetyldiaminopentane Pathway for Improved Production of Diaminopentane by *Corynebacterium glutamicum*'. *Appl. Environ. Microbiol.* 76: 5175–5180.
- [105] Kind, S., Kreye, S. and Wittmann, C. (2011). 'Metabolic engineering of cellular transport for overproduction of the platform chemical 1,5-diaminopentane in *Corynebacterium glutamicum*'. *Metab. Eng.* 13: 617–627.
- [106] Kind, S., Neubauer, S., Becker, J., Yamamoto, M., Völkert, M., Abendroth, G. von, Zelder, O. and Wittmann, C. (2014). 'From zero to hero Production of bio-based nylon from renewable resources using engineered *Corynebacterium glutamicum*'. *Metab. Eng.* 25: 113–123.
- [107] Kinoshita, S., Najayama, K. and Kitada, S. (1958). 'L-Lysine production using microbial auxotroph'. *J. Gen. Appl. Microbiol.* 4: 128–129.
- [108] Kinoshita, S., Udaka, S. and Shimono, M. (1957). 'Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms.' J. Gen. Appl. Microbiol. 3: 193–205.
- [109] Kjeldsen, K. R. and Nielsen, J. (2009). 'In silico genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network'. *Biotechnol. Bioeng.* 102: 583–597.
- [110] Klingner, A., Bartsch, A., Dogs, M., Wagner-Döbler, I., Jahn, D., Simon, M., Brinkhoff, T., Becker, J. and Wittmann, C. (2015). 'Large-Scale 13C Flux Profiling Reveals Conservation of the Entner-Doudoroff Pathway as a Glycolytic Strategy among Marine Bacteria That Use Glucose'. Appl. Environ. Microbiol. 81: 2408–2422.
- [111] Kohlstedt, M. (2014). A multi-omics perspective on osmoadaptation and osmoprotection in Bacillus subtilis. PhD thesis. Universität des Saarlandes.
- [112] Kohlstedt, M., Becker, J. and Wittmann, C. (2010). 'Metabolic fluxes and beyond—systems biology understanding and engineering of microbial metabolism'. *Appl. Microbiol. Biotechnol.* 88: (5), 1065–1075.

- [113] Krause, F. S., Henrich, A., Blombach, B., Krämer, R., Eikmanns, B. J. and Seibold, G. M. (2010). 'Increased Glucose Utilization in *Corynebacterium glutamicum* by Use of Maltose, and its Application for the Improvement of L-Valine Productivity'. *Appl. Environ. Microbiol.* 76: 370–374.
- [114] Krings, E., Krumbach, K., Bathe, B., Kelle, R., Wendisch, V. F., Sahm, H. and Eggeling, L. (2006). 'Characterization of myo-Inositol Utilization by *Corynebacterium glutamicum*: the Stimulon, Identification of Transporters, and Influence on L-Lysine Formation'. J. Bacteriol. 188: 8054–8061.
- [115] Kröger, B., Zelder, O., Klopprogge, C., Schröder, H. and Haefner, S. (2005). *Pgro expression units*.
- [116] Krömer, J. O., Fritz, M., Heinzle, E. and Wittman, C. (2005). 'In vivo quantification of intracellular amino acids and intermediates of the methionine pathway in *Corynebacterium glutamicum*'. *Anal. Biochem.* 340: 171–173.
- [117] Krömer, J. O., Sorgenfrei, O., Klopprogge, K., Heinzle, E. and Wittmann, C. (2004). 'In-Depth Profiling of Lysine-Producing *Corynebacterium glutamicum* by Combined Analysis of the Transcriptome, Metabolome, and Fluxome'. *J. Bacteriol.* 186: 1769–1784.
- [118] Kuhlmann, A. U. and Bremer, E. (2002). 'Osmotically Regulated Synthesis of the Compatible Solute Ectoine in *Bacillus pasteurii* and Related Bacillus spp.' *Appl. Environ. Microbiol.* 68: 772–783.
- [119] Kuhlmann, A. U., Bursy, J., Gimpel, S., Hoffmann, T. and Bremer, E. (2008).
   'Synthesis of the Compatible Solute Ectoine in *Virgibacillus pantothenticus* Is Triggered by High Salinity and Low Growth Temperature'. *Appl. Environ. Microbiol.* 74: 4560–4563.
- [120] Kunte, H. J., Lentzen, G. and Galinski, E. A. (2014). 'Industrial Production of the Cell Protectant Ectoine: Protection Mechanisms, Processes, and Products'. *Current Biotechnology* 3: 10–25.
- [121] Lang, Y.-j., Bai, L., Ren, Y.-n., Zhang, L.-h. and Nagata, S. (2011). 'Production of ectoine through a combined process that uses both growing and resting cells of *Halomonas salina* DSM 5928T'. *Extremophiles* 15: 303–310.
- [122] Lee, J. W., Na, D., Park, J. M., Lee, J., Choi, S. and Lee, S. Y. (2012). 'Systems metabolic engineering of microorganisms for natural and nonnatural chemicals'. *Nat. Chem. Biol.* 8: 536–546.
- [123] Lee, S. J., Song, H. and Lee, S. Y. (2006). 'Genome-Based Metabolic Engineering of *Mannheimia succiniciproducens* for Succinic Acid Production'. *Appl. Environ. Microbiol.* 72: 1939–1948.
- [124] Lentzen, G. and Schwarz, T. (2006). 'Extremolytes: natural compounds from extremophiles for versatile applications'. *Appl. Microbiol. Biotechnol.* 72: 623–634.
- [125] Leßmeier, L., Zahoor ul Hassan, A., Lindner, S. and Wendisch, V. F. (2015). Metabolic Engineering of Corynebacterium glutamicum for Alternative Carbon Source Utilization. In: ed. by Burkovski, A. Corynebacterium glutamicum: From Systems Biology to Biotechnological Applications. Caister Academic Press. Chap. 5, 57–70.

- [126] Leuchtenberger, W. (2008). *Amino Acids Technical Production and Use*. In: *Biotechnology*. Wiley-VCH Verlag GmbH, 465–502.
- [127] Liebl, W. (2005). 'Corynebacterium taxonomy'. In: Handbook of Corynebacterium glutamicum. CRC Press, 9–35.
- [128] Lindner, S. N., Seibold, G. M., Henrich, A., Krämer, R. and Wendisch, V. F. (2011). 'Phosphotransferase System-Independent Glucose Utilization in *Co-rynebacterium glutamicum* by Inositol Permeases and Glucokinases'. *Appl. Environ. Microbiol.* 77: 3571–3581.
- [129] Lindner, S. N., Niederholtmeyer, H., Schmitz, K., Schoberth, S. M. and Wendisch, V. F. (2010). 'Polyphosphate/ATP-dependent NAD kinase of *Corynebacterium glutamicum*: biochemical properties and impact of ppnK overexpression on lysine production'. *Appl. Microbiol. Biotechnol.* 87: 583–593.
- [130] Liu, S., Lu, H., Hu, R., Shupe, A., Lin, L. and Liang, B. (2012). 'A sustainable woody biomass biorefinery'. *Biotechnol. Adv.* 30: 785–810.
- [131] Lo, C.-C., Bonner, C. A., Xie, G., D'Souza, M. and Jensen, R. A. (2009).
   'Cohesion Group Approach for Evolutionary Analysis of Aspartokinase, an Enzyme That Feeds a Branched Network of Many Biochemical Pathways'. *Microbiol. Mol. Biol. R.* 73: 594–651.
- [132] Louis, P. and Galinski, E. A. (1997). 'Characterization of genes for the biosynthesis of the compatible solute ectoine from Marinococcus halophilus and osmoregulated expression in *Escherichia coli*'. *Microbiology* 143: 1141–1149.
- [133] Mainzer, S. E. and Hempfling, W. P. (1976). 'Effects of growth temperature on yield and maintenance during glucose-limited continuous culture of *Escherichia coli*.' *J. Bacteriol.* 126: 251–256.
- [134] Maris, A. van, Abbott, D., Bellissimi, E., Brink, J. van den, Kuyper, M., Luttik, M., Wisselink, H., Scheffers, W., Dijken, J. van and Pronk, J. (2006). 'Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status'. *Antonie van Leeuwenhoek* 90: (4), 391–418.
- [135] Marx, A., Hans, S., Möckel, B., Bathe, B. and Graaf, A. A. de (2003). 'Metabolic phenotype of phosphoglucose isomerase mutants of *Corynebacterium glutamicum*'. J. Biotechnol. 104: 185–197.
- [136] Marx, A., Striegel, K., Graaf, A. A. de, Sahm, H. and Eggeling, L. (1997). 'Response of the central metabolism of *Corynebacterium glutamicum* to different flux burdens'. *Biotechnol. Bioeng.* 56: 168–180.
- [137] Meier-Wagner, J., Nolden, L., Jakoby, M., Siewe, R., Krämer, R. and Burkovski, A. (2001). 'Multiplicity of ammonium uptake systems in *Corynebacterium glutamicum*: role of Amt and AmtB'. *Microbiology* 147: 135–143.
- [138] Meiswinkel, T. M., Gopinath, V., Lindner, S. N., Nampoothiri, K. M. and Wendisch, V. F. (2013). 'Accelerated pentose utilization by *Corynebacterium glutamicum* for accelerated production of lysine, glutamate, ornithine and putrescine'. *Microb. Biotechnol.* 6: 131–140.
- [139] Meiswinkel, T. M., Rittmann, D., Lindner, S. N. and Wendisch, V. F. (2013).
   'Crude glycerol-based production of amino acids and putrescine by *Corynebacterium glutamicum*'. *Bioresour. Technol.* 145: 254–258.

- [140] Melzer, G., Esfandabadi, M., Franco-Lara, E. and Wittmann, C. (2009). 'Flux Design: In silico design of cell factories based on correlation of pathway fluxes to desired properties'. *BMC Syst. Biol.* 3: 1–16.
- [141] Moon, M.-W., Park, S.-Y., Choi, S.-K. and Lee, J.-K. (2007). 'The Phosphotransferase System of *Corynebacterium glutamicum*: Features of Sugar Transport and Carbon Regulation'. *J. Mol. Microbiol. Biotechnol.* 12: 43–50.
- [142] N/A (2013). Lysine and other Amino Acids (Methionine, Threonine & Tryptophan) Market by Application (Animal Feed, Food & Dietary Supplements, Pharmaceuticals), by Livestock (Swine, Poultry, Others): Global Industry Analysis, Size, Share, Growth, Trends and Forecast, 2012 2018. Tech. rep. Transparency Market Research.
- [143] Nakayama, K. (1972). *Lysine and Diaminopimelic acid*. In: *The microbial production of amino acids*. Halsted Press, 369–397.
- [144] Nara, T., Samejima, H., Fujita, C., Ito, M., Nakayama, K. and Kinoshita, S. (1961). 'L-Homoserine Fermentation'. *Agric. Biol. Chem.* 25: 532–541.
- [145] Neuner, A. (2012). *Metabolic engineering of Corynebacterium glutamicum for* L-lysine production on silage. PhD thesis. Universität des Saarlandes.
- [146] Neuner, A. and Heinzle, E. (2011). 'Mixed glucose and lactate uptake by *Corynebacterium glutamicum* through metabolic engineering'. *Biotech. J.* 6: 318–329.
- [147] Ning, Y., Wu, X., Zhang, C., Xu, Q., Chen, N. and Xie, X. (2016). 'Pathway construction and metabolic engineering for fermentative production of ectoine in *Escherichia coli*'. *Metab. Eng.* 36: 10–18.
- [148] Noor, E., Eden, E., Milo, R. and Alon, U. (2010). 'Central Carbon Metabolism as a Minimal Biochemical Walk between Precursors for Biomass and Energy'. *Mol. Cell* 39: 809–820.
- [149] Ohnishi, J., Hayashi, M., Mitsuhashi, S. and Ikeda, M. (2003). 'Efficient 40°C fermentation of L-lysine by a new *Corynebacterium glutamicum* mutant developed by genome breeding'. *Appl. Microbiol. Biotechnol.* 62: (1), 69–75.
- [150] Ohnishi, J., Mitsuhashi, S., Hayashi, M., Ando, S., Yokoi, H., Ochiai, K. and Ikeda, M. (2002). 'A novel methodology employing *Corynebacterium glutamicum* genome information to generate a new L-lysine-producing mutant'. *Appl. Microbiol. Biotechnol.* 58: (2), 217–223.
- [151] Ohnishi, J. and Ikeda, M. (2006). 'Comparisons of Potentials for L-Lysine Production among Different *Corynebacterium glutamicum* Strains'. *Biosci., Biotechnol., Biochem.* 70: 1017–1020.
- [152] Ohnishi, J., Katahira, R., Mitsuhashi, S., Kakita, S. and Ikeda, M. (2005). 'A novel *gnd* mutation leading to increased L-lysine production in *Corynebacterium glutamicum*'. *FEMS Microbiol. Lett.* 242: 265–274.
- [153] Okafor, N. (2007). *Modern Industrial Microbiology and Biotechnology*. CRC Press.
- [154] Okino, S., Inui, M. and Yukawa, H. (2005). 'Production of organic acids by *Corynebacterium glutamicum* under oxygen deprivation'. *Appl. Microbiol. Biotechnol.* 68: 475–480.

- [155] Okino, S., Noburyu, R., Suda, M., Jojima, T., Inui, M. and Yukawa, H. (2008).
  'An efficient succinic acid production process in a metabolically engineered *Corynebacterium glutamicum* strain'. *Appl. Microbiol. Biotechnol.* 81: (3), 459–464.
- [156] Okino, S., Suda, M., Fujikura, K., Inui, M. and Yukawa, H. (2008). 'Production of D-lactic acid by *Corynebacterium glutamicum* under oxygen deprivation'. *Appl. Microbiol. Biotechnol.* 78: 449–454.
- [157] Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M. and Murooka, Y. (1999). 'Characterization of Biosynthetic Enzymes for Ectoine as a Compatible Solute in a Moderately Halophilic Eubacterium, *Halomonas elongata*'. J. Bacteriol. 181: 91–99.
- [158] Ooyen, J. van, Noack, S., Bott, M., Reth, A. and Eggeling, L. (2012). 'Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity'. *Biotechnol. Bioeng.* 1–12.
- [159] Papini, M., Salazar, M. and Nielsen, J. (2010). 'Systems Biology of Industrial Microorganisms'. In: *Biosystems Engineering I*. Ed. by Wittmann, C. and Krull, R. 120: Advances in Biochemical Engineering/Biotechnology. Springer Berlin / Heidelberg, 51–99.
- [160] Pastor, J. M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L. N., Iborra, J. L., Vargas, C., Nieto, J. J. and Cánovas, M. (2010).
  'Ectoines in cell stress protection: Uses and biotechnological production'. *Biotechnol. Adv.* 28: 782–801.
- [161] Peters-Wendisch, P., Stolz, M., Etterich, H., Kennerknecht, N., Sahm, H. and Eggeling, L. (2005). 'Metabolic Engineering of *Corynebacterium glutamicum* for L-Serine Production'. *Appl. Environ. Microbiol.* 71: 7139–7144.
- [162] Peters-Wendisch, P., Wendisch, V. F., Paul, S., Eikmanns, B. J. and Sahm, H. (1997). 'Pyruvate carboxylase as an anaplerotic enzyme in *Corynebacterium glutamicum*'. *Microbiology* 143: 1095–1103.
- [163] Pfefferle, W., Möckel, B., Bathe, B. and Marx, A. (2003). 'Biotechnological Manufacture of Lysine'. In: *Microbial Production of L-Amino Acids*. Ed. by Faurie, R., Thommel, J., Bathe, B., Debabov, V., Huebner, S., Ikeda, M., Kimura, E., Marx, A., Möckel, B., Mueller, U. and Pfefferle, W. 79: Advances in Biochemical Engineering/Biotechnology. Springer Berlin / Heidelberg, 59–112.
- [164] Pompejus, M., Schröder, H., Kröger, B. and Zelder, O. (2008). *Method for altering the genome of corynebacteria*.
- [165] Prabhu, J., Schauwecker, F., Grammel, N., Keller, U. and Bernhard, M. (2004).
   'Functional Expression of the Ectoine Hydroxylase Gene (thpD) from *Streptomyces chrysomallus* in *Halomonas elongata*'. *Appl. Environ. Microbiol.* 70: 3130–3132.
- [166] Quek, L.-E. and Nielsen, L. (2014). 'Steady-State 13C Fluxomics Using Open-FLUX'. In: *Metabolic Flux Analysis*. Ed. by Krömer, J. O., Nielsen, L. K. and Blank, L. M. 1191: Methods in Molecular Biology. Springer New York, 209– 224.

- [167] Quek, L., Wittmann, C., Nielsen, L. and Krömer, J. (2009). 'OpenFLUX: efficient modelling software for <sup>13</sup>C-based metabolic flux analysis'. *Microb. Cell. Fact.* 8: 1–15.
- [168] Razak, M. A. and Viswanath, B. (2014). 'Optimization of fermentation upstream parameters and immobilization of *Corynebacterium glutamicum* MH 20-22 B cells to enhance the production of L-lysine'. *3 Biotech* 5: 531–540.
- [169] Reshetnikov, A. S., Khmelenina, V. N. and Trotsenko, Y. A. (2006). 'Characterization of the ectoine biosynthesis genes of haloalkalotolerant obligate methanotroph "Methylomicrobium alcaliphilum 20Z". Arch. Microbiol. 184: 286–297.
- [170] Reuter, K., Pittelkow, M., Bursy, J., Heine, A., Craan, T. and Bremer, E. (2010).
   'Synthesis of 5-Hydroxyectoine from Ectoine: Crystal Structure of the Non-Heme Iron(II) and 2-Oxoglutarate-Dependent Dioxygenase EctD'. *PLoS One* 5: 1–11.
- [171] Rittmann, D., Lindner, S. N. and Wendisch, V. F. (2008). 'Engineering of a Glycerol Utilization Pathway for Amino Acid Production by *Corynebacterium glutamicum*'. *Appl. Environ. Microbiol.* 74: 6216–6222.
- [172] Rodríguez-Moya, J., Argandoña, M., Iglesias-Guerra, F., Nieto, J. J. and Vargas, C. (2013). 'Temperature- and Salinity-Decoupled Overproduction of Hydroxyectoine by *Chromohalobacter salexigens*'. *Appl. Environ. Microbiol.* 79: 1018–1023.
- [173] Sakai, S., Tsuchida, Y., Okino, S., Ichihashi, O., Kawaguchi, H., Watanabe, T., Inui, M. and Yukawa, H. (2007). 'Effect of Lignocellulose-Derived Inhibitors on Growth of and Ethanol Production by Growth-Arrested *Corynebacterium glutamicum* R'. *Appl. Environ. Microbiol.* 73: 2349–2353.
- [174] Sasaki, M., Jojima, T., Kawaguchi, H., Inui, M. and Yukawa, H. (2009).
   'Engineering of pentose transport in *Corynebacterium glutamicum* to improve simultaneous utilization of mixed sugars'. *Appl. Microbiol. Biotechnol.* 85: (1), 105–115.
- [175] Sauer, T. and Galinski, E. A. (1998). 'Bacterial milking: A novel bioprocess for production of compatible solutes'. *Biotechnol. Bioeng.* 57: 306–313.
- [176] Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G. and Pühler, A. (1994). 'Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*'. *Gene* 145: 69–73.
- [177] Schaffer, S. and Burkovski, A. (2005). 'Proteomics'. In: *Handbook of Coryne*bacterium glutamicum. CRC Press, 99–118.
- [178] Schiraldi, C., Maresca, C., Catapano, A., Galinski, E. A. and Rosa, M. D. (2006). 'High-yield cultivation of *Marinococcus* M52 for production and recovery of hydroxyectoine'. *Res. Microbiol.* 157: 693–699.
- [179] Schneider, J., Niermann, K. and Wendisch, V. (2011). 'Production of the amino acids L-glutamate, L-lysine, L-ornithine and L-arginine from arabinose by recombinant *Corynebacterium glutamicum*'. *J. Biotechnol.* 154: 191–198.

- [180] Schneider, K., Schütz, V., John, G. and Heinzle, E. (2010). 'Optical device for parallel online measurement of dissolved oxygen and pH in shake flask cultures'. *Bioproc. Biosyst. Eng.* 33: (5), 541–547.
- [181] Schnoor, M., Voß, P., Cullen, P., Böking, T., Galla, H.-J., Galinski, E. A. and Lorkowski, S. (2004). 'Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer'. *Biochem. Biophys. Res. Commun.* 322: 867–872.
- [182] Schrumpf, B., Schwarzer, A., Kalinowski, J., Puhler, A., Eggeling, L. and Sahm, H. (1991). 'A functionally split pathway for lysine synthesis in *Corynebacterium glutamicum*'. *J. Bacteriol.* 173: 4510–4516.
- [183] Schubert, T., Maskow, T., Benndorf, D., Harms, H. and Breuer, U. (2007). 'Continuous Synthesis and Excretion of the Compatible Solute Ectoine by a Transgenic, Nonhalophilic Bacterium'. *Appl. Environ. Microbiol.* 73: 3343– 3347.
- [184] Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., Rampp, M., Schuster, S. C., Klenk, H.-P., Pfeiffer, F., Oesterhelt, D. and Kunte, H. J. (2011). 'A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581T'. *Environ. Microbiol.* 13: 1973–1994.
- [185] Seibold, G., Auchter, M., Berens, S., Kalinowski, J. and Eikmanns, B. J. (2006).
  'Utilization of soluble starch by a recombinant *Corynebacterium glutamicum* strain: Growth and lysine production'. *J. Biotechnol.* 124: 381–391.
- [186] Seip, B., Galinski, E. A. and Kurz, M. (2011). 'Natural and Engineered Hydroxyectoine Production Based on the *Pseudomonas stutzeri ectABCD-ask* Gene Cluster'. *Appl. Environ. Microbiol.* 77: 1368–1374.
- [187] Shah, A. H., Hameed, A., Ahmad, S. and Khan, G. M. (2002). 'Optimization of Culture Conditions for L-Lysine Fermentation by *Corynebacterium glutamicum*'. *OnLine Journal of Biological Sciences* 2: 151–156.
- [188] Shimizu, H. and Hirasawa, T. (2007). 'Production of Glutamate and Glutamate-Related Amino Acids: Molecular Mechanism Analysis and Metabolic Engineering'. In: Amino Acid Biosynthesis – Pathways, Regulation and Metabolic Engineering. Ed. by Wendisch, V. 5: Microbiology Monographs. Springer Berlin Heidelberg, 1–38.
- [189] Shinfuku, Y., Sorpitiporn, N., Sono, M., Furusawa, C., Hirasawa, T. and Shimizu, H. (2009). 'Development and experimental verification of a genomescale metabolic model for *Corynebacterium glutamicum*'. *Microb. Cell Fact.* 8: 1–15.
- [190] Shiratsuchi, M., Kuronuma, H., Kawahara, Y., Yoshihara, Y., Miwa, H. and Nakamori, S. (1995). 'Simultaneous and High Fermentative Production of L-Lysine and L-Glutamic Acid Using a Strain of *Brevibacterium lactofermentum*'. *Biosci, Biotechnol, Biochem* 59: 83–86.
- [191] Siewe, R. M., Weil, B., Burkovski, A., Eikmanns, B. J., Eikmanns, M. and Krämer, R. (1996). 'Functional and Genetic Characterization of the (Methyl)-

ammonium Uptake Carrier of *Corynebacterium glutamicum*'. *J. Biol. Chem.* 271: 5398–5403.

- [192] Sindelar, G. and Wendisch, V. (2007). 'Improving lysine production by *Corynebacterium glutamicum* through DNA microarray-based identification of novel target genes'. *Appl. Microbiol. Biotechnol.* 76: 677–689.
- [193] Sjostrom, E. (1981). Wood Chemistry: Fundamentals and Applications. Academic Press Inc.
- [194] Sonntag, K., Eggeling, L., Graaf, A. A. de and Sahm, H. (1993). 'Flux partitioning in the split pathway of lysine synthesis in Corynebacterium glutamicum'. *Eur. J. Biochem.* 213: 1325–1331.
- [195] Spaans, S. K., Weusthuis, R. A., Van Der Oost, J. and Kengen, S. W. (2015). 'NADPH-generating systems in bacteria and archaea'. *Front. Microbiol.* 6:
- [196] Stäbler, N., Oikawa, T., Bott, M. and Eggeling, L. (2011). '*Corynebacterium glutamicum* as a Host for Synthesis and Export of D-Amino Acids'. *J. Bacteriol.* 193: 1702–1709.
- [197] Stephanopoulos, G. and Vallino, J. (1991). 'Network rigidity and metabolic engineering in metabolite overproduction'. *Science* 252: 1675–1681.
- [198] Stöveken, N., Pittelkow, M., Sinner, T., Jensen, R. A., Heider, J. and Bremer, E. (2011). 'A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501'. *J. Bacteriol.* 193: 4456–4468.
- [199] Street, T. O., Bolen, D. W. and Rose, G. D. (2006). 'A molecular mechanism for osmolyte-induced protein stability'. *P. Natl. Acad. Sci.* 103: 13997–14002.
- [200] Street, T. O., Krukenberg, K. A., Rosgen, J., Bolen, D. W. and Agard, D. A. (2010). 'Osmolyte-induced conformational changes in the Hsp90 molecular chaperone'. *Protein Sci.* 19: 57–65.
- [201] Subramaniyan, S. and Prema, P. (2002). 'Biotechnology of Microbial Xylanases: Enzymology, Molecular Biology, and Application'. *Crit. Rev. Biotechnol.* 22: 33–64.
- [202] Sun, Y. and Cheng, J. (2002). 'Hydrolysis of lignocellulosic materials for ethanol production: a review'. *Bioresour. Technol.* 83: 1–11.
- [203] Suzuki, N. and Inui, M. (2013). 'Genome Engineering of Corynebacterium glutamicum'. In: Corynebacterium glutamicum. Ed. by Yukawa, H. and Inui, M. 23: Microbiology Monographs. Springer Berlin Heidelberg, 89–105.
- [204] Sydlik, U., Gallitz, I., Albrecht, C., Abel, J., Krutmann, J. and Unfried, K. (2009). 'The Compatible Solute Ectoine Protects against Nanoparticleinduced Neutrophilic Lung Inflammation'. *Am. J. Respir. Crit. Care Med.* 180: 29–35.
- [205] Tateno, T., Fukuda, H. and Kondo, A. (2007). 'Direct production of L-lysine from raw corn starch by *Corynebacterium glutamicum* secreting *Streptococcus bovis* α-amylase using *cspB* promoter and signal sequence'. *Appl. Microbiol. Biotechnol.* 77: 533–541.

- [206] Tateno, T., Fukuda, H. and Kondo, A. (2007). 'Production of L-Lysine from starch by *Corynebacterium glutamicum* displaying α-amylase on its cell surface'. *Appl. Microbiol. Biotechnol.* 74: (6), 1213–1220.
- [207] Tauch, A., Homann, I., Mormann, S., Rüberg, S., Billault, A., Bathe, B., Brand, S., Brockmann-Gretza, O., Rückert, C., Schischka, N., Wrenger, C., Hoheisel, J., Möckel, B., Huthmacher, K., Pfefferle, W., Pühler, A. and Kalinowski, J. (2002). 'Strategy to sequence the genome of *Corynebacterium glutamicum* ATCC 13032: use of a cosmid and a bacterial artificial chromosome library'. *J. Biotechnol.* 95: 25–38.
- [208] Tauch, A., Pühler, A., Kalinowski, J. and Thierbach, G. (2003). 'Plasmids in *Corynebacterium glutamicum* and their molecular classification by comparative genomics'. *J. Biotechnol.* 104: 27–40.
- [209] UN (1992). Convention on biological diversity. URL: https://www.cbd. int/convention/articles/default.shtml?a=cbd-02.
- [210] UN (2015). Poultry data. URL: http://faostat3.fao.org/browse/Q/ QL/E.
- [211] Vallino, J. J. and Stephanopoulos, G. (1993). 'Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction'. *Biotechnol. Bioeng.* 41: 633–646.
- [212] Van-Thuoc, D., Guzmán, H., Quillaguamán, J. and Hatti-Kaul, R. (2010).
   'High productivity of ectoines by *Halomonas boliviensis* using a combined two-step fed-batch culture and milking process'. *J. Biotechnol.* 147: 46–51.
- [213] Villadsen, J., Nielsen, J. and Lidén, G. (2011). 'Thermodynamics of Bioreactions'. In: *Bioreaction Engineering Principles*. Springer US, 119–150.
- [214] Vogt, M., Haas, S., Klaffl, S., Polen, T., Eggeling, L., Ooyen, J. van and Bott, M. (2014). 'Pushing product formation to its limit: Metabolic engineering of *Corynebacterium glutamicum* for L-leucine overproduction'. *Metab. Eng.* 22: 40–52.
- [215] Voronovsky, A. Y., Ryabova, O. B., Verba, O. V., Ishchuk, O. P., Dmytruk, K. V. and Sibirny, A. A. (2005). 'Expression of *xylA* genes encoding xylose isomerases from *Escherichia coli* and *Streptomyces coelicolor* in the methylotrophic yeast *Hansenula polymorpha*'. *FEMS Yeast Res.* 5: 1055–1062.
- [216] Vrljić, M., Garg, J., Bellmann, A., Wachi, S., Freudi, R., Malecki, M., Sahm, H., Kozina, V., Eggeling, L. and Saier Jr, M. (1999). 'The LysE superfamily: topology of the lysine exporter LysE of *Corynebacterium glutamicum*, a paradyme for a novel superfamily of transmembrane solute translocators'. J. *Mol. Microbiol. Biotechnol.* 1: 327–336.
- [217] Vrljić, M., Sahm, H. and Eggeling, L. (1996). 'A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*'. *Mol. Microbiol.* 22: 815–826.
- [218] Wachi, M. (2013). 'Amino Acid Exporters in Corynebacterium glutamicum'. In: Corynebacterium glutamicum. Ed. by Yukawa, H. and Inui, M. 23: Microbiology Monographs. Springer Berlin Heidelberg, 335–349.

- [219] Wang, C., Cai, H., Zhou, Z., Zhang, K., Chen, Z., Chen, Y., Wan, H. and Ouyang, P. (2014). 'Investigation of *ptsG* gene in response to xylose utilization in *Corynebacterium glutamicum*'. *J. Ind. Microbiol. Biot.* 41: 1249–1258.
- [220] Wang, C., Zhang, H., Cai, H., Zhou, Z., Chen, Y., Chen, Y. and Ouyang, P. (2014). 'Succinic Acid Production from Corn Cob Hydrolysates by Genetically Engineered Corynebacterium glutamicum'. Appl. Biochem. Biotechnol. 172: 340–350.
- [221] Website (2016). URL: http://www.tradingeconomics.com/commodity/ sugar.
- [222] Website (2016). URL: https://www.kdbdw.com/bbs/board/message/ view.do?categoryId=76&messageId=195321.
- [223] Wei, L., Wedeking, A., Büttner, R., Kalff, J., Tolba, R. and Echten-Deckert, G. van (2009). 'A Natural Tetrahydropyrimidine Protects Small Bowel from Cold Ischemia and Subsequent Warm in vitro Reperfusion Injury'. *Pathobiology* 76: 212–220.
- [224] Weinand, M., Krämer, R. and Morbach, S. (2007). 'Characterization of compatible solute transporter multiplicity in *Corynebacterium glutamicum*'. *Appl. Microbiol. Biotechnol.* 76: 701–708.
- [225] Welsh, D. T. (2000). 'Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate'. *FEMS Microbiol. Rev.* 24: 263–290.
- [226] Wendisch, V. F., Bott, M., Kalinowski, J., Oldiges, M. and Wiechert, W. (2006).
   'Emerging *Corynebacterium glutamicum* systems biology'. J. Biotechnol. 124: 74–92.
- [227] Werpy, T., Petersen, G., Aden, A., Bozell, J., Holladay, J., White, J., Manheim, A., Eliot, D., Lasure, L., Jones, S. et al. (2004). *Top Value Added Chemicals From Biomass. Volume 1-Results of Screening for Potential Candidates From Sugars and Synthesis Gas.*
- [228] Winden, W. A. van, Wittmann, C., Heinzle, E. and Heijnen, J. J. (2002).'Correcting mass isotopomer distributions for naturally occurring isotopes'. *Biotechnol. Bioeng.* 80: 477–479.
- [229] Wittmann, C. (2007). 'Fluxome analysis using GC-MS'. *Microb. Cell Fact.* 6: 1–17.
- [230] Wittmann, C. and Becker, J. (2007). 'The L-Lysine Story: From Metabolic Pathways to Industrial Production'. In: Amino Acid Biosynthesis - Pathways, Regulation and Metabolic Engineering. Ed. by Wendisch, V. 5: Microbiology Monographs. Springer Berlin / Heidelberg, 39–70.
- [231] Wittmann, C. and Graaf, A. A. de (2005). 'Metabolic Flux Analysis in *Corynebacterium glutamicum*'. In: *Handbook of Corynebacterium glutamicum*. Ed. by Eggeling, L. and Bott, M. CRC Press. Chap. 12, 277–304.
- [232] Wittmann, C. and Heinzle, E. (2001). 'Application of MALDI-TOF MS to lysine-producing *Corynebacterium glutamicum*'. *Eur. J. Biochem.* 268: 2441–2455.

- [233] Wittmann, C. and Heinzle, E. (2001). 'Modeling and experimental design for metabolic flux analysis of lysine-producing *Corynebacteria* by mass spectrometry'. *Metab. Eng.* 3: 173–191.
- [234] Wittmann, C. and Heinzle, E. (2002). 'Genealogy Profiling through Strain Improvement by Using Metabolic Network Analysis: Metabolic Flux Genealogy of Several Generations of Lysine-Producing Corynebacteria'. *Appl. Environ. Microbiol.* 68: 5843–5859.
- [235] Wittmann, C., Kiefer, P. and Zelder, O. (2004). 'Metabolic fluxes in *Corynebacterium glutamicum* during lysine production with sucrose as carbon source.' *Appl. Environ. Microbiol.* 70: 7277–7287.
- [236] Wittmann, C. and Lee, S. Y. (2012). 'Preface'. In: *Systems Metabolic Engineering*. Ed. by Wittmann, C. and Lee, S. Y. Springer Netherlands, vii–ix.
- [237] Wood, J. M. (2015). 'Bacterial responses to osmotic challenges'. J. Gen. Physiol. 145: 381–388.
- [238] Xu, J., Han, M., Zhang, J., Guo, Y., Qian, H. and Zhang, W. (2013). 'Improvement of L-lysine production combines with minimization of by-products synthesis in *Corynebacterium glutamicum*'. *J. Chem. Technol. Biotechnol.* 1924–1933.
- [239] Xu, J., Han, M., Zhang, J., Guo, Y. and Zhang, W. (2014). 'Metabolic engineering *Corynebacterium glutamicum* for the L-lysine production by increasing the flux into L-lysine biosynthetic pathway'. *Amino Acids* 46: 2165–2175.
- [240] Yan, Y., Yang, J., Dou, Y., Chen, M., Ping, S., Peng, J., Lu, W., Zhang, W., Yao, Z., Li, H., Liu, W., He, S., Geng, L., Zhang, X., Yang, F., Yu, H., Zhan, Y., Li, D., Lin, Z., Wang, Y., Elmerich, C., Lin, M. and Jin, Q. (2008). 'Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501'. P. Natl. Acad. Sci. 105: 7564–7569.
- [241] Yancey, P. H. (2005). 'Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses'. *J. Exp. Biol.* 208: 2819–2830.
- [242] Yokota, A. and Lindley, N. (2005). 'Central Metabolism: Sugar uptake and conversion'. In: Handbook of Corynebacterium glutamicum. CRC Press, 215– 240.
- [243] Yoosin, S. and Sorapipatana, C. (2007). 'A study of ethanol production cost for gasoline substitution in Thailand and its competitiveness'. *Thammasat Int J Sci Technol* 12: 69–80.
- [244] Yoshida, A., Tomita, T., Kurihara, T., Fushinobu, S., Kuzuyama, T. and Nishiyama, M. (2007). 'Structural Insight into Concerted Inhibition of  $\alpha 2\beta 2$ -Type Aspartate Kinase from *Corynebacterium glutamicum*'. *J. Mol. Biol.* 368: 521–536.
- [245] Yoshikawa, K., Furusawa, C., Hirasawa, T. and Shimizu, H. (2012). 'Design of Superior Cell Factories Based on Systems Wide Omics Analysis'. In: Systems Metabolic Engineering. Ed. by Wittmann, C. and Lee, S. Y. Springer Netherlands, 57–81.

- [246] Young, V. R. and Pellett, P. L. (1994). 'Plant proteins in relation to human protein and amino acid nutrition.' *Amer J Clin Nutr* 59: 1203–1212.
- [247] Zahoor, A., Lindner, S. and Wendisch, V. (2012). 'Metabolic engineering of *Corynebacterium glutamicum* aimed at alternative carbon sources and new products'. *Comput. Struct. Biotechnol. J.* 3: 1–11.