Carbon Catabolism in *Bacillus subtilis*: Global and Molecular Views on the Control of Gene Expression

Dissertation zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Oliver Schilling
aus Bamberg

Göttingen 2007

D 7

Referent: Prof. Dr. Jörg Stülke

Korreferent: Prof. Dr. Ralf Ficner

Tag der mündlichen Prüfung:

Danksagung

An erster Stelle möchte ich ganz herzlich Prof. Dr. Jörg Stülke für die Stellung des interessanten Themas und für die gute Betreuung der Arbeit danken. Er hat mir für die Verwirklichung meiner Ideen alle nötigen Freiheiten gelassen und brachte mir dadurch großes Vertrauen entgegen. Gleichwohl konnte ich immer auf seine Hilfsbereitschaft zählen, wenn sie gebraucht wurde.

Desweiteren bedanke ich mich bei Herrn Prof. Dr. Ralf Ficner für die freundliche Übernahme des Korreferats und für die Zusammenarbeit bei einem Projekt zur strukturellen Aufklärung des GlcT-RAT Komplexes. Für die finanzielle Unterstützung der Arbeit bedanke ich mich beim Fonds der Chemischen Industrie und bei der Deutschen Forschungsgemeinschaft.

Bei Dr. Armin Ehrenreich bedanke ich mich für die Einführung in die Welt der Microarrays und für seine stete Hilfe bei allen technischen Fragen rund ums Spotten, Labeln und Hybridisieren. Prof. Dr. Elmar Heinzle, PD Dr. Christoph Wittmann und Oliver Frick danke ich für die wertvolle Zusammenarbeit, aber auch dafür, dass ich eine Woche in ihrer Arbeitsgruppe an der Universität Saarbrücken mitarbeiten durfte. In diesem Zusammenhang bedanke ich mich auch bei Michael Fritz und Tae Hoon Yang für die freundliche Mithilfe bei diesem Projekt. Für die Bereitstellung von Bakterienstämmen bedanke ich mich bei Prof. Dr. Richard Herrmann (Universität Heidelberg), Dr. Ulf Gerth (Universität Greifswald) und Dominique Le Coq (Laboratoire de Génétique des Microorganismes, Thiverval-Grignon). Bei Dr. Matthias Görlach und Dr. Matthias Stoldt (Fritz-Lipmann-Institut, Jena) bedanke ich mich für die hilfreichen Gespräche zur NMR-Strukturanalyse. Für ihren experimentellen Beitrag durch Zuckertransport-Messungen möchte ich mich bei PD Dr. Fritz Titgemeyer und Tina Hertrich bedanken.

Ein ganz besonderes Dankeschön geht an Christina Herzberg, die sich in besonderer Weise engagierte und damit wesentlich zu den Ergebnissen in dieser Arbeit beigetragen hat. Sie gehört sicher in die Top-Ten der TAs dieser Welt. Vielen Dank auch für die mentale Unterstützung und all die praktischen Tipps und Ratschläge fürs Leben.

Mein Dank geht auch an Dr. Thorsten Mascher und seine Arbeitsgruppe, welche man immer wieder für ein Spielchen am Kicker begeistern kann. Es war interessant zu beobachten, welch große Fortschritte man dabei in kurzer Zeit erzielen kann. Desweiteren möchte ich mich bei Thorsten für die Einführung in die Tiefen der LFH-PCR bedanken. Bei Dr. Boris Görke bedanke ich mich für die interessanten Diskussionen zu allerlei molekularbiologischen

Methoden. Seine Art der praktischen Durchführung ist dabei äußerst konträr zu der, die ich aus dem ursprünglichen „Botchers Lab“ kennengelernt habe.
Prof. Dr. Wolfgang Hillen, an dessen Lehrstuhl an der Universität Erlangen einige Experimente für diese Arbeit durchgeführt wurden, danke ich für die Bereitstellung der dortigen Laborausstattung und Arbeitsumgebung. Den „Erlangern“ Steffi Bachem, Hans-Matti Blencke, Christian Detsch, Ines Langbein, Michael Müller, Irene Reif, Matthias Schmalisch, Shane Hanson und Ingrid Wacker danke ich für all die kleinen und großen Beiträge zu dieser Arbeit.
Vielen Dank auch an den Erlanger Kern, der mit mir gemeinsam die Reise ins norddeutsche Ausland gewagt hat: Dank an Hans-Matti, der mir in den ersten Wochen in Göttingen ein Dach über dem Kopf gegeben hat, und von dem ich auch gerne etwas von seiner nordischen Gelassenheit übernommen hätte. Dank auch an Fabian Commichau, der mir als erster gezeigt hat, dass es auch außerhalb von Bayern intelligentes Leben gibt. Last but not least danke ich Matthias Schmalisch für seine ausgedehnte Diskussionsbereitschaft zu nahezu jedem Thema unter der Sonne und für seine Überredungskünste, ohne die ich einige Dinge sicher noch nicht ausprobiert hätte (Rollersoccer???). Auch an unseren USA-Aufenthalt werde ich noch lange zurückdenken.
Ein grosses Dankeschön geht aber auch an die "alteingesessenen Göttinger" Birgit Veith und Mark Hoffmeister, die mir zu Beginn in ihrem Labor Asyl gewährt und mich schrittweise in die neuen Gegebenheiten eingeführt haben. Bedanken möchte ich mich natürlich auch bei meinen Praktikanten Timo Hupfeld, Dirk Jessen, Desiree Krausse und Hanna Vörsmann, die ich betreuen durfte, und die teilweise erheblichen Beitrag zu dieser Arbeit geleistet haben. Ich möchte auch meinem ersten und einzigen Diplomanden Sebastian Hübner dafür danken, dass er durch seine selbständige Arbeitsweise wenig Betreuungszeit in Anspruch genommen hat, wovon eigene Projekte profitieren konnten.
Mein Dank gilt selbstverständlich in gleicher Weise auch allen nicht namentlich genannten Kollegen für das gute und lockere Laborklima.
Zu guter Letzt möchte ich mich noch bei meiner Familie für ihre beständige Unterstützung bedanken. Vielen Dank auch an Kathrin, die mich aufopferungsvoll auch durch schwere Zeiten begleitet und mir die nötige mentale Kraft verliehen hat.

Table of contents

List of abbreviations

A	adenine
asRNA	antisense RNA
ATP	adenosine triphosphate
BSA	bovine serum albumin
bp	base pair
C	carbon
C	cytosine
cDNA	copy DNA
CCR	carbon catabolite repression
CCR	combined chain reaction
cm	centimeter
Cm	chloramphenicol
CTT	cytidine triphosphate
Cy3/Cy5	fluorescent dyes of the cyanine dye family
DHAP	dihydroxyacetone phosphate
DIG	digoxigenin
DNA	desoxyribonucleic acid
DNase	desoxyribonuclease I
dsRNA	double stranded RNA
EI	enzyme I
EII	enzyme II
EDTA	ethylenediaminetetraacetic acid
Em	erythromycin
FAD	flavin adenine dinucleotide
FBP	fructose-1,6-bisphosphate
FMN	flavin mononucleotide
Fig.	figure
G	guanine
Glc	glucose
GlcN6P	glucosamine-6-phosphate
GMP	guanosine monophosphate
GTP	guanosine triphosphate
H	hydrogen
h	hour
IMP	inosine monophosphate
IPTG	isopropyl-β-D-1-thiogalactopyranoside
kb	kilo base
kDa	kilodalton
Km	kanamycin
LB	Luria Bertani
LFH	long flanking homology
Lin	lincomycin
µl	microliters
µm	micrometers
M	molar
mCi	milli Curie
MCS	multiple cloning site
mg	milligram

min	minutes
miRNA	micro RNA
mj	milli Joule
ml	milliliter
mM	millimolar
MMR	multiple mutation reaction
mRNA	messenger RNA
mU	milliunit
m/z	ratio of mass to charge
NAD^+/NADH	β-nicotinamide adenine dinucleotide
NADPH	β-nicotinamide adenine dinucleotide phosphate
ncRNA	non-coding RNA
ng	nanogram
nm	nanometer
NTA	nitrilotriacetic acid
OD	optical density
PAA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEP	phosphoenolpyruvate
pg	picograms
pH	power of hydrogen
pl	picoliter
pM	pico molar
pmol	picomole
PPP	pentose phosphate pathway
PRD-I	PTS regulation domain I
PRD-II	PTS regulation domain II
PTS	phosphoenolpyruvate:sugar phosphotransferase system
pyr	pyrimidine
PVDF	polyvinylidene difluoride membrane
RAT	RNA antiterminator
RBD	RNA-binding domain
RNA	ribonucleic acid
RNAi	RNA interference
ROSE	repression of heat-shock gene expression
rRNA	ribosomal RNA
s	second
Sal	salicin
SAM	S-adenosylmethionine
SD	Shine Dalgarno
SDS	sodium dodecyl sulfate
SELEX	systematic evolution of ligands by exponential enrichment
siRNA	small interfering RNA
Spc	spectinomycin
sRNA	small RNA
Suc	sucrose
T	thymine
TAE	tris acetate EDTA
TCA	tricarboxylic acid
TE	Tris-EDTA

Tm	melting temperature
TPP	thiamine pyrophosphate
TRAP	*trp* RNA-binding attenuation protein
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
trp	tryptophan
TTP	thymidine triphosphate
U	uracile
U	unit
UMP	uridyl monophosphate
UTP	uridyl triphosphate
UTR	untranslated region
UV	ultraviolet
X-Gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside
WT	wild type
wt/vol	weight per volume

List of publications

Schilling, O., I. Langbein, M. Müller, M. H. Schmalisch and J. Stülke. 2004. A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res* **32:**2853-2864.

Hames, C., S. Halbedel, O. Schilling and J. Stülke. 2005. Multiple-mutation reaction: a method for simultaneous introduction of multiple mutations into the *glpK* gene of *Mycoplasma pneumoniae*. *Appl Environ Microbiol* **71:**4097-4100.

Schilling, O., C. Herzberg, T. Hertrich, H. Vorsmann, D. Jessen, S. Hübner, F. Titgemeyer and J. Stülke. 2006. Keeping signals straight in transcription regulation: specificity determinants for the interaction of a family of conserved bacterial RNA-protein couples. *Nucleic Acids Res* **34:**6102-6115.

Schilling, O., O. Frick, C. Herzberg, A. Ehrenreich, E. Heinzle, C. Wittmann and J. Stülke. 2007. Transcriptional and metabolic responses of *Bacillus subtilis* to the availability of organic acids: Transcription regulation is important but not sufficient to account for metabolic adaptation. *Appl Environ Microbiol* **73:**499-507.

Vogl, C., S. Grill, O. Schilling, J. Stülke, M. Mack and J. Stolz. 2007. Characterization of riboflavin (vitamin B_2) transport proteins from *Bacillus subtilis* and *Corynebacterium glutamicum*. Submitted to *J. Bacteriol*.

Summary

The soil bacterium *Bacillus subtilis* is exposed to frequently changing environmental conditions. Moreover, in its natural habitat, *B. subtilis* is surrounded by numerous microorganisms that are competing for the available resources. Therefore, the survival strategy of *B. subtilis* is optimized towards a rapid utilization of the available nutrients. The uptake of glucose, which is the preferred carbon source of *B. subtilis*, is mediated by the phosphoenolpyruvate:phosphotransferase system (PTS) encoded by the *ptsGHI* operon. The expression of this operon is controlled by transcriptional antitermination. The mRNA of the *ptsGHI* operon can adopt one of two alternative secondary structures in its 5'-untranslated region. The formation of a thermodynamic more stable terminator causes premature transcription termination. However, when the antiterminator protein GlcT binds to this region, it stabilizes an antiterminator structure called RAT and thus prevents terminator formation. The activity of GlcT is controlled by the availability of glucose. Only in the presence of glucose, GlcT can bind and stabilize the RAT structure. There are three more PTS-dependent antitermination systems in *B. subtilis*. These systems are controlled by the availability of sucrose or β-glucosides. Both the antiterminator proteins as well as the respective RATs are all very similar in structure and sequence. Nevertheless, the regulation by the antiterminator proteins is highly specific.

The aim of this work was the detection of all specificity determinants of the related antitermination systems. Site directed mutagenesis of two different RATs should help to identify nucleotides that are essential for the specific interaction with the respective antiterminator proteins. Furthermore, it was tested if the specificity of certain RATs can be changed towards other antiterminator proteins. For the simultaneous introduction of multiple point mutations, a method called MMR (multiple mutation reaction) was established. All nucleotides essential for specificity are located in the lower loop regions of the related RAT structures. Site directed mutagenesis of this region could also change specificity of certain RATs towards other antiterminator proteins. The interaction of the antiterminator proteins with their respective permeases is also specific and contributes to the overall specificity of the systems. Furthermore, it was found that carbon catabolite repression is also essential for the maintenance of specificity and for preventing cross-talk among the different systems.

For the utilization of its preferred carbon source glucose, *B. subtilis* features a complete set of enzymes for glycolysis, the pentose phosphate shunt, the tricarboxylic acid cycle and the respiratory chain. When available, glucose causes the repression of genes

needed for the utilization of alternative carbon sources. Only when glucose is completely exhausted, the genes necessary for the utilization of these carbon sources are expressed. This effect is known as carbon catabolite repression and has already been intensively studied.

The second aim of this work was to determine the effect of the organic acids glutamate and succinate on the central metabolism of *B. subtilis* when given in addition to glucose. This was achieved by a combined approach using transcriptomic data and metabolic flux analysis. The results of both studies were in good agreement for most of the studied genes. The addition of glutamate and succinate had no major effect on the genes of glycolysis and the pentose phosphate pathway. However, the flux of acetyl-CoA into the tricarboxylic acid cycle was severely reduced. On the other hand, the overflow metabolic pathways of lactate and acetate synthesis were significantly induced. For some of the genes, the change in transcription had not the expected consequence on the metabolic fluxes. This is in particular true for the genes involved in acetoin biosynthesis. Although the relevant genes were highly induced, no formation of acetoin was observed. These differences in the two data sets could indicate that the relevant enzymes are regulated on the level of protein activity.

Zusammenfassung

Als Bodenbakterium ist *Bacillus subtilis* in seinem natürlichen Lebensraum häufig wechselnden Umweltbedingungen ausgesetzt. Zudem ist *B. subtilis* in seinem Habitat von zahlreichen weiteren Mikroorganismen umgeben, welche mit ihm um die verfügbaren Ressourcen konkurrieren. Zur Überlebensstrategie von *B. subtilis* gehört es daher, die vorhandenen Nährstoffe möglichst schnell aufzunehmen und zu verwerten. Hierbei ist Glukose seine bevorzugte Kohlenstoffquelle. Die Aufnahme der Glukose erfolgt über das Phosphoenolpyruvat:Zucker Phosphotransferase System (PTS), welches durch das *ptsGHI*-Operon kodiert wird. Die Expression dieses Operons sowie die Expression einiger weiterer zuckerspezifischer Permeasen werden durch transkriptionelle Antitermination kontrolliert. Hierbei kann die mRNA dieser Gene im 5'-Bereich eine von zwei alternativen Sekundärstrukturen annehmen. Bei Ausbildung des thermodynamisch stabileren Terminators kommt es zum vorzeitigen Abbruch der Transkription. Bindet jedoch das Antiterminatorprotein GlcT an diesen Bereich der mRNA, so wird die Antiterminator-Struktur RAT stabilisiert und die Ausbildung des Terminators verhindert. Die Aktivität von GlcT wird durch die Verfügbarkeit von Glukose reguliert. Nur wenn Glukose vorhanden ist, kann GlcT an das RAT binden und es stabilisieren. In *B. subtilis* gibt es drei weitere PTS-abhängige Antiterminationssysteme, welche durch das Vorhandensein von Saccharose oder β-Glucosiden reguliert werden. Sowohl die Antiterminatorproteine als auch die zugehörigen RATs dieser Systeme weisen eine starke Homologie zueinander auf. Dennoch erkennen die Antiterminatorproteine spezifisch ihre jeweilige RAT-Struktur.

In dieser Arbeit sollten die Spezifitätsdeterminanten aller verwandten Antiterminationssysteme gefunden werden. Durch ortsgerichtete Mutagenesen zweier unterschiedlicher RATs sollten die Nukleotide identifiziert werden, welche für die spezifische Interaktion mit dem zugehörigen Antiterminatorprotein essenziell sind. Außerdem sollte getestet werden, ob es möglich ist, die Spezifität zugunsten eines anderen Antiterminatorproteins zu verändern. Um mehrere ortsgerichtete Mutagenesen parallel durchführen zu können, wurde die MMR (Multiple Mutation Reaction) als neue Methode etabliert. Es konnte gezeigt werden, dass sich die für die Spezifität wichtigen Nukleotide ausschließlich im unteren Schleifen-Bereich der verwandten RAT-Strukturen befinden. Durch gezielte Mutationen in diesem Bereich war es zudem möglich, die Spezifität der RATs zugunsten anderer Antiterminatorproteine zu verändern. Die Interaktion der Antiterminatorproteine mit ihrer zugehörigen Permease, welche sich für die direkte

Regulation des jeweiligen Antiterminatorproteins verantwortlich zeigt, ist spezifisch und trägt zur Spezifität des Gesamtsystems bei. Darüber hinaus konnte gezeigt werden, dass auch die Katabolitenrepression essenziell für die Aufrechterhaltung der Spezifität ist und Überschneidungen unter den verwandten Systemen verhindert.

Zur Verwertung seiner bevorzugten C-Quelle Glukose steht *B. subtilis* der komplette Satz an Enzymen für Glykolyse, Pentose-Phosphat-Weg, Zitronensäure-Zyklus und Atmungskette zur Verfügung. Gleichzeitig reprimiert die Anwesenheit von Glukose die Expression der Gene, welche für die Verwertung von anderen C-Quellen benötigt werden. Erst wenn Glukose verbraucht ist, werden die Gene für die Verwertung anderer C-Quellen induziert. Dieser als Katabolitenrepression bezeichnete Effekt war bereits Mittelpunkt zahlreicher Untersuchungen.

Eine weitere Zielsetzung dieser Arbeit war es daher, bei gleichzeitiger Anwesenheit von Glukose, den Einfluss der organischen Säuren Glutamat und Succinat auf den zentralen Stoffwechsel von *B. subtilis* zu untersuchen. Hierzu wurden Transkriptomdaten, welche mit DNA-Microarrays gewonnen wurden, sowie Daten aus metabolischen Flussanalysen kombiniert. Die Ergebnisse beider Methoden stimmen für den Großteil der untersuchten Stoffwechselwege sehr gut überein. Die Zugabe von Glutamat und Succinat hat keinen Effekt auf die Gene der Glykolyse und des Pentose-Phosphat-Wegs. Der Zufluss von Acetyl-CoA in den Zitronensäure-Zyklus ist jedoch stark reduziert, wohingegen die Aktivität der Enzyme für die Überfluss-Stoffwechselwege der Lactat- und Acetat-Synthese stark erhöht ist. Für einige wenige Gene wirkt sich die Änderung in der Genexpression jedoch nicht wie erwartet auf den metabolischen Fluss aus. Dies gilt besonders für die Gene der Acetoin-Biosynthese. Trotz einer starken Induktion der beteiligten Gene konnte keine Synthese von Acetoin beobachtet werden. Derartige Abweichungen können Hinweise darauf geben, dass die Aktivität der beteiligten Enzyme auf Proteinebene reguliert wird.

1 Introduction

1.1 Sugar metabolism in *Bacillus subtilis*

1.1.1 An overview

All living organisms face the same challenge: they need energy to maintain their vital functions such as motility, transport of molecules, maintenance of proton motive force and many anabolic reactions. The central metabolism does not only provide the supply of energy rich molecules, but also transforms and modifies them, interconnects different chemical pathways, and feeds the supply of all building blocks of the living cell. Nitrogen metabolism, for example, is of great importance for the supply of amino acids and nitrogen containing molecules like purines and pyrimidines for DNA and RNA synthesis. Thus, the central metabolism plays a key role in all processes of life.

Many metabolic pathways, some organism-specific, catalyze the conversion of various metabolites to obtain a continuous synthesis of ATP, NADH and NADPH. These pathways, among them glycolysis, the pentose phosphate pathway, the citric acid cycle, the Calvin cycle, and photosynthesis, can be found in many organisms of prokaryotic and eukaryotic origin.

Among bacteria, the Gram-positive soil bacterium *Bacillus subtilis* is one of the best studied organisms. In its natural habitat, the upper part of the soil and the rhizosphere, the bacterium has access to a variety of plant remains, containing diverse energy rich molecules. *B. subtilis* features a complete genomic configuration to utilize carbohydrates as its preferred carbon and energy source. Excreted enzymes, such as amylases and glucanases, help to make long chain sugar molecules like starch and cellulose available for the organism. Sugar uptake is mediated by facilitators, primary and secondary active transporters, and the phosphoenolpyruvate:sugar phosphotransferase system (PTS) (Postma *et al.*, 1993; Saier *et al.*, 2002; Simoni *et al.*, 1967).

Among these, the PTS is the most important sugar uptake system. It consists of the two general proteins enzyme I (EI) and HPr, and several sugar specific enzymes (see chapter 1.1.2 and 1.3). These membrane-bound sugar specific permeases (enzyme II or EII) not only mediate sugar uptake, but also modify the sugar by phosphorylation (Reizer *et al.*, 1988). Phosphorylated sugars are negatively charged and cannot leave the cell anymore. Thus, they are captured inside the cell. The PTS represents an energy efficient way for sugar uptake and

for making sugars available for downstream metabolic pathways like glycolysis and the pentose phosphate pathway.

Among the sugars, glucose is the preferred carbon source of *B. subtilis*. Its uptake is mainly mediated by the glucose specific EII^{Glc} of the PTS. The resulting glucose-6-phosphate can directly feed into glycolysis. *B. subtilis* features all glycolytic enzymes needed to convert glucose to pyruvate (Ludwig *et al.*, 2001). Glucose, that was not transported via the PTS or that is derived from hydrolysis of other oligo- or disaccharides, e.g. sucrose, has to be phosphorylated by the glucose kinase (Skarlatos and Dahl, 1998). The next step is the conversion of glucose-6-phosphate to fructose-1,6-bisphosphate which is then split into two C-3 molecules, dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (see Fig. 1.1). Accumulated DHAP is also converted to glyceraldehyde-3-P. The enzymes needed for the conversion of glyceraldehyde-3-phosphate to phosphoenolpyruvate are organized in an operon (*gapA* operon) on the *B. subtilis* chromosome (Fillinger *et al.*, 2000; Leyva-Vazquez and Setlow, 1994) (see Chapter 1.1.2). The last step is the conversion of PEP to pyruvate, and the net outcome of the glycolysis are 2 molecules of ATP and 2 molecules NADH.

Glucose-6-phosphate is important for the synthesis of many polysaccharides e.g. peptidoglycan which is needed for cell wall synthesis. The glycolytic intermediates 3-phosphoglycerate, PEP, and pyruvate are significant as sources for the synthesis of all aromatic amino acids, and for the amino acids of the alanine and serine families. Many enzymes of glycolysis work also reversibly and are therefore part of the gluconeogenesis. The only glycolytic enzymes that do not work reversibly are *ptsG*, *gapA, pfk,* and *pykA* (see Fig. 1.1).

In the absence of glucose, gluconeogenesis forms new glucose from metabolic intermediates such as oxaloacetic acid or dihydroxyacetone phosphate. Gluconeogenesis starts with the conversion of oxaloacetate to PEP, carried out by the phosphoenolpyruvate carboxykinase (*pckA*). This gluconeogenetic reaction is essential when the organism grows on TCA cycle intermediates (Diesterhaft and Freese, 1973; Yoshida *et al.*, 2001). Although the *B. subtilis* genome codes for a phosphoenolpyruvate synthase (*pps*), the bacterium, in contrast to *Escherichia coli*, does not seem to be able to directly convert pyruvate to PEP (Sauer and Eikmanns, 2005). To synthesize glucose-6-P from PEP two further enzymes are necessary, since *gapA* and *pfk* do not work reversibly (see Fig. 1.1).

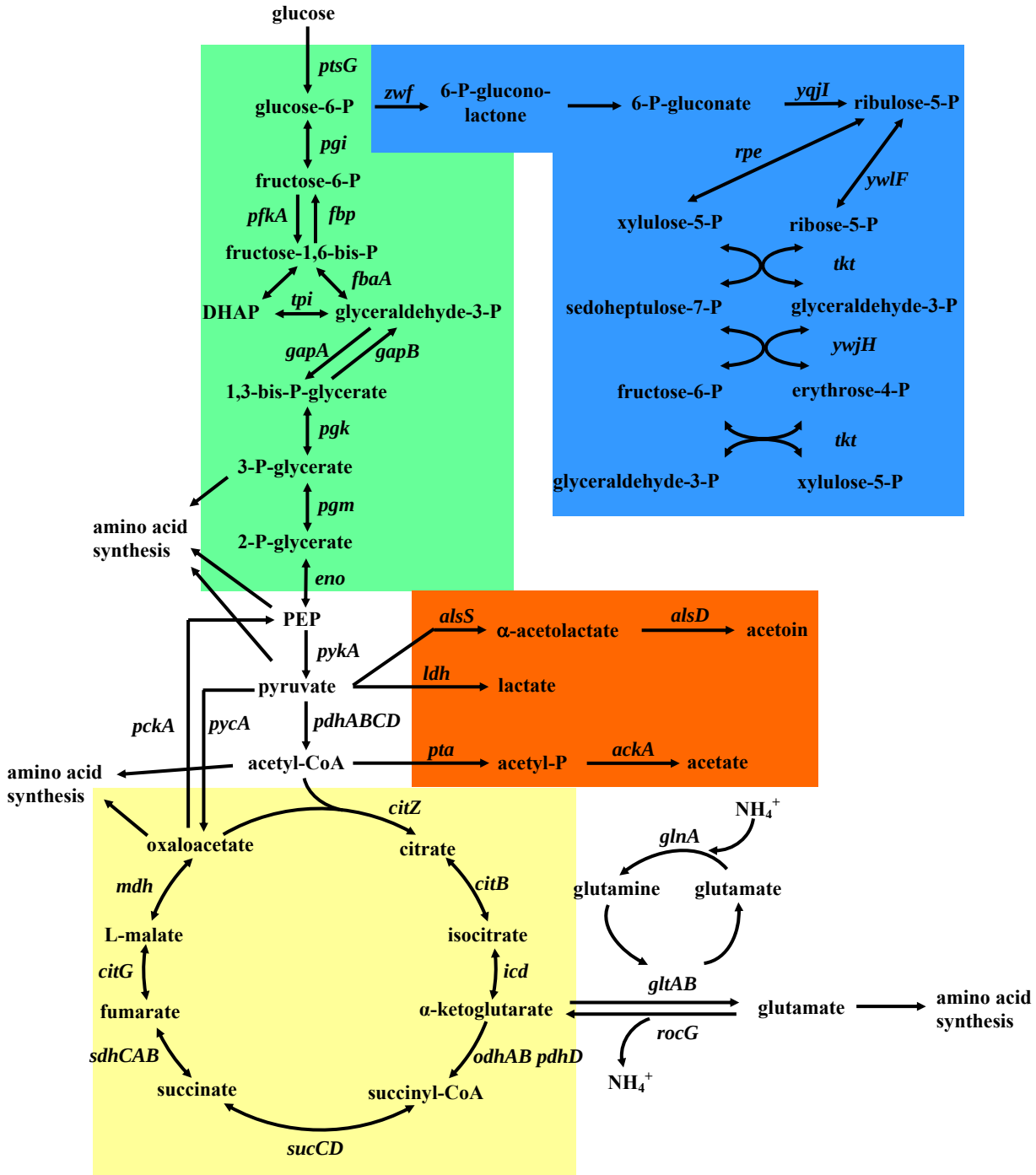

Figure 1.1. Overview on the genes of central metabolic pathways in *B. subtilis*.
Glycolysis is indicated in green, the pentose phosphate shunt is indicated in blue, overflow metabolic pathways are indicated in red, and the citric acid cycle is indicated in yellow. Modified from Blencke *et al.*, 2003.

The synthesis of glyceraldehyde-3-P from 1,3-bisphosphoglycerate is catalyzed by the glyceraldehyde-3-phosphate dehydrogenase (*gapB*), and fructose-6-phosphate is formed from fructose-1,6-bisphosphate by the activity of the fructose-1,6-bisphosphatase (*fbp*) (Fillinger *et al.*, 2000; Fujita *et al.*, 1998; Servant *et al.*, 2005).

The pentose phosphate pathway (PPP) provides precursors for DNA, RNA, and amino acid syntheses (Tozzi *et al.*, 2006). Furthermore, it has an important role in providing NADPH for anabolic reactions. The pentose phosphate pathway can be subdivided into an oxidative and a non-oxidative part. In the oxidative part, glucose-6-phosphate is converted to ribulose-5-phosphate and 2 molecules of NADPH are synthesized. In the non-oxidative stage, it can interconvert different types of pentoses and additionally converts pentoses to hexoses (fructose-6-phosphate) and trioses (glyceraldehyde-3-phosphate) that can feed into glycolysis (see Fig. 1.1). Depending on what metabolic intermediates are needed, the pentose phosphate pathway can flow into the one or the other direction. In contrast to its important role for *B. subtilis*, the PPP was not too much in the focus of scientific research. Therefore, it was only recently discovered that the phosphogluconate dehydrogenase is coded by *yqjI*, and not *gntZ*, as assumed previously (Zamboni *et al.*, 2004).

Pyruvate is the final product of glycolysis, but it is also generated when PEP is used as a phosphate donor to power the sugar transport of the PTS. Pyruvate can be used to produce lactate via the lactate dehydrogenase (*ldh*) or acetoin via the acetolactate synthase and acetolactate decarboxylase (*alsSD* operon) (Renna *et al.*, 1993). These reactions are part of the overflow metabolism, which helps the bacterium to regenerate NAD^+ from NADH when no other electron acceptors like oxygen or nitrate are available (Cruz Ramos *et al.*, 2000). *B. subtilis* uses overflow metabolic pathways also under aerobic conditions, when excess of carbon and nitrogen sources are present (Shivers *et al.*, 2006).

Pyruvate can also be converted to acetyl coenzyme A (acetyl-CoA) by the pyruvate dehydrogenase complex (*pdhABCD*) (Hemilä *et al.*, 1991). Acetyl-CoA can feed into the TCA cycle, is used for synthesis of the amino acids lysine and methionine, or is converted to acetyl phosphate and further to acetate via phosphotransacetylase (*pta*) and acetate kinase (*ackA,* see Fig. 1.1) (Grundy *et al.*, 1993; Presecan-Siedel *et al.*, 1999; Shin *et al.*, 1999). The latter reactions are also part of overflow metabolism.

B. subtilis features the complete enzymatic configuration for a cyclic TCA pathway (Huynen *et al.*, 1999). Under aerobic conditions, the TCA cycle catalyzes the complete oxidation of pyruvate to CO_2, and hydride ions are transferred to electron carrier molecules. One molecule of pyruvate leads to the formation of 4 NADH (or 3 NADH and 1 NADPH),

and 1 FADH. Moreover, 1 GTP molecule is formed per cycle. The energy bound in these hydrogen atom carrier molecules (NAD^+, FAD^+) is used to generate ATP via the respiratory chain (Hederstedt, 1993; Sonenshein, 2002). The provided NADPH can be used in anabolic reactions.

Another important task of the TCA cycle is the supply of precursors for amino acid biosynthesis such as α-ketoglutarate, and oxaloacetate (see Fig. 1.1). α-Ketoglutarate is the precursor of glutamate, which itself is the precursor of many other amino acids (Schreier, 1993). The synthesis of glutamate is performed by the transfer of an amino group to α-ketoglutarate by the glutamate synthase (*gltAB* operon). This reaction represents the link between C- and N-metabolism in *B. subtilis* (Commichau *et al.*, 2006). Oxaloacetate is used as a precursor for the synthesis of the amino acids of the aspartate family (Akashi and Gojobori, 2002). To ensure a continuous function of the TCA cycle, the flow of TCA intermediates has to be substituted. This can be achieved either by the reversion of the above-mentioned reactions, or by directly carboxylating pyruvate to oxaloacetate, a reaction that requires energy in form of ATP. This so called anaplerotic reaction is carried out by the pyruvate carboxylase (*pycA*) in *B. subtilis* (Bernhardt *et al.*, 2003; Sonenshein, 2002). A third way to refill TCA cycle intermediates is the generation of succinyl-CoA by the β-oxidation of odd-chain fatty acids. Due to the lack of a glyoxylate cycle, *B. subtilis* is not able to grow on acetate as its only carbon source (Huynen *et al.*, 1999).

1.1.2 Regulation of sugar metabolism in *B. subtilis*: cause and effect

B. subtilis can utilize a huge variety of different sugars as carbon and energy source. The transcription of the respective gene needed to metabolize a certain sugar is only initiated if the particular carbohydrate is available (Stülke and Hillen, 2000). However, not all carbon sources deliver the same amount of energy. Therefore, if a multitude of sugars is available, only the optimal carbon source is metabolized. This mode of gene regulation is described as carbon catabolite repression (CCR) (Chambliss, 1993).

The regulation of sugar metabolism in *B. subtilis* already starts at the uptake process, and here the PTS plays a crucial role (see chapter 1.3). Besides its important role in sugar uptake, the PTS also has a major impact on the regulation of downstream metabolic processes. Among the general PTS components, the HPr protein plays the key role. It can be phosphorylated at positions His-15 and Ser-46 (Galinier *et al.*, 1998; Stülke and Hillen, 2000). HPr (His-P) serves as a phosphate donor for the phosphorylation of proteins involved in the regulation of sugar permeases and inducer exclusion (Stülke and Schmalisch, 2004). In

contrast to HPr (His-P), HPr (Ser-P) does not serve as a phosphate donor, but acts as a co-repressor of the global transcription regulatory protein CcpA (see below). The Ser-46 phosphorylation is mediated by the enzyme HPr kinase/phosphorylase (*hprK*) (Deutscher *et al.*, 1986; Deutscher and Saier, 1983). HPrK/P senses the concentration of intracellular ATP and glycolytic intermediates such as fructose-1,6-bisphosphate (FBP). FBP stimulates the kinase activity of HPrK/P, whereas inorganic phosphate stimulates the phosphorylase activity (Hanson *et al.*, 2002; Jault *et al.*, 2000; Ramström *et al.*, 2003). Thus, HPrK/P controls the HPr phosphorylation state subject to the energy charge of the cell (Galinier *et al.*, 1998; Ludwig *et al.*, 2002b; Reizer *et al.*, 1998). There is also indication that the activity of HPrK/P is influenced by the intracellular pH value of the cell (Ramström *et al.*, 2003).

CcpA (carbon catabolite control protein A) is the main regulatory protein of carbon metabolism in *B. subtilis* and other Gram-positive bacteria. CcpA belongs to the LacI/GalR family of transcription regulatory proteins (Henkin, 1996; Weickert and Adhya, 1992). It consists of an N-terminal DNA binding domain and two C-terminal subdomains. The CcpA protein is a dimer of two identical subunits (Jones *et al.*, 1997; Schumacher *et al.*, 1994). It is assumed that over 300 genes are subjected to CcpA dependent regulation, which is almost 10% of the total number of *B. subtilis* genes (Blencke *et al.*, 2003; Moreno *et al.*, 2001; Yoshida *et al.*, 2001).

There are two modes of CcpA dependent regulation. CcpA can influence the transcription of catabolic genes and operons when it binds to a palindromic sequence called catabolite responsive element (*cre* sequence). This was first discovered in studies on the regulation of the α–amylase gene in *B. subtilis* and a consensus sequence for *cre* was determined (Henkin *et al.*, 1991; Hueck *et al.*, 1994; Kim *et al.*, 1995; Miwa *et al.*, 2000). The intrinsic interaction of *cre* and CcpA is stimulated when HPr (Ser-P) binds to CcpA. The HPr analog Crh can also bind to CcpA. Hence, HPr and Crh act both as cofactors in CcpA dependent regulation. Unlike HPr, Crh can not be phosphorylated at position His-15 and is therefore not part of the PTS (Galinier *et al.*, 1999; Galinier *et al.*, 1997). Regulation that is mediated by the binding of CcpA to a *cre* sequence is defined as class I regulation. The complex of CcpA and HPr/Crh (Ser-P) can act both as a repressor and as an activator of gene expression (see below). In a CcpA mutant strain, differential gene expression occurs also for genes that do not have a *cre* sequence. Since a functional CcpA protein is necessary for efficient sugar transport, the absence of CcpA affects the function of the PTS and hence the glucose uptake rate. Genes controlled by regulators that sense glucose or its metabolic intermediates are therefore indirectly influenced by CcpA activity. This indirect mode of

CcpA dependent regulation caused by limited sugar uptake is defined as class II regulation. Additionally, CcpA seems to be involved in the regulation of genes that do not belong to class I or class II. For most of these genes, the mode of operation is not clear.

When phosphorylated at position Ser-46, HPr can not be phosphorylated at His-15, and therefore can not function as a phosphate donor within the PTS (Deutscher and Saier, 1983). Thus, HPrK/P exerts also an indirect influence on the regulation of PTS mediated sugar uptake.

Many of the genes involved in glycolysis are constitutively expressed. Except for phosphoglucoisomerase and fructose-1,6-bisphosphate aldolase (*pgi, fbaA*) all glycolytic enzymes are organized in operons (Kunst *et al.*, 1997). The *gapA* operon includes the glycolytic genes *gapA*, *pgk*, *tpi*, *pgm*, and *eno*, coding for glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, triose phosphate isomerase, phosphoglycerate mutase, and enolase (see Fig. 1.2) (Kunst *et al.*, 1997; Leyva-Vazquez and Setlow, 1994).

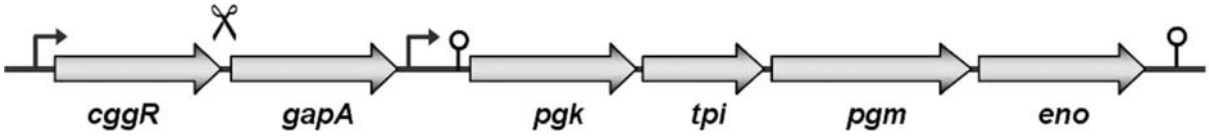

Figure 1.2. Overview on the *gapA* operon of *B. subtilis*. The promoter upstream of *cggR* is feedback regulated by CggR. The constitutive promoter upstream of *pgk* ensures the availability of *pgk*, *tpi*, *pgm*, and *eno* under both glycolytic and glyconeogenetic conditions. The processing site between *cggR* and *gapA* modulates the stability of the polycistronic mRNA. The processed *gapA* transcript or the processed transcript comprising *gapA*, *pgk*, *tpi*, *pgm* are more stable compared to the unprocessed mRNAs.

Additionally, the gene coding for the transcription regulatory protein CggR (central glycolytic gene regulator) is also part of the operon and is located upstream of *gapA* (see Fig. 1.2) (Fillinger *et al.*, 2000). This regulatory protein represses the *gapA* operon when no substrate is present to feed glycolysis. The activity of CggR is regulated by the glycolytic intermediate FBP (Doan and Aymerich, 2003; Zorrilla *et al.*, 2007). Additionally, the expression of the *gapA* operon is also modulated by the availability of amino acids (Ludwig *et al.*, 2001). The *gapA* operon also seems to be indirectly regulated by CcpA in a class II dependent manner (see above). In a CcpA mutant strain, glucose dependent regulation of *gapA* is abolished, but no functional *cre* sequence is present in the promoter region of the operon (Ludwig *et al.*, 2002b).

The glycolytic enzymes located downstream of *gapA* are constitutively expressed (Ludwig *et al.*, 2001). This is achieved by a constitutive promoter located between *gapA* and *pgk* (see Fig. 1.2). As these enzymes work reversibly (see Fig. 1.1), they are also part of gluconeogenesis. GapA, however, catalyzes exclusively the conversion of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate. Thus, it is only part of glycolysis. Besides the direct repression by CggR and the indirect regulation by CcpA, the *gapA* mRNA stability is modulated by an endonuclease (see Fig. 1.2) (Ludwig *et al.*, 2001; Meinken *et al.*, 2003).

As mentioned above (see chapter 1.1.1), GapB catalyzes the reverse reaction of GapA, which is the conversion of 1,3-bisphosphoglycerate to glyceraldehyde-3-phosphate. This reaction is part of gluconeogenesis. In contrast to *gapA*, *gapB* is repressed in the presence of glucose (Fillinger *et al.*, 2000; Yoshida *et al.*, 2001). Thus, *gapA* and *gapB* are also regulated in a reverse manner. More details about the regulation of the gluconeogenetic genes *gapB* and *pckA* (see chapter 1.1.1 and Fig. 1.1) have been uncovered recently. In addition to the reversible glycolytic enzymes, both are needed for the efficient formation of glucose-6-phosphate from TCA cycle intermediates. The expression of both genes is repressed in the presence of glucose, but this regulation is independent or, in the case of *pckA*, only indirectly dependent on CcpA (Blencke *et al.*, 2003; Moreno *et al.*, 2001). Instead, CcpN (control catabolite protein of gluconeogenic genes) was identified to bind in a cooperative manner to the promoter regions of *pckA* and *gapB*, causing transcriptional repression in the presence of glucose (Licht and Brantl, 2006; Servant *et al.*, 2005). In addition, CcpN was shown to bind to the promoter region of *sr1*, coding for a small regulatory RNA (see also chapter 1.2.2) (Licht *et al.*, 2005).

The genes involved in overflow metabolism (*pta*, *ackA*, *alsSD*) and the *pdhABCD* operon (see chapter 1.1.1) are also under the control of CcpA. The glucose dependent induction of *ackA* and *pta* is directly related to the binding of CcpA to a *cre* sequence upstream of the respective promoter region (Grundy *et al.*, 1993; Presecan-Siedel *et al.*, 1999; Shin *et al.*, 1999; Turinsky *et al.*, 1998). For the activation of the *alsSD* operon, CcpA seems to play an indirect role, since no *cre* sequence could be found in the promoter region (Tobisch *et al.*, 1999b; Turinsky *et al.*, 2000). The direct regulator of the *alsSD* operon is the transcription activator AlsR, but the exact trigger for its activity is not yet known (Cruz Ramos *et al.*, 2000; Renna *et al.*, 1993). There is not much known about the regulation of the *pdhABDC* operon, but there is also indication for an indirect role of CcpA (Blencke *et al.*, 2003).

The TCA cycle is an important link between carbon- and nitrogen metabolism (see chapter 1.1.1 and Fig. 1.1). Its activity is therefore modulated by the availability of carbon and nitrogen sources in a synergistic manner. Only in the presence of both a good carbon and good nitrogen source the TCA cycle is repressed (Rosenkrantz *et al.*, 1985; Sonenshein, 2002). The main regulation is mediated by CcpC, a transcription regulatory protein of the LysR family (Jourlin-Castelli *et al.*, 2000). It controls the expression of the *citZ* operon which codes for the citrate synthase, isocitrate dehydrogenase, and malate dehydrogenase (*citZ, icd, mdh*). Additionally, *icd* and *mdh* contain also CcpC independent promoters (Jin *et al.*, 1996). In addition, the aconitase (*citB*) is under the control of CcpC (Jourlin-Castelli *et al.*, 2000). It was shown that CcpC represses the respective genes depending on the availability of glucose as well as glutamine, glutamate, and ammonium. *citZ* and *citB* are also subject to CcpA dependent carbon catabolite repression. However, only the promoter region of *citZ* contains a CcpA binding site (*cre* sequence) (Sonenshein, 2002). The influence of CcpA on the regulation of *citB* is an indirect effect resulting from the CcpA dependent regulation of *citZ* (Kim *et al.*, 2002). CitZ catalyzes the formation of citrate from oxaloacetate and acetyl-CoA. In the presence of glucose, *citZ* is under CcpA dependent carbon catabolite repression and therefore the intracellular citrate concentration is decreased. Since citrate is the inducer of CcpC repressed genes, the repression of *citZ* by CcpA causes the indirect repression of *citB* (Kim *et al.*, 2003a; Kim *et al.*, 2002). *CitB* is also repressed in the presence of good nitrogen sources. It was proposed that glutamine or α-ketoglutarate is the trigger of this repression, but newer studies indicate that arginine could be the nitrogen signal causing repression of *citB* (Blencke *et al.*, 2006; Fisher and Magasanik, 1984; Rosenkrantz *et al.*, 1985).

In contrast to the important role of the pentose phosphate pathway for *B. subtilis*, there is not much known about the regulation of the involved genes. A recent study revealed that the genes encoding the enzymes of this pathway might be constitutively expressed (Blencke *et al.*, 2003).

1.2 RNA and its role in gene regulation

The idea that RNAs may serve as regulatory elements was not born just recently. The first were Jacob and Monod who proposed a possible role for RNA in direct gene regulation (Jacob and Monod, 1961). However, when the first regulatory systems were elucidated and it became clear that gene expression was controlled on the level of transcription initiation by

proteins, they discarded this idea completely. Today, gene regulation by RNA is a field of growing interest and the modes by which RNA regulates gene expression are extremely diverse.

1.2.1 The structural features of single stranded RNA

Folding prediction for a protein from its amino acid sequence still remains an unsolved problem. This is due to the large number of 22 codogenic amino acids (when including selenocysteine and pyrolysine) and the fact that folding is not only dependent on sequence but also on the pH and the salt concentration. Despite the increasing power of computer hardware over the last decade, it is still not possible to predict the right folding structure of even small proteins (Zhang and Skolnick, 2004). Although RNA is made up of only four ribonucleotides, structure prediction is a likewise complicated task. Especially single stranded RNA can adopt secondary structures with a complexity similar to that of proteins. Its structural flexibility is also responsible for its functional diversity (Schroeder *et al.*, 2004). The high flexibility in structure derives from the fact that the four ribonucleotides not only interact by Watson-Crick base pairing to stabilize a secondary structure (Watson and Crick, 1953), but also by so called Hoogsteen, "C-H", and sugar edge interactions. In a Hoogsteen base pairing, the hydrogen in purines at position C-8 instead of position C-2 is involved in hydrogen bond formation. Similarly, the hydrogen in pyrimidines at position C-5 instead of N-3 is involved in "C-H" base pairing interaction. In sugar edge interactions, the hydroxyl group of the ribose interacts with the hydrogen atoms of the RNA bases (see Fig. 1.3). The bases can furthermore interact in either of two different orientations, *cis* or *trans*, relative to the hydrogen bonds which altogether results in more than 16 basic types of interaction (see Fig. 1.3) (Leontis *et al.*, 2002; Leontis and Westhof, 2001; Yang *et al.*, 2003). Single stranded RNA can also interact with other RNAs and thereby modulate translation efficiency or the half-life of these molecules.

This flexibility also led to the "RNA world" hypothesis where RNA was responsible for both the maintenance of genetic information and the catalysis of chemical reactions (Gilbert, 1986). In an ancient world where neither DNA nor proteins have existed, RNA could have catalyzed its own replication and this error-prone process might have favoured a quick evolution of life (Jeffares *et al.*, 1998; Johnston *et al.*, 2001; Long *et al.*, 2003). Today, four billion years after the evolution of life started, there are still some hints present in modern organisms indicating RNA as a multifunctional molecule apart from being a messenger for genetic information.

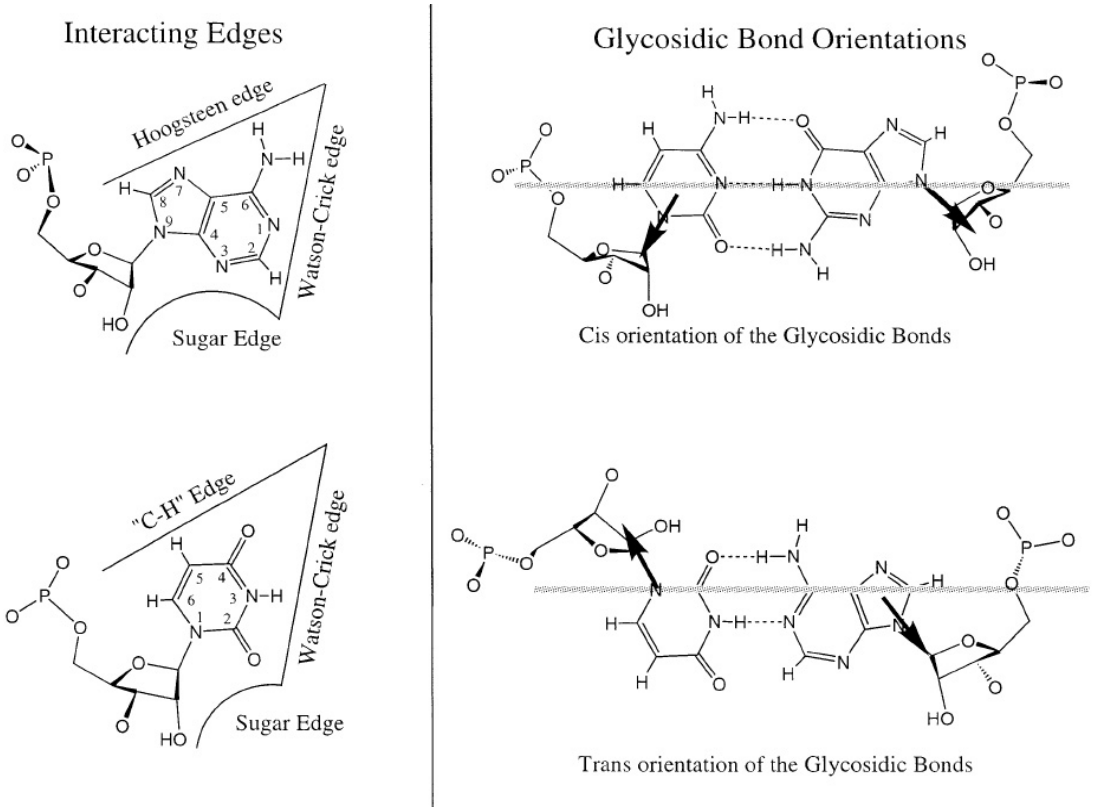

Figure 1.3. Left: Illustration of the possible interaction sites of the RNA bases (Watson-Crick, Hoogsteen, C-H, and sugar edge). Right: Illustration of possible orientations of the glycosidic bonds in a Watson-Crick base pair interaction. The upper picture shows an uracile-guanine interaction with glycosidic bonds in cis orientation. The lower picture shows an uracile-adenine interaction with trans orientation of the glycosidic bonds (Leontis *et al.*, 2002).

In some viruses, RNA is still the long term storage of genetic information (Domingo and Holland, 1997; Strauss and Strauss, 1988) and the process of DNA replication can not be initiated without an RNA primer (Eigen *et al.*, 1981). Additionally, many ribozymes, like the spliceosome and the ribosome, are evidence of the catalytic power of RNA. In these complexes the RNA exhibits the catalytic function and the protein part is mostly responsible for giving the proper shape to the RNA catalytic center (Noller, 1993; Wise, 1993). Finally, the fact that ribonucleotides serve as precursors for deoxynucleotide synthesis points to the important role that RNA carried out prior to DNA in the history of life (Joyce, 1989).

Another important characteristic of RNA is its much lower half-life as compared to DNA. This is due to the hydroxyl group at the 2'-C atom of the ribose. Thus, an RNA polymer is susceptible to spontaneous hydrolysis leading to chain cleavage (Winkler and Breaker, 2003). As a result, RNA is also unstable against acidic or alkaline hydrolysis (Larralde *et al.*, 1995). However, exactly this feature makes RNA so catalytically reactive and thus broadly applicable.

The catalytic power of RNA is even more impressive when considering that RNA is not only made up of just four different components, but that these are also very similar in size and chemistry. Even proteins, although having access to an arsenal of amino acids that vary significantly in size, charge, and polarity, can not fulfill their job without the use of lower molecular weight cofactors (Jadhav and Yarus, 2002; Reichow and Varani, 2006). Therefore, the affinity of some RNAs to lower molecular weight agents like CoA, cyanocobalamin, and biotin might be inherited from the RNA-world, where RNA needed the help of this molecules to exert its catalytic functions (Jadhav and Yarus, 2002).

As there are more complete genome sequences published every year, comparative sequence analysis of RNAs that might exhibit catalytic or regulatory functions seems to be feasible very easily. However, not only the actual sequence, but also the structural features of catalytic RNAs are important. For further research, the accurate structure prediction becomes essential for the RNA science community. There are elaborated algorithms for RNA secondary structure prediction. A very common tool for this purpose is “Mfold” (Mathews *et al.*, 1999; Zuker, 2003). The software includes Watson-Crick base pairing to calculate the secondary structure of an RNA sequence at a given temperature. A viable method to verify the correctness of *in silico* predicted structures is the introduction of mutations. If the structure prediction is accurate, both mutations that are tolerated and mutations that destroy the secondary structure and consequently the activity of the catalytic RNA can be predicted and tested (Barash, 2003). The programs for RNA structure prediction become also more accurate by incorporating more sophisticated algorithms and combining them with data obtained from high-resolution RNA structures (Gautheret and Gutell, 1997; Lillo *et al.*, 2002; Mathews *et al.*, 2004; Sykes and Levitt, 2005). Thus, it is very likely that with progressive and optimized *in silico* methods by far more catalytic and regulatory RNAs will be found in all sequenced organisms.

1.2.2 Non-coding RNAs

It was first discovered in bacteria that small non-coding RNA molecules (ncRNA) had a negative effect on the expression of other genes (Mizuno *et al.*, 1984; Tomizawa *et al.*, 1981). The replication control of a plasmid (ColE1) and the transcription of an *Escherichia coli* outer membrane protein (*ompF*) were found to be under the control of small RNA transcripts. Later on, small RNA molecules were used to influence gene regulation in eukaryotes. These artificially designed RNA molecules were complementary to dedicated mRNA target sequences, and were therefore called “antisense RNAs” (asRNA). Although the mechanism

was not fully understood, it became a useful tool for influencing gene expression in organisms that are not fully accessible to genetic manipulations (Aigner, 2006; Amarzguioui *et al.*, 2005).

In eukaryotes, many types of small RNAs have influence on transcription and stability of different kinds of RNAs. Micro RNAs (miRNAs) play a role in the silencing of coding genes, control chromosome organisation, inhibit viruses and transposable elements, and have influence on proliferation, apoptosis, and differentiation (Ambros, 2004; Aravin and Tuschl, 2005; Xu *et al.*, 2004; Ying *et al.*, 2006). Small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and small guide RNAs (gRNAs) are involved in mRNA and rRNA processing (Carthew, 2006; Mattick and Makunin, 2005; Stuart and Panigrahi, 2002; Valadkhan, 2005; Wagner and Flärdh, 2002).

Although bacterial ncRNAs, which are also called small RNAs (sRNAs) (Gottesman, 2004), play a similar role in gene regulation as eukaryotic miRNAs, they differ in many aspects significantly from each other. While their eukaryotic counterparts are processed from double stranded regions of stem loop structures, bacterial ncRNAs are single stranded transcripts. Processed eukaryotic miRNAs also feature a size of about 21 to 23 nucleotides (Sewer *et al.*, 2005). In contrast, bacterial ncRNAs are not processed and are generally 80-100 bps in size (Gottesman, 2004; Gottesman, 2005). They fold into stem-loop secondary structures which help to stabilize the molecules and prevent degradation (Massé *et al.*, 2003). Although protein cofactors play a crucial role for the activity of regulatory RNAs in both kinds of organisms, the mode of regulation differs. In bacteria, a protein factor does not seem to be necessary in all cases, but for some ncRNAs the protein Hfq is important for their regulatory activity (Valentin-Hansen *et al.*, 2004; Zhang *et al.*, 2003). Hfq, first discovered in *E. coli*, seems to be necessary for the structural reorganisation of the interacting RNA molecules, and thus mediates binding (Møller *et al.*, 2002). More than 20 ncRNAs are known to directly interact with Hfq (Zhang *et al.*, 2003). For the degradation process of the target mRNA, RNase E seems to play a crucial role (Moll *et al.*, 2003).

The regulatory RNAs can cause inhibition of translation and RNA degradation, but in contrast to miRNAs and siRNAs, also stimulation of translation and mRNA stabilisation was observed (Majdalani *et al.*, 1998; Massé *et al.*, 2003; Opdyke *et al.*, 2004; Repoila and Gottesman, 2003). Bacterial non-coding RNAs are involved in the regulation of virulence, stress response, quorum sensing, uptake systems, and plasmid replication (Brantl, 2002; Guillier *et al.*, 2006; Lenz *et al.*, 2004; Massé *et al.*, 2003; Massé *et al.*, 2005; Repoila and Gottesman, 2003; Romby *et al.*, 2006).

In contrast to *E. coli*, there is not much known about ncRNAs in *B. subtilis*. Recently, it was found that the small regulatory RNA *sr1* is involved in the fine tuning of arginine catabolism (see chapter 1.1.2) (Heidrich *et al.*, 2006). Another small regulatory RNA (*ratA*) was found to act as an antitoxin (Silvaggi *et al.*, 2005). Although YmaH was identified as an Hfq analog in *B. subtilis* it does not seem to be involved in *sr1* and *ratA* mediated regulation (Heidrich *et al.*, 2006; Silvaggi *et al.*, 2005). A recent study also revealed that ncRNAs take part in the complex regulatory process of sporulation in *B. subtilis* (Silvaggi *et al.*, 2006).

1.2.3 RNA switches

The ability of single stranded RNA to adopt complex three-dimensional structures is essential for gene regulation mediated by RNA switches (riboswitches). The underlying principle is the possibility of an RNA sequence to fold into two distinct alternative structures (Mironov *et al.*, 2002). This structural change is mediated by the presence of an external effector. The "switch" can be turned on or off by a lower molecular agent, a protein, another RNA molecule, or even in response to higher or lower temperature (Grundy and Henkin, 1993; Narberhaus, 2002; Stülke, 2002; Winkler, 2005). Thus, RNA-switches are *cis* acting elements that control gene regulation by folding into mutually exclusive structures.

Functional RNA-switching elements are composed of two domains: The first domain is a ligand-binding domain that serves as a natural aptamer, and the second one serves as a genetic control element also named expression platform. The genetic control element senses the status of the binding domain and exerts genetic control onto the subsequent coding genes (Winkler and Breaker, 2003). Therefore, RNA-switching elements are generally located in the 5'-untranslated region of their regulated genes.

The ligand-binding sites of RNA switches bind their ligands very efficiently and independently from the presence of the genetic control sequence. Thereby, they are highly selective and discriminate between closely related structures (Nudler, 2006; Reichow and Varani, 2006). This selectivity is mediated by a mix of standard and non standard base pairs, hydrogen bonds, van der Waals, and electrostatic and stacking interactions (Nudler, 2006).

The regulation by the RNA genetic control elements can occur on the level of transcription and translation (Nudler and Mironov, 2004; Vitreschak *et al.*, 2002). In the first case, the transcription of a coding gene is prematurely stopped by an intrinsic terminator. The formation or the suppression of the terminator structure is dependent on the binding of a certain ligand to the ligand-binding domain.

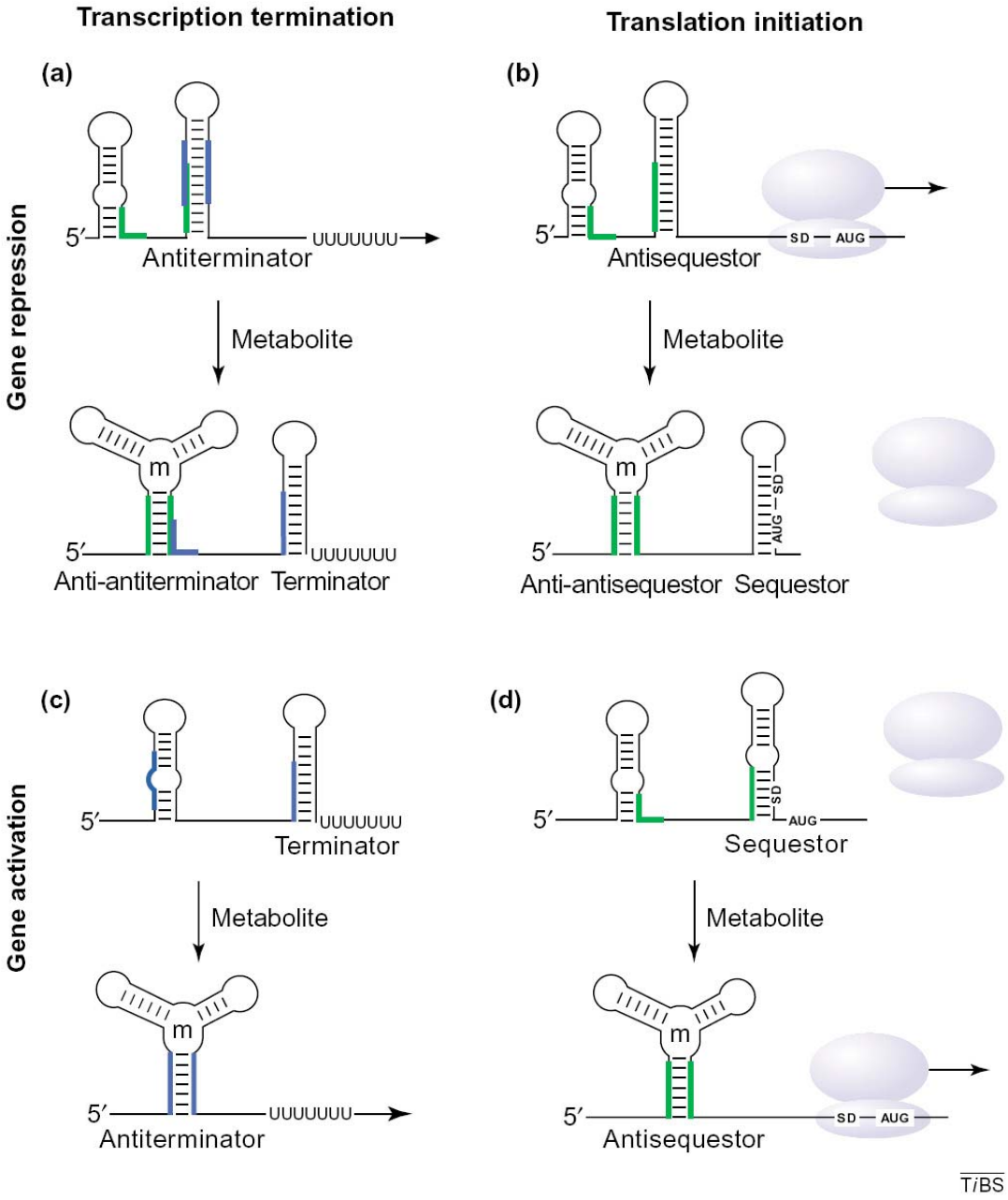

Figure 1.4. Possible regulation modes of RNA switches. Left (a, c): Regulation by transcription termination. The binding of a metabolite causes the formation of a terminator which leads to transcription termination (a), or metabolite binding represses the formation of a stable terminator and enables readthrough into the coding genes (c).
Right (b, d): Regulation by translation initiation. Metabolite binding causes formation of an RNA structure (sequestor) that comprises Shine Dalgarno sequence and start codon which inhibits translation initiation (b). Metabolite binding favours a secondary structure that comprises the part of the sequestor sequence. This represses the structure of the sequestor and gives access to Shine Dalgarno sequence and start codon (d) (Nudler and Mironov, 2004).

The binding of the ligand can theoretically result in repression or activation of the coding gene dependent on the overall architecture of the RNA switch (see Fig. 1.4). As for the second case, regulation occurs when the whole mRNA is already synthesized. The RNA switch controls the accessibility of the Shine Dalgarno (SD) sequence for the ribosomes and thus controls gene regulation on the level of translation initiation (sequestration) (Nudler and Mironov, 2004; Winkler, 2005; Winkler and Breaker, 2003). Examples for both possibilities have already been found in different organisms. Whether the RNA switch represses or activates gene translation upon binding of the ligand, as mentioned above, depends also on the overall design of the riboswitch (see Fig. 1.4). However, most of the known regulatory elements that act on the level of translation initiation repress gene regulation upon ligand binding. There are also known cases where the mRNA is processed upon ligand binding and therefore, these RNA switches can be considered as metabolite dependent ribozymes (Jansen *et al.*, 2006; Kubodera *et al.*, 2003; Sudarsan *et al.*, 2003a; Winkler *et al.*, 2004).

It is possible that RNA switches are actually not able to "switch" their conformation back and forth in response to ligand availability. Perhaps the decision which structure to adopt is only made once in the lifetime of such an RNA molecule. The differentiated structure folding is a time critical process and occurs during transcription elongation depending on the availability of the ligand (Nudler, 2006). Thus, riboswitch folding might be a one way street.

Among bacteria, *E. coli* and *B. subtilis* are the best studied organisms. This also holds true for the research on gene regulation by RNA switches. It is notable, that riboswitches found in *E. coli* are mostly regulating on the level of translation initiation whereas riboswitches found in *B. subtilis* are generally regulating on the level of transcription termination (Nudler and Mironov, 2004).

The number of genes controlled by regulatory RNA elements is hard to evaluate. In 2003, the number of genes that were known to be regulated by riboswitches was 2% of the total genome of *B. subtilis* (Winkler and Breaker, 2003). Two years later the number of genes known to be regulated in that way has increased to 4.1% (Winkler, 2005). Thus, gene regulation by riboswitches seems to be common in bacteria (Barrick *et al.*, 2004). These *cis* acting regulatory RNA elements have been found in many other bacteria (Miranda-Rios *et al.*, 2001; Rodionov *et al.*, 2002; Vitreschak *et al.*, 2002), archaea (Rodionov *et al.*, 2002), and even in eukaryotes like fungi and plants (Sudarsan *et al.*, 2003a; Thore *et al.*, 2006), whereas in humans only one example has been found so far (Bugala *et al.*, 2005).

In all domains of life, the ligand-binding domain of a particular RNA switch is highly conserved, whereas the regulatory element may vary. For example, the thiamine

pyrophosphate binding riboswitch is found in Gram-negative and Gram-positive bacteria as well as in eukaryotes. Although the ligand binding domain is highly conserved among all these organisms (Serganov *et al.*, 2006; Thore *et al.*, 2006), the expression platform does not only vary in sequence but even differs in the mode of genetic control (Kubodera *et al.*, 2003; Mironov *et al.*, 2002; Winkler *et al.*, 2002b). In Gram-positive bacteria, gene regulation is mediated by control of premature transcription termination while in Gram-negative bacteria the genes are controlled by translation initiation. In eukaryotes however, the riboswitch acts as a modulator of splicing. There are also cases known where different genes of the same metabolic pathway are regulated differently in the same organism. The *ypaA* gene of the *Bacillus* and *Clostridium* group is regulated both at the level of transcription and translation (Lee *et al.*, 2001; Winkler *et al.*, 2002a).

Most of the metabolite binding riboswitches are widespread among evolutionary distant organisms. Considering this together with the high conservation of the ligand binding sites of these regulators could lead to the assumption that riboswitches are of ancient origin and relics of the RNA world (Joyce, 2002) (see chapter 1.2.1). The particular metabolite binding sites were optimised throughout evolution and thus show this high similarities among distant organisms. The fact that no additional factors like proteins are involved also supports the theory of an ancient origin of riboswitches (Vitreschak *et al.*, 2004; Winkler and Breaker, 2003). However, the same argument can also be quoted to suggest that RNA switches have been spread by numerous horizontal gene transfers. The lack of necessity for any additional factor makes the transfer from one organism to another a comparatively easy task.

Most of these RNA control elements have been found by comparative *in silico* analyses and are 70 to 200 bases in size. Recently, an RNA switch has been found whose functional region is only 35 bases in size (Winkler, 2005). This indicates the possibility that much more RNA regulatory elements will be found with more precise and better adapted algorithms.

1.2.3.1 Regulation by RNA thermometers

RNA thermometers are RNA structures that are located in the 5'-untranslated region (UTR) of the mRNA of a coding gene, but in contrast to all other RNA switches they are capable of regulating genes without the need of any additional factors (Lai, 2003; Narberhaus, 2002; Narberhaus *et al.*, 2006). Instead of sensing the concentration of molecules, RNA thermometers respond to a physical signal, i.e. temperature. Like other RNA switches, the sequence in the 5'-UTR of an RNA thermometer can adopt two distinct structures. One

structure blocks the Shine Dalgarno sequence and usually also the start codon by intramolecular base pairing, whereas the second structure allows access of the ribosome. The alteration of the RNA structure depends on a temperature shift that can be in some cases as small as 8°C.

A conserved RNA element called ROSE (*R*epression *O*f heat-*S*hock gene *E*xpression) in the 5'-UTR of multiple heat shock genes of *Bradyrhizobium japonicum* was found to control genes by sensing temperature (Münchbach *et al.*, 1999; Narberhaus *et al.*, 1998; Narberhaus *et al.*, 1996). In the 5'-leader of *hspA*, ROSE represses expression at temperatures of 30°C and below, but deletion of this region abolished regulation and resulted in constitutive expression. The ROSE RNA thermometer does also control heat-shock genes in other Rhizobiales and many α- and γ-proteobacteria (Nocker *et al.*, 2001; Waldminghaus *et al.*, 2005). Some of these genes are under dual temperature regulation, as they are also controlled by a σ^{32} dependent promoter.

Not only heat but also cold is a physical condition that can be sensed by RNA structures. Analysis of the *cspA* mRNA, coding for a cold shock gene in *E. coli*, revealed that this gene is controlled by alternative RNA structures in the 5'-leader region of its mRNA. In this case, accessibility for ribosome binding is possible at low temperatures, whereas higher temperatures block translation of this mRNA (Yamanaka *et al.*, 1999). Additionally, rapid RNaseE degradation of the *cspA* mRNA occurs at higher temperatures (Fang *et al.*, 1997). It is notable that all RNA thermometers found so far are regulating gene expression rather on the level of translation than transcription.

1.2.3.2 RNA mediated regulation

Gene regulation that depends on specific RNA-RNA interactions was already discussed in chapter 1.2.2. Here, only gene regulation where a second RNA molecule switches the folding status of an RNA dependent riboswitch will be described.

In many Gram-positive bacteria, a conserved structural RNA element called T-box can be found in the leader region of many aminoacyl-tRNA synthetases, and genes of amino acid biosynthesis and amino acid transport (Gerdeman *et al.*, 2003; Grundy and Henkin, 1993; Grundy and Henkin, 2003). The T-box regulates genes by an intrinsic terminator causing premature transcription termination. However, a conserved sequence in the 5'-region of the controlled mRNAs can directly interact with the appropriate uncharged tRNA. If, for example, the intramolecular concentration of glycine gets low, the ratio of charged to uncharged $tRNA^{Gly}$ changes towards the latter. The intermolecular base pairings between the

uncharged tRNA and the T-box mRNA stabilizes an antiterminator structure, which leads to the transcription of the downstream genes (Grundy *et al.*, 2002; Putzer *et al.*, 2002). The RNA-RNA interaction is specific for the T-Box RNA of a certain gene and the associated uncharged tRNA. This specificity is mainly mediated by two points: First, a specifier sequence in the leader of the mRNA forms base pairing interactions with the tRNA anticodon sequence (Grundy and Henkin, 2003; Grundy *et al.*, 1997). Second, the overall structure of the tRNA is recognized by the leader RNA (Grundy *et al.*, 2000; Yousef *et al.*, 2003). Additionally, the 3'-terminal region of the tRNA pairs with the mRNA and stabilizes a conformation that promotes expression of the coding genes (Grundy *et al.*, 1994). These interactions work in a complex 3D-structural manner and there is also evidence that other factors apart from base specific interactions play a role in target recognition (Fauzi *et al.*, 2005). Genes that are regulated by a tRNA-binding riboswitch are best studied in *B. subtilis*. Among others, the tRNA synthetase genes *cysES* (cysteine), *leuS* (leucine), *pheS* (phenylalanine), *thrS*, *thrZ* (threonine), *glyQS* (glycine), and the *ilv-leu*-operon (Condon *et al.*, 1996; Grundy and Henkin, 2004) are regulated by the T-box riboswitches.

1.2.3.3 Metabolite mediated regulation

In addition to RNA thermometers, the metabolite binding riboswitches might represent the most ancient form of RNA switches, and probably even the oldest form of gene regulation (Vitreschak *et al.*, 2004). Without the need of any additional protein factors, regulation is mediated by direct sensing the concentration of metabolites, e. g. the product of the regulated pathway (Henkin and Yanofsky, 2002). Among the best-studied metabolite binding riboswitches are those that regulate metabolic pathways of vitamin, amino acid, and purine biosynthesis.

The *thi*-box (Miranda-Rios *et al.*, 1997; Miranda-Rios *et al.*, 2001) is an RNA sequence of 39 nucleotides which is able to bind thiamine pyrophosphate (TPP). It is highly specific for TPP binding, as it binds neither its precursor thiamine (vitamine B1) nor thiamine monophosphate (Mironov *et al.*, 2002; Winkler *et al.*, 2002b). This regulatory element is very well documented in *E. coli*, *Rhizobium etli*, and *B. subtilis* (Begley *et al.*, 1999; Petersen and Downs, 1997) and has been found in over 100 other bacteria (Rodionov *et al.*, 2002; Sudarsan *et al.*, 2003a). Furthermore, it has been found in archaea, fungi, and plants. Thus, the *thi*-element is the most widespread riboswitch found so far. It is involved in the regulation of many genes of thiamine biosynthesis and transport. The ligand binding domain of this riboswitch has been found to interact with the negatively charged phosphate groups of TPP

(Serganov *et al.*, 2006; Thore *et al.*, 2006). As RNA itself is a negatively charged molecule, this was unexpected. Interestingly, the structures of TPP binding riboswitches from the different organisms are almost identical, while the mode of regulation can vary between transcriptional or translational level or can even occur by modulating mRNA splicing (Kubodera *et al.*, 2003; Sudarsan *et al.*, 2003a; Winkler *et al.*, 2004).

Another riboswitch involved in the biosynthesis of a vitamin is the *rfn* element (Gelfand *et al.*, 1999; Vitreschak *et al.*, 2002; Winkler *et al.*, 2002a). It has been found in more than 40 different bacterial groups and controls genes involved in the biosynthesis or transport of riboflavin (vitamin B_2), a precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). It is best studied in *B. subtilis* where the *rfn* element is controlling the *ribDEAHT* operon by directly sensing the concentration of FMN (Kil *et al.*, 1992; Mironov *et al.*, 1994). The affinity to FMN is 100-fold stronger than to riboflavin, which is chemically almost identical but lacks an additional phosphate group (Winkler *et al.*, 2002a). The *ypaA* gene of *B. subtilis*, coding for a riboflavin transporter, is regulated at the level of translation as binding of FMN blocks the accessibility of the SD sequence for the ribosomes (Winkler *et al.*, 2002a).

Some genes coding for amino acid synthetic operons are also controlled by riboswitch mechanisms. The S-box is a conserved riboswitch element involved in sulfur metabolism and the synthesis of the sulfur containing amino acids methionine and cysteine (Grundy and Henkin, 1998). This regulatory element has been almost exclusively found in Gram-positive bacteria and is best studied in *B. subtilis* (Grundy and Henkin, 2003). The regulation of these genes is dependent on the availability of S-adenosylmethionine (SAM), an essential coenzyme directly synthesized from methionine (Grundy and Henkin, 1998; Winkler *et al.*, 2003).

The *lys*-element is part of another riboswitch controlling amino acid synthetic genes. Like the S-box, it has been found mainly in Gram-positive bacteria. In *B. subtilis*, it controls the *lysC* gene coding for the first enzyme of lysine biosynthesis (Kochhar and Paulus, 1996; Mandal *et al.*, 2003). The binding domain of this RNA switch is able to discriminate between L- and D-lysine and also rejects binding of ornithine, homolysine, and 5-hydroxylysine which all differ very slightly from each other (Sudarsan *et al.*, 2003b). However, the compound aminoethylcysteine (AEC), which is an analog of L-lysine, is bound by this riboswitch. This molecule is toxic to *B. subtilis* cells and this could be due to the repression of lysine synthetic genes by premature transcription termination caused by the "wrong" ligand (Lu *et al.*, 1992; Winkler and Breaker, 2005).

A glycine dependent riboswitch has been found in a variety of bacteria (Barrick *et al.*, 2004). The conserved element of this riboswitch is called *gcvT*-motif and has been studied in *B. subtilis* and *Vibrio cholerae* (Mandal *et al.*, 2004). In *B. subtilis*, this RNA-motif is involved in regulation of a glycine catabolic operon (*gcvT-gcvPA-gcvPB*) which is necessary to use glycine as an energy source. Interestingly, this riboswitch element has a tandem aptamer configuration with a 1:1 stoichiometry between glycine and each individual aptamer (Mandal *et al.*, 2004). By binding of a single glycin molecule to one of the two aptamer sites, the glycine binding affinity of the second site is improved by 1000-fold. Thus, this data indicates a cooperative binding mode of the riboswitch resulting in an almost "digital" regulation of the controlled operon.

Riboswitches that control purine synthetic operons are also best studied in *B. subtilis*. Genes involved in guanine or adenine metabolism are controlled by a conserved RNA switching element (G-box) (Mandal *et al.*, 2003; Mandal and Breaker, 2004). This regulatory element directly binds guanine (G-switch RNA) or adenine (A-switch RNA) specifically and although both compounds are very similar to each other, no cross-talk has been observed. The *pur* operon, consisting of 12 genes coding for *de novo* synthesis of inosine monophosphate (IMP), and the *xpt-pbuX* operon, encoding xanthine phosphoribosyltransferase and a xanthin transporter, are regulated by the G-box riboswitch (Mandal *et al.*, 2003). The conserved G-box element has also been found in the leader region of three other operons: *pbuG*, *nupG* and *pbuE* which are purine nucleoside transport proteins. The riboswitch of the *pbuE* operon, coding for a purine efflux pump, binds adenine rather than guanine and upon binding of the ligand, gene expression is activated. The difference in ligand specificity is mediated by a single base pair exchange in the ligand-binding domain of these RNA switches. This single base exchange transforms a guanine-binding switch into an adenine-binding one (Lescoute and Westhof, 2005; Mandal and Breaker, 2004; Noeske *et al.*, 2005).

The *glmS*-riboswitch is mainly found in Gram-positive bacteria where it regulates expression of the GlmS protein (Winkler *et al.*, 2004), which catalyzes the conversion of fructose-6-phosphate to glucosamine-6-phosphate (GlcN6P). The *glmS*-riboswitch has a high affinity to GlcN6P and regulates this gene by cleavage of the *glmS*-mRNA upon binding of its ligand. Thus, the *glmS*-riboswitch acts like a GlcN6P dependent ribozyme and is therefore also called *glmS*-ribozyme (Winkler *et al.*, 2004). Newer studies indicate that the *glmS* cleavage site is very close to the ligand binding domain of the riboswitch and GlcN6P could have a coenzyme-like role in the catalytic core of this ribozyme (Jansen *et al.*, 2006).

1.2.3.4 Protein mediated regulation

Protein mediated regulation of an RNA switch can be divided into two subclasses. The first class works similar to the metabolite binding riboswitches but instead of sensing the concentration of a small molecule it measures the concentration of a certain protein. In most cases, the protein is the product of a feedback-regulated pathway. A good example for this class is the feedback repression of some ribosomal proteins (r-proteins) (Zengel and Lindahl, 1994). These proteins are normally associated with ribosomal RNAs, but if synthesised in excess, they bind to the 5'-UTR of their coding gene and negatively affect translation (Springer *et al.*, 1997). There is indication that the leader sequence of the regulated gene mimics the structure of the rRNA r-protein binding site (Merianos *et al.*, 2004; Nomura *et al.*, 1980; Stelzl *et al.*, 2003). This negative feedback regulation ensures that no energy is wasted for r-protein synthesis if its concentration is already in excess to ribosomal RNA.

The second class of protein mediated RNA switches differs from the first as the proteins are specifically designed to regulate gene expression by RNA binding. For an additional level of regulation, these proteins are not just sensed by their concentration but are also regulated by their binding activity. Their binding activity can be regulated either by lower molecular weight factors, by another regulatory protein, or by protein modification events, such as phosphorylation (see also chapter 1.3). As they control transcription termination, they are named terminator or antiterminator proteins, depending on their mode of operation.

A well studied example is TRAP (*trp* RNA-binding attenuation protein), which plays a role in regulation of tryptophan biosynthesis in *B. subtilis* (Babitzke, 2004; Gollnick, 1994; Otridge and Gollnick, 1993). TRAP is composed of 11 identical subunits that form a ring-like structure and it controls expression of the *trpEDCFBA* operon and of *trpG*. The leader sequence of this operon is wrapped around the protein, which leads to transcription termination. TRAP binds to a stretch of 11 (G/U)AG repeats in the leader sequence (Antson *et al.*, 1999; Antson *et al.*, 1995; Gollnick *et al.*, 1995). Additionally, it negatively regulates translation of some transport and efflux genes by blocking ribosomal binding to the mRNA (Babitzke and Gollnick, 2001; Yakhnin *et al.*, 2006; Yakhnin *et al.*, 2004). Activity of TRAP is activated by tryptophan binding. One molecule of tryptophan binds to each TRAP subunit in a cooperative manner. Its activity is negatively regulated when binding to anti-TRAP (AT) which is expressed when the concentration of uncharged $tRNA^{trp}$ is high (Sarsero *et al.*, 2000; Shevtsov *et al.*, 2005; Valbuzzi and Yanofsky, 2001). Therefore, TRAP activity is regulated by a lower molecular weight factor, i.e. tryptophan, as well as by a regulatory protein. Gene

regulation by TRAP is highly nuanced in a way that would not be possible just with a tryptophan dependent riboswitch.

Another example of an RNA switch controlling protein whose activity is regulated by a lower molecular weight factor is PyrR. In contrast to the purine-binding riboswitches (see chapter 1.2.3.3), the pyrimidine concentration is not directly sensed by a regulatory RNA. Instead, the PyrR protein binds as a dimer to the leader sequences in the *pyr*-operon coding for pyrimidine biosynthetic genes (Quinn *et al.*, 1991). The binding sites are located in the 5'-UTR and within the *pyrR-pyrP* and *pyrP-pyrB* intercistronic regions of this operon (Grabner and Switzer, 2003; Zhang and Switzer, 2003). Without binding of PyrR, an antiterminator structure is formed, enabling transcription. Binding of active PyrR to its target sequence forms and stabilizes an anti-antiterminator structure, which prevents antiterminator formation. Under these conditions a terminator structure leads to transcription termination (Switzer *et al.*, 1999). Activity of PyrR is dependent on the availability of pyrimidines. If the intracellular concentration of UMP and UTP is high, PyrR is active and binds as a dimer to its target sequence. In *Bacillus caldolyticus* GMP was found to bind to PyrR and antagonizing its RNA binding ability (Chander *et al.*, 2005). As PyrR itself has uracil phosphoribosyltransferase activity it can be assumed that it was not originally designed to bind RNA structures. Therefore, this might be a good example of how a protein can evolve to exert new functionality and how the RNA coevolved in that case (Tomchick *et al.*, 1998).

There are many more examples of RNA-binding proteins, which modulate RNA-structure upon binding, e.g. HutP (Kumar *et al.*, 2006; Oda *et al.*, 2000; Wray and Fisher, 1994), GlpP (Rutberg, 1997), AmiR (Wilson and Drew, 1995; Wilson *et al.*, 1993), NasR (Chai and Stewart, 1998; Chai and Stewart, 1999), and BglG (Amster-Choder and Wright, 1992; Mahadevan and Wright, 1987; Schnetz and Rak, 1990). They can be found in many different kinds of bacteria. Antiterminator proteins that are controlled by reversible phosphorylation events, like the BglG family of antiterminator proteins, are discussed in chapter 1.3.

1.2.3.5 RNA aptamers: control by artificial regulatory RNA elements

Genetic control elements, like the *lac*-operon regulatory system (Kercher *et al.*, 1997), are common tools for controllable expression of gene constructs. Natural riboswitches are able to regulate gene expression without the need of any additional protein factors and could therefore expand the repertoire of genetic control elements with a more general applicable system. There was not much known about these natural regulatory elements when some

scientists started to make use of the specific binding affinity that RNA provides. With an *in vitro* selection method called SELEX (Ellington and Szostak, 1990) they were able to screen for RNA sequences (aptamers) that specifically bind to small organic molecules. Even small proteins can bind to RNAs selected by this method (Hermann and Patel, 2000). With lessons learned from natural RNA switches it was possible to engineer artificial but functional RNA-elements capable of switching gene expression dependent on ligand binding (Werstuck and Green, 1998).

A good example is the Tet-repressor system (Orth *et al.*, 2000; Saenger *et al.*, 2000). This system derived from *Escherichia coli* where it controls tetracycline dependent gene expression. TetR binds to the *tet*-operator if no tetracycline is available and thereby blocks transcription. If the antibiotic is present, it binds TetR and induces a structural change that releases the protein from the DNA. The system is used as a genetic switch to control gene expression in bacteria and eukaryotes. However, this system has some limitations in eukaryotes. Therefore, a system was engineered to control expression without the need of a protein factor. A tetracycline binding aptamer, found by SELEX, was placed in the 5'-UTR of a reporter gene and this construct was functional in controlling gene expression in yeast dependent on the availability of tetracycline. If the antibiotic is available, it binds to the RNA and interferes with its accessibility for the ribosomes (Hanson *et al.*, 2005; Müller *et al.*, 2006; Suess *et al.*, 2003).

Another example is a synthetic theophylline responsive riboswitch. A theophylline binding aptamer was combined with a translational control element that exerts regulation. Normally, translation is blocked but binding of theophylline causes a minor structural reorganisation of the control element, which leads to efficient translation (Desai and Gallivan, 2004; Suess *et al.*, 2004).

Aptamers are also used as biosensor elements for molecular analysis ranging from small molecules to proteins (Hesselberth *et al.*, 2003; Seetharaman *et al.*, 2001). However, the binding affinity of these artificially isolated RNA molecules to their targets is far from that observed for natural riboswitches (Nudler, 2006).

1.3 PTS-controlled antitermination in *B. subtilis*

1.3.1 The PTS and its role in sugar transport

The phosphoenolpyruvate:sugar phosphotransferase system (PTS) plays a central role in the regulation of catabolic operons in many bacteria (Meadow *et al.*, 1990; Postma *et al.*, 1993; Reizer *et al.*, 1988). As mentioned above (see chapter 1.1.1), the PTS consists of general and sugar specific components. The sugar specific enzymes (enzyme II or EII) consist of three or four domains. These domains can be separate proteins or fused together forming single polypeptides (Postma *et al.*, 1993). The general proteins are the cytoplasmatic proteins enzyme I (EI) and HPr. EI transfers phosphate from phosphoenolpyruvate (PEP) to HPr. HPr features two phosphorylation sites. The phosphate coming from EI is transferred to position His-15 of HPr, and then passed to a sugar specific EII. When the relevant sugar is available, it is phosphorylated by EIIBA, whereas EIIC is involved in the uptake process (see Fig 1.5). Thus, the membrane spanning sugar permeases accomplish sugar uptake paired with simultaneous phosphorylation. The phosphate derived from PEP is transferred from phosphoenolpyruvate to the sugar via EI, HPr (His15), and EII (see Fig. 1.5).

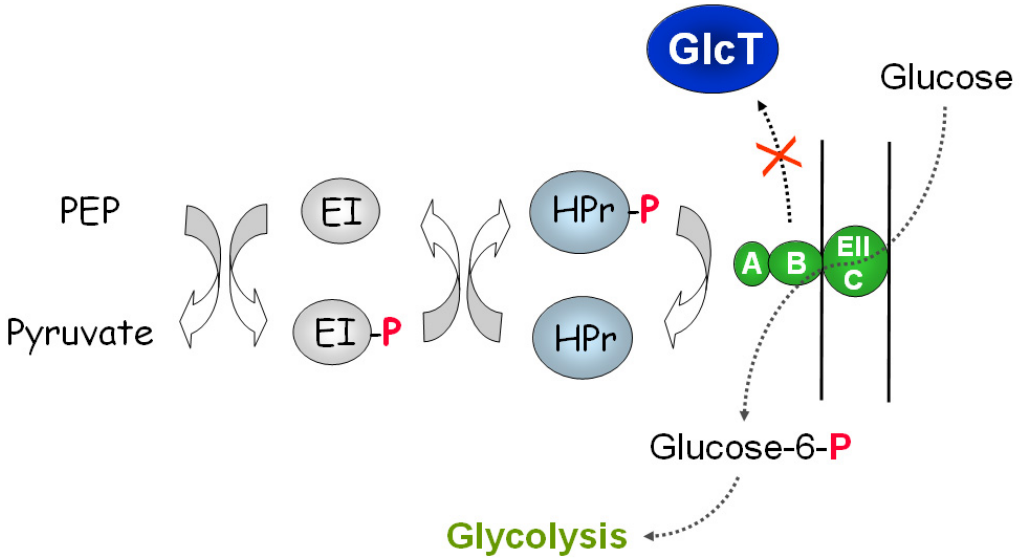

Figure 1.5. Schematic overview on the glucose-specific phosphotransferase system (PTS) of *B. subtilis*. The phosphoryl residue (-P) is transferred from PEP via enzyme I (EI), the HPr protein (His-15 residue), and enzymes EIIA (A) and EIIB (B) to the incoming glucose. The glucose uptake is mediated by the membrane spanning enzyme EIIC. If glucose is available, the phosphate residue is transferred to the incoming glucose. In the absence of glucose, the phosphate residue is transferred to the transcriptional antiterminator protein GlcT. Modified from Bachem and Stülke, 1998.

In the presence of a PTS sugar the general components and the sugar specific permease are mainly in a non-phosphorylated state. If the sugar for a certain permease is not present, the phosphate accumulates at that permease. In such a case, the phosphate can be transferred to a regulatory protein.

The first of these proteins described was BglG from *E. coli* (Amster-Choder *et al.*, 1989; Amster-Choder and Wright, 1992; Mahadevan and Wright, 1987). It modulates transcriptional antitermination of the *bgl*-operon and is therefore called "antiterminator protein". The activity of this "antiterminator" protein is controlled by the phosphorylation state of its cognate EII, i.e. BglF (Amster-Choder *et al.*, 1989). This control mechanism involves a phosphoryl-transfer from EII to the antiterminator protein, leading to its inactivity in the absence of β-glucosides. In the presence of β-glucosides, the antiterminator protein remains in its unphosphorylated state and prevents transcription termination of its target RNA, the *bglP*-operon (Amster-Choder and Wright, 1993). In *B. subtilis*, there are four antiterminator proteins of the BglG family. These regulatory proteins can bind to RNA structures found in the 5'-UTR of some genes involved in the uptake and metabolism of the appropriate sugars. Upon binding, they prevent the formation of intrinsic terminators and enable transcription of the coding genes.

1.3.2 Regulation of the *ptsGHI* operon by transcriptional antitermination

The *ptsGHI* operon of *B. subtilis* codes for the sugar specific permease EII^{Glc} (*ptsG*), and the general components HPr (*ptsH*) and Enzyme I (*ptsI*). Although there are constitutive promoters in front of *ptsG* and *ptsHI*, only HPr and EI are constitutively synthesized (Gonzy-Tréboul *et al.*, 1989; Stülke *et al.*, 1997). High amounts of EII^{Glc} can only be found if glucose is available. The *ptsG* mRNA is only found when glucose is present, which indicates that *ptsG* might be controlled at the level of transcription. Further studies revealed that *ptsG* expression is controlled by premature transcription termination (Stülke *et al.*, 1997). If glucose is not available, only a short transcript of *ptsG* is made. A terminator stem-loop located between the promoter region and the coding region of *ptsG* is formed, causing transcriptional antitermination. In the presence of glucose, a second structure called RAT (RNA antiterminator) is formed, preventing terminator formation (see Fig. 1.6) (Langbein *et al.*, 1999). Thus, the regulation mode of the *ptsG* RNA switch is similar as described in Figure 1.4c. However, unlike metabolite binding riboswitches, the regulation of the *ptsG* RNA switch is not mediated by a direct glucose-RNA interaction. Instead, the regulatory

protein GlcT causes transcriptional antitermination in the presence of glucose (Stülke *et al.*, 1997).

The *glcT* gene is located upstream of the *ptsGHI* operon, and its transcription is under the control of a constitutive promoter (see Fig. 1.6). The activity of GlcT is controlled by a phosphorylation event that indirectly correlates to the glucose concentration. When glucose is not present, GlcT is phosphorylated. Phosphorylated GlcT is inactive and does not bind the RAT RNA sequence. When glucose is present, GlcT remains in its unphosphorylated but active state (see Fig. 1.5).

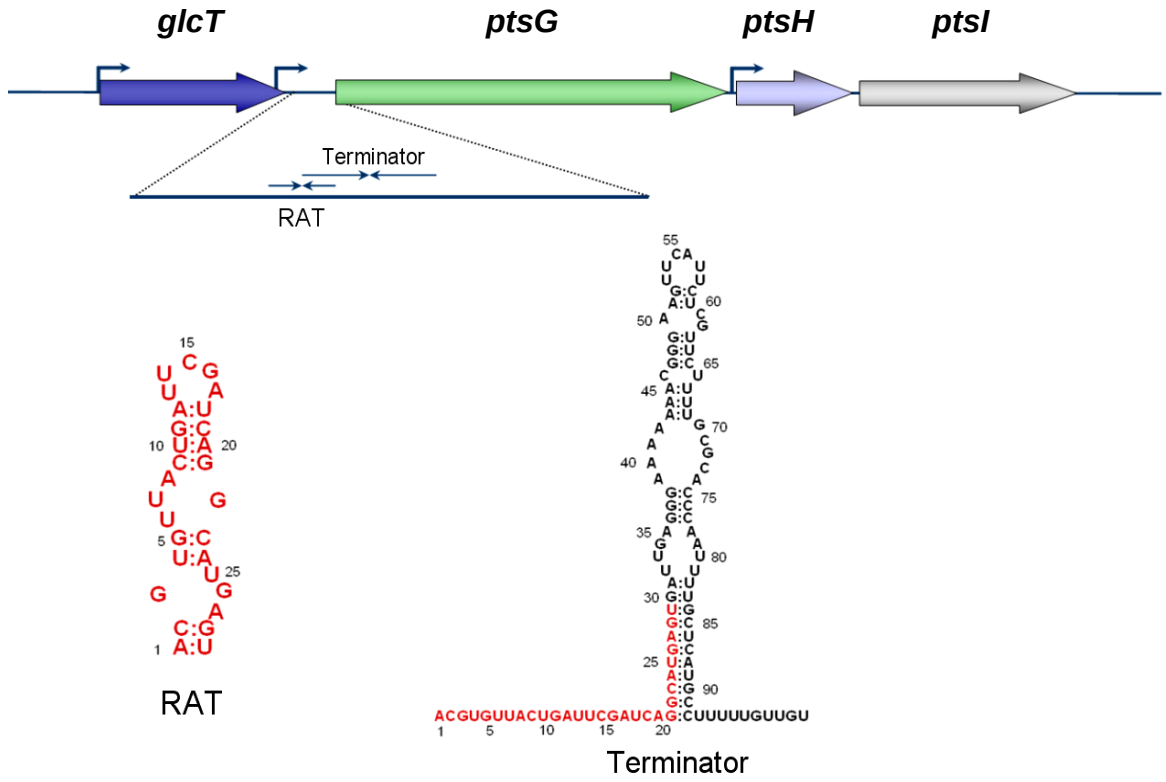

Figure 1.6. The *ptsGHI* operon and a proposed model of the antiterminator and terminator structures of the *ptsG* leader mRNA. The *glcT* gene is located upstream of the *ptsGHI* operon. The *ptsGHI* operon codes for the glucose specific permease *ptsG*, and the general PTS enzymes HPr and Enzyme I (*ptsI*). The promoters shown upstream of *glcT* and *ptsGHI* as well as the promoter upstream of *ptsH* are constitutively expressed.
In the untranslated region of *ptsG*, the RNA sequence can form two different secondary structures. The terminator is the thermodynamically more stable structure and causes transcriptional termination when formed. The formation of the RAT (RNA antiterminator) structure needs to be stabilized by the protein factor GlcT. Since RAT and terminator are overlapping sequences, both structures are mutually exclusive structures. Modified from Stülke and Hillen, 2000.

GlcT consists of an RNA-binding domain (RBD), and two PTS regulatory domains (PRD-I and PRD-II, see Fig. 1.7) (Stülke *et al.*, 1998). If no glucose is present, GlcT is phosphorylated at PRD-I by EII^{Glc} and thereby inactivated (Schmalisch *et al.*, 2003). However, if glucose is present, the sugar is phosphorylated rather than PRD-I by EII^{Glc} and GlcT becomes active. Active GlcT binds as a dimer of two identical subunits to its RNA target sequence and stabilizes the RAT secondary structure. As the terminator sequence overlaps the RAT sequence partially, both structures are mutually exclusive (see Fig. 1.6). Thus, the formation of the RAT structure with the help of GlcT suppresses the thermodynamically more stable terminator and enables the transcription into the coding gene. It was also demonstrated that a protein consisting only of the RBD is sufficient to cause antitermination at the *ptsGHI* operon (Bachem and Stülke, 1998). This shortened protein can dimerize and binds to the RAT-sequence. As it can not be inactivated due to the lack of PRD-I, it shows constitutive activity and thus causes constitutive expression of *ptsG*. Besides the negative phosphorylation of GlcT by the sugar permease EII^{Glc} at PRD-I, the antiterminator protein can also be phosphorylated at PRD-II by HPr. HPr dependent phosphorylation of active GlcT at PRD-II has a slightly positive effect on the GlcT activity (Schmalisch *et al.*, 2003).

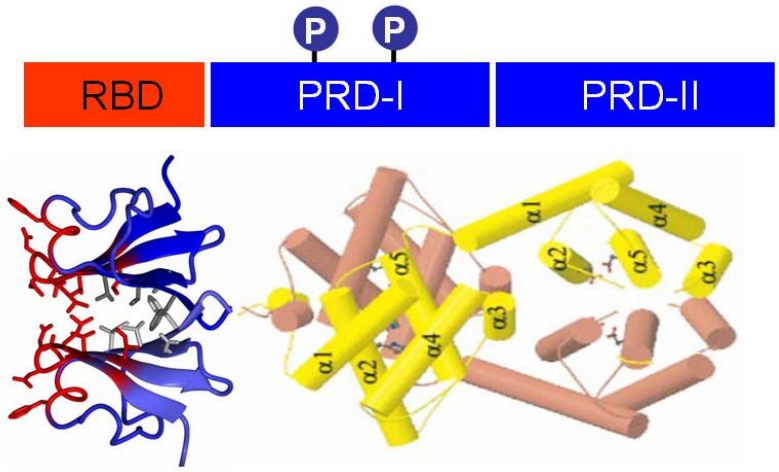

Figure 1.7. Schematic illustration of the antiterminator protein GlcT. Up to now, the structure of the full length GlcT protein could not be obtained. The picture of the RNA-binding domain (RBD) was taken from an NMR structure of the GlcT-RBD (Müller, 2001). PRD-I and PRD-II are taken from the crystal structure of LicT (van Tilbeurgh *et al.*, 2001). The phosphorylation sites of PRD-II are not shown as they are not relevant for GlcT activity.

1.3.3 The family of homologous antiterminator systems of *B. subtilis*

In *B. subtilis*, there are four PTS controlled antiterminator proteins regulating the transcription of several genes (see Fig. 1.8). These antiterminator proteins have all the same domain organization as GlcT (see Fig. 1.7) (Manival *et al.*, 1997; Stülke *et al.*, 1998; Tortosa *et al.*, 1997). The PRD-I is negatively phosphorylated by the cognate EII if the appropriate sugar is not available, whereas the PRD-II is positively phosphorylated by HPr if no better carbon source is available. Similar than GlcT does for the *ptsGHI* operon, LicT controls the transcription of the *bglPH* operon and *licS* (β-glucoside permease, phospho-β-glucoside hydrolase, and β-1,3-1,4-glucanase) by transcriptional antitermination. These genes are involved in the utilization of aryl-β-glucosides and β-glucans (Krüger and Hecker, 1995; Schnetz *et al.*, 1996). The sugar permease BglP controls LicT activity by phosphorylation of the PRD-I, depending on the availability of β-glucosides (Lindner *et al.*, 1999). The antiterminator protein SacT controls the *sacPA* operon (sucrose permease and sucrase). SacP mediates the uptake of sucrose, and SacT activity is controlled by SacP dependent phosphorylation of PRD-I (Arnaud *et al.*, 1996; Arnaud *et al.*, 1992).

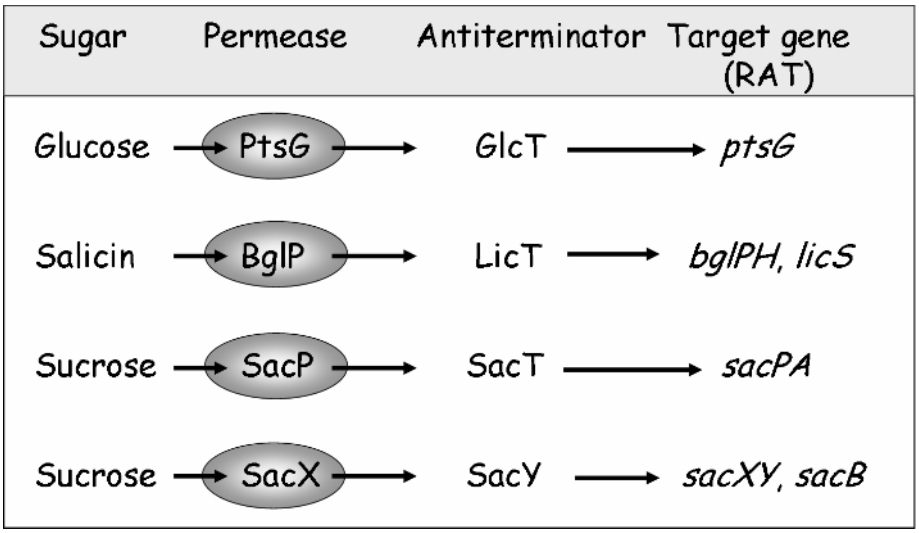

Figure 1.8. A family of antitermination systems controls sugar uptake and metabolism in *B. subtilis*. The design of the four signaling systems. The target genes encode the following proteins: *ptsG*, glucose permease of the PTS; *bglP*, β-glucoside permease of the PTS; *bglH*, phospho-β-glucoside hydrolase; *licS*, β-1,3-1,4-glucanase; *sacP*, sucrose permease of the PTS; *sacA*, sucrase; *sacX*, sucrose permease of the PTS (low affinity); *sacY*, antitermination protein; *sacB*, extracellular levansucrase.

SacY activity is controlled by SacX, which codes for a low affinity sucrose permease. SacX only transports sucrose when the extracellular sucrose concentration is high. Thus, SacY is only active when plenty of sucrose is available. It controls the *sacXY* operon and *sacB* (extracellular levansucrase) (Arnaud *et al.*, 1996; Crutz *et al.*, 1990; Tortosa *et al.*, 1997). To become active, LicT and SacT need to be dephosphorylated at PRD-I and positively phosphorylated at PRD-II (Arnaud *et al.*, 1996; Lindner *et al.*, 1999). When glucose is simultaneously available with sucrose or β-glucosides, the PRD-I of the relevant antiterminator protein is dephosphorylated, but the HPr dependent positive phosphorylation is lacking as most of the HPr molecules phosphorylate the E-II^{Glc} and not the PRD-II of the respective antiterminator proteins. This prevents the activity of SacT or LicT in the presence of the prefered carbon source glucose. The PRD-II of GlcT and SacY can also be phosphorylated by HPr, but the activity of GlcT and SacY is not severely influenced by the phosphorylation state of PRD-II (Bachem and Stülke, 1998; Schmalisch *et al.*, 2003; Tortosa *et al.*, 1997).

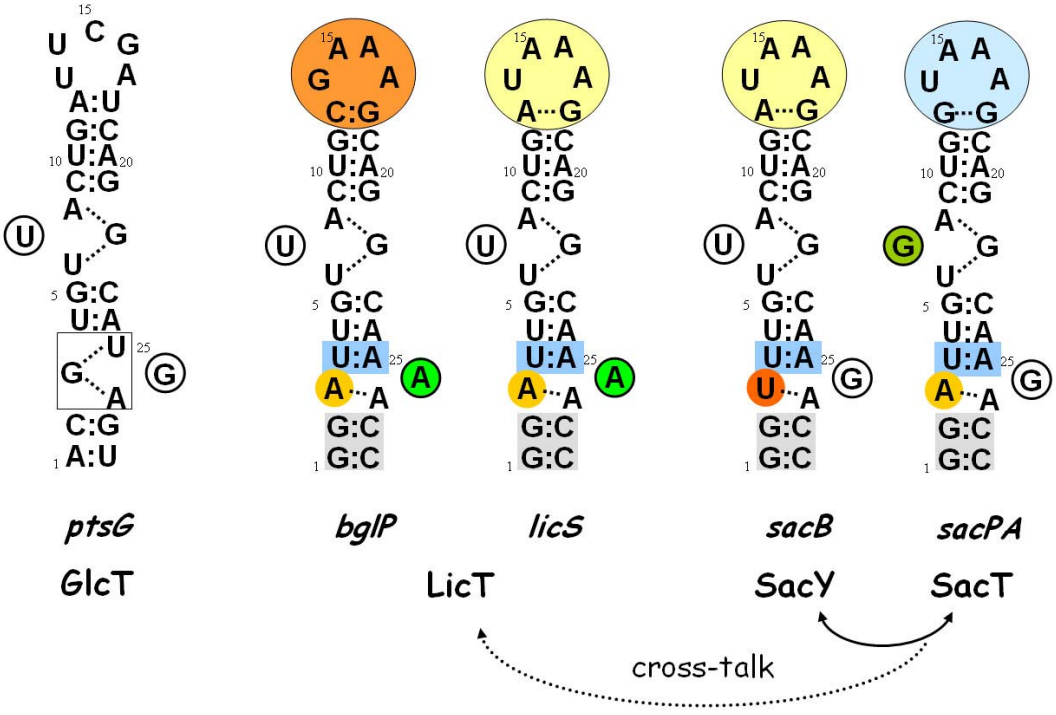

Figure 1.9. Proposed secondary structures for the related RAT RNAs in *B. subtilis* (Yang *et al.*, 2002). The relevant antiterminator proteins are indicated below their cognate RAT structures. Boxes and circles indicate nucleotides that differ from the *ptsG* RAT. Dashed lines indicate bases that are proposed to be in direct contact to each other. The main differences are located in the top loop and lower loop structures. The *ptsG* RAT holds an outstanding position due to its unique lower stem loop and the triple base pairing in the lower loop region. Binding of sucrose antiterminator proteins to RAT sequences of *bgl* genes was reported previously. Also cross-talk within the sucrose related antiterminator systems has been revealed (Aymerich and Steinmetz, 1992).

Since all these proteins bind their target RNA as a dimer, it is most likely that the phosphorylation state of the PRDs modulates the dimerization ability of the two identical subunits. Although the four antiterminator proteins are part of four separate regulatory systems, the proteins are highly similar to each other. Not only the basic structures of the proteins are identical (see Fig. 1.7), but also the domains RBD, PRD-I, and PRD-II are each very similar among the related antiterminator proteins (Débarbouillé *et al.*, 1990; Greenberg *et al.*, 2002; Stülke *et al.*, 1998; Tortosa *et al.*, 1997). But not only the proteins share high homology, also the cognate RAT structures of the different systems share high similarity in sequence and structure (see Fig. 1.9) (Aymerich and Steinmetz, 1992; Yang *et al.*, 2002). The antiterminator proteins of the sucrose systems, SacT and SacY, seem to be able to recognize the RAT structure of their respective counterpart. It was also reported that the antiterminator proteins of the sucrose systems can bind to the RATs of the *bgl* genes (Aymerich and Steinmetz, 1992; Steinmetz *et al.*, 1989). However, neither can LicT bind to *sacB* or *sacP* RATs, nor does sucrose induce *licS* or the *bglPH* operon. The GlcT/*ptsG* RAT couple is the most distant system and no cross-talk to other systems has been observed.

In general, not only the RAT/antiterminator protein recognition, but also the positive and negative phosphorylation of the PRDs, and also carbon catabolite repression, which limits the expression of some systems under certain conditions, contribute to the regulatory specificity.

1.4 Aim of this work

B. subtilis is one of the best-studied organisms so far. Although much effort has been devoted to the elucidation of the genes involved in central metabolism, not much is known about their regulation and the regulatory networks by which each individual pathway is connected to the others. In this work, a global approach combining DNA microarrays and metabolic flux analysis should allow to draw conclusions not only about gene regulation on a transcriptomic level, but also about regulatory events on protein level. Therefore, the first aim was the design and establishment of working methods for global screening systems based on transcriptome and fluxome analysis. As a first proof of concept, the influence of organic acids, like glutamate and succinate, on the metabolism of *B. subtilis* in addition to glucose should be determined in minimal medium.

The PTS is part and also trigger of regulatory events in carbohydrate metabolism in *B. subtilis*. Some of the involved genes are regulated by premature transcription termination controlled by a protein dependent RNA switch. In *B. subtilis*, there is a family of four systems sharing the same mode of regulation. The components of these systems are highly conserved among each other and cross-talk for some systems was reported *in vitro*. The GlcT/*ptsG* system holds an outstanding position. Neither does GlcT bind any other riboswitch structures than that of *ptsG* RAT, nor do any other antiterminator proteins than GlcT bind to this riboswitch. *In vivo* only the genes needed for the utilization of the currently available carbon source are activated. Moreover, if several sugars are available, the bacterium only metabolizes the best carbon source. The second aim of this work was the characterization of the specificity determinants of the GlcT/*ptsG* antiterminator system and in a further approach the characterization of all factors needed to maintain specificity for the four different systems.

As site-directed mutagenesis was a major tool to determine the specificity determinants of the different protein/RAT couples, a fast and efficient method for the introduction of multiple mutations in the RAT sequences should be developed.

2 Transcriptional and metabolic responses of *Bacillus subtilis* to the availability of organic acids: Transcription regulation is important but not sufficient to account for metabolic adaptation

The results described in this chapter were published in:

Schilling, O., O. Frick, C. Herzberg, A. Ehrenreich, E. Heinzle, C. Wittmann, and J. Stülke. 2007. Transcriptional and metabolic responses of *Bacillus subtilis* to the availability of organic acids: Transcription regulation is important but not sufficient to account for metabolic adaptation. *Appl Environ Microbiol* **73:**499-507.

Author contributions:

This study was designed by O.S., O.F., C.W., E.H., and J.S.. O.S. performed and interpreted all reporter gene and transcriptome analyses. O.F. performed and interpreted all fluxome analyses. C.H. contributed substantially to all lab work related to transcriptome analysis. A.E. provided the equipment needed for microarray design and analysis. J.S. and O.S. wrote the paper.

Abstract:

The soil bacterium *Bacillus subtilis* can use sugars or organic acids as sources of carbon and energy. These nutrients are metabolized by glycolysis, the pentose phosphate pathway, and the Krebs citric acid cycle. While the response of *B. subtilis* to the availability of sugars is well understood, much less is known about the changes in metabolism if organic acids feeding into the Krebs cycle are provided. If *B. subtilis* is supplied with succinate and glutamate in addition to glucose, the cells readjust their metabolism as determined by transcriptome and metabolic flux analyses. The portion of glucose-6-phosphate that feeds into the pentose phosphate pathway is significantly increased in the presence of organic acids. Similarly, important changes were detected at the level of pyruvate and acetyl coenzyme A (acetyl-CoA). In the presence of organic acids, oxaloacetate formation is strongly reduced, whereas the formation of lactate is significantly increased. The *alsSD* operon required for acetoin formation is strongly induced in the presence of organic acids; however, no acetoin formation was observed. The recently discovered phosphorylation of acetolactate decarboxylase may provide an additional level of control of metabolism. In the presence of organic acids, both types of analyses suggest that acetyl-CoA was catabolized to acetate rather than used for feeding the Krebs cycle. Our results suggest that future work has to concentrate on the posttranslational mechanisms of metabolic regulation.

Introduction

Bacillus subtilis and its close relatives are among the most important industrial microorganisms. They are used for the production of antibiotics, enzymes such as proteases, amylases, lipases, or cellulases, vitamins such as riboflavin, and even insecticidal proteins. Due to this practical relevance, much research has been devoted to many aspects of these bacteria, making *B. subtilis* one of the best-studied living organisms.

The basis for understanding the potential of *B. subtilis* is a firm knowledge of its metabolism. These bacteria prefer sugars and ammonia as carbon and nitrogen sources, respectively. The preferred source of carbon and energy is glucose. This sugar is taken up and concomitantly phosphorylated by the bacterial phosphoenolpyruvate:sugar phosphotransferase system. The further metabolism involves glycolysis, the pentose phosphate pathway, and the Krebs cycle, resulting in the oxidation of glucose to carbon

dioxide and the generation of ATP, a proton-motive force, reducing power, and intermediates for all anabolic reactions. Similarly, other sugars and polyols can be phosphorylated and are catabolized in a similar manner (for a review see Stülke and Hillen, 2000). The utilization of organic acids requires the conversion of these acids to intermediates of the Krebs cycle, gluconeogenesis, and the pentose phosphate pathway to obtain both the energy and the metabolic precursors needed by the cell (Doan *et al.*, 2003). The central intermediate of nitrogen metabolism is glutamate, the universal amino-group donor for the biosynthesis of amino acids, nucleotides, and all other nitrogen-containing compounds in the cells (for a review see Fisher, 1999).

The details of the central carbon and nitrogen metabolic pathways in *B. subtilis* and their regulation have been the subject of extensive investigations. The *ptsGHI* operon, encoding the general and glucose-specific phosphotransferase system proteins, is induced by glucose due to a mechanism of transcriptional antitermination (Schmalisch *et al.*, 2003; Stülke *et al.*, 1997). The glycolytic enzymes that are needed for reversible reactions are constitutively transcribed, whereas those that catalyze irreversible reactions (*pfkA*, *gapA*, and *pyk*, encoding phosphofructokinase, glyceraldehyde-3-phosphate dehydrogenase, and pyruvate kinase, respectively) are induced in the presence of glucose (Fillinger *et al.*, 2000; Ludwig *et al.*, 2001). The complex *gapA* operon, containing six genes encoding enzymes of triose phosphate interconversions, is repressed by the CggR repressor in the absence of glycolytic substrates, and the mRNA is in addition subject to posttranscriptional processing. Moreover, an internal promoter ensures that those enzymes catalyzing reversible reactions are present under both glycolytic and gluconeogenic conditions (Doan and Aymerich, 2003; Meinken *et al.*, 2003). The *pdhABCD* operon, encoding pyruvate dehydrogenase, the link between glycolysis and the Krebs cycle, is also induced by glucose by an as-yet-unknown mechanism (Blencke *et al.*, 2003). Several genes encoding enzymes of the Krebs cycle, such as *citZ* and *citB*, coding for the major citrate synthase and aconitase, respectively, are subject to a synergistic repression by glucose and a preferred nitrogen source such as glutamine. This repression is mediated by the LysR-type repressor protein CcpC (Blencke *et al.*, 2006; Jourlin-Castelli *et al.*, 2000). Even though it is important under all conditions, not much is known about the regulation of the pentose phosphate pathway. A recent transcriptome study suggested that the genes encoding the enzymes of this pathway are constitutively expressed (Blencke *et al.*, 2003).

The assimilation of ammonium is initiated by its uptake by diffusion and by the AmtB (also designated NrgA) ammonium transporter at high and low ammonium concentrations, respectively (Detsch and Stülke, 2003). In *B. subtilis*, ammonium is assimilated exclusively

by the glutamine synthetase/glutamate synthase pathway. Briefly, glutamine is synthesized from glutamate and ammonium, and the glutamine is used as the amino-group donor for the production of two molecules of glutamate with 2-oxoglutarate as the acceptor (Belitsky *et al.*, 2000). The gene encoding glutamine synthetase (*glnA*) is controlled by the availability of glutamine through the GlnR repressor (Schreier *et al.*, 1989). The *gltAB* operon encoding glutamate synthase is subject to a dual control by the carbon and nitrogen sources: in the absence of ammonium, expression of the operon is repressed by the TnrA transcription factor, whereas transcription is activated by the GltC protein in the presence of sugars (Belitsky *et al.*, 2000; Commichau *et al.*, 2006; Wacker *et al.*, 2003).

Table 1. *B. subtilis* strains used in this study

Strain	Genotype	Source[a]
168	*trpC2*	Laboratory collection
BSIP1114	*trpC2 amyE*::(*pta-lacZ cat*)	Presecan-Siedel *et al.*, 1999
GP133	*trpC2 amyE*::(*fbaA-lacZ cat)*	Ludwig *et al.*, 2001
GP205	*trpC2 amyE*::(*citB-lacZ cat*)	Blencke *et al.*, 2006
GP216	*trpC2 amyE*::(*pdhA-lacZ cat*)	pGP721 → 168[a]
GP250	*trpC2 amyE*::(*nrgA-lacZ aphA3*)	Detsch and Stülke, 2003
GP268	*trpC2 amyE*::(*glnR-lacZ aphA3*)	pGP189 → 168[a]
GP313	*trpC2 amyE*::(*cggR-lacZ cat*)	Ludwig *et al.*, 2001
GP314	*trpC2 amyE*::(*pgi-lacZ cat*)	Ludwig *et al.*, 2001
GP315	*trpC2 amyE*::(*pfkA-lacZ cat*)	Ludwig *et al.*, 2001
GP317	*trpC2 amyE*::(*pgk-lacZ cat*)	Ludwig *et al.*, 2001
GP342	*trpC2 amyE*::(*gltA-lacZ aphA3*)	Wacker *et al.*, 2003
QB5556	*trpC2 amyE*::*(rocD'-'lacZ cat)*	Gardan *et al.*, 1995
QB7041	*trpC2 amyE*::(ΔCA*ptsG-lacZ cat*)	Stülke *et al.*, 1997

[a] Arrows indicate construction by transformation.

While much work has been devoted to the effects of the presence of glucose on gene expression and the cellular physiology of *B. subtilis*, relatively little is known about the response towards organic acids that feed into the Krebs cycle. These intermediates of the Krebs cycle and glutamate are known to induce specific transporters (Asai *et al.*, 2000; Tanaka *et al.*, 2003; Yamamoto *et al.*, 2000). Moreover, glutamate represses the genes of the Krebs cycle as long as ammonium and glucose are also available (Blencke *et al.*, 2006;

Sonenshein, 2002). The impact of these Krebs cycle intermediates on the general metabolism has not been studied so far.

In this work, we investigated the response of *B. subtilis* to the presence of organic acids. This response was determined at the level of the subtranscriptome of genes of the central carbon and nitrogen metabolism and by assaying the carbon fluxes. Our results indicate that overflow metabolism, the Krebs cycle, and the reactions involved in the interconversions between glutamate and 2-oxoglutarate are strongly affected by the presence of the organic acids. Importantly, the data allow us to derive conclusions relating to possible posttranscriptional regulation events that had escaped attention until now.

Materials and Methods

Bacterial strains, growth conditions, and enzyme assays.

The *B. subtilis* strains used in this study are listed in Table 1. They are all isogenic derivatives of the strain 168. Bacteria were grown in C minimal medium supplemented with the auxotrophic requirement (tryptophan at 50 mg liter^{-1}) (Faires *et al.*, 1999). Other media consisted of C medium supplemented with 99% [1-^{13}C]glucose (CG medium) or C medium supplemented with a mixture of 99% [1-^{13}C]glucose, succinate, and glutamate (CGSE medium). The labeled glucose was obtained from Campro Scientific (Veenendaal, The Netherlands). Glucose, succinate, and glutamate were added to final concentrations of 0.5%, 0.6%, and 0.8%, respectively (Faires *et al.*, 1999). Parallel cultivations were performed in 500-ml shake flasks using a culture volume of 50 ml. The resulting physiological data on CG and CGSE are listed below (Table 2). Aliquots from identical cultures grown to mid-log phase (optical density at 600 nm of 1.5) were used for transcriptome and metabolic flux analyses. Quantitative assays of *lacZ* expression in *B. subtilis* were performed with cell extracts using *o*-nitrophenyl-ß-galactopyranoside as the substrate, as described previously (Martin-Verstraete *et al.*, 1992).

Plasmid constructions.

Translational fusions with the *lacZ* gene were constructed using the vector pAC5 (Martin-Verstraete *et al.*, 1992), which carries the pC194 chloramphenicol resistance gene *cat*, or the vector pAC7 (Weinrauch *et al.*, 1991), containing the kanamycin resistance gene *aphA3*.

Table 2. Specific rates of growth, substrate consumption, and product formation and yields of *B. subtilis* cultivated on CG and CGSE

Measurement	Result for cultivation on:	
	CG	CGSE
Specific growth rate μ [h^{-1}]	0.64	0.81
Specific rates [mmol g^{-1} h^{-1}] for:		
Glucose	-9.1	-10.0
Succinate	0	-1.0
Glutamate	0	-4.2
Acetate	5.3	8.3
Lactate	0	2.5
Pyruvate	0.1	0.6
Yield coefficients [mol mol^{-1}] for:		
Acetate	0.58	0.83
Lactate	0	0.25
Pyruvate	0.02	0.06
2-Oxoglutarate	0.01	0.06
Cell dry mass [g $mmol^{-1}$]	0.07	0.08

Both plasmids harbor a *lacZ* gene without a promoter located between two fragments of the *B. subtilis amyE* gene. Plasmid pGP189, containing a *glnR-lacZ* fusion, was constructed as follows. A 331-bp fragment containing the promoter region of the *glnRA* operon was amplified by PCR using the primers CD43 (5' AAAGAATTCCATTATGGCA GCAGGGACGTT) and CD44 (5' TCCTATTGGATCCAAAGGCATTGAGCG). The PCR products were digested with *EcoR*I and *BamH*I (restriction sites were introduced by PCR and are underlined in the primer sequences) and ligated with pAC7 linearized with the same enzymes. Similarly, plasmid pGP721, used for the construction of the *pdhA-lacZ* fusion, was obtained using the primers HMB49 (5' CGGAATTCGATGCTGCAGGCTATCGTG) and HMB50 (5' CGGGATCCGCTTTTTTCGTTTTTGCAGCC) with pAC5 as the cloning vector.

RNA isolation and quality assessment.

To isolate high-quality RNA suited for the detection of long transcripts and their precursors, the following procedure was used. The cells were harvested at the exponential phase. For RNA preparation, 12 ml of a cell suspension (optical density at 600 nm of 1.5) was used. After mechanical cell disruption, the frozen powder was instantly resuspended in 3 ml lysis buffer (4 M guanidine isothiocyanate, 0.025 M sodium acetate, pH 5.3, 0.5% [wt/vol] *N*-laurylsarcosine). Subsequently, total RNA extraction with acid phenol solution was performed as described previously (Ludwig *et al.*, 2001). The quality of the isolated RNA was tested by PCR to exclude the presence of traces of DNA and by Northern blot analysis to detect a large nonabundant mRNA (7.4 kb mRNA of the *gapA* operon). Digoxigenin RNA probes specific for *gapA* (Tobisch *et al.*, 1999b) were used to detect the transcript of the *gapA* operon. The sizes of the RNA molecular weight markers (GIBCO BRL) were as follows: 9.49, 7.46, 4.40, 2.37, 1.35, and 0.24 kb.

Transcriptome analysis.

The microarray slides contained 93 different PCR products of 300 to 500 bp in length, representing 70 genes of the central carbon and nitrogen metabolic pathways of *B. subtilis*. The PCR products correspond to the 3' end of each gene. Additionally, selected genes were also present on the slides with 300 to 500 bp of the 5' ends as a quality control. To avoid PCR artifacts due to unspecific priming, all PCR products were checked on an agarose gel and sequenced. The PCR products were purified using a QIAquick 96 PCR purification kit (QIAGEN, Hilden, Germany) on a BioRobot 8000 (QIAGEN, Hilden, Germany). The slides were produced by spotting the PCR products using a Lucidea (Amersham Biosciences, Freiburg, Germany) contact printer. Each gene was present with eight spots organized in two rows with four spots on the microarray slide. Each row originated from an independent PCR. For laser calibration, a custom-made scorecard, consisting of 10 PCR products of *Escherichia coli* and *Mycoplasma pneumoniae* genes (400 to 500 bp each), was used. These DNA fragments were spotted, and the absence of cross-hybridization with any of the labeled *B. subtilis* cDNAs was verified. Each spot contained 30 pg of probe DNA and was generated by placing 100 pl of a 50% aqueous dimethyl sulfoxide solution on amino-silane-coated Type 7* slides (Amersham Biosciences, Freiburg, Germany). The probes were cross-linked onto the slides by UV illumination (50 mJ/cm^2) after the total evaporation of the spotting solution. The arrays were stored in the dark in a vacuum desiccator.

cDNA labeling.

The isolated RNA was used to generate cDNAs labeled with fluorescent Cy-3 and Cy-5 dyes according to a direct labeling protocol (Hovey *et al.*, 2005) with the following modifications. Annealing was performed with 25 µg of total *B. subtilis* RNA, 40 mU of RNA Guard (Amersham Biosciences, Germany), in vitro transcribed RNA of the 10 *E. coli* and *M. pneumoniae* genes (ranging from 1.5 to 25 ng) with defined ratios for the Cy-3 and Cy-5 reaction mixes (1:10, 10:1, 1:5, 5:1, and 1:1), and 1 µl specific primer mix (including primers of the 10 scorecard genes). Each primer was present in the primer mix at a concentration of 5 pM. Primer annealing was performed in a thermocycler by asymptotic cooling from 70°C to 22°C within 30 min. Then, fivefold reaction buffer, 200 U Superscript III reverse transcriptase (Invitrogen GmbH, Karlsruhe, Germany), 2 µl dithiothreitol (0.1 M), 1 µl deoxynucleoside triphosphate mix (10 mM ATP, GTP, TTP, 3 mM CTP), and 1 µl Cy-3 or Cy-5 dyes (Amersham Biosciences, Germany) were added. Incubation was performed for 2 hours at 42°C. The labeled cDNA was purified with a CyScribe GFX purification kit (Amersham Biosciences, Germany). Finally, the amount of incorporated fluorescent nucleotides was determined by measuring the absorption of the purified labeled cDNA at 550 nm for Cy-3 and at 650 nm for Cy-5. In each hybridization, equal amounts of incorporated fluorescent dye, ranging from 80 to 150 pmol, were used.

Hybridization.

For hybridization, the fluorescently labeled cDNA was denatured for 2 min at 95°C followed by the addition of 50 µl microarray hybridization buffer (CyScribe first-strand cDNA labeling kit; Amersham Biosciences, Germany) and 110 µl formamide. Hybridization was carried out at 42°C overnight using an automatic sample processor (Lucidea SlidePro hybridization chamber; Amersham Biosciences, Germany) (Hovey *et al.*, 2005).

Microarray analysis: data normalization and evaluation.

The signal intensities from each spot on the array were collected using a GenePix 4000B scanner and GenPix Pro software version 4.0 from Axon Instruments (Union City, CA). During a prescan at lower resolution, approximate normalization using the known concentrations of the scorecard was performed. The scan for data acquisition was then done at a 10-µm/pixel resolution. The calculation of the expression ratios for all genes was performed using the GenPix Pro 4.0 software. Three mathematically different ratios were calculated: ratio of medians, ratio of means, and regression ratio. To ensure data quality, the following

two criteria were applied: (i) the signal intensities of both fluorescent dyes of a spot corrected for its calculated background had to be greater than the standard deviation of the background of the corresponding fluorescent dye, and (ii) the deviation of the expression ratio calculated by the three different methods had to be within a 30% interval. To generate two biologically independent experiments, RNA was extracted from two independent cultures grown in CG or CGSE minimal medium. One pair of RNA samples was additionally subjected to a dye-flip experiment. As each open reading frame was spotted eight times on each microarray slide (see above), 24 expression values were generated for evaluation if each passed the quality control criteria. From at least 12 of the 24 potential values for each open reading frame, the average value of the ratio of median was calculated, and the standard deviations are shown elsewhere (see the supplemental material).

Mass spectrometric ^{13}C labeling analysis.

Mass isotopomer fractions of amino acids from the cell protein were determined by gas chromatography-mass spectrometry (Kiefer *et al.*, 2004; Wittmann *et al.*, 2004). For this purpose, cells (about 1 mg dry cell mass) were harvested from the culture and washed twice with deionized water. The pellet was then incubated with 50 µl 6 M HCl for 24 h at 105°C, subsequently neutralized with 6 M NaOH, and cleared from insoluble matter by centrifugation (5 min; Ultrafree-MC filter units and 0.22-µm Durapore membrane; Millipore, Bedford, MA). The remaining clear solution was lyophilized. Analysis of the amino acids was performed after derivatization into the *t*-butyl-dimethylsilyl derivative (Krömer *et al.*, 2004; Wittmann *et al.*, 2002). All samples were first measured in scan mode to check for potential isobaric interference between analytes and other sample components. The labeling patterns of the amino acids were then determined in triplicate via selective ion monitoring of selected ion clusters, representing [M-57] fragments with the complete carbon skeletons of the amino acids. The mean experimental error for the mass isotopomer fractions was about 0.15%.

Metabolic modeling and parameter estimation.

The metabolic network for growth of *B. subtilis* on CG or CGSE comprised all central metabolic pathways, i.e., glycolysis, the pentose phosphate pathway, the Krebs cycle, and anaplerotic carboxylation. Additionally, the pathways for the formation of lactic acid, acetic acid, acetoine, and glycerol and the anabolic pathways from intermediary precursors into biomass were implemented. For glycine synthesis, two possible routes, i.e., via serine and via threonine aldolase, were considered (Simic *et al.*, 2002).

Table 3. Calculated and experimentally detected ^{13}C labeling patterns of amino acids in cell hydrolysates of *B. subtilis* cultivated on CG and CGSE

Measured Metabolite [a]	Result for cultivation on:					
	CG			CGSE		
	M_0	M_1	M_2	M_0	M_1	M_2
Alanine (*m/z* 260, C_{1-3})						
calc.	0.463	0.394	0.110	0.516	0.355	0.100
exp.	0.462	0.398	0.110	0.517	0.357	0.098
Valine (*m/z* 288, C_{1-5})						
calc.	0.280	0.416	0.212	0.345	0.404	0.183
exp.	0.279	0.419	0.212	0.344	0.405	0.183
Threonine (*m/z* 404, C_{1-4})						
calc.	0.305	0.388	0.206	0.560	0.273	0.123
exp.	0.305	0.385	0.206	0.560	0.273	0.122
Aspartate (*m/z* 418, C_{1-4})						
calc.	0.305	0.388	0.206	0.559	0.273	0.124
exp.	0.303	0.391	0.205	0.556	0.275	0.124
Glutamate (*m/z* 432, C_{1-5})						
calc.	0.205	0.365	0.264	0.629	0.231	0.107
exp.	0.204	0.364	0.264	0.629	0.231	0.107
Lysine (*m/z* 431, C_{1-6})						
calc.	0.192	0.357	0.271	0.376	0.366	0.173
exp.	0.198	0.352	0.267	0.374	0.367	0.174
Phenylalanine (*m/z* 336, C_{1-9})						
calc.	0.2396	0.3946	0.2476	0.2478	0.3847	0.2449
exp.	0.2389	0.3942	0.2477	0.2424	0.3943	0.2433
Tyrosine (*m/z* 466, C_{1-9})						
calc.	0.2063	0.3641	0.2616	0.2134	0.3564	0.2586
exp.	0.2071	0.3645	0.2621	0.2123	0.3630	0.2558

[a] calc., calculated values predicted by the solution of the mathematical model corresponding to the optimized set of fluxes; exp., experimental values obtained by gas chromatography-mass spectrometry analysis of *t*-butyl-dimethylsilyl-derivatized amino acids.

Calculation of the anabolic demands for the different precursors was based on literature data on the biomass composition of *B. subtilis*, previously determined as a function of growth rate (Dauner and Sauer, 2001; Dauner *et al.*, 2002). The exact demands for both studied conditions were interpolated from these literature data depending on the specific growth rate measured in the present work (Table 2).

Labeling data for proteinogenic amino acids and for glucose from all carbohydrates and the mean values of the stoichiometric data from two parallel cultivations were combined for a calculation of metabolic flux. The set of fluxes that gave minimum deviation between experimental ($M_{i,exp}$) and simulated ($M_{i,calc}$) mass isotopomer fractions was taken as the best estimate for the intracellular flux distribution. The network was overdetermined, so that a least-squares approach was possible. As the error criterion, a weighted sum of least squares was used (Wittmann and Heinzle, 2002). All metabolic simulations were carried out on a personal computer by use of Matlab 7.0 (Mathworks Inc., Natick, MA). Details on the computational tools are given elsewhere (Wittmann and Heinzle, 2001; Wittmann and Heinzle, 2002; Wittmann *et al.*, 2004). For the flux distributions, an excellent agreement between experimentally detected and calculated labeling patterns was achieved. The deviation between measured and calculated mass isotopomers was typically below 3% and thus rather small (Table 3) . Flux distributions, including deviations within intervals of 90% confidence, are shown in Fig. 2.1. The confidence intervals were calculated by a Monte Carlo approach as described previously (Wittmann and Heinzle, 2002).

Results and Discussion

Regulation of glycolysis.

The expression of glycolytic genes was assayed by microarray analysis. The promoter activities of several glycolytic genes and operons were additionally studied by measuring ß-galactosidase activities driven by these promoters. The transcriptome data suggested that the expression of glycolytic enzymes was only marginally affected by the addition of succinate and glutamate to the medium. These results are well supported by the comparison of the expression of *lacZ* fusions to promoters of glycolytic genes (Table 4). As judged from these expression data, the *ptsGHI* and *gapA* operons are highly transcribed under both conditions studied here, whereas the other promoters (*pgi*, *pfkA*, *fbaA*, and *pgk*) exhibit an intermediate activity (between 50 and 100 units of ß-galactosidase/mg of protein).

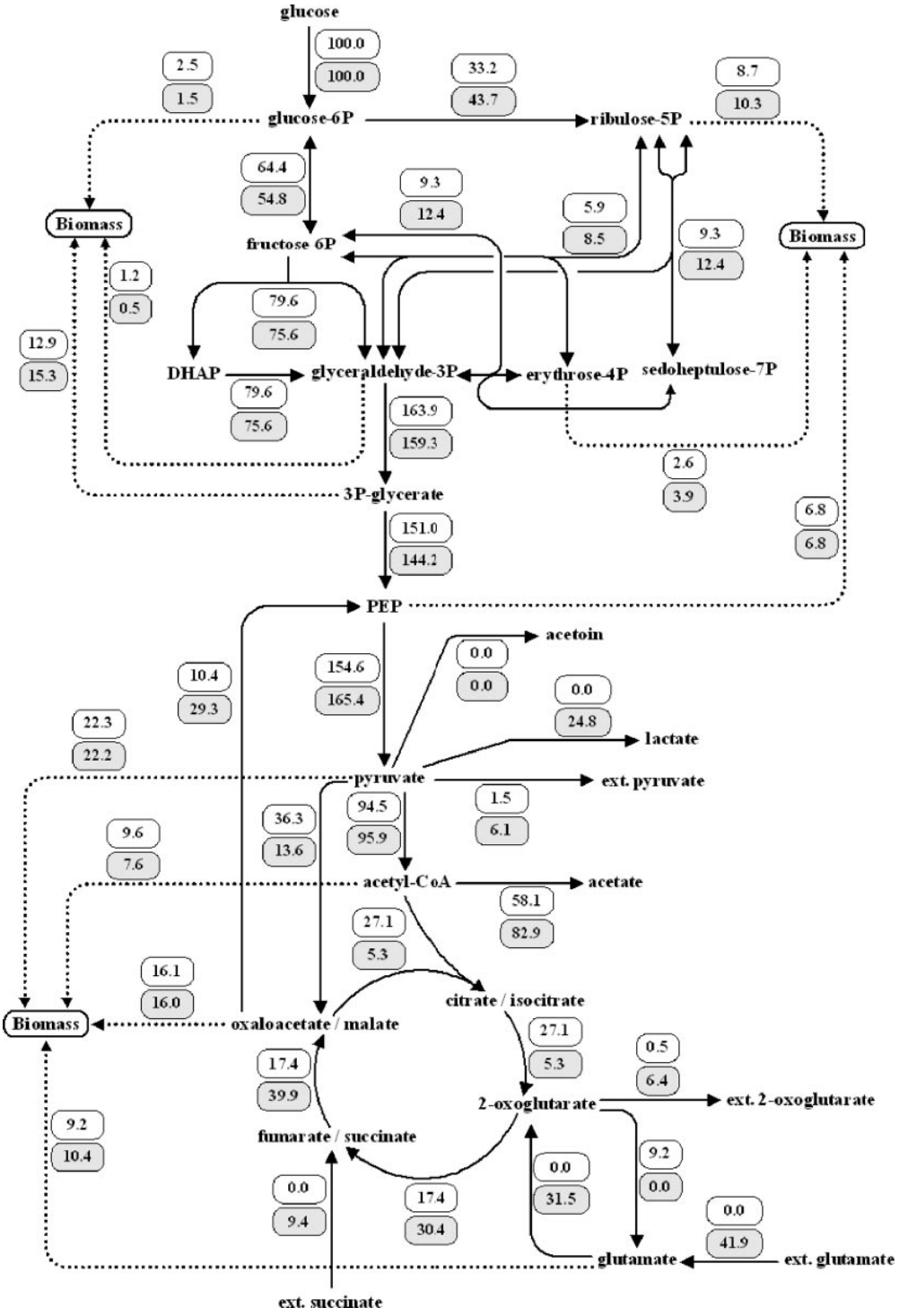

Figure 2.1. Intracellular carbon flux distribution of *B. subtilis* cultivated on 99% [1-^{13}C]glucose under aerobic conditions in CG medium (white squares) and CGSE medium (gray squares). All fluxes are given as relative fluxes normalized to the specific glucose uptake rate (Table 2) including deviation values for 90% confidence. The fluxes correspond to the optimal fit between experimentally determined ^{13}C labeling patterns of amino acids of the cell protein and of glucose from cell carbohydrates and ^{13}C labeling patterns simulated via isotopomer modeling. PEP, phosphoenolpyruvate; DHAP, dihydroxyacetone phosphate; ext., external.

Table 4. Expression of genes of central metabolism in *B. subtilis* as judged from transcriptional fusions to the *lacZ* gene[a]

Strain	Gene/ Operon	β-Galactosidase activity[b]		CGSE/CG ratio by:	
		CG	CGSE	*lacZ* assay	transcriptome analysis
QB7041	*ptsG*	603	649	1.07	0.99
GP314	*pgi*	56	53	0.94	1.06
GP315	*pfkA*	63	93	1.47	1.30
GP133	*fbaA*	95	83	0.88	1.20
GP313	*cggR* (*gapA*)	393	496	1.26	1.13
GP317	*pgk*	35	55	1.56	1.25
GP216	*pdhA*	110	145	1.32	1.69
BSIP1114	*pta*	50	220	4.30	1.67
GP205	*citB*	318	29	0.11	0.17
GP250	*nrgA*	2	1.8	0.90	0.98
GP268	*glnRA*	282	195	0.69	0.67
GP342	*gltA*	295	172	0.58	0.42
QB5556	*rocD*	12	12	1.00	1.17

[a] The regulation of the indicated genes was determined using the results of the reporter gene studies as well as those of the transcriptome analysis.

[b] Bacteria were grown in CG or CGSE minimal medium containing glucose. All strains contain transcriptional fusions of the promoter regions of the indicated genes to a promoterless lacZ gene. β-Galactosidase activities are given as units/mg of protein. Experiments were done in duplicate or triplicate. Representative results are shown.

These results are in good agreement with previous findings, which demonstrated induction of *ptsGHI*, *gapA*, and *pfkA-pyk* expression in the presence of glucose to levels similar to those observed in this study (Blencke *et al.*, 2003; Ludwig *et al.*, 2001).

The results of the transcription analysis demonstrate that the transcription of glycolytic genes and operons is unaffected by the presence of glutamate and succinate as long as glucose is present in the medium. This conclusion is supported and extended by the metabolic flux analysis: with the notable exception of the reaction catalyzed by phosphoglucoisomerase (Pgi), the flux of the carbon skeletons derived from glucose through glycolysis is not affected by the presence of the organic acids. Thus, the activity of the glycolytic enzymes remains constant in addition to their expression (with the exception of phosphoglucoisomerase [see below]).

The partition of glucose-6-phosphate and the pentose phosphate pathway.

Starting with glucose as the external substrate, glucose-6-phosphate is the first intracellular metabolite that is used by more than one enzyme. In CG minimal medium, glucose-6-phosphate is used by the glycolytic enzyme phosphoglucoisomerase (Pgi), and about one-third is catabolized to ribulose-5-phosphate due to the activity of glucose-6-phosphate dehydrogenase (Zwf) and the further enzymes of the pentose phosphate pathway. Finally, a small amount (about 2.5%) is used for the synthesis of biomass such as the cell wall. In the presence of glutamate and succinate, significantly less glucose-6-phosphate is converted to fructose-6-phosphate by the phosphoglucoisomerase than in the absence of the organic acids (54.8% compared to 64.4% relative to the total amount of glucose) (Fig. 2.1). Similarly, relatively less glucose-6-phosphate is used for biosynthetic purposes under these conditions. This might result from the availability of glutamate and succinate as precursors for many amino acids that have to be produced from the glucose in CG but not in CGSE medium. In contrast, the portion of glucose-6-phosphate that is oxidized by glucose-6-phosphate dehydrogenase is significantly increased (43.7% compared to 33.2%). This regulation of glucose-6-phosphate partition is not reflected by the transcriptional regulation of the genes encoding the two major enzymes that use glucose-6-phosphate as a substrate, i.e., *pgi* and *zwf*, which are constitutively expressed under the conditions used here. Previous studies indicated that the expression of these genes is also not affected by the presence of glucose in the medium (Blencke *et al.*, 2003; Ludwig *et al.*, 2001). A candidate that might cause the increased flux through the pentose phosphate pathway in the presence of succinate and glutamate is the *ywlF* gene, tentatively believed to encode ribulose-5-phosphate epimerase. The expression of this gene increased weakly but significantly upon the addition of the organic acids (Fig. 2.2). Moreover, both the transketolase (Tkt) and transaldolase (YwjH) were recently shown to be phosphorylated in vivo in *B. subtilis* (Lévine *et al.*, 2006). It is, however, unknown under which conditions these phosphorylations occur and how they affect the activities of the two enzymes. Due to the convergence of the pentose phosphate pathway with glycolysis at the fructose-6-phosphate and glyceraldehyde-3-phosphate levels, the changes in the partition of glucose-6-phosphate do not affect the fluxes through the lower part of glycolysis, which are essentially the same in both media (see above; Fig. 2.1).

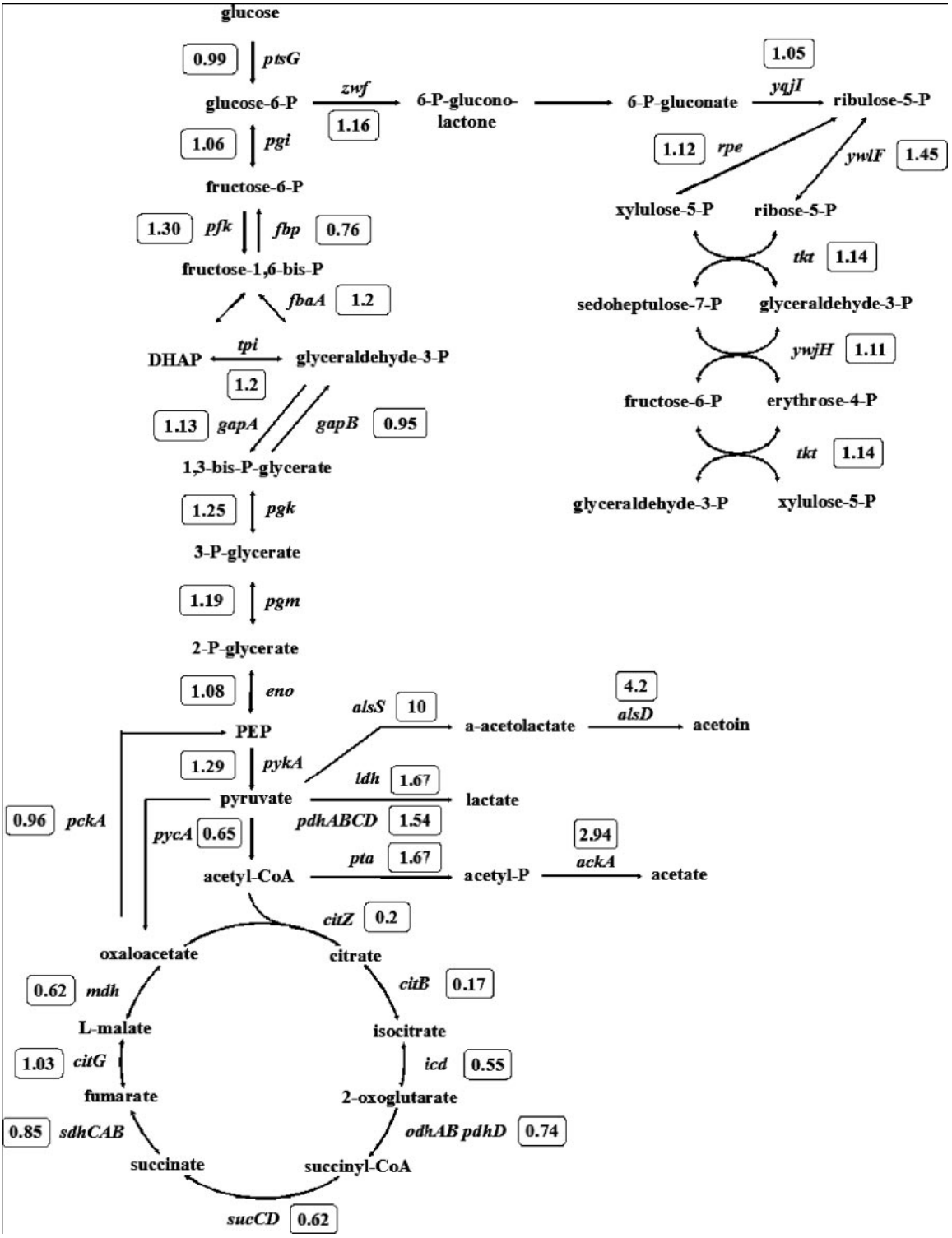

Figure 2.2. Overview on the regulation of genes encoding enzymes of glycolysis, the pentose phosphate shunt, and the Krebs cycle in *B. subtilis*. The result of the transcriptome analysis is given next to each gene. The numbers correspond to the ratios of gene expression under the two conditions (CGSE/CG). For multisubunit enzymes encoded by operons, the mean factors of regulation of the genes of the operon are shown. A complete list of all studied genes is given elsewhere (see the supplemental material). PEP, phosphoenolpyruvate; DHAP, dihydroxyacetone phosphate.

The partition of pyruvate and acetyl-CoA and the regulation of overflow metabolism. Pyruvate is among the most versatile metabolic intermediates. It can be oxidized to generate acetyl coenzyme A (acetyl-CoA) to enter the Krebs cycle (pyruvate dehydrogenase complex; *pdhABCD*), it can be converted to acetoin (acetolactate synthase and acetolactate decarboxylase; *alsS* and *alsD*, respectively), it can be reduced for the regeneration of NAD^+ (lactate dehydrogenase; *ldh*), and it can be used as an acceptor of carbon dioxide to replenish the Krebs cycle (pyruvate carboxylase; *pycA*). Moreover, the synthesis of several amino acids requires pyruvate as a precursor. Finally, pyruvate can be excreted into the medium.

The transcriptome data revealed that expression of the pyruvate dehydrogenase complex (*pdhABCD*) is weakly but significantly increased in the presence of the organic acids. This observation is supported by an assay of the ß-galactosidase activity driven by the *pdhA* promoter (Table 4). However, this increase is not paralleled by an increased formation of acetyl-CoA from pyruvate in the presence of glutamate and succinate (probably due to the high activity of lactate dehydrogenase under these conditions [see below]).

Acetyl-CoA can be transferred to oxaloacetate to initiate the Krebs cycle (citrate synthase [CitZ]) for substrate level phosphorylation (generation of acetate involving phosphotransacetylase Pta and acetate kinase AckA [see below]) or for biosynthetic purposes. In the presence of glucose, *B. subtilis* is known to excrete large amounts of acetate. Indeed, significant acetate production was observed for CG medium. This acetate production was even strongly increased if glutamate and succinate were present (CGSE medium) (Fig. 2.1). This increased acetate excretion corresponds well with the increased transcription of the *pta* and *ackA* genes, encoding phosphotransacetylase and acetate kinase, respectively, as detected in the transcriptome analysis. Finally, the *pta* promoter was found to be stimulated fourfold by the presence of glutamate and succinate (Table 4). In contrast, the flux of acetyl-CoA into the tricarboxylic acid branch of the Krebs cycle was severely reduced in the presence of glutamate and succinate (Fig. 2.1). The reduction of citrate formation from acetyl-CoA is probably due to the fivefold reduced expression of *citZ*, encoding the major citrate synthase of *B. subtilis* in CGSE medium. Since the citrate synthase works according to the Michaelis-Menten kinetics, the fivefold repression in *citZ* expression results in a fivefold reduced flux of acetyl-CoA into the Krebs cycle. This may reflect the lack of need of the Krebs cycle if glucose, as a source of carbon and energy, and glutamate and succinate, as Krebs cycle intermediates, are available for the cell.

As long as glucose is present in excess, *B. subtilis* produces acetate to benefit from the additional substrate-level phosphorylation. However, the accumulation of acetate results in (i)

the acidification of the medium (Tobisch *et al.*, 1999b) and (ii) the induction of the *alsSD* operon, encoding acetolactate synthase and acetolactate decarboxylase, and the subsequent synthesis of acetoin (Cruz Ramos *et al.*, 2000). It is not known whether the decrease of the pH or the accumulation of acetate triggers AlsR, the transcription activator of the *alsSD* operon. Our microarray study revealed that transcription of the *alsSD* operon is strongly induced by the presence of glutamate and succinate (10-fold induction of *alsS*) (Fig. 2.2). In this experiment, the pH value of the culture medium decreased from 7.17 to 7.01 on CG and from 7.33 to 7.04 on CGSE. The nearly constant pH of the culture medium suggests that the accumulation of acetate is the primary trigger of AlsR activity. However, no excretion of acetoin was observed under either condition. This may be due to the fact that acetoin might have just started to accumulate and thus been below the detection limit even if the genes were actively transcribed. Alternatively, the translation or the activity of acetolactate synthase and/or acetolactate decarboxylase might be controlled at an additional level. Indeed, the phosphorylation of acetolactate decarboxylase was recently reported (Lévine *et al.*, 2006). The expression of the *ldh* gene encoding lactate dehydrogenase is induced more weakly in CGSE medium than in CG medium. However, no excretion of lactate was detected in CG medium, whereas significant amounts of lactate were detected in the medium of the culture grown in the presence of glutamate and succinate.

If glucose is provided as the only carbon source (as in CG medium), the growth of *B. subtilis* depends on the activity of pyruvate carboxylase to replenish the Krebs cycle. Indeed, a significant portion of pyruvate is used for oxaloacetate production. In the presence of glutamate and succinate, the anaplerotic reactions are less important. This idea is supported by the finding that the flux of pyruvate to oxaloacetate is reduced about threefold in CGSE medium. This reduced flux is paralleled by a reduced expression of the *pycA* gene, which encodes pyruvate carboxylase, in CGSE medium relative to what is seen for CG medium. However, the regulation of *pycA* expression is much weaker than that of the flux through the reaction catalyzed by the pyruvate carboxylase. Thus, additional factors that control the activity of the enzyme may be involved. In both media, an important portion of pyruvate (about 12 to 15% of the total pyruvate pool) is used for the production of biomass. Finally, a small amount of pyruvate is excreted into the medium. Excretion is increased fourfold in CGSE medium; however, the proteins and gene products involved in pyruvate excretion have not yet been identified

Regulation of the Krebs cycle.

The Krebs cycle is central to the physiology of most heterotrophic organisms because it generates reducing power for respiration, precursors for anabolism, and an additional ATP in substrate-level phosphorylation. In *B. subtilis*, the expression of the genes encoding the initial enzymes of the Krebs cycle, citrate synthase and aconitase (*citZ* and *citB*, respectively), is synergistically repressed by glucose and glutamate/ammonium (Blencke *et al.*, 2006; Jourlin-Castelli *et al.*, 2000). Our transcriptome analysis confirmed this regulatory pattern: both genes were repressed fivefold in CGSE medium (Fig. 2.2). Moreover, the activity of the *citB* promoter was repressed about tenfold by the presence of the organic acids in addition to glucose (Table 4). This repression of initial enzymes of the Krebs cycle is paralleled by the 4.6-fold reduction of carbon flow from acetyl-CoA to 2-oxoglutarate in CGSE from the level seen for CG medium. In the presence of externally provided glutamate and succinate, these acids can be taken up and glutamate can be converted to 2-oxoglutarate. As a result, the flux from glutamate to oxaloacetate was higher in CGSE than in CG medium. The genes of the enzymes catalyzing the reactions from 2-oxoglutarate to oxaloacetate were weakly regulated by the presence of organic acids. This reflects the need for these enzymes under both conditions studied in this work.

The genes encoding the transporters for succinate and glutamate, *dctP*, *gltP*, and *gltT*, are only weakly affected by the presence of organic acids. It should be noted, however, that *dctP* and *gltP* are subject to a catabolite repression in the presence of glucose (Blencke *et al.*, 2003 and our unpublished results), suggesting that the two major transporters for succinate and glutamate are relatively weakly expressed under both conditions addressed in this study. In any case, significant amounts of glutamate and succinate were transported into the cells if these acids were available.

Regulation of nitrogen assimilation.

The basic C medium used in this study contained ammonium, which can be transported into the cell by the NrgA (AmtB) transport protein and assimilated by the glutamine synthetase/glutamate synthase cycle made up of glutamine synthetase and glutamate synthase. In the presence of glutamate, neither the ammonium transporter nor the glutamate synthase is required for growth. The determination of the *nrgA* promoter activity demonstrated that the gene was not expressed in the two media studied here (Table 4). This is due to the lack of transcription activation at high ammonium concentrations, which allow the entry of ammonia by diffusion (Detsch and Stülke, 2003). Similarly, the transcriptome data revealed that *nrgA*

expression was not affected by the presence of glutamate. The expression of *glnA* was slightly repressed in CGSE medium. This may result from the reduced need for glutamine synthesis if glutamate is present in the medium: under these conditions, glutamine synthesis is required only to meet the need for this amino acid, while it is an intermediate in glutamate biosynthesis in CG medium. The *gltAB* operon, encoding the two subunits of glutamate synthase, was repressed about twofold by the addition of glutamate and succinate, as judged from both transcriptome and promoter fusion data (Table 4). Accordingly, no conversion of 2-oxoglutarate to glutamate was detectable in CGSE medium. The imported glutamate was used for biomass production (about 25% of the glutamate), and the remaining 75% was converted to 2-oxoglutarate by the action of glutamate dehydrogenase encoded by *rocG*. Interestingly, *rocG* expression was somewhat lower in the presence of glutamate and succinate. However, it should be mentioned that only basal *rocG* expression occurs in both CG and CGSE media, since the expression of *rocG* requires induction by arginine and moreover is repressed by glucose (Belitsky and Sonenshein, 1998). Obviously, this low expression of glutamate dehydrogenase is sufficient to make a substantial contribution to feed the Krebs cycle in CGSE-grown bacteria. Thus, even though only weakly regulated at the level of transcription, the activities of the two enzymes interconverting 2-oxoglutarate and glutamate are mutually absolutely exclusive.

Conclusion.

The metabolism of *B. subtilis* has evolved to adapt to a broad range of conditions of nutrient supply. This adaptation often involves changes in the transcription of the relevant genes. However, the *ldh* gene, encoding lactate dehydrogenase, is constitutively transcribed under the conditions studied here, even though lactate formation in CGSE medium is strongly increased compared to that in CG medium. In this case, the enzymatic activity of lactate dehydrogenase might be controlled, perhaps by the intracellular redox state. This suggestion is in good agreement with the observation that the expression of the *ldh* gene is controlled by the redox state of the cell in the presence of nitrate (Reents *et al.*, 2006). On the other hand, some transcriptional regulation events are not paralleled by changes in the carbon fluxes. This was observed for the *alsSD* operon, required for acetoin biosynthesis. In this case, posttranslational control of AlsD activity by protein phosphorylation might account for an additional level of regulation. Phosphorylation is a common means to adjust protein activities to the metabolic requirements. However, only with the advance of proteomic methods can these modifications

be detected (Lévine *et al.*, 2006). The role of protein phosphorylation in the control of metabolic pathways is well documented in the case of *E. coli* isocitrate dehydrogenase, and this phosphorylation also occurs in *B. subtilis* (El-Mansi *et al.*, 2006; Lévine *et al.*, 2006). Moreover, the control of 2-oxoglutarate dehydrogenase activity by phosphorylation was recently discovered for *Corynebacterium glutamicum* (Niebisch *et al.*, 2006). Thus, much work remains to be done to study the details of the regulatory events that occur at the different stages after transcription involving changes in mRNA stability, differential translation, and posttranslational modifications. This work will be essential for the development of systems biology.

3 Multiple-mutation reaction: A method for the simultaneous introduction of multiple mutations into the *glpK* gene of *Mycoplasma pneumoniae*

The results described in this chapter were published in:

Hames, C., S. Halbedel, O. Schilling, and J. Stülke. 2005. Multiple-mutation reaction: a method for simultaneous introduction of multiple mutations into the *glpK* gene of *Mycoplasma pneumoniae*. *Appl Environ Microbiol* **71:**4097-4100.

Author contributions:

This study was designed by J.S., S.H. and O.S.. O.S. established the basic principles of the method. C.H. performed all experiments. S.H. designed the oligonucleotides and supervised the experimental work that was accomplished by C.H. during her diploma thesis. All authors were involved in the interpretation of the collected data and in the writing of the manuscript.

Abstract

In *Mycoplasma pneumoniae*, the UGA opal codon specifies tryptophan rather than a translation stop site. This often makes it difficult to express *Mycoplasma* proteins in *E. coli* isolates. In this work, we developed a strategy for the one-step introduction of several mutations. This method, the multiple-mutation reaction, is used to simultaneously replace nine opal codons in the *M. pneumoniae glpK* gene.

Introduction

Mycoplasma pneumoniae is a pathogen that lives on mucosal surfaces and causes diseases such as mild pneumonia, tracheobronchitis, and complications affecting the central nervous system, the skin, and mucosal surfaces (Jacobs, 1997; Lind, 1983; Waites and Talkington, 2004). This bacterium possesses one of the smallest genomes of any free-living organism known so far. This reduced genome makes *Mycoplasma* spp. interesting from two points of view: (i) the analysis of these bacteria may help to identify the minimal set of genes that is required for independent life (Hutchison III *et al.*, 1999), and (ii) *M. pneumoniae* and its close relative *M. genitalium* are well suited for the development of the methods of the postgenomic era (Jaffe *et al.*, 2004; Wasinger *et al.*, 2000). Another interesting aspect of the small genome is the observation that several enzymes of *Mycoplasma* spp. are "moonlighting"; i.e., they have multiple unrelated functions (Jeffery, 1999). This was discovered for glycolytic kinases, which are also active as nucleoside diphosphate kinases in *M. pneumoniae* and other *Mycoplasma* spp. (Pollack *et al.*, 2002).

However, the analysis of proteins from *Mycoplasma* spp. is hampered by a peculiarity of the genetic code of these bacteria: they use the UGA opal codon to incorporate tryptophan rather than as a stop codon as in the universal genetic code (Inamine *et al.*, 1990; Simoneau *et al.*, 1993). Thus, if cloned into *Escherichia coli* or other hosts, the genes from *M. pneumoniae* may contain many stop codons that prevent heterologous expression. Several strategies have been developed to solve this problem. For example, some *M. pneumoniae* genes, such as *ptsH* or *hprK*, do not possess UGA codons and thus require no special care (Steinhauer *et al.*, 2002). Expression of mollicute genes in *Spiroplasma* spp. that read the UGA as a tryptophan codon was reported, but these bacteria are difficult to handle (Stamburski *et al.*, 1991). *E. coli* suppressor strains expressing an opal suppressor tRNA were developed, but they fail if

multiple opal codons are present (Smiley and Minion, 1993). *M. pneumoniae* genes containing few UGA codons have been expressed in *Bacillus subtilis* with low efficiency (Kannan and Baseman, 2000). In cases with only a few opal codons, these were changed by site-directed mutagenesis to allow expression in *E. coli* (Knudtson *et al.*, 1997; Noh *et al.*, 2002). The *M. pneumoniae* P1 adhesin gene contains 21 opal codons, and a large-scale purification of the protein, though highly desired, has so far not been possible. In this case, protein fragments were expressed and purified (Chaudhry *et al.*, 2005). Finally, *Mycoplasma* genes could be synthesized *in vitro* from oligonucleotides; this strategy is, however, quite expensive. In this work, we present a PCR-based method that allows the simultaneous introduction of several mutations in a single step. Using this strategy, 9 of the 10 opal codons of the *glpK* gene from *M. pneumoniae* were modified, leading to expression of glycerol kinase in *E. coli*.

Outline of the multiple-mutation reaction (MMR) strategy.

Several methods for PCR-based site-directed mutagenesis have been developed. Among these, the combined chain reaction method (Bi and Stambrook, 1997; Bi and Stambrook, 1998) proved to be very rapid and reliable. The principle of this method is the use of mutagenic primers that hybridize more strongly to the template than the external primers. The mutagenic primers are phosphorylated at their 5' ends, and these are ligated to the 3' OH groups of the extended upstream primers by the action of a thermostable DNA ligase. Moreover, the DNA polymerase employed must not exhibit 5' → 3' exonuclease activity, to prevent the degradation of the extended primers. In our view, *Pfu* and *Pwo* polymerases are both well suited (Meinken *et al.*, 2003; Schilling *et al.*, 2004). The original protocol describes the introduction of two mutations simultaneously. In a previous study, we used a combined chain reaction to mutagenize four distant bases in a DNA fragment in a one-step reaction (our unpublished results).

For the introduction of up to nine mutations in a single experiment, we developed the MMR. This method requires the efficient binding of all the mutagenic primers to the target DNA. To ensure that extension of a PCR product is not possible beyond the next (i.e., more downstream) mutation site without ligation to the corresponding mutagenic primer, special care needs to be taken in primer design. This reaction is based on an accurate calculation of melting temperatures. For this purpose, the formula T_m (melting temperature in °C) = 81.5 + 16.6(log[Me+]) + 0.41 x %G+C – (500/oligonucleotide length) – 0.61 x % formamide was used (Meinkoth and Wahl, 1984). Only bases that match between primer and template were

used for the calculation. One consideration was made when designing the mutagenic primers: ligation was facilitated by placing a G or C at the 5' end of the oligonucleotide to favor close duplex formation between the primer and the target DNA.

Table 5. Oligonucleotides used in this study

Oligonucleotide	Sequence (5'-3')[a]	Mutation	T_m (°C)
CH7	AAAAGTCGACATGGATCTAAAACAACAATAC ATTCTTG	None	59
CH8	TATAAAGCTTGTCTTAGTCTAAGCTAGCCCAT TTTAG	A1512G	63
CH9	AAAAGTCGACATGGATCTAAAACAAC	None	57
CH10	TATAAAGCTTGTCTTAGTCTAAGCTAG	None	59
CH11	P-GATCCCTTAGAAATTTGGTCAGTCCAATTAG	A165G	64
CH12	P-CCATTGTGTTATGGAACAAAGAAAATGGTT TG	A273G	63
CH13	P-CACTAAGATTGCTTGGATCTTGGAAAATGTT C	A438G	64
CH14	P-CCTGGTTAATTTGGAAACTAACGGGTG	A522G	64
CH15	P-CCATGACATGGTCACAAGAGTTAGGC	A606G	65
CH16	P-TACCGAGTCATTGGTCTACTAGTGC	A705G	65
CH17	P-CCTTAAAGTGGTTAAGGGATAGTCTTAAGG	A966G	65
CH18	P-GCAGTTAATTATTGGAAGGACACTAAACAA C	A1386G	63
CH19	P-GAAATCAAAGCGTTGGAACGAAGCTG	A1482G	65

[a] The "P" at the 5' end of oligonucleotide sequences indicates phosphorylation.

The external primers were selected to have melting temperatures considerably lower (about 4°C) than those of the mutagenic primers. The MMR was performed with 2.5 units of *Pfu* DNA polymerase (MBI Fermentas, Lithuania) and 15 units of Ampligase® (Epicentre, Madison, WI) in MMR buffer (20 mM Tris-HCl [pH 8.5], 3 mM $MgCl_2$, 50 mM KCl, 0.4 mg/ml bovine serum albumin, and 0.5 mM NAD^+) in a total volume of 50 µl. Conditions for MMR included denaturation at 95°C for 30 s, primer annealing at 57°C for 30 s, and elongation at 65°C for 6 min, for 35 cycles. Initially, the DNA fragment (100 ng) was denatured for 5 min at 95°C. Ten picomoles of each primer were used. The sequences and the arrangements of the oligonucleotides used in this study are shown in Table 5 and Figure 3.1, respectively.

Cloning of *M. pneumoniae glpK* and expression of the protein in *E. coli*.

An analysis of growth behavior and the *in vivo* protein phosphorylation pattern identified glycerol as a key carbon source associated with regulatory phenomena. This substrate triggered *in vivo* phosphorylation of the HPr protein of the phosphotransferase system by the metabolite-sensitive HPr kinase/phosphorylase (Halbedel *et al.*, 2004; Steinhauer *et al.*, 2002). We were therefore interested in studying glycerol metabolism and its regulation in more detail. As a first step, we intended to purify the glycerol kinase. This enzyme is known to be a key target of catabolite regulation in Gram-positive bacteria (Darbon *et al.*, 2002; Stülke and Hillen, 2000). However, the corresponding *glpK* gene contains 10 opal codons and was therefore a good subject for MMR in order to change these codons to tryptophan codons for *E. coli*. The *glpK* gene was amplified using the oligonucleotides CH7 and CH8 and chromosomal DNA of *M. pneumoniae* M129 (ATCC 29342) as a template. With CH8, the most C-terminal opal codon was replaced by a TGG codon. The amplicon was cloned between the *Sal*I and *Hind*III sites of the expression vector pWH844 (Schirmer *et al.*, 1997). The resulting plasmid, pGP253, was used as a template for MMR with CH9 and CH10 as external primers and CH11 through CH19 as mutagenesis primers. Five independent MMRs were carried out, and the MMR products were individually cloned as a *Sal*I/*Hind*III fragment into pWH844. The inserts of one clone resulting from each MMR were sequenced. Out of the five candidates, three contained the nine desired mutations without any additional mutations. One plasmid contained seven out of nine mutations, and the fifth plasmid bore all nine mutations and one additional undesired 1-bp deletion in one of the primer regions.

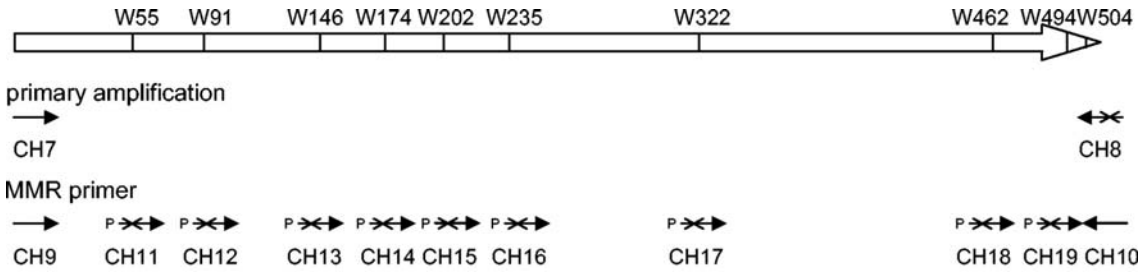

Figure 3.1. Strategy for amplification and mutagenesis of the *M. pneumoniae glpK* gene (MPN050 (Himmelreich *et al.*, 1996)). The positions of the opal codons in the wild-type *glpK* gene (indicated by a W followed by the number corresponding to the amino acid) and the position and orientation of the external and mutagenic oligonucleotides are shown. The annealing site of each oligonucleotide is indicated by an arrow. Oligonucleotides bearing an A→G transition are depicted by crossed arrows.

Plasmids bearing all nine desired mutations but no additional mutations were designated pGP254. pGP254 allows the expression of *M. pneumoniae* glycerol kinase fused to an N-terminal hexahistidine sequence under the control of an isopropyl-ß-D-thiogalactopyranoside (IPTG)-inducible promoter. To test the success of the mutagenesis, we compared the protein contents of *E. coli* cultures carrying either pWH844, pGP253, or pGP254.

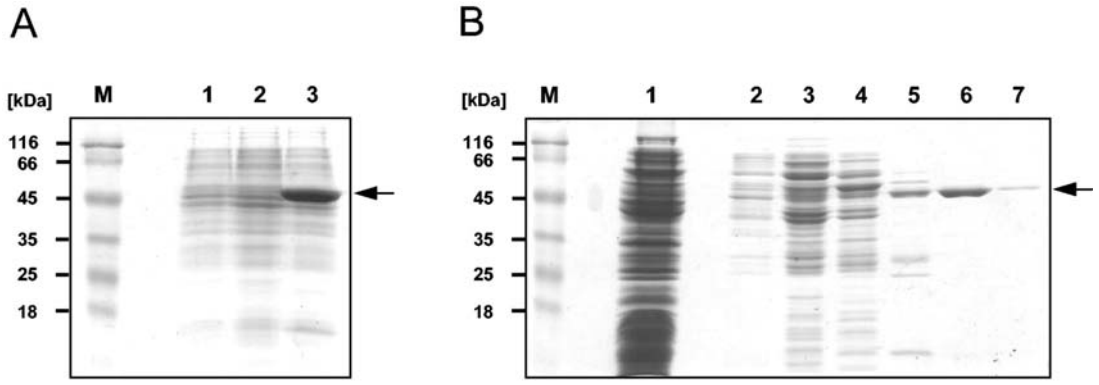

Figure 3.2. Overproduction and purification of *M. pneumoniae* GlpK. (A) Sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis for the detection of His_6-tagged GlpK in crude extracts of *E. coli* DH5α bearing either the empty expression vector pWH844 (lane 1), the expression vector including the wild-type *glpK* allele, pGP253 (lane 2), or the vector including the mutated *glpK* allele, pGP254 (lane 3). Cells were grown to an optical density at 600 nm of 0.8, and expression from the IPTG-inducible promoter was induced by addition of 1 mM IPTG (final concentration). After 2 h, cells were harvested and disrupted by sonication. The insoluble fraction was pelleted in a centrifugation step and solubilized using 6 M urea, and sample aliquots were separated on an SDS-12% polyacrylamide gel. (B) SDS-polyacrylamide gel electrophoresis to monitor the purification of His_6-tagged GlpK. Crude extract of the GlpK expression strain (*E. coli* DH5α bearing the plasmid pGP254) that had been grown in the presence of 1 mM IPTG was passed over a Ni^{2+}-nitrilotriacetic acid superflow column (5 ml bed volume; QIAGEN) and washed extensively with a buffer containing 10 mM Tris-HCl (pH 7.4) and 200 mM NaCl, followed by elution with an imidazole gradient (from 10 to 500 mM imidazole). Aliquots of the individual fractions were separated on SDS-12% polyacrylamide gels. A prestained protein molecular mass marker (Fermentas) served as a standard (lane M). Lane 1, flowthrough; lane 2, 10 mM imidazole; lane 3, 20 mM imidazole; lane 4, 50 mM imidazole; lane 5, 100 mM imidazole; lane 6, 200 mM imidazole; lane 7, 500 mM imidazole.

A prominent band corresponding to an approximate molecular mass of 56 kDa is detectable in the strain bearing pGP254, while no such protein is expressed from pGP253 encoding the unmutated *glpK* gene (see Fig. 3.2A). The glycerol kinase was purified to apparent homogeneity by Ni^{2+}-nitrilotriacetic acid chromatography as described previously (see Fig. 3.2B) (Meinken *et al.*, 2003). Thus, MMR was successful in achieving efficient overproduction of *M. pneumoniae* glycerol kinase for biochemical studies.

This study demonstrates that MMR can be used for the rapid and highly efficient introduction of multiple mutations into a gene. Out of five individual clones, four had the desired mutations. Of these four, only one candidate contained an extra mutation, which was most probably due to an impure oligonucleotide mix. Indeed, other experiments indicated that the quality of the oligonucleotides is the limiting factor for MMR. Obviously, this method is useful not only for the expression of *Mycoplasma* species genes, but also to change codon usage patterns or for any other purpose that requires the introduction of many mutations or combinations of mutations at the same time. What is the maximum number of mutations that can be introduced by MMR in a single step? Our results suggest that the target of nine mutations is still far from a theoretical limit, and we are confident that this method can be made even more effective by taking care of the quality of the oligonucleotides (see above) and by using mutagenic primers that alternate between the two strands of the DNA. With this method at hand, even the expression of a functional P1 adhesin gene in *E. coli*, which has so far been beyond imagination (Chaudhry *et al.*, 2005), now seems feasible.

4 A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity

The results described in this chapter were published in:

Schilling, O., I. Langbein, M. Müller, M. H. Schmalisch, and J. Stülke. 2004. A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res* **32:**2853-2864.

Author contributions:

This study was designed by O.S., I.L., M.H.S., and J.S.. O.S. performed and interpreted all reporter gene analyses and most of the gel shift experiments. O.S. constructed all mutant strains from GP366 to GP392. I.L. perfomed the RNase T1 footprint analysis and one of the gel shift experiments. I.L. constructed the mutant strains from GP150 to GP191. M.M. constructed the GlcT-RBD overexpression vector pGP230. M.H.S. supervised O.S. in his diploma work. O.S. and J.S. wrote the paper.

Abstract

The Gram-positive soil bacterium *Bacillus subtilis* transports glucose by the phosphotransferase system. The genes for this system are encoded in the *ptsGHI* operon. The expression of this operon is controlled at the level of transcript elongation by a protein-dependent riboswitch. In the absence of glucose a transcriptional terminator prevents elongation into the structural genes. In the presence of glucose, the GlcT protein is activated and binds and stabilizes an alternative RNA structure that overlaps the terminator and prevents termination. In this work, we have studied the structural and sequence requirements for the two mutually exclusive RNA structures, the terminator and the RNA antiterminator (the RAT sequence). In both cases, the structure seems to be more important than the actual sequence. The number of paired and unpaired bases in the RAT sequence is essential for recognition by the antiterminator protein GlcT. In contrast, mutations of individual bases are well tolerated as long as the general structure of the RAT is not impaired. The introduction of one additional base in the RAT changed its structure and resulted in complete loss of interaction with GlcT. In contrast, this mutant RAT was efficiently recognized by a different *B. subtilis* antitermination protein, LicT.

Introduction

In bacteria, gene expression is most commonly regulated at the level of transcription initiation. This control is achieved by interactions of specific DNA sequences with regulatory proteins. A second mechanism of control of gene expression is mediated at the RNA level by riboswitches. This type of control may affect either transcription or translation. The formation of riboswitches can be triggered by low molecular weight effectors, tRNAs or by regulatory proteins (for reviews see Henkin, 2000; Stülke, 2002; Winkler and Breaker, 2003).

Regulatory protein–RNA interactions control the expression of genes involved in diverse physiological functions. In *Bacillus subtilis* and other Gram-positive bacteria, genes encoding enzymes for the biosyntheses of tryptophan and pyrimidines are controlled by a termination/antitermination mechanism (Babitzke and Gollnick, 2001; Switzer *et al.*, 1999). Genes required for the utilization of several sugars are subject to regulation by antitermination in many bacteria (Rutberg, 1997). Similarly, the utilization of aliphatic amides in *Pseudomonas aeruguniosa* and assimilatory nitrate reduction in *Klebsiella oxytoca* is

controlled by RNA-binding antitermination proteins (Chai and Stewart, 1998; Wilson *et al.*, 1996). Recently, an antitermination mechanism that controls gene expression after a cold shock in *Escherichia coli* was discovered (Bae *et al.*, 2000).

As with all other riboswitches, protein-dependent antitermination/termination systems are based on the existence of alternative, mutually exclusive RNA structures. One of these structures is a transcriptional terminator. The alternative structure, also called the RNA antiterminator (RAT) (Aymerich and Steinmetz, 1992), prevents formation of the terminator and allows transcription elongation to proceed. Depending on the nature of the system, one of the structures is energetically favored and can form in the absence of any other factor. In contrast, the less favored structure depends on stabilization, which is caused by binding of the regulatory protein. While the antiterminators are more stable in the anabolic systems resulting in transcription elongation as the default state, formation of the terminators is energetically favored in many catabolic systems. Thus, the respective genes are not expressed unless the inducer is present and binding of the antiterminator protein allows formation of the antitermination structure (Aymerich and Steinmetz, 1992; Henkin, 2000; Stülke, 2002).

We are interested in the control of glucose utilization in *B. subtilis*. Glucose is the preferred source of carbon and energy for these bacteria. The sugar is transported by the bacterial sugar:phosphoenolpyruvate phosphotransferase system (PTS) and is subsequently catabolized in the glycolytic and pentose phosphate pathways (for a review see Stülke and Hillen, 2000). Two operons encoding enzymes of glucose catabolism are inducible by the presence of glucose in the medium, the *ptsGHI* operon and the *gapA* operon coding for the components of the PTS and for glycolytic enzymes of triose phosphate interconversions, respectively (Ludwig *et al.*, 2001; Stülke *et al.*, 1997; Tobisch *et al.*, 1999b). While induction of the *gapA* operon is governed at the level of transcription initiation by a specific repressor (Doan and Aymerich, 2003; Fillinger *et al.*, 2000; Ludwig *et al.*, 2001), the *ptsGHI* operon is controlled by transcriptional antitermination (Stülke *et al.*, 1997). Transcription of this operon is constitutively initiated but stops at a factor-independent terminator upstream of the *ptsG* structural gene. In the presence of glucose, an antiterminator protein, GlcT, is activated and is thought to bind a RAT sequence overlapping the transcription terminator (see Fig. 4.1). GlcT binding to the RAT prevents formation of the terminator and allows transcription elongation (Langbein *et al.*, 1999). The RNA-binding activity of GlcT is controlled by reversible phosphorylation in response to the presence of the inducer, glucose.

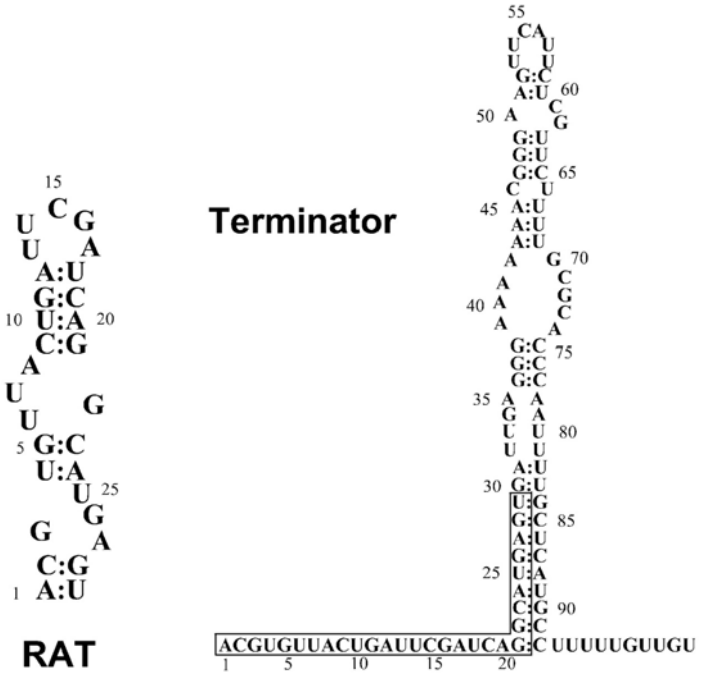

Figure 4.1. Proposed schematic model of the antiterminator and terminator structures of the *ptsG* leader mRNA. The secondary structure of the antiterminator is based on known RAT structures (Yang *et al.*, 2002). The RAT sequence is boxed in the terminator structure to highlight the overlap between the RAT and the terminator that constitutes the riboswitch.

In the presence of glucose, the glucose-specific enzyme II of the PTS may transfer its phosphate group to the sugar. In contrast, enzyme II is permanently phosphorylated in the absence of glucose and this form may phosphorylate and thereby inactivate the antiterminator protein GlcT (Bachem and Stülke, 1998; Schmalisch *et al.*, 2003).

GlcT is a member of a family of transcriptional antiterminators, the prototypes of which are BglG from *E. coli* and SacY and LicT from *B. subtilis*. These proteins are all composed of an N-terminal RNA-binding domain (Langbein *et al.*, 1999; Manival *et al.*, 1997) and two reiterated PTS regulation domains (PRDs) which control the RNA-binding activity of the antiterminators via phosphorylation-dependent dimerization (Amster-Choder and Wright, 1992; Fux *et al.*, 2003; Stülke *et al.*, 1998; Tortosa *et al.*, 1997). The RNA sequences bound by these antiterminators are all very similar to each other. However, while antiterminators such as SacT and SacY from *B. subtilis* have a relaxed specificity, i.e. they bind RAT sequences of the *bgl* type in addition to their cognate targets, binding of the *B. subtilis* LicT and *E. coli* BglG antiterminators is restricted to their cognate *bgl* RAT sequences (Manival *et al.*, 1997; Schnetz *et al.*, 1996). The structures of the RNA-binding domains of SacY and LicT were determined and revealed a novel fold for RNA-binding proteins (Declerck *et al.*, 1999; Manival *et al.*, 1997; van Tilbeurgh *et al.*, 1997). Until the

recent determination of the solution structure of the LicT–RNA complex, the question of how a symmetric dimeric protein would recognize an apparently asymmetric RNA has remained enigmatic. It now turned out that the *bglP* RAT sequence adopts a quasi-symmetric structure which allows binding of the protein dimer (Yang *et al.*, 2002). In contrast, the RAT sequence recognized by GlcT is only distantly related to the *bgl*- and *sac*-type RAT sequences and no cross-talk was observed (Langbein *et al.*, 1999). Homologs of GlcT and the respective RAT sequence were identified in *Staphylococcus carnosus* (Knezevic *et al.*, 2000).

In this work, we studied the interaction of the RNA-binding domain of GlcT with its cognate RAT target. Several mutations in the RAT sequence were found to affect the antitermination efficiency *in vivo*. Mutations that restored expression of the *ptsGHI* operon in the presence of a non-functional RAT all destroyed the transcriptional terminator in the *ptsGHI* leader mRNA. The sites of interaction between the RNA-binding domain of GlcT and its RAT were studied by *in vitro* footprinting and mutagenesis. As predicted by Yang *et al.* (Yang *et al.*, 2002), our results suggest that the RAT sequence may form a stem–loop structure with a nearly perfect symmetry, which is the recognition site for GlcT. Moreover, we identified a structural specificity determinant which decides the interaction partner.

Materials and Methods

Bacterial strains and growth conditions

The *B. subtilis* strains used in this study are shown in Table 6. Strains used in the random and site-directed mutagenesis studies are listed in Tables 7 and 8. All *B. subtilis* strains are derivatives of the wild-type strain 168. *Escherichia coli* DH5α (Sambrook *et al.*, 1989) was used for cloning experiments and for expression of recombinant proteins.

Bacillus subtilis was grown in SP medium or in CSE minimal medium (Faires *et al.*, 1999). The media were supplemented with auxotrophic requirements (at 50 mg/l) and carbon sources and inducers as indicated. *Escherichia coli* was grown in LB medium and transformants were selected on plates containing ampicillin (100 µg/ml). LB and SP plates were prepared by the addition of 17 g/l Bacto agar (Difco) to LB or SP medium, respectively.

Table 6. *B. subtilis* strains used in this study

Strain	Genotype	Source[a]
168	*trpC2*	Laboratory collection
BGW10	*trpC2 lys-3* Δ*licTS*::*erm*	Krüger and Hecker, 1995
GM1010	*sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	Le Coq *et al.*, 1995
QB5448	*trpC2 amyE*::*(ΔLA ptsG'-'lacZ aphA3)*	Stülke *et al.*, 1997
GP109	*trpC2* Δ*glcT8 amyE*::(ΔLA *ptsG*'-'*lacZ aphA3*)	Bachem and Stülke, 1998
GP150	*trpC2* Δ*glcT8 amyE*::('*lacZ cat*)	Langbein *et al.*, 1999
GP151	*trpC2 amyE*::($\Delta LA_{C75G\ C76G}$ *ptsG*'-'*lacZ aphA3*)	Langbein *et al.*, 1999
GP169	*trpC2 amyE*::($\Delta LA_{C75G\ C76G\ G37C\ G38C}$ *ptsG*'-'*lacZ aphA3*)	pGP332→168
GP172	*trpC2 amyE*::($\Delta CA_{G11A\ G26A}$ *ptsG*'-'*aphA3*-'*lacZ cat*)	pGP339→168
GP173	*trpC2 amyE*::($\Delta CA_{G11C\ G26A}$ *ptsG*'-'*aphA3*-'*lacZ cat*)	pGP340→168
GP174	*trpC2 amyE*::(ΔCA *ptsG*'-'*aphA3*-'*lacZ cat*)	pGP342→168
GP385	*trpC2 amyE*::($\Delta LA_{A\text{-}ins}$ *ptsG*'-'*lacZ aphA3*)	pGP549→168
GP386	*trpC2 amyE*::($\Delta LA_{A\text{-}ins}$ *ptsG*'-'*lacZ aphA3*) Δ*glcT8*	pGP549→GP150
GP387	*amyE*::(ΔLA *ptsG*'-'*lacZ aphA3*) *sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	pGP66→GM1010
GP388	*amyE*::($\Delta LA_{A\text{-}ins}$ *ptsG*'-'*lacZ aphA3*) *sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	pGP549→GM1010
GP389	*trpC2 lys-3 amyE*::(ΔLA *ptsG*'-'*lacZ aphA3*) Δ*licTS*::*erm*	pGP66→BGW10
GP390	*trpC2 lys-3 amyE*::($\Delta LA_{A\text{-}ins}$ *ptsG*'-'*lacZ aphA3*) Δ*licTS*::*erm*	pGP549→BGW10

[a] Arrows indicate construction by transformation.

DNA manipulation

Transformation of *E. coli* and plasmid DNA extraction were performed using standard procedures (Sambrook *et al.*, 1989). Restriction enzymes, T4 DNA ligase and DNA polymerases were used as recommended by the manufacturers. DNA fragments were purified from agarose gels using a Nucleospin extract kit (Macherey and Nagel). *Pfu* DNA polymerase was used for PCR, as recommended by the manufacturer. The combined chain reaction was performed with *Pfu* DNA polymerase and thermostable DNA ligase (Ampligase®; Epicentre, Madison, WI). DNA sequences were determined using the dideoxy chain termination method (Sambrook *et al.*, 1989). Chromosomal DNA of *B. subtilis* was isolated as described (Kunst and Rapoport, 1995).

Site-directed mutagenesis of the *ptsG* terminator and RAT sequences

Translational fusions of mutant variants of the *ptsG* regulatory region with the *lacZ* gene were constructed using the vector pAC7 (Weinrauch *et al.*, 1991) containing the kanamycin resistance gene *aphA3*. The plasmid harbors a *lacZ* gene without a promoter located between two fragments of the *B. subtilis amyE* gene. To study the effect of point mutations in the conditional terminator and the RAT sequence preceding the *ptsG* structural gene the following strategy was applied. A DNA fragment carrying the mutant form of the terminator or *ptsG*-RAT was constructed by site-directed mutagenesis using PCR-based approaches as outlined previously (Bachem *et al.*, 1997; Bi and Stambrook, 1998). Plasmid pGP66 (Stülke *et al.*, 1997) containing the *ptsG* promoter region served as a template. Multiple mutations were inserted using plasmids containing one or two mutations as templates. The mutagenic primers and the resulting plasmids are available upon request. The oligonucleotides JS11 (Stülke *et al.*, 1997) and IL5 (Langbein *et al.*, 1999) were used as outer primers. The final PCR products were purified and cut at the *BamH*I and *Mfe*I sites introduced by the PCR primers. To introduce the constructed *lacZ* fusions into the chromosome of *B. subtilis*, competent cells of the wild-type strain 168 were transformed with plasmids carrying the respective mutations linearized with *Sca*I.

Transformation and characterization of phenotype

Bacillus subtilis was transformed with plasmid DNA according to the two-step protocol described previously (Kunst and Rapoport, 1995). Transformants were selected on SP plates containing chloramphenicol (5 µg/ml) or kanamycin (5 µg/ml). In *B. subtilis*, amylase activity was detected after growth on SP medium supplemented with 10 g/l hydrolyzed starch (Connaught). Starch degradation was detected by sublimating iodine onto the plates.

Quantitative studies of *lacZ* expression in *B. subtilis* in liquid medium were performed as follows. Cells were grown in CSE medium supplemented with the carbon sources indicated. Cells were harvested at OD_{600} = 0.6–0.8. Cell extracts were obtained by treatment with lysozyme and DNase. ß-Galactosidase activity was determined as previously described using *o*-nitrophenyl-galactoside as substrate (Kunst and Rapoport, 1995). One unit is defined as the amount of enzyme which produces 1 nmol *o*-nitrophenol per min at 28°C.

Protein purification

To purify the RNA-binding domain of GlcT as a native protein, plasmid pGP230 was constructed as follows. Plasmid pGP114 (Langbein *et al.*, 1999) containing the DNA

fragment corresponding to amino acids 1–60 of GlcT fused to an N-terminal hexahistidine sequence was linearized with *BamH*I. The oligonucleotides MM6 (5'-GATCTCTGGTTCCGCGTGGTTCCATGA) and MM7 (5'-GATCTCATGGAACC ACGCGGAACCAGA) carrying a DNA fragment encoding a thrombin cleavage site were hybridized at 80°C and ligated to the linearized plasmid pGP114. Positive candidates were verified by sequencing.

Escherichia coli DH5α was used as host for overexpression of the recombinant proteins. Expression was induced by the addition of IPTG (final concentration 1 mM) to logarithmically growing cultures (OD_{600} = 0.8). The crude extracts were passed over Ni–NTA Superflow (Qiagen), followed by elution with an imidazole gradient. The Bio-Rad dye-binding assay was used to determine protein concentration. Bovine serum albumin was used as the standard. Purified His::GlcT-RBD protein was concentrated using a centriprep concentrator unit (Millipore). The GlcT-RBD peptidic fragment (7498 Da) was generated by thrombin cleavage (Pharmacia) of the His::GlcT-RBD protein according to the supplier's instructions (Pharmacia). After thrombin cleavage the protein preparation was loaded onto a Superdex 75 prep grade HR16/60 column (Pharmacia) for size exclusion chromatography in 300 mM NaCl, 50 mM Na_2HPO_4 (pH 7.8). The purity of the protein was determined on 15% Tris–Tricine gels using Coomassie brilliant blue staining (Schägger and von Jagow, 1987).

***In vitro* transcription**

To obtain a template for *in vitro* synthesis of the wild-type RAT RNA, a 84 bp PCR product was generated using pGP66 (Stülke *et al.*, 1997) as template and primers IL59 (5'-CCAAGTAATACGACTCACTATAGGACGTGTTACTGATTCG) and IL60 (5'-CAAGAATTGGGACAACTCTTCTTCTCCTTTTTTTTCCTCAATCACTCATGCC). To generate a template for *in vitro* transcription of mutant RAT RNAs, the primers OS25 (5'-CCAAGTAATACGACTCACTATAGGAATTCAGTTTATCCTTAT) and OS26 (5'-TTGAGGGAAAAAA ACGGGAAGTTC) were used to amplify a 99 bp PCR product. The presence of a T7 RNA polymerase recognition site in primers IL59 and OS25 (underlined) allowed the use of the PCR product as a template for *in vitro* transcription with T7 RNA polymerase (Roche Diagnostics). As a non-specific RNA, a 350 bp *gapA* transcript was prepared as described previously. The integrity of the RNA transcripts was analyzed by denaturating agarose gel electrophoresis (Ludwig *et al.*, 2001).

Assay of interaction between GlcT-RBD and RAT RNA

Binding of GlcT-RBD to RAT RNA was analyzed by gel retardation experiments. The RAT RNA (in water) was denatured by incubation at 90°C for 2 min and renatured by dilution 1:1 with ice-cold water and subsequent incubation on ice. If required, non-specific RNA was renatured separately and mixed with the RAT RNA prior to protein addition. Purified GlcT-RBD was added to the RAT RNA and the samples were incubated for 10 min at room temperature. After this incubation, glycerol was added to a final concentration of 10% (wt/vol). The samples were then analyzed on 12% Tris–acetate polyacrylamide gels.

RNase T1 footprinting

The method used for RNase T1 digestion of the RAT RNA was as follows. Radiolabeled RNA was denatured by heating to 90°C and was subsequently allowed to cool down at room temperature. Diluted GlcT-RBD (7 pmol in 1 µl) was added to 5'-radiolabeled RAT RNA (140 000 c.p.m., gel purified) in the presence of 20 mM $MgCl_2$ and 300 mM NaCl (final concentration). The reaction was incubated at 37°C for 10 min, after which 20 µg yeast tRNA (Sigma) and appropriate diluted RNase T1 (1 or 2 U) was added, followed by 15 min incubation at 37°C. The reaction was stopped by adding another 20 µg tRNA. The control reactions were performed in the absence of protein or RNase T1 under corresponding conditions. The RAT RNA was extracted with acid phenol/chloroform, precipitated with ethanol, resuspended and analyzed by electrophoresis.

Screening system to isolate random terminator mutants

Plasmids containing the mutated *ptsG*-RAT and a promoterless kanamycin resistance cassette were obtained as follows. The mutated *ptsG*-RAT alleles were amplified via PCR using the plasmids pGP329 and pGP330 as template DNA (obtained by site-directed mutagenesis of the RAT sequence; see above) and the oligonucleotides IL5 (Langbein *et al.*, 1999) and JS11 (Stülke *et al.*, 1997). The PCR products were digested with *Mfe*I and *BamH*I and the 0.5 kb fragment was cloned into pAC6 (Stülke *et al.*, 1997) cut with the same enzymes. The resulting plasmids were designated pGP337 and pGP338. In a second step a promoterless aphA3 kanamycin resistance cassette (Trieu-Cuot and Courvalin, 1983) was amplified using the primers IL46 (5'-CGGGATCCTAATGTTAGAAAAGAGGAA GGAAATAA) and IL48 (5'-CGGGATCCCTACTAAAACAATTCATCCAGTAA) and pAC7 as template DNA. The DNA fragment was cut with *BamH*I and cloned into the target vectors pGP337 and pGP338 linearized with the same enzyme. The resulting plasmids were designated pGP339 and

pGP340, respectively. Moreover, by the same strategy we constructed plasmid pGP342 harboring a *ptsG–aphA3 lacZ* fusion under the control of the wild-type regulatory elements. These plasmids were used to introduce a *ptsG-aphA3 lacZ* fusion into the chromosome of *B. subtilis* 168 (see Table 6).

Liquid cultures of GP172 and GP173 containing the artificial *ptsG–aphA3 lacZ* operon were grown overnight in the absence of kanamycin. Aliquots of 100 µl of independently grown cultures were plated onto SP plates containing kanamycin at a final concentration of 30 and 60 µg/ml, respectively. The plates were incubated for 2 days at 28°C to select for colonies able to grow on kanamycin and express the *ptsG–aphA3 lacZ* fusion. To verify the localization of the selected mutations in the control region of the artificial operon, *B. subtilis* 168 was transformed with chromosomal DNA of the mutant strains and transformants were selected for chloramphenicol resistance. Transformants resistant to both chloramphenicol and kanamycin contained the mutation in close proximity to the cat resistance gene, which is located downstream of the *lacZ* gene. The *ptsG* promoter regions of these strains were analyzed by sequencing.

Results

Mutational analysis of the *ptsG* terminator sequence

In a previous work, we demonstrated that a mutation disrupting the presumed stem structure of the transcriptional terminator in the *ptsG* leader region resulted in constitutive expression of *ptsG* (Langbein *et al.*, 1999). To ensure that it is the structure of the terminator rather than its sequence we constructed a strain, GP169, carrying a compensatory mutation in the terminator region. In this strain, the potential stem structure of the terminator is restored without correcting the sequence. The *ptsG* promoter region containing this mutation was inserted into the chromosome of *B. subtilis* in front of a promoterless *lacZ* gene. The synthesis of ß-galactosidase after growth of the strains in CSE minimal medium in the absence and presence of glucose was monitored to assay transcription driven by the wild-type and mutant *ptsG* promoter regions (Table 7). While transcription was inducible by glucose in the wild-type strain QB5448, constitutive expression was observed in the terminator mutant GP151.

Table 7. Effect of mutations in the terminator on expression of a *ptsG–lacZ* fusion

Strain	Terminator	ß-Galactosidase activity (U/mg protein)[a]	
		– Glucose	+ Glucose
QB5448	Wild-type	19	294
GP151	C75G C76G	779	360
GP169	G37C G38C C75G C76G	18	204
GP174	Wild-type	5	450
GP172	'Wild-type' (G11A G26A)	3	12
GP186	GP172 Δ4	824	584
GP187	GP172 G72U	49	120
GP188	GP172 Δ3	841	651
GP189	GP172 G28U	497	445
GP190	GP172 C75A	133	177
GP191	GP172 G47U	405	368
GP173	'Wild-type' (G11C G26A)	7	11
GP179	GP173 Δ1	1635	1725
GP181	GP173 G38A	71	95
GP182	GP173 G47U	475	337
GP183	GP173 Δ2	796	510
GP184	GP173 C75U	98	113
GP185	GP173 G90A	170	150

[a] Representative values of *lacZ* expression. All measurements were performed at least twice.

These results are in good agreement with previous findings (Langbein *et al.*, 1999). The newly constructed compensatory mutation restored the inducibility of *ptsG* expression and therefore the functionality of the transcriptional terminator. Thus, the structure of the terminator seems to be important rather than its actual sequence. The restoration of the terminator structure results in transcription termination in the absence of the inducer, glucose.

To provide further evidence for the functional role of the terminator, we developped a screening and selection system that allowed us to select for spontaneous mutations resulting in constitutive *ptsG* expression. Briefly, an artificial *aphA3 lacZ* operon was created and placed under control of the *ptsG* expression signal. The two strains used here (GP172 and GP173) carried mutations of the RAT sequence that prevented their recognition by GlcT. Mutant strains allowing expression of the artificial operon were selected as decribed in Materials and Methods.

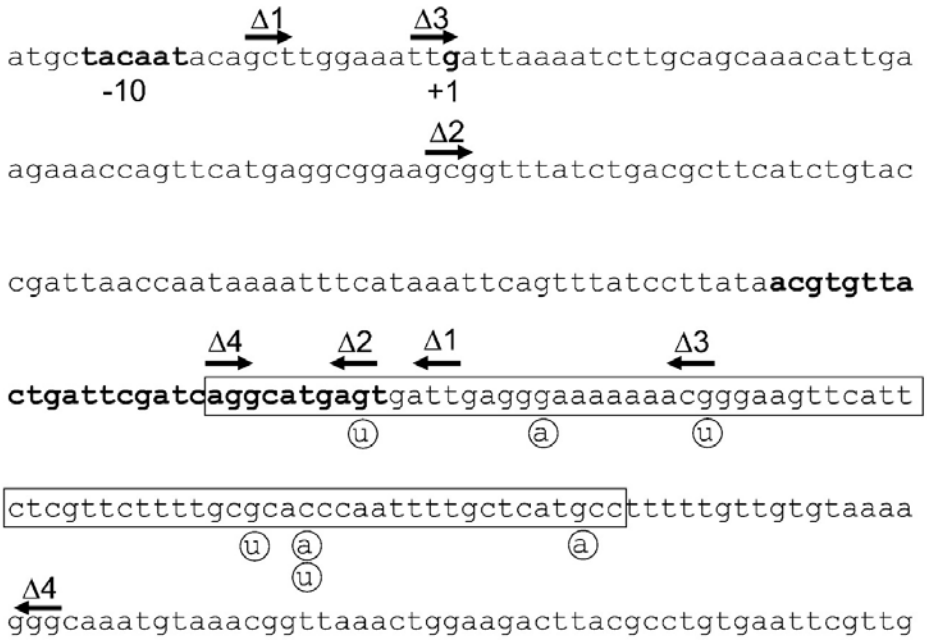

Figure 4.2. Sequence of the *ptsG* promoter region. The RAT sequence is shown in bold and the terminator is boxed. The –10 promoter region and the transcription start site are shown in bold. The identified point mutations are shown in circles below the original nucleotides. Arrows with numbers indicate the deleted regions found in the sequenced mutants. The strain designations, the 'numbering' of the positions and the effects of the mutations are given in Table 7.

Several mutants exhibiting both resistance to kanamycin and constitutive expression of the *lacZ* gene (as monitored by the blue color of colonies on X-Gal plates) were obtained and their *ptsG* promoter regions upstream of the *aphA3 lacZ* operon were sequenced. The mutations all affected the transcriptional terminator in the *ptsG* leader region. A total of 12 different spontaneous mutations, six derived from each strain, were identified. Among these mutations were four deletions encompassing the terminator region and eight point mutations. One of these point mutations ($G_{47}T$) was obtained with both original strains, whereas two different mutants obtained with either GP172 and GP173 affected the same position of the terminator ($C_{75}A$ and $C_{75}T$, respectively). A summary of the sequence information concerning the mutations is shown in Figure 4.3. The effect of the mutations was quantified by assaying the ß-galactosidase activity driven by the mutant *ptsG* control elements. As expected, inducible expression was observed in the wild-type strain GP174, while no expression of the *lacZ* gene was detected in the mutant strains GP172 and GP173 that were used to select for terminator mutants. The mutations of the terminator all resulted in constitutive expression of the artificial *aphA3 lacZ* operon. However, the absolute level of expression was different for

the individual mutants. The deletions and the mutation of G_{47} resulted in high constitutive expression (400–1700 U ß-galactosidase/mg protein; see Table 7). A point mutation at position C_{75} or G_{90} resulted in intermediate expression (100–200 U ß-galactosidase/mg protein), while the mutations of G_{38} and G_{72} resulted in weak constitutive expression (<100 U ß-galactosidase/mg protein; see Table 7).

Taken together, our findings indicate that the terminator is the only negatively acting regulatory element in the *ptsG* promoter region. Moreover, it is the secondary structure rather than the actual sequence of the terminator which is important for the control of *ptsG* expression.

Binding of the RNA-binding domain of GlcT to the RAT sequence

Previous genetic and biochemical evidence suggested direct binding of GlcT to the RAT sequence that overlaps the terminator in the *ptsG* leader mRNA. To substantiate these observations and to study the molecular details of protein–RNA interaction, we first purified the N-terminal domain of GlcT devoid of a His tag. In a previous study, binding of His-tagged GlcT to the RAT RNA was demonstrated (Langbein *et al.*, 1999). However, subsequent gel retardation experiments suggested that the affinity tag might cause some non-specific RNA-binding (our unpublished results). We therefore constructed a system that allowed removal of the His tag after initial purification of the protein.

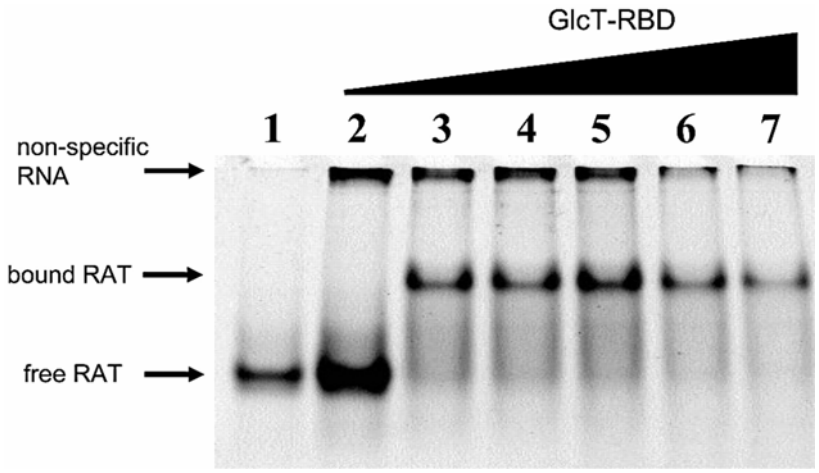

Figure 4.3. Electrophoretic mobility band shift analysis of the interaction between the RAT and the RBD. Lane 1 shows 125 pmol of the free RAT RNA. In lanes 2–7, 250 pmol RAT RNA and 30 pmol non-specific RNA were used. Increasing concentrations of RBD-GlcT were added to the RNA prior to electrophoresis. Aliquots of 250, 500, 750, 1000 and 1500 pmol RBD-GlcT were used. The arrows indicate the different RNA species.

The purified protein was efficiently cleaved at the introduced thrombin cleavage site. The resulting protein was purified to apparent homogeneity by size exclusion chromatography as described in Materials and Methods.

The ability of the *in vitro* generated RNA-binding domain of GlcT to bind its RNA target was tested by electrophoretic mobility shift analysis. As shown in Figure 4.4, addition of the RNA-binding domain to the RAT RNA resulted in a shift of the apparent size of the RAT fragment. At a high concentration of the RNA-binding domain (3-fold molar excess), the RAT RNA was completely shifted (see Fig. 4.3, lane 7). The shifted band contained a complex of both the RNA and the RNA-binding domain, as suggested by the differential detection of both components with ethidium bromide (detection of RNA) and Coomassie brilliant blue (detection of protein) (data not shown).

Structural analysis of the GlcT–RAT interaction by RNase T1 protection footprint analysis

Specificity determinants have already been studied for the RAT sequences of the *bgl* and the *sac* type and the structure of the complex between the *bglP* RAT and the LicT antiterminator was determined (Aymerich and Steinmetz, 1992; Yang *et al.*, 2002). However, there is no cross-talk between the antiterminators of the Bgl and Sac types and the *ptsG* RAT sequence and GlcT is not capable of causing antitermination at *bgl* or *sac* RAT sequences (Langbein *et al.*, 1999). Thus GlcT and the *ptsG* RAT may be the most distant members of this large family of protein-dependent riboswitches. This prompted us to initiate an analysis of the molecular details of the interaction between the RNA-binding domain of GlcT and the *ptsG* RAT sequence by RNase T1 footprint analysis.

RNase T1 specifically cleaves adjacent to G residues in single-stranded regions of the RNA. Binding of the G residue to a protein or formation of a double-stranded RNA structure protects the RNA from degradation by RNase T1. As expected, incubation of denatured RAT RNA with RNase T1 resulted in RNA cleavage at each of the G residues in the RAT sequence (see Fig. 4.4, lane 2). Similarly, all G residues are available for cleavage by RNase T1 in the native form of the RAT RNA (see Fig. 4.4, lanes 3 and 5). Thus, the RAT does not seem to adopt a secondary structure in the absence of GlcT. In the presence of the RNA-binding domain of GlcT several G residues were protected from nucleolytic degradation, with only one G residue remaining accessible.

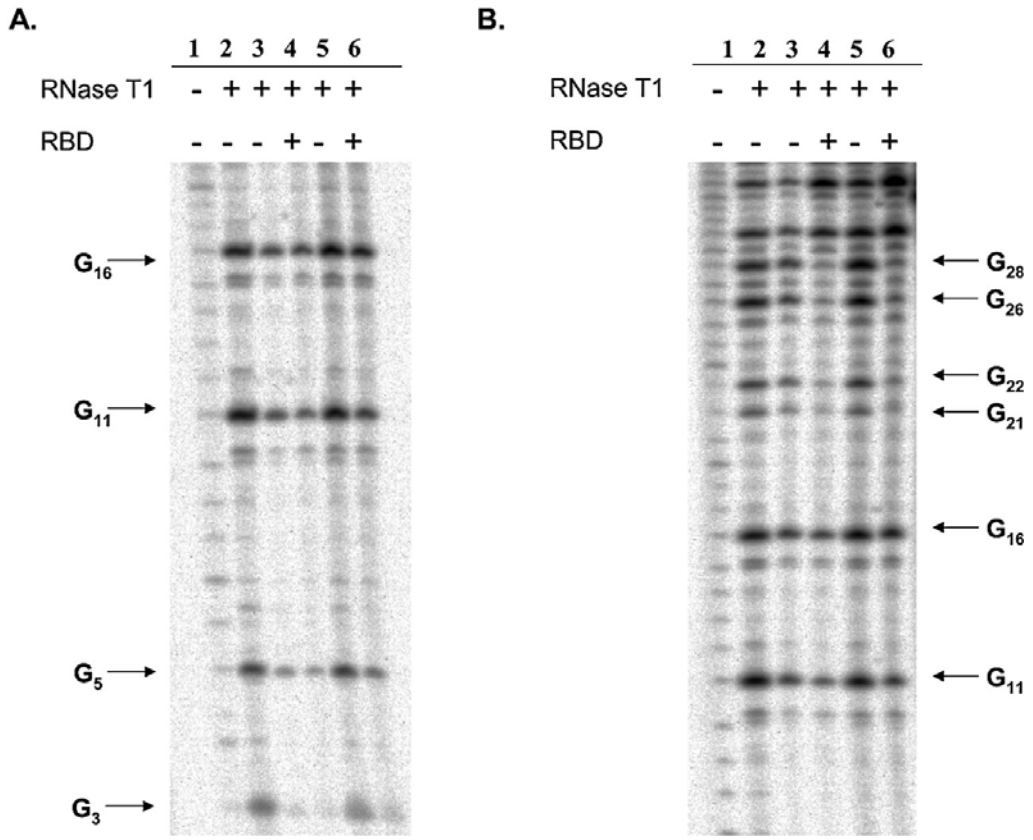

Figure 4.4. RNase T1 footprint of the *ptsG* RAT region. Two different polyacrylamide concentrations were used to resolve the complete region: (**A**) 26%; (**B**) 20%. Weak hydrolysis of the RNA in the absence of the RNase and the antiterminator protein shows every nucleotide of the RAT (lane 1). High concentrations of RNase T1 under denaturing conditions leads to cleavage at every guanine nucleotide (lane 2). In the other assays, the RAT RNA and the RBD were used as indicated. In lanes 5 and 6, twice as much radiolabeled RNA was used as compared to lanes 3 and 4. Arrows indicate the position of the guanine nucleotides in the RAT.

This residue, G_{16}, is probably located in the top loop of the RAT secondary structure and thus part of a single-stranded region in the absence and in the presence of GlcT. Among the residues protected by GlcT, G_{22}, G_{26} and G_{28} were most strongly protected. According to the secondary structure model of the RAT RNA (see Figs. 4.1 and 4.7), G_{22} and G_{26} are located in loop regions of the RAT RNA and may thus be protected by interaction with GlcT. G_3, G_{11} and G_{21} were weakly protected in the presence of GlcT. This may reflect their location in a double-stranded region rather than interaction with GlcT. Indeed, G_{11} and G_{21} are both thought to be located in the third double-stranded region of the RAT structure. Thus, the footprinting data indicate that the RAT RNA exists in the single-stranded form in the absence of GlcT and adopts the double-stranded conformation upon GlcT addition. Moreover, GlcT seems to interact with residues in the two loops of the RAT structure.

Probing of the RAT structure and of protein–RNA interactions by site-directed mutagenesis

Since only the guanine residues can be probed by RNase T1 footprinting, we attempted to study the role of the individual bases of the RAT sequence in more detail by site-directed mutagenesis. Mutations of the RAT sequence were introduced into the *ptsG* promoter region, located in front of a promoterless *lacZ* gene. If the mutations in the RAT sequence were expected to affect the structural integrity of the overlapping terminator, we introduced compensatory mutations in the terminator to restore the presumed wild-type secondary structure (see Fig. 4.1).

To assay the effects of the different mutations, the strains were grown in CSE minimal medium in the absence and presence of glucose and the ß-galactosidase activities were determined as an indicator of termination/antitermination activity. The results are summarized in Table 8. As observed previously, a strong induction of the *ptsG–lacZ* fusion was detected for the wild-type strain QB5448.

Mutations preventing formation of the first double-stranded region of the RAT RNA resulted in loss of induction (GP367, C_2G; GP368, $G_{28}C$). Interestingly, a strain with an inverse arrangement of the C and G residues (and an additional compensatory mutation in the terminator, GP366; see Table 8) resulted in constitutive expression rather than in restoration of inducibility. Constitutive expression is a strong indication of inactivity of the transcriptional terminator. Thus, the overall structure of the RAT/terminator region may have changed in such a way that transcription termination no longer occurs in the triple mutant GP366. To address this hypothesis, we assayed the activity of the *ptsG* control region present in GP366 in a *glcT* background. While no expression of the wild-type *ptsG–lacZ* fusion was observed in the *glcT* mutant GP109, constitutive expression was detected for the C_2G $G_{28}C$ $C_{85}G$ triple mutant (data not shown). This result confirms that antitermination is not required to express ß-galactosidase in this construct. To test whether GlcT binds the RAT sequence with the inverse C:G base pair, we performed a gel mobility shift assay. As shown in Figure 4.6, this fragment was as efficiently bound by GlcT as the wild-type RAT fragment. In contrast, a fragment containing only the C_2G mutation that interferes with formation of the lowest part of the stem was barely shifted by GlcT. Thus, formation of this stem is important for RNA-binding by GlcT, whereas the actual sequence does not seem to be critical.

Table 8. Effect of mutations in the RAT sequence on expression of a *ptsG–lacZ* fusion

Strain	RAT	Compensatory mutation (terminator)	ß-Galactosidase activity (U/mg protein)[a]	
			– Glucose	+ Glucose
QB5448	Wild-type		11	610
GP367	C2G		17	7
GP368	G28C	C85G	29	18
GP366	C2G G28C	C85G	360	302
GP375	G3C		14	3
GP376	G3A		20	4
GP380	A27U	U86A	9	7
GP381	A27G	U86C	686	1053
GP370	G26U	G87A	64	1038
GP379	U25G	A88C	8	12
GP369	G5C C23G	G90C	14	277
GP377	U6G		7	3
GP378	U6A		26	4
GP371	U7G		15	225
GP175	U7C		9	580
GP176	A8G		8	58
GP177	U7C A8G		33	60
GP373	G22C	C91G	12	1
GP374	G22U	C91A	7	2
GP178	C9U		5	21
GP372	C9G G21C	C92G	376	418
GP160	G11A		6	82
GP166	A12U U13A		5	353
GP156	U14A		12	367
GP158	C15U		7	492
GP159	C15A		7	562

[a] Representative values of *lacZ* expression. All measurements were performed at least twice.

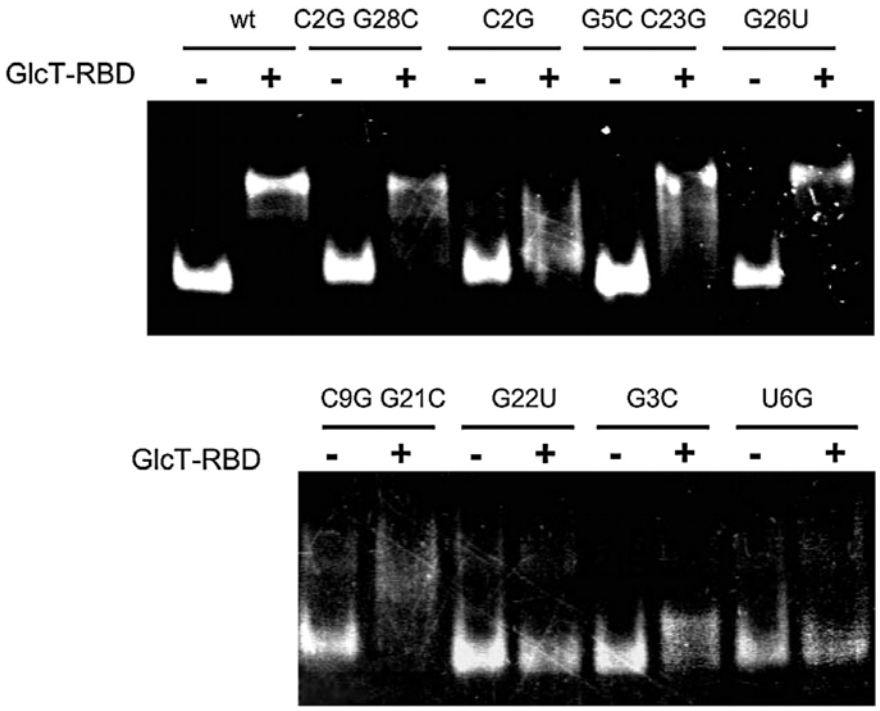

Figure 4.5. Electrophoretic mobility band shift analysis of the interaction between different mutant RAT species and the RBD.

Mutations affecting loop 1 of the RAT had different effects: two different mutations affecting the G_3 residue resulted in complete loss of induction. This finding is in good agreement with the observed role of this residue in the interaction with GlcT as inferred from the RNase T1 protection experiment. Moreover, a RAT sequence containing the G_3C mutation is not bound by GlcT *in vitro* (see Fig. 4.5). However, both mutations G_3C and G_3A may alternatively affect the structure of the RAT RNA, as both introduced bases would find pairing partners on the opposing side of the RAT (U_{25} or G_{26}). Similarly, mutations affecting the residues U_{25} and A_{27} resulted in loss of antitermination. These mutations were not expected to alter the secondary structure of the RAT. However, the compensatory mutation for one of the mutations ($U_{86}C$ in GP381) results in loss of terminator activity and thus in constitutive expression of ß-galactosidase. Finally, the mutation $G_{26}U$ present in GP370 resulted in increased expression of the *ptsG–lacZ* fusion, both in the absence and presence of glucose. Here, the terminator may be leaky due to the compensatory mutation (A:U instead of G:C base pair) and allow some basal expression even in the absence of glucose. Alternatively, the perfectly symmetrical structure of the RAT generated by the $G_{26}U$ mutation might result in improved binding of GlcT, including some weak antitermination in the absence of glucose. Indeed, the mutated RAT seems to be an efficient target for GlcT, as judged from the clear

induction of ß-galactosidase in this strain. In addition, the RAT RNA containing this mutation was efficiently bound by GlcT *in vitro* (see Fig. 4.5). Moreover, expression of this *ptsG–lacZ* fusion was completely lost in a *glcT* mutant strain (see Table 9). Taken together, our data indicate that loop 1 is of great importance for the proper interaction of the RAT RNA with GlcT.

One mutant affected in the second short double-stranded region of the RAT was studied. In this strain, GP369, the predicted G:C base pair was inverted to be C:G, and efficient induction of ß-galactosidase activity was detected. Additionally, the RAT RNA carrying the inversion interacted with the RNA-binding domain of GlcT *in vitro* (see Fig. 4.5). Thus, the structure rather than the actual nucleotide at this position is important for GlcT binding and antitermination. This is in good agreement with the observed weak protection of G_5 from RNase T1 digestion in the presence of GlcT.

Several mutations in loop 2 resulted in loss of GlcT binding to the RAT and thus in loss of antitermination, namely those affecting U_6 and G_{22}. Results obtained with the *ptsG-lacZ* fusion (Table 8) and the *in vitro* RAT–GlcT binding assay (see Fig. 4.5) are in perfect agreement for these mutations. In contrast, replacement of U_7 and/or A_8 showed no or only slight effects. The importance of G_{22} for efficient transcriptional antitermination is consistent with the strong RNase T1 protection obtained for this base in the presence of GlcT. However, similarly to loop 1, not all bases seem to play an important role in GlcT recognition.

As we approached the top of the proposed RAT structure, the mutations had lesser effects. The first base pair of the third double-stranded region is important for antitermination, as judged from lack of expression of the *ptsG–lacZ* fusion in the corresponding mutant strain GP178. An inversion of this base pair along with the required compensatory mutation in the terminator caused constitutive expression (GP372), thus we are unable to decide whether this inversion would affect antitermination. The constitutive expression of ß-galactosidase in this strain was not abolished by deletion of the *glcT* gene (data not shown), indicating that the mutation inactivated the terminator. An *in vitro* assay of GlcT binding to the mutant RAT suggested that an interaction occurred, albeit the complex seemed to be less stable (see Fig. 4.5). A mutation of G_{11} had a mild effect whereas the $A_{12}U$ $U_{13}A$ double mutation present in GP166 did not influence the antitermination activity of GlcT. Similarly, three mutations in the top loop of the RAT did not affect induction by glucose and, therefore, binding of GlcT. These results indicate that the upper part of the RAT may only be required to obtain the correct RAT structure and is not involved in the interaction with GlcT.

Table 9. Effect of mutations in the *ptsG* RAT sequence on induction of a *ptsG–lacZ* fusion

Strain	Relevant genotype	ß-Galactosidase activity (U/mg protein)[a]				
		CSE	CSE Glc	CSE Suc (0.1%)	CSE Suc (2%)	CSE Sal
QB5448	Wild-type	9	548	231	329	267
GP109	Δ*glcT*	4	5	2	8	9
GP387	Δ*sacT* Δ*sacY*	10	327	85	166	335
GP389	Δ*licT*	12	412	222	422	312
GP385	RAT-A_{ins}	6	8	37	34	102
GP386	RAT-A_{ins} Δ*glcT*	10	18	48	49	121
GP388	RAT-A_{ins} Δ*sacT* Δ*sacY*	6	11	5	6	110
GP390	RAT-A_{ins} Δ*licT*	16	11	14	17	13
GP370	RAT-$G_{26}U$	38	790	622	740	665
GP392	RAT-$G_{26}U$ Δ*glcT*	8	7	13	44	29

[a] Representative values of *lacZ* expression. All measurements were performed at least twice.

Converting the recognition specificity of the RAT sequence to LicT binding

A comparison of RAT sequences of the glucose, sucrose, ß-glucoside and lactose classes revealed that all but the glucose RAT sequences contain 3 bp in the second double-stranded region whereas the *ptsG* RAT sequences of *B. subtilis* and *S. carnosus* contain only 2 bp (see Fig. 4.6). Moreover, loop 1 of the *ptsG* RAT is made up of four putatively non-paired bases whereas three non-paired bases form this loop in the other RAT sequences (Knezevic *et al.*, 2000; Stülke *et al.*, 1997; Yang *et al.*, 2002) (see Fig. 4.6). Our mutagenesis studies suggested an important role of RAT structure in the loop regions for the interaction with GlcT. We therefore asked whether this unique arrangement of the *ptsG* RAT is important for induction specificity by GlcT.,

To address this question we constructed a *B. subtilis* strain with a *ptsG–lacZ* fusion with an additional adenine nucleotide between G_3 and U_4 of the RAT. This insertion is predicted to result in the formation of an extra A:U base pair in the second paired region. Concomitantly, the mutant loop 1 would contain three bases instead of four. Thus, the structure of this RAT is expected to be similar to those of the sucrose and ß-glucoside classes (see Fig. 4.6).

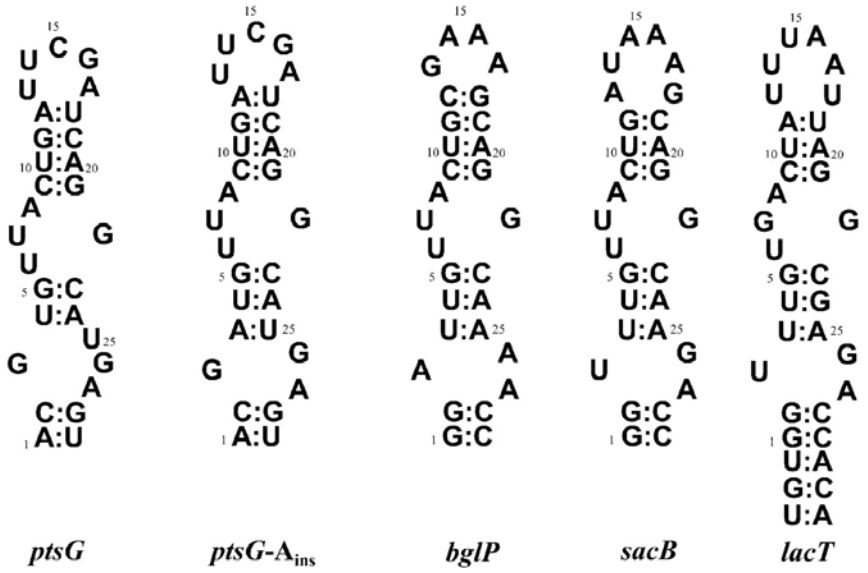

Figure 4.6. Proposed secondary structure for different RAT RNAs (Yang *et al.*, 2002). The fold proposed for the *ptsG*-A$_{ins}$ RAT is based on similarity to the *bglP* and *sacB* RAT RNAs.

Binding of antitermination proteins to this mutant RAT was assayed by determining the activity of the *ptsG–lacZ* fusion under conditions when the different antiterminator proteins are known to be active. In the wild-type strain QB5448, glucose, sucrose and salicin induced expression of the fusion. This finding is in agreement with previous reports and results from the activity of GlcT due to non-specific regulation by glucose permease (compare QB5448 and GP109 in Table 9) (Görke *et al.*, 2004; Langbein *et al.*, 1999; Stülke *et al.*, 1997). In contrast, the mutant *ptsG* control region present in GP385 did not allow induction by glucose, suggesting that the altered RAT was not recognized by GlcT. However, significant induction was observed in the presence of sucrose or the ß-glucoside salicin. This induction was independent of a functional *glcT* gene (compare GP385 and GP386 in Table 9), suggesting that the mutant RAT was bound by a different antiterminator. To test this possibility, we constructed strains carrying the wild-type or mutant RAT sequences and deletions of the genes encoding the other *B. subtilis* antiterminators of the BglG family, i.e. *sacT*, *sacY* and *licT*. These deletions had no effect on the activity of a *ptsG–lacZ* fusion in the presence of the wild-type RAT sequence. This observation is in good agreement with the idea that induction of *ptsG* by sucrose and salicin is mediated by GlcT (Görke *et al.*, 2004; Langbein *et al.*, 1999; see above). However, the *lacZ* fusion controlled by the mutant RAT was not induced by sucrose and salicin in the *licT* mutant strain GP390. Thus, LicT is essential to overcome

transcription termination in the presence of the altered RAT sequence. Similarly, induction by sucrose is lost in the *sacT sacY* double mutant strain GP388. SacT and SacY recognize identical RAT sequences and therefore the effect of these mutations was not assayed individually. Interestingly, induction is still possible in the presence of salicin in GP388, confirming that this induction is caused by LicT.

Discussion

As a target of regulation, RNA differs substantially from DNA in its structural diversity and in its ability to adopt alternative, mutually exclusive structures. Moreover, these structures, the riboswitches, are capable of interacting with a wide variety of regulatory partners, such as metabolites, other RNAs or proteins. Due to this huge variability, the themes in RNA–protein recognition are much more diverse than those identified in DNA–protein interactions.
GlcT and the *ptsG* riboswitch are members of an expanding family of transcription regulatory systems (Stülke *et al.*, 1998). A recent analysis of the evolution of the regulatory domains of the antitermination proteins of this family revealed that GlcT from *B. subtilis* and its ortholog from *S. carnosus* form a distinct subgroup whereas the antiterminators of the sucrose and ß-glucoside classes exhibit a close relationship with each other (Greenberg *et al.*, 2002). This is supported by the finding that there is no cross-talk between the regulatory components of the glucose type on the one hand and those of the sucrose and ß-glucoside type on the other (Langbein *et al.*, 1999). In contrast, binding of sucrose antiterminator proteins to RAT sequences of *bgl* genes was reported (Aymerich and Steinmetz, 1992).

Our studies regarding the overall structure of the *ptsG* RAT suggest that it folds similarly to the RAT RNAs of the *bgl* and *sac* classes, thus confirming a proposal of Yang *et al.* (Yang *et al.*, 2002). Interestingly, this structure deviates from that obtained by the Zuker algorithm proposed earlier (Langbein *et al.*, 1999). However, both mutational and footprint analyses are in disagreement with the RAT structure suggested previously, whereas they confirm the model depicted here (see Figs. 4.1, 4.6, and 4.7). Mutations in the proposed stem regions had different consequences: single mutations prevented antitermination while compensatory mutations that restored the base pairing restored GlcT binding at the same time. Similar results were obtained with the *sacB* RAT sequence (Aymerich and Steinmetz, 1992).

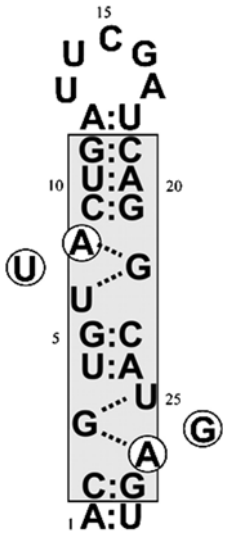

Figure 4.7. Proposed model of the base pairing in the RAT RNA. The circled bases are proposed to directly interact with GlcT.

While the sequence of the *ptsG* RAT differs substantially from those of the RAT sequences of the *sac* and *bgl* classes, we detected a perfect match of loop 2 of the *ptsG*, *bglP* and *sacB* RAT sequences and the regions surrounding loop 2 (see Fig. 4.6). It therefore seems safe to conclude that this region of the RAT is structurally identical to that of the *bglP* RAT as determined by NMR spectroscopy (Yang *et al.*, 2002). In loop 2, G_{22} forms contacts with two bases, U_6 and A_8. Thus, only U_7 extrudes from the stem structure (see Fig. 4.7). Analysis of the LicT interaction with the *bglP* RAT revealed that G_{22} is not involved in direct interactions with the protein. However, the mutations of G_{22} and its partners studied in this work abolished antitermination. This confirms the critical importance of these residues for the structure of the RAT. The *ptsG* RAT is unique since it is the only RAT sequence with a symmetrical arrangement of the two loops. Therefore, it is tempting to speculate that G_3 is involved in contacts with U_{25} and A_{27}, with G_{26} being exposed. This idea is again supported by the mutational analysis: mutations disrupting the 'ménage à trois' in loop 1 strongly interfere with antitermination and are therefore expected to interfere with the RAT structure, as observed for loop 2.

Our studies concerning the interaction of GlcT with the RAT revealed that loops 1 and 2, as well as the stems, are critical for antitermination. This is in good agreement with previous findings for the SacY-*sacB* RAT and LicT-*bglP* RAT interactions (Aymerich and Steinmetz, 1992; Yang *et al.*, 2002). Our mutation and footprint analyses demonstrated that the top loop is important neither for adoption of the proper RAT structure nor for the

interaction with GlcT. In contrast, the RNA-binding coat protein of bacteriophage MS2 and the phage λ N protein requires the top loop of their RNA targets for binding (Grahn *et al.*, 1999; Legault *et al.*, 1998). As far as the paired regions are concerned, our analysis fails to identify the bases and amino acids in the RAT and GlcT, respectively, that are directly involved in the interaction. However, in our footprint analysis G_{26} in loop 1 was strongly protected from RNase T1 cleavage in the presence of GlcT, suggesting that it is contacted by the protein. The corresponding base in loop 2 was shown to be involved in interaction of the *bglP* RAT with several amino acids of LicT (Yang *et al.*, 2002). We have analyzed mutations of these two exposed nucleotides in loop 1 (G_{26}) and loop 2 (U_7). In both cases a strong induction by glucose was observed. We may therefore propose that although these exposed bases are contact partners for GlcT, their actual nature does not seem to be crucial. This fits well with previous observations on the interaction of the *bglP* RAT and LicT (Yang *et al.*, 2002).

The idea that the structure of the RAT rather than its sequence is important for recognition by GlcT is strongly supported by the observation that insertion of a single nucleotide that makes the RAT structure more similar to that of the *bgl* and *sac* RATs resulted in complete loss of GlcT-dependent induction. In contrast, the novel RAT is recognized by LicT even though the sequence remains quite different from that of the cognate LicT RAT sequences. On the other hand, a small sequence alteration with a probably important structural consequence results in complete loss of antitermination by GlcT. So far, it is known that SacT and SacY bind to RAT sequences of the *bgl* and *sac* classes whereas LicT (and its *E. coli* counterpart BglG) binds only to *bgl* RAT sequences (Aymerich and Steinmetz, 1992; Manival *et al.*, 1997). The structures of the RAT sequences of these two classes are very similar (see Fig. 4.6). Accordingly, the recognition specificity for LicT is determined in the sequence of the RAT, and mutations in loop 1 of the *sacB* RAT sequence allowed binding of BglG in addition to SacY. The effects of the corresponding mutations were cumulative (Aymerich and Steinmetz, 1992). This is in agreement with the idea that SacY and LicT recognize RAT RNAs with the same structure. In contrast, we report here the first example of a complete change of partner specificity. A single mutation that changes the structure of the RAT RNA is sufficient for altered recognition and results in an all-or-nothing effect. Obviously, a RAT can be a member of the *glc* or the *sac/bgl* family, with no intermediate possible. The RNA-binding domains of LicT and GlcT are well conserved (for alignments see Langbein *et al.*, 1999; Yang *et al.*, 2002. Five amino acids of LicT were shown to be involved in the direct interaction with the internal loops of the RAT RNA (Yang *et al.*, 2002). Of these, three (N10,

G26 and F31) are conserved in all *B. subtilis* antiterminator proteins of this family. The residue K5 is conserved in all proteins but SacY. Only R27 is not conserved at all and may thus be involved in providing interaction specificity.

The evolution of regulation by adaptation of DNA-binding proteins to new DNA-binding sites is well documented. In the case of protein-dependent riboswitches, the question arises what was first: does the protein evolve to recognize an altered RNA or does the RNA adapt to protein variations? Our results hint at the second idea: one mutation in the RNA is sufficient to gain a novel regulatory specificity. In contrast, many mutations in the protein would be required to recognize an RNA target with a similar sequence but different structure. It will be interesting to study the molecular details of GlcT–RAT interaction by determining the structure of the complex.

5 Keeping signals straight in transcription regulation: specificity determinants for the interaction of a family of conserved bacterial RNA-protein couples

The results described in this chapter were published in:

Schilling, O., C. Herzberg, T. Hertrich, H. Vörsmann, D. Jessen, S. Hübner, F. Titgemeyer, and J. Stülke. 2006. Keeping signals straight in transcription regulation: specificity determinants for the interaction of a family of conserved bacterial RNA-protein couples. *Nucleic Acids Res* **34:**6102-6115

Author contributions:

This study was designed by O.S. and J.S.. O.S. performed and interpreted the reporter gene analyses and gel shift experiments. O.S. constructed all mutant strains. S.H. performed the Northern Blot analysis. T.H. performed the glucose uptake measurements and F.T. supervised T.H. during her diploma work. H.V. and D.J. contributed to some of the mutant strain constructions and reporter gene analyses during their practical training under the supervision of O.S.. C.H. contributed substantially to all lab work related to this project. O.S. and J.S. wrote the paper.

Abstract

Regulatory systems often evolve by duplication of ancestral systems and subsequent specialization of the components of the novel signal transduction systems. In the Gram-positive soil bacterium *Bacillus subtilis*, four homologous antitermination systems control the expression of genes involved in the metabolism of glucose, sucrose and β-glucosides. Each of these systems is made up of a sensory sugar permease that does also act as phosphotransferase, an antitermination protein, and a RNA switch that is composed of two mutually exclusive structures, a RNA antiterminator (RAT) and a transcriptional terminator. We have studied the contributions of sugar specificity of the permeases, carbon catabolite repression, and protein–RAT recognition for the straightness of the signaling chains. We found that the β-glucoside permease BglP does also have a minor activity in glucose transport. However, this activity is irrelevant under physiological conditions since carbon catabolite repression in the presence of glucose prevents the synthesis of the β-glucoside permease. Reporter gene studies, *in vitro* RNA–protein interaction analyzes and northern blot transcript analyzes revealed that the interactions between the antiterminator proteins and their RNA targets are the major factor contributing to regulatory specificity. Both structural features in the RATs and individual bases are important specificity determinants. Our study revealed that the specificity of protein–RNA interactions, substrate specificity of the permeases as well as the general mechanism of carbon catabolite repression together allow to keep the signaling chains straight and to avoid excessive cross-talk between the systems.

Introduction

To sense their environment and to adapt to changing conditions, all organisms possess signal transduction systems which are composed of a sensor that perceives the signal, a regulator that can modify its activity in response to the signal, and a target of regulation. This general scheme can be modified in many ways: The sensor and the regulator are often combined in the same molecule as in the Lac repressor. The sensor and the regulator are usually proteins, but regulatory RNAs continue to be uncovered. The target of the regulation may be a protein, i.e. an enzyme, but for the control of gene expression, specific DNA or RNA sequences are the most common targets.

In bacteria, the number of environmental or internal signals that need to be sensed is much higher than the number of non-related regulatory systems. Thus, large families of regulation systems are present in bacteria. Among the most common families are the two-component regulatory systems, sigma factors with their anti-sigma factors as well as several families of repressor and activator proteins (Helmann, 1999; Huffman and Brennan, 2002; Pané-Farré *et al.*, 2005; Ramos *et al.*, 2005; Stock *et al.*, 2000). All these families can be divided into sub-families that do often respond to similar signals. The evolution of signaling families is still in progress and can be observed in the transcriptional regulation of biodegradation pathways. Even more, new regulatory systems can be generated artificially (Galvao and de Lorenzo, 2006; Garmendia *et al.*, 2001). The similarity of the components of many families of signal transduction systems raises the question how the bacteria avoid excessive cross-talk, i.e. the activation of a regulatory protein by gratuitous inducers or the induction of a gene by a non-cognate regulator protein that recognizes a similar DNA sequence. This problem was the subject of extensive analyzes for the two-component regulatory systems in the Gram-positive soil bacterium *Bacillus subtilis* (Hoch and Varughese, 2001).

We are interested in the control of glucose utilization in *B. subtilis*. This sugar is transported by a specific permease of the phosphotransferase system (PTS) encoded by *ptsG* and is subsequently catabolized via the glycolytic pathway (Stülke and Hillen, 2000). The expression of the *ptsG* gene and of several glycolytic genes is inducible by glucose, however, the mechanisms differ. While *ptsG* expression is induced by transcriptional antitermination, the glycolytic *gapA* operon is controlled by the repressor CggR (Doan and Aymerich, 2003; Fillinger *et al.*, 2000; Ludwig *et al.*, 2001; Stülke *et al.*, 1997). Induction of *ptsG* expression involves a RNA switch which is the target of the antitermination protein GlcT, and the sensory glucose permease, PtsG. As part of the PTS, the glucose permease possesses two soluble domains that are involved in the phosphate transfer from phosphoenolpyruvate to the incoming sugar, the domains IIA and IIB (Bachem *et al.*, 1997). If glucose is present, the phosphate groups are immediately transferred to the sugar, whereas they accumulate on the glucose permease as well as on the two general proteins of the PTS, enzyme I and HPr, in the absence of glucose. Under these conditions, the glucose permease can transfer a phosphate residue to GlcT thereby inactivating the antitermination protein (Bachem and Stülke, 1998; Schmalisch *et al.*, 2003). GlcT is made up of three domains, an N-terminal RNA-binding domain, and two homologous PTS-regulation domains called PRD-I and PRD-II (Bachem and Stülke, 1998; Manival *et al.*, 1997; Stülke *et al.*, 1998).

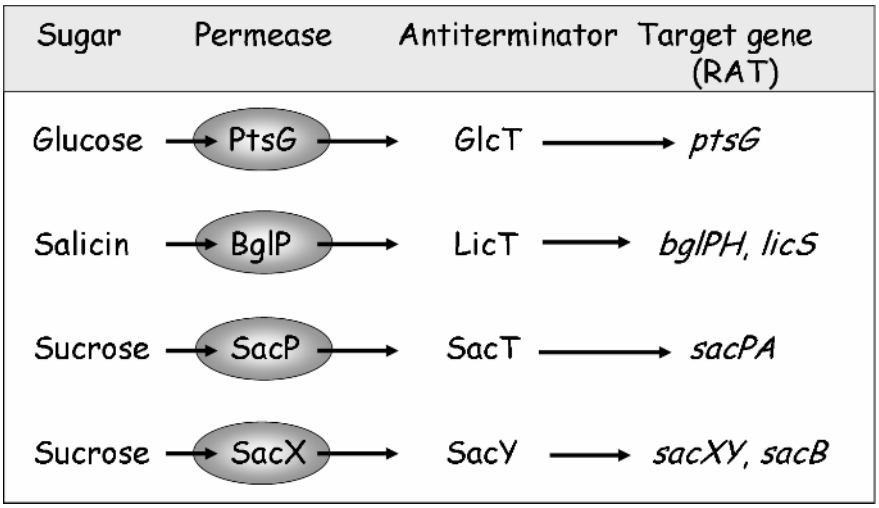

Figure 5.1. A family of antitermination systems controls sugar uptake and metabolism in *B. subtilis*. The design of the four signaling systems. The target genes encode the following proteins: *ptsG*, glucose permease of the PTS; *bglP*, β-glucoside permease of the PTS; *bglH*, phospho-β-glucoside hydrolase; *licS*, β-1,3-1,4-glucanase; *sacP*, sucrose permease of the PTS; *sacA*, sucrase; *sacX*, sucrose permease of the PTS (low affinity); *sacY*, antitermination protein; *sacB*, extracellular levansucrase.

Phosphorylation of a conserved histidine residue in PRD-I by the glucose permease results in GlcT inactivation in the absence of glucose. Biochemical studies revealed that PRD-II of GlcT can also be phosphorylated on a conserved histidine residue. However, this phosphorylation is catalyzed by the HPr protein of the PTS and has only a very minor impact on the activity of GlcT (Schmalisch *et al.*, 2003). If in the right phosphorylation state, i.e. if non-phosphorylated in PRD-I, GlcT can bind its target site on the *ptsG* mRNA called RNA antiterminator (RAT) (Aymerich and Steinmetz, 1992; Bachem and Stülke, 1998). The RAT overlaps a transcriptional terminator located in the leader region of the *ptsG* mRNA and the two structures form a RNA switch since they are mutually exclusive. Binding of GlcT to the RAT is thought to prevent the formation of the terminator and to allow transcription elongation into the *ptsG* structural gene. This regulatory system couples the availability of the inducer glucose to the phosphorylation state of the sensor permease and the antitermination protein GlcT resulting in either of two states of the *ptsG* RNA switch and subsequently in *ptsG* gene expression.

The regulatory system controlling *ptsG* expression is part of a family made up of highly conserved components, i.e. sensor permeases, antitermination proteins and RAT targets for the regulatory proteins (see Fig. 5.1). The additional permeases transport sucrose and the β-glucoside salicin. Two antitermination proteins, SacT and SacY, regulate expression of sucrose catabolic genes. While SacT is thought to be active at low sucrose concentrations, high concentrations of sucrose are required to activate SacY since its cognate permease SacX

has a very weak transport activity (Stülke and Hillen, 2000). LicT controls the expression of the *bglPH* operon and the *licS* gene in the presence of salicin (Schnetz *et al.*, 1996). SacT and LicT are only active if (i) their inducers are present and if (ii) no glucose is present in the medium. This allows their phosphorylation by HPr in the PRD-II leading to activation of the antiterminator proteins.

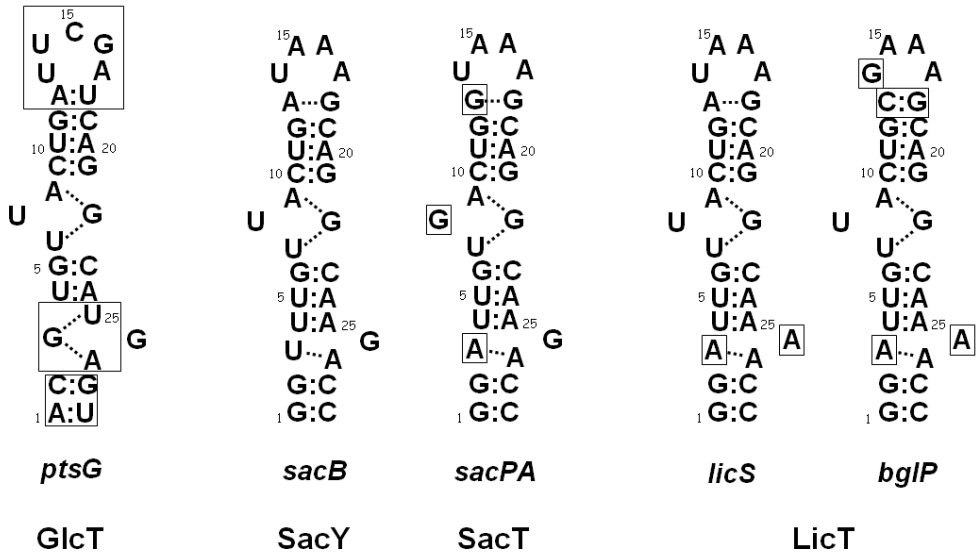

Figure 5.2. Comparison of the secondary structures of the related RAT-RNAs of *B. subtilis* (Schilling *et al.*, 2004; Yang *et al.*, 2002). The relevant antiterminator proteins are indicated below their cognate RAT structures. Boxes indicate nucleotides that differ from the sacB RAT. Dashed lines indicate bases that are proposed to be in direct contact to each other.

In contrast, SacY and GlcT, which are active in the presence of high sugar concentrations (or with the preferred sugar), are independent on a HPr-dependent activation even though HPr can phosphorylate these proteins. The glucose-dependent control of the antiterminator proteins' activity by HPr is part of the phenomenon of carbon catabolite repression which results in the preferential utilization of easily metabolizable carbon sources (Lindner *et al.*, 1999; Stülke *et al.*, 1998; Tortosa *et al.*, 1997).

The RAT targets of the four antiterminator proteins are all similar to each other (see Fig. 5.2). Some determinants causing specificity of protein–RNA interaction have been identified in a pioneering work (Aymerich and Steinmetz, 1992). The determination of the structure of the complex between the RNA-binding domain of LicT and its cognate RAT RNA suggested that the essential contacts between the protein and the RNA are made in the loop regions of the RAT (Yang *et al.*, 2002). The *ptsG* RAT recognized by GlcT is most

different from all other RAT structures, and neither this RAT nor GlcT are involved in any cross-talk (Schilling *et al.*, 2004). In this work, we identified determinants that result in the regulatory specificity of the four distinct antitermination systems.

Materials and Methods

Bacterial strains and growth conditions

The *B. subtilis* strains used in this study are shown in Table 10. All *B. subtilis* strains are derivatives of the wild type strain 168. Strains used in the cause of site-directed mutagenesis studies are listed in Tables 11 and 12. These strains were all derived by transformation from the basal mutant strains listed in Table 10. *Escherichia coli* DH5α and BL21 (DE3) (Sambrook *et al.*, 1989) were used for cloning experiments and for expression of recombinant proteins, respectively.

B. subtilis was grown in SP medium or in CSE minimal medium (Faires *et al.*, 1999). The media were supplemented with auxotrophic requirements (at 50 mg/l), carbon sources and inducers as indicated. *E.coli* was grown in Luria–Bertani medium (LB medium) and transformants were selected on plates containing ampicillin (100 μg/ml). LB and SP plates were prepared by the addition of 17 g Bacto agar/l (Difco) to LB or SP medium, respectively.

Transformation and characterization of the phenotype

B. subtilis was transformed with plasmid DNA according to the two-step protocol described previously (Kunst and Rapoport, 1995). Transformants were selected on SP plates containing kanamycin (Km 5 μg/ml), chloramphenicol (Cm 5 μg/ml), spectinomycin (Spc 100 μg/ml), or erythromycin plus lincomycin (Em 1 μg/ml and Lin 10 μg/ml).

In *B. subtilis*, amylase activity was detected after growth on SP medium supplemented with 10 g hydrolyzed starch/l (Connaught). Starch degradation was detected by sublimating iodine onto the plates.

Quantitative studies of *lacZ* expression in *B. subtilis* in liquid medium were performed as follows: cells were grown in CSE medium supplemented with the carbon sources indicated. Cells were harvested at OD_{600} 0.6–0.8. Cell extracts were obtained by treatment with lysozyme and DNase. β-Galactosidase activities were determined as previously described using *o*-nitrophenyl-galactoside as a substrate (Kunst and Rapoport, 1995). One unit is defined as the amount of enzyme which produces 1 nmol of *o*-nitrophenol per min at 28°C.

DNA manipulation

Transformation of *E. coli* and plasmid DNA extraction were performed using standard procedures (Sambrook *et al.*, 1989). Restriction enzymes, T4 DNA ligase and DNA polymerases were used as recommended by the manufacturers. DNA fragments were purified from agarose gels using the QIAquick gel extraction kit (Qiagen®, Hilden, Germany). *Pfu* DNA polymerase was used for the PCR as recommended by the manufacturer.

Table 10. *B. subtilis* strains used in this study

Strain	Genotype	Source[a]
168	*trpC2*	Laboratory collection
BGW10	*trpC2 lys-3* Δ*licTS* ::*erm*	Krüger and Hecker, 1995
GM1112	*sacXY*Δ*3 sacB*Δ*23 sacT*Δ*4 bglP*::Tn*10 erm amyE*::(*sacB-lacZ phl*)	Le Coq *et al.*, 1995
QB5435	*trpC2* Δ*ptsG*::*cat*	Stülke *et al.*, 1997
QB5448	*trpC2 amyE*::(ΔLA *ptsG´-´lacZ aphA3*)	Stülke *et al.*, 1997
GP109	*trpC2 ΔglcT8 amyE*::(ΔLA *ptsG´-´lacZ aphA3*)	Bachem and Stülke, 1998
GP150	*trpC2 ΔglcT8 amyE*::(*´lacZ cat*)	Langbein *et al.*, 1999
GP385	*trpC2 amyE*::(ΔLA *ptsG-R1´-´lacZ aphA3*)	Schilling *et al.*, 2004
GP425	*trpC2* Δ*sacY*::*cat*	see Materials and Methods
GP427	*trpC2* Δ*licTS*::*erm*	BGW10 → 168
GP429	*trpC2* Δ*sacT*::*spc*	see Materials and Methods
GP430	*trpC2* Δ*sacY*::*cat* Δ*sacT*::*spc*	GP429 → GP425
GP431	*trpC2* Δ*licTS* ::*erm*Δ *sacT*::*spc*	GP429 → GP427
GP432	*trpC2* Δ*licTS* ::*erm* Δ*sacY*::*cat*	GP425 → GP427
GP433	*trpC2* Δ*licTS* ::*erm* Δ*sacY*::*cat sacT*::*spc*	GP429 → GP432
GP437	*trpC*2 *amyE*::(*sacB-lacZ aphA3*)	pGP564 → 168
GP470	*trpC2* Δ*ptsG*::*cat bglP*::Tn*10 erm*	GM1112 → QB5435

[a] Arrows indicate construction by transformation.

The combined chain reaction and the multiple mutation reaction were performed with *Pfu* DNA polymerase and thermostable DNA ligase (Ampligase®, Epicentre, Wisconsin, USA). DNA sequences were determined using the dideoxy chain termination method (Sambrook *et al.*, 1989). Chromosomal DNA of *B. subtilis* was isolated as described (Kunst and Rapoport, 1995).

Construction of sacT and sacY mutant strains by allelic replacement

To construct *sacT* and *sacY* mutant strains, the long flanking homology PCR (LFH-PCR) technique was used (Wach, 1996). Briefly, cassettes carrying the *cat* and *spc* resistance genes were amplified from the plasmids pGEM-*cat* and pDG1726, respectively (Guérout-Fleury *et al.*, 1995; Youngman, 1990). DNA fragments of ~1000 bp flanking the regions to be deleted at their 5′ and 3′ ends were amplified. The 3′ end of the upstream fragment as well as the 5′ end of the downstream fragment extended into the gene(s) to be deleted in a way that all expression signals of genes up- and downstream of the targeted genes remained intact. The joining of the two fragments to the resistance cassette was performed in a second PCR as described previously (Jordan *et al.*, 2006). In these reaction we used the primer pairs cat-fwd (5′-CGGCAATAGTTACCCTTATTATCAAG)/cat-rev (5′-CCAGCGTGGACCGGCGAGG CTAGTTACCC) and spec-fwd/spec-rev (Jordan *et al.*, 2006) for the amplification and joining of the cat and spc cassettes, respectively. The PCR products were directly used to transform *B. subtilis*. The integrity of the regions flanking the integrated resistance cassettes was verified by sequencing PCR products of ~1000 bp amplified from chromosomal DNA of the resulting mutants. The resulting strains were GP425 (Δ*sacY*::*cat*) and GP429 (Δ*sacT*::*spc*).

Site-directed mutagenesis

Translational fusions of variants of the *ptsG* and *sacB* regulatory regions with the *lacZ* gene were constructed using the vector pAC7 (Weinrauch *et al.*, 1991) containing the kanamycin resistance gene *aphA3*. The plasmid harbours a *lacZ* gene without a promoter located between two fragments of the *B. subtilis amyE* gene. To construct a translational *sacB-lacZ* fusion the DNA upstream from the *sacB* gene [−464 to +15 nt relative to the translational start point of *sacB* (Steinmetz *et al.*, 1985)] was amplified by PCR using the primers OS49 (5′-AAAGAATTCGATCCTTTTTAACCCATCACATATAC) and OS50 (5′-TTTGGATCCT TTTTGATGTTCATCGTTCATGTC). The primers introduced *BamH*I and *EcoR*I cloning sites at the ends of the amplified fragment and created an in-frame translational fusion of the *lacZ* gene with the 5th codon of *sacB*. The PCR product was inserted into pAC7, both linearized with the same enzymes producing plasmid pGP437.

To study the effect of point mutations in the RAT sequences the following strategy was applied: a DNA fragment carrying the mutant form of the RAT was constructed by site-directed mutagenesis using either the combined chain reaction or the multiple mutation reaction (to introduce three or more mutations simultaneously) as outlined previously (Bi and

Stambrook, 1998; Hames *et al.*, 2005). Plasmids pGP66 (Stülke *et al.*, 1997) and pGP437 containing the *ptsG* and *sacB* promoter regions, respectively, served as templates. The mutagenic primers and the resulting plasmids are available upon request. The oligonucleotides JS11 (Stülke *et al.*, 1997)/IL5 (Langbein *et al.*, 1999) and OS49/OS50 (see above) were used as outer primers for *ptsG* and *sacB*, respectively. The final PCR products were purified and cut by *BamH*I and *Mfe*I (for *ptsG*) or *BamH*I and *EcoR*I (for *sacB*) sites introduced by the PCR primers. To introduce the constructed *lacZ* fusions into the chromosome of *B. subtilis*, competent cells of the wild type strain 168 were transformed with the plasmids carrying the respective mutations linearized with *Sca*I.

Construction of expression vectors for the RNA-binding domains of antiterminator proteins

A plasmid allowing the fusion of any protein to a Strep tag at the C-terminus was constructed as follows: First, the expression vector pET3C (Novagen) was digested with *Nde*I and *BamH*I. The insert containing a small multiple cloning site and the Strep tag was prepared by annealing the complementary oligonucleotides OS91 (5′-TATGGAGCTCGG ATCCTGGAGCCACCCGCAGTTCGAAAAATGATAGT) and OS92 (5′-GATCACTATC ATTTTTCGAACTGCGGGTGGCTCCAGGATCCGAGCTCCA). The resulting DNA fragment carries ends compatible with *Nde*I and *BamH*I. Upon ligation, the *Nde*I site was conserved whereas the *BamH*I site was lost. The resulting plasmid, pGP574, carries an IPTG-inducible promoter, a small cloning site (*Nde*I–*Sac*I–*BamH*I) for the insertion of the coding sequences, and the sequence encoding the Strep tag followed by two stop codons.

To fuse the RNA-binding domains of GlcT, LicT and SacT to a Strep tag at their C-termini, plasmids pGP575, pGP576 and pGP577 were constructed: DNA fragments corresponding to amino acids 1–60 of GlcT, and 1–57 of LicT and SacT were amplified by PCR using chromosomal DNA of *B. subtilis* QB5448 and the primer pairs OS93/OS94, OS95/OS96, and OS97/OS98, respectively (the primer sequences are available upon request). The PCR products were digested with *Nde*I and *BamH*I, and the resulting fragments were cloned into the expression vector pGP574 cut with the same enzymes.

Protein purification

E. coli BL21(DE3)/pLysS was used as host for the overexpression of recombinant proteins. Expression was induced by the addition of IPTG (final concentration 1 mM) to exponentially growing cultures (OD_{600} of 0.8). Cells were lysed using a french press. After lysis the crude

extracts were centrifuged at 15 000 *g* for 30 min and then passed over a Streptactin column (IBA, Göttingen, Germany). The recombinant protein was eluted with desthiobiotin (Sigma, final concentration 2.5 mM). After elution, the fractions were tested for the desired protein using 12.5% SDS-PAGE gels. The relevant fractions were combined and dialyzed overnight. Purified proteins were concentrated using Microsep™ Microconcentrators with a molecular weight cut-off of 3 kDa (Pall Filtron, Northborough, MA). The protein concentration was determined according to the method of Bradford using the Bio-rad dye-binding assay and BSA as the standard.

Assay of interaction between the RNA-binding domains and RAT RNA

To obtain templates for the *in vitro* synthesis of the *ptsG* RAT-RNA, the primers OS25/OS26 (Schilling *et al.*, 2004) were used to amplify a 99 bp PCR product using pGP66 or the plasmid carrying the desired mutation as template. Similarly, a 99 bp DNA fragment encompassing the *sacB* RAT was amplified using the oligonucleotides OS86 (5′-CCAAGTAATACGACTCACTATAGGCGAAAAGTAAATCGCGCG) and OS87 (5′-GTATACACTTTGCCCTTTACAC) and pGP437 or a mutant variant as template. The presence of a T7 RNA polymerase recognition site on primers OS25 and OS86 (underlined) allowed the use of the PCR product as a template for *in vitro* transcription with T7 RNA polymerase (Roche Diagnostics). The integrity of the RNA transcripts was analyzed by denaturating agarose gel electrophoresis (Ludwig *et al.*, 2001).

Binding of the RNA-binding domains to RAT RNA was analyzed by gel retardation experiments. The RAT RNA (in water) was denatured by incubation at 90°C for 2 min and renatured by dilution 1:1 with ice cold water and subsequent incubation on ice. Purified protein was added to the RAT-RNA and the samples were incubated for 10 min at room temperature in TAE buffer in the presence of 300 mM NaCl. After this incubation, glycerol was added to a final concentration of 10% (wt/vol). The samples were then analyzed on 10% Tris-acetate PAA gels.

Northern blot analysis

RNA was prepared by the modified ‘mechanical disruption protocol’ described previously (Ludwig *et al.*, 2001). Briefly, 20 ml of cells were harvested at the exponential phase. After mechanical cell disruption, the frozen powder was instantly resuspended in 3 ml lysis buffer [4 M guanidine isothiocyanate; 0.025 M sodium acetate, pH 5.3; 0.5% *N*-laurylsarcosine (wt/vol)]. Subsequently, total RNA was extracted using the RNeasy Mini kit (Quiagen,

Germany). Digoxigenin RNA probes specific for the *E. coli lacZ* gene were obtained by *in vitro* transcription with T7 RNA polymerase (Roche Diagnostics) using a PCR-generated fragment as templates. The primers used for PCR were SHU55 (5′-GTTTTACAACGTCGTGACTGG) and SHU56 (5′-CTAATACGACTCACTAT AGGGAGGTGTGCAGTTCAACCACCG). The reverse primers contained a T7 RNA polymerase recognition sequence. *In vitro* RNA labelling, hybridization and signal detection were carried out according to the manufacturer's instructions according to the instructions of the manufacturer (DIG RNA labelling kit and detection chemicals; Roche Diagnostics).

Uptake of radioactive glucose in vivo

B. subtilis strains were grown in CSE medium with glucose (10 g/l). Sugar uptake assays were performed as described previously (Stülke *et al.*, 1997). Exponentially growing cells were harvested at an OD_{600} of 0.6–0.8 and washed once with the incorporation medium. Labelled [^{14}C] glucose (184 mCi mmol^{-1}) and non-labelled glucose (final concentration 0.4 mM) were added. Samples were taken and treated as described (Stülke *et al.*, 1997).

Results

Analysis of the loop structures in the ptsG RAT

The *ptsG* RAT differs from all other RAT sequences recognized by antiterminator proteins of the BglG/SacY family in the structure of the lower loop (see Fig. 5.2). In a previous work, we have demonstrated that the insertion of one base into the lower loop of the *ptsG* RAT (the *ptsG*-R1 mutation, see Fig. 5.3) makes its structure similar to that recognized by the other antiterminator proteins and results in exclusive binding of LicT to this structure, whereas it is not bound by GlcT (Schilling *et al.*, 2004). From this result, it was concluded that structure rather than the nucleotide sequence is important for antiterminator protein-RAT recognition.

Since the antitermination proteins bind as dimers to the RAT, and LicT contacts different structures of the lower and the upper loop, we asked whether a RAT with an ‘inversion’ of the lower and upper RAT structures might be recognized by any of the antitermination proteins (see Fig. 5.3).

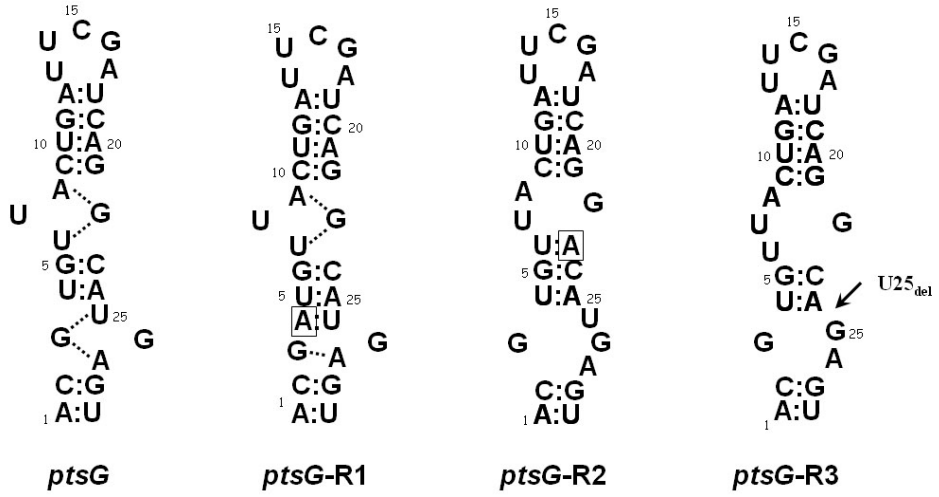

Figure 5.3. Predicted secondary structures of *ptsG*-RAT and the *ptsG*-RAT mutants *ptsG*-R1, *ptsG*-R2, and *ptsG*-R3. The insertions of an adenine at position 4 in the *ptsG*-R1 and at position 23 in the *ptsG*-R2 RAT mutants are boxed. The deletion of the base U25 in *ptsG*-R3 mutant is indicated by an arrow.

The activity of this mutant RAT, *ptsG*-R2, was assayed by analyzing the expression of a translational fusion of the mutated *ptsG* control region to a promoterless *lacZ* gene (see Table 11). While the presence of glucose in the growth medium resulted in a strong GlcT-dependent induction of the *ptsG* promoter in the wild type, salicin induced the *ptsG*-R1 promoter region in a LicT-dependent manner. In contrast, the *ptsG*-R2 promoter region did not allow expression of the *lacZ* fusion irrespective of the potential inducing carbohydrate present in the medium. Thus, this RAT is not bound by any of the antitermination proteins in *B. subtilis* (Table 11).

The *ptsG*-R1 structure was obtained by inserting an A after position 3 of the RAT sequence. This did not only create a lower loop structure similar to those present in RAT structures bound by LicT, SacT, and SacY, but did also generate an additional base pair between the lower and upper loops (see Figs. 5.1, 5.2, and 5.3). To rule out any effect of this extra base pair we constructed the *ptsG*-R3 RAT mutant by deleting the U at position 25. This results in a lower loop identical to that in the *ptsG*-R1 RAT, but separated from the upper loop by only 2 bp (Fig. 5.3).

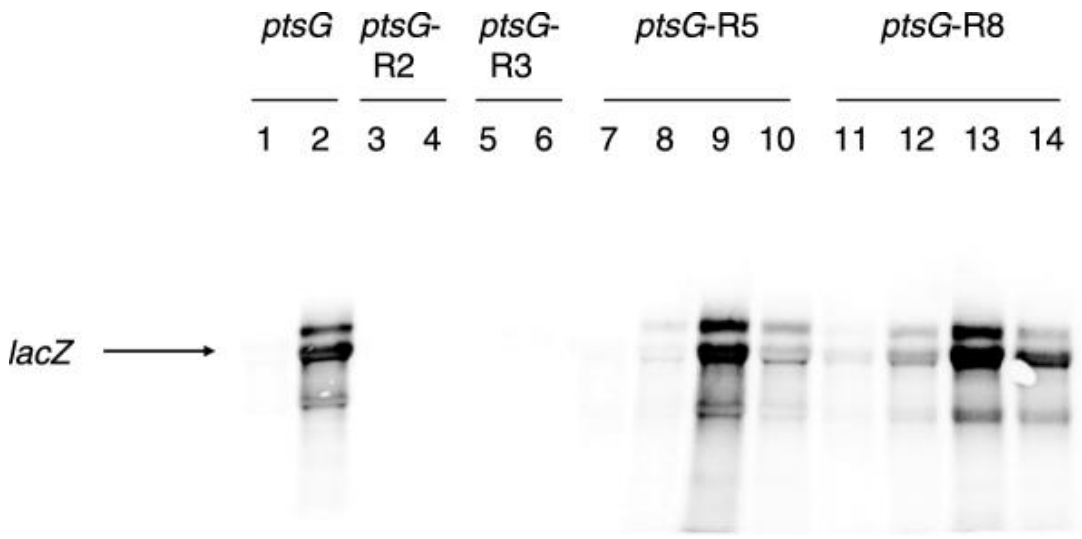

Figure 5.4. Northern Blot analysis of the expression of the *lacZ* gene under control of wild type and mutant RATs. Total RNA was separated by electrophoresis in 1.0% agarose gels and, after blotting, nylon membranes were hybridized to a riboprobe specific for *lacZ*. The mRNA corresponding to the *lacZ* gene is marked by an arrow. Note that a larger transcript was detected due to imperfect termination of *lacZ* transcription. 5µg RNA per lane were applied. The RNAs were isolated from the wild type strain QB5448 (lanes 1, 2), and the RAT mutant strains GP413 (*ptsG*-R2, lanes 3, 4), GP415 (*ptsG*-R3, lanes 5, 6), GP404 (*ptsG*-R5, lanes 7 to 10), and GP419 (*ptsG*-R8, lanes 11 to 14). The cultures for RNA isolation were grown in CSE minimal medium (lanes 1, 3, 5, 7, 11), in CSE medium supplemented with glucose (lanes 2, 4, 6, 8, 12), salicin (lanes 9, 13), or sucrose (lanes 10, 14).

The biological activity of this RAT mutant was determined by studying its effect on the expression of a *ptsG-lacZ* fusion. As shown in Table 11, the *ptsG*-R3 RAT did not confer induction under any of the conditions tested suggesting that this structure is not recognized by any of the antitermination proteins (see Discussion). The inability of GlcT to bind to the *ptsG*-R2 and *ptsG*-R3 RATs was verified by a northern blot analysis. The amounts of *lacZ* mRNA were compared in the wild type strain QB5448 and the two mutant strains GP413 and GP415. As can be seen in Figure 5.4, the *lacZ* mRNA was strongly induced in cells grown in the presence of glucose whereas no induction was observed in the two mutant strains. This result is in perfect agreement with those obtained by the reporter gene assays (Table 11).

Table 11. Effect of mutations in the *ptsG* RAT on recognition by the different antiterminator proteins.

Strain	RAT	Relevant genotype	ß-Galactosidase activity (U/mg protein)				
			CSE	CSE-Glc	CSE Suc (0.1%)	CSE Suc (2%)	CSE Sal
QB5448	*ptsG*	wild type	9	548	231	329	267
GP109	*ptsG*	*glcT*	4	5	2	8	9
GP387	*ptsG*	Δ*sacT* Δ*sacY*	10	327	85	166	335
GP389	*ptsG*	Δ*licT*	12	412	222	422	312
GP385	*ptsG*-R1	wild type	6	8	34	37	102
GP386	*ptsG*-R1	*glcT*	10	18	48	49	121
GP390	*ptsG*-R1	Δ*licT*	16	11	17	14	13
GP413	*ptsG*-R2	wild type	3	2	2	3	4
GP415	*ptsG*-R3	wild type	14	11	12	7	14
GP416	*ptsG*-R4	wild type	11	3	8	6	94
GP396	*ptsG*-R4	*glcT*	7	9	34	20	90
GP417	*ptsG*-R4	Δ*licT*	10	10	11	10	9
GP404	*ptsG*-R5	wild type	34	33	174	121	877

GP400	*ptsG*-R5	*glcT*	24	100	457	178	633
GP402	*ptsG*-R5	Δ*licT*	12	18	66	80	16
GP455	*ptsG*-R5	Δ*licT* Δ*sacY*	7	9	42	40	8
GP454	*ptsG*-R5	Δ*licT* Δ*sacT*	6	10	10	9	9
GP456	*ptsG*-R5	Δ*licT* Δ*sacT* Δ*sacY*	10	9	10	11	8
GP434	*ptsG*-R5	*glcT* Δ*licT*	9	8	-	-	6
GP436	*ptsG*-R5	*glcT* Δ*sacT*	32	70	-	-	-
GP435	*ptsG*-R5	*glcT* Δ*sacY*	33	73	-	-	-
GP408	*ptsG*-R6	wild type	34	16	45	50	702
GP399	*ptsG*-R6	*glcT*	19	50	239	156	514
GP409	*ptsG*-R6	Δ*licT*	14	16	20	20	12
GP464	*ptsG*-R7	wild type	37	32	36	37	60
GP419	*ptsG*-R8	wild type	147	188	157	238	1038
GP420	*ptsG*-R8	*glcT*	160	174	660	798	786
GP421	*ptsG*-R8	Δ*licT*	209	150	300	287	217

[a] Representative values of *lacZ* expression. All measurements were performed at least twice.

Contribution of individual bases to the recognition of the RAT sequence by antitermination proteins

The cognate RATs bound by LicT, SacT, and SacY are very similar to each other both in terms of structure and sequence (see Fig. 5.2). However, the *ptsG*-R1 RAT is recognized by LicT only and not by SacY or SacT. Therefore, we decided to introduce further mutations into the *ptsG*-R1 RAT that allow the evaluation of the contribution of individual nucleotides to protein–RNA recognition.

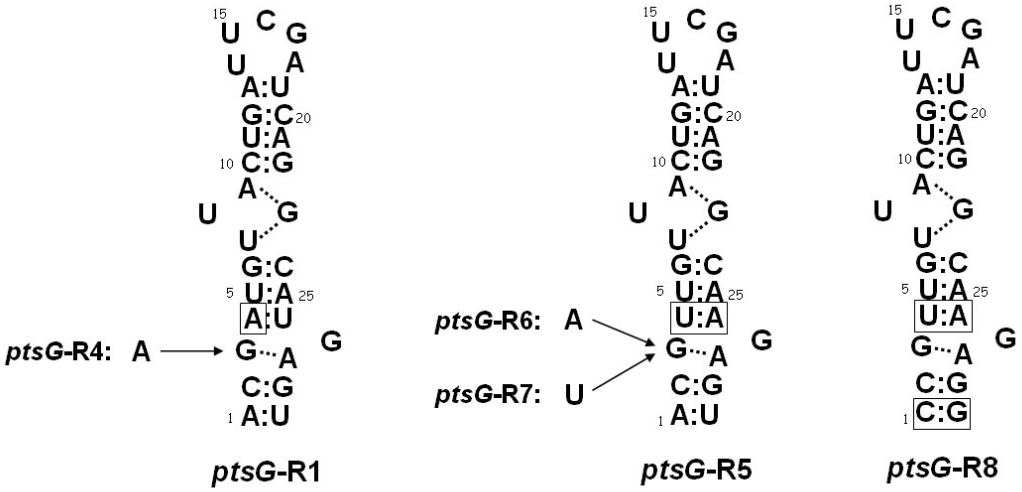

Figure 5.5. Secondary structures of the *ptsG*-R1, *ptsG*-R5, and *ptsG*-R8 RAT mutants. Bases that differ from *ptsG*-RAT are boxed.

The RATs recognized by LicT and SacT contain an A at position 3 in the lower loop rather than a G as in the *ptsG*-R1 RAT. Therefore, we exchanged the G3 for an A. The effect of this mutation, present in the *ptsG*-R4 RAT (see Fig. 5.5), was tested by the analysis of a *ptsG*-R4-*lacZ* fusion. This mutation resulted in a *lacZ* expression comparable to that observed with the *ptsG*-R1 RAT. As determined for *ptsG*-R1, the expression driven by the *ptsG*-R4 promoter region was completely dependent on a functional *licT* gene (Table 11). Thus, LicT is the only antiterminator protein binding to both the *ptsG*-R1 and *ptsG*-R4 RATs.

Another important difference between the *ptsG*-R1 RAT and all other RATs recognized by LicT, SacT, or SacY is the U:A base pair above the lower loop, which is A:U in *ptsG*-R1 (see Figs. 5.2 and 5.3). Previous results suggested that inversions of base pairs in the stems of the RAT are tolerated as long as the general structure is conserved (Schilling *et al.*, 2004). However, due to the strict conservation of the U:A pair in this position in all RATs except *ptsG*-R1, we addressed the effect of such a base pair inversion. As can be seen in Table 11, this inversion, present in the *ptsG*-R5 mutation (see Fig. 5.5), resulted in an

increased expression of the fusion under all conditions tested. However, the *ptsG*-R5 RAT conferred a strong induction in the presence of salicin, and this induction was dependent on the presence of the LicT antiterminator protein. In contrast, glucose did not induce this fusion suggesting that GlcT is unable to bind this RAT (Table 11).

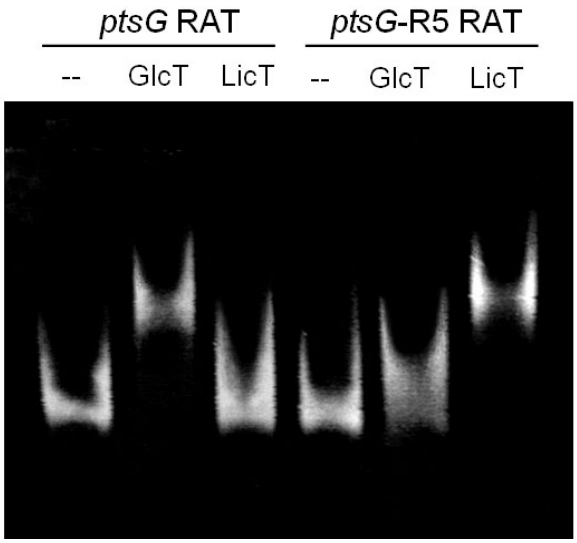

Figure 5.6. Electrophoretic mobility shift analysis of the interaction between the wild type *ptsG* and *ptsG*-R5 RATs, and the RNA-binding domains of GlcT or LicT. 100 pmol of the *ptsG* and *ptsG*-R5 RAT RNAs were used. GlcT or LicT (250 pmol) were added to the RNA as indicated prior to electrophoresis.

These observations were verified by an electrophoretic mobility shift analysis using the purified RNA-binding domains of GlcT and LicT (see Fig. 5.6). As reported previously (Schilling *et al.*, 2004), GlcT efficiently bound the *ptsG* RAT RNA. In contrast, LicT was unable to bind this RNA. In good agreement with the reporter gene analysis, LicT but not GlcT was capable of binding the *ptsG*-R5 RAT *in vitro* (see Fig. 5.6). For the *ptsG*-R1 RAT, induction in the presence of salicin and, to a lesser extent, sucrose, is strictly LicT-dependent (Table 11) (Schilling *et al.*, 2004). We tested therefore whether the induction of the *ptsG*-R5 RAT by sucrose was also due to binding of LicT. As mentioned above, a deletion of the *licT* gene resulted in loss of *ptsG*-R5 induction by salicin. However, the *licT* mutation did not abolish the induction by sucrose at the *ptsG*-R5 RAT (Table 11) suggesting that either SacT or SacY (or both) bind this RNA and cause antitermination. To test this possibility, we assayed the expression of the *ptsG*-R5-*lacZ* fusion in strains containing combinations of mutations of the three antiterminator genes. In a *licT sacY* double mutant, a slight reduction of sucrose induction of the *ptsG*-R5-*lacZ* fusion was observed as compared to the *licT* mutant strain (Table 11).

In contrast, induction was completely lost in the *licT sacT* and the *licT sacT sacY* double and triple mutant strains carrying deletions of two and of three antiterminator protein-encoding genes, respectively. From this result we may conclude that SacT can recognize the *ptsG*-R5 RAT in addition to LicT. A northern blot analysis of the *lacZ* mRNA confirmed the strong induction by salicin and, to a lesser extent, by sucrose, conferred by the *ptsG*-R5 RAT (see Fig. 5.4). Taken together, these data demonstrate that the U:A base pair just above the lower loop is important to facilitate binding of LicT to this structure (compare the high β-galactosidase activity to that driven by the *ptsG*-R1 fusion, Table 11) and to allow binding of SacT.

The role of the base at position 3 in the lower loop was also analyzed in the context of the *ptsG*-R5 RAT. However, as observed with *ptsG*-R4, only minor effects of a substitution of G3 by A were observed (Table 11, see *ptsG*-R6, see Fig. 5.5). Induction by salicin was slightly decreased, and the induction with sucrose was also completely dependent on LicT indicating that SacT did not bind the *ptsG*-R6 RAT. A substitution of G3 by U (*ptsG*-R7, as present in the *sacB* RAT recognized by SacY, see Fig. 5.5) resulted in loss of induction by sucrose and only weak induction upon the addition of salicin (15-fold reduction as compared to *ptsG*-R5, see Table 11). Taken together, these results indicate that the G at position 3 facilitates binding of the antitermination proteins. In contrast, an U at this position strongly diminishes binding by LicT. These conclusions are validated by an analysis of the *sacB* RAT (see below).

Binding of the antitermination proteins to their RAT targets allows the formation of otherwise non-favoured RAT structures and prevents concomitantly the formation of the transcription terminators. The relative stability of the RAT structures may therefore be important for the level of gene expression. The RATs recognized by LicT, SacT, and SacY contain two G:C base pairs in the bottom stem whereas the *ptsG* RAT contains a A:U and a C:G base pair at this position. It seemed therefore possible that the replacement of the A:U base pair by a C:G base pair would result in a more stable RAT structure and thus affect transcription. To test this idea, the *ptsG*-R8 RAT was constructed based on *ptsG*-R5 and analyzed (see Fig. 5.5, Table 11). While the *ptsG*-R5 RAT allowed only a weak basal expression in the absence of any inducer (CSE medium), a strongly increased basal expression was found for *ptsG*-R8 (34 versus 147 U of β-galactosidase). This was also reflected in a northern blot analysis of *lacZ* mRNA if expressed under the control of the *ptsG*-R5 and *ptsG*-R8 RAT (compare Fig. 5.4, lanes 7 and 11). The *ptsG*-R8-*lacZ* fusion was also induced by salicin, and the induced expression was the sum of read through (~150 U) and

real induction (~900 U, see *ptsG*-R5, Table 11). However, the increased read through might also result from a destabilization of the terminator even though an extra mutation was introduced in the terminator to restore base pairing.

Carbon catabolite repression interferes with the transport of glucose by BglP

The analysis of the *ptsG*-R5 RAT revealed that this structure is efficiently bound by LicT but not by GlcT. The disruption of the *glcT* gene in a strain carrying the *ptsG*-R5-*lacZ* fusion resulted in induction of β-galactosidase by salicin and sucrose (see Table 11) as expected due to the binding of LicT and SacT, respectively (see above). Surprisingly, glucose did also activate expression of this fusion in a *glcT* mutant strain. Since GlcT is not available in this mutant, LicT or SacT must be activated in the presence of glucose in the *glcT* mutant. To test this idea, we studied the activity of the *ptsG*-R5 control region in *glcT licT* or *glcT sacT* double mutants. As shown in Table 11, only a minor effect of the *sacT* deletion was observed, whereas the deletion of the *licT* gene resulted in complete loss of induction by glucose. Thus, glucose can activate LicT in a *glcT* mutant strain.

Two scenarios for the activation of LicT by glucose can be envisaged. First, there might be some non-specificity in BglP that results in the transport of glucose by this permease and the subsequent dephosphorylation and activation of the cognate antiterminator LicT. Second, there might be some cross-talk between the glucose permease PtsG and the LicT antiterminator that results in LicT activation upon glucose transport. Several lines of evidence demonstrate that the former possibility reflects the truth: (i) The glucose permease PtsG is not expressed in a *glcT* mutant strain and is therefore unable to activate LicT in a *glcT* mutant (Bachem and Stülke, 1998; Stülke *et al.*, 1997). (ii) BglP phosphorylates and thereby inactivates LicT in the absence of the substrate salicin, and this regulation would be dominant over any minor PtsG-dependent dephosphorylation of LicT (Lindner *et al.*, 1999; Tortosa *et al.*, 1997). (iii) To provide direct evidence for glucose uptake by BglG we measured the glucose transport of glucose-grown cells of a wild type strain (*B. subtilis* 168), a *ptsG* mutant (QB5435) and a *ptsG bglP* double mutant (GP470). As shown in Fig. 5.7, glucose was efficiently transported by the wild type strain (initial uptake rate 620 ± 110 pmol glucose per minute and OD_{600}), whereas a significant reduction was observed in the *ptsG* mutant (initial uptake rate 62 ± 5 pmol glucose per minute and OD_{600}).

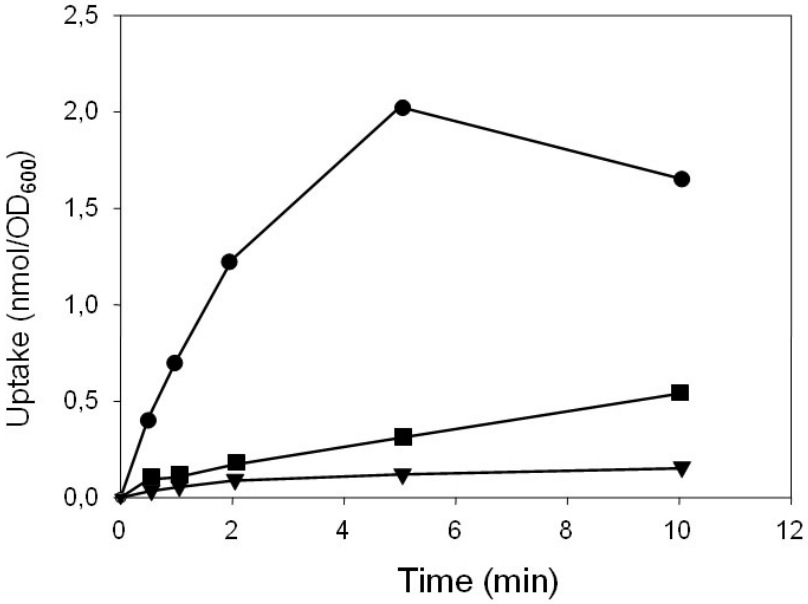

Figure 5.7. The implication of BglP in glucose uptake. For uptake measurements of radioactively labelled glucose the *B. subtilis* strains 168 (wild type, circles), QB5435 (*ptsG*, squares), and GP470 (*ptsG bglP*, triangles) were grown in CSE minimal medium supplemented with 0.5% of glucose.

These results are in good agreements with previous studies of glucose transport in *ptsG* mutants (Bachem *et al.*, 1997; Stülke *et al.*, 1997). In the *ptsG bglP* double mutant GP470, the transport of glucose was further reduced (see Fig. 5.7, initial uptake rate 24 ± 1 pmol glucose per minute and OD_{600}), confirming that BglP has some minor glucose transport activity which may explain glucose-dependent activation of LicT in the *glcT* mutant background (see Discussion).

Conversion analysis of the *sacB* RAT sequence towards new recognition specificities

The similarity of the *ptsG*-R6 RAT to that of the *sacPA* operon (see Figs. 5.1, 5.2 and 5.5) suggests that both RNA structures might be recognized by the same proteins. However, as shown in Table 11, *ptsG*-R6 is bound exclusively by LicT whereas the *sacPA* RAT is the target of SacT and is not recognized by LicT (Manival *et al.*, 1997). Similarly, the *ptsG*-R7 RAT which is poorly recognized by LicT but by none of the other antiterminator proteins resembles strongly the *sacB* RAT which is the target of SacY (see Figs. 5.1, 5.2, and 5.5). Thus, additional components seem to play a role in RAT-antiterminator protein recognition. To unravel these factors, we decided to perform an in-depth conversion analysis of the *sacB* RAT to mutate it and shift it gradually to sequences that are not longer recognized by SacB but rather by one of the three other family members. We chose the *sacB* RAT for this purpose

since *sacB* lacks any additional regulation by carbon catabolite repression (Steinmetz *et al.*, 1989). This analysis was aimed at the identification of bases that are responsible for the specificity for one or the other antiterminator protein.

Discrimination between SacY and SacT.

First, we determined the regulation mediated by the wild type *sacB* RAT. If the *lacZ* gene was expressed under the control of this RAT, induction was observed only in the presence of sucrose confirming that neither GlcT nor LicT bind the *sacB* RAT. Induction by sucrose occurred both at low and high sucrose concentrations which activate SacT and SacY, respectively. Indeed, induction at a low sucrose concentration was lost in the *sacT* mutant. In the *sacY* mutant strain, induction was still visible at both concentrations suggesting that SacT is active under both conditions. In the *sacT sacY* double mutant, the *sacB* RAT-terminator couple did not allow induction under all the condition tested (Table 12).

The *ptsG*-R5 RAT, which is recognized by LicT and SacT, closely resembles the *sacB* RAT but contains a G at position 3 rather than a U as in the *sacB* RAT. We constructed therefore the *sacB*-R1 RAT by replacing U3 by a G (see Fig. 5.8). This single mutation resulted in a significant specificity shift. The *sacB*-R1 RAT was not longer a target for SacY, whereas the activation of SacT allowed a higher β-galactosidase expression as compared to the wild type *sacB* RAT (Table 12).

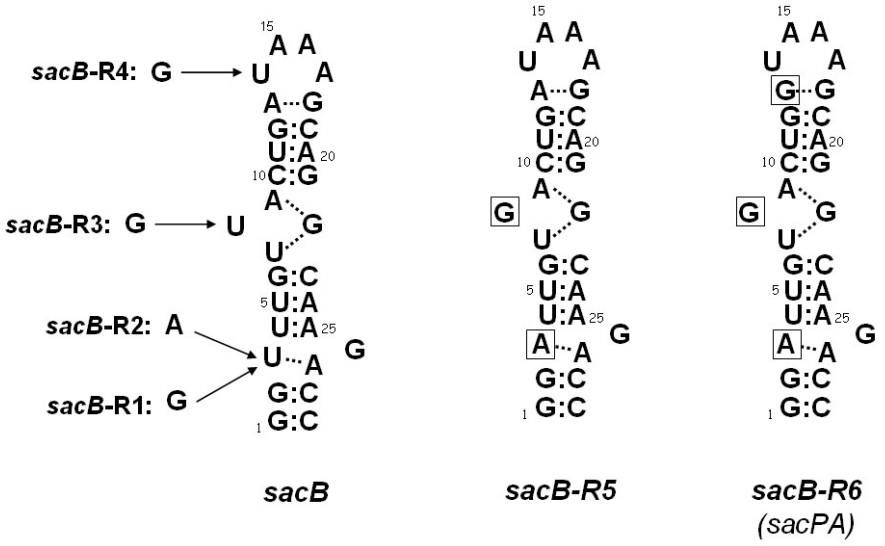

Figure 5.8. **Secondary structures of the wild type *sacB*, *sacB*-R5, and *sacB*-R6 mutant RATs.** Bases that differ from *sacB* RAT are boxed in the *sacB*-R5 double- and in the *sacB*-R6 triple mutant. The *sacB*-R1 to R4 mutant RATs are indicated as single base exchanges into *sacB* RAT. The *sacB*-R6 triple mutant RAT is identical to the *sacPA* wild type RAT (see Fig. 5.2).

Moreover, the *sacB*-R1 RAT allowed LicT-dependent induction by salicin. These results are in good general agreement with the observed affinity of the similar *ptsG*-R5 RAT for SacT and LicT, however, the preference for the two antitermination proteins was inverse. The *sacPA* RAT, which is the cognate target of SacT, contains also a purine base at position 3, i.e. an A (see Fig. 5.2). Therefore, the *sacB*-R2 containing an A at position 3 was constructed (see Fig. 5.8). The presence of this RAT conferred induction by sucrose but neither by salicin nor glucose (Table 12). Thus, this RAT is not bound by LicT. To distinguish whether it is recognized by SacY or SacT we analyzed the expression driven by the *sacB*-R2-*lacZ* fusion in *licT sacT* and *licT sacY* double mutant strains. As shown in Table 12, the *sacB*-R2 RAT is efficiently bound by SacT at both low and high sucrose concentrations whereas it is not recognized by SacY. The results obtained with the *sacB*-R1 and -R2 mutants suggest that the U at position 3 of the RAT is important for recognition by SacY. In contrast, SacT tolerates all three tested bases at this position.

The data presented above demonstrate that U3 is important for SacY binding in the context of the *sacB* RAT. However, since SacY is capable of recognizing the *sacPA* RAT (Manival *et al.*, 1997), it seems to be able to accept bases different from U at position 3. To address this question we exchanged the three bases in the *sacB* RAT that are different from the *sacPA* RAT. A replacement of U8 in the middle loop by a G as in *sacPA* (*sacB*-R3, see Fig. 5.8) resulted in loss of binding by SacY whereas SacT bound this RAT as judged from loss of sucrose induction in the *sacT* mutant (see Table 12). Both the *sacB* and *sacPA* RATs contain a UAAA tetraloop at the top. This loop is flanked by A-G and G-G pairs in *sacB* and *sacPA*, respectively (see Fig. 5.2). Therefore, we constructed the *sacB*-R4 mutant RAT with a G-G pair at the bottom of the top loop (A13G exchange, see Fig. 5.8). This mutation did not affect binding by SacY and SacT as compared to the wild type *sacB* RAT. Moreover, it did not confer induction by salicin indicating that it is no target for LicT (see Table 12). A combination of the two mutations of *sacB*-R2 and *sacB*-R3 (U3A and U8G) present in the *sacB*-R5 RAT (Fig. 5.8) resulted in enhanced induction by sucrose which was exclusively dependent on SacT as determined using a *sacT* mutant strain (Table 11). Indeed, the *sacB*-R5 RAT was efficiently bound by the RNA-binding domain of SacT (see Fig. 5.9).

Table 12. Conversion analysis of the *sacB*-RAT.

Strain	RAT	Relevant genotype	ß-Galactosidase activity (U/mg protein)[a]				
			CSE	CSE-Glc	CSE Suc (0.1%)	CSE Suc (2%)	CSE Sal
GP437	*sacB*	wild type	9	6	54	78	5
GP440	*sacB*	Δ*sacT*	5	4	5	28	4
GP438	*sacB*	Δ*sacY*	6	5	50	49	6
GP441	*sacB*	Δ*sacT* Δ*sacY*	6	3	5	4	3
GP461	*sacB*-R1	wild type	7	8	132	104	35
GP465	*sacB*-R1	Δ*sacT* Δ*licT*	5	4	3	4	2
GP466	*sacB*-R1	Δ*sacY* Δ*licT*	5	5	80	74	3
GP463	*sacB*-R1	Δ*sacT* Δ*sacY*	11	8	10	12	28
GP462	*sacB*-R1	Δ*licT*	7	9	85	82	5
GP460	*sacB*-R2	wild type	7	6	93	145	8
GP472	*sacB*-R2	Δ*sacY* Δ*licT*	3	5	97	112	5
GP471	*sacB*-R2	Δ*sacT* Δ*licT*	4	3	4	7	3
GP519	*sacB*-R3	wild type	5	2	96	88	4
GP521	*sacB*-R3	Δ*sacT*	4	-	4	9	4
GP540	*sacB*-R4	wild type	8	8	70	117	7
GP541	*sacB*-R4	Δ*sacT*	7	8	9	35	7
GP542	*sacB*-R4	Δ*sacT* Δ*sacY*	4	3	3	4	2
GP520	*sacB*-R5	wild type	4	3	216	208	6
GP522	*sacB*-R5	Δ*sacT*	4	2	5	7	2

GP537	*sacB*-R6	wild type	4	6	349	463	11
GP538	*sacB*-R6	Δ*sacT*	6	5	6	22	9
GP539	*sacB*-R6	Δ*sacT* Δ*sacY*	3	4	4	4	3
GP476	*sacB*-R7	wild type	3	6	176	145	92
GP544	*sacB*-R7	Δ*sacT*	5	4	7	8	136
GP484	*sacB*-R7	Δ*licT*	3	2	195	185	2
GP536	*sacB*-R7	Δ*sacT* Δ*licT*	2	2	3	5	2
GP477	*sacB*-R8	wild type	6	7	248	294	218
GP486	*sacB*-R8	Δ*licT*	2	1	305	188	2
GP487	*sacB*-R8	Δ*sacT*	5	4	9	8	376
GP480	*sacB*-R9	wild type	7	11	280	253	510
GP492	*sacB*-R9	Δ*licT*	3	4	291	189	3
GP493	*sacB*-R9	Δ*licT* Δ*sacT*	4	2	3	4	2
GP494	*sacB*-R9	Δ*licT* Δ*sacY*	3	3	213	193	3
GP444	*sacB*-R10	wild type	6	3	24	18	10
GP543	*sacB*-R10	Δ*sacT*	5	-	5	8	12
GP446	*sacB*-R11	wild type	48	32	43	34	23
GP448	*sacB*-R12	wild type	5	4	4	6	4
GP450	*sacB*-R13	wild type	28	95	112	129	33
GP453	*sacB*-R13	Δ*licT* Δ*sacT* Δ*sacY*	48	86	104	80	28
GP451	*sacB*-R13	*glcT*	31	33	20	36	18

[a] Representative values of *lacZ* expression. All measurements were performed at least twice.

In contrast, the wild type *sacB* RAT was only weakly bound by SacT. These observations are in very good agreement with the high SacT-dependent induction of gene expression mediated by *sacB*-R5 as compared to induction conferred by the wild type *sacB* RAT. An additional mutation of the base pair at the bottom of the top loop (A13G) made the resulting *sacB*-R6 RAT identical to that of *sacPA*, but in a sequence context of *sacB* (see Figs. 5.2 and 5.8).

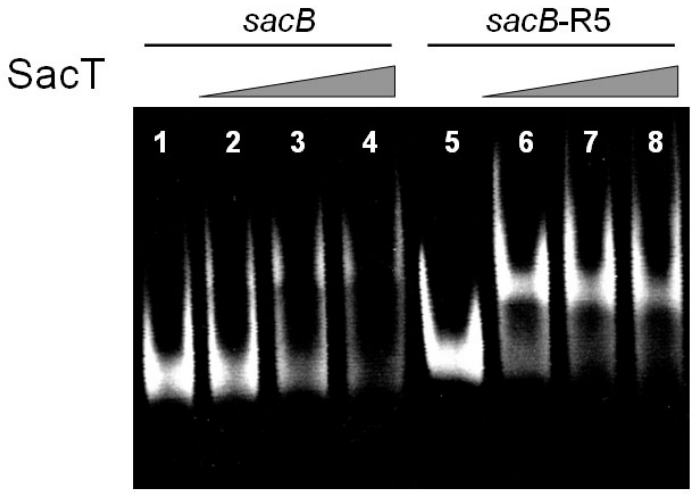

Figure 5.9. Electrophoretic mobility shift analysis of the interaction between the *sacB* and *sacB*-R5 RATs, and the RNA-binding domain of SacT. Lanes 1-4 and 5-8 contain 100 pmol of *sacB* and *sacB*-R5 RAT RNAs, respectively. Increasing concentrations of SacT were added to the RNA in lanes 2-4 and 6-8 prior to electrophoresis. Aliquots of 75, 150 and 300 pmol SacT were used.

As expected, this RAT is most efficiently recognized by SacT. In the *sacT* mutant, only a very weak induction by sucrose was observed which was lost in the *sacT sacY* double mutant strain (Table 12). Thus, the A3 and G8 do both discriminate against binding by SacY. However, as demonstrated using the *sacB*-R6 RAT, the G-G base pair at the bottom of the top loop seems to weaken this discrimination and does thus allow weak binding by SacY.

Discrimination between SacY and LicT.

The *sacB* RAT differs from the *licS* and *bglPH* RATs that are the cognate targets of LicT by two bases in the lower loop. Additionally, the *bglP* RAT contains a C-G base pair at the bottom of the top loop and a GAAA tetraloop at the top (see Fig. 5.2). The first step in the conversion of the *sacB* RAT to a structure expected to be recognized by LicT was the *sacB*-R2 mutation (U3A) described above. This RAT was bound by SacT but not by LicT (see Table 12). With the introduction of a second mutation in the lower loop (G26A) the resulting *sacB*-R7 RAT was identical to that of *licS* (see Figs. 5.2 and 5.10).

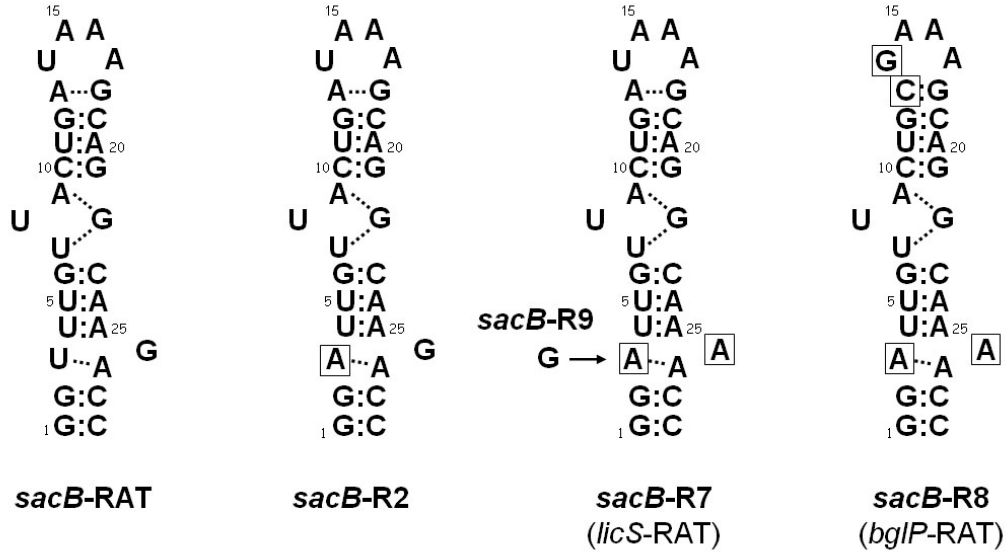

Figure 5.10. Secondary structures of the wild type *sacB*, *sacB*-R2, *sacB*-R7 and *sacB*-R8 RATs. Bases that differ from the *sacB* RAT are boxed. The mutations introduced into the *sacB* RAT convert it gradually to the *licS* (*sacB*-R7) and the *bglP* RATs (*sacB*-R8).

The determination of β-galactosidase regulation conferred by this RAT demonstrated induction not only by sucrose but also by salicin (see Table 12). Induction by salicin was completely lost in a LicT mutant whereas sucrose induction was lost in the *sacT* mutant strain. Binding of LicT to the *sacB*-R7 RAT was verified by an electrophoretic mobility shift assay. While the RNA-binding domain of LicT was unable to retard the wild type RAT of *sacB*, the *sacB*-R7 RAT was bound by this protein (see Fig. 5.11). Thus, the A at position 26 is an important feature that makes the RAT a target for LicT. A mutation of 2 bp affecting the top loop converts the *licS*-RAT (*sacB*-R7) to the *bglP* RAT (*sacB*-R8). These mutations increase the affinity of both LicT and SacT as inferred from β-galactosidase activities of the *sacB*-R8-*lacZ* fusion strains (see Table 12). This finding is in good agreement with the observed stronger salicin-dependent induction of the *bglPH* operon as compared to the *licS* gene by the antiterminator LicT (Schnetz *et al.*, 1996). As observed for the *sacB*-R7 RAT, the *sacB*-R8 RAT was bound by the RNA-binding domain of LicT *in vitro* (see Fig. 5.11). The importance of A26 for recognition by LicT is underlined by the analysis of the *sacB*-R9 RAT in which the A at position 3 (present in *sacB*-R7 and the cognate targets of LicT) is replaced by a G (see Fig. 5.10).

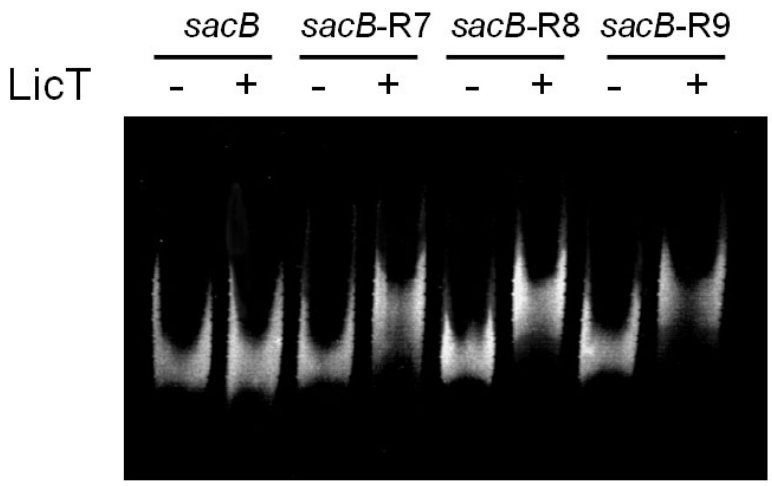

Figure 5.11. Electrophoretic mobility shift analysis of the interaction between the wild type *sacB* and several mutant RAT RNAs (see Fig. 5.9) with the RNA-binding domain of LicT. In all lanes, 100 pmol of RNA were used. In the lanes labelled with „+“, 250 pmol of LicT were added prior to electrophoresis.

As shown in Table 12, the presence of this RAT allows even higher LicT-dependent induction by salicin. Again, this RAT was recognized by LicT *in vitro* (see Fig. 5.11). However, this mutation did not affect binding specificity since *sacB*-R9 was also a target for SacT. Thus, for LicT and SacT, position 3 seems to be most important to maintain the proper RAT structure, whereas the opposing A at position 26 is important for allowing efficient binding by LicT.

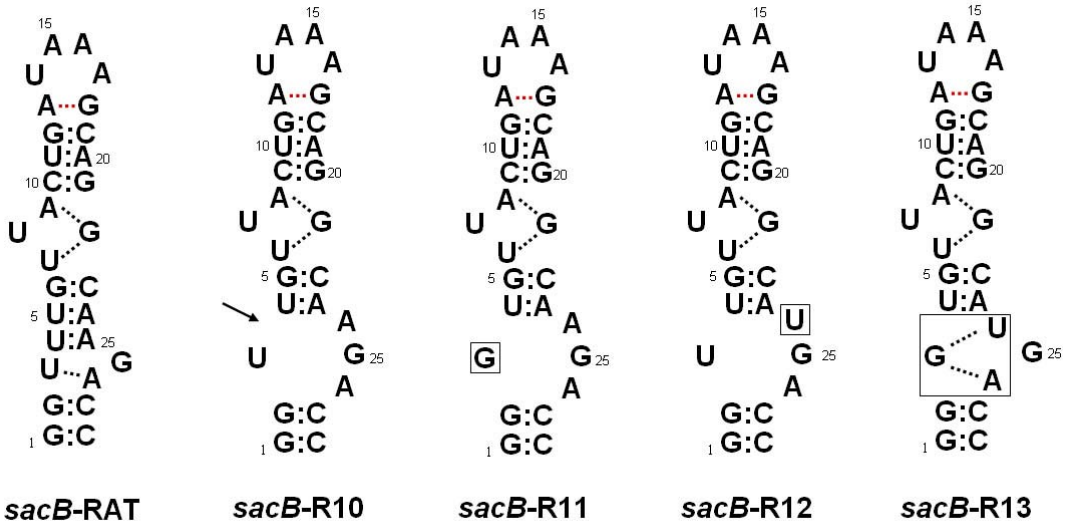

Figure 5.12. Gradual conversion of the lower loop region of the wild type *sacB* RAT to a stucture similar to that found in the *ptsG* RAT. A deletion of an U in *sacB*-R10 is indicated by an arrow. Bases that differ from *sacB*-RAT are boxed. The lower loop region in the *sacB*-R13 mutant is identical to that of the *ptsG* RAT (see Fig. 5.2)

Discrimination between SacY and GlcT.

The *ptsG* RAT is most different from all other RAT structures in *B. subtilis* due to the triple base pairing in the lower loop region (see above, Fig. 5.2; Schilling *et al.*, 2004). It has been proposed that this distinct structure rather than the details of the actual nucleotide sequence is important for recognition by GlcT. To verify this assumption we introduced mutations into the *sacB* RAT that made its structure gradually more similar to that of the *ptsG* RAT. In a first step, the U at position 4 was deleted (Fig. 5.12). This mutation is a reversal of the conversion from the *ptsG* to the *ptsG*-R1 RAT (insertion of one base at position 4, see Fig. 5.3), but in the context of the *sacB* RAT. The resulting *sacB*-R10 RAT allowed a very weak SacT-dependent induction by sucrose. In contrast, this RAT was not at all recognized by GlcT as concluded from the absence of induction by glucose. Here, the U3 might form a base pair with either A24 or A26 thus forming a structure weakly recognized and sufficiently stabilized by SacT to allow antitermination.

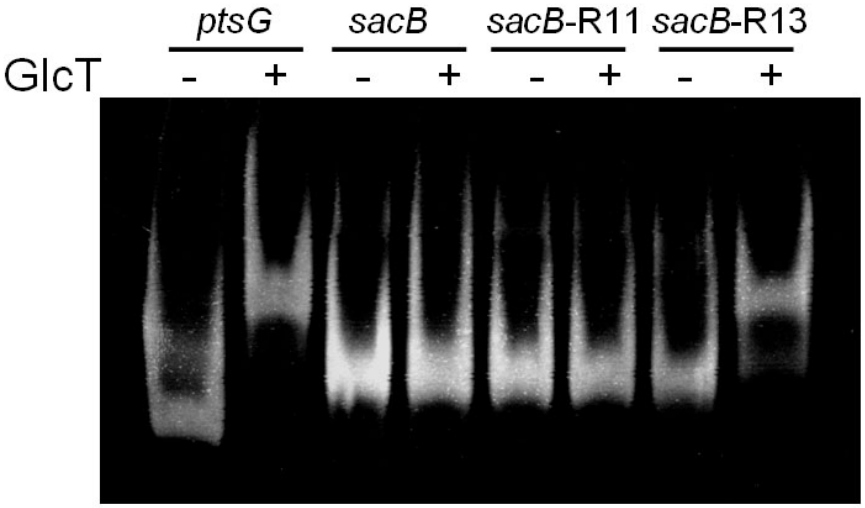

Figure 5.13. Electrophoretic mobility shift analysis of the interaction between the *ptsG* and *sacB* RAT-RNAs, several *sacB* mutant RAT-RNAs and the RNA-binding domain of GlcT. In all lanes, 100 pmol of RAT RNA were used. In the lanes labelled with „+“, 250 pmol of GlcT were added prior to electrophoresis.

The *sacB*-R10 mutant RAT was then parent to two further variants. In the *sacB*-R11-RAT, the U3 was replaced by a G (see Fig. 5.12). In the *ptsG* RAT, a G at this position contacts the nucleotides at positions 24 (U) and 26 (A). However, this does not seem to be the case with the two adenines in *sacB*-R11. This RAT does not confer induction to the *lacZ* gene, however, the read through was somewhat increased even in the absence of any inducer (see Table 12). A replacement of A24 present in *sacB*-R10 by a U (as in the *ptsG* RAT at this position)

resulted in complete loss of expression of the reporter gene (see Table 12, *sacB*-R12, see Fig. 5.12). By replacing the U3 of *sacB*-R12 by a G, we obtained a lower loop that is identical to that found in the *ptsG* RAT (*sacB*-R13, Fig. 5.12). Indeed, the *sacB*-R13-*lacZ* fusion was induced by glucose and sucrose (see Table 12). Since salicin and sucrose are known to activate GlcT (Langbein *et al.*, 1999), we tested the expression of this fusion in a *licT sacT sacY* triple mutant as well as in a *glcT* mutant strain. As expected, the combined deletion of *licT*, *sacT*, and *sacY* did not affect the induction by any of the sugars whereas no induction was observed in the *glcT* mutant strain (see Table 12). Thus, this RAT is exclusively recognized by GlcT. To verify this observation we performed electrophoretic mobility shift assays in the presence of the RNA-binding domain of GlcT using the RATs of *ptsG* and *sacB* as controls as well as the *sacB*-R11- and *sacB*-R13-RATs (see Fig. 5.13). As previously observed, GlcT is capable of binding its cognate *ptsG* RAT. In contrast, the *sacB* RAT was not recognized by GlcT. Similarly, the *sacB*-R11-RAT was not retarded. As expected from the transcription regulation conferred by GlcT and the *sacB*-R13 RAT, an RNA-fragment containing this RAT was bound by the RNA-binding domain of GlcT. These findings confirm the important role of G3, U24, and A26 for the formation of the structure in the lower loop and in GlcT binding.

Discussion

Several distinct mechanisms contribute to the specificity of the four antitermination systems present in *B. subtilis*. These include, first, the sugar permeases and their interactions with their substrates and with the cognate antitermination proteins. Second, carbon catabolite repression limits the conditions under which certain systems are expressed and the antiterminator proteins active. Finally, the interaction between the antiterminator proteins and the RAT RNAs makes a major contribution to regulatory specificity.

The sugar permeases of the PTS can transport and phosphorylate only one substrate, or they can exhibit a relaxed specificity, i.e. they may transport more than one sugar. The glucose permease PtsG is known to transport sucrose and salicin in addition to glucose thus explaining the induction of *ptsG* expression by these sugars (Langbein *et al.*, 1999) (see Table 11). A relaxed specificity has also been observed for the GlcB permease from *Staphylococcus carnosus* which is also capable of transporting salicin in addition to glucose (Christiansen and Hengstenberg, 1999). Similarly, the β-glucoside permease BglP is able to transport glucose, although with a low efficiency (see Fig. 5.7). In wild type strains, the *bglP* gene is strongly repressed in the presence of glucose, thus, this relaxed specificity has no

biological consequence. In contrast, the two sucrose permeases seem to be highly specific for sucrose, and the SacX permease is regarded as being inactive since it does not contribute to sucrose transport (Kunst *et al.*, 1974). All experiments with the different antitermination systems published so far did not provide any indication that a permease might interact with a non-cognate antiterminator protein. This might reflect the parallel evolution of the permeases and their targets, the PRD-I domains of the antitermination proteins (Greenberg *et al.*, 2002). Indeed, the control of the antitermination proteins by the corresponding sugar-specific permeases works beyond the species barrier as shown for *B. subtilis* LicT in *E. coli* or *S. carnosus* GlcT in *B. subtilis* (Knezevic et al., 2000; Schnetz et al., 1996).

Bacteria use carbon sources in a hierarchical order, i.e. those that are most easily metabolized with a maximum yield of energy are preferred. In *B. subtilis*, glucose is the preferred carbon source, and the presence of glucose prevents the activity of many enzymes as well as the expression of genes and operons that are required for the utilization of alternative carbon sources. Among the genes studied here, only *ptsG* is induced by glucose (via antitermination) whereas *sacPA*, *bglPH*, and *licS* are repressed. This repression is achieved by two independent mechanisms: First, the CcpA repressor protein binds target sites in the promoter regions of these genes and prevents their expression if glucose is present. Second, the antiterminator proteins SacT and LicT, which are required for the expression of these genes, are inactive as long as glucose is present. In the absence of glucose they are phosphorylated at their PRD-II and thereby activated by HPr (Krüger *et al.*, 1996; Lindner *et al.*, 1999; Stülke *et al.*, 1998). These two mechanisms result in the absence of the BglP permease if glucose is available. Only in the *glcT* mutant strain, if glucose is unable to exert carbon catabolite repression since it can not efficiently be transported into the cell, the *bglP* gene can escape carbon catabolite repression and the BglP protein may exert a weak glucose transport activity which is sufficient for the activation of LicT (see Fig. 5.7, Table 11).

A major specificity determinant in transcription regulation by the four antiterminator proteins is the RNA-protein interaction. As shown previously, the loop structures of the RATs are crucial for the specific recognition (Schilling *et al.*, 2004; Yang *et al.*, 2002). In this work, we have identified all the factors that determine the specificity for any of the four antiterminator proteins (see Fig. 5.14).

GlcT is unique in that it requires two identical and nearly symmetrical triple base pairings in the RAT. In contrast, LicT, SacT, and SacY bind RAT structures that resemble the upper triple base pair in the *ptsG* RAT, but differ significantly in the lower loop.

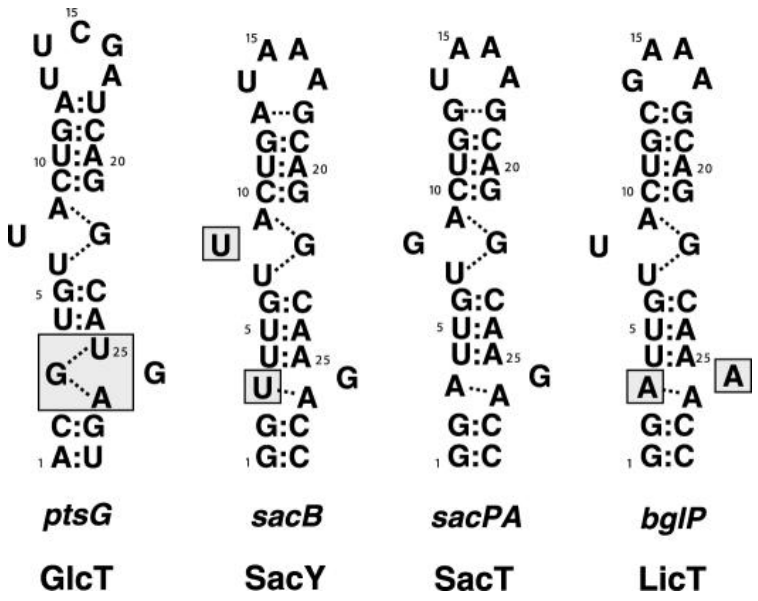

Figure 5.14. Summary of the relevant features that cause protein–RNA recognition specificity of the RAT RNAs of *B. subtilis*. Boxes indicate nucleotides that switch specificity towards the appropriate antiterminator proteins. For *ptsG*, the structure of the lower loop region discriminates it from all the other RAT-structures and facilitates exclusive GlcT binding. SacY binding depends on the Us found at positions 3 and 8 in the *sacB* RAT. No specificity determinant could be found for SacT, as it binds to all RAT structures except for that of *ptsG*. SacT signaling specificity is achieved by the control of the protein's activity by dual PTS-dependent phosphorylation. The LicT targets (*bglP*, *licS*) are characterized by the essential As found at positions 3 and 26.

A mutation that changes the lower loop of the *ptsG* RAT towards that found in the RAT recognized by the other antiterminator proteins (*ptsG*-R1, *ptsG*-R5, see Table 11) prevented recognition by GlcT and allowed binding by SacT and LicT. Similarly, a mutation of the *sacB* RAT, which affected the lower loop and allowed the formation of a triple base pair, resulted in loss of SacY and SacT binding whereas GlcT recognized such a structure (*sacB*-R13, see Figs. 5.12 and 5.13; Table 12).

LicT recognizes structures that are highly similar to the targets of SacT and SacY. An inspection of the RAT structures reveals that the LicT targets are unique in having an A at position 26 (see Fig. 5.2). The importance of this position is underlined by our mutation analysis of the *sacB* RAT. The importance of A26 for recognition by LicT is also supported by the structure of the LicT–RAT complex. There are several contacts of LicT with A26 and the sugar phosphate backbone in its immediate neighbourhood (Yang *et al.*, 2002). The differential role of guanine and adenine residues for recognition of nucleic acids by proteins is well established (Nobeli *et al.*, 2001). Moreover, the data indicate that the A at position 26 is necessary but not sufficient for LicT binding. In addition, a purine base is required at the opposing position 3 of the RAT. This is in good agreement with a previous study (Aymerich

and Steinmetz, 1992). Interestingly, a G at position 3 (*sacB*-R9) allows much higher LicT-, but also SacT-dependent antitermination as compared to a similar RAT containing an A at this position (*sacB*-R7).

There are conflicting reports on the recognition of the *sacB* RAT by SacT (Aymerich and Steinmetz, 1992; Manival *et al.*, 1997). We observed that the *sacB* RAT is recognized by both SacT and SacY. Interestingly, the SacT-dependent induction of *sacB* is stronger than the induction mediated by the cognate antiterminator, SacY. Thus, SacT induces both the *sacPA* operon and the *sacB* gene encoding levansucrase. In contrast, SacY exerts only a very minor effect at the *sacPA* RAT (identical to *sacB*-R6). Since SacT is active at both high and low sucrose concentrations whereas SacY is active only in the presence of large amounts of sucrose, SacT may be regarded as the major antiterminator protein controlling sucrose utilization. The minor role of SacY is also illustrated by the weak affinity of this protein to the *sacB* RAT which is two orders of magnitude lower than the affinities observed for LicT and GlcT with their respective targets (Declerck *et al.*, 1999, I. Langbein and J. Stülke, unpublished data). A step-wise conversion of the *sacB* RAT into a *sacPA* RAT like structure revealed the following observations: Single base mutations (see Fig. 5.8, *sacB*-R2, *sacB*-R3, *sacB*-R4, Table 12) all enhanced the binding of SacT, whereas the double and triple mutations (see Fig. 5.8, *sacB*-R5, *sacB*-R6) had additive effects. Most single mutations and the double mutation prevented SacY binding. However, the triple mutation which did also affect the top loop neutralized the negative effect of the two other nucleotides and restored binding of SacY.

Taken together, our data indicate that SacT is the most promiscuous of the antitermination proteins whereas GlcT at the other end of the spectrum is strictly confined to its cognate *ptsG* RAT due to its specific structural demands. In the living cell, glucose plays a special role as the by far most preferred carbon source. Therefore it is advantageous for the bacteria to have a regulatory system for glucose utilization that avoids any risk of cross-talk. On the other hand, SacT induces both sucrose catabolic systems, but it does not mediate antitermination at the *bglP* RAT in wild type bacteria (Krüger and Hecker, 1995). As shown here and in previous publications, SacT can bind *bglP*-like RAT structures (Manival *et al.*, 1997) (Table 12). It is so far unknown why SacT does not induce the *bglPH* operon in wild type bacteria. More factors such as the sequence context surrounding the RAT, the top loop, and the overall stability of the different RAT/terminator couples may provide additional levels for controlling the effective interaction with the antiterminator proteins. A careful analysis of the data presented here indicates that this is indeed the case. More work will be required to study the contributions of these factors.

6 Discussion

6.1 Effects of organic acids on the central metabolism of *B. subtilis*

In its natural habitat, *B. subtilis* faces a highly competitive environment. The bacterium is subjected to rapidly changing conditions, e.g. temperature and nutrient supply. Nevertheless, *B. subtilis* is also competing with other organisms for the available resources. Some soil organisms avoid direct competition by occupying exclusive ecological niches, others developed ingenious strategies to defend their habitat against intruders, e.g. by the production of antibiotics. In addition, metabolism of the organisms needs to adapt to the above mentioned challenges. The environmental conditions are sensed, and the obtained information is interpreted and transferred into gene and protein regulatory processes.

As *B. subtilis* is widely used for the industrial production of amylases, proteases, lipases, and vitamins, there is much interest in optimizing the metabolism of industrially used *B. subtilis* strains towards higher production yields (Harwood, 1992; Sauer *et al.*, 1996). To monitor the effects of different nutrient sources or genomic modifications on a global level, new screening methods have been developed over the last decade. *B. subtilis* was among the first fully sequenced bacteria (Kunst *et al.*, 1997), which enabled scientists to study this organism on global transcriptomic and proteomic levels (Hecker, 2003; Hecker and Engelmann, 2000; Yoshida *et al.*, 2001). Much effort has been devoted to study the influence of glucose and other sugars on the central metabolism of different eukaryotic and prokaryotic organisms. In this work, the methods of DNA-microarray and fluxome analyses were used to study the influences of the organic acids glutamate and succinate on the central metabolism of glucose grown *B. subtilis*. The combined evaluation of data derived from these two methods was expected to indicate for so far unknown regulatory contexts.

Metabolic pathways that show accurate correlation between the two global screening methods

With only few exceptions, the results obtained from the two global screening methods, i.e. microarray analysis and metabolic flux analysis, are in good agreement for most of the studied genes. The transcriptomic data revealed that the expression of most of the glycolytic genes is not affected by the availability of glutamate and succinate. This is also true for the genes coding for the pentose phosphate pathway (see Figs. 2.2 and 6.1). The glycolytic genes coding for enzymes that catalyze reversible reactions are known to be constitutively expressed, and it is therefore no surprise that no significant regulation was observed in the

transcriptomic analysis. The highest regulatory effect on genes of these two pathways has been observed for the *ptsGHI* and *gapA* operon in previous studies (Ludwig *et al.*, 2001; Stülke *et al.*, 1997). Both were highly induced in the presence of glucose, and were not expressed when no sugar was available. The transcription of the *gapA* operon has also been found to be influenced by the availability of amino acids (Ludwig *et al.*, 2001). Full induction was only observed when the medium was supplemented with casein hydrolysate. However, the addition of glutamate and succinate had no effect on the transcription of this operon in this study. The highest upregulation of glycolytic genes was observed for *pfkA* and *pykA* when glutamate and succinate were present in the medium, but only a weak 1.3-fold induction was observed. *PykA* and *pfkA* are clustered in the same operon. The phosphofructokinase is known to be allosterically controlled by ADP and fructose-6-phosphate in Bacilli (Evans *et al.*, 1981; Yoshida, 1972), but its regulation on a transcriptional or translational level is a controversial issue. On the one hand, it was reported that *pfkA* is slightly upregulated in the presence of glucose when using β-galactosidase assays and macroarrays (Blencke *et al.*, 2003; Ludwig *et al.*, 2001). On the other hand, no higher protein amount was detectable in a proteomic study under the same conditions (Tobisch *et al.*, 1999b), but the method of 2D gel electrophoresis, that was used to obtain the proteomic data, is not sensitive enough to detect such a small regulatory effect. In *Lactococcus lactis*, *pfk* and *pyk* are also organized in an operon, i.e. the *las* operon (Llanos *et al.*, 1993). Transcription of the *las* operon is activated by the presence of sugars transported via the PTS by means of CcpA (Luesink *et al.*, 1998). In a *ccpA* mutant strain, the transcription level of this operon is 4-fold reduced. A *cre* site was found upstream of the promoter region of the *las* operon, confirming the assumption that CcpA induces this operon when PTS sugars are available. Also, in *L. casei pfk* and *pyk* are clustered in an operon. However, in a *ccpA* mutant the transcription of these genes is slightly increased, indicating a PTS/CcpA dependent regulation mode that differs from that of *L. lactis* (Viana *et al.*, 2005). Similar to the situation in *B. subtilis*, the mode of *pfkA* regulation is not clear in *L. casei*. Data obtained from microarray analysis alone are not accurate enough to suggest a significant upregulation with the observed 1.3-fold change for *pfkA*. Due to the extensive tests that were carried out prior to this study, all transcription changes above 1.5-fold are considered as significant.

But the transcriptomic data can only give ratios for the gene expression under the two studied conditions. They can neither give information about the absolute expression rate of the analyzed genes, nor do they reflect the activity of the respective enzymes. The β-galactosidase assays that were carried out in this work can also not reflect the activity of a

certain enzyme, but in addition to microarray analysis, it can give indication if a gene is highly expressed or not, even if the ratio for the gene examined under the two conditions is 1. The β-galactosidase assays revealed that *ptsG* and the *gapA* operon are highly expressed under both conditions studied here. This high expression is caused by the presence of glucose under both conditions and is also in good agreement with previous findings (Ludwig *et al.*, 2001; Stülke *et al.*, 1997). It was also reported, that casein hydrolysate has a positive effect on the transcription of this operon (Ludwig *et al.*, 2001). Thus, glucose and amino acids seem to have a synergistic effect on the expression of the *gapA* operon. However, no additional effect upon addition of glutamate and succinate was observed in the microarray analysis.

Contrary to the constant expression of the glycolytic genes, the flux into glycolysis is slightly decreased in the presence of organic acids. The reason for that is a reduced flux from glucose-6-phosphate to fructose-6-phosphate, carried out by the glucose-6-phosphate isomerase (Pgi), in the presence of glutamate and succinate. This can be explained with the higher portion of glucose-6-phosphate entering the pentose phosphate pathway via glucose-6-phosphate dehydrogenase (Zwf) under this condition. For both enzymes, glucose-6-phosphate is the direct substrate. With the used methods, it can not be distinguished whether a lower enzymatic activity of glucose-6-phosphate isomerase or a higher activity of glucose-6-phosphate dehydrogenase is the reason for the change in metabolic fluxes. The transcription of both genes is not altered under any of the analyzed conditions. It was reported in previous studies that the expression of these genes is also not affected by the presence or absence of glucose (Blencke *et al.*, 2003; Ludwig *et al.*, 2001). In *E. coli*, it was shown that the deletion of *pgi* resulted in a significant higher flux of glucose-6-phosphate into the PPP when grown on glucose. At the same time, the maximum growth rate was greatly reduced (Canonaco *et al.*, 2001). The impact on growth rate could be traced back to a disturbed reducing power balance due to the difficulty of *E. coli* metabolism to reoxidize NADPH in the *pgi* mutant. When grown on fructose, which does not need the glucose-6-phosphate isomerase to enter glycolysis, the *pgi* mutant grew with maximum specific growth rate (Canonaco *et al.*, 2001). This demonstrates that a reduced activity of glucose-6-phosphate isomerase could indeed lead to a higher flux of glucose-6-phosphate into the PPP. However, an enzyme catalyzing a downstream reaction could also lead to the observed metabolic fluxes. A good candidate gene would be *ywlF,* coding for ribulose-5-phosphate epimerase.

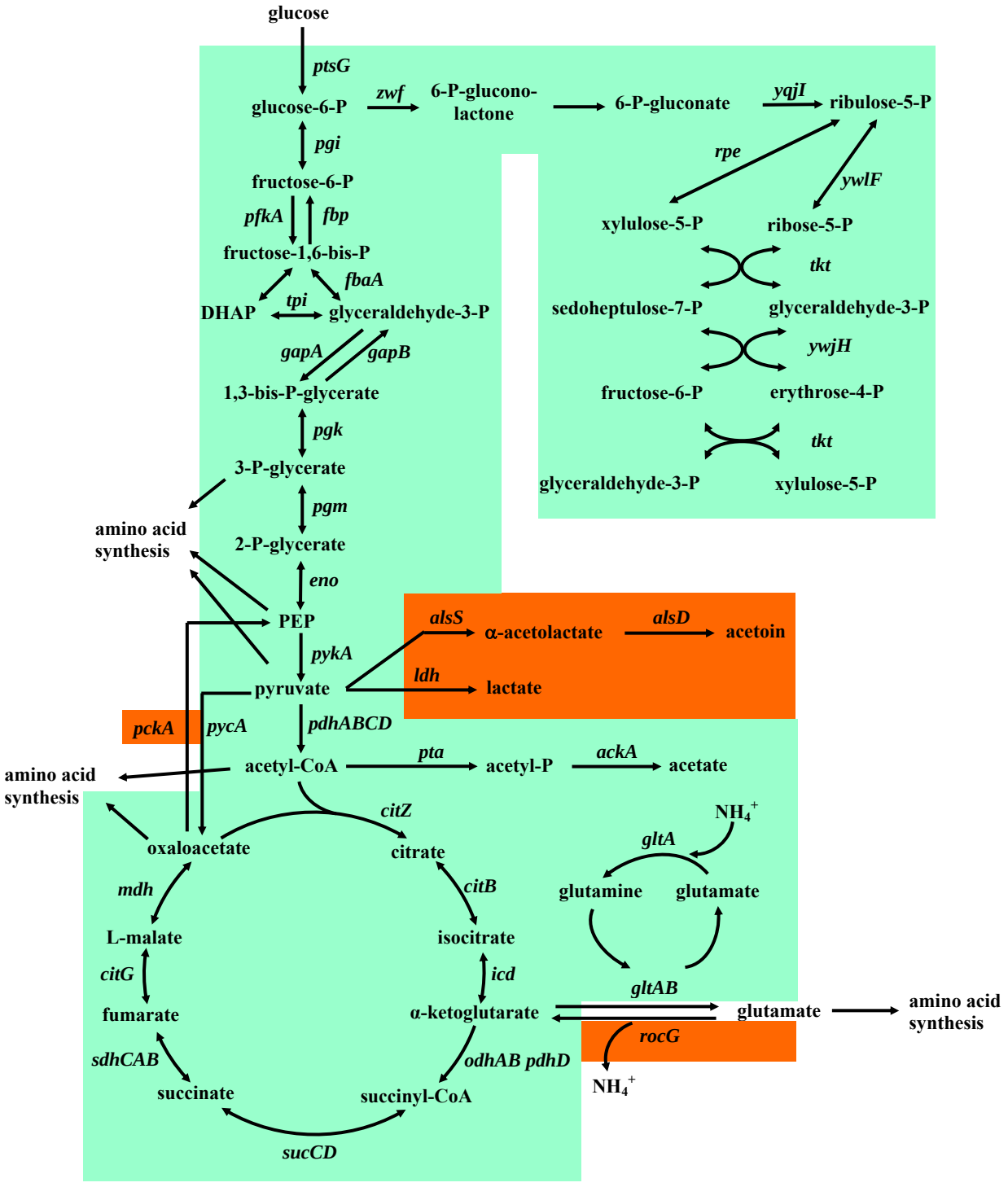

Figure 6.1 Overview on the genes of central metabolic pathways in *B. subtilis*.
Metabolic pathways that showed good correlation for the used screening methods are indicated in green. Pathways that showed no correlation between the used methods are indicated in red.

The transcriptomic analysis revealed a weak induction of *ywlF* in the presence of organic acids (1.45-fold), and this could be a possible reason for the increased flux into the pentose phosphate pathway. It is interesting to note that the flux into the PPP was relatively high under both tested conditions. In a previous study, it was reported that the metabolic flux into the PPP increases with the growth rate of *B. subtilis* (Dauner and Sauer, 2001). Also, a relatively high flux into this pathway, which is between 30 and 50 percent of the metabolized glucose, has been observed (Sauer *et al.*, 1997; Sauer *et al.*, 1996). Therefore, the pentose phosphate pathway can be considered as one of the major pathways for glucose catabolism besides glycolysis. At higher growth rate, and therefore also higher biomass production, *B. subtilis* would have to increase the flux into the PPP to meet its higher demand of NADPH for anabolic reactions. In agreement with this theory, the growth rate was increased by 27% when glutamate and succinate were available, and the data gained from the fluxome analysis confirmed that a higher portion of the metabolic intermediates is used for biomass generation.

The TCA cycle shows a strong repression in the presence of organic acids. The transcriptomic data revealed a 5- and 6-fold repression of *citZ* and *citB,* respectively (see Fig. 2.2). These genes code for enzymes that catalyze the initial reactions of the TCA cycle. This significant repression was observed not only on the transcriptomic level, but also the fluxome data revealed a 5-fold lower flux of acetyl-CoA into the TCA cycle when glutamate and succinate were available (see Fig. 2.1). Moreover, the reporter gene analysis of a *citB-lacZ* fusion showed a 9-fold reduced expression upon addition of these organic compounds. The observed repression is a synergistic effect coming from both glucose and glutamate. It was shown previously, that glucose alone causes a repression of *citZ* and *citB*. This repression is increased when glucose and amino acids are present simultaneously (Rosenkrantz *et al.*, 1985; Sonenshein, 2002). Similarly to the synergistic induction of the *gapA* operon by glucose and amino acids, the TCA cycle is synergistically repressed by the same compounds (Blencke *et al.*, 2003; Blencke *et al.*, 2006; Ludwig *et al.*, 2001). Since the TCA cycle has a dual role, i.e. the catabolism of carbon sources and the anabolism of amino acids, it makes sense that its regulation is affected from both directions (see also chapter 1.1.2).

The regulation mode of *citZ* and *citB* might not be the only reason for the reduced flux from acetyl-CoA through the TCA cycle. A significant increase of acetate production due to an induced transcription of the respective enzymes, phosphotransacetylase (Pta) and acetate kinase (AckA), lowers the availability of acetyl-CoA for the TCA cycle and therefore leads to a “by-pass” of this pathway. It was shown in a *pta* deletion mutant that *citB* expression is 2-fold increased, and when overexpressing phosphotransacetylase, *citB* is even repressed

when no PTS sugar but gluconate is available (Blencke *et al.*, 2006). This indicates that the flux from acetyl-CoA to acetate has indeed an influence on the regulation of the TCA cycle.

High amounts of acetate are already produced without the addition of glutamate and succinate, since *pta* and *ackA* are induced in the presence of glucose by the means of CcpA (Grundy *et al.*, 1993; Presecan-Siedel *et al.*, 1999; Turinsky *et al.*, 1998). Similar as observed for the TCA cycle, the presence of glutamate and succinate has a synergistic effect on gene transcription, but now in the opposite direction. *AckA* was found to be 3-fold induced and the flux from acetyl-CoA to acetate was significant higher in the presence of glutamate and succinate. It was recently found that *ackA*, in addition to CcpA, is also regulated by the transcriptional regulator CodY (Molle *et al.*, 2003; Shivers *et al.*, 2006). The activity of CodY is regulated by the intracellular levels of GTP and branched-chain amino acids (BCAA), i.e. leucine, isoleucine, and valine, which are synthesized from pyruvate. CodY is the repressor of more than 100 genes (Molle *et al.*, 2003). However, *ackA* represents the first gene that is positively regulated by CodY and this positive effect is exclusively dependent on the availability of BCAA (Shivers *et al.*, 2006). It is known that CcpA and CodY are both involved in the repression of the TCA cycle and genes of BCAA biosynthesis (Kim *et al.*, 2003a; Kim *et al.*, 2002; Kim *et al.*, 2003b; Ludwig *et al.*, 2002a; Shivers and Sonenshein, 2004; Shivers and Sonenshein, 2005; Tojo *et al.*, 2005). The fact that both regulators are also involved in the regulation of *ackA* suggests that CodY plays also a significant role in controlling cellular carbon flow. It was found that CcpA and CodY have an additive effect and both regulators are necessary for full activation of *ackA* expression (Shivers *et al.*, 2006). This would fit very well to the regulatory pattern observed in this study. However, it remains unclear how the presence of glutamate and succinate could lead to the activation of CodY. Site directed mutageneses of the *ackA* promoter region, affecting CodY binding, could indicate whether CodY is involved in the observed induction of *ackA* or not.

The transcription of TCA cycle genes other than *citZ* and *citB* is not severely affected by the availability of glutamate and succinate. This reflects the need of the respective enzymes under both studied conditions. The fluxome analysis revealed that a high portion of the external glutamate is converted to 2-oxoglutarate and thus entering the TCA cycle (see Fig. 2.1). Also external succinate feeds into this pathway. Under these conditions, the TCA cycle does only seem to fulfill a biosynthetic function. This behaviour of the TCA cycle in excess of nutrient sources was also observed for *Saccharomyces cerevisiae*. The performance of the catabolic and anabolic function was found to be related to the glucose uptake rate and extracellular pH (Blank and Sauer, 2004).

More than just the sum of its individual parts: Data that show no correlation between the two global screening methods

The comparison of the data derived from transcriptome and fluxome analyses revealed that in some cases, there is no close correlation for the two global screening methods. The genes coding for acetolactate synthase and acetolactate decarboxylase, converting pyruvate to α-acetolactate and acetoin (Cruz Ramos *et al.*, 2000), are highly induced without a notable change in flux to acetoin in the presence of organic acids. In fact, these metabolites were undetectable under any of the studied conditions (see Figs. 2.1 and 6.1). The transcription of *pckA* did not change, while a notable change in flux from oxaloacetate to pyruvate occurred. The ratio found by microarray analysis was close to 1, but the conversion from oxaloacetate to PEP was threefold increased. For *ldh*, a change towards the same direction was measured for the transcriptome and fluxome analyses, but the fold change of 1.7 does not meet the observed 25-fold increase in metabolic fluxes from pyruvate to lactate when glutamate and succinate are present (see Figs. 2.1, 2.2, and 6.1).

These observations do not necessarily mean that the data obtained by any of these two methods is wrong. For a better interpretation of the information, it is important to know the limitations of the two methods. Microarray analysis gives only information about the presence of a certain transcript under a certain condition. But it does not give information for the absolute amount of a transcript. In fact, only the ratio in relation to another condition is obtained (Harrington *et al.*, 2000). For the analysis, it is also assumed that the amount of a certain transcript reflects the amount of the respective protein. But this does not necessarily need to be the case. Due to alternative splicing, a single mRNA can code for several different proteins in eukaryotic cells (Black, 2000). Although this is not the case in prokaryotic organisms, it can not be considered that the amount of transcripts found in a cell completely reflects the amount of the corresponding proteins. A transcribed mRNA can interact with several factors that regulate its translation. These factors can stimulate or inhibit translation initiation (see also chapter 1.2). Also the complete protein synthesis machinery can interact differently with the same mRNA under varied conditions (Laursen *et al.*, 2005). Protein activity can also be modulated by modifications, e.g. phosphorylation events, and nevertheless, the half-life of a protein can also be altered under different environmental conditions (Jenal and Hengge-Aronis, 2003). Therefore, the transcriptome can only partly reflect the regulatory processes of an organism.

A way to determine the activity of metabolic enzymes is the combination of transcriptome analysis with metabolic flux analysis. Metabolic flux analysis is a method to determine the proportion that each metabolic pathway exerts when an organism is metabolizing a given substrate under a certain condition (Christensen *et al.*, 2002; Nielsen, 2003; Sauer, 2004). However, this method also has some drawbacks. Metabolic pathways that are not coupled to measurable fluxes can not be resolved (Wiechert, 2001). Also in case of "parallel metabolic pathways" where none of the branches is coupled to a measurable variable, it is impossible to resolve the "two branch fluxes" (Sonntag *et al.*, 1993; Wiechert, 2001). Moreover, the whole method is based on a computer model of the studied organism. Therefore, the data obtained with this method can only be as accurate as the underlying model (Covert *et al.*, 2001; Wiechert, 2001).

Both methods are established and already extensively used to analyze the metabolism of *B. subtilis* (Blencke *et al.*, 2003; Dauner *et al.*, 2001; Koburger *et al.*, 2005; Moreno *et al.*, 2001; Sauer *et al.*, 1996). The combination of data sets from both methods can indicate so far unknown regulatory processes on the protein level. This could be the case for the above mentioned genes. The regulation of the *alsSD* operon and *ldh* are emphasized in this discussion, since they represent both extremes observed in this study, i.e. a high induction of transcription without changed enzymatic activity, and a significant higher enzymatic activity at almost unchanged transcription rate.

The *alsSD* operon was found to be 10-fold induced in the presence of organic acids. However, no acetoin formation was detectable under both studied conditions (see Figs. 2.1 and 2.2). It was reported, that detectable amounts of acetoin are only produced when *B. subtilis* is in the stationary phase. Interestingly, a 10-fold induction was also observed when the bacteria were in stationary phase (Renna *et al.*, 1993). The *alsSD* operon is controlled by the transcription activator AlsR (Cruz Ramos *et al.*, 2000; Renna *et al.*, 1993). As mentioned above, high amounts of acetate are produced when the bacterium grows on glucose, and it was reported that either acetate or a lowered pH value activates AlsR (Holtzclaw and Chapman, 1975; Renna *et al.*, 1993). There is also evidence that the *alsSD* operon is indirectly induced by the redox regulator Fnr under anaerobic conditions (Cruz Ramos *et al.*, 2000). Additionally, CcpA plays an indirect role in the induction of this operon in the presence of high amounts of glucose (Renna *et al.*, 1993; Tobisch *et al.*, 1999b). The data indicate that no acetoin is formed although the *alsSD* operon is highly induced upon addition of glutamate and succinate. Since glucose is present under both conditions, it can also be assumed that the operon is already induced without the addition of organic acids. The absence

of the respective product under both conditions might indicate an additional regulatory level. On the one hand, translation initiation could be controlled by an additional factor, e.g. a riboswitch, ncRNA, or a protein factor. On the other hand, activity of acetolactate synthase and/or acetolactate decarboxylase might be controlled by allosteric inhibitors or by protein modification. Indeed, the phosphorylation of acetolactate decarboxylase was recently reported (Lévine *et al.*, 2006). The trigger might be the intracellular acetate concentration, since the pH value of the medium was constant and no oxygen limitations occurred.

The transcription of the *ldh* gene, encoding lactate dehydrogenase, was 1.7-fold induced when glutamate and succinate were present. Although no lactate was detectable before, significant amounts of lactate were produced upon addition of organic acids. On the other hand, the flux from pyruvate to acetyl-CoA was not affected by the higher activity of lactate dehydrogenase, which also uses pyruvate as its direct substrate (see Figs. 2.1, 2.2, 6.1). It was demonstrated in *E. coli* that overexpression of *ldh* did also not alter the flux from pyruvate to acetyl-CoA, and when the organism was grown under the same conditions, a higher level of lactate dehydrogenase did also not automatically result in higher lactate production (Yang *et al.*, 1999). This implies that the amount of lactate dehydrogenase is not a limiting factor for the respective enzymatic reaction, and the enzyme might be controlled on the protein level. In addition, a significant higher amount of acetate production was observed when organic acids were present (see above). By producing acetate, *B. subtilis* yields additional ATP, but more NAD^+ is reduced to NADH as well. Since much higher acetate production was observed in the presence of organic acids, the lactate production might be necessary to maintain the redox balance since NAD^+ is regenerated in this reaction. It was shown for *Lactococcus lactis* that a functional lactate dehydrogenase is essential for the regeneration of NAD^+ under aerobic and anaerobic conditions (Neves *et al.*, 2002). Moreover, it was observed that the redox state of *B. subtilis* controls *ldh* gene expression in the presence of nitrate (Reents *et al.*, 2006). This supports the idea that the redox state of the cell might be responsible for the higher activity of the lactate dehydrogenase in the presence of organic acids. The increased flux from pyruvate to lactate when glutamate and succinate are present (see above) is paralleled by a reduced flux from pyruvate to oxaloacetate, which is catalyzed by the pyruvate carboxylase (see Fig 2.1). This reciprocal ratio of these two fluxes explains why the flux from pyruvate to acetyl-CoA is not altered under both tested conditions.

To be or not to be…: Metabolic adaption to a highly competitive environment

From a global point of view, a general strategy becomes visible. In excess of glucose, glycolysis is highly induced whereas the flux of acetyl-CoA is directed to the overflow metabolic pathways rather than to the TCA cycle. This regulation pattern is intensified upon addition of glutamate and succinate. The TCA cycle is almost exclusively used for anabolism of amino acids and glycolysis, apart from delivering precursors for amico acid synthesis, provides the energy for the cell. The PPP delivers NADPH and building blocks for cell wall and nucleotide synthesis, but the relatively high flux of glucose-6-phosphate might exceed the necessary amount for the supply of these compounds. Thus, the PPP might be also a major pathway for glucose catabolism.

Although this regulation pattern seems to be reasonable, it is noteworthy that the bacterium wastes much energy with this strategy. Normally, bacteria are noted for their efficiency. If the metabolism of *B. subtilis* is not oxygen limited, it could yield much more ATP per mole glucose when using the TCA cycle for catabolism of acetyl-CoA and oxidative phosphorylation for NAD^+ recovery instead of using fermentation pathways. To see the benefit of its metabolic strategy, the natural habitat of *B. subtilis* has to be considered. The bacterium lives in a very competitive environment and struggles with many other organisms for the available nutrient resources. All organisms have to make a tradeoff between rate and yield of ATP production (Pfeiffer *et al.*, 2001; Stucki, 1980). In general, single cellular microorganisms benefit from a higher ATP production rate when they have to share resources with other organisms. This is true for both eukaryotes and prokaryotes. Lactobacilli and Fungi, like *Saccharomyces*, use fermentation for ATP production even in the presence of oxygen (Poolman, 1993; van Dijken *et al.*, 1993). The result is a rapid exhaustion of the given resources, and the aim of these organisms is to get the most from the dish. This strategy can be seen as a selfish contest in biomass production. However, when the glucose uptake rate becomes limited, *Saccharomyces* increases respiration and catabolic TCA cycle activity (Blank and Sauer, 2004).

In multicellular organisms, the cells learned to cooperate with each other and to share the available resources. Since they are mostly metabolizing internal resources, they are also not in direct competition with other organisms for nutrient sources. As a result, they aim for a high yield of ATP production rather than a high rate (Pfeiffer *et al.*, 2001). In consequence, the metabolism of *B. subtilis* is optimized for a high production rate and therefore is well adapted to a highly competitive environment.

6.2 Specificity in gene regulatory systems

Although the specificity determinants of the related RAT-antiterminator couples were in the focus of this study, it is obvious that the interplay of several mechanisms results in the specific activation of the relevant systems. This starts already with the more or less specific interaction of the permeases with their appropriate substrate. However, relaxed specificity was found for BglP in this study and was previously reported for PtsG (Langbein *et al.*, 1999). In addition to transcription control by antitermination, the expression of the permeases, except for PtsG and SacX, is also under the control of carbon catabolite repression. This helps to prevent the expression of any other permeases than PtsG and SacX in the presence of glucose. Finally, the negative and positive regulation of the antiterminator proteins at PRD-I and PRD-II, respectively, also helps to keep signals straight in transcriptional antitermination (see also chapter 1.3.3). PRD-I directly interacts with the appropriate permease. The interaction of the permeases with their corresponding antiterminator protein is very specific and no cross-talk was reported so far. On the other hand, it has been reported that HPr, which normally phosphorylates PRD-II, can phosphorylate the PRD-I of LicT *in vitro*. However, this effect was led back to the non-physiologically high HPr concentrations used in that study (Lindner *et al.*, 1999; Tortosa *et al.*, 2001). Furthermore, high SacT concentration resulted in dimerization and RAT-binding of SacT *in vitro* independent from its phosphorylation state (Arnaud *et al.*, 1996). Both observations indicate that specificity is also a matter of the concentration of the potential interaction partners.

As described here, a biological gene regulatory system is composed of many different elements that have to interact specifically with each other. An external or internal signal has to be sensed and the gained information is directly or indirectly transmitted to a regulator. The RNA thermometers seem to be the most basic regulatory systems in biological organisms (Narberhaus, 2002; Narberhaus *et al.*, 2006). A certain sequence in the 5'-UTR of an mRNA can act as both sensor and regulator. The only needed interactions are RNA-intramolecular base pairings. This is also true for metabolite dependent riboswitches, but additionally, the RNA has to interact specifically with the sensed molecule.

As soon as a regulatory protein gets involved, the complexity of a regulatory system increases by a considerable amount. This increases also the necessity for specific interactions to prevent any cross-talk with other regulatory systems in the cell. The proteins of the four antiterminator systems interact specifically with their cognate RAT and with two other protein factors, i.e. the respective sugar permease and HPr. The RNA binding domains of the four

related antiterminator proteins seem to recognize both structure and sequence of their target. However, the structure seems to be more relevant for the specific target recognition.

In contrast, TRAP is a good example of an antiterminator protein that exclusively seems to recognize the sequence rather than the structure of its target RNA (see chapter 1.2.3.4). This protein is composed of 11 identical subunits and controls expression of the *trpEDCFBA* operon and *trpG* (Antson *et al.*, 1995; Babitzke, 2004; Gollnick, 1994). TRAP forms a donut-like structure and the target RNA is wrapped around the protein ring. It was shown that 11 subunits interact specifically with 11 (G/U)AG repeats in the RNA target sequence, each repeat separated by two or three nucleotide spacers (Babitzke *et al.*, 1996; Baumann *et al.*, 1997; Yang *et al.*, 1997). In addition to the target RNA sequence, TRAP binds to free tryptophan and to a protein called anti-TRAP. Both ligands modulate TRAP activity.

The fact that most regulatory proteins are made up of distinct subunits, whereas each subunit has a dedicated function and interaction specificity, facilitates the adoption of existing proteins to other functions and thus the evolution of new regulatory systems. The overall composition of LevR, controlling *levDEFGsacC* operon, and LicR, controlling *licBCAH* operon, is very similar to that of the four PTS controlled antiterminator proteins (Martin-Verstraete *et al.*, 1998; Martin-Verstraete *et al.*, 1990; Stülke *et al.*, 1995; Tobisch *et al.*, 1999a). Likewise, they have two C-terminal PRDs that can be phosphorylated by components of the PTS system or sugar permeases. However, they feature an N-terminal DNA binding domain instead of an RNA binding domain. Thus, they are transcription factors rather than antiterminator proteins. This shows how the exchange of a single subunit could alter the functionality of a regulatory protein.

In contrast to RNA, double stranded DNA does not form complex secondary structures. For a specific interaction with a certain target sequence, a regulatory protein can only interact with the side chains and the edges of the base pairs by hydrogen bonding interactions. The distance of the protein interacting base pairs, the minor and major groove, and the DNA bending and torque are used as specificity determinants (Pabo and Sauer, 1984). The helix-turn-helix motive is the most common DNA binding motive, and DNA binding proteins bind their target mostly as symmetric dimers. The LysR family of transrciptional regulators is a good example how more than 50 autoregulatory transcriptional regulators, each binding to different DNA sequences and regulating the target genes by transcription activation or repression, can evolve from a distant ancestor. The activity of these transcription regulators is modulated by direct interaction with lower molecular weight factors (Schell, 1993).

Two-component systems make specific signal transduction even more crucial. At least two proteins, a sensor kinase and a response regulator, lead to the specific activation of a target gene mostly based on an external signal. The signal is always transmitted by phosphate group transfer among the proteins of the regulatory system (Albright *et al.*, 1989; Goudreau and Stock, 1998; Hoch and Varughese, 2001). Bacteria with a large genome, such as *E. coli* and *B. subtilis*, possess up to 40 pairs of two component systems, each dedicated to individual signals and genes. *Nostoc punctiformis* was found to have even 145 sensor kinases and 103 response regulators (Fabret *et al.*, 1999; Mizuno, 1997). The signaling pathways of chemotaxis or sporulation increase the amount of needed sensor and signaling proteins by a considerable amount and specific protein-protein interaction as well as the concentration of the involved proteins are critical to keep signal transduction straight within the regulatory system (Burbulys *et al.*, 1991; Levit *et al.*, 1998; Trach *et al.*, 1991).

Specificity determinants of the PTS controlled antiterminator systems

RNA can form complex secondary and tertiary structures similar to that of proteins (see chapter 1.2.1). Because of the possibility of alternative intramolecular base pairings and the resulting structural complexity of even small RNAs, the exact folding of the RATs as well as the RNA bases involved in RAT-antiterminator interactions are not predictable. Besides NMR and X-ray structural analysis, single nucleotide exchanges can give an indication for secondary structure and specificity determinants. The exchange of nucleotides essential for RNA-protein interaction or secondary structure formation might affect the functionality of the antiterminator system.

The overall secondary structure of the RATs can be subdivided into a top loop and upper and lower loop regions. Furthermore, there is an upper, middle, and lower stem region (see Fig. 6.2). Not only the basic structures of *ptsG*, *sacB*, *sacPA*, *licS*, and *bglP* RAT are very similar, but GlcT, LicT, SacT, and SacY are also recognizing structures that are very similar in sequence (see Fig. 6.3). Therefore, it is not obvious how these proteins can distinguish their specific target from other RATs. The major differences among these sequences can be found in the top loop and lower loop regions (see Fig. 6.3). Hence, it is very likely that the loop regions are involved in specificity determination.

It was found in previous studies that bases located in loop regions are often involved in RNA-protein recognition, whereas bases in stem regions are responsible for stability and structure formation (Nagai, 1996).

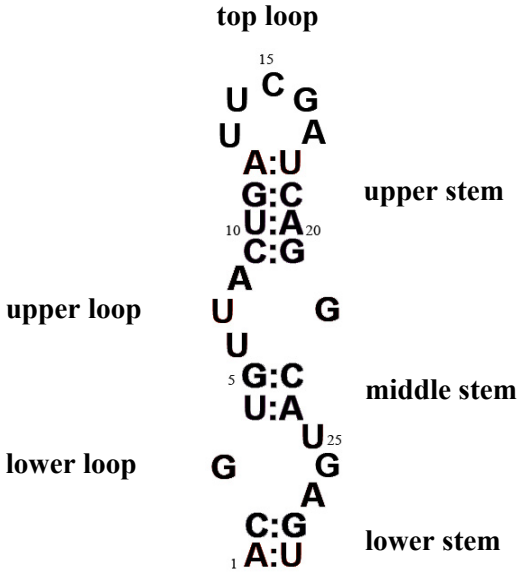

Figure 6.2. Overall secondary structure of the related RATs at the basis of the *ptsG* RAT (Yang *et al.*, 2002). All RATs share the same basic structure, which can be subdivided into three loop regions (top loop, upper and lower loop) and three stem regions (upper, middle, and lower stem).

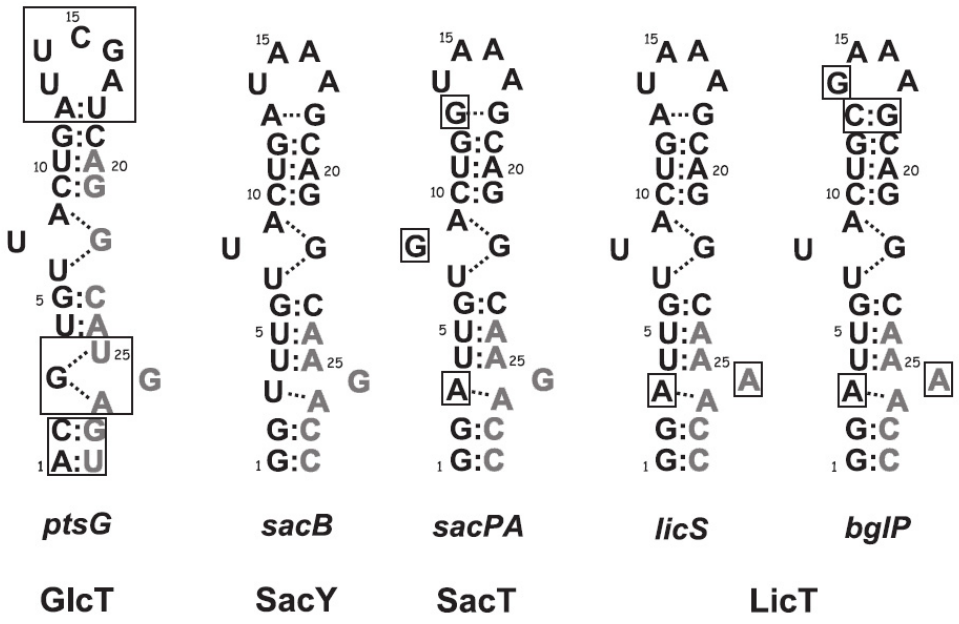

Figure 6.3. Comparison of the secondary structures of the related RAT RNAs of *B. subtilis* (Schilling *et al.*, 2004; Yang *et al.*, 2002). The relevant antiterminator proteins are indicated below their cognate RAT structures. Boxes indicate nucleotides that differ from the *sacB* RAT. Dashed lines indicate bases that are proposed to be in direct contact to each other. Grey nucleotides are overlapping with the terminator structure. The main differences are located in the top loop and lower loop regions. The *ptsG* RAT holds an outstanding position due to its unique lower stem loop and the triple base pairing in the lower loop region.

The N-peptide of phage λ and the MS2 coat protein of phage R17 recognize the top loop structure of their respective target RNA for binding (Grahn *et al.*, 1999; Legault *et al.*, 1998; Valegård *et al.*, 1997). In contrast, site directed mutagenesis of the *sacB* and *ptsG* top loop regions had no effect on target recognition or specificity Aymerich and Steinmetz, 1992; this work). As the upper loop regions of most of the related RAT structures are identical (except for *bglP* and *ptsG* RAT, see Fig 6.3), it seems to be reasonable that the upper loop cannot contribute to specificity determination.

When analyzing the NMR structure of the LicT-*bglP* RAT complex, it turned out that the top loop region is indeed not involved in RNA-protein interaction, whereas the upper and lower loop regions of *bglP* RAT are both bound by LicT (Yang *et al.*, 2002) (see Figs. 6.4 and 6.5). As the upper loop regions (with the exception of G8 in *sacPA* RAT, see Fig. 6.3) are also identical in the related RAT structures, it seems also unlikely that it can contribute to the specificity of the RAT–protein interaction.

In contrast to the top loop and upper loop regions, the lower loop features more diversity among the RAT sequences (see Fig. 6.3). Accordingly, the site directed mutagenesis of *ptsG* and *sacB* RAT revealed that the specificity determinants of the related RNA-protein couples are indeed located in the lower loop region of the respective RAT structure. The adenines at position 3 and 26 are essential for RAT recognition by LicT. SacT has a relaxed specificity as it binds to all *sacB* derived RAT mutants. The data indicate that SacT, aside from binding to its cognate RAT, i.e. *sacPA*, is also able to bind *sacB*, *bglP*, and *licS* RAT. However, SacT dependent transcriptional antitermination of *bglP* or *licS* RAT has not been reported in the presence of sucrose. The s*acPA* RAT seems to have the optimal configuration for SacT binding, as it is bound by SacT with the highest efficiency of all tested RAT sequences. SacY, apart from binding *sacX* and *sacB* RAT, also binds to the *sacPA* RAT sequence. In contrast to SacT, SacY has a more stringent specificity since it does not tolerate the exchange of any of the bases that differ in *sacB* and *sacPA*. These bases are located at positions 3, 8, and 13 in the respective RAT structure (see Fig. 6.3). Only the strict combination of these bases, as found in *sacB* (U3, U8, A13) or *sacPA* RAT (A3, G8, G13), is recognized by SacY (see Fig. 6.3). A single nucleotide exchange in these combinations is not tolerated by SacY.

According to the NMR structure of the LicT-*bglP* RAT complex, some bases of the middle stem loop are directly interacting with certain amino acids of LicT. Additionally, the protein is directly interacting with the sugar phosphate backbone of the nucleotides of the middle stem loop (see Fig. 6.4).

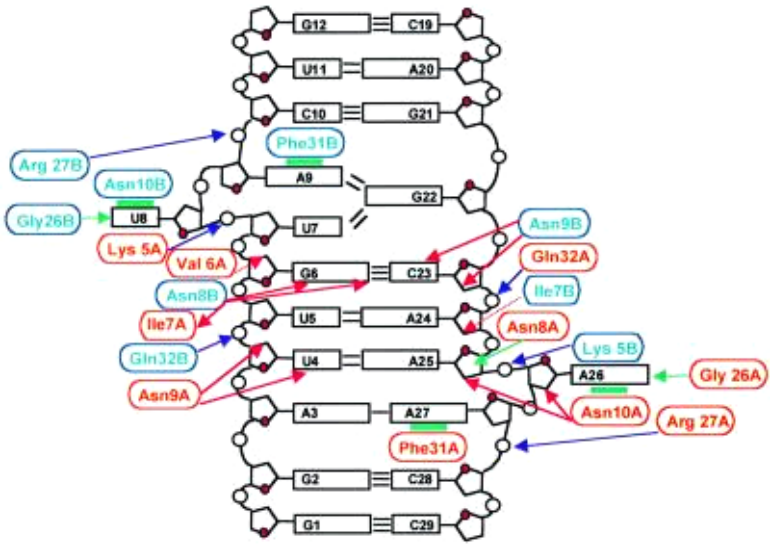

Figure 6.4. Schema of the direct molecular interactions of the LicT-*bglP* RAT complex. Solid red lines indicate hydrogen bonds involving an amino acid side chain atom, dashed red lines indicate hydrogen bonds involving a protein backbone atom, green lines indicate van der Waals interactions and blue lines indicate possible interactions with the RNA phosphodiester backbone (Yang *et al.*, 2002).

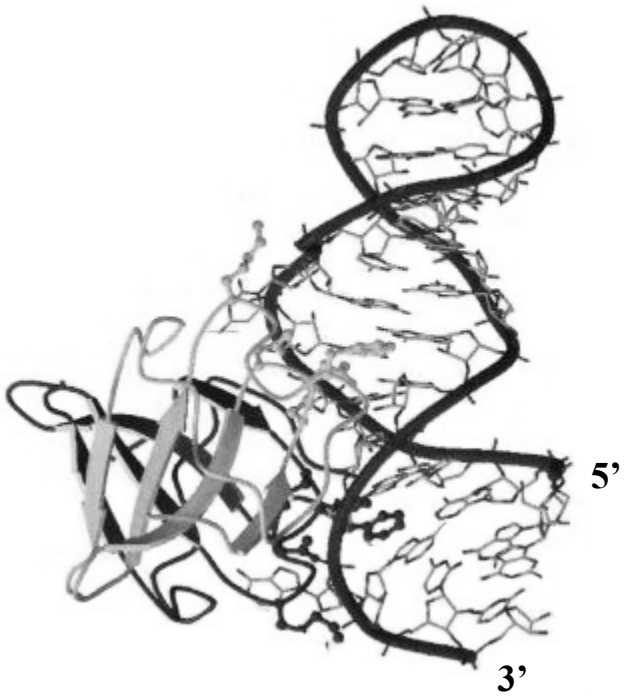

Figure 6.5. MOLSCRIPT (Kraulis, 1991) representation of the LicT dimer interacting with its RAT hairpin target. Modified from Yang *et al.*, 2002.

This is similar to the viral Rev-Peptide from HIV-1, which induces the splicing of a viral mRNA when binding to the RRE-sequence (Rev responsive element) (Daly *et al.*, 1989; Zapp and Green, 1989). Here, the protein interacts also with the sugar phosphate backbone and several nucleotides located in the stem region of the RRE secondary structure. This interaction induces stem structure formation and is also essential for RNA-protein recognition (Bartel *et al.*, 1991; Battiste *et al.*, 1996; Battiste *et al.*, 1994). Whether the interaction of LicT with the *bglP* middle stem-region is essential for RNA structure formation or not cannot be determined with the methods used in this study. However, with the exception of *ptsG*, the middle stem structure and sequence is identical for all the related RAT structures and is therefore not considered to be a specificity determinant. Thus, the lower loop region of the RATs seems to be the central element for the specific RNA-protein interaction.

Interestingly, a single nucleotide exchange in *sacB* RAT or the introduction of an additional nucleotide into *ptsG* RAT could change specificity towards another antiterminator protein. Something similar has been previously observed for metabolite dependent riboswitches. Guanine and adenine binding riboswitches are also very similar in structure and sequence and bind to their respective target with high affinity (Mandal *et al.*, 2003; Mandal and Breaker, 2004) (see also chapter 1.2.3.3). The exchange of a single cytosine to uridine in its RNA sequence converts the guanine specific riboswitch to an adenine binding riboswitch (Mandal and Breaker, 2004). Similar to *ptsG* RAT, an unusual triple base pairing is responsible for specificity and target interaction (Noeske *et al.*, 2005).

The exceptional position of GlcT and *ptsG*-RAT

GlcT and *ptsG* RAT hold an outstanding position since neither binds GlcT to any of the other related RAT sequences nor do any of the other antiterminator proteins bind to the *ptsG* RAT sequence. To understand the high specificity of the GlcT-*ptsG* RAT couple, a detailed comparison of all antiterminator systems is required. When having a closer look to the predicted secondary structures of the related RATs, it becomes apparent that *sacB*, *sacPA*, *bglP*, and *licS* RAT are all very similar in structure and sequence. On the other hand, *ptsG* RAT is more distant from the other RATs and exhibits some exclusive features.

First, the top loop features an additional nucleotide and a sequence (UUCGA) that is very different from the other four RAT top loop sequences (UAAA or GAAA, see Fig 6.3). Secondly, the middle stem region consists of only two base pairings instead of the three found in the other RATs. As the unpaired nucleotide at position U25 is part of the lower loop region, the *ptsG* RAT lower loop comprises four nucleotides instead of three. Thirdly, the lower stem

of *ptsG* RAT is unique and differs completely from the lower stem region of the other RATs (see Figs. 6.2, 6.3). Finally, a larger amount of nucleotides of the *ptsG* RAT sequence is overlapping with the terminator sequence (see Fig. 6.3).

The NMR structure of the LicT-*bglP* RAT complex clearly shows that the upper and lower loop regions are actually no open loop structures as shown in Figure 6.2 for *ptsG* RAT. Most nucleotides located in the loop regions are in unusual base pairings with opposing nucleotides. In the lower loop A3 interacts with A27 and in the upper loop region U7, A9, and G22 are forming a triple base pairing. Therefore, only the two nucleotides U8 and A26 remain unpaired. Thus, most of the nucleotides in the loop regions are participating in stem structure formation (see Fig. 6.4) (Yang *et al.*, 2002). Since all related RAT structures, including *ptsG* RAT, feature the same nucleotide configuration in the top loop region, it is safe to propose the same triple base pairing for all RATs. The *ptsG* RAT features this nucleotide composition also for the lower loop region (see Figs. 6.2 and 6.3). It is therefore predicted that the same triple base pairing formation is also present in the lower loop region forming a highly symmetrical RAT structure. When introducing this structural element into *sacB* RAT, binding of GlcT to the mutant RAT is observed. This confirms that this structural feature is indeed the specificity determinant for GlcT-*ptsG* RAT interaction.

When analyzing the evolution of the antiterminator protein regulatory domains it was found that GlcT of *B. subtilis* and *Staphylococcus carnosus* (Christiansen and Hengstenberg, 1996) is more distant from the antiterminator proteins of the sucrose and β-glucoside classes. The latter exhibit a closer relationship with each other and form a distinct subgroup (Greenberg *et al.*, 2002). Thus, both the *ptsG* RAT and the GlcT antiterminator protein are the most distant components of the related antiterminator systems.

Considering all this, it becomes more clear why the GlcT-*ptsG* RAT interaction is the most specific of the related antitermination systems. The sugar transported by this system holds also an exclusive position. Glucose is the preferred carbon source of *B. subtilis* and when available, carbon catabolite repression suppresses genes needed for the utilization of alternative sugars. Eliminating any cross-talk is economical for the bacterium and helps to repress the expression of genes that are not required. *Escherichia coli* follows the same strategy. However, the control of PTS dependent glucose uptake differs fundamentally from that of *B. subtilis*. In *E. coli*, external glucose is also transported and phosphorylated via the PTS system (Meadow *et al.*, 1990; Postma *et al.*, 1993). The expression of the glucose permease PtsG is negatively controlled by the global regulator Mlc on the level of transcription initiation (Kimata *et al.*, 1998; Tanaka *et al.*, 1999). When glucose is transported

into the cell, the membrane spanning domain of the glucose permease (enzyme $IIBC^{Glc}$) is mainly in a dephosphorylated state. Under these conditions, $IIBC^{Glc}$ binds Mlc resulting in sequestration of the transcription factor to the membrane (Lee *et al.*, 2000; Tanaka *et al.*, 2000). The membrane localization is the actual reason for Mlc inactivation (Tanaka *et al.*, 2004). Additionally, PtsG synthesis is also regulated on a post-transcriptional level by modulating mRNA stability. The small non-coding RNA SgrS is expressed and, with the help of Hfq (see chapter 1.2.2), interacts with the 5'-end of the *ptsG* mRNA resulting in SgrS-*ptsG* base pairing (Vanderpool and Gottesman, 2004). The SgrS-*ptsG* complex is then subjected to RNase E-dependent degradation (Kawamoto *et al.*, 2005; Kimata *et al.*, 2001; Morita *et al.*, 2004).

Although the implementation differs significantly, the PTS and regulatory RNA elements play a crucial role in both bacteria. Considering the special role of glucose within the PTS system, it can be assumed that the PTS evolved primarily as a glucose regulon. This was already proposed previously for *E. coli* (Plumbridge, 2002) and would give a plausible explanation for the outstanding position of glucose within the PTS in both bacteria. In addition, the PTS dependent antitermination systems of *B. subtilis* might also have evolved originally for the regulation of glucose uptake. However, the presence of PRD-II in GlcT, which is not of big regulatory relevance, could be an evidence for another origin of this system (Greenberg *et al.*, 2002). Maybe an existing antitermination system regulating a sugarpermease that is under catabolite repression in the presence of glucose (like *sacPA* or *bglPH*) was the origin of the GlcT-*ptsG* system. In the original system, PRD-II might be needed for switching the system off when both the respective sugar and glucose are present simultaneously (see also chapter 1.3.3). When the system adapted to glucose uptake regulation, the PRD-II lost its relevance.

More than just nucleotide-amino acid interaction mediates specificity

Other factors than specific nucleotide-amino acid interaction seem to have influence on the specificity of the four antiterminator systems. Although not bound by the antiterminator proteins, the top loops seem to have indirect influence on gene expression. The nucleotide sequences and especially the nucleotides found at position 13 and 18 (or 12 and 18 in the case of *ptsG*), which are proposed to be in direct base pairing (see Fig. 6.3), are found to have influence on the expression level of the respective genes (data not shown). The GAAA tetraloop motive found in *bglP* RAT is very common in RNA secondary structures and is known to stabilize adjacent stem structures (Conn and Draper, 1998; Moore, 1999; Murphy

and Cech, 1994). The exchange of a single nucleotide within the tetraloop can result in an overall destabilization of the RNA secondary structure. Thus, the configuration of the top loop has an indirect influence on the stability of the stem region and therefore on the expression level of the coding gene. The same was found for some nucleotides that are bound by the antiterminator proteins. It is proposed that U7 and G26 in *ptsG* RAT are in direct contact with amino acids of GlcT. Although the individual base is not relevant for interaction specificity, the type of nucleotide had an influence on the level of expression. An U7 U26 configuration resulted in a twofold higher expression rate, whereas a G7 G26 configuration resulted in a much lower expression rate in comparison to the wildtype RAT.

It was also found that nucleotides in the upper and lower loop region have a dual role by stabilizing the overall RAT structure and interacting with the respective protein. All RAT structures feature an adenine at position 27 (see Fig. 6.3) and this adenine is known to be in direct contact with the protein in the LicT-*bglP* RAT complex (see Fig. 6.4). Additionally, this nucleotide is in base pairing with an opposing G, U, or A and thus participating in stem structure formation. The Rev peptide of HIV-I is also recognizing an unusual GG base pairing in its target RNA. Furthermore, it was shown that binding of Rev induces the GG base pairing formation (Battiste *et al.*, 1994). It was also reported that non Watson-Crick base pairings are necessary to maintain stem formation in double stranded RNA structures and induce RNA bending (Baeyens *et al.*, 1996). The NMR structure of the LicT-*bglP* RAT complex clearly shows that the RNA structure is twisted (see Fig. 6.5) and the type of pairing at position 3 and 27 in the respective RAT might affect the torsion. Finally, the thermodynamic stability of the terminator structure has also influence on the overall expression rate of the coding genes. Considering all this, it is hard to separate the effects coming from a single nucleotide exchange. It could affect the stability of the RAT secondary structure, the strength of RAT-protein interaction, the torsion of the RAT structure, or, when compensatory mutations were introduced, the stability of the terminator structure.

In their active state, all four antiterminator proteins form dimers from identical subunits. As the monomers of the cognate proteins share a high homology among each other, it is reasonable to wonder whether the belonging subunits specifically interact with each other or whether heterodimers formed from different subunits might exist. The data obtained from site directed mutagenesis of *sacB* RAT revealed that one of the RAT mutants could only cause antitermination when both SacT and SacY were present and active. The deletion of either SacT or SacY resulted in transcription termination of the relevant RAT mutant (data not shown). Although this proposed heterodimer does not seem to interact with any of the wild

type RAT structures, it might be possible that this interaction helps to fine-tune the activity of SacT and SacY under certain conditions by forming inactive heterodimers.

Furthermore, there still seem to be unknown factors involved in specificity determination. SacT was found to bind to all *sacB* derived RAT structures. However, it is not reported that the presence of sucrose results in SacT dependent expression of *bglPH* or *licS*. On the other hand, a mutant RAT structure that was derived from *ptsG* RAT sequence and was changed towards *sacPA* RAT sequence exhibited strong LicT binding and only minor interaction with SacT was observed. This could indicate that sequences other than RAT and terminator might be relevant for the specific interaction of some RAT-protein couples. It is also noteworthy that the RNA-protein recognition is a highly dynamic process occurring while the RNA is just originating. The sequence upstream of the respective RAT might be involved in RAT structure formation or in guiding the antiterminator protein to the nascent RAT sequence. However, the dynamic RNA synthesis and folding process cannot be monitored in detail.

6.3 Outlook

The results in this work clearly show the influence of the organic acids glutamate and succinate on the central metabolism of *B. subtilis* when given in addition to glucose. The data obtained from the two methods used, cDNA microarray analysis and metabolic flux analysis, are in good agreement for most metabolic pathways. However, the data obtained with the microarray analysis for *pckA* and some overflow metabolic pathways are not congruent with the data gained from the metabolic flux analysis. This could indicate that the relevant enzymes are regulated rather on the level of protein activity than on the level of transcription or translation. Classical molecular biological methods and 2D gelelectrophoresis could give further evidence for the mode of regulation of these enzymes.

Furthermore, two modern and powerful screening methods were established and optimized for the analysis of the central metabolic pathways of *B. subtilis*. With these methods at hand, further analysis of the influence of nutrients on the central metabolism of *B subtilis* is a comparatively easy task.

The mutational analysis of the *ptsG* and *sacB* RAT sequences revealed the specificity determinants of all related RAT-protein couples. However, it is still unclear if other sequences than RAT and terminator might be involved in the dynamic process of RAT folding and protein recognition. The potential of site directed mutagenesis also comes to its limits as the consequence of a single nucleotide exchange can not be studied isolated from side effects (see chapter 6.2). Therefore, it is desirable to obtain crystal or NMR structures of the RAT mutants in complex with the recognized antiterminator protein. Moreover, a mutational analysis of the antiterminator proteins could give further insights into RNA-protein recognition and could elucidate how RNA and protein adapted to target recognition.

B. subtilis is able to grow in the range of 12°C to 52°C. The temperature sensitive RNA switches (see chapter 1.2.3.1) clearly indicate that even a small shift in temperature can alter RNA secondary structure. It might be interesting to study whether the PTS controlled antitermination systems are functional within the entire temperature range and how the systems are affected by heavy temperature shifts. The RATs and antiterminator proteins of thermophilic bacilli could be studied to address these questions.

7 References

Aigner, A. 2006. Gene silencing through RNA interference (RNAi) *in vivo*: strategies based on the direct application of siRNAs. *J Biotechnol* **124:**12-25.

Akashi, H. and T. Gojobori. 2002. Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci USA* **99:**3695-3700.

Albright, L. M., E. Huala and F. M. Ausubel. 1989. Prokaryotic signal transduction mediated by sensor and regulator protein pairs. *Annu Rev Genet* **23:**311-336.

Amarzguioui, M., J. J. Rossi and D. Kim. 2005. Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Lett* **579:**5974-5981.

Ambros, V. 2004. The functions of animal microRNAs. *Nature* **431:**350-355.

Amster-Choder, O., F. Houman and A. Wright. 1989. Protein phosphorylation regulates transcription of the beta-glucoside utilization operon in *E. coli*. *Cell* **58:**847-855.

Amster-Choder, O. and A. Wright. 1992. Modulation of the dimerization of a transcriptional antiterminator protein by phosphorylation. *Science* **257:**1395-1398.

Amster-Choder, O. and A. Wright. 1993. Transcriptional regulation of the *bgl* operon of *Escherichia coli* involves phosphotransferase system-mediated phosphorylation of a transcriptional antiterminator. *J Cell Biochem* **51:**83-90.

Antson, A. A., E. J. Dodson, G. Dodson, R. B. Greaves, X. Chen and P. Gollnick. 1999. Structure of the *trp* RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* **401:**235-242.

Antson, A. A., J. Otridge, A. M. Brzozowski, E. J. Dodson, G. G. Dodson, K. S. Wilson, T. M. Smith, M. Yang, T. Kurecki and P. Gollnick. 1995. The structure of *trp* RNA-binding attenuation protein. *Nature* **374:**693-700.

Aravin, A. and T. Tuschl. 2005. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett* **579:**5830-5840.

Arnaud, M., M. Débarbouillé, G. Rapoport, M. H. Saier and J. Reizer. 1996. *In vitro* reconstitution of transcriptional antitermination by the SacT and SacY proteins of *Bacillus subtilis*. *J Biol Chem* **271:**18966-18972.

Arnaud, M., P. Vary, M. Zagorec, A. Klier, M. Débarbouillé, P. Postma and G. Rapoport. 1992. Regulation of the *sacPA* operon of *Bacillus subtilis*: identification of phosphotransferase system components involved in SacT activity. *J Bacteriol* **174:**3161-3170.

Asai, K., S. H. Baik, Y. Kasahara, S. Moriya and N. Ogasawara. 2000. Regulation of the transport system for C4-dicarboxylic acids in *Bacillus subtilis*. *Microbiology* **146:**263-271.

Aymerich, S. and M. Steinmetz. 1992. Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. *Proc Natl Acad Sci USA* **89:**10410-10414.

Babitzke, P. 2004. Regulation of transcription attenuation and translation initiation by allosteric control of an RNA-binding protein: the *Bacillus subtilis* TRAP protein. *Curr Opin Microbiol* **7:**132-139.

Babitzke, P. and P. Gollnick. 2001. Posttranscription initiation control of tryptophan metabolism in *Bacillus subtilis* by the *trp* RNA-binding attenuation protein (TRAP), anti-TRAP, and RNA structure. *J Bacteriol* **183:**5795-5802.

Babitzke, P., J. Yealy and D. Campanelli. 1996. Interaction of the *trp* RNA-binding attenuation protein (TRAP) of *Bacillus subtilis* with RNA: effects of the number of GAG repeats, the nucleotides separating adjacent repeats, and RNA secondary structure. *J Bacteriol* **178:**5159-5163.

Bachem, S., N. Faires and J. Stülke. 1997. Characterization of the presumptive phosphorylation sites of the *Bacillus subtilis* glucose permease by site-directed mutagenesis: implication in glucose transport and catabolite repression. *FEMS Microbiol Lett* **156:**233-238.

Bachem, S. and J. Stülke. 1998. Regulation of the *Bacillus subtilis* GlcT antiterminator protein by components of the phosphotransferase system. *J Bacteriol* **180:**5319-5326.

Bae, W., B. Xia, M. Inouye and K. Severinov. 2000. *Escherichia coli* CspA-family RNA chaperones are transcription antiterminators. *Proc Natl Acad Sci USA* **97:**7784-7789.

Baeyens, K. J., H. L. De Bondt, A. Pardi and S. R. Holbrook. 1996. A curved RNA helix incorporating an internal loop with G·A and A·A non-Watson-Crick base pairing. *Proc Natl Acad Sci U S A* **93:**12851-12855.

Barash, D. 2003. Deleterious mutation prediction in the secondary structure of RNAs. *Nucleic Acids Res* **31:**6578-6584.

Barrick, J. E., K. A. Corbino, W. C. Winkler, A. Nahvi, M. Mandal, J. Collins, M. Lee, A. Roth, N. Sudarsan, I. Jona, J. K. Wickiser and R. R. Breaker. 2004. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proc Natl Acad Sci USA* **101:**6421-6426.

Bartel, D. P., M. L. Zapp, M. R. Green and J. W. Szostak. 1991. HIV-1 Rev regulation involves recognititon of non-Watson-Crick base pairs in viral RNA. *Cell* **67:**529-536.

Battiste, J. L., H. Mao, N. S. Rao, R. Tan, D. R. Muhandiram, L. E. Kay, A. D. Frankel and J. R. Williamson. 1996. α Helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. *Science* **273:**1547-1551.

Battiste, J. L., R. Y. Tan, A. D. Frankel and J. R. Williamson. 1994. Binding of an HIV Rev peptide to Rev responsive element RNA induces formation of purine-purine base pairs. *Biochemistry* **33:**2741-2747.

Baumann, C., S. Xirasagar and P. Gollnick. 1997. The *trp* RNA-binding attenuation protein (TRAP) from *Bacillus subtilis* binds to unstacked *trp* leader RNA. *J Biol Chem* **272:**19863-19869.

Begley, T. P., D. M. Downs, S. E. Ealick, F. W. McLafferty, A. P. Van Loon, S. Taylor, N. Campobasso, H. J. Chiu, C. Kinsland, J. J. Reddick and J. Xi. 1999. Thiamin biosynthesis in prokaryotes. *Arch Microbiol* **171:**293-300.

Belitsky, B. R. and A. L. Sonenshein. 1998. Role and regulation of *Bacillus subtilis* glutamate dehydrogenase genes. *J Bacteriol* **180:**6298-6305.

Belitsky, B. R., L. V. Wray, Jr., S. H. Fisher, D. E. Bohannon and A. L. Sonenshein. 2000. Role of TnrA in nitrogen source-dependent repression of *Bacillus subtilis* glutamate synthase gene expression. *J Bacteriol* **182:**5939-5947.

Bernhardt, J., J. Weibezahn, C. Scharf and M. Hecker. 2003. *Bacillus subtilis* during feast and famine: visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. *Genome Res* **13:**224-237.

Bi, W. and P. J. Stambrook. 1997. CCR: a rapid and simple approach for mutation detection. *Nucleic Acids Res* **25:**2949-2951.

Bi, W. and P. J. Stambrook. 1998. Site-directed mutagenesis by combined chain reaction. *Anal Biochem* **256:**137-140.

Black, D. L. 2000. Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* **103:**367-370.

Blank, L. M. and U. Sauer. 2004. TCA cycle activity in *Saccharomyces cerevisiae* is a function of the environmentally determined specific growth and glucose uptake rates. *Microbiology* **150:**1085-1093.

Blencke, H.-M., G. Homuth, H. Ludwig, U. Mäder, M. Hecker and J. Stülke. 2003. Transcriptional profiling of gene expression in response to glucose in *Bacillus subtilis*: regulation of the central metabolic pathways. *Metab Eng* **5:**133-149.

Blencke, H.-M., I. Reif, F. M. Commichau, C. Detsch, I. Wacker, H. Ludwig and J. Stülke. 2006. Regulation of *citB* expression in *Bacillus subtilis*: integration of multiple metabolic signals in the citrate pool and by the general nitrogen regulatory system. *Arch Microbiol* **185:**136-146.

Brantl, S. 2002. Antisense RNAs in plasmids: control of replication and maintenance. *Plasmid* **48:**165-173.

Bugala, K., M. Zywicki, E. Wyszko, M. Z. Barciszewska and J. Barciszewski. 2005. Riboswitches. *Postepy Biochem* **51:**111-119.

Burbulys, D., K. A. Trach and J. A. Hoch. 1991. Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell* **64:**545-552.

Canonaco, F., T. A. Hess, S. Heri, T. Wang, T. Szyperski and U. Sauer. 2001. Metabolic flux response to phosphoglucose isomerase knock-out in *Escherichia coli* and impact of overexpression of the soluble transhydrogenase UdhA. *FEMS Microbiol Lett* **204:**247-252.

Carthew, R. W. 2006. Gene regulation by microRNAs. *Curr Opin Genet Dev* **16:**203-208.

Chai, W. and V. Stewart. 1998. NasR, a novel RNA-binding protein, mediates nitrate-responsive transcription antitermination of the *Klebsiella oxytoca* M5al *nasF* operon leader *in vitro*. *J Mol Biol* **283:**339-351.

Chai, W. and V. Stewart. 1999. RNA sequence requirements for NasR-mediated, nitrate-responsive transcription antitermination of the *Klebsiella oxytoca M5al nasF* operon leader. *J Mol Biol* **292:**203-216.

Chambliss, G. H. 1993. Carbon source-mediated catabolite repression. *In*: *Bacillus subtilis and other gram-positive bacteria, biochemistry, physiology, and molecular genetics.* A. L. Sonenshein, J. A. Hoch and R. Losick. (eds.), ASM press, Washington DC, pp. 213-219.

Chander, P., K. M. Halbig, J. K. Miller, C. J. Fields, H. K. Bonner, G. K. Grabner, R. L. Switzer and J. L. Smith. 2005. Structure of the nucleotide complex of PyrR, the *pyr* attenuation protein from *Bacillus caldolyticus*, suggests dual regulation by pyrimidine and purine nucleotides. *J Bacteriol* **187:**1773-1782.

Chaudhry, R., N. Nisar, B. Hora, S. R. Chirasani and P. Malhotra. 2005. Expression and immunological characterization of the carboxy-terminal region of the P1 adhesin protein of *Mycoplasma pneumoniae*. *J Clin Microbiol* **43:**321-325.

Christensen, B., A. K. Gombert and J. Nielsen. 2002. Analysis of flux estimates based on ^{13}C-labelling experiments. *Eur J Biochem* **269:**2795-2800.

Christiansen, I. and W. Hengstenberg. 1996. Cloning and sequencing of two genes from *Staphylococcus carnosus* coding for glucose-specific PTS and their expression in *Escherichia coli* K-12. *Mol Gen Genet* **250:**375-379.

Christiansen, I. and W. Hengstenberg. 1999. Staphylococcal phosphoenolpyruvate-dependent phosphotransferase system–two highly similar glucose permeases in *Staphylococcus carnosus* with different glucoside specificity: protein engineering *in vivo*? *Microbiology* **145:**2881-2889.

Commichau, F. M., K. Forchhammer and J. Stülke. 2006. Regulatory links between carbon and nitrogen metabolism. *Curr Opin Microbiol* **9:**167-172.

Condon, C., M. Grunberg-Manago and H. Putzer. 1996. Aminoacyl-tRNA synthetase gene regulation in *Bacillus subtilis*. *Biochimie* **78:**381-389.

Conn, G. L. and D. E. Draper. 1998. RNA structure. *Curr Opin Struct Biol* **8:**278-285.

Covert, M. W., C. H. Schilling, I. Famili, J. S. Edwards, Goryanin, II, E. Selkov and B. O. Palsson. 2001. Metabolic modeling of microbial strains *in silico*. *Trends Biochem Sci* **26:**179-186.

Crutz, A. M., M. Steinmetz, S. Aymerich, R. Richter and D. Le Coq. 1990. Induction of levansucrase in *Bacillus subtilis*: an antitermination mechanism negatively controlled by the phosphotransferase system. *J Bacteriol* **172:**1043-1050.

Cruz Ramos, H., T. Hoffmann, M. Marino, H. Nedjari, E. Presecan-Siedel, O. Dreesen, P. Glaser and D. Jahn. 2000. Fermentative metabolism of *Bacillus subtilis*: physiology and regulation of gene expression. *J Bacteriol* **182:**3072-3080.

Daly, T. J., K. S. Cook, G. S. Gray, T. E. Maione and J. R. Rusche. 1989. Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element *in vitro*. *Nature* **342:**816-819.

Darbon, E., P. Servant, S. Poncet and J. Deutscher. 2002. Antitermination by GlpP, catabolite repression via CcpA and inducer exclusion triggered by P~GlpK dephosphorylation control *Bacillus subtilis glpFK* expression. *Mol Microbiol* **43:**1039-1052.

Dauner, M., J. E. Bailey and U. Sauer. 2001. Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. *Biotechnol Bioeng* **76:**144-156.

Dauner, M. and U. Sauer. 2001. Stoichiometric growth model for riboflavin-producing *Bacillus subtilis*. *Biotechnol Bioeng* **76:**132-143.

Dauner, M., M. Sonderegger, M. Hochuli, T. Szyperski, K. Wüthrich, H. P. Hohmann, U. Sauer and J. E. Bailey. 2002. Intracellular carbon fluxes in riboflavin-producing *Bacillus subtilis* during growth on two-carbon substrate mixtures. *Appl Environ Microbiol* **68:**1760-1771.

Débarbouillé, M., M. Arnaud, A. Fouet, A. Klier and G. Rapoport. 1990. The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. *J Bacteriol* **172:**3966-3973.

Declerck, N., F. Vincent, F. Hoh, S. Aymerich and H. van Tilbeurgh. 1999. RNA recognition by transcriptional antiterminators of the BglG/SacY family: functional and structural comparison of the CAT domain from SacY and LicT. *J Mol Biol* **294:**389-402.

Desai, S. K. and J. P. Gallivan. 2004. Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation. *J Am Chem Soc* **126:**13247-13254.

Detsch, C. and J. Stülke. 2003. Ammonium utilization in *Bacillus subtilis*: transport and regulatory functions of NrgA and NrgB. *Microbiology* **149:**3289-3297.

Deutscher, J., B. Pevec, K. Beyreuther, H. H. Kiltz and W. Hengstenberg. 1986. Streptococcal phosphoenolpyruvate-sugar phosphotransferase system: amino acid sequence and site of ATP-dependent phosphorylation of HPr. *Biochemistry* **25:**6543-6551.

Deutscher, J. and M. H. Saier, Jr. 1983. ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Streptococcus pyogenes*. *Proc Natl Acad Sci USA* **80:**6790-6794.

Diesterhaft, M. D. and E. Freese. 1973. Role of pyruvate carboxylase, phosphoenolpyruvate carboxykinase, and malic enzyme during growth and sporulation of *Bacillus subtilis*. *J Biol Chem* **248:**6062-6070.

Doan, T. and S. Aymerich. 2003. Regulation of the central glycolytic genes in *Bacillus subtilis*: binding of the repressor CggR to its single DNA target sequence is modulated by fructose-1,6-bisphosphate. *Mol Microbiol* **47:**1709-1721.

Doan, T., P. Servant, S. Tojo, H. Yamaguchi, G. Lerondel, K. Yoshida, Y. Fujita and S. Aymerich. 2003. The *Bacillus subtilis ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate. *Microbiology* **149:**2331-2343.

Domingo, E. and J. J. Holland. 1997. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* **51:**151-178.

Eigen, M., W. Gardiner, P. Schuster and R. Winkler-Oswatitsch. 1981. The origin of genetic information. *Sci Am* **244:**88-92.

El-Mansi, M., A. J. Cozzone, J. Shiloach and B. J. Eikmanns. 2006. Control of carbon flux through enzymes of central and intermediary metabolism during growth of *Escherichia coli* on acetate. *Curr Opin Microbiol* **9:**173-179.

Ellington, A. D. and J. W. Szostak. 1990. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346:**818-822.

Evans, P. R., G. W. Farrants and P. J. Hudson. 1981. Phosphofructokinase: structure and control. *Philos Trans R Soc Lond B Biol Sci* **293:**53-62.

Fabret, C., V. A. Feher and J. A. Hoch. 1999. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *J Bacteriol* **181:**1975-1983.

Faires, N., S. Tobisch, S. Bachem, I. Martin-Verstraete, M. Hecker and J. Stülke. 1999. The catabolite control protein CcpA controls ammonium assimilation in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* **1:**141-148.

Fang, L., W. Jiang, W. Bae and M. Inouye. 1997. Promoter-independent cold-shock induction of *cspA* and its derepression at 37 °C by mRNA stabilization. *Mol Microbiol* **23:**355-364.

Fauzi, H., K. D. Jack and J. V. Hines. 2005. *In vitro* selection to identify determinants in tRNA for *Bacillus subtilis tyrS* T box antiterminator mRNA binding. *Nucleic Acids Res* **33:**2595-2602.

Fillinger, S., S. Boschi-Muller, S. Azza, E. Dervyn, G. Branlant and S. Aymerich. 2000. Two glyceraldehyde-3-phosphate dehydrogenases with opposite physiological roles in a nonphotosynthetic bacterium. *J Biol Chem* **275:**14031-14037.

Fisher, S. H. 1999. Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la différence! *Mol Microbiol* **32:**223-232.

Fisher, S. H. and B. Magasanik. 1984. 2-Ketoglutarate and the regulation of aconitase and histidase formation in *Bacillus subtilis*. *J Bacteriol* **158:**379-382.

Fujita, Y., K. Yoshida, Y. Miwa, N. Yanai, E. Nagakawa and Y. Kasahara. 1998. Identification and expression of the *Bacillus subtilis* fructose-1,6-bisphosphatase gene (*fbp*). *J Bacteriol* **180:**4309-4313.

Fux, L., A. Nussbaum-Shochat and O. Amster-Choder. 2003. Interactions between the PTS regulation domains of the BglG transcriptional antiterminator from *Escherichia coli*. *J Biol Chem* **278:**46203-46209.

Galinier, A., J. Deutscher and I. Martin-Verstraete. 1999. Phosphorylation of either crh or HPr mediates binding of CcpA to the *Bacillus subtilis xyn cre* and catabolite repression of the *xyn* operon. *J Mol Biol* **286:**307-314.

Galinier, A., J. Haiech, M. C. Kilhoffer, M. Jaquinod, J. Stülke, J. Deutscher and I. Martin-Verstraete. 1997. The *Bacillus subtilis crh* gene encodes a HPr-like protein involved in carbon catabolite repression. *Proc Natl Acad Sci USA* **94:**8439-8444.

Galinier, A., M. Kravanja, R. Engelmann, W. Hengstenberg, M. C. Kilhoffer, J. Deutscher and J. Haiech. 1998. New protein kinase and protein phosphatase families mediate signal transduction in bacterial catabolite repression. *Proc Natl Acad Sci USA* **95:**1823-1828.

Galvao, T. C. and V. de Lorenzo. 2006. Transcriptional regulators *à la carte*: engineering new effector specificities in bacterial regulatory proteins. *Curr Opin Biotechnol* **17:**34-42.

Gardan, R., G. Rapoport and M. Débarbouillé. 1995. Expression of the *rocDEF* operon involved in arginine catabolism in *Bacillus subtilis*. *J Mol Biol* **249:**843-856.

Garmendia, J., D. Devos, A. Valencia and V. de Lorenzo. 2001. *À la carte* transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylR to non-natural effectors. *Mol Microbiol* **42:**47-59.

Gautheret, D. and R. R. Gutell. 1997. Inferring the conformation of RNA base pairs and triples from patterns of sequence variation. *Nucleic Acids Res* **25:**1559-1564.

Gelfand, M. S., A. A. Mironov, J. Jomantas, Y. I. Kozlov and D. A. Perumov. 1999. A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends Genet* **15:**439-442.

Gerdeman, M. S., T. M. Henkin and J. V. Hines. 2003. Solution structure of the *Bacillus subtilis* T-box antiterminator RNA: seven nucleotide bulge characterized by stacking and flexibility. *J Mol Biol* **326:**189-201.

Gilbert, W. 1986. The RNA world. *Nature (London)* **319:**618.

Gollnick, P. 1994. Regulation of the *Bacillus subtilis trp* operon by an RNA-binding protein. *Mol Microbiol* **11:**991-997.

Gollnick, P., C. Baumann, M. Yang, J. Otridge and A. Antson. 1995. Interaction of the 11-subunit *trp* RNA-binding attenuation protein (TRAP) with its RNA target. *Nucleic Acids Symp Ser* 43-45.

Gonzy-Tréboul, M. Zagorec, M. C. Rain-Guion and M. Steinmetz. 1989. Phosphoenolpyruvate:sugar phosphotransferase system of *Bacillus subtilis*: nucleotide sequence of *ptsX*, *ptsH* and the 5'-end of *ptsI* and evidence for a *ptsHI* operon. *Mol Microbiol* **3:**103-112.

Görke, B., L. Fraysse and A. Galinier. 2004. Drastic differences in Crh and HPr synthesis levels reflect their different impacts on catabolite repression in *Bacillus subtilis*. *J Bacteriol* **186:**2992-2995.

Gottesman, S. 2004. The small RNA regulators of *Escherichia coli*: roles and mechanisms. *Annu Rev Microbiol* **58:**303-328.

Gottesman, S. 2005. Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet* **21:**399-404.

Goudreau, P. N. and A. M. Stock. 1998. Signal transduction in bacteria: molecular mechanisms of stimulus-response coupling. *Curr Opin Microbiol* **1:**160-169.

Grabner, G. K. and R. L. Switzer. 2003. Kinetic studies of the uracil phosphoribosyltransferase reaction catalyzed by the *Bacillus subtilis* pyrimidine attenuation regulatory protein PyrR. *J Biol Chem* **278:**6921-6927.

Grahn, E., N. J. Stonehouse, J. B. Murray, S. van den Worm, K. Valegård, K. Fridborg, P. G. Stockley and L. Liljas. 1999. Crystallographic studies of RNA hairpins in complexes with recombinant MS2 capsids: implications for binding requirements. *RNA* **5:**131-138.

Greenberg, D. B., J. Stülke and M. H. Saier, Jr. 2002. Domain analysis of transcriptional regulators bearing PTS regulatory domains. *Res Microbiol* **153:**519-526.

Grundy, F. J., J. A. Collins, S. M. Rollins and T. M. Henkin. 2000. tRNA determinants for transcription antitermination of the *Bacillus subtilis tyrS* gene. *RNA* **6:**1131-1141.

Grundy, F. J. and T. M. Henkin. 1993. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* **74:**475-482.

Grundy, F. J. and T. M. Henkin. 1998. The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in gram-positive bacteria. *Mol Microbiol* **30:**737-749.

Grundy, F. J. and T. M. Henkin. 2003. The T box and S box transcription termination control systems. *Front Biosci* **8:**20-31.

Grundy, F. J. and T. M. Henkin. 2004. Regulation of gene expression by effectors that bind to RNA. *Curr Opin Microbiol* **7:**126-131.

Grundy, F. J., S. E. Hodil, S. M. Rollins and T. M. Henkin. 1997. Specificity of tRNA-mRNA interactions in *Bacillus subtilis tyrS* antitermination. *J Bacteriol* **179:**2587-2594.

Grundy, F. J., S. M. Rollins and T. M. Henkin. 1994. Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the *Bacillus subtilis tyrS* gene: a new role for the discriminator base. *J Bacteriol* **176:**4518-4526.

Grundy, F. J., D. A. Waters, S. H. Allen and T. M. Henkin. 1993. Regulation of the *Bacillus subtilis* acetate kinase gene by CcpA. *J Bacteriol* **175:**7348-7355.

Grundy, F. J., W. C. Winkler and T. M. Henkin. 2002. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. *Proc Natl Acad Sci USA* **99:**11121-11126.

Guérout-Fleury, A. M., K. Shazand, N. Frandsen and P. Stragier. 1995. Antibiotic-resistance cassettes for *Bacillus subtilis*. *Gene* **167:**335-336.

Guillier, M., S. Gottesman and G. Storz. 2006. Modulating the outer membrane with small RNAs. *Genes Dev* **20:**2338-2348.

Halbedel, S., C. Hames and J. Stülke. 2004. *In vivo* activity of enzymatic and regulatory components of the phosphoenolpyruvate:sugar phosphotransferase system in *Mycoplasma pneumoniae*. *J Bacteriol* **186:**7936-7943.

Hames, C., S. Halbedel, O. Schilling and J. Stülke. 2005. Multiple-mutation reaction: a method for simultaneous introduction of multiple mutations into the *glpK* gene of *Mycoplasma pneumoniae*. *Appl Environ Microbiol* **71:**4097-4100.

Hanson, K. G., K. Steinhauer, J. Reizer, W. Hillen and J. Stülke. 2002. HPr kinase/phosphatase of *Bacillus subtilis*: expression of the gene and effects of mutations on enzyme activity, growth and carbon catabolite repression. *Microbiology* **148:**1805-1811.

Hanson, S., G. Bauer, B. Fink and B. Suess. 2005. Molecular analysis of a synthetic tetracycline-binding riboswitch. *RNA* **11:**503-511.

Harrington, C. A., C. Rosenow and J. Retief. 2000. Monitoring gene expression using DNA microarrays. *Curr Opin Microbiol* **3:**285-291.

Harwood, C. R. 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol* **10:**247-256.

Hecker, M. 2003. A proteomic view of cell physiology of *Bacillus subtilis*–bringing the genome sequence to life. *Adv Biochem Eng Biotechnol* **83:**57-92.

Hecker, M. and S. Engelmann. 2000. Proteomics, DNA arrays and the analysis of still unknown regulons and unknown proteins of *Bacillus subtilis* and pathogenic gram-positive bacteria. *Int J Med Microbiol* **290:**123-134.

Hederstedt, L. 1993. The Krebs Citric Acid Cycle. *In*: *Bacillus subtilis and other gram-positive bacteria, biochemistry, physiology, and molecular genetics.* A. L. Sonenshein, J. A. Hoch and R. Losick. (eds.), ASM press, Washington DC, pp. 181-197.

Heidrich, N., A. Chinali, U. Gerth and S. Brantl. 2006. The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism. *Mol Microbiol* **62:**520-536.

Helmann, J. D. 1999. Anti-sigma factors. *Curr Opin Microbiol* **2:**135-141.

Hemilä, H., A. Palva, L. Paulin, L. Adler, S. Arvidson and I. Palva. 1991. The secretory S complex in *Bacillus subtilis* is identified as pyruvate dehydrogenase. *Res Microbiol* **142:**779-785.

Henkin, T. M. 1996. The role of CcpA transcriptional regulator in carbon metabolism in *Bacillus subtilis*. *FEMS Microbiol Lett* **135:**9-15.

Henkin, T. M. 2000. Transcription termination control in bacteria. *Curr Opin Microbiol* **3:**149-153.

Henkin, T. M., F. J. Grundy, W. L. Nicholson and G. H. Chambliss. 1991. Catabolite repression of alpha-amylase gene expression in *Bacillus subtilis* involves a trans-acting gene product homologous to the *Escherichia coli lacl* and *galR* repressors. *Mol Microbiol* **5:**575-584.

Henkin, T. M. and C. Yanofsky. 2002. Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *Bioessays* **24:**700-707.

Hermann, T. and D. J. Patel. 2000. Adaptive recognition by nucleic acid aptamers. *Science* **287:**820-825.

Hesselberth, J. R., M. P. Robertson, S. M. Knudsen and A. D. Ellington. 2003. Simultaneous detection of diverse analytes with an aptazyme ligase array. *Anal Biochem* **312:**106-112.

Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkl, B.-C. Li and R. Herrmann. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res* **24:**4420-4449.

Hoch, J. A. and K. I. Varughese. 2001. Keeping signals straight in phosphorelay signal transduction. *J Bacteriol* **183:**4941-4949.

Holtzclaw, W. D. and L. F. Chapman. 1975. Degradative acetolactate synthase of *Bacillus subtilis*: purification and properties. *J Bacteriol* **121:**917-922.

Hovey, R., S. Lentes, A. Ehrenreich, K. Salmon, K. Saba, G. Gottschalk, R. P. Gunsalus and U. Deppenmeier. 2005. DNA microarray analysis of *Methanosarcina mazei* Gö1 reveals adaptation to different methanogenic substrates. *Mol Genet Genomics* **273:**225-239.

Hueck, C. J., W. Hillen and M. H. Saier, Jr. 1994. Analysis of a cis-active sequence mediating catabolite repression in gram-positive bacteria. *Res Microbiol* **145:**503-518.

Huffman, J. L. and R. G. Brennan. 2002. Prokaryotic transcription regulators: more than just the helix-turn-helix motif. *Curr Opin Struct Biol* **12:**98-106.

Hutchison III, C. A., S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith and J. C. Venter. 1999. Global transposon mutagenesis and a minimal mycoplasma genome. *Science* **286:**2165-2169.

Huynen, M. A., T. Dandekar and P. Bork. 1999. Variation and evolution of the citric-acid cycle: a genomic perspective. *Trends Microbiol* **7:**281-291.

Inamine, J. M., K.-C. Ho, S. Loechel and P.-C. Hu. 1990. Evidence that UGA is read as a tryptophan codon rather than as a stop codon by *Mycoplasma pneumoniae*, *Mycoplasma genitalium*, and *Mycoplasma gallisepticum*. *J Bacteriol* **172:**504-506.

Jacob, F. and J. Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3:**318-356.

Jacobs, E. 1997. *Mycoplasma* infections of the human respiratory tract. *Wien Klin Wochenschr* **109:**574-577.

Jadhav, V. R. and M. Yarus. 2002. Coenzymes as coribozymes. *Biochimie* **84:**877-888.

Jaffe, J. D., H. C. Berg and G. M. Church. 2004. Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* **4:**59-77.

Jansen, J. A., T. J. McCarthy, G. A. Soukup and J. K. Soukup. 2006. Backbone and nucleobase contacts to glucosamine-6-phosphate in the *glmS* ribozyme. *Nat Struct Mol Biol* **13:**517-523.

Jault, J. M., S. Fieulaine, S. Nessler, P. Gonzalo, A. Di Pietro, J. Deutscher and A. Galinier. 2000. The HPr kinase from *Bacillus subtilis* is a homo-oligomeric enzyme which exhibits strong positive cooperativity for nucleotide and fructose-1,6-bisphosphate binding. *J Biol Chem* **275:**1773-1780.

Jeffares, D. C., A. M. Poole and D. Penny. 1998. Relics from the RNA world. *J Mol Evol* **46:**18-36.

Jeffery, C. J. 1999. Moonlighting proteins. *Trends Biochem Sci* **24:**8-11.

Jenal, U. and R. Hengge-Aronis. 2003. Regulation by proteolysis in bacterial cells. *Curr Opin Microbiol* **6:**163-172.

Jin, S., M. De Jesus-Berrios and A. L. Sonenshein. 1996. A *Bacillus subtilis* malate dehydrogenase gene. *J Bacteriol* **178:**560-563.

Johnston, W. K., P. J. Unrau, M. S. Lawrence, M. E. Glasner and D. P. Bartel. 2001. RNA-catalyzed RNA polymerization: accurate and general RNA-templated primer extension. *Science* **292:**1319-1325.

Jones, B. E., V. Dossonnet, E. Küster, W. Hillen, J. Deutscher and R. E. Klevit. 1997. Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. *J Biol Chem* **272:**26530-26535.

Jordan, S., A. Junker, J. D. Helmann and T. Mascher. 2006. Regulation of LiaRS-dependent gene expression in *Bacillus subtilis*: identification of inhibitor proteins, regulator binding sites, and target genes of a conserved cell envelope stress-sensing two-component system. *J Bacteriol* **188:**5153-5166.

Jourlin-Castelli, C., N. Mani, M. M. Nakano and A. L. Sonenshein. 2000. CcpC, a novel regulator of the LysR family required for glucose repression of the *citB* gene in *Bacillus subtilis*. *J Mol Biol* **295:**865-878.

Joyce, G. F. 1989. RNA evolution and the origins of life. *Nature* **338:**217-224.

Joyce, G. F. 2002. The antiquity of RNA-based evolution. *Nature* **418:**214-221.

Kannan, T. R. and J. B. Baseman. 2000. Expression of UGA-containing *Mycoplasma* genes in *Bacillus subtilis*. *J Bacteriol* **182:**2664-2667.

Kawamoto, H., T. Morita, A. Shimizu, T. Inada and H. Aiba. 2005. Implication of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*. *Genes Dev* **19:**328-338.

Kercher, M. A., P. Lu and M. Lewis. 1997. *Lac* repressor-operator complex. *Curr Opin Struct Biol* **7:**76-85.

Kiefer, P., E. Heinzle, O. Zelder and C. Wittmann. 2004. Comparative metabolic flux analysis of lysine-producing *Corynebacterium glutamicum* cultured on glucose or fructose. *Appl Environ Microbiol* **70:**229-239.

Kil, Y. V., V. N. Mironov, I. Gorishin, R. A. Kreneva and D. A. Perumov. 1992. Riboflavin operon of *Bacillus subtilis*: unusual symmetric arrangement of the regulatory region. *Mol Gen Genet* **233:**483-486.

Kim, H. J., S. I. Kim, M. Ratnayake-Lecamwasam, K. Tachikawa, A. L. Sonenshein and M. Strauch. 2003a. Complex regulation of the *Bacillus subtilis* aconitase gene. *J Bacteriol* **185:**1672-1680.

Kim, H. J., A. Roux and A. L. Sonenshein. 2002. Direct and indirect roles of CcpA in regulation of *Bacillus subtilis* Krebs cycle genes. *Mol Microbiol* **45:**179-190.

Kim, J. H., Z. T. Guvener, J. Y. Cho, K. C. Chung and G. H. Chambliss. 1995. Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. *J Bacteriol* **177:**5129-5134.

Kim, S. I., C. Jourlin-Castelli, S. R. Wellington and A. L. Sonenshein. 2003b. Mechanism of repression by *Bacillus subtilis* CcpC, a LysR family regulator. *J Mol Biol* **334:**609-624.

Kimata, K., T. Inada, H. Tagami and H. Aiba. 1998. A global repressor (Mlc) is involved in glucose induction of the *ptsG* gene encoding major glucose transporter in *Escherichia coli*. *Mol Microbiol* **29:**1509-1519.

Kimata, K., Y. Tanaka, T. Inada and H. Aiba. 2001. Expression of the glucose transporter gene, *ptsG*, is regulated at the mRNA degradation step in response to glycolytic flux in *Escherichia coli*. *Embo J* **20:**3587-3595.

Knezevic, I., S. Bachem, A. Sickmann, H. E. Meyer, J. Stülke and W. Hengstenberg. 2000. Regulation of the glucose-specific phosphotransferase system (PTS) of *Staphylococcus carnosus* by the antiterminator protein GlcT. *Microbiology* **146:**2333-2342.

Knudtson, K. L., M. Manohar, D. E. Joyner, E. A. Ahmed and B. C. Cole. 1997. Expression of the superantigen *Mycoplasma arthritidis* mitogen in *Escherichia coli* and characterization of the recombinant protein. *Infect Immun* **65:**4965-4971.

Koburger, T., J. Weibezahn, J. Bernhardt, G. Homuth and M. Hecker. 2005. Genome-wide mRNA profiling in glucose starved *Bacillus subtilis* cells. *Mol Genet Genomics* **274:**1-12.

Kochhar, S. and H. Paulus. 1996. Lysine-induced premature transcription termination in the *lysC* operon of *Bacillus subtilis*. *Microbiology* **142:**1635-1639.

Kraulis, P. J. 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* **24:**946-950.

Krömer, J. O., O. Sorgenfrei, K. Klopprogge, E. Heinzle and C. Wittmann. 2004. In-depth profiling of lysine-producing *Corynebacterium glutamicum* by combined analysis of the transcriptome, metabolome, and fluxome. *J Bacteriol* **186:**1769-1784.

Krüger, S., S. Gertz and M. Hecker. 1996. Transcriptional analysis of *bglPH* expression in *Bacillus subtilis*: evidence for two distinct pathways mediating carbon catabolite repression. *J Bacteriol* **178:**2637-2644.

Krüger, S. and M. Hecker. 1995. Regulation of the putative *bglPH* operon for aryl-ß-glucoside utilization in *Bacillus subtilis*. *J Bacteriol* **177:**5590-5597.

Kubodera, T., M. Watanabe, K. Yoshiuchi, N. Yamashita, A. Nishimura, S. Nakai, K. Gomi and H. Hanamoto. 2003. Thiamine-regulated gene expression of *Aspergillus oryzae thiA* requires splicing of the intron containing a riboswitch-like domain in the 5'-UTR. *FEBS Lett* **555:**516-520.

Kumar, P. K., T. Kumarevel and H. Mizuno. 2006. Structural basis of HutP-mediated transcription anti-termination. *Curr Opin Struct Biol* **16:**18-26.

Kunst, F., N. Ogasawara, I. Moszer, A. M. Albertini, G. Alloni, V. Azevedo, M. G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S. C. Brignell, S. Bron, S. Brouillet, C. V. Bruschi, B. Caldwell, V. Capuano, N. M. Carter, S. K. Choi, J. J. Codani, I. F. Connerton, A. Danchin and et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390:**249-256.

Kunst, F., M. Pascal, J. A. Lepesant, J. Walle and R. Dedonder. 1974. Purification and some properties of an endocellular sucrase from a constitutive mutant of *Bacillus subtilis* Marburg 168. *Eur J Biochem* **42:**611-620.

Kunst, F. and G. Rapoport. 1995. Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*. *J Bacteriol* **177:**2403-2407.

Lai, E. C. 2003. RNA sensors and riboswitches: self-regulating messages. *Curr Biol* **13:**R285-R291.

Langbein, I., S. Bachem and J. Stülke. 1999. Specific interaction of the RNA-binding domain of the *Bacillus subtilis* transcriptional antiterminator GlcT with its RNA target, RAT. *J Mol Biol* **293:**795-805.

Larralde, R., M. P. Robertson and S. L. Miller. 1995. Rates of decomposition of ribose and other sugars: implications for chemical evolution. *Proc Natl Acad Sci USA* **92:**8158-8160.

Laursen, B. S., H. P. Sørensen, K. K. Mortensen and H. U. Sperling-Petersen. 2005. Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* **69:**101-123.

Le Coq, D., C. Lindner, S. Krüger, M. Steinmetz and J. Stülke. 1995. New ß-glucoside (*bgl*) genes in *Bacillus subtilis*: the *bglP* gene product has both transport and regulatory functions similar to those of BglF, its *Escherichia coli* homolog. *J Bacteriol* **177:**1527-1535.

Lee, J. M., S. Zhang, S. Saha, S. Santa Anna, C. Jiang and J. Perkins. 2001. RNA expression analysis using an antisense *Bacillus subtilis* genome array. *J Bacteriol* **183:**7371-7380.

Lee, S. J., W. Boos, J. P. Bouche and J. Plumbridge. 2000. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of *Escherichia coli*. *Embo J* **19:**5353-5361.

Legault, P., J. Li, J. Mogridge, K. E. Lewis and J. Greenblatt. 1998. NMR structure of the bacteriophage λ N peptide/*boxB* RNA complex: recognition of a GNRA fold by an arginine-rich motive. *Cell* **93:**289-299.

Lenz, D. H., K. C. Mok, B. N. Lilley, R. V. Kulkarni, N. S. Wingreen and B. L. Bassler. 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118:**69-82.

Leontis, N. B., J. Stombaugh and E. Westhof. 2002. The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res* **30:**3497-3531.

Leontis, N. B. and E. Westhof. 2001. Geometric nomenclature and classification of RNA base pairs. *RNA* **7:**499-512.

Lescoute, A. and E. Westhof. 2005. Riboswitch structures: purine ligands replace tertiary contacts. *Chem Biol* **12:**10-13.

Lévine, A., F. Vannier, C. Absalon, L. Kuhn, P. Jackson, E. Scrivener, V. Labas, J. Vinh, P. Courtney, J. Garin and S. J. Séror. 2006. Analysis of the dynamic *Bacillus subtilis* Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes. *Proteomics* **6:**2157-2173.

Levit, M. N., Y. Liu and J. B. Stock. 1998. Stimulus response coupling in bacterial chemotaxis: receptor dimers in signalling arrays. *Mol Microbiol* **30:**459-466.

Leyva-Vazquez, M. A. and P. Setlow. 1994. Cloning and nucleotide sequences of the genes encoding triose phosphate isomerase, phosphoglycerate mutase, and enolase from *Bacillus subtilis*. *J Bacteriol* **176:**3903-3910.

Licht, A. and S. Brantl. 2006. Transcriptional repressor CcpN from *Bacillus subtilis* compensates asymmetric contact distribution by cooperative binding. *J Mol Biol* **364:**434-448.

Licht, A., S. Preis and S. Brantl. 2005. Implication of CcpN in the regulation of a novel untranslated RNA (SR1) in *Bacillus subtilis*. *Mol Microbiol* **58:**189-206.

Lillo, F., S. Basile and R. N. Mantegna. 2002. Comparative genomics study of inverted repeats in bacteria. *Bioinformatics* **18:**971-979.

Lind, K. 1983. Manifestations and complications of *Mycoplasma pneumoniae* disease: a review. *Yale J Biol Med* **56:**461-468.

Lindner, C., A. Galinier, M. Hecker and J. Deutscher. 1999. Regulation of the activity of the *Bacillus subtilis* antiterminator LicT by multiple PEP-dependent, enzyme I- and HPr-catalysed phosphorylation. *Mol Microbiol* **31:**995-1006.

Llanos, R. M., C. J. Harris, A. J. Hillier and B. E. Davidson. 1993. Identification of a novel operon in *Lactococcus lactis* encoding three enzymes for lactic acid synthesis: phosphofructokinase, pyruvate kinase, and lactate dehydrogenase. *J Bacteriol* **175:**2541-2551.

Long, M. B., J. P. Jones, 3rd, B. A. Sullenger and J. Byun. 2003. Ribozyme-mediated revision of RNA and DNA. *J Clin Invest* **112:**312-318.

Lu, Y., T. N. Shevtchenko and H. Paulus. 1992. Fine-structure mapping of *cis*-acting control sites in the *lysC* operon of *Bacillus subtilis*. *FEMS Microbiol Lett* **71:**23-27.

Ludwig, H., G. Homuth, M. Schmalisch, F. M. Dyka, M. Hecker and J. Stülke. 2001. Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. *Mol Microbiol* **41:**409-422.

Ludwig, H., C. Meinken, A. Matin and J. Stülke. 2002a. Insufficient expression of the *ilv-leu* operon encoding enzymes of branched-chain amino acid biosynthesis limits growth of a *Bacillus subtilis ccpA* mutant. *J Bacteriol* **184:**5174-5178.

Ludwig, H., N. Rebhan, H. M. Blencke, M. Merzbacher and J. Stülke. 2002b. Control of the glycolytic *gapA* operon by the catabolite control protein A in *Bacillus subtilis*: a novel mechanism of CcpA-mediated regulation. *Mol Microbiol* **45:**543-553.

Luesink, E. J., R. E. van Herpen, B. P. Grossiord, O. P. Kuipers and W. M. de Vos. 1998. Transcriptional activation of the glycolytic *las* operon and catabolite repression of the *gal* operon in *Lactococcus lactis* are mediated by the catabolite control protein CcpA. *Mol Microbiol* **30:**789-798.

Mahadevan, S. and A. Wright. 1987. A bacterial gene involved in transcription antitermination: regulation at a *rho*-independent terminator in the *bgl* operon of *E. coli*. *Cell* **50:**485-494.

Majdalani, N., C. Cunning, D. Sledjeski, T. Elliott and S. Gottesman. 1998. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* **95:**12462-12467.

Mandal, M., B. Boese, J. E. Barrick, W. C. Winkler and R. R. Breaker. 2003. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. *Cell* **113:**577-586.

Mandal, M. and R. R. Breaker. 2004. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nat Struct Mol Biol* **11:**29-35.

Mandal, M., M. Lee, J. E. Barrick, Z. Weinberg, G. M. Emilsson, W. L. Ruzzo and R. R. Breaker. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* **306:**275-279.

Manival, X., Y. Yang, M. P. Strub, M. Kochoyan, M. Steinmetz and S. Aymerich. 1997. From genetic to structural characterization of a new class of RNA-binding domain within the SacY/BglG family of antiterminator proteins. *EMBO J* **16:**5019-5029.

Martin-Verstraete, I., V. Charrier, J. Stülke, A. Galinier, B. Erni, G. Rapoport and J. Deutscher. 1998. Antagonistic effects of dual PTS-catalysed phosphorylation on the *Bacillus subtilis* transcriptional activator LevR. *Mol Microbiol* **28:**293-303.

Martin-Verstraete, I., M. Débarbouillé, A. Klier and G. Rapoport. 1990. Levanase operon of *Bacillus subtilis* includes a fructose-specific phosphotransferase system regulating the expression of the operon. *J Mol Biol* **214:**657-671.

Martin-Verstraete, I., M. Débarbouillé, A. Klier and G. Rapoport. 1992. Mutagenesis of the *Bacillus subtilis* "-12, -24" promoter of the levanase operon and evidence for the existence of an upstream activating sequence. *J Mol Biol* **226:**85-99.

Massé, E., F. E. Escorcia and S. Gottesman. 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* **17:**2374-2383.

Massé, E., C. K. Vanderpool and S. Gottesman. 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* **187:**6962-6971.

Mathews, D. H., M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker and D. H. Turner. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci USA* **101:**7287-7292.

Mathews, D. H., J. Sabina, M. Zuker and D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* **288:**911-940.

Mattick, J. S. and I. V. Makunin. 2005. Small regulatory RNAs in mammals. *Hum Mol Genet* **14 Spec No 1:**R121-132.

Meadow, N. D., D. K. Fox and S. Roseman. 1990. The bacterial phosphoenolpyruvate: glycose phosphotransferase system. *Annu Rev Biochem* **59:**497-542.

Meinken, C., H.-M. Blencke, H. Ludwig and J. Stülke. 2003. Expression of the glycolytic *gapA* operon in *Bacillus subtilis*: differential syntheses of proteins encoded by the operon. *Microbiology* **149:**751-761.

Meinkoth, J. and G. Wahl. 1984. Hybridization of nucleic acids immobilized on solid supports. *Anal Biochem* **138:**267-284.

Merianos, H. J., J. Wang and P. B. Moore. 2004. The structure of a ribosomal protein S8/spc operon mRNA complex. *RNA* **10:**954-964.

Miranda-Rios, J., C. Morera, H. Taboada, A. Davalos, S. Encarnacion, J. Mora and M. Soberon. 1997. Expression of thiamin biosynthetic genes (*thiCOGE*) and production of symbiotic terminal oxidase *cbb3 in Rhizobium etli*. *J Bacteriol* **179:**6887-6893.

Miranda-Rios, J., M. Navarro and M. Soberon. 2001. A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proc Natl Acad Sci USA* **98:**9736-9741.

Mironov, A. S., I. Gusarov, R. Rafikov, L. E. Lopez, K. Shatalin, R. A. Kreneva, D. A. Perumov and E. Nudler. 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111:**747-756.

Mironov, V. N., A. S. Kraev, M. L. Chikindas, B. K. Chernov, A. I. Stepanov and K. G. Skryabin. 1994. Functional organization of the riboflavin biosynthesis operon from *Bacillus subtilis SHgw*. *Mol Gen Genet* **242:**201-208.

Miwa, Y., A. Nakata, A. Ogiwara, M. Yamamoto and Y. Fujita. 2000. Evaluation and characterization of catabolite-responsive elements (*cre*) of *Bacillus subtilis*. *Nucleic Acids Res* **28:**1206-1210.

Mizuno, T. 1997. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res* **4:**161-168.

Mizuno, T., M. Y. Chou and M. Inouye. 1984. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci USA* **81:**1966-1970.

Mogk, A., R. Hayward and W. Schumann. 1996. Integrative vectors for constructing single-copy transcriptional fusions between *Bacillus subtilis* promoters and various reporter genes encoding heat-stable enzymes. *Gene* **182:**33-36.

Moll, I., T. Afonyushkin, O. Vytvytska, V. R. Kaberdin and U. Blasi. 2003. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **9:**1308-1314.

Molle, V., Y. Nakaura, R. P. Shivers, H. Yamaguchi, R. Losick, Y. Fujita and A. L. Sonenshein. 2003. Additional targets of the *Bacillus subtilis* global regulator CodY identified by chromatin immunoprecipitation and genome-wide transcript analysis. *J Bacteriol* **185:**1911-1922.

Møller, T., T. Franch, P. Hojrup, D. R. Keene, H. P. Bachinger, R. G. Brennan and P. Valentin-Hansen. 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9:**23-30.

Moore, P. B. 1999. Structural motifs in RNA. *Annu Rev Biochem* **68:**287-300.

Moreno, M. S., B. L. Schneider, R. R. Maile, W. Weyler and M. H. Saier, Jr. 2001. Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. *Mol Microbiol* **39:**1366-1381.

Morita, T., H. Kawamoto, T. Mizota, T. Inada and H. Aiba. 2004. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol Microbiol* **54:**1063-1075.

Müller, M. 2001. NMR-Untersuchungen an einer gefalteten RNA. Diplomarbeit, Friedrich-Alexander-Universität Erlangen-Nürnberg.

Müller, M., J. E. Weigand, O. Weichenrieder and B. Suess. 2006. Thermodynamic characterization of an engineered tetracycline-binding riboswitch. *Nucleic Acids Res* **34:**2607-2617.

Münchbach, M., A. Nocker and F. Narberhaus. 1999. Multiple small heat shock proteins in *rhizobia*. *J Bacteriol* **181:**83-90.

Murphy, F. L. and T. R. Cech. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J Mol Biol* **236:**49-63.

Nagai, K. 1996. RNA-protein complexes. *Curr Opin Struct Biol* **6:**53-61.

Narberhaus, F. 2002. mRNA-mediated detection of environmental conditions. *Arch Microbiol* **178:**404-410.

Narberhaus, F., R. Käser, A. Nocker and H. Hennecke. 1998. A novel DNA element that controls bacterial heat shock gene expression. *Mol Microbiol* **28:**315-323.

Narberhaus, F., T. Waldminghaus and S. Chowdhury. 2006. RNA thermometers. *FEMS Microbiol Rev* **30:**3-16.

Narberhaus, F., W. Weiglhofer, H. M. Fischer and H. Hennecke. 1996. The *Bradyrhizobium japonicum rpoH*$_1$ gene encoding a σ^{32}-like protein is part of a unique heat shock gene cluster together with *groESL*$_1$ and three small heat shock genes. *J Bacteriol* **178:**5337-5346.

Neves, A. R., R. Ventura, N. Mansour, C. Shearman, M. J. Gasson, C. Maycock, A. Ramos and H. Santos. 2002. Is the glycolytic flux in *Lactococcus lactis* primarily controlled by the redox charge? Kinetics of NAD^+ and NADH pools determined *in vivo* by ^{13}C NMR. *J Biol Chem* **277:**28088-28098.

Niebisch, A., A. Kabus, C. Schultz, B. Weil and M. Bott. 2006. Corynebacterial protein kinase G controls 2-oxoglutarate dehydrogenase activity via the phosphorylation status of the OdhI protein. *J Biol Chem* **281:**12300-12307.

Nielsen, J. 2003. It is all about metabolic fluxes. *J Bacteriol* **185:**7031-7035.

Nobeli, I., R. A. Laskowski, W. S. Valdar and J. M. Thornton. 2001. On the molecular discrimination between adenine and guanine by proteins. *Nucleic Acids Res* **29:**4294-4309.

Nocker, A., N. P. Krstulovic, X. Perret and F. Narberhaus. 2001. ROSE elements occur in disparate *rhizobia* and are functionally interchangeable between species. *Arch Microbiol* **176:**44-51.

Noeske, J., C. Richter, M. A. Grundl, H. R. Nasiri, H. Schwalbe and J. Wöhnert. 2005. An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. *Proc Natl Acad Sci USA* **102:**1372-1377.

Noh, E. J., S. W. Kang, Y. J. Shin, D. C. Kim, I. S. Park, M. Y. Kim, B. G. Chun and B. H. Min. 2002. Characterization of *Mycoplasma* arginine deiminase expressed in *E. coli* and its inhibitory regulation of nitric oxide synthesis. *Mol Cells* **13:**137-143.

Noller, H. F. 1993. Peptidyl transferase: protein, ribonucleoprotein, or RNA? *J Bacteriol* **175:**5297-5300.

Nomura, M., J. L. Yates, D. Dean and L. E. Post. 1980. Feedback regulation of ribosomal protein gene expression in *Escherichia coli*: structural homology of ribosomal RNA and ribosomal protein mRNA. *Proc Natl Acad Sci USA* **77:**7084-7088.

Nudler, E. 2006. Flipping riboswitches. *Cell* **126:**19-22.

Nudler, E. and A. S. Mironov. 2004. The riboswitch control of bacterial metabolism. *Trends Biochem Sci* **29:**11-17.

Oda, M., N. Kobayashi, A. Ito, Y. Kurusu and K. Taira. 2000. *cis*-acting regulatory sequences for antitermination in the transcript of the *Bacillus subtilis hut* operon and histidine-dependent binding of HutP to the transcript containing the regulatory sequences. *Mol Microbiol* **35:**1244-1254.

Opdyke, J. A., J. G. Kang and G. Storz. 2004. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* **186:**6698-6705.

Orth, P., D. Schnappinger, W. Hillen, W. Saenger and W. Hinrichs. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7:**215-219.

Otridge, J. and P. Gollnick. 1993. MtrB from *Bacillus subtilis* binds specifically to *trp* leader RNA in a tryptophan-dependent manner. *Proc Natl Acad Sci USA* **90:**128-132.

Pabo, C. O. and R. T. Sauer. 1984. Protein-DNA recognition. *Annu Rev Biochem* **53:**293-321.

Pané-Farré, J., R. J. Lewis and J. Stülke. 2005. The RsbRST stress module in bacteria: a signalling system that may interact with different output modules. *J Mol Microbiol Biotechnol* **9:**65-76.

Petersen, L. A. and D. M. Downs. 1997. Identification and characterization of an operon in *Salmonella typhimurium* involved in thiamine biosynthesis. *J Bacteriol* **179:**4894-4900.

Pfeiffer, T., S. Schuster and S. Bonhoeffer. 2001. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292:**504-507.

Plumbridge, J. 2002. Regulation of gene expression in the PTS in *Escherichia coli*: the role and interactions of Mlc. *Curr Opin Microbiol* **5:**187-193.

Pollack, J. D., M. A. Myers, T. Dandekar and R. Herrmann. 2002. Suspected utility of enzymes with multiple activities in the small genome *Mycoplasma* species: the replacement of the missing "household" nucleoside diphosphate kinase gene and activity by glycolytic kinases. *OMICS* **6:**247-258.

Poolman, B. 1993. Energy transduction in lactic acid bacteria. *FEMS Microbiol Rev* **12:**125-147.

Postma, P. W., J. W. Lengeler and G. R. Jacobson. 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* **57:**543-594.

Presecan-Siedel, E., A. Galinier, R. Longin, J. Deutscher, A. Danchin, P. Glaser and I. Martin-Verstraete. 1999. Catabolite regulation of the *pta* gene as part of carbon flow pathways in *Bacillus subtilis*. *J Bacteriol* **181:**6889-6897.

Putzer, H., C. Condon, D. Brechemier-Baey, R. Brito and M. Grunberg-Manago. 2002. Transfer RNA-mediated antitermination *in vitro*. *Nucleic Acids Res* **30:**3026-3033.

Quinn, C. L., B. T. Stephenson and R. L. Switzer. 1991. Functional organization and nucleotide sequence of the *Bacillus subtilis* pyrimidine biosynthetic operon. *J Biol Chem* **266:**9113-9127.

Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan and R. Tobes. 2005. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* **69:**326-356.

Ramström, H., S. Sanglier, E. Leize-Wagner, C. Philippe, A. Van Dorsselaer and J. Haiech. 2003. Properties and regulation of the bifunctional enzyme HPr kinase/phosphatase in *Bacillus subtilis*. *J Biol Chem* **278:**1174-1185.

Reents, H., R. Münch, T. Dammeyer, D. Jahn and E. Härtig. 2006. The Fnr regulon of *Bacillus subtilis*. *J Bacteriol* **188:**1103-1112.

Reichow, S. and G. Varani. 2006. Structural biology: RNA switches function. *Nature* **441:**1054-1055.

Reizer, J., C. Hoischen, F. Titgemeyer, C. Rivolta, R. Rabus, J. Stülke, D. Karamata, M. H. Saier, Jr. and W. Hillen. 1998. A novel protein kinase that controls carbon catabolite repression in bacteria. *Mol Microbiol* **27:**1157-1169.

Reizer, J., M. H. Saier, Jr., J. Deutscher, F. Grenier, J. Thompson and W. Hengstenberg. 1988. The phosphoenolpyruvate:sugar phosphotransferase system in gram-positive bacteria: properties, mechanism, and regulation. *Crit Rev Microbiol* **15:**297-338.

Renna, M. C., N. Najimudin, L. R. Winik and S. A. Zahler. 1993. Regulation of the *Bacillus subtilis alsS*, *alsD*, and *alsR* genes involved in post-exponential-phase production of acetoin. *J Bacteriol* **175:**3863-3875.

Repoila, F. and S. Gottesman. 2003. Temperature sensing by the *dsrA* promoter. *J Bacteriol* **185:**6609-6614.

Rodionov, D. A., A. G. Vitreschak, A. A. Mironov and M. S. Gelfand. 2002. Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J Biol Chem* **277:**48949-48959.

Romby, P., F. Vandenesch and E. G. Wagner. 2006. The role of RNAs in the regulation of virulence-gene expression. *Curr Opin Microbiol* **9:**229-236.

Rosenkrantz, M. S., D. W. Dingman and A. L. Sonenshein. 1985. *Bacillus subtilis citB* gene is regulated synergistically by glucose and glutamine. *J Bacteriol* **164:**155-164.

Rutberg, B. 1997. Antitermination of transcription of catabolic operons. *Mol Microbiol* **23:**413-421.

Saenger, W., P. Orth, C. Kisker, W. Hillen and W. Hinrichs. 2000. The Tetracycline Repressor-A Paradigm for a Biological Switch. *Angew Chem Int Ed Engl* **39:**2042-2052.

Saier, M. H., Jr., S. R. Goldman, R. R. Maile, M. S. Moreno, W. Weyler, N. Yang and I. T. Paulsen. 2002. Transport capabilities encoded within the *Bacillus subtilis* genome. *J Mol Microbiol Biotechnol* **4:**37-67.

Sambrook, J., E. F. Fritsch and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sarsero, J. P., E. Merino and C. Yanofsky. 2000. A *Bacillus subtilis* operon containing genes of unknown function senses tRNATrp charging and regulates expression of the genes of tryptophan biosynthesis. *Proc Natl Acad Sci USA* **97:**2656-2661.

Sauer, U. 2004. High-throughput phenomics: experimental methods for mapping fluxomes. *Curr Opin Biotechnol* **15:**58-63.

Sauer, U. and B. J. Eikmanns. 2005. The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol Rev* **29:**765-794.

Sauer, U., V. Hatzimanikatis, J. E. Bailey, M. Hochuli, T. Szyperski and K. Wüthrich. 1997. Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. *Nat Biotechnol* **15:**448-452.

Sauer, U., V. Hatzimanikatis, H. P. Hohmann, M. Manneberg, A. P. van Loon and J. E. Bailey. 1996. Physiology and metabolic fluxes of wild-type and riboflavin-producing *Bacillus subtilis*. *Appl Environ Microbiol* **62:**3687-3696.

Schägger, H. and G. von Jagow. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166:**368-379.

Schell, M. A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu Rev Microbiol* **47:**597-626.

Schilling, O., O. Frick, C. Herzberg, A. Ehrenreich, E. Heinzle, C. Wittmann and J. Stülke. 2007. Transcriptional and Metabolic Responses of *Bacillus subtilis* to the Availability of Organic Acids: Transcription Regulation Is Important but Not Sufficient To Account for Metabolic Adaptation. *Appl Environ Microbiol* **73:**499-507.

Schilling, O., C. Herzberg, T. Hertrich, H. Vorsmann, D. Jessen, S. Hübner, F. Titgemeyer and J. Stülke. 2006. Keeping signals straight in transcription regulation: specificity determinants for the interaction of a family of conserved bacterial RNA-protein couples. *Nucleic Acids Res* **34:**6102-6115.

Schilling, O., I. Langbein, M. Müller, M. Schmalisch and J. Stülke. 2004. A protein-dependent riboswitch controlling *ptsGHI* operon expression in *Bacillus subtilis*: RNA structure rather than sequence provides interaction specificity. *Nucleic Acids Res* **32:**2853-2864.

Schirmer, F., S. Ehrt and W. Hillen. 1997. Expression, inducer spectrum, domain structure, and function of MopR, the regulator of phenol degradation in *Acinetobacter calcoaceticus* NCIB8250. *J Bacteriol* **179:**1329-1336.

Schmalisch, M. H., S. Bachem and J. Stülke. 2003. Control of the *Bacillus subtilis* antiterminator protein GlcT by phosphorylation: elucidation of the phosphorylation chain leading to inactivation of GlcT. *J Biol Chem* **278:**51108-51115.

Schnetz, K. and B. Rak. 1990. Beta-glucoside permease represses the *bgl* operon of *Escherichia coli* by phosphorylation of the antiterminator protein and also interacts with glucose-specific enzyme III, the key element in catabolite control. *Proc Natl Acad Sci USA* **87:**5074-5078.

Schnetz, K., J. Stülke, S. Gertz, S. Krüger, M. Krieg, M. Hecker and B. Rak. 1996. LicT, a *Bacillus subtilis* transcriptional antiterminator protein of the BglG family. *J Bacteriol* **178:**1971-1979.

Schreier, H. J. 1993. Biosynthesis of glutamine and glutamate and the assimilation of ammonia. *In*: *Bacillus subtilis and other gram-positive bacteria, biochemistry, physiology, and molecular genetics.* A. L. Sonenshein, J. A. Hoch and R. Losick. (eds.), ASM press, Washington DC, pp. 281-298.

Schreier, H. J., S. W. Brown, K. D. Hirschi, J. F. Nomellini and A. L. Sonenshein. 1989. Regulation of *Bacillus subtilis* glutamine synthetase gene expression by the product of the *glnR* gene. *J Mol Biol* **210:**51-63.

Schroeder, R., A. Barta and K. Semrad. 2004. Strategies for RNA folding and assembly. *Nat Rev Mol Cell Biol* **5:**908-919.

Schumacher, M. A., K. Y. Choi, H. Zalkin and R. G. Brennan. 1994. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by alpha helices. *Science* **266:**763-770.

Seetharaman, S., M. Zivarts, N. Sudarsan and R. R. Breaker. 2001. Immobilized RNA switches for the analysis of complex chemical and biological mixtures. *Nat Biotechnol* **19:**336-341.

Serganov, A., A. Polonskaia, A. T. Phan, R. R. Breaker and D. J. Patel. 2006. Structural basis for gene regulation by a thiamine pyrophosphate-sensing riboswitch. *Nature* **441:**1167-1171.

Servant, P., D. Le Coq and S. Aymerich. 2005. CcpN (YqzB), a novel regulator for CcpA-independent catabolite repression of *Bacillus subtilis* gluconeogenic genes. *Mol Microbiol* **55:**1435-1451.

Sewer, A., N. Paul, P. Landgraf, A. Aravin, S. Pfeffer, M. J. Brownstein, T. Tuschl, E. van Nimwegen and M. Zavolan. 2005. Identification of clustered microRNAs using an *ab initio* prediction method. *BMC Bioinformatics* **6:**267.

Shevtsov, M. B., Y. Chen, P. Gollnick and A. A. Antson. 2005. Crystal structure of *Bacillus subtilis* anti-TRAP protein, an antagonist of TRAP/RNA interaction. *Proc Natl Acad Sci USA* **102:**17600-17605.

Shin, B. S., S. K. Choi and S. H. Park. 1999. Regulation of the *Bacillus subtilis* phosphotransacetylase gene. *J Biochem* **126:**333-339.

Shivers, R. P., S. S. Dineen and A. L. Sonenshein. 2006. Positive regulation of *Bacillus subtilis ackA* by CodY and CcpA: establishing a potential hierarchy in carbon flow. *Mol Microbiol* **62:**811-822.

Shivers, R. P. and A. L. Sonenshein. 2004. Activation of the *Bacillus subtilis* global regulator CodY by direct interaction with branched-chain amino acids. *Mol Microbiol* **53:**599-611.

Shivers, R. P. and A. L. Sonenshein. 2005. *Bacillus subtilis ilvB* operon: an intersection of global regulons. *Mol Microbiol* **56:**1549-1559.

Silvaggi, J. M., J. B. Perkins and R. Losick. 2005. Small untranslated RNA antitoxin in *Bacillus subtilis*. *J Bacteriol* **187:**6641-6650.

Silvaggi, J. M., J. B. Perkins and R. Losick. 2006. Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. *J Bacteriol* **188:**532-541.

Simic, P., J. Willuhn, H. Sahm and L. Eggeling. 2002. Identification of *glyA* (encoding serine hydroxymethyltransferase) and its use together with the exporter ThrE to increase L-threonine accumulation by *Corynebacterium glutamicum*. *Appl Environ Microbiol* **68:**3321-3327.

Simoneau, P., C.-M. Li, S. Loechel, R. Wenzel, R. Herrmann and P.-C. Hu. 1993. Codon reading scheme in *Mycoplasma pneumoniae* revealed by the analysis of the complete set of tRNA genes. *Nucleic Acids Res* **21:**4967-4974.

Simoni, R. D., M. Levinthal, F. D. Kundig, W. Kundig, B. Anderson, P. E. Hartman and S. Roseman. 1967. Genetic evidence for the role of a bacterial phosphotransferase system in sugar transport. *Proc Natl Acad Sci USA* **58:**1963-1970.

Skarlatos, P. and M. K. Dahl. 1998. The glucose kinase of *Bacillus subtilis*. *J Bacteriol* **180:**3222-3226.

Smiley, B. K. and F. C. Minion. 1993. Enhanced readthrough of opal (UGA) stop codons and production of *Mycoplasma pneumoniae* P1 epitopes in *Escherichia coli*. *Gene* **134:**33-40.

Sonenshein, A. L. 2002. The Krebs citric acid cycle. *In*: *Bacillus subtilis and its closest relatives: from genes to cells.* A. L. Sonenshein, J. A. Hoch and R. Losick. (eds.), ASM Press, Washington, DC, pp. 151-162.

Sonntag, K., L. Eggeling, A. A. De Graaf and H. Sahm. 1993. Flux partitioning in the split pathway of lysine synthesis in *Corynebacterium glutamicum*. Quantification by ^{13}C- and ^{1}H-NMR spectroscopy. *Eur J Biochem* **213:**1325-1331.

Springer, M., C. Protier and M. Grunberg-Manago. 1997. RNA mimicry in the translational apparatus. *In*: *RNA structure and function.* R. W. Simons and M. Grunberg-Manago. (eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., pp. 377-414.

Stamburski, C., J. Renaudin and J. M. Bove. 1991. First step toward a virus-derived vector for gene cloning and expression in spiroplasmas, organisms which read UGA as a tryptophan codon: synthesis of chloramphenicol acetyltransferase in *Spiroplasma citri*. *J Bacteriol* **173:**2225-2230.

Steinhauer, K., T. Jepp, W. Hillen and J. Stülke. 2002. A novel mode of control of *Mycoplasma pneumoniae* HPr kinase/phosphatase activity reflects its parasitic lifestyle. *Microbiology* **148:**3277-3284.

Steinmetz, M., D. Le Coq and S. Aymerich. 1989. Induction of saccharolytic enzymes by sucrose in *Bacillus subtilis*: evidence for two partially interchangeable regulatory pathways. *J Bacteriol* **171:**1519-1523.

Steinmetz, M., D. Le Coq, S. Aymerich, G. Gonzy-Tréboul and P. Gay. 1985. The DNA sequence of the gene for the secreted *Bacillus subtilis* enzyme levansucrase and its genetic control sites. *Mol Gen Genet* **200:**220-228.

Stelzl, U., J. M. Zengel, M. Tovbina, M. Walker, K. H. Nierhaus, L. Lindahl and D. J. Patel. 2003. RNA-structural mimicry in *Escherichia coli* ribosomal protein L4-dependent regulation of the *S10* operon. *J Biol Chem* **278:**28237-28245.

Stock, A. M., V. L. Robinson and P. N. Goudreau. 2000. Two-component signal transduction. *Annu Rev Biochem* **69:**183-215.

Strauss, J. H. and E. G. Strauss. 1988. Evolution of RNA viruses. *Annu Rev Microbiol* **42:**657-683.

Stuart, K. and A. K. Panigrahi. 2002. RNA editing: complexity and complications. *Mol Microbiol* **45:**591-596.

Stucki, J. W. 1980. The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation. *Eur J Biochem* **109:**269-283.

Stülke, J. 2002. Control of transcription termination in bacteria by RNA-binding proteins that modulate RNA structures. *Arch Microbiol* **177:**433-440.

Stülke, J., M. Arnaud, G. Rapoport and I. Martin-Verstraete. 1998. PRD - a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol Microbiol* **28:**865-874.

Stülke, J. and W. Hillen. 2000. Regulation of carbon catabolism in *Bacillus* species. *Annu Rev Microbiol* **54:**849-880.

Stülke, J., I. Martin-Verstraete, V. Charrier, A. Klier, J. Deutscher and G. Rapoport. 1995. The HPr protein of the phosphotransferase system links induction and catabolite repression of the *Bacillus subtilis* levanase operon. *J Bacteriol* **177:**6928-6936.

Stülke, J., I. Martin-Verstraete, M. Zagorec, M. Rose, A. Klier and G. Rapoport. 1997. Induction of the *Bacillus subtilis ptsGHI* operon by glucose is controlled by a novel antiterminator, GlcT. *Mol Microbiol* **25:**65-78.

Stülke, J. and M. H. Schmalisch. 2004. The bacterial phosphotransferase system: a perfect link of sugar transport and signal transduction. *Topics Curr Genet* **9:**179-205.

Sudarsan, N., J. E. Barrick and R. R. Breaker. 2003a. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9:**644-647.

Sudarsan, N., J. K. Wickiser, S. Nakamura, M. S. Ebert and R. R. Breaker. 2003b. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev* **17:**2688-2697.

Suess, B., B. Fink, C. Berens, R. Stentz and W. Hillen. 2004. A theophylline responsive riboswitch based on helix slipping controls gene expression *in vivo*. *Nucleic Acids Res* **32:**1610-1614.

Suess, B., S. Hanson, C. Berens, B. Fink, R. Schroeder and W. Hillen. 2003. Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucleic Acids Res* **31:**1853-1858.

Switzer, R. L., R. J. Turner and Y. Lu. 1999. Regulation of the *Bacillus subtilis* pyrimidine biosynthetic operon by transcriptional attenuation: control of gene expression by an mRNA-binding protein. *Prog Nucleic Acid Res Mol Biol* **62:**329-367.

Sykes, M. T. and M. Levitt. 2005. Describing RNA structure by libraries of clustered nucleotide doublets. *J Mol Biol* **351:**26-38.

Tanaka, K., K. Kobayashi and N. Ogasawara. 2003. The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YflS, and is essential for utilization of malate in minimal medium. *Microbiology* **149:**2317-2329.

Tanaka, Y., F. Itoh, K. Kimata and H. Aiba. 2004. Membrane localization itself but not binding to IICB is directly responsible for the inactivation of the global repressor Mlc in *Escherichia coli*. *Mol Microbiol* **53:**941-951.

Tanaka, Y., K. Kimata and H. Aiba. 2000. A novel regulatory role of glucose transporter of *Escherichia coli*: membrane sequestration of a global repressor Mlc. *Embo J* **19:**5344-5352.

Tanaka, Y., K. Kimata, T. Inada, H. Tagami and H. Aiba. 1999. Negative regulation of the *pts* operon by Mlc: mechanism underlying glucose induction in *Escherichia coli*. *Genes Cells* **4:**391-399.

Thore, S., M. Leibundgut and N. Ban. 2006. Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **312:**1208-1211.

Tobisch, S., J. Stülke and M. Hecker. 1999a. Regulation of the *lic* operon of *Bacillus subtilis* and characterization of potential phosphorylation sites of the LicR regulator protein by site-directed mutagenesis. *J Bacteriol* **181:**4995-5003.

Tobisch, S., D. Zühlke, J. Bernhardt, J. Stülke and M. Hecker. 1999b. Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. *J Bacteriol* **181:**6996-7004.

Tojo, S., T. Satomura, K. Morisaki, J. Deutscher, K. Hirooka and Y. Fujita. 2005. Elaborate transcription regulation of the *Bacillus subtilis ilv-leu* operon involved in the biosynthesis of branched-chain amino acids through global regulators of CcpA, CodY and TnrA. *Mol Microbiol* **56:**1560-1573.

Tomchick, D. R., R. J. Turner, R. L. Switzer and J. L. Smith. 1998. Adaptation of an enzyme to regulatory function: structure of *Bacillus subtilis* PyrR, a *pyr* RNA-binding attenuation protein and uracil phosphoribosyltransferase. *Structure* **6:**337-350.

Tomizawa, J., T. Itoh, G. Selzer and T. Som. 1981. Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc Natl Acad Sci USA* **78:**1421-1425.

Tortosa, P., S. Aymerich, C. Lindner, M. H. Saier, Jr., J. Reizer and D. Le Coq. 1997. Multiple phosphorylation of SacY, a *Bacillus subtilis* transcriptional antiterminator negatively controlled by the phosphotransferase system. *J Biol Chem* **272:**17230-17237.

Tortosa, P., N. Declerck, H. Dutartre, C. Lindner, J. Deutscher and D. Le Coq. 2001. Sites of positive and negative regulation in the *Bacillus subtilis* antiterminators LicT and SacY. *Mol Microbiol* **41:**1381-1393.

Tozzi, M. G., M. Camici, L. Mascia, F. Sgarrella and P. L. Ipata. 2006. Pentose phosphates in nucleoside interconversion and catabolism. *Febs J* **273:**1089-1101.

Trach, K., D. Burbulys, M. Strauch, J. J. Wu, N. Dhillon, R. Jonas, C. Hanstein, P. Kallio, M. Perego, T. Bird and et al. 1991. Control of the initiation of sporulation in *Bacillus subtilis* by a phosphorelay. *Res Microbiol* **142:**815-823.

Trieu-Cuot, P. and P. Courvalin. 1983. Nucleotide sequence of the *Streptococcus faecalis* plasmid gene encoding the 3'5"-aminoglycoside phosphotransferase type III. *Gene* **23:**331-341.

Turinsky, A. J., F. J. Grundy, J. H. Kim, G. H. Chambliss and T. M. Henkin. 1998. Transcriptional activation of the *Bacillus subtilis ackA* gene requires sequences upstream of the promoter. *J Bacteriol* **180:**5961-5967.

Turinsky, A. J., T. R. Moir-Blais, F. J. Grundy and T. M. Henkin. 2000. *Bacillus subtilis ccpA* gene mutants specifically defective in activation of acetoin biosynthesis. *J Bacteriol* **182:**5611-5614.

Valadkhan, S. 2005. snRNAs as the catalysts of pre-mRNA splicing. *Curr Opin Chem Biol* **9:**603-608.

Valbuzzi, A. and C. Yanofsky. 2001. Inhibition of the *B. subtilis* regulatory protein TRAP by the TRAP-inhibitory protein, AT. *Science* **293:**2057-2059.

Valegård, K., J. B. Murray, N. J. Stonehouse, S. van den Worm, P. G. Stockley and L. Liljas. 1997. The three-dimensional structures of two complexes between recombinant MS2 capsids and RNA operator fragments reveal sequence-specific protein-RNA interactions. *J Mol Biol* **270:**724-738.

Valentin-Hansen, P., M. Eriksen and C. Udesen. 2004. The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Mol Microbiol* **51:**1525-1533.

van Dijken, J. P., R. A. Weusthuis and J. T. Pronk. 1993. Kinetics of growth and sugar consumption in yeasts. *Antonie Van Leeuwenhoek* **63:**343-352.

van Tilbeurgh, H., D. Le Coq and N. Declerck. 2001. Crystal structure of an activated form of the PTS regulation domain from the LicT transcriptional antiterminator. *EMBO J* **20:**3789-3799.

van Tilbeurgh, H., X. Manival, S. Aymerich, J. M. Lhoste, C. Dumas and M. Kochoyan. 1997. Crystal structure of a new RNA-binding domain from the antiterminator protein SacY of *Bacillus subtilis*. *EMBO J* **16:**5030-5036.

Vanderpool, C. K. and S. Gottesman. 2004. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54:**1076-1089.

Viana, R., G. Pérez-Martinez, J. Deutscher and V. Monedero. 2005. The glycolytic genes *pfk* and *pyk* from *Lactobacillus casei* are induced by sugars transported by the phosphoenolpyruvate:sugar phosphotransferase system and repressed by CcpA. *Arch Microbiol* **183:**385-393.

Vitreschak, A. G., D. A. Rodionov, A. A. Mironov and M. S. Gelfand. 2002. Regulation of riboflavin biosynthesis and transport genes in bacteria by transcriptional and translational attenuation. *Nucleic Acids Res* **30:**3141-3151.

Vitreschak, A. G., D. A. Rodionov, A. A. Mironov and M. S. Gelfand. 2004. Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet* **20:**44-50.

Wach, A. 1996. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. *Yeast* **12:**259-265.

Wacker, I., H. Ludwig, I. Reif, H.-M. Blencke, C. Detsch and J. Stülke. 2003. The regulatory link between carbon and nitrogen metabolism in *Bacillus subtilis*: regulation of the *gltAB* operon by the catabolite control protein CcpA. *Microbiology* **149:**3001-3009.

Wagner, E. G. and K. Flärdh. 2002. Antisense RNAs everywhere? *Trends Genet* **18:**223-226.

Waites, K. B. and D. F. Talkington. 2004. *Mycoplasma pneumoniae* and its role as a human pathogen. *Clin Microbiol Rev* **17:**697-728.

Waldminghaus, T., A. Fippinger, J. Alfsmann and F. Narberhaus. 2005. RNA thermometers are common in α- and γ-proteobacteria. *Biol Chem* **386:**1279-1286.

Wasinger, V. C., J. D. Pollack and I. Humphery-Smith. 2000. The proteome of *Mycoplasma genitalium*. Chaps-soluble component. *Eur J Biochem* **267:**1571-1582.

Watson, J. D. and F. H. C. Crick. 1953. Molecular structure of nucleic acids. *Nature* **4356:**737-742.

Weickert, M. J. and S. Adhya. 1992. A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem* **267:**15869-15874.

Weinrauch, Y., T. Msadek, F. Kunst and D. Dubnau. 1991. Sequence and properties of *comQ*, a new competence regulatory gene of *Bacillus subtilis*. *J Bacteriol* **173:**5685-5693.

Werstuck, G. and M. R. Green. 1998. Controlling gene expression in living cells through small molecule-RNA interactions. *Science* **282:**296-298.

Wiechert, W. 2001. ^{13}C metabolic flux analysis. *Metab Eng* **3:**195-206.

Wilson, S. A. and R. E. Drew. 1995. Transcriptional analysis of the amidase operon from *Pseudomonas aeruginosa*. *J Bacteriol* **177:**3052-3057.

Wilson, S. A., S. J. Wachira, R. E. Drew, D. Jones and L. H. Pearl. 1993. Antitermination of amidase expression in *Pseudomonas aeruginosa* is controlled by a novel cytoplasmic amide-binding protein. *EMBO J* **12:**3637-3642.

Wilson, S. A., S. J. Wachira, R. A. Norman, L. H. Pearl and R. E. Drew. 1996. Transcription antitermination regulation of the *Pseudomonas aeruginosa* amidase operon. *EMBO J* **15:**5907-5916.

Winkler, W. C. 2005. Metabolic monitoring by bacterial mRNAs. *Arch Microbiol* **183:**151-159.

Winkler, W. C. and R. R. Breaker. 2003. Genetic control by metabolite-binding riboswitches. *ChemBioChem* **4:**1024-1032.

Winkler, W. C. and R. R. Breaker. 2005. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **59:**487-517.

Winkler, W. C., S. Cohen-Chalamish and R. R. Breaker. 2002a. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci USA* **99:**15908-15913.

Winkler, W. C., A. Nahvi and R. R. Breaker. 2002b. Thiamine derivates bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419:**890-891.

Winkler, W. C., A. Nahvi, A. Roth, J. A. Collins and R. R. Breaker. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428:**281-286.

Winkler, W. C., A. Nahvi, N. Sudarsan, J. E. Barrick and R. R. Breaker. 2003. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10:**701-707.

Wise, J. A. 1993. Guides to the heart of the spliceosome. *Science* **262:**1978-1979.

Wittmann, C., M. Hans and E. Heinzle. 2002. *In vivo* analysis of intracellular amino acid labelings by GC/MS. *Anal Biochem* **307:**379-382.

Wittmann, C. and E. Heinzle. 2001. Application of MALDI-TOF MS to lysine-producing *Corynebacterium glutamicum*: a novel approach for metabolic flux analysis. *Eur J Biochem* **268:**2441-2455.

Wittmann, C. and E. Heinzle. 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.

Wittmann, C., P. Kiefer and O. Zelder. 2004. Metabolic fluxes in *Corynebacterium glutamicum* during lysine production with sucrose as carbon source. *Appl Environ Microbiol* **70:**7277-7287.

Wray, L. V., Jr. and S. H. Fisher. 1994. Analysis of *Bacillus subtilis hut* operon expression indicates that histidine-dependent induction is mediated primarily by transcriptional antitermination and that amino acid repression is mediated by two mechanisms: regulation of transcription initiation and inhibition of histidine transport. *J Bacteriol* **176:**5466-5473.

Xu, P., M. Guo and B. A. Hay. 2004. MicroRNAs and the regulation of cell death. *Trends Genet* **20:**617-624.

Yakhnin, H., A. V. Yakhnin and P. Babitzke. 2006. The *trp* RNA-binding attenuation protein (TRAP) of *Bacillus subtilis* regulates translation initiation of *ycbK*, a gene encoding a putative efflux protein, by blocking ribosome binding. *Mol Microbiol* **61:**1252-1266.

Yakhnin, H., H. Zhang, A. V. Yakhnin and P. Babitzke. 2004. The trp RNA-binding attenuation protein of *Bacillus subtilis* regulates translation of the tryptophan transport gene *trpP* (*yhaG*) by blocking ribosome binding. *J Bacteriol* **186:**278-286.

Yamamoto, H., M. Murata and J. Sekiguchi. 2000. The CitST two-component system regulates the expression of the Mg-citrate transporter in *Bacillus subtilis*. *Mol Microbiol* **37:**898-912.

Yamanaka, K., M. Mitta and M. Inouye. 1999. Mutation analysis of the 5' untranslated region of the cold shock *cspA* mRNA of *Escherichia coli*. *J Bacteriol* **181:**6284-6291.

Yang, H., F. Jossinet, N. Leontis, L. Chen, J. Westbrook, H. Berman and E. Westhof. 2003. Tools for the automatic identification and classification of RNA base pairs. *Nucleic Acids Res* **31:**3450-3460.

Yang, M., X. Chen, K. Militello, R. Hoffman, B. Fernandez, C. Baumann and P. Gollnick. 1997. Alanine-scanning mutagenesis of *Bacillus subtilis trp* RNA-binding attenuation protein (TRAP) reveals residues involved in tryptophan binding and RNA binding. *J Mol Biol* **270:**696-710.

Yang, Y., N. Declerck, X. Manival, S. Aymerich and M. Kochoyan. 2002. Solution structure of the LicT-RNA antitermination complex: CAT clamping RAT. *EMBO J* **21:**1987-1997.

Yang, Y. T., K. Y. San and G. N. Bennett. 1999. Redistribution of metabolic fluxes in *Escherichia coli* with fermentative lactate dehydrogenase overexpression and deletion. *Metab Eng* **1:**141-152.

Ying, S. Y., D. C. Chang, J. D. Miller and S. L. Lin. 2006. The microRNA: overview of the RNA gene that modulates gene functions. *Methods Mol Biol* **342:**1-18.

Yoshida, K., K. Kobayashi, Y. Miwa, C. M. Kang, M. Matsunaga, H. Yamaguchi, S. Tojo, M. Yamamoto, R. Nishi, N. Ogasawara, T. Nakayama and Y. Fujita. 2001. Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in *Bacillus subtilis*. *Nucleic Acids Res* **29:**683-692.

Yoshida, M. 1972. Allosteric nature of thermostable phosphofructokinase from an extreme thermophilic bacterium. *Biochemistry* **11:**1087-1093.

Youngman, P. 1990. Use of transposons and integrational vectors for mutagenesis and construction of gene fusions in *Bacillus subtilis*. *In*: *Molecular Biological Methods for Bacillus.* C. R. Harwood and S. M. Cutting. (eds.), John Wiley & Sons, Chichester, United Kingdom, pp. 221-266.

Yousef, M. R., F. J. Grundy and T. M. Henkin. 2003. tRNA requirements for *glyQS* antitermination: a new twist on tRNA. *RNA* **9:**1148-1156.

Zamboni, N., E. Fischer, D. Laudert, S. Aymerich, H. P. Hohmann and U. Sauer. 2004. The *Bacillus subtilis yqjI* gene encodes the $NADP^+$-dependent 6-P-gluconate dehydrogenase in the pentose phosphate pathway. *J Bacteriol* **186:**4528-4534.

Zapp, M. L. and M. R. Green. 1989. Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature* **342:**714-716.

Zengel, J. M. and L. Lindahl. 1994. Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog Nucleic Acid Res Mol Biol* **47:**331-370.

Zhang, A., K. M. Wassarman, C. Rosenow, B. C. Tjaden, G. Storz and S. Gottesman. 2003. Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50:**1111-1124.

Zhang, H. and R. L. Switzer. 2003. Transcriptional pausing in the *Bacillus subtilis pyr* operon *in vitro*: a role in transcriptional attenuation? *J Bacteriol* **185:**4764-4771.

Zhang, Y. and J. Skolnick. 2004. Automated structure prediction of weakly homologous proteins on a genomic scale. *Proc Natl Acad Sci USA* **101:**7594-7599.

Zorrilla, S., T. Doan, C. Alfonso, E. Margeat, A. Ortega, G. Rivas, S. Aymerich, C. Royer and N. Declerck. 2007. Inducer-Modulated Cooperative Binding of the Tetrameric CggR Repressor to Operator DNA. *Biophys J* **92:**3215-3227.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31:**3406-3415.

8 Supplementary material

Oligonucleotides

Table 13. Oligonucleotides

Name	Sequence 5' → 3' [#]	Description
cat-fwd	CGGCAATAGTTACCCTTATTATCAAG	LFH-PCR[*]; amplification of *cat* cassette
cat-rev	CCAGCGTGGACCGGCGAGGCTAGTTACCC	LFH-PCR[*]; amplification of *cat* cassette
CH7	AAAAGTCGACATGGATCTAAAACAACAATACATTCTTG	cloning of *M. pneumoniae glpK* gene fw (*Sal*I)
CH8	TATAAAGCTTGTCTTAGTCTAAGCTAGCCCATTTTAG	cloning of *M. pneumoniae glpK* gene rev (*Hind*III), A1512→G
CH9	AAAAGTCGACATGGATCTAAAACAAC	cloning of *M. pneumoniae glpK* gene fw (*Sal*I)
CH10	TATAAAGCTTGTCTTAGTCTAAGCTAG	cloning of *M. pneumoniae* glpK gene rev (*Hind*III)
CH11	P-GATCCCTTAGAAATTTGGTCAGTCCAATTAG	MMR: *M. pneumoniae glpK* A165→G
CH12	P-CCATTGTGTTATGGAACAAAGAAAATGGTTTG	MMR: *M. pneumoniae glpK* A273→G
CH13	P-CACTAAGATTGCTTGGATCTTGGAAAATGTTC	MMR: *M. pneumoniae glpK* A438→G
CH14	P-CCTGGTTAATTTGGAAACTAACGGGTG	MMR: *M. pneumoniae glpK* A522→G
CH15	P-CCATGACATGGTCACAAGAGTTAGGC	MMR: *M. pneumoniae glpK* A606→G
CH16	P-TACCGAGTCATTGGTCTACTAGTGC	MMR: *M. pneumoniae glpK* A705→G
CH17	P-CCTTAAAGTGGTTAAGGGATAGTCTTAAGG	MMR: *M. pneumoniae glpK* A966→G
CH18	P-GCAGTTAATTATTGGAAGGACACTAAACAAC	MMR: *M. pneumoniae glpK* A1386→G
CH19	P-GAAATCAAAGCGTTGGAACGAAGCTG	MMR: *M. pneumoniae glpK* A1482→G
IL5	AAACAATTGAGGCATATCAGATTTACAATAGAAAG	amplification of *ptsG* upstream region; (*Mfe*I)
IL59	*CCAAGTAATACGACTCACTATAGG*ACGTGTTACTGATTCG	*ptsG*-PCR-fragment for *in vitro*-transcription; includes T7-promoter
IL60	CAAGAATTGGGACAACTCTTCTTCTCCTTTTTTTTCCTCAATCACTCATGCC	*ptsG*-PCR-fragment for *in vitro*-transcription

Name	Sequence 5' → 3' [#]	Description
JS11	TTTGGATCCGCAACTGGAAGCATAAGCGC	amplification of *ptsG* upstream region; (*BamH*I)
JS39	TCTATCAACAGGAGTCCA	sequencing of pWH844 constructs
spec-fwd	GACTGGCTCGCTAATAACGTAACGTGACT GGCAAGAG	LFH-PCR[*]; amplification of *spc* cassette
spec-rev	CGTAGCGAGGGCAAGGGTTTATTGTTTTCT AAAATCTG	LFH-PCR[*]; amplification of *spc* cassette
OS25	*CCAAGTAATACGACTCACTATAGG*AATTCAG TTTATCCTTAT	*ptsG*-PCR-fragment for *in vitro*-transcription of mutant RATs; includes T7-promoter
OS26	TTGAGGGAAAAAA ACGGGAAGTTC	*ptsG*-PCR-fragment for *in vitro*-transcription of mutant RATs
OS28	P-TTATCCTTATAACG**A**TGTTACTGATTCG	mutagenesis of *ptsG*-RAT: A4ins
OS29	P-TCGATCAGGCAGAGTGATTGAGGGA	mutagenesis of *ptsG*-RAT: U25del
OS30	P-GTTTATCCTTATA**C**CG**T**TGTTACTG	mutagenesis of *ptsG*-RAT: A1→C; U4ins
OS31	P-CAGGCA**A**GAG**G**GATTGAGG	mutagenesis of *ptsG*-RAT: U25→A; U29→G
OS32	P-GCACCCAATTTT**T**CTC**T**TGCCTTTTTG	mutagenesis of *ptsG*-RAT: terminator region, G84→U; A88→U
OS33	AAGGATCCCGTTTTATGAATGGCGGATG	*sacT* knockout; (*BamH*I)
OS34	TTTAAGCTTCATCTCGCTTTTGATCGAGC	*sacT* knockout; (*Hind*III)
OS35	AAGGATCCGGCATTCCGATGATTTCAATC G	*sacY* knockout; (*BamH*I)
OS36	TTTAAGCTTGGAGAAAAGCCTTGGAACTC	*sacY* knockout; (*Hind*III)
OS37	CGTAATGCCTGTGGATAACG	*sacT*-forw1 (LFH-PCR[*] for knockout)
OS38	CCACGCTTACTACGTTGATAAGCCTATATG CTCCTCCGGCAGTGTTTGC	*sacT*-rev1 (LFH-PCR[*] for knockout); complementary to *tet*-cassette
OS39	GGGATCAACTTTGGGAGAGAGTTCGTTGG AATCAAATGAAGCGCTCCATC	*sacT*-forw2 (LFH-PCR [*] for knockout); complementary to *tet* cassette
OS40	CTTGCCTGCTTTACCAGTC	*sacT*-rev2 (LFH-PCR[*] for knockout)
OS41	CCTATCACCTCAAATGGTTCGCTGCTTTAT TCATAGGGATGAGATC	*sacT*-rev1 (LFH-PCR[*] for knockout); complementary to *kan* cassette
OS42	CGAGCGCCTACGAGGAATTTGTATCGCAT TCGCCTATCAATGTGCG	*sacT*-forw2 (LFH-PCR[*] for knockout); complementary to *aphA3* cassette
OS43	GGCAATAACAGCAAGCGG	*sacY*-forw1 (LFH-PCR [*] for knockout)
OS44	CTTGATAATAAGGGTAACTATTGCCGATC CTTTACGACGATAGC	*sacY*-rev1 (LFH-PCR [*] for knockout); complementary to *cat*

Name	Sequence 5' → 3' [#]	Description
OS45	GGGTAACTAGCCTCGCCGGTCCACGGCTA AGCATCGGCACCTTTGTG	*sacY*-forw2 (LFH-PCR[*] for knockout); complementary to *cat*-cassette
OS46	CGGTTCCTCTCTCTTAGATG	*sacY*-rev2 (LFH-PCR[*] for knockout)
OS47	P-GCGCGGG**G**TGTTACTGATAAAG	mutagenesis of *sacB*-RAT: U4-deletion+U3→G
OS48	P-GCAGGCA**T**GACCTAAAATGTG	mutagenesis of *sacB*-RAT: A24→U24 (U4del)
OS49	AAAGAATTCGATCCTTTTTAACCCATCACA TATAC	amplification of *sacB* upstream region; (*EcoR*)
OS50	TTTGGATCCTTTTTGATGTTCATCGTTCAT GTC	amplification of *sacB* upstream region; (*BamH*I)
OS52	P-GCACCCAATTTT**C**CTC**T**TGCCTTTTTG	mutagenesis of *ptsG*-terminator: G84→C; A88→U
OS53	P-GTTTATCCTTATA**C**C**TT**TGTTACTG	mutagenesis of *ptsG*-RAT: A1→C; G3→U; U4ins
OS54	P-CATATTTTAGGTC**A**TTTTTTATTGTGCG	mutagenesis of *sacB*-RAT, terminator region U85→A
OS55	P-GCGCGGGTTGTTACTGATAAAG	mutagenesis of *sacB*-RAT: U4-deletion
OS56	AAAGGATCCCACAATGGCTGGTGACGATC	amplification of *pdhB* (*BamH*I)
OS57	AAAGGGCCCGGTGAAGGAATTTCTACTAC TGC	amplification of *pdhB* (*Apa*I)
OS58	TTTAAGCTTCTAGGGTAAGTAAATTGAGT A	amplification of *spc*-cassette (*Hind*III)
OS59	TTTCAATTGATTGAGAGAAGTTTCTATAGA	amplification of *spec*-cassette (*Mfe*I)
OS60	P-TATCCTTATAAC**TT**TGTTACTGATTCG	mutagenesis of *ptsG*-RAT (plasmid pGP557 as template): RAT G3→U
OS61	P-CGCGCGGG**A**TTGTTACTG	mutagenesis of *sacB*-RAT: U3→A
OS62	P-CGCGCGGG**G**TTGTTACTG	mutagenesis of *sacB*-RAT: U3→G
OS63	GCTGATCAACTCCTGCGG	knockout of *ydaP* with LFH-PCR[*]; complementary to *ydaO*
OS64	CCTATCACCTCAAATGGTTCGCTGCATTGC TCAAGCAGCTCTGTC	knockout of *ydaP* with LFH-PCR[*]; complementary to *ydaP*
OS65	CGAGCGCCTACGAGGAATTTGTATCGGAA ACGAACCGCCGCTTC	knockout of *ydaP* with LFH-PCR[*]; complementary to *ydaP*
OS66	GATTGCAGGGGCTATTAATGC	knockout of *ydaP* with LFH-PCR[*]; complementary to *mntH*
OS67	GGTTTAACAACGACGCCTGC	knockout of *csrA* with LFH-PCR[*]; complementary to *yviE*
OS68	CCTATCACCTCAAATGGTTCGCTGGCACCT ATTTGAATCGCTTCG	knockout of *csrA* with LFH-PCR[*]; complementary to *csrA*

Name	Sequence 5' → 3' [#]	Description
OS69	CGAGCGCCTACGAGGAATTTGTATCGCAG GAAGAAAATAACCGTGCAG	knockout of *csrA* with LFH-PCR[*]; complementary to *csrA*
OS70	CTGTTGGTTTGCTTGAGCAAG	knockout of *csrA* with LFH-PCR[*]; complementary to *hagP*
OS71	P-GCAGGCAA**A**ACCTAAAATGTG	mutagenesis of *sacB*-RAT: G26→A
OS72	P-CATATTTTAGGT**T**TTTTTTTATTGTGCG	mutagenesis of *sacB*-terminator: C85→U
OS73	P-TTGTTACTG**CG**AAAGCAGGCAA	mutagenesis of *sacB*-RAT: A13U14→CG
OS74	AAAGGATCCGGAGGTGATAAGCCATGGCG CAAATGACAATG	amplification of *pdhB* (*BamH*I)
OS75	AAAGGATCCGCAGTTT**C**ATTAAAATTCAA GCAC	amplification of *pdhB* (*BamH*I)
OS76	TTGTTACTG**CG**AAAGCAGGCAA**A**ACC	mutagenesis of *sacB*-RAT: G26→A; A13U14→CG
OS77	TTGTTACTG**CG**AAAGCAGGCAAGACC	mutagenesis of *sacB*-RAT: A13U14→CG
OS78	GATCTCTGGTTCCGCGTGGTTCCATGGGAT CCGTCGACCTGCA	introducution of thrombin restriction site and MCS in pWH844 (*BamH*I / *Pst*I); complementary to OS79
OS79	GGTCGACGGATCCCATGGAACCACGCGGA ACCAGA	see OS78
OS80	AAAGTCGACAAAATCTATAAAGTATTAAA CAACAATGCG	cloning of SacT-RBD (forw); (*Sal*I)
OS81	TTTAAGCTTCTAGAACTTTTCATTCTCGTC GCGCAC	cloning of SacT-RBD (rev); (*Hind*III)
OS82	AAAGTCGACAAAATTGCGAAGGTGATCAA CAATAATGTG	cloning of LicT-RBD (forw); (*Sal*I)
OS83	TTTAAGCTTCTATTCTGATACATCCTTGTT ATCGAGC	cloning of LicT-RBD (rev); (*Hind*III)
OS84	AAAGTCGACAAAATTAAAAGAATCTTAAA TCATAATGCTATCG	cloning of SacY-RBD (forw); (*Sal*I)
OS85	TTTAAGCTTCTACTTATAGTCAGGTGTATC TTTTCTG	cloning of SacY-RBD (rev); (*Hind*III)
OS86	*CCAAGTAATACGACTCACTATAGG*CGAAAAG TAAATCGCGCG	*sacB*-PCR-fragment for *in vitro*-transcription; includes T7-promoter
OS87	GTATACACTTTGCCCTTTACAC	*sacB*-PCR-fragment for *in vitro*-transcription
OS88	P-GGGTTTGT**G**ACTGATAAAGCAGGC	mutagenesis of *sacB*-RAT: U8→G
OS89	P-GGG**A**TTGT**G**ACTGATAAAGCAGGC	mutagenesis of *sacB*-RAT: U8→G; U3→A

Name	Sequence 5' → 3' [#]	Description
OS90	P-CAGTTTATCCTTATAACGTTGTTACTG**AT AAAG**CAGGCAAGAGTGATTGAGGG	mutagenesis of *ptsG*-RAT mutates top loop similar to *sacB*-RAT
OS91	TATGGAGCTCGGATCCTGGAGCCACCCGC AGTTCGAAAAATGATAGT	introduction of Strep-tag into pET3C, complementary to OS92
OS92	GATCACTATCATTTTTCGAACTGCGGGTGG CTCCAGGATCCGAGCTCCA	see OS91
OS93	AAACATATGGTGAATGGGTCCTTCACAGT G	cloning of GlcT-RBD into pGP574; (*NdeI*)
OS94	AAAGGATCCTTGTTCCTTCTCGTCTTTTAA AATGAAC	cloning of GlcT-RBD into pGP574; (*BamH*I)
OS95	AAACATATGAAAATTGCGAAGGTGATCAA CAATAATGTG	cloning of LicT-RBD into pGP574; (*NdeI*)
OS96	AAAGGATCCTGATACATCCTTGTTATCGA GCG	cloning of LicT-RBD into pGP574; (*BamH*I)
OS97	AAACATATGAAAATCTATAAAGTATTAAA CAACAATGCG	cloning of SacT-RBD into pGP574; (*NdeI*)
OS98	AAAGGATCCGAACTTTTCATTCTCGTCGCG CAC	cloning of SacT-RBD into pGP574; (*BamH*I)
OS99	AAACATATGAAAATTAAAAGAATCTTAAA TCATAATGCTATCG	cloning of SacY-RBD into pGP574; (*NdeI*)
OS100	AAAGGATCCCTTATAGTCAGGTGTATCTTT TCTG	cloning of SacY-RBD into pGP574; (*BamH*I)
OS101	NNNNNNNNN	nonamere random primer for labeling
OS102	P-GG**A**TTGT**G**ACTG**G**TAAAGCAGGCAAG	mutagenesis of *sacB*-RAT: A13→G; U8→G; U3→A
OS103	P-CAGTTTATCCTTATAAC**AT**TGT**G**ACTGA T**AA**A**G**CAGGC	mutagenesis of *ptsG*-RAT; mutates top loop similar to *sacB*-RAT +G3→A; U8→G
OS104	P-CAGTTTATCCTTATAAC**TT**TGTTACTGAT AAAGCAGG	mutagenesis of *ptsG*-RAT; mutates top loop similar to *sacB*-RAT+G3U
OS105	P-GGTTTGTTACTG**G**TAAAGCAGGCAAG	mutagenesis of *sacB*-RAT: A13→G
OS106	AAAGGATCCTTATTCCGAGAGATCTTGAA GAC	cloning of SacT with C-terminal Strep-tag
OS107	AAAGGATCCTCAGCGTGCGACTGACC	cloning of SacY with C-terminal Strep-tag
OS108	AAAGGATCCTTCCGAGAGATCTTGAAGAC G	cloning of SacT-full with C-terminal Strep-tag;
OS109	AAAGGATCCGCGTGCGACTGACCG	cloning of SacY-full with C-terminal Strep-tag.
OS110	TATCCTTATAACG**TT**GTTACTGATTCG	mutagenesis of *ptsG*-RAT: $U4_{ins}$,
OS111	TATCCTTATAAC**AT**TGTTACTGATTCG	mutagenesis of *ptsG*-RAT: G3→A, $U4_{ins}$

Name	Sequence 5' → 3' [#]	Description
OS112	CAGTTTATCCTTATAAC**AA**TGTTACTGATT CG	mutagenesis of *ptsG*-RAT: $A4_{ins}$, G3→A
Name	Sequence 5' → 3' [#]	Description
OS113	CCAATTTTGCTCTGCCTTTTTGTTG	mutagenesis of *ptsG*-RAT: terminator region $A88_{del}$
OS114	CGATCAGGCAGAGTGATTGAGG	mutagenesis of *ptsG*-RAT: $U25_{del}$
OS115	CCAATTTTGCTCATG**T**CCTTTTTGTTG	mutagenesis of *ptsG*-RAT: terminator region $T91_{ins}$
OS116	CTGATTCGATCAGG**A**CATGAGTGATTG	mutagenesis of *ptsG*-RAT: $A23_{ins}$
OS117	TGTTACTGATTCGA**G**CAGGCATGAGT	mutagenesis of *ptsG*-RAT: U18→G
OS118	CAGGCA**A**GAGTGATTGAGGGAA	mutagenesis of *ptsG*-RAT: U25→A
OS119	CCAATTTTGCTC**T**TGCCTTTTTGTTG	mutagenesis of *ptsG*-RAT: terminator region A88→U
SHU55	GTTTTACAACGTCGTGACTGG	template for RNA probe *(E. coli lacZ)*
SHU56	*CTAATACGACTCACTATAGGGAGG*TGTGCAG TTCAACCACCG	template for RNA probe *(E. coli lacZ)*; includes T7-promoter

[#] restriction sites are underlined, introduced mutations are bold, promoters are italic, P indicates 5'-phosphorylation

[*] LFH-PCR: long flanking homology PCR

Table 14. Oligonucleotides used for microarray analysis

Name	forward primer sequence 5' → 3' [#]	reverse primer sequence 5' → 3' [#]
ptsG	GGTTCGCGGAAAAATTCTCAATG	CTTCTTGTTCTCTGTTGACTGAAC
ptsH	AAAGTCGACTGCATGGCACAAAAA ACATTTAAAGTAACT (MS22)	AAAGTCGACATGGCACAAAAAACA TTTAAAGTAACT (MS25)
ptsI	TCGCGCGCTGCTTCGTG (AM4)	GTGTAATTAAGCGAAGGATTGCC
pgi	AAGGATCCATGACGCATGTACGCTT GACTA	TTTCTGCAGTTAATCTTCCAGACGT TTTTCAAGC
pgi-2	CTTGACGGCTTGAACTATTTAGC	
pfkA	AAGGATCCATGAAACGAATAGGGG TATTAACG	TTTCTGCAGTTAGATAGACAGTTCT TTTGAAAGCTG
pfkA-2	CACAGTATTATTATTGTTGCCGAAG G	
fbaA	AAGGATCCATGCCTTTAGTTTCTAT GACG	TTTCTGCAGTTAAGCTTGGTTTGAA GAACCAAATTC
fbaA-2	CTTACAAAGGTGAGCCAAACCTTG	
tpiA	AAGGATCCATGAGAAAACCAATTA TCGCCG	TTTCTGCAGTTACTCATATTGACCTT CCTCCA

Name	forward primer sequence 5' → 3' #	reverse primer sequence 5' → 3' #
tpiA-2	CTTGCTGGTCTTTCTGAAGAAC	
gapA	CAATGCGTGTTCCAACTCCAAAC	TTTAAGCTTTTAAAGACCTTTTTTTGCGATGTAAGC
gapA-2	AAAAGTCGACATGGCAGTAAAAGTCGGTATTAACG	
pgk	AAAAGTCGACAATAAAAAAACTCTCAAAGACATCGACG	TTTAAGCTTTTATTTATCGTTCAGTGCAGCTAC
pgk-2	CAATCGACATTGGTACGAAAACAC	
pgm	AAGGATCCAGTAAAAAACCAGCTGCACTCAT	TTTCTGCAGTTATTTTTGAATTAAAGATGTTCCTGTCATTTC
pgm-2	GGAATGGTTGAACCAACAATTAAAGC	
eno	AAGGATCCCCATACATTGTTGATGTTTATGCAC	TTTCTGCAGTTACTTGTTTAAGTTGTAGAAAGAGTTG
eno-2	GGTATTAAAAACGGCGTAGGCAAC	
pyk	AAGGATCCAGAAAAACTAAAATTGTTTGTACCATCGG	TTTCTGCAGTTAAAGAACGCTCGCACGG
pyk-2	CATTGGACGCAAATCAGCTTAC	
pdhA	CTAATTGAAACACTTACATTCCGTTATG	TTACTTCGACTCCTTCTGTGTATA
pdhB	CATTAAAAGCTGCTGATGAGCTTG	TTAAAATTCAAGCACTTTTCTTGCTG
pdhC	GAAATCAATGGCCTTGCAACAAAAG	TTACGCCTCCATTAAAATTAATTGTGG
pdhD	GAACTTGCATCAGTTGGTTACAC	TTATTTTACGATGTGAATCGGACTTC
alsS	GTTTCTTATTCTCAGCAATGGAATTAG	CTAGAGAGCTTTCGTTTTCATGAG
alsS-2	CAAGGTGTCACACATGTATTTGG	CTACACTGTATTTTGTAATCGGC
alsD	CAAGAAAAACCTTACGTGCCAATG	TTATTCAGGGCTTCCTTCAGTTG
alsD-2	GTCAAAAAGATCAGCCTGTGAG	GCATAAAATAAGTTTCTGCTTGGC
ldh	CAAATTGTAGATGATGTGAAAAACGC	TTAGTTGACTTTTTGTTCTGCAAAATG
pta	CAGTGAAAATCGCGAAAGAAAAAGC	TTACAGTGCTTGCGCCGCT
ackA	GTTGAAGCTACGAAAGAAGGAAATG	TTATTTTGCTAAACGAACAACGTCG
citA	CAGGGTGTACAAAACAAAAGACC	TTATGAAAGAACTTCCTCGGGAAT
citZ	CGAAGCATCTGAAAGAAATGAGC	TTAGGCTCTTTCTTCAATCGGAA
icd	GAGATTACATTTCTGATGCTCTTGC	TTAGTCCATGTTTTTGATCAGTTCTTC

Name	forward primer sequence 5' → 3' [#]	reverse primer sequence 5' → 3' [#]
mdh	CTTGAAACTCTTATTCCGAAAGAAC	TTAGGATAATACTTTCATGACATTTTTGAC
citB	GCAAAAGGAACAAACCTTCTCG	TCAGGACTGCTTCATTTTTTCAC
citB-2	CAAGCGAGAAAAACGTTTACTAC	CCGCTTTATCTACCTGTACAG
odhA	GGTAAAGTGTCTATAGACATTAGC	TTAGTTTTTGCGAGTCAAGCTATC
odhB	CAGGAATCGAAAAAGAGATCGG	TTATCCTTCTAATAAAAGCTGTTCAGG
sucC	CGATGGATATTATCAAGCATTATGG	TTAAACTAAGGATACGATTTTCTGCG
sucC-2	GCTTTTACAGCAGAAGAAGCAG	GTGTTTTCTCCGCAACCTCTTC
sucD	GACGTTTTAAAAGCGTTTAACGAAG	TTAATGCGTTTTACAAGTTTCGAAC
sdhC	GTCAATACAGCTACATGAGAAACTG	TTAAACAAATGCAAAAATCGCTTTTAAGC
sdhA	GATGACAAAATTCAGGAGTTAATGG	TTATTTCGCCACCTTCTTCTTC
sdhB	GTCGTCAATGGGCATATGAATTATC	TTATACTCTGTCGCTTCCGAAG
citG	GAACTGAACGTGTTTAAGCCTG	TTACGCCTTTGGTTTTACCATG
zwf	AAGGATCCGTGAAAACAAACCAACAACCAAAAG	TTATATGTTCCACCAGTGTAAGC
zwf-2	CGCTTTACTTAAATGCTAAAAAGCTTG	CCATTGCCAAGTAGAACATTCTG
gntZ	AAGGATCCTTCAATTCGATTGGTGTCATAGG	TTTCTGCAGTTATTCAGACCAATTCGTATGGAAG
gntZ-2	CAAAGACATCGCCTTGATTTTCC	
rpe	AAAAGTCGACATAAAGGTTGCACCATCTATTCTTTC	TTTAAGCTTTTATTTACTTCCTCTGATTTCAGAAATTG
rpe-2	CTCATTAAAGAGCAAGGAGTGAAG	
ywlF	AAAAGTCGACAAAGTAGCCATTGCATCGGATC	TTTAAGCTTCTACAGGTTTTTCTCTTCATAATCG
ywlF-2	CGATTATCCGGATTATGCTTTTCC	GTTTGTGTCATTATGCTCCCTC
tkt	AAGGATCCGATACAATTGAAAAGAAATCAGTTGC	CGAGTTAAGGCATTAATCAATAAGTAAGCTTTTGAAA
tkt-2	GAGGTAGGTCTTGCAATTGAAG	TTACTTATTGATTAATGCCTTAACTCG
ywjH	AAAAGTCGACATGTTATTCTTTGTTGATACAGCCA	TTTAAGCTTTTATTTGTTCCAGTCTGCCAGG
ywjH-2	GCATCAAAACAAACGTTACATTGATC	
fbp	AAGGATCCTTTAAAAATAATGTCATACTTTTAAATTCACC	TTTCTGCAGTCACTTCATATACCGATACTCC

Name	forward primer sequence 5' → 3' #	reverse primer sequence 5' → 3' #
fbp-2	CATACACCTGTAAAAGAAATCGAAG	CGCTTCTTTCGCATCAAAGTC
gapB	AAAAGTCGACAAGGTAAAAGTAGCGATCAACG	TTTAAGCTTTTATACAGCAGACGGATGTTTCA
gapB-2	GCATCTGAAAGGAAAACTTCACG	CCATCACAATGGTAACGTCTTC
pckA	AAAAGATCTAACTCAGTTGATTTGACCGCTG	TTTCTGCAGTTATACGAGAGGGCCGCCTG
pckA-2	CGTTTTCTTAGTCAATACCGGATG	CGCGCAAATAAATTGTGCCATG
pycA	AAAAGTCGACTTGTCTCAGCAATCGATACAAAAAG	TTTTGTCGACTTATGCTTTTTCAATTTCAAGGAGCAG
pycA-2	CGTCGTTTATTTCGAACTCAACG	GACGGCTTCAAGCGTTTCG
gltA	AAGGATCCACGTACAATCAAATGCCAAAAGCT	TTTTGTCGACTTACTGTACTACCGCTTGTTTTTG
gltA-2	GGTTTAAACGCAAATGCAACCTTG	CAAACGGACAGCTTTTTTGCG
gltB	AAGGATCCATGGGGAAACCAACTGGATTTAT	TTTCTGCAGTTACGGAAGAACTGAACTCCCA
gltB-2	CTTCATACCATCCAAATGGAAAAAGTG	CTTTATCGATAATGGTCCGCTC
glnA	CAAAACAGAAGCCCGCTTATC	GAGACATATACTGTTCGCGTTC
nrgB	AAACATATGAGCGGTCAAATGTTCAAGGTA (CD35)	GATGTTGATCGTATTGCTGATTTC
rocG	GATATCCCTTATTTGCTCGAT (PT10)	CTCTTCTTCCGACCAATAATATC
rocG-2	GCAAAGCAAGTCTCGAAAGATG	CCGTAAGGAAGATTGGCAATC
rocA	CAGATGTTGATGAAAACGCCC	AAAGCTTCGGACGTTGTTTTGG
rocD	AGAATGTCTTGTTTATTGCGGATGA	TCTGCCGCGGACTTCTTTAAT
gudB	CATAATATCCGGGCTAAAATTGTC	TATCCAGCCTCTAAAACGCGAA
argC	TTCGTTATCGTTGACGCGAAGA	ATTGCTGCCGTACACTTCTTTC
argG	ACTGGTGAAAGAGGTAGCC	GCTTTAATCTGCTCCTTCTTTTTC
carA	TAAACATATCGCTCTCATTGATTTTGG	CCACATAGCTGTGGTTTTGG
aspB	ACCCGACATCAATCGCACAATA	GCTATGTTTTTCTACAAAACGCTTG
ansA	CAGATGTATGTCTGCTGAAGC	CAGGACAACGTCATCAGCTAT
pyrAA	AATGTGACGTCATCGTTGTG	TATGCGTCACTTCCAGTTCT
nrgA	ATGTACGCGTTCATCAACACAAA	TAAAATAACCCATCTGCGCCC
citM	CTTTGGTGGTTCAATTTATTGTTAACG	GTCATTCGGCATCAAAAATGTG
citH	CACGTTCACTCATCCCTGAT	GAAAGAAATAATCCCAATCAAAAGCG
dctP	CTTTACGGATTCAGCCTGTG	CGTTGAGGCAAGGACGATAAAT

Name	forward primer sequence 5' → 3' [#]	reverse primer sequence 5' → 3' [#]
gltP	CTCCGGCTTGTCATTGAACT	CTTTCAGTTTGGCTGTTTGCG
gltT	AATCTATTGCCGCACTCTTTTTAG	ACCAGAAATGGTTGCGTTTTGTT
glnQ	GCCATTAACGACAGAAAAACTGAT	CTAGCACCTCACCGATCATTT
pps	CTATCTCCACCACGTGTTATC	CAATGTAGCCTTCTGTACCATG
yqjI	CCAAATGAAAGCTGCTTCTGAAG	CCATTCAGTATGGAAGATGCC
mleA	GCTGTATACGGAACAGACAAATC	GGACAACATCACCTTGTTTAGTG
malS	CAACTCTTTCTGAAGCAAAGCC	CCTGTATAATGTCTTCAGGCTC
ydaP	GAACGCGAATCAGGATTTCATC	CAATCATGACATCCACTACGAC
hns-1	*CCAAGTAATACGACTCACTATAGG*GA AACGCTGGAAGAAATGCTG	GCTTGATCAGGAAATCGTCGAG
ompA-1	*CCAAGTAATACGACTCACTATAGG*CA ACTTCAACAAAGCAACCCTG	CTGAGTTACAACGTCTTTGATACC
ompA-2	*CCAAGTAATACGACTCACTATAGG*CA TGACACTGGTTTCATCAAC	GTTGTTGGTCCACTGGTATTC
ompC-1	*CCAAGTAATACGACTCACTATAGG*GA AATACGACGCTAACAACATC	CGATGTTATCAGTGTTGATGCC
ompC-2	*CCAAGTAATACGACTCACTATAGG*CA TTCGCAGGTCTGAAATTC	GTTCTGAGCATCAGTACGTTTG
mpn474-1	GACCAAACACAACTGCTTGAC	*CTAATACGACTCACTATAGGGAGA*CTT GCAATTGTTG (SH75)
mpn474-2	*CCAAGTAATACGACTCACTATAGG*CA GCAACAATTCCATCAAGC	GTAACAACTCGTAAGCGGAG
mpn474-3	CCAAGTAATACGACTCACTATAGGC CACTAACTTTGAGTTAGGTC	CATCACTCTGTTGCAAGTGG
mpn474-4	ACTCCTTCAACCCAACAAGTC (SH74)	GTAGTTGGTGCAACTGTGTC
mpn474-5	CAACTCCAAGCTGCTTACAATAAC	GCTTGCCATTGATCTTAGCTAG

[#] restriction sites are underlined, promoters are italic

Plasmids

Table 15. Plasmids used in this work

Plasmid	Resistance	Description	Reference
pAC7	ampicillin/ kanamycin	construction of translational *lacZ* fusions	Weinrauch *et al.*, 1991
pBgaB	ampicillin/ neomycin	construction of translational *lacZ* fusions with thermostable β-galactosidase	Mogk *et al.*, 1996
pBluescript II SK^+	ampicillin	cloning	Stratagene, USA

Plasmid	Resistance	Description	Reference
pDG1726	spectinomycin	PCR template for *spc*-cassette	Guérout-Fleury *et al.*, 1995
pET3C	ampicillin	construction of overexpression vector with C-terminal Strep-tag (pGP574)	Novagen, Germany
pGEM-*cat*	chloramphenicol	PCR template for *cat*-cassette	Youngman, 1990
pGP66	ampicillin/ kanamycin	*ptsG*-lacZ fusion: wild type RAT	Stülke *et al.*, 1997
pGP230	ampicillin	overexpression of GlcT-RBD	Schilling *et al.*, 2004
pGP253	pWH844 + *glpK* *Sal*I/*Hind*III	PCR: *gplK* CH7 + CH8; A1512→G	Hames *et al.*, 2005
pGP254	pWH844 + *glpK* *Sal*I/*Hind*III	PCR: *gplK* CH7 + CH10; all TGA codons converted into TGG	Hames *et al.*, 2005
pGP332	ampicillin/ kanamycin	*ptsG-lacZ* fusion: pAC6/*Mfe*I/ *BamH*I; terminator G37C G38C	Schilling *et al.*, 2004
pGP339	ampicillin/ chloramphenicol	*ptsG-lacZ* fusion pAC7/*BamH*I RAT G11A G26A	Schilling *et al.*, 2004
pGP340	ampicillin/ chloramphenicol	*ptsG-lacZ* fusion: pAC7/*BamH*I RAT G11C G26A	Schilling *et al.*, 2004
pGP342	ampicillin/ chloramphenicol	*ptsG-lacZ* fusion: pAC7/*BamH*I wild type RAT	Schilling *et al.*, 2004
pWH844	ampicillin	overexpression vector	Schirmer *et al.*, 1997

Table 16. Plasmids constructed in this work

Plasmid	Construction	Description	Reference
pGP549	pAC7 / *BamH*I + *Eco*RI	CCR: *ptsG*-RAT IL5 + JS11 + OS28/ *BamH*I + *MfeI* (*ptsG*-R1)	Schilling *et al.*, 2004
pGP550	pAC7 / *BamH*I + *Eco*RI	MMR: *ptsG*-RAT IL5 + JS11 + OS116 + OS115 / *BamH*I + *Mfe*I (*ptsG*-R2)	Schilling *et al.*, 2006
pGP551	pAC7 / *BamH*I + *Eco*RI	MMR: *ptsG*-RAT IL5 + JS11 + OS28 + OS117 / *BamH*I + *Mfe*I	
pGP552	pAC7 / *BamH*I + *Eco*RI	MMR: *ptsG*-RAT IL5 + JS11 + OS114 + OS113 / *BamH*I + *Mfe*I (*ptsG*-R3)	Schilling *et al.*, 2006
pGP553	pAC7 / *BamH*I + *Eco*RI	CCR: *ptsG*-RAT IL5 + JS11 + OS112 / *BamH*I + *Mfe*I (*ptsG*-R4)	Schilling *et al.*, 2006

Plasmid	Construction	Description	Reference
pGP554	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG*-RAT IL5 + JS11 + OS117 + OS111 + OS118 + OS119 / *BamH*I + *Mfe*I	
pGP555	pAC7 / *BamH*I + *EcoR*I	CCR: *ptsG*-RAT IL5 + JS11 + OS117 + OS112 / *BamH*I + *Mfe*I	
pGP556	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG*-RAT IL5 + JS11 + OS111 + OS118 + OS119 / *BamH*I + *Mfe*I (*ptsG*-R6)	Schilling *et al.*, 2006
pGP557	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG*-RAT IL5 + JS11 + OS110 + OS118 + OS119 / *BamH*I + *Mfe*I (*ptsG*-R5)	Schilling *et al.*, 2006
pGP558	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG*-RAT IL5 + JS11 + OS117 + OS110 + OS118 + OS119 / *BamH*I + *Mfe*I	
pGP559	pAC7 / BamHI + EcoRI	CCR: *ptsG*-RAT IL5 + JS11 + OS6 + OS7 + OS8 / *BamH*I + *Mfe*I	
pGP560	*pyk* in pBluescript II SK^+ / *BamH*I + *Pst*I	*pyk*-forw1 + *pyk*-rev	
pGP561	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG*-RAT IL5 + JS11 + OS30 + OS31 + OS32 / *BamH*I + *Mfe*I (*ptsG*-R8)	Schilling *et al.*, 2006
pGP562	pBgaB /*BamH*I + *EcoR*I	PCR: *ptsG*-RAT IL5 + JS11 / *BamH*I + *Mfe*I	
pGP563	pWH844 / *BamH*I + *Pst*I	*eno* forw1 + *eno* rev	
pGP564	pAC7 / *BamH*I + *EcoR*I	PCR: *sacB* OS49 + OS50 / *BamH*I + *EcoR*I	Schilling *et al.*, 2006
pGP565	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS55 / *BamH*I + *EcoR*I (*sacB*-R10)	Schilling *et al.*, 2006
pGP566	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS47 / *BamH*I + *EcoR*I (*sacB*-R11)	Schilling *et al.*, 2006
pGP567	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS55 + OS48 + OS54 / *BamH*I + *EcoR*I (*sacB*-R12)	
pGP568	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS47 + OS48 + OS54 / *BamH*I + *EcoR*I (*sacB*-R13)	Schilling *et al.*, 2006
pGP569	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS61 / *BamH*I + *EcoR*I (*sacB*-R2)	Schilling *et al.*, 2006
pGP570	pWH844 / *BamH*I + *Pst*I	OS78 + OS79; introduction of thrombin restriction site; insert is obtained by oligonucleotide annealing	
pGP571	pGP570 / *Sal*I + *Hind*III	SacT-RBD; N-terminal fusion to His_6-tag; can be cleaved off by Thrombin	

Plasmid	Construction	Description	Reference
pGP572	pGP570 / *Sal*I + *Hind*III	LicT-RBD; N-terminal fusion to His_6-tag; can be cleaved off by Thrombin	
pGP573	pGP570 / *Sal*I + *Hind*III	SacY-RBD; N-terminal fusion to His_6-tag; can be cleaved off by Thrombin	
pGP574	pET3C / *Nde*I + *BamH*I	OS91+OS92; cloning of C-terminal Strep-tag into pET3C; introduction of *Sac*I and displacing of *BamH*I restriction site; insert is obtained by oligonucleotide annealing	Schilling *et al.*, 2006
pGP575	pGP574 / *Nde*I + *BamH*I	GlcT-RBD; C-terminal fusion to Strep-tag with OS93 + OS94	Schilling *et al.*, 2006
pGP576	pGP574 / *Nde*I + *BamH*I	LicT-RBD C-terminal fusion to Strep-tag with OS95 + OS96	Schilling *et al.*, 2006
pGP577	pGP574 / *Nde*I + *BamH*I	SacT-RBD C-terminal fusion to Strep-tag with OS97 + OS98	Schilling *et al.*, 2006
pGP578	pGP574 / *Nde*I + *BamH*I	SacY-RBD C-terminal fusion to Strep-tag with OS99 + OS100	
pGP579	pBluescript / *BamH*I + *Apa*I	PCR OS56 + OS57 (*pdhB* knock out)	
pGP580	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS62 / *BamH*I + *EcoR*I (*sacB*-R1)	Schilling *et al.*, 2006
pGP581	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS71 + OS72 / *BamH*I + *EcoR*I	
pGP582	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS61 + OS71 + OS72 / *BamH*I + *EcoR*I (*sacB*-R7)	Schilling *et al.*, 2006
pGP583	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS62 + OS71 + OS72 / *BamH*I + *EcoR*I (*sacB*-R9)	Schilling *et al.*, 2006
pGP584	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS77 / *BamH*I + *EcoR*I	
pGP585	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS62 + OS77 / *BamH*I + *EcoR*I	
pGP586	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS76 + OS72 / *BamH*I + *EcoR*I	
pGP587	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS61 + OS76 + OS72 (*sacB*-R8) / *BamH*I + *EcoR*I	Schilling *et al.*, 2006
pGP588	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS62 + OS76 + OS72	
pGP589	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS88 (*sacB*-R3) / *BamH*I + *EcoR*I	Schilling *et al.*, 2006

Plasmid	Construction	Description	Reference
pGP590	pAC7 / *BamH*I + *EcoR*I	MMR: *sacB* OS49 + OS50 + OS89 (*SacB*-R5) / *BamH*I + *EcoR*I	Schilling *et al.*, 2006
pGP591	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG* OS66 + OS118 + OS119 / *BamH*I + *Mfe*I (*ptsG*-R7)	Schilling *et al.*, 2006
pGP592	pAC7 / *BamH*I + *EcoR*I	MMR: *ptsG* IL5 + JS11 + OS90 + OS119 an pGP557 / *BamH*I + *Mfe*I	
pGP593	pGP574 / *Nde*I + *BamH*I	SacT; C-terminal fusion to Strep-tag with OS97 + OS108	
pGP594	pGP574 / *Nde*I + *BamH*I	SacY; C-terminal fusion to Strep-tag with OS99 + OS109	
pGP595	pAC7 / *BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS102 / *BamH*I + *EcoR*I (*sacB*-R6)	Schilling *et al.*, 2006
pGP596	pAC7 /*BamH*I + *EcoR*I	CCR: *sacB* OS49 + OS50 + OS105 / *BamH*I + *EcoR*I (*sacB*-R4)	Schilling *et al.*, 2006
pGP597	pAC7 /*BamH*I + *EcoR*I	MMR: sacB OS49 + OS50 + OS61 + OS77 / *BamH*I + *EcoR*I	Schilling *et al.*, 2006
pGP598	pAC7 /*BamH*I + *EcoR*I	CCR: pGP592 + IL5 + JS11 + OS103 / *BamH*I + *Mfe*I	
pGP599	pAC7 /*BamH*I + *EcoR*I	CCR: pGP592 + IL5 + JS11 + OS104 / *BamH*I + *Mfe*I	
pGP1100	pWH844 / *BamH*I + *Pst*I	*pyk* forw1 + *pyk* rev	
pGP1101	pWH844 / *BamH*I + *Pst*I	*pgm* forw1 + *pgm* rev	
pGP1102	pWH844 / *Sal*I + *Hind*III	*pgk* forw1 + *pgk* rev	

Strains

Table 17. Strains used in this work

Strain	Genotype	Reference
Escherichia coli		
DH5α	*recA1 endA1 gyrA96 thi hsdR17*rK-mK+*relA1 supE44* Φ80Δ*lacZ*Δ*M15* Δ(*lacZYA-argF*)U169	Sambrook *et al.*, 1989
BL21 (DE3)	F- *lon ompT rBmB hsdS gal* (cIts857*ind*1 Sam7 *nin*5 *lac*UV5- T7 gene1)	Sambrook *et al.*, 1989
Bacillus subtilis		
168	*trpC2*	Laboratory collection
BGW10	*trpC2 lys-3* Δ*licTS* ::*erm*	Krüger and Hecker, 1995

Strain	Genotype	Reference
BSIP1114	*trpC2 amyE*::(*pta-lacZ cat*)	Presecan-Siedel *et al.*, 1999
GM1010	*sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	Le Coq *et al.*, 1995
GM1112	*sacXY*Δ*3 sacB*Δ*23 sacT*Δ*4 bglP*::Tn*10 erm amyE*::(*sacB-lacZ phl*)	Le Coq *et al.*, 1995
GP109	*trpC2* Δ*glcT8 amyE*::(ΔLA *ptsG*'-'*lacZ aphA3*)	Bachem and Stülke, 1998
GP133	*trpC2 amyE*::(*fbaA-lacZ cat)*	Ludwig *et al.*, 2001
GP150	*trpC2 ΔglcT8 amyE*::('*lacZ cat*)	Langbein *et al.*, 1999
GP151	*trpC2 amyE*::(ΔLA$_{C75G\ C76G}$ *ptsG´-´lacZ aphA3*)	Langbein *et al.*, 1999)
GP169	*trpC2 amyE*::(ΔLA$_{C75G\ C76G\ G37C\ G38C}$ *ptsG´-´lacZ aphA3*)	Schilling *et al.*, 2004
GP172	*trpC2 amyE*::(ΔCA$_{G11A\ G26A}$ *ptsG*'-'*aphA3*-'*lacZ cat*)	Schilling *et al.*, 2004
GP173	*trpC2 amyE*::(ΔCA$_{G11C\ G26A}$ *ptsG*'-'*aphA3*-'*lacZ cat*)	Schilling *et al.*, 2004
GP174	*trpC2 amyE*::(ΔCA *ptsG*'-'*aphA3*-'*lacZ cat*)	Schilling *et al.*, 2004
GP175	*trpC2 amyE*::(ΔLA $_{U7C}$ *ptsG´-´lacZ aphA3*)	Schilling *et al.*, 2004
GP176	*trpC2 amyE*::(ΔLA $_{A8G}$ *ptsG´-´lacZ aphA3*)	Schilling *et al.*, 2004
GP177	*trpC2 amyE*::(ΔLA $_{U7C\ A8G}$ *ptsG´-´lacZ aphA3*)	Schilling *et al.*, 2004
GP179	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ \Delta1}$ *ptsG´- aphA-lacZ cat*)	Schilling *et al.*, 2004
GP181	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G38A}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP182	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G47U}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP183	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ \Delta2}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP184	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ C75U}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP185	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G90A}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP186	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ \Delta4}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP187	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G72U}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP188	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ \Delta3}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP189	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G28U}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP190	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ C75A}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP191	*trpC2 amyE*::(ΔLA $_{G11C\ G26A\ G47U}$ *ptsG´- aphA3-lacZ cat*)	Schilling *et al.*, 2004
GP205	*trpC2 amyE*::(*citB-lacZ cat*)	Blencke *et al.*, 2006
GP216	*trpC2 amyE*::(*pdhA-lacZ cat*)	Schilling *et al.*, 2007
GP250	*trpC2 amyE*::(*nrgA-lacZ aphA3*)	Detsch and Stülke, 2003
GP268	*trpC2 amyE*::(*glnR-lacZ aphA3*)	Schilling *et al.*, 2007
GP313	*trpC2 amyE*::(*cggR-lacZ cat*)	Ludwig *et al.*, 2001
GP314	*trpC2 amyE*::(*pgi-lacZ cat*)	Ludwig *et al.*, 2001
GP315	*trpC2 amyE*::(*pfkA-lacZ cat*)	Ludwig *et al.*, 2001
GP317	*trpC2 amyE*::(*pgk-lacZ cat*)	Ludwig *et al.*, 2001
GP342	*trpC2 amyE*::(*gltA-lacZ aphA3*)	Wacker *et al.*, 2003

Strain	Genotype	Reference
QB5435	*trpC2* Δ*ptsG*::*cat*	Stülke *et al.*, 1997
QB5448	*trpC2 amyE*::*(ΔLA ptsG'-'lacZ aphA3)*	Stülke *et al.*, 1997
QB5449	*trpC2 ptsG*::*cat amyE*::(Δ*LA ptsG-lacZ aphA3*)	Stülke *et al.*, 1997
QB5556	*trpC2 amyE*::*(rocD'-'lacZ cat)*	Gardan *et al.*, 1995
QB7041	*trpC2 amyE*::(ΔCA*ptsG-lacZ cat*)	Stülke *et al.*, 1997

Table 18. *B. subtilis* strains constructed in this work

Name	Genotype	Construction [a]	Reference
GP366	*trpC2 amyE*::(ΔLA$_{C2G\ G28C\ C85G}$ *ptsG´-´lacZ aphA3*)	pGP532 → GP368	Schilling *et al.*, 2004
GP367	*trpC2 amyE*::(ΔLA$_{C2G}$ *ptsG´-´lacZ aphA3*)	pGP533 → 168	Schilling *et al.*, 2004
GP368	*trpC2 amyE*::(ΔLA$_{G28C\ C85G}$ *ptsG´-´lacZ aphA3*)	pGP534 → 168	Schilling *et al.*, 2004
GP369	*trpC2 amyE*::(ΔLA$_{G5C\ C23G\ G90C}$ *ptsG´-´lacZ aphA3*)	pGP535 → 168	Schilling *et al.*, 2004
GP370	*trpC2 amyE*::(ΔLA$_{G26U\ G87A}$ *ptsG´-´lacZ aphA3*)	pGP536 → 168	Schilling *et al.*, 2004
GP371	*trpC2 amyE*::(ΔLA$_{U7G}$ *ptsG´-´lacZ aphA3*)	pGP537 → 168	Schilling *et al.*, 2004
GP375	*trpC2 amyE*::(ΔLA$_{G3C}$ *ptsG´-´lacZ aphA3*)	pGP541 → 168	Schilling *et al.*, 2004
GP376	*trpC2 amyE*::(ΔLA (RAT*[3])*ptsG´-´lacZ aphA3*)	pGP542 → 168	Schilling *et al.*, 2004
GP377	*trpC2 amyE*::(ΔLA$_{U6G}$ *ptsG´-´lacZ aphA3*)	pGP543 → 168	Schilling *et al.*, 2004
GP378	*trpC2 amyE*::(ΔLA$_{U6A}$ *ptsG´-´lacZ aphA3*)	pGP544 → 168	Schilling *et al.*, 2004
GP379	*trpC2 amyE*::(ΔLA$_{U25G\ A88C}$ *ptsG´-´lacZ aphA3*)	pGP545 → 168	Schilling *et al.*, 2004
GP380	*trpC2 amyE*::(ΔLA$_{A27U\ U86A}$ *ptsG´-´lacZ aphA3*)	pGP546 → 168	Schilling *et al.*, 2004
GP381	*trpC2 amyE*::(ΔLA$_{A27G\ U86C}$ *ptsG´-´lacZ aphA3*)	pGP547 → 168	Schilling *et al.*, 2004
GP385	*trpC2 amyE*::(ΔLA$_{R1}$ *ptsG´-´lacZ aphA3*)	pGP549 → 168	Schilling *et al.*, 2004
GP386	*trpC2 amyE*::(Δ LA$_{R1}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP549 → GP150	Schilling *et al.*, 2004
GP387	*amyE*::(ΔLA *ptsG´-´lacZ aphA3*) *sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	pGP66 → GM1010	Schilling *et al.*, 2004

Name	Genotype	Construction [a]	Reference
GP388	*amyE*::(ΔLA $_{R1}$ *ptsG´-´lacZ aphA3*) *sacXY*Δ3 *sacB*Δ23 *sacT*Δ4	pGP549 → GM1010	Schilling *et al.*, 2004
GP389	*trpC2 lys-3 amyE*::(ΔLA *ptsG´-´lacZ aphA3*) Δ*licTS* ::*erm*	pGP66 → BGW10	Schilling *et al.*, 2004
GP390	*trpC2 lys-3 amyE*::(Δ LA $_{R1}$ *ptsG´-´lacZ aphA3*) Δ*licTS* ::*erm*	pGP549 → BGW10	Schilling *et al.*, 2004
GP391	prototrophic	GM1010 chrom. → 168	
GP392	*trpC2 amyE*::(ΔLA$_{G26T\ G87A}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP536 → GP150	Schilling *et al.*, 2004
GP393	*trpC2 amyE*::(ΔLA $_{R2}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP550 → GP150	
GP394	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ U18G}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP551 → GP150	
GP395	*trpC2 amyE*::(ΔLA $_{R3}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP552 → GP150	
GP396	*trpC2 amyE*::(ΔLA $_{R4}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP553 → GP150	Schilling *et al.*, 2006
GP397	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U25A\ A88U\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP554 → GP150	
GP398	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP555 → GP150	
GP399	*trpC2 amyE*::(ΔLA $_{R6}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP556 → GP150	Schilling *et al.*, 2006
GP400	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP557 → GP150	Schilling *et al.*, 2006
GP401	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U25A\ A88U\ U18G}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP558 → GP150	
GP402	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*) Δ*licTS* ::*erm*	pGP557 → GP427	Schilling *et al.*, 2006
GP403	*trpC2 lys-3 amyE*::(ΔLA $_{A4\text{-}ins\ A4U\ U25A\ A88U\ U18G}$ *ptsG´-´lacZ aphA3*) Δ*licTS*::*erm*	pGP558 → GP427	
GP404	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*)	pGP557 → 168	Schilling *et al.*, 2006
GP405	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U25A\ A88U\ U18G}$ *ptsG´-´lacZ aphA3*)	pGP558 → 168	
GP406	*trpC2 amyE*::(ΔLA$_{U7G\ G26U\ C87A}$ *ptsG´-´lacZ aphA3*)	pGP559 → 168	

Name	Genotype	Construction [a]	Reference
GP407	*trpC2 lys-3 amyE*::(ΔLA$_{A4\text{-}ins\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*) Δ*licTS* ::*erm*	pGP555 → GP427	
GP408	*trpC2 amyE*::(ΔLA $_{R6}$ *ptsG´-´lacZ aphA3*)	pGP556 → 168	Schilling *et al.*, 2006
GP409	*trpC2 amyE*::(ΔLA $_{R6}$ *ptsG´-´lacZ aphA3*) Δ*licTS*::*erm*	pGP556 → GP427	Schilling *et al.*, 2006
GP410	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U25A\ A88U\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*)	pGP554 → 168	
GP411	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U25A\ A88U\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*) Δ*licTS* ::*erm*	pGP554 → GP427	
GP412	*trpC2 amyE*::(ΔLA *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP342 → GP109	
GP413	*trpC2 amyE*::(ΔLA $_{R2}$ *ptsG´-´lacZ aphA3*)	pGP550 → GP168	Schilling *et al.*, 2006
GP414	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ U18G}$ *ptsG´-´lacZ aphA3*)	pGP551 → GP168	
GP415	*trpC2 amyE*::(ΔLA $_{R3}$ *ptsG´-´lacZ aphA3*)	pGP552 → GP168	Schilling *et al.*, 2006
GP416	*trpC2 amyE*::(ΔLA $_{R4}$ *ptsG´-´lacZ aphA3*)	pGP553 → GP168	Schilling *et al.*, 2006
GP417	*trpC2 amyE*::(ΔLA $_{R4}$ *ptsG´-´lacZ aphA3*) *ΔlicTS*::*erm*	pGP553 → GP427	Schilling *et al.*, 2006
GP418	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ G3A\ U18G}$ *ptsG´-´lacZ aphA3*)	pGP555 → GP168	
GP419	*trpC2 amyE*::(ΔLA $_{R8}$ *ptsG´-´lacZ aphA3*)	pGP561 → GP168	Schilling *et al.*, 2006
GP420	*trpC2 amyE*::(ΔLA $_{R8}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	pGP561 → GP150	Schilling *et al.*, 2006
GP421	*trpC2 amyE*::(ΔLA $_{R8}$ *ptsG´-´lacZ aphA3*) *ΔlicTS* ::*erm*	pGP561 → GP427	Schilling *et al.*, 2006
GP422	*trpC2 amyE*::*(ΔLA ptsG'- 'lacZ neo)*	pGP562 → 168	
GP425	*trpC2* Δ*sacY*::*cat*	Long Flanking →168	Schilling *et al.*, 2006
GP427	*trpC2* Δ*licTS*::*erm*	BGW10 chrom. → 168	Schilling *et al.*, 2006
GP429	*trpC2* Δ*sacT*::*spec*	Long Flanking → 168	Schilling *et al.*, 2006
GP430	*trpC2* Δ*sacY*::*cat* Δ*sacT*::*spec*	GP429 chrom. → GP425	Schilling *et al.*, 2006
GP431	*trpC2* Δ*licTS*::*erm*Δ *sacT*::*spec*	GP429 chrom. → GP427	Schilling *et al.*, 2006

Name	Genotype	Construction [a]	Reference
GP432	*trpC2* Δ*licTS*::*erm* Δ*sacY*::*cat*	GP425 chrom. → GP427	Schilling *et al.*, 2006
GP433	*trpC2* Δ*licTS*::*erm* Δ*sacY*::*cat* *sacT::spec*	GP429 chrom. → GP432	Schilling *et al.*, 2006
GP434	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*) *ΔglcT8* Δ*licTS*::*erm*	BGW10 chrom. → GP400	Schilling *et al.*, 2006
GP435	*trpC2* Δ*sacY*::*cat amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*) *ΔglcT8*	GP425 chrom. → GP400	Schilling *et al.*, 2006
GP436	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG´-´lacZ aphA3*) *ΔglcT8* Δ*sacT*::spec	GP429 chrom. → GP400	Schilling *et al.*, 2006
GP437	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*)	pGP564 → 168	Schilling *et al.*, 2006
GP438	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) Δ*sacY::cat*	pGP564 → GP425	Schilling *et al.*, 2006
GP439	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) Δ*licTS*::*erm*	pGP564 → GP427	
GP440	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) Δ*sacT*::spec	pGP564 → GP49	Schilling *et al.*, 2006
GP441	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) Δ*sacT*::spec Δ*sacY::cat*	pGP564 → GP430	Schilling *et al.*, 2006
GP442	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) *sacB* Δ*licTS::erm*Δ *sacT::spec*	pGP564 → GP431	
GP443	*trpC*2 *amyE*::(ΔLA *sacB'- 'lacZ aphA3*) *sacB* Δ*licTS*::*erm* Δ*sacY::cat*	pGP564 → GP432	
GP444	*trpC*2 *amyE::*(ΔLA $_{R10}$ *sacB'- 'lacZ aphA3*)	pGP565 → 168	Schilling *et al.*, 2006
GP445	*trpC*2 *amyE*::(ΔLA $_{R10}$ *sacB'- 'lacZ aphA3*) Δ*licTS*::*erm*	pGP565 → GP427	
GP446	*trpC*2 *amyE*::(ΔLA $_{R11}$ *sacB'- 'lacZ aphA3*)	pGP566 → 168	Schilling *et al.*, 2006
GP447	*trpC*2 *amyE*::(ΔLA $_{R11}$ *sacB'- 'lacZ aphA3*) Δ*licTS*::*erm*	pGP566 → GP427	
GP448	*trpC*2 *amyE*::(ΔLA $_{R12}$ *sacB'- 'lacZ aphA3*)	pGP567 → 168	Schilling *et al.*, 2006
GP449	*trpC*2 *amyE*::(ΔLA $_{R12}$ *sacB'- 'lacZ aphA3*) Δ*licTS*	pGP567 → GP427	
GP450	*trpC*2 *amyE*::(ΔLA $_{R13}$ *sacB'- 'lacZ aphA3*)	pGP568 → 168	Schilling *et al.*, 2006
GP451	*trpC*2 *amyE*::(ΔLA $_{R13}$ *sacB'- 'lacZ aphA3*) *ΔglcT8*	pGP568 → GP150	Schilling *et al.*, 2006

Name	Genotype	Construction [a]	Reference
GP452	*trpC*2 *amyE*::(ΔLA $_{R13}$ *sacB'-'lacZ aphA3*) Δ*licTS*	pGP568 → GP427	
GP453	*trpC*2 *amyE*::(ΔLA $_{R13}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat sacT*::*spec*	pGP568 → GP433	Schilling *et al.*, 2006
GP454	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm*Δ *sacT*::*spec*	pGP557 → GP431	Schilling *et al.*, 2006
GP455	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat*	pGP557 → GP432	Schilling *et al.*, 2006
GP456	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat sacT*::*spec*	pGP557 → GP433	Schilling *et al.*, 2006
GP457	*trpC*2 Δ*ydap*::*spec*	LFH → 168	
GP458	*trpC2 amyE*::*(citB-`lacZ cat)* Δ*ydap*::*spec* Δ*pta*::*aphA3*	LFH → GP241	
GP459	*trpC*2 Δ*pdhB*::*spec*	pGP579 → 168	
GP460	*trpC*2 *amyE*::(ΔLA $_{R2}$ *sacB'-'lacZ aphA3*)	pGP569 → 168	Schilling *et al.*, 2006
GP461	*trpC*2 *amyE::*(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*)	pGP580 → 168	Schilling *et al.*, 2006
GP462	*trpC*2 *amyE*::(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP580 → GP427	Schilling *et al.*, 2006
GP463	*trpC*2 *amyE*::(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*) Δ*sacY::cat sacT*::*spec*	pGP580 → GP430	Schilling *et al.*, 2006
GP464	*trpC2 amyE*::(ΔLA $_{R7}$ *ptsG'-'lacZ aphA3*)	pGP591 → 168	Schilling *et al.*, 2006
GP465	*trpC*2 *amyE*::(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm sacT*::*spec*	pGP580 → 431	Schilling *et al.*, 2006
GP466	*trpC*2 *amyE*::(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY::cat*	pGP580 → 432	Schilling *et al.*, 2006
GP467	*trpC*2 *amyE*::(ΔLA $_{R1}$ *sacB'-'lacZ aphA3*) *ΔglcT8*	pGP580 → GP150	
GP468	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*sacY*::*cat* Δ*sacT*::*spec*	pGP557 → GP430	
GP469	*trpC*2 Δ*csrA*::*spec*	LFH → 168	
GP470	*trpC2* Δ*ptsG::cat* Δ*bglP::*Tn*10 erm*	chrom. DNA GM1112→ QB5435	Schilling *et al.*, 2006
GP471	*trpC*2 *amyE*::(ΔLA $_{R2}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm sacT*::*spec*	pGP569 → GP431	Schilling *et al.*, 2006

Name	Genotype	Construction [a]	Reference
GP472	*trpC*2 *amyE*::(ΔLA $_{R2}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat*	pGP569 → GP 432	Schilling *et al.*, 2006
GP473	*trpC*2 *amyE*::(ΔLA $_{R2}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP569 → GP427	
GP474	*trpC2* Δ*ptsG*::*cat*	chrom. DNA QB5449 → 168	
GP475	*trpC2* Δ*bglP*::*erm*	chrom. DNA GM1112 → 168	
GP476	*trpC*2 *amyE*::(ΔLA $_{R7}$ *sacB'-'lacZ aphA3*)	pGP582 → 168	Schilling *et al.*, 2006
GP477	*trpC*2 *amyE*::(ΔLA $_{R8}$ *sacB'-'lacZ aphA3*)	pGP587 → 168	Schilling *et al.*, 2006
GP478	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*)	pGP581 → 168	
GP479	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G\ G26A}$ *sacB'- 'lacZ aphA3*)	pGP586 → 168	
GP480	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*)	pGP583 → 168	Schilling *et al.*, 2006
GP481	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G\ G26A}$ *sacB'- 'lacZ aphA3*)	pGP588 → 168	
GP482	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G}$ *sacB'- 'lacZ aphA3*)	pGP584 → 168	
GP483	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G}$ *sacB'- 'lacZ aphA3*)	pGP585 → 168	
GP484	*trpC*2 *amyE*::(ΔLA $_{R7}$ *sacB'-'lacZ aphA3*) Δ*licTS* ::*erm*	pGP582 → GP427	Schilling *et al.*, 2006
GP485	*trpC*2 *amyE*::(ΔLA $_{R7}$ *sacB'-'lacZ aphA3*) Δ*sacY*::*cat sacT*::*spec*	pGP582 → GP430	
GP486	*trpC*2 *amyE*::(ΔLA $_{R8}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP587 → GP427	Schilling *et al.*, 2006
GP487	*trpC*2 *amyE*::(ΔLA $_{R8}$ *sacB'- 'lacZ aphA3*) *sacT*::*spec*	pGP587 → GP429	Schilling *et al.*, 2006
GP488	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*) *sacT*::*spec*	pGP581 → GP429	
GP489	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*) Δ*sacY*::*cat*	pGP581 → GP425	
GP490	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G\ G26A}$ *sacB'- 'lacZ aphA3*) *sacT*::*spec*	pGP586 → GP429	
GP491	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G\ G26A}$ *sacB'- 'lacZ aphA3*) Δ*sacY*::*cat*	pGP586 → GP425	
GP492	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP583 → GP427	Schilling *et al.*, 2006

Name	Genotype	Construction [a]	Reference
GP493	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* *sacT*::*spec*	pGP583 → GP431	Schilling *et al.*, 2006
GP494	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat*	pGP583 → GP432	Schilling *et al.*, 2006
GP495	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*) *sacT*::*spec* Δ*sacY*::*cat*	pGP583 → GP430	
GP496	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G\ G26A}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP588 → GP427	
GP497	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G\ G26A}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP588 → GP429	
GP498	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP584 → GP429	
GP499	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G}$ *sacB'-'lacZ aphA3*) Δ*sacY*::*cat*	pGP584 → GP425	
GP518	*trpC*2 *amyE*::(ΔLA$_{A13C\ U14G\ G26A}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat* Δ*sacT*::*spec*	pGP586 → GP433	
GP519	*trpC*2 *amyE*::(ΔLA $_{R3}$ *sacB'-'lacZ aphA3*)	pGP589 → 168	Schilling *et al.*, 2006
GP520	*trpC*2 *amyE*::(ΔLA $_{R5}$ *sacB'-'lacZ aphA3*)	pGP590 → 168	Schilling *et al.*, 2006
GP521	*trpC*2 *amyE*::(ΔLA $_{R3}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP589 → GP429	Schilling *et al.*, 2006
GP522	*trpC*2 *amyE*::(ΔLA $_{R5}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP590 → GP429	Schilling *et al.*, 2006
GP523	*trpC*2 *amyE*::(ΔLA $_{R3}$ *sacB'-'lacZ aphA3*) Δ*sacY*::*cat*	pGP589 → GP429	
GP524	*trpC*2 *amyE*::(ΔLA $_{R5}$ *sacB'-'lacZ aphA3*) Δ*sacY*::*cat*	pGP590 → GP425	
GP525	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U14del\ C15A\ G16A\ U18G\ U25A\ A88U}$ *ptsG'-'lacZ aphA3*)	pGP592 → 168	
GP526	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U14del\ C15A\ G16A\ U18G\ U25A\ A88U}$ *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP592 → GP427	
GP527	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ A4U\ U14del\ C15A\ G16A\ U18G\ U25A\ A88U}$ *ptsG'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP592 → GP429	
GP528	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*ptsG*::*cat*	pGP557 → GP474	

Name	Genotype	Construction [a]	Reference
GP529	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*ptsG*::*cat*	pGP557 → QB5435	
GP530	*trpC2 amyE*::(ΔLA $_{R5}$ *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP557 → GP473	
GP531	*amyE*::(ΔLA *ptsG'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP66 → GP473	
GP532	*trpC*2 *amyE*::(ΔLA $_{R9}$ *sacB'-'lacZ aphA3*) *ΔglcT8*	pGP583 → GP150	
GP533	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP581 → GP427	
GP534	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*Δ *sacT*::*spec*	pGP581 → GP431	
GP535	*trpC*2 *amyE*::(ΔLA$_{G26A}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm* Δ*sacY*::*cat*	pGP581 → GP432	
GP536	*trpC*2 *amyE*::(ΔLA $_{R7}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*Δ *sacT*::*spec*	pGP582 → GP431	Schilling *et al.*, 2006
GP537	*trpC*2 *amyE*::(ΔLA $_{R6}$ *sacB'-'lacZ aphA3*)	pGP595 → GP168	Schilling *et al.*, 2006
GP538	*trpC*2 *amyE*::(ΔLA $_{R6}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP595 → GP429	Schilling *et al.*, 2006
GP539	*trpC*2 *amyE*::(ΔLA $_{R6}$ *sacB'-'lacZ aphA3*) *sacT*::*spec* Δ*sacY*::*cat*	pGP595 → GP430	Schilling *et al.*, 2006
GP540	*trpC*2 *amyE*::(ΔLA $_{R4}$ *sacB'-'lacZ aphA3*)	pGP596 → GP168	Schilling *et al.*, 2006
GP541	*trpC*2 *amyE*::(ΔLA $_{R4}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP596 → GP429	Schilling *et al.*, 2006
GP542	*trpC*2 *amyE*::(ΔLA $_{R4}$ *sacB'-'lacZ aphA3*) *sacT*::*spec* Δ*sacY*::*cat*	pGP596 → GP430	Schilling *et al.*, 2006
GP543	*trpC*2 *amyE*::(ΔLA $_{R10}$ *sacB'-'lacZ aphA3*) *sacT*::*spec*	pGP565 → GP429	Schilling *et al.*, 2006
GP544	*trpC*2 *amyE*::(ΔLA $_{R7}$ *sacB'-'lacZ aphA3*) *sacT*::*spec*	pGP582 → GP429	Schilling *et al.*, 2006
GP545	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G}$ *sacB'-'lacZ aphA3*) Δ*licTS*::*erm*	pGP597 → GP427	
GP546	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G}$ *sacB'-'lacZ aphA3*) Δ*sacT*::*spec*	pGP597 → GP429	
GP547	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ G3A\ A4U\ U8G\ U14del\ C15A\ G16A\ U18G\ U25A\ A88U}$ *ptsG'-'lacZ aphA3*)	pGP598→ GP168	

Name	Genotype	Construction [a]	Reference
GP548	*trpC2 amyE*::(ΔLA$_{A4\text{-}ins\ G3U\ A4U\ U14del\ C15A\ G16A\ U18G\ U25A\ A88U}$ *ptsG′-′lacZ aphA3*)	pGP599 → GP168	
GP549	*trpC*2 *amyE*::(ΔLA$_{U3G\ A13C\ U14G}$ *sacB'- 'lacZ aphA3*) Δ*sacT*::*spec* Δ*licTS*::*erm*	pGP597 → GP430	

[a] Arrows indicate construction by transformation.

Table 19. Regulation of genes encoding enzymes of the central metabolic pathways as determined by microarray analysis

Gene	Operon	Description	Regulation factor a	Standard deviation b
Glycolysis				
ptsG	*ptsGHI*	glucose-specific enzyme II	0.99	0.071
ptsH	*ptsGHI*	HPr	1.02	0.11
ptsI	*ptsGHI*	PTS enzyme I	0.98	0.75
pgi	*pgi*	phosphoglucoisomerase	1.06	0.204
pfk	*pfk pykA*	phosphofructokinase	1.30	0.228
fbaA	*fbaA*	fructose-1,6-bisphosphate aldolase	1.20	0.199
tpi	*cggR gapA pgk tpiA pgm eno*	triose phosphate isomerase	1.23	0.128
gapA	*cggR gapA pgk tpiA pgm eno*	glyceraldehyde-3-phosphate dehydrogenase	1.13	0.183
pgk	*cggR gapA pgk tpiA pgm eno*	phosphoglycerate kinase	1.25	0.125
pgm	*cggR gapA pgk tpiA pgm eno*	phosphoglycerate mutase	1.19	0.167
eno	*cggR gapA pgk tpiA pgm eno*	enolase	1.08	0.184
pykA	*pfk pykA*	pyruvate kinase	1.29	0.175
Pyruvate dehydrogenase				
pdhA	*pdhABCD*	pyruvate dehydrogenase (pyruvate decarboxylase, subunit α)	1.69	0.18

pdhB	*pdhABCD*	pyruvate dehydrogenase (pyruvate decarboxylase, subunit β)	1.74	0.214
pdhC	*pdhABCD*	pyruvate dehydrogenase (dihydrolipoamide acetyltransferase)	1.42	0.173
pdhD	*pdhABCD*	pyruvate dehydrogenase (lipoamide dehydrogenase)	1.32	0.145
Fermentative and overflow metabolism				
alsS	*alsSD*	α-acetolactate synthase	10.04	1.372
alsD	*alsSD*	α-acetolactate decarboxylase	4.2	1.039
ldh	*ldh lctP*	lactate dehydrogenase	1.67	0.276
pta	*pta*	phosphotransacetylase	1.67	0.287
ackA	*ackA*	acetate kinase	2.94	0.65
mleA		malate oxidoreductase	1.27	0.371
Krebs acid cycle				
citA		citrate synthase (minor)	1.27	0.231
citZ	*citZ icd mdh*	citrate synthase (major)	0.20	0.023
icd	*citZ icd mdh*	isocitrate dehydrogenase	0.55	0.033
mdh	*citZ icd mdh*	malate dehydrogenase	0.62	0.057
citB		aconitase	0.17	0.023

odhA	*odhAB*	2-oxoglutarate dehydrogenase (E1 subunit)	0.75	0.126
odhB	*odhAB*	2-oxoglutarate dehydrogenase (E2 subunit)	0.73	0.098
sucC	*sucCD*	succinyl-CoA synthetase (beta subunit)	0.58	0.105
sucD	*sucCD*	succinyl-CoA synthetase (alpha subunit)	0.65	0.077
sdhC	*sdhCAB*	succinate dehydrogenase (cytochrome *b*-558 subunit)	0.92	0.081
sdhA	*sdhCAB*	succinate dehydrogenase (flavoprotein subunit)	0.82	0.056
sdhB	*sdhCAB*	succinate dehydrogenase (iron-sulfur protein)	0.82	0.091
citG		fumarate hydratase	1.03	0.144
Pentose phosphate pathway				
zwf		glucose-6-phosphate 1-dehydrogenase	1.16	0.216
yqjI		6-phosphogluconate dehydrogenase	1.05	0.177
rpe		ribulose-5-phosphate epimerase	1.12	0.15
ywlF		ribose-5-phosphate isomerase (putative)	1.45	0.13
tkt		transketolase	1.14	0.186
ywjH		transaldolase (putative)	1.11	0.142
Gluconeogenesis				
fbp		fructose-1.6-bisphosphatase	0.76	0.161
gapB		glyceraldehyde-3-phosphate dehydrogenase	1.06	0.348

pckA		phosphoenolpyruvate carboxykinase	0.96	0.216
malS		malate dehydrogenase (decarboxylating)	0.9	0.143
Anaplerotic reaction				
pycA		pyruvate carboxylase	0.65	0.127
Transporter of organic acids				
citM	*citM (yflO)*	Mg^{2+}/citrate transporter	1.13	0.382
citH	*citH (yxiQ)*	Ca^{2+}/citrate transporter	1.21	0.332
dctP	*dctP(ydbH)*	C4-dicarboxylate transport protein	0.82	0.303
gltP	*gltP*	glutamate-aspartate carrier protein	1.03	0.343
gltT	*gltT (yhfG)*	H^{+}/Na^{+}-glutamate symport protein	1.25	0.249
glnQ	*glnQHMP*	glutamine ABC transporter	1.20	0.327
Nitrogen metabolism				
gltA	*gltAB*	glutamate synthase (large subunit)	0.42	0.047
gltB	*gltAB*	glutamate synthase (small subunit)	0.40	0.055
glnA	*glnRA*	glutamine synthetase	0.67	0.060
nrgA	*nrgAB*	ammonium transporter (AmtB)	0.98	0.311
nrgB	*nrgAB*	PII-like protein (GlnK)	1.11	0.301
rocG	*rocG*	glutamate dehydrogenase	0.65	0.120

gudB	*gudB*	glutamate dehydrogenase (cryptic)	1.06	0.267
Arginine metabolism				
rocA	*rocABC*	pyrroline-5-carboxylate dehydrogenase	0.7	0.106
rocD	*rocDEF*	ornithine aminotransferase	1.17	0.246
argC	*ArgCJBD-cpa-F*	N-acetylglutamate-γ-semialdehyde dehydrogenase	1.05	0.15
argG	*argGH*	argininosuccinate synthase	1.19	0.231
carA	*carAB*	carbamoyl-P-transferase	1.29	0.307
Urea cycle				
aspB	*aspB*	aspartate aminotransferase	1.10	0.121
ansA	*ansAB*	L-asparaginase	1.09	0.272
Other genes				
pyrAA	*pyrAAAB*	carbamoyl-phosphate synthetase	0.94	0.06
pps		phosphoenolpyruvate synthase	1.08	0.188
crh	*yvcIJKL crh yvcN*	catabolite repression HPr-like protein	0.94	0.091

[a] The relative gene expression in the presence of succinate and glutamate is shown (CG medium: 1.0).

[b] The standard deviation includes values from three arrays that represent two biological experiments and one dye-flip experiment.

Curriculum vitae

Personal Information

Date of Birth	May 2nd, 1973
Place of Birth	Bamberg, Germany
Nationality	German
Parents	Johann Schilling and Rosemarie Schilling

School Education

1979 – 1982	Primary School (Grundschule Walsdorf)
1982 – 1985	Secondary School (Hauptschule Walsdorf)
1985 – 1993	ETA-Hoffmann-Gymnasium Bamberg
1993	General University Qualification (Abitur)

Activities before studies

1993 – 1994	Civil Service, Caritas Verband Erlangen
1994 – 1996	Trainee in Financial Services, Sparkasse Bamberg

Scientific Education

1996 – 2001	Studies of Biology and Chemistry at the University of Erlangen-Nuremberg
2001 – 2003	Diploma Studies of Biology
2002 – 2003	Diploma thesis in the group of PD Dr. Jörg Stülke, at the University of Erlangen-Nuremberg
Thesis title:	"Specific determinants of the interaction of the Antiterminator Protein GlcT with its target mRNA"
10/2003 – 08/2006	Ph.D. Student in the group of Prof. Dr. Jörg Stülke, at the University of Göttingen
Thesis title:	"Carbon Catabolism in *Bacillus subtilis*: Global and Molecular Views on the Control of Gene Expression
since 08/2006	Employee at Miltenyi Biotec GmbH