Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Ahmed-Hocine Boumezbeur

born in: Constantine, Algeria

Oral examination: 27.05.2020

Riboswitches as targets for metabolic engineering

in *Bacillus subtilis*

Abstract

Many metabolic pathways in bacteria are modulated by metabolite-sensing riboswitches, which regulate gene expression at the level of transcription elongation or translation initiation. Riboswitches represent promising targets to modulate expression of genes and operons relevant for the biotechnological production of commercially relevant compounds. In Firmicutes, approximately 70% of all putative and validated riboswitches (are predicted to) act exclusively at the transcriptional level using a termination-antitermination mechanism.

In a first attempt to interfere with purine-sensing riboswitches and deregulate purine metabolism in *Bacillus subtilis*, a set of synthetic small RNAs (sRNAs) targeting the purine-sensing aptamers were designed to impair ligand binding using rational design combined with *in silico* evolution. However, the designed sRNAs did not show any activity *in vivo* on the riboswitch controlling purine biosynthesis (*pur* operon riboswitch). The effect of the antisense RNA (asRNA) perfectly complementary to the aptamer of the *pur* operon riboswitch was also tested; The asRNA did not affect negatively expression of a riboswitch-regulated *lacZ* gene, yet similarly to the partially complementary sRNAs, the asRNA did not impair the downregulation exerted by the riboswitch in the presence of ligand. Finally, expression of the small RNAs in *B. subtilis*

was quantified, and the kinetic limitations for their hybridization with the aptamer and their competition with the ligand are discussed.

A second metabolic engineering strategy based on editing the genome of *B. subtilis* with regard to transcriptional riboswitches was investigated. Removal of the riboswitches that control purine biosynthesis and riboflavin biosynthesis in *B. subtilis* led to auxotrophic strains. As an alternative, a rational approach was developed for engineering transcriptional riboswitches independently from the availability of their 3D structures. This approach consists in the identification and deletion of a key nucleotide sequence exclusively involved in transcription termination without affecting formation of other secondary and tertiary structures potentially involved in other roles. To demonstrate the efficacy of the approach, it was applied to derepress the purine and the riboflavin biosynthetic pathways in *B. subtilis*. Following the proof of concept using specialized reporter strains, the approach was implemented into a *B. subtilis* wild-type strain employing CRISPR-Cas genome editing. The CRISPR-Cas9 system displayed an efficiency of 61% in editing the genome, and the resulting purine and riboflavin production strains were characterized at the level of gene expression, metabolite synthesis, and growth. With a substantial enhancement observed at each level, the strategy established here represents a powerful tool for deregulating pathways modulated by transcriptional riboswitches. Finally, applying this strategy to derepress the purine pathway of an industrial riboflavin overproducing strain, with impaired growth, led to an increase in biomass by 53% and resulted in an enhanced total production of riboflavin in the culture.

Zusammenfassung

Viele bakterielle Stoffwechselwege werden von RNA-Schaltern (*riboswitches*) reguliert. Diese binden bestimmte Metabolite und regulieren die Expression der dazugehörigen Gene durch Elongation der Transkription oder Initiation der Translation. RNA-Schalter sind vielversprechende Zielstrukturen im Hinblick auf die Regulation der Expression bestimmter Gene und Transkriptionseinheiten, die für die biotechnologische Herstellung kommerziell relevanter Produkte von Bedeutung sind. Ungefähr 70% aller erforschten und vermuteten RNA-Schalter in den *Firmicutes* agieren ausschließlich auf transkriptionaler Ebene und basieren auf einem Termination-Antiterminations-Mechanismus.

In einem ersten Versuch die Purin-Biosynthese in *Bacillus subtilis* durch Manipulation Purin-regulierter RNA-Schalter zu deregulieren, wurden verschiedene synthetische kleine RNA-Moleküle (*small-RNAs*; sRNAs) eingesetzt. Diese sRNAs wurden so konstruiert, dass sie das Purin-sensitive Aptamer der RNA-Schalter binden und die Bindung des Purin-Liganden verhindern. Anders als erwartet zeigten die konstruierten sRNAs *in vivo* keine Aktivität im Hinblick auf die Expression der durch den Ziel-RNA-Schalter kontrollierten Purin-Biosynthesegene (*pur*-Operon). Darüber hinaus wurde die *anti-sense* RNA (asRNA), die komplementär zum Aptamer des pur-Operon

Riboswitch ist, auf ihre Wirkung hin untersucht. Diese asRNA hatte jedoch keinen negativen Einfluss auf die Expression der Ziel-mRNA. Außerdem hatte sie, vergleichbar mit den partiell komplementären sRNAs, keine Auswirkung auf die Regulation des RNA-Schalters. Schlussendlich wurde die Expression der sRNAs in *B. subtilis* quantifiziert. Die kinetische Limitierung bezüglich der Hybridisierung sRNA/Aptamer, sowie ihre kompetitiven Wechselwirkungen mit entsprechenden Liganden, werden diskutiert.

Die zweite Strategie zur Deregulierung transkriptionaler RNA-Schalter basiert auf der Editierung des Genoms von *B. subtilis*. Die Entfernung der (gesamten) RNA-Schalter, die die Purin- und Riboflavin-Biosynthesegene kontrollieren, erzeugte Purin- und Riboflavin-auxotrophe Stämme. Aus diesem Grund scheint die Entfernung von RNA-Schaltern keine gute Option zu sein, um diese Stoffwechselwege zu deregulieren. Als Alternative wurde eine Strategie entwickelt, mit der transkriptionale RNA-Schalter unabhängig von der Verfügbarkeit ihrer 3D-Struktur verändert werden können. Dieser Ansatz besteht aus der Identifizierung und Ausschaltung einer bestimmten Nukleotidsequenz, die für den Abbruch der Transkription verantwortlich ist. Dabei sollte jedoch die Bildung anderer relevanter Sekundär- und Tertiärstrukturen nicht beeinflusst werden, da diese potenziell noch andere Aufgaben haben. Um die Wirksamkeit dieser Strategie zu demonstrieren, wurde sie sowohl auf den Purin- als auch den Riboflavin-Biosyntheseweg in *B. subtilis* angewandt. Ziel war in beiden Systemen die negative Regulierung der jeweiligen RNA-Schalter zu unterdrücken. Diese Strategie wurde zuerst in einem speziellen Reporterstamm demonstriert. Danach wurde sie in einem *B. subtili*s Wildtypstamm implementiert, dessen Genom mit Hilfe von CRISPR-

Cas9 editiert wurde. Die Genomeditierung zeigte eine Effizienz von 61%. Die daraus resultierenden Purin- und Riboflavin-überproduzierenden Stämme wurden auf den Ebenen der Genexpression, Metabolitkonzentration und des Wachstum untersucht. Die hier verfolgte Strategie hat auf allen untersuchten Ebenen zu Steigerungen geführt und stellt deshalb eine vielversprechende Methode zur Deregulierung von Stoffwechselwegen dar, sofern diese durch RNA-Schalter gesteuert werden. Angewandt auf den Purin-Stoffwechselweg eines wachstumslimitierten, industriellen, Riboflavin-überproduzierenden Produktionsstamms führte diese Strategie sogar zu einer erhöhten Produktion (53%) von Biomasse bei gleichzeitiger Steigerung der Riboflavin-Produktion im Kulturüberstand.

Contents

List of Figures

Supplementary figures

List of Tables

1 Introduction

1.1 Bacterial riboswitches

A riboswitch is a *cis*-acting RNA element present in the 5'-untranslated region (5'-UTR) of some messenger RNAs (mRNAs). Riboswitches regulate expression of downstream gene(s) via their ability to specifically sense metabolites or inorganic ions and undergo a conformational change [1-3]. At present, about 40 different classes of riboswitches were identified, mainly in bacteria but they are also present in archaea, fungi, plants, and algae [4]. Bacterial riboswitches are involved in the regulation of several fundamental metabolic pathways such as the biosynthesis and transport of nucleotides, vitamins and amino acids [5]. In the Gram-positive bacterium *Bacillus subtilis*, a GRAS (generally recognized as safe) organism employed in a variety of biotechnological processes, at least 2% of the genes are under riboswitch control [5]. Comparative genomics analyses predict that a much higher number of riboswitches, whose ligands are undetermined, is present in bacteria [6, 7]. Hence, these RNA molecules represent promising engineering targets for the overproduction of small molecules employing specialized prokaryotic production strains.

1.2 Mechanisms of action of riboswitches

Riboswitches are RNA molecules containing an aptamer domain capable of specifically binding a ligand, commonly a small molecule related to the function of the regulated gene(s). The binding event induces a conformational change in the downstream structure and results in a switch from an ON to an OFF conformation or inversely, leading either to repression or to activation of downstream genes depending on the type of the riboswitch [8] (repressive or activating) (**Fig. 1a,b** or **1c,d** respectively). Thereby, the system works as a molecular switch that can turn expression OFF or ON in response to changes in metabolite levels. Riboswitch modulation of gene expression can be exerted at the transcriptional level (**Fig. 1a,c**) and/or at the translational level (**Fig. 1b,d**). In transcriptional riboswitches, the ligand-induced structural rearrangement of the nascent RNA leads to formation of a transcriptional terminator or antiterminator structure [9, 10], whereas in translational riboswitches, the structural rearrangement leads to formation of alternative secondary structures that modify the accessibility of the ribosomal binding site (RBS) [11].
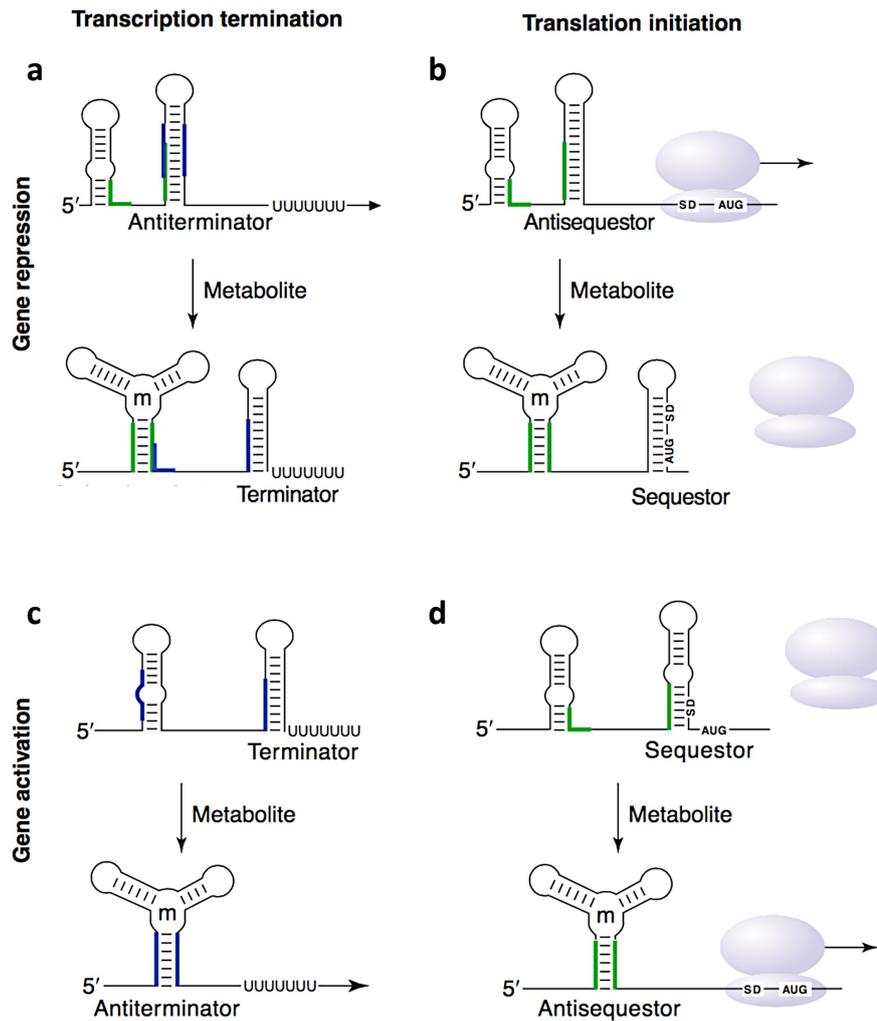
Figure 1 | Riboswitch-mediated control of gene expression. Bacterial riboswitches repress or activate their genes depending on the configuration of the corresponding leader RNA sequence (shown by the line growing from the 5' end). In response to the change of the metabolite concentration, they control transcription termination (a,c) or translation initiation (b,d) or both if the stem–loop structure of the terminator also serves as a sequestor of the ribosome-binding site (RBS). In each case, the binding of a specific metabolite (m) to the conserved RNA-sensor (aptamer) domain triggers a conformational change that leads to the formation of an alternative RNA structure that could be a terminator (a), sequestor (b), antiterminator (c), or antisequestor of RBS (d). The ribosome is shown in pale blue. The complementary RNA regions are indicated in green and blue. SD: Shine-Dalgarno sequence. AUG: start codon. Figure adapted from Nudler and Mironov (2004) [8].

The mode of action of transcriptional riboswitches being linked to the transcriptional process, researchers have proposed that their activity is governed by a kinetic regulation regime [12-16]. In this model, a functional interaction between the aptamer and the ligand is limited to the time that the RNA polymerase (RNAP) takes to progress between the end of aptamer synthesis and when it reaches the regulatory decision point to continue or abort transcription at the level of the terminator. This time frame was shown to be too short for the ligand-binding reaction to reach thermodynamic equilibrium. In translational riboswitches however, the time available for aptamer-ligand interaction and for mRNA interrogation by the ribosome lasts for the lifetime of the mRNA. Consequently, the activity of riboswitches that act on the translational level is governed by a thermodynamic regulation regime [15].

Approximately 70% of all putative and validated riboswitches in Firmicutes (are predicted to) act exclusively at the transcriptional level [11]. Despite differences in the structures and degrees of conformational change between riboswitches of different classes [17], the proposed transcription termination-antitermination mechanism described above represents a common mode of action in transcriptional riboswitches [9].

1.3 Regulation of purine and riboflavin biosyntheses in *Bacillus subtilis*

Metabolites derived from the riboswitch-controlled purine and riboflavin biosynthetic pathways, such as inosine, guanosine, riboflavin (vitamin B2), and folic acid (vitamin B9) are of considerable importance for a sustainable biomanufacturing industry, and the corresponding biotechnological production processes employing *B. subtilis* as a cell factory are well established and of broad interest [18].

Expression of the genes responsible for *de novo* biosynthesis of purine and riboflavin is regulated by repressive transcriptional riboswitches in *B. subtilis*. The purine biosynthesis operon *purEKBCSQLFMNHD* [5] is controlled by a guanine-sensing riboswitch (*pur operon* riboswitch) (**Fig. 2a**), and the riboflavin biosynthesis operon *ribDEAHT* [10] is controlled by a flavin mononucleotide (FMN)-sensing riboswitch (*rib operon* riboswitch) (**Fig. 2b**). Notably, expression of the *pur* operon is subjected to an additional negative feedback control at the DNA level mediated by the repressor PurR [19, 20], which binds to the so called PurBoxes upstream of the promoter P*pur* and affects initiation of transcription of the *pur* operon (**Fig. 2a**). Two PurBoxes are present, a proximal and a distal sequence (box) relatively to the promoter P*pur* [21, 22]. PurR binding to PurBoxes is inhibited by the purine biosynthesis substrate phosphoribosyl-α-1-pyrophosphate (PRPP) [20], whose synthesis is inhibited by adenosine diphosphate (ADP) and guanosine diphosphate (GDP) [23]. Therefore, the repressor PurR is functional when the levels of purine metabolites are high. The *rib* operon is also controlled by a regulatory protein

(RibR) (**Fig. 2b**). Synthesis of RibR is induced in the presence of organic sulfur sources such as methionine and taurine [24, 25]. RibR represents a superordinate regulator that keeps the *rib* operon riboswitch in an ON state even in the presence of high levels of FMN. Since the growth media used in this study do not contain high levels of methionine or taurine, RibR-mediated regulation is not relevant with regard to the present work.
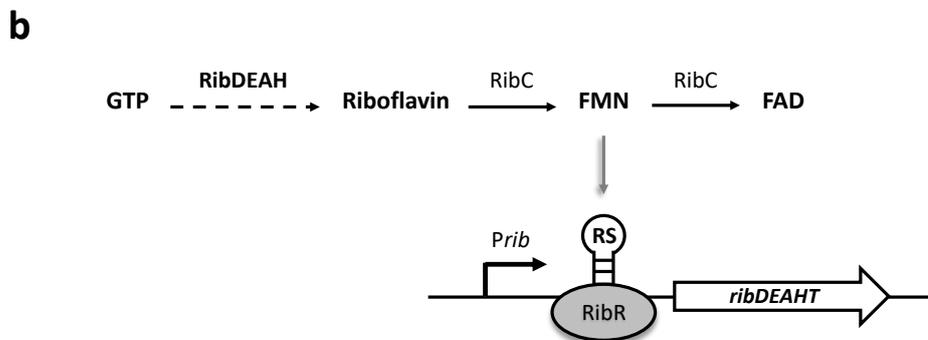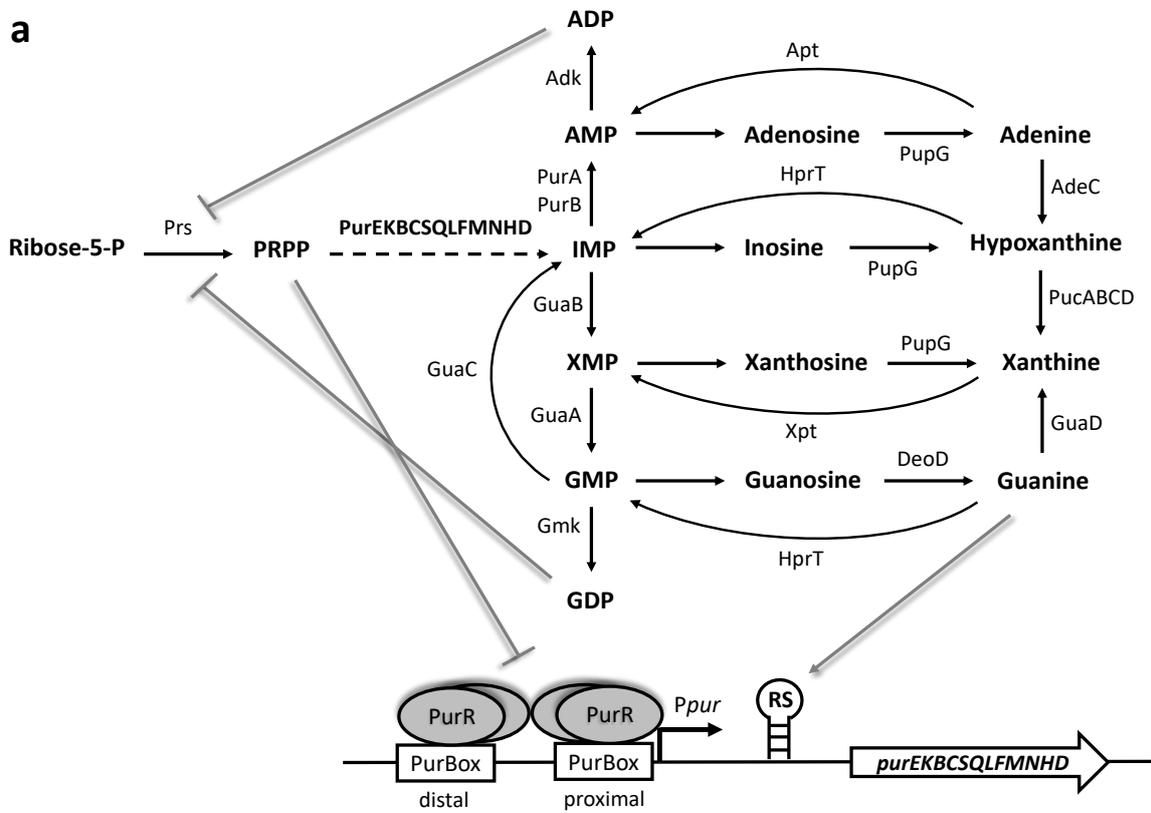
Figure 2 | Regulation of the *pur* operon responsible for *de novo* biosynthesis of purines (a) and the *rib* operon responsible for *de novo* biosynthesis of riboflavin (b) in *B. subtilis*.

(**a**) PurR repressor binding to the two PurBoxes is inhibited by PRPP, whose synthesis is repressed by ADP and GDP. High levels of guanine lead to the termination of transcription of the genes *purEKBCSQLFMNHD* mediated by a riboswitch (RS) (*pur* operon riboswitch). IMP inosine monophosphate, AMP adenosine monophosphate, XMP xanthosine monophosphate, GMP guanosine monophosphate, ADP adenosine diphosphate, GDP guanosine diphosphate, Prs phosphoribosylpyrophosphate synthetase, *PurEKBCSQLFMNHD* purine *de novo* biosynthesis enzymes, PurA adenylosuccinate synthetase, PurB adenylosuccinate lyase, GuaB IMP dehydrogenase, GuaA GMP synthetase, GuaC GMP reductase, Adk adenylate kinase, Gmk guanylate kinase, PupG purine nucleoside phosphorylase, DeoD purine nucleoside phosphorylase, HprT hypoxanthine phosphoribosyltransferase, Xpt xanthine phosphoribosyltransferase, Apt adenine phosphoribosyltransferase, GuaD guanine deaminase, AdeC adenine deaminase, PucABCD xanthine dehydrogenases. P*pur* promoter of the *pur* operon.

(**b**) High levels of FMN lead to the termination of transcription of the genes *ribDEAHT* mediated by a riboswitch (RS) (*rib* operon riboswitch). An additional control is provided by RibR, which locks the riboswitch ON in the presence of methionine and taurine. FMN flavin mononucleotide, FAD flavin adenine dinucleotide, RibDEAH enzymes of riboflavin *de novo* biosynthesis, RibC bifunctional flavokinase/FAD synthetase, P*rib* promoter of the *rib* operon.

1.4   A riboswitch-based regulon controls purine metabolism in *Bacillus subtilis*

In addition to the *pur* operon riboswitch regulating expression of the purine biosynthetic genes, four additional purine-sensing transcriptional riboswitches with conserved aptamer domains were identified in *B. subtilis* [26]. They modulate transcription of genes involved in purine salvage (the *xpt* gene encoding a phosphoribosyltransferase), uptake (*pbuX*, *pbuG*, and *nupG* genes), and export (*pbuE* gene) (**Fig. 3**). The *xpt* and *pbuX* genes are organized in a riboswitch-regulated transcriptional unit forming the *xpt-pbuX* operon and the remaining *pbuG*, *nupG*, and *pbuE* genes are present as single genes and regulated by distinct riboswitches. Similarly to the *pur* operon riboswitch, the *xpt-pbuX*, *pbuG*, and *nupG* riboswitches are guanine-sensing repressive riboswitches, whereas the *pbuE* riboswitch is an adenine-sensing activating riboswitch [27]. Under conditions of high intracellular concentration of purine metabolites, the riboswitch-mediated regulation leads to a simultaneous repression of biosynthesis, salvage, and uptake genes while the exporter gene is concurrently activated. As a result of such changes, purine concentration decreases within the cytoplasm [28]. In addition to binding guanine, the *xpt-pbuX* riboswitch was shown to also bind *in vitro* xanthine and hypoxanthine, yet with a lower affinity (dissociation constant lower than 5 nM for guanine and approximately 50 nM for xanthine and hypoxanthine) [5].
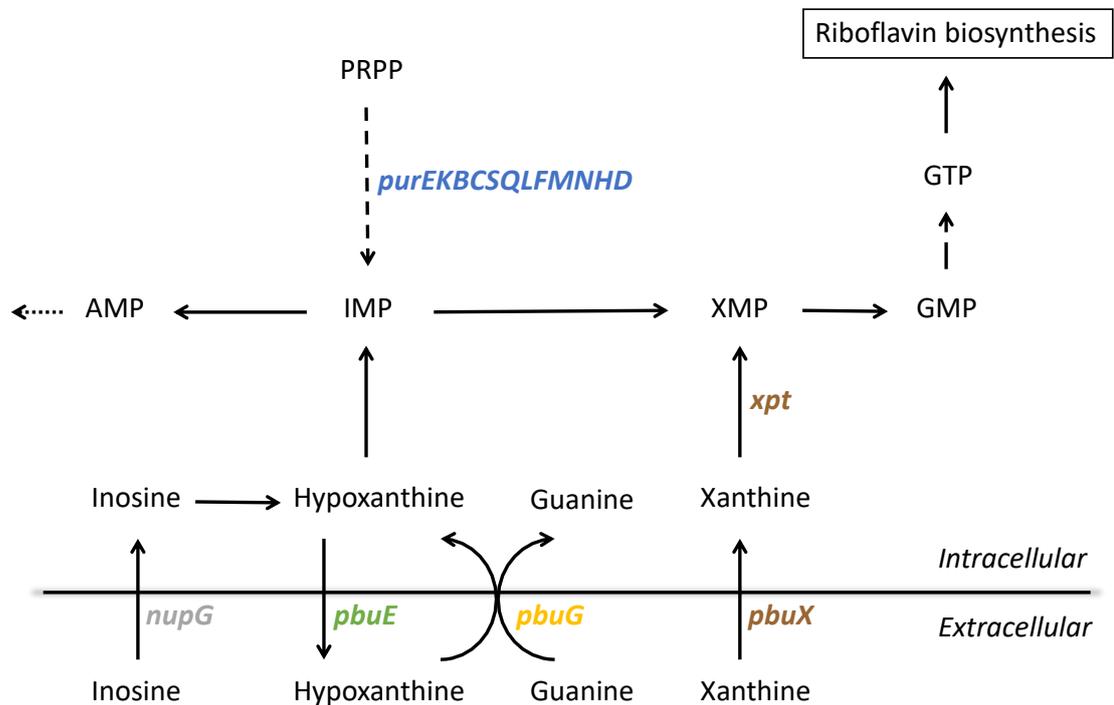
Figure 3 | Riboswitch-mediated control of purine metabolism in *B. subtilis*. The genes involved in purine biosynthesis (*purEKBCSQLFMNHD*), salvage (*xpt*), and uptake (*pbuX*, *pbuG*, and *nupG*) are negatively regulated by a guanine-binding riboswitch, whereas the *pbuE* gene encoding a hypoxanthine exporter is positively regulated by an adenine-binding riboswitch. Each gene(s) color represents a purine riboswitch-controlled transcriptional unit. As the purine nucleotide GTP is a precursor of riboflavin biosynthesis, the purine and the riboflavin metabolic pathways are directly connected. For clarity, only relevant purine metabolites and interconversion reactions are shown in the figure. IMP inosine monophosphate, AMP adenosine monophosphate, XMP xanthosine monophosphate, GMP guanosine monophosphate, GTP guanosine triphosphate.

1.5 Challenges in derepressing riboswitch-controlled genes

1.5.1 Riboswitch and terminator deletions

Deletion of the *rib* operon riboswitch in a riboflavin overproducing *B. subtilis* strain unexpectedly led to reduced transcript levels and lower production of riboflavin [29]. Furthermore, conversely to what was found in a previous experiment employing reporter gene fusions [9], deletion of the chromosomal sequence that forms the termination hairpin of the *rib* operon riboswitch, and that is partly involved in formation of the antiterminator structure, resulted in a decrease in transcript level and in riboflavin production similarly to the deletion of the full-length riboswitch [30]. This can be explained by the fact that beside their known metabolite-sensing function, riboswitches can be involved in other processes. Indeed, their 5' end secondary and tertiary structures can play a role in transcript stability [31-34], interaction with transcription factors [24, 25], or regulation in *trans* of the expression of other messenger RNAs [35]. These considerations raise the challenge of preserving a protective and/or functional structure when engineering a riboswitch to derepress a certain pathway.

1.5.2 Random and targeted aptamer mutagenesis

Classical random mutants selection led to the identification of deregulated *B. subtilis* riboflavin production strains carrying mutations in the aptamer of the *rib* operon riboswitch [36]. This approach for identifying mutations that derepress riboswitch-controlled genes requires the availability of growth

reducing antimetabolites for selection. In addition, a very high number of mutants has to be analyzed in a time-consuming process. Targeted rational mutagenesis of the aptamer moiety could successfully disable a specific riboswitch from binding its ligand [5]. However, it can be tedious to concurrently preserve a functional fold, and the strategy requires the availability of detailed structural information about the riboswitch (i.e., its crystal structure).

1.6 Small regulatory RNAs in bacteria

Small regulatory RNAs (sRNAs) are short noncoding RNAs found in many bacteria to regulate gene expression post-transcriptionally by sequence-specific base pairing with target mRNAs [37]. Similarly to regulatory RNAs in eukaryotes, bacterial sRNAs can be perfectly complementary to the target mRNA or can act with partial complementarity [38]. In bacteria, sRNAs act typically by pairing with 5' ends of target mRNAs around the RBS region, modulating initiation of translation [39-41] and/or by directing the stability of target transcripts [42-44].

Perfectly complementary antisense RNAs (asRNAs) are also referred to as *cis*-encoded sRNAs as they are naturally encoded on the complementary strand of their target gene within the same genomic locus [45-48]. The typical function of characterized bacterial asRNAs is to repress expression of toxic proteins [49, 50] but few cases were reported where asRNAs base pairing stabilizes the target mRNA leading to increased expression [51].

Partially complementary small RNAs are also referred to as *trans*-encoded

sRNAs as they are usually located on a distinct genomic locus from their target gene. The partial complementarity with mRNA targets confers to this sRNA type the flexibility to often target multiple transcripts simultaneously [40]. The base pairing is initiated at the level of a seed sequence made of a minimum of six to eight perfectly complementary nucleotides. sRNAs with partial complementarity are generally expressed under specific growth conditions and are involved in many cellular processes in bacteria, where they act either as positive or as negative regulators of mRNA translation. In some cases, a sRNA targeting multiple mRNAs can have both positive and negative regulation effects depending on the mRNA target [52, 53]. The RNA chaperone Hfq is required for the activity of partially complementary sRNAs in proteobacterial species including *Escherichia coli*, *Salmonella enterica*, and *Vibrio cholerae* [54-56]. However, in Gram-positive bacteria such as *Staphylococcus aureus* and *Bacillus subtilis*, sRNA activity is independent from Hfq [57-60].

Synthetic small regulatory RNAs were successfully used in bacteria to inhibit translation by targeting mRNAs at the level of the Shine-Dalgarno (SD) sequence, preventing thereby ribosome binding [61]. Inversely, they were also used as translation activators in cases where the SD sequence is sequestered by a *cis* complementary sequence as it is the case in mRNAs regulated by translational riboswitches. The binding of the synthetic sRNA to the complementary sequence in that case unmasks the RBS and promotes translation [62].

1.7 The CRISPR-Cas9 system

The clustered regularly interspaced short palindromic repeats (CRISPR) associated (Cas) systems are natural bacterial defense systems against invasive genetic elements [63]. The mechanism of action of the most studied CRISPR-Cas system, the type II CRISPR-Cas9 [64], consists in the recognition and the processing of foreign DNA from which twenty nucleotide fragments (protospacers) juxtaposing a specific three nucleotides sequence, termed protospacer adjacent motif (PAM), get integrated into the CRISPR array locus in the genome of the bacterial host [65]. These integrated fragments are then transcribed as parts of RNAs termed CRISPR RNAs (crRNAs). crRNAs are recognized by a protein-RNA complex composed of the CRISPR-associated nuclease Cas9 and the *trans*-activating CRISPR RNA (tracrRNA), which contain a sequence complementary to the crRNA [66]. Thereby, the crRNA serves as a guide to the complex for recognition and specific double strand cleavage of the foreign target DNA by Cas9 [67, 68] (**Fig. 4**).

Figure 4 | Clustered regularly interspaced short palindromic repeats (CRISPR)−CRISPR-associated 9 (Cas9) targeting of DNA. Cas9 interrogates DNA and binds reversibly to protospacer adjacent motif (PAM) sequences with stabilization of Cas9 at the target occurring via formation of the *trans*-activating crRNA (tracrRNA)::CRISPR RNA (crRNA) duplex. Mature crRNA anneals to the target with base-pairing complementarity and tracrRNA functional modules govern Cas9 activity and define orthogonality. Activation of Cas9 causes simultaneous cleavage of each strand by the RuvC and HNH domains, represented here as black wedges. Figure adapted from Selle and Barrangou (2015) [69].

The system was adapted and increasingly used during the last decade for editing the genome of numerous eukaryotic cell lines by expressing Cas9 and the CRISPR RNAs in target cells, together with the introduction of a DNA homology repair template containing the desired genetic modification [70, 71]. Instead of expressing the crRNA and the tracrRNA separately, the system was simplified by the fusion of the two RNAs into a chimeric single guide RNA (gRNA) containing the scaffold that interacts with Cas9 and the twenty nucleotides spacer that guides the complex and binds to the target chromosomal sequence [72]. The specific double strand break caused by Cas9 in the chromosome leads to the integration of the DNA homology repair template into the corresponding locus by homologous recombination.

The system was also applied to the genome editing of some bacterial species including *Escherichia coli*, *Streptococcus pneumoniae* [73], and *Lactobacillus reuteri* [74], allowing markerless and precise genetic modifications with no prerequisites such as specific mutations or counterselection genes as it is the case using classical selection-counterselection tools [75]. The type II CRISPR-Cas9 system from *Streptococcus pyogenes* was recently adapted for the genetic engineering of the *B. subtilis* chromosome with editing efficiencies ranging from 33% up to 100% depending on the technic used, the strain, the targeted locus and the type of genetic modification [76].

1.8 Objective of the work

The present work was initiated to invent a versatile method for impairing the regulation exerted by transcriptional riboswitches on metabolic pathways. In the first chapter, the possibility of using synthetic small RNAs to prevent ligand binding to aptamer domains and thereby impair riboswitch regulation was evaluated. The strategy was investigated employing small RNAs generated by rational design combined with *in silico* evolution to target purine riboswitches in *B.* subtilis. In the second chapter, the physiological effects of riboswitch deletion were studied, and a rational approach for editing transcriptional riboswitches to activate gene expression was developed. The approach was applied to the metabolic engineering of the purine and the riboflavin pathways in *B. subtilis* employing the CRISPR-Cas9 technology.

2 Materials

2.1 Lab equipment

Table 1 | List of used lab equipment for standard procedures.

Device	Model	Manufacturer
Autoclave	Varioklav® Dampfsterilisator 75S	H+P Labortechnik, Habermoos, Germany
Bench centrifuge	Eppendorf 5415 R	Eppendorf, Hamburg, Germany
Centrifuge (30k rpm)	Sorvall WX Ultra	Thermo Fischer, Waltham, USA
Centrifuge (6k rpm)	Heraeus Fresco 21	Heraeus, Hanau, Germany
Dialysis membrane	Tubing-Visking size 11	Medicell Int. LTD, London, UK
Electroporation device	Gene Pulser® II	BIO-RAD, Munich, Germany
French Press	Constant Cell Disruption System	Constant System, Daventry, UK
Gel electrophoresis (horizontal)	Mini-Sub Cell GT + Wide, Power Pac 200, Power Pac 300	BIO-RAD, Munich, Germany

Device	Model	Manufacturer
Gel Imager	Molecular Imager® GelDoc™ XR, Quantity One 1D-Analysis software	BIO-RAD, Munich, Germany
Heating block	Thermomixer comfort	Eppendorf, Hamburg, Germany
Incubator	15 L NLF22	Bioengineering, Wald, Switzerland
LC/MS-system	1200 Infinity series with DAD/FLD and API-ESI 6130 Quadrupole	Agilent Technologies, Waldbronn, Germany
Optical microscope	BH-2	Olympus Europe, Hamburg, Germany
PCR	C1,000™ Thermal Cycler, dual	BIO-RAD, Munich, Germany
Photometer DNA quantification	NanoVue	GE Healthcare, Little Chalfont, UK
Photometer	Uvikon 933	BioTek Kontron Instruments AG, Neufahrn
Shaking incubator	Certomat® IS	Sartorius Stedim GmbH, Göttingen, Germany
Sterile bench	Variolab Mobilien W 90	Waldner Laboreinrichtungen GmbH & Co KG, Wangen, Germany
Water bath	IKA® HBR4 digital	IKA GmbH, Staufen im Breisgau, Germany

2.2 Enzymes and chemicals

Enzymes (Phusion Hot Start II DNA polymerase used for PCR amplification, FastDigest restriction enzymes, FastAP Thermosensitive Alkaline Phosphatase, T4 Polynucleotide Kinase and T4 DNA ligase), kits for DNA extraction and purification, and oligonucleotides and DNA fragments were purchased from Thermo Fisher Scientific (Darmstadt, Germany), unless otherwise stated. Kits for RNA extraction and purification were purchased from Zymo Research (Freiburg, Germany). Chemicals were purchased from Sigma Aldrich (Taufkirchen, Germany), unless otherwise stated. Xanthine, adenine, guanosine, riboflavin, FMN, and FAD were dissolved in dimethyl sulfoxide (DMSO). The antibiotics ampicillin (Amp), kanamycin (Kan), and spectinomycin (Spc) were dissolved in distilled water, and chloramphenicol (Cm) was dissolved in 95% ethanol. Isopropyl β-D-1-thiogalactopyranoside (IPTG) and mannose were dissolved in distilled water.

2.3 Growth media

Solutions were prepared in ultrapure water and media were heat sterilized in an autoclave (Varioklav R 75S, H+P Labortechnik, Habermoos) before use.

Table 2 | Lysogeny Broth (LB), pH 7.2 [77].

Component	Final Concentration
Tryptone	10 g L^{-1}
Yeast extract	5 g L^{-1}
NaCl	10 g L^{-1}

Table 3 | 5x C salts stock solution for C minimal medium [78].

Component	Amount
KH_2PO_4	20 g
K_2HPO_4, 3 H_2O	80 g
$(NH_4)_2SO_4$	16.5 g
H_2O	1000 ml

Table 4 | III' salts stock solution for C minimal medium [78].

Component	Amount
$MnSO_4$, 4 H_2O	232 mg
$MgSO_4$, 7 H_2O	12.3 g
$ZnCl_2$ solution (100 mM)	12.5 ml
H_2O	1000 ml

Table 5 | C minimal medium (CMM) [78], 3% glucose.

Component	Amount
5x C salts solution (Table 3)	20 ml
Tryptophan (5 mg/ml)	1 ml
Iron Ferric Ammonium Citrate (2.2 mg/ml)[a]	1 ml
III' salts solution (Table 4)	1 ml
Glucose	3 g
H_2O	100 ml

[a]*Sensitive to light*

2.4 Buffers

Solutions were prepared in ultrapure water. For RNA analysis, RNase-free water was prepared by the addition of 0.1% v/v DEPC, stirring for a minimum of 2 hours, and heat inactivation in the autoclave.

Table 6 | 10x MOPS buffer, pH 7.

Component	Amount	Final concentration
MOPS	41.9 g	200 mM
Sodium acetate anhydrous	2.9 g	50 mM
EDTA, 3 Na, 3 H_2O	3.94 g	10 mM
RNase-free H_2O	1000 ml	

Table 7 | 20x SSC buffer, pH 7.

Component	Amount	Final concentration
NaCl	175.3 g	3 M
Trisodium citrate dihydrate	88.2 g	300 mM
RNase-free H_2O	1000 ml	

Table 8 | Maleic acid buffer, pH 7.5.

Component	Amount	Final concentration
Maleic acid	11.6 g	100 mM
NaCl	8.8 g	150 mM
NaOH	7.9 g	200 mM
RNase-free H_2O	1000 ml	

Table 9 | Detection buffer, pH 9.5.

Component	Amount	Final concentration
Tris-HCl	7.9 g	100 mM
NaCl	2.9 g	100 mM
RNase-free H_2O	500 ml	

2.5 Oligonucleotides

Table 10 | Oligonucleotides used in this study. Restriction endonuclease sites are underlined.

Name	Sequence (5'-3')
S01	CACTG<u>GCATGCGGTCTC</u>GTGGAAGTTACTGACGTAAGAT
S02	ATATCTAAGTT<u>GCGGCCGC</u>TGGTCTGATCGATGGGATGT
S03	CCGGTTTGTATTGCTTCCTCATAAGTTTTTTATTTTCTGAAAACAAAAGC
S04	GCTTTTGTTTTCAGAAAATAAAAAACTTATGAGGAAGCAATACAAACCGG
S05	AGTTCATGAAGTTTCGTCGCAGCGGAAAGACAATCTTTTAAAGAATGGAG
S06	TAATCTAGAAAGGCCTTATTGGCCCGAATATCATGTACAACAAATTCGTG
S07	ATAGGGTCGACGGCCAACGAGGCCCACCAAAGAAACTTAGCTTATCAGG
S08	CCGCTGCGACGAAACTTCATGAACTCACCCCCAAAATCCG
S09	TTTTTATTTTCTGAAAACAAAAGCATTAGAAGGTGGGGAAC
S10	TAATCTAGAAAGGCCTTATTGGCCCTATAATGAGCGGTAATTTCGACATC
S11	ATAGGGTCGACGGCCAACGAGGCCCTATTATGGTCAATGTGATTACGAAG
S12	CCTTCTAATGCTTTTGTTTTCAGAAAATAAAAAACTTATGAGGAAGCAATACAAACCG
S13	CAG<u>GTCGAC</u>TATTATGGTCAATGTGATTACGAAG
S14	GCC<u>TCTAGA</u>TATAATGAGCGGTAATTTCGACATC
S15	AGT<u>GGTCTC</u>TTTTTTATTTTCTGAAAACAAAAGCATTAGAAGG
S16	GCC<u>TCTAGA</u>TATAATGAGCGGTAATTTCGACATC

Name	Sequence (5'-3')

S17   ATGGTCGACTGATTTTAACGCTGAAAGGACAG

S18   GCTGGTCTCTAAAAGTTTATCTTAACAACGGACATGGA

S19   GCTGGTCTCATTTGACGGTAAATAACAAAAGAGGGGA

S20   TATCTAGACGTCCTCAGGAATCGAGAGTACGGT

S21   ACTGTCGACGAATAGATTCATATTGGCTGGAGGT

S22   GCTGGTCTCTCAAAATAAAAAATTCGGGGCTTTAGGT

S23   ATGGTCGACGGTTAAAGTGCCTGATAACAAG

S24   ACTGGTCTCACAAAATTGGTTATTATAGCGATCCT

S25   TACGCTTGCAGCTGCGTACGCAAG

S26   AAACCTTGCGTACGCAGCTGCAAG

S27   TACGCTCATAAGTGCAATGCAGAG

S28   AAACCTCTGCATTGCACTTATGAG

S29   TACGCCCGAATTTTTTATAAATTC

S30   AAACGAATTTATAAAAAATTCGGG

S31   TCCTCCCACAACCGATTTGTGC

S32   TGCGTTTACTCGGCATCGCAAC

S33   GCAAACCATCTTATCAAATGGGAT

S34   AACTCATATGTGATGATATCGCTGA

S35   GGTTTTGACCATTTGACCGATGATT

S36   AATCCATGTCGGCGCTATTTGATC

S37   GCGTTAAATATATTCCGAAAATGAA

Name	Sequence (5'-3')
S38	AGCGACTGTCCAAGTGTCTG
S39	GCATTTGACGAAGACCTTGC
S40	CTGCTTTCTAACACTGTCTGTTTTG
S41	GGGTGCCGACATTTACGTT
S42	GCACATGGCGGTGTTTTT
S43	GAAAGCCACGGCTAACTACG
S44	GACAACGCTTGCCACCTAC
S45	DIG-CGTGCGGTTCCATTGCTCACCCATAGTCGG-DIG
S46	DIG-CAGCTGGCGAAAGGGGGATGTGCTGCAAGG-DIG
S47	DIG-TCTAGAATTGTTATCCGCTCACAATTCCAC-DIG

2.6 Plasmids

Table 11 | Plasmids used in this study.

Name	Characteristics	Source or reference
sRNA expression plasmids		
pIDTSMART	Amp^R (*E. coli*)	Integrated DNA Technologies
pT7sRNA$_{control}$	Amp^R (*E. coli*), P_{T7} sRNA$_{control}$	This work
pT7sRNA30	Amp^R (*E. coli*), P_{T7} sRNA30	This work
pT7sRNA80	Amp^R (*E. coli*), P_{T7} sRNA80	This work
pT7sRNA110	Amp^R (*E. coli*), P_{T7} sRNA110	This work
pT7sRNA40$_{xpt}$	Amp^R (*E. coli*), P_{T7} sRNA40$_{xpt}$	This work
pHT01	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac	[79]
pHTsRNA$_{random}$	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA$_{random}$	This work
pHTsRNA30	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA30	This work
pHTsRNA70	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA70	This work
pHTsRNA110	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA110	This work
pHTsRNA30$_{pur}$	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA30$_{pur}$	This work

Name	Characteristics	Source or reference
pHTsRNA71$_{pur}$	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac sRNA71$_{pur}$	This work
pHTasRNApur	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac asRNApur	This work
pHTasRNAnupG	Amp^R (*E. coli*), Cm^R (*B. subtilis*), Pgrac asRNAnupG	This work

In vitro transcription plasmids

Name	Characteristics	Source or reference
pPrib-luc	Amp^R (*E. coli*), P*ribB luc*	[25]
pPrib-RSxpt-luc	Amp^R (*E. coli*), P*ribB* RS*xpt-pbuX luc*	This work
pMA	Amp^R (*E. coli*)	Thermo Fisher Scientific
pPtac-RSxpt-luc'	Amp^R (*E. coli*), PtacI RS*xpt-pbuX luc'*	This work

Integrative plasmids for the generation of reporter strains

Name	Characteristics	Source or reference
pDG268	Amp^R (*E. coli*), 3'-*amyE*, Cm^R (*B. subtilis*), P*spoIIG lacZ*, 5'-*amyE*	[80]
pSG1729	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), P*xyl gfpmut1'*, 5'-*amyE*	[81]
pBGAB	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *lacZ*, 5'-*amyE*	This work
pBGABpur	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *purE'–lacZ*, 5'-*amyE*	This work

Name	Characteristics	Source or reference
pBGABxpt	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *xpt'–lacZ*, 5'-*amyE*	This work
pBGABnupG	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *nupG'–lacZ*, 5'-*amyE*	This work
pBGABpbuG	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *pbuG'–lacZ*, 5'-*amyE*	This work
pBGABpbuE	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *pbuE'–lacZ*, 5'-*amyE*	This work
pBGAB11	Amp^R (*E. coli*), 3'-*amyE*, Spc^R (*B. subtilis*), *purE'*(mPB)–*lacZ*, 5'-*amyE*	This work
pBGAB12	Amp^R (*E. coli*), 3'-*amyE,* Spc^R (*B. subtilis*), *purE'*(mPB & ΔRS)–*lacZ*, 5'-*amyE*	This work
pBGAB13	Amp^R (*E. coli*), 3'-*amyE,* Spc^R (*B. subtilis*), *purE'*(mPB & eRS)–*lacZ*, 5'-*amyE*	This work

CRISPR-Cas9 plasmids

Name	Characteristics	Source or reference
pJOE8999	Kan^R (*E. coli* & *B. subtilis*), PE194^ts, P*manP cas9*, P*vanP lacPOZ'-*'gRNA	[82]
pPurRHT	Kan^R (*E. coli* & *B. subtilis*), PE194^ts, P*manP cas9*, Δ*purR* homology template, P*vanP lacPOZ'-*'gRNA	This work
peRSpurHT	Kan^R (*E. coli* & *B. subtilis*), PE194^ts, P*manP cas9*, eRS*pur* homology template, P*vanP lacPOZ'-*'gRNA	This work

Name	Characteristics	Source or reference
peRSribHT	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, eRS*rib* homology template, P*vanP lacPOZ'-*'gRNA	This work
pABpurR	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, Δ*purR* homology template, P*vanP* sgRNA$_{\Delta purR}$	This work
pABeRSpur	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, eRS*pur* homology template, P*vanP* sgRNA$_{eRS pur}$	This work
pABdRSpur	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, ΔRS*pur* homology template, P*vanP* sgRNA$_{\Delta RS pur}$	This work
pABeRSrib	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, eRS*rib* homology template, P*vanP* sgRNA$_{eRS rib}$	This work
pABdRSrib	Kan^R (*E. coli* & *B. subtilis*), PE194^{ts}, P*manP cas9*, ΔRS*rib* homology template, P*vanP* sgRNA$_{\Delta RS rib}$	This work

2.7   Bacterial strains

Table 12 | Bacterial strains used in this study.

Name	Genotype	Source or construction
Escherichia coli		
DH5α	F⁻ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *rec*A1 *end*A1 *hsd*R17(r$_K^-$, m$_K^+$) *pho*A *sup*E44 λ⁻ *thi*-1 *gyr*A96 *rel*A1	Invitrogen
XL10-Gold	Tet^{r}Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* Hte [F′ *pro*AB *lacI^q*ZΔM15 Tn10 (Tet^r) Amy Cam^r]	Agilent
NM538	*supF hsdR trpR lacY*	J. Altenbuchner, University of Stuttgart
Bacillus subtilis		
168	*trpC2*	Bacillus Genetic Stock Center
BGABpur	*trpC2 amyE*::[pBGABpur *purE'–lacZ spc*]	pBGABpur into 168
BGABxpt	*trpC2 amyE*::[pBGABxpt *xpt'–lacZ spc*]	pBGABxpt into 168

Name	Genotype	Source or construction
BGABnupG	*trpC2 amyE*::[pBGABnupG *nupG'–lacZ spc*]	pBGABnupG into 168
BGABpbuG	*trpC2 amyE*::[pBGABpbuG *pbuG'–lacZ spc*]	pBGABpbuG into 168
BGABpbuE	*trpC2 amyE*::[pBGABpbuE *pbuE'–lacZ spc*]	pBGABpbuE into 168
BGAB11	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*]	pBGAB11 into 168
BGABempty	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHT01	pHT01 into BGAB11
BGABsRNA$_{random}$	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA$_{random}$	pHTsRNA$_{random}$ into BGAB11
BGABsRNA30	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA30	pHTsRNA30 into BGAB11
BGABsRNA70	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA70	pHTsRNA70 into BGAB11
BGABsRNA110	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA110	pHTsRNA110 into BGAB11
BGABsRNA30$_{pur}$	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA30$_{pur}$	pHTsRNA30$_{pur}$ into BGAB11
BGABsRNA71$_{pur}$	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)*–lacZ spc*] pHTsRNA71$_{pur}$	pHTsRNA71$_{pur}$ into BGAB11

Name	Genotype	Source or construction
BGABasRNApur	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)–*lacZ spc*] pHTasRNApur	pHTasRNApur into BGAB11
BGABasRNAnupG	*trpC2 amyE*::[pBGAB11 *purE'*(mPB)–*lacZ spc*] pHTasRNAnupG	pHTasRNAnupG into BGAB11
168sRNArandom	*trpC2* pHTsRNArandom	pHTsRNArandom into 168
BGAB12	*trpC2 amyE*::[pBGAB12 *purE'*(mPB ΔRS)–*lacZ spc*]	pBGAB12 into 168
BGAB13	*trpC2 amyE*::[pBGAB13 *purE'*(mPB eRS)–*lacZ spc*]	pBGAB13 into 168
BSpurR	*trpC2 ΔpurR*	pABpurR into 168
eBSpur2	*trpC2 ΔpurR* eRS*pur*	pABeRSpur into BSpurR
BSdRSp	*trpC2 ΔpurR* ΔRS*pur*	pABdRSpur into BSpurR
BGAB11ΔpurR	*trpC2 ΔpurR amyE*::[pBGAB11 *purE'*(mPB)–*lacZ spc*]	pBGAB11 into BSpurR
BGAB13ΔpurR	*trpC2 ΔpurR amyE*::[pBGAB13 *purE'*(mPB eRS)–*lacZ spc*]	pBGAB13 into BSpurR
eBSflv	*trpC2* eRS*rib*	pABeRSrib into 168
BSdRSr	*trpC2* ΔRS*rib*	pABdRSrib into 168
BGBSpurR	*trpC2 ΔpurR amyE*::[pBGAB11 *purE'*(mPB)–*lacZ spc*]	pBGAB11 into BSpurR

Name	Genotype	Source or construction
BGeBSpur2	*trpC2 ΔpurR* eRS*pur* *amyE*::[pBGAB11 *purE′*(mPB)–*lacZ* *spc*]	pBGAB11 into eBSpur2
BSRF	P*rib*->Pveg RS*rib** *ribC** *tkt**	M. Mack, Hochschule Mannheim
eBSRFpur2	P*rib*->Pveg RS*rib** *ribC** *tkt** *ΔpurR* eRS*pur*	pABpurR, pABeRSpur into BSRF

3 Methods

3.1 Bacterial growth conditions

Bacterial strains used in this study are listed in Table 12.

For plasmid propagation, *E. coli* was aerobically grown in lysogeny broth (LB) supplemented with 100 µg/ml ampicillin or 50 µg/ml kanamycin depending on the plasmid. *B. subtilis* 168 and derivatives were aerobically cultivated in C minimal medium supplemented with 50 µg/ml tryptophan and 3% glucose (Table 5), unless otherwise stated. When required, antibiotics were added to a final concentration of 100 µg/ml for spectinomycin, 5 µg/ml for chloramphenicol, or 50 µg/ml for kanamycin. For induction of sRNA expression from pHTsRNAx plasmids, IPTG was added at inoculation to a final concentration of 1 mM. For activation of the *pbuE* riboswitch-controlled *lacZ* in the strain BGABpbuE, the medium was additionally supplemented with 3.7 mM adenine. For validation of the riboswitch engineering approach in reporter strains, the medium was additionally supplemented with 1 mM guanosine. To allow growth of auxotrophic strains, the medium was supplemented either with 500 µM guanosine and 500 µM adenosine for purine auxotrophic strains, or 20 µM riboflavin for riboflavin auxotrophic strains. Preparation of *B. subtilis* competent cells and transformation of *B. subtilis* were performed according to

the standard procedures described in Molecular Biological Methods for Bacillus [83]. For β-galactosidase assays, *B. subtilis* cells were cultivated to the mid-exponential phase of growth, unless otherwise stated. For RNA extractions and riboflavin/FMN/FAD analyses, *B. subtilis* cells were cultivated to the early stationary growth phase.

3.2 Cell lysis

Cells were lysed employing a FastPrep-24T M 5G Instrument (MP Biomedicals, Santa Ana, CA, USA) using 0.2-0.3 µm glass beads and the following settings: 8 cycles of 30 sec at 6.5 m/s with 2 min pause on ice after each cycle.

3.3 RNA extraction

Total RNA was extracted from *B. subtilis* strains using the Quick-RNA Miniprep Plus Kit (Zymo Research) with in-column DNase I treatment following the manufacturer's protocol. Extracted RNA was quantified using NanoVue Plus Spectrophotometer (GE Healthcare, Freiburg, Germany) and stored at -80 °C.

3.4 Construction of plasmids

The sequences of oligonucleotides used in this work for the construction of plasmids, proof sequencing, and colony PCR are listed in Table 10, and the plasmids used in this study are listed in Table 11.

3.4.1 sRNA expression plasmids

For the expression *in vitro* of synthetic sRNAs (sequences in Fig. 6, Fig. 7, and Fig. 8), DNA fragments corresponding to each sRNA directed upstream by the T7 promoter and followed by the T7 terminator sequences [84] were synthesized and inserted into pIDTSMART by Integrated DNA technologies (IDT), resulting into the set of plasmids pT7sRNAx. For sRNA expression in *B. subtilis,* the DNA fragments corresponding to each sRNA were digested from pT7sRNAx and inserted into a predigested pHT01 *B. subtilis/E.coli* shuttle vector using the restriction endonuclease pair XbaI/AatII, generating the set of plasmids pHTsRNAx.

3.4.2 *In vitro* transcription plasmids

For the *in vitro* transcription assay to study the activity of the *xpt-pbuX* riboswitch, a DNA fragment coding for the riboswitch sequence and flanked by HindIII and BamHI restriction sites was synthesized (due to the difficulty of PCR-amplifying riboswitch sequences that contain highly strong secondary structures), digested with the corresponding endonucleases, and inserted into a predigested pPrib-luc plasmid, resulting into pPrib-RSxpt-luc. The improved plasmid pPtac-RSxpt-luc' was generated by the synthesis of the DNA fragment corresponding to "PtacI−*xpt-pbuX* riboswitch−truncated *luc"* and its insertion into the GeneArt pMA plasmid by Invitrogen (Thermo Fisher Scientific).

3.4.3 Integrative plasmids for the generation of reporter strains

For the construction of *B. subtilis* integration plasmids that generated β-galactosidase reporter strains, the *lacZ* gene was PCR-amplified using the

primer pair S01/S02 and pDG268 as a template. The resulting DNA fragment was introduced into the *B. subtilis*/*E.coli* shuttle vector pSG1729 replacing a fragment containing the P*xyl* promoter and the gene *gfp*mut1' by *lacZ*. This cloning step was performed by digesting both the PCR product and pSG1729 with the restriction endonucleases SphI/NotI followed by ligation, leading to pBGAB.

   DNA fragments corresponding to the full-length regulatory regions (including the start codon) of the five purine riboswitch-controlled transcriptional units from *B. subtilis* (see Fig. 3) were synthesized. They were flanked by the restriction sites SphI upstream and BsaI downstream, the latter allowing to achieve scarless cloning [85]. Insertion of these synthetic DNA fragment into pBGAB using *Sph*I and *Bsa*I endonucleases generated the plasmids pBGABpur, pBGABxpt, pBGABnupG, pBGABpbuG, and pBGABpbuE respectively. For the generation of a reporter strain where the *pur* operon riboswitch is the sole modulator of LacZ expression, a DNA fragment corresponding to the *purE* regulatory sequence with inactivated PurBoxes was synthesized. PurBox inactivation consisted in the lack of the promoter-distal PurBox and the substitution C-44G in the proximal PurBox [21], which was not merely deleted due to a potential sequence overlap with the promoter P*pur*. Introduction of the resulting fragment into pBGAB using the same procedure as described above resulted in the plasmid pBGAB11. A final DNA fragment corresponding to the *purE* regulatory sequence lacking the purine riboswitch was synthesized and inserted into pBGAB following the procedure described above, leading to the plasmid pBGAB12. Finally, the editing of the purine riboswitch in pBGAB11 (deletion of nucleotides 181GCAATGCAGAGCGGGTAT198 within *purE* 5'-UTR)

generated the plasmid pBGAB13. This deletion was achieved by site directed mutagenesis using the complementary oligonucleotide pair S03/S04 and the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent, Waldbronn, Germany).

3.4.4 CRISPR-Cas9 plasmids

For CRISPR-Cas9 genome editing, the repair template moieties flanking target sequences were amplified by PCR from *B. subtilis* 168 chromosomal DNA using the primer pairs S05/S06 and S07/S08 for *purR* deletion, S09/S10 and S11/S12 for editing the *pur* operon riboswitch, S15/S16 and S17/S18 for *pur* operon riboswitch deletion, S19/S20 and S21/S22 for editing the *rib* operon riboswitch, and S19/S20 and S23/S24 for *rib* operon riboswitch deletion. The resulting PCR fragments flanking *purR* were introduced into SmaI-linearized pJOE8999 using a single step CPEC reaction as described by Quan and Tian (2009) [86], generating pPurRHT. The repair template moieties flanking the target sequence for *pur* operon riboswitch editing were assembled using overlap extension PCR, and the resulting fragment was amplified using the primer pair S13/S14, which introduced terminal SalI and XbaI sites. The resulting fragment and pJOE8999 were treated with SalI/XbaI and ligation generated peRSpurHT. SalI/BsaI or BsaI/XbaI-treated repair template moieties flanking the target sequence for *rib* operon riboswitch editing were assembled and ligated to SalI/XbaI-treated pJOE8999 in one step generating peRSribHT. Spacer DNA sequences were generated by annealing the complementary oligonucleotide pairs S25/S26 for targeting *purR*, S27/S28 for targeting the *pur* operon riboswitch, and S29/S30 for targeting the *rib* operon

riboswitch. Thereafter, they were respectively ligated into BsaI-digested pPurRHT, peRSpurHT, and peRSribHT, generating the final CRISPR-Cas9 vectors pABpurR, pABeRSpur, and pABeRSrib that were used for chromosomal *purR* deletion, *pur* operon riboswitch editing, and *rib* operon riboswitch editing, respectively. Final vectors for deletion of the *pur* operon riboswitch pABdRSpur and deletion of the *rib* operon riboswitch pABdRSrib were generated via the assembly and cloning of their homology templates respectively into pABeRSpur and pABeRSrib using SalI/BsaI/XbaI in one step as described above.

3.5 *In vitro* transcription

In vitro transcription assays were performed employing T7 RNA polymerase (Promega, Mannheim, Germany) for the production of sRNAs from linearized pT7sRNAx plasmids (**Table 13**) or employing *E. coli* RNA Polymerase Holoenzyme (New England Biolabs, Frankfurt, Germany) for transcription of the *xpt-pbuX* riboswitch−*luc* fusion from pPrib-RSxpt-luc (**Table 14**) or pPtac-RSxpt-luc' (**Table 15**). Both type of reactions were performed in the presence of Recombinant RNasin Ribonuclease Inhibitor (Promega) for preventing RNA degradation. Reaction mixtures were prepared as described in the following tables and incubated 12 h at 37 °C.

Table 13 | Reaction mixture for sRNA production *in vitro* (Fig. 9).

Component	Amount
Transcription Optimized 5x buffer	4 µl
DTT, 100 mM	2 µl
Ribonuclease Inhibitor, 100 U/µl	0.5 µl (50 units)
rNTP mix, 10 mM	0.4 µl
Linear pT7sRNA	500 ng
T7 RNA Polymerase, 20 U/µl	1 µl (20 units)
Nuclease-free water	Up to 20 µl

Table 14 | Reaction mixture for *in vitro* transcription of the *xpt-pbuX* riboswitch−*luc* fusion driven by Prib (Fig. 10 & 11).

Component	Amount
5x *E. coli* RNA Polymerase buffer	2 µl
DTT, 100 mM	1 µl
Ribonuclease Inhibitor, 100 U/µl	0.25 µl (25 units)
rNTP mix, 10 mM[a]	0.2 µl
pPrib-RSxpt-luc	250 ng
Purified sRNA	100 ng
E. coli RNA Polymerase, 1 U/µl	2 µl (2 units)
Nuclease-free water	Up to 10 µl

[a]*Different concentrations of the rNTP mix were tested (Fig. 11b, lanes A, B, C)*

Table 15 | Reaction mixture for *in vitro* transcription of the *xpt-pbuX* riboswitch−*luc'* fusion driven by PtacI (Fig. 13). Conditiion 1: purified sRNAs were added to the riboswitch transcription reaction. Condition 2: sRNAs were co-transcribed with the riboswitch in a combined *E. coli*/T7 RNA polymerases reaction.

Component	Condition 1	Condition 2
5x *E. coli* RNA Polymerase buffer	2 µl	1 µl
Transcription Optimized 5x buffer	/	1 µl
DTT, 100 mM	1 µl	1 µl
Ribonuclease Inhibitor, 100 U/µl	0.25 µl	0.25 µl
rNTP mix, 10 mM	0.2 µl	0.2 µl
Linear pPtac-RSxpt-luc'	250 ng	250 ng
Linear pT7sRNA	/	250 ng
Purified sRNA	150-200 ng	/
Xanthine (1 mM) or DMSO	1 µl	1 µl
E. coli RNA Polymerase, 1 U/µl	2 µl	2 µl
T7 RNA Polymerase, 20 U/µl	/	0.5 µl
Nuclease-free water	Up to 10 µl	Up to 10 µl

Transcribed RNAs were purified using RNA Clean & Concentrator-5 kit (Zymo Research) with in-column DNase I treatment following the manufacturer's protocol. Purified RNA was quantified using NanoVue Plus Spectrophotometer (GE Healthcare) and stored at -80 °C.

3.6 Northern Blot

500 ng of RNA purified from *in vitro* transcription reactions or 10 µg total RNA extracted from *B. subtilis* was mixed with 2x RNA Gel Loading Dye (Thermo Fisher Scientific) and incubated 10 min at 70 °C followed by 3 min on ice. RNAs were then separated by denaturing agarose gel electrophoresis (1.5% agarose in 1x MOPS buffer (Table 6), 3.7% formaldehyde) and blotted overnight with 20x SSC buffer (Table 7) onto a nylon membrane (Hybond-N+, Merck/Sigma Aldrich) by capillary force using The CP-1526 mini capillary blotter (Scie-Plas, Cambridge, UK).

RNA was crosslinked to the membrane by exposure to UV (e = 1200 x 100 µJ/cm^2). The membrane was pre-hybridized 30 min with DIG Easy Hyb solution (Roche) at 50 °C, then incubated overnight at the same temperature with a DIG Easy Hyb solution containing 20 nM double digoxigenin (DIG)-labeled DNA probe for hybridization. The DIG-labeled oligonucleotide S45 (Table 10) binds to the *xpt-pbuX* riboswitch, allowing thereby detection of both terminated and unterminated transcripts. The labeled oligonucleotide S46 binds to *lacZ* mRNA, and S47 binds to the *lacO* sequence transcribed upstream of sRNAs. The membrane was washed twice for 5 min in 2x SSC, 0.1% SDS solution at room temperature (RT) and twice for 15 min in 0.1x SSC, 0.1% SDS solution at 55 °C. The membrane was rinsed with washing buffer (maleic acid buffer (Table 8) containing 0.3% polysorbate 20) and incubated 30 min at RT in blocking solution (1% m/v Blocking Reagent (Roche) dissolved in maleic acid buffer) followed by another 30 min incubation in antibody solution (Anti-DIG-AP 150 mU/ml (Roche) diluted 1:5000 in blocking solution). The membrane was

washed twice for 15 min in washing buffer and equilibrated 2-5 min in detection buffer (Table 9). CDP-*Star* solution (0.25 mM, Roche) was added to the membrane, and after 5 min incubation, the chemiluminescence signal was detected.

3.7 RNA-Seq

RNA samples extracted from three independent cultures of *B. subtilis* 168 transformed with pHTsRNA$_{random}$ were pooled in one sample to an equimolar ratio. The following steps of library preparation, sequencing, and data processing were performed by Tobias Busche (Kalinowski lab, CeBiTec, Bielefeld University) as described in Schaffert et al. (2019) [87]. rRNA was depleted using the Ribo-Zero rRNA Removal Kit for bacteria (Illumina, San Diego, CA, United States). Successful depletion was verified using an Agilent RNA 6000 Pico chip in a Bioanalyzer (Agilent, Böblingen, Germany). cDNA libraries were prepared following protocols from Pfeifer-Sancar et al. (2013) [88] and Irla et al. (2015) [89] using the TruSeq stranded mRNA kit (Illumina). The libraries were quantified by a DNA High Sensitivity Assay chip in the Bioanalyzer and sequenced on a 2 × 75 nt HiSeq 1500 run (Illumina). Raw reads were quality-trimmed using Trimmomatic [90]. Trimmed reads were mapped to the respective reference sequence using bowtie2 in the paired-end mode [91].

3.8 Chromosomal integration and amylase assay

pBGAB-derived integration vectors were linearized using the endonuclease ScaI, purified, and introduced into *B. subtilis* 168. Cells were plated on LB-agar plates supplemented with spectinomycin and incubated overnight at 37 °C.

Successful double crossover integration of pBGAB-derived constructs into the *amyE* chromosomal locus of *B. subtilis* results into a loss of amylase activity. This was tested as follows. Single colonies were suspended in 20 µL sterile water and plated on fresh LB-agar plates supplemented with spectinomycin and 1 % (m/v) potato starch. Following an overnight incubation at 37°C, colonies were scraped from the plate with a scraper. Plates were flush with 1.1 ml Lugol's iodine solution. Plates corresponding to colonies with disrupted *amyE* were fully stained, whereas plates containing colonies with intact *amyE* displayed unstained halos due to the absence of stainable amylose in colony spots [92].

3.9 β-galactosidase assay

B. subtilis strains with different versions of regulatory regions from purine riboswitch-controlled transcriptional units fused to *lacZ* and integrated in the *amyE* locus were cultivated to the mid-exponential growth phase, except for BGAB11, BGBSpurR, and BGeBSpur2 (Fig. 28b), which were cultivated to the early stationary phase for assessing purine production levels during their complete growth cycle. Cultures were centrifuged, cells were washed, centrifuged again, and cell pellets suspended in buffer PM (10 mM NaH_2PO_4,

90 mM Na_2HPO_4, 1 mM $MgSO_4$; pH 7.8) in which the cells were lysed. Using *o*-nitrophenyl-β-D-galactopyranoside (ONPG) as a substrate, β-galactosidase activity was measured by monitoring ONPG hydrolysis over time. Activity values expressed in nanomoles of *o*-nitrophenol produced per minute per milliliter of culture (mU/ml) were normalized to the optical density values of corresponding cultures.

3.10 CRISPR-Cas9 genome editing

The CRISPR-Cas9 vectors were generated using *E. coli* DH5α and propagated using the *recA*-proficient strain *E. coli* NM538 [93] from which they were isolated and employed for *B. subtilis* transformation. *B. subtilis* strains transformed with CRISPR-Cas9 vectors were plated on LB-agar plates supplemented with 0.2% mannose and 50 µg/ml kanamycin. With a kanamycin concentration of 5 µg/ml, about 10% of the colonies corresponded to the original untransformed strain, indicating that the applied selective pressure was too low. Hence, kanamycin concentration was increased to 50 µg/ml. More than a thousand transformants per microgram of DNA were obtained on every transformation plate. Plates were incubated 24 to 36 hours at 30 °C until colonies were visible. For plasmid curing, cells were transferred to LB-agar plates without kanamycin and incubated overnight at 50 °C. Subsequently, colonies were streaked on fresh LB-agar plates without kanamycin and incubated overnight at 42 °C to generate single colonies. A part of these colonies was transferred to LB plates containing kanamycin and incubated at 30 °C to test them for plasmid loss. Simultaneously, the

remaining colony material was transferred to 40 µL sterile water for colony PCR.

Spacer design was performed using the GPP sgRNA designer tool [94] for identifying spacers with high predicted ON-target activity and Cas-OFFinder [95] for *in silico* testing of selected spacer candidates for target specificity in *B. subtilis*.

3.11 Colony PCR from *B. subtilis* cells

Single colonies were suspended in 40 µL water, cells were disrupted by three freeze-boil cycles, samples were centrifuged, and 5 µL of the supernatant was used as DNA template in a 40 µL PCR reaction. Primer pairs S31/S32, S33/S34, and S35/S36, which anneal outside the sequences that correspond to the homology repair templates, were used for amplification of genomic regions around *purR*, the *pur* operon riboswitch, and the *rib* operon riboswitch respectively. Amplified fragments were analyzed by agarose gel electrophoresis and DNA sequencing.

3.12 Quantitative RT-PCR

1 µg of total RNA extracted from *B. subtilis* was used for cDNA synthesis using M-MLV Reverse Transcriptase (Promega) in the presence of random hexamer primers. The reaction was diluted 1:5 in water and used for quantitative PCR with PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) and 0.4 µM of each primer. The primer pairs S37/S38, S39/S40, S41/S42, and S43/S44 were used for amplification of cDNA corresponding to *purR*, *purE*, *ribD* and

rrnA-16S, respectively. The qPCR reaction was performed in a 10 µL final volume in a 384 well-plate using a Quantstudio 5 device (Applied Biosystems, Thermo Fisher Scientific).

3.13 HPLC analysis of flavins

Aliquots were removed from *B. subtilis* 168, eBSflv, BSRF, and eBSRFpur2 cultures, and the cells were disrupted into the growth medium to determine total production of riboflavin, FMN, and FAD. Samples were treated with 10% (m/v) trichloroacetic acid, filtered, and flavins were separated using an Agilent 1220 Infinity LC System (Agilent Technologies) with a reverse phase biphenyl column (2.6 µm particle size, 150 mm x 2.1 mm. Phenomenex, Aschaffenburg, Germany) at a flow rate of 0.2 ml/min. The column was equilibrated into 15% (v/v) methanol, 20 mM formic acid, 20 mM ammonium formate; pH 3.7. A gradient from 15% to 42% methanol was applied over 8 minutes during which absorption at 445 nm was monitored.

CHAPTER I. Targeting purine riboswitches with

synthetic small RNAs

4 Results and Discussions

4.1 Design of regulatory small RNAs (sRNAs)

4.1.1 General design strategy

Following a strategy based on rational design combined with *in silico* evolution to target purine-binding aptamers from *B. subtilis* and interfere with the binding of their natural ligands, a set of designer regulatory small RNAs with various properties were generated using a computational automated tool termed Ribomaker, which was manually fed with structural and rational input. Ribomaker is based on an algorithm that implements an evolutionary design strategy for bacterial riboregulation, where random mutations are selected according to a physicochemical model based on the minimization of free energies [96]. The number of iterations was set to 9000 for the design of all sRNAs in the present study. The length of generated sRNAs varies between 30 and 110 nucleotides. A 30 nucleotide sequence corresponding to a bacterial Rho-independent terminator was added to the 3' end of each sRNA. Structural constrains were fed into the program by specifying secondary structures of the targeted aptamers into the input data. An additional constrain was added by specifying an intermolecular seed sequence, which was set in the 5' end of the aptamer sequence in order to enable the sRNA to start annealing with the

nascent aptamer as soon as the transcription process of the riboswitch starts, giving thereby a kinetic advantage to the sRNA over the ligand, which is only able to bind to the riboswitch when the aptamer is fully synthesized and folded. This approach is in accordance with the kinetic regulation model proposed for riboswitches that control expression at the transcriptional level [12-15]. To make sure that sRNA binding to the aptamer does not affect formation of the downstream structures, an additional constrain was integrated by specifying that the aptamer's 3' end sequence that is involved in the formation of the antiterminator (see Fig. 1a) or the terminator (see Fig. 1c) should stay unpaired in the aptamer-sRNA complex, except for $sRNA40_{xpt}$, which was intended to be used as a control for this specific approach.

4.1.2 Generation of a consensus sequence and a consensus structure representing the purine-binding aptamers

Considering the high conservation of the purine aptamer domains at both the sequence and the structure levels, it is possible to identify and target the most conserved nucleotides. Therefore, a unique consensus sequence representing the five purine-binding aptamers from *B. subtilis* was generated using the software LocARNA [97, 98] (**Fig. 5a**), and its predicted secondary structure (**Fig. 5b**) was manually optimized by comparison with the available crystal structure of the *xpt-pbuX* aptamer [99] (**Fig. 5c**).

Figure 5 | Generation of a consensus sequence and a consensus secondary structure representing the five conserved purine-sensing aptamers of *B. subtilis*.

(**a**) Multiple sequence alignment of the purine aptamers showing sequence and structural conservation. Sequence conservation is represented by grey columns

under the alignment. The alignment is annotated with its consensus structure, which is represented as a string of dots and brackets on the top of the alignment; the string is well-bracketed, such that base pairs in the structure are indicated by corresponding opening and closing brackets. Predicted intramolecular pairings are highlighted with a color code; the hue shows sequence conservation and the saturation shows structural conservation. Alignment generated using the LocARNA program. Predicted consensus structure in LocARNA is generated by the tool RNAalifold [100] of the Vienna Package [101].

(**b**) 2D layout of the consensus secondary structure of purine aptamers alignment as predicted by RNAalifold. Base pairs are colored using the same color code as in (**a**). Per alignment position, the set of nucleotides with frequency higher than average are represented using IUPAC ambiguity codes for nucleotides.

(**c**) Secondary structure of the *xpt–pbuX* aptamer of *B. subtilis* confirmed by X-ray crystallography. Nucleotides conserved in more than 90% of known guanine riboswitches are shown in red. The numbering is consistent with that of the full-length mRNA. Figure adapted from Batey et al. (2008).

(**d**) Final consensus sequence with an optimized consensus structure of the purine aptamers. The secondary structure, represented as a string of dots and brackets as described in (**a**), was optimized by comparison of the predicted consensus structure in (**b**) with the confirmed structure in (**c**).

4.1.3   Generation and validation *in silico* of sRNAs

Targeting conserved nucleotides would additionally allow a multitargeting strategy where a single sRNA could be used to deregulate the function of all conserved riboswitches simultaneously. Four sRNAs (sRNA30, sRNA70, sRNA80, and sRNA110) were thereby designed following this strategy by setting the consensus sequence and structure (**Fig. 5d**) as targets for sRNAs in Ribomaker (**Fig. 6a**), whereas three additional sRNAs were designed to target specifically either the *xpt-pbuX* riboswitch (sRNA40$_{xpt}$) or the *pur* operon riboswitch (sRNA30$_{pur}$ and sRNA71$_{pur}$) using their individual corresponding aptamer sequence as a target in Ribomaker (**Fig. 7a**). Generated sRNAs were validated *in silico* by simulating their interaction with the aptamer target using the NUPACK program [102] (**Fig. 6b** and **Fig. 7b**).

Experimentally, the aim is to first test and compare the activity of the generated sRNAs in order to gain insight about the effect of different design parameters such as the size of sRNAs, the targeted regions in the aptamer, the type of target (consensus sequence vs individual aptamer), and the level of pairing, and to subsequently optimize the design strategy accordingly to generate an improved sRNA if necessary.

a

sRNA30
```
UUCGGGCAAGGAGUGUGAGUGUUUCUUUAGcaaagcccgccgaaaggcgggcuuuucugu
((((.((.....)).)))).........(((.(((((((((((....)))))))))).))).
```

Consensus aptamer + sRNA30
```
(((((((((((((((....(((.(....((((..............)))....)))))....((
(((((..............&.))))))))....))))))))))))))).(((.((((((((((
....)))))))))))).))).
```

hybridization free energy = -20.8 kcal/mol
Seed-based free energy = -8.4 kcal/mol

sRNA70
```
CGCGGCGUGACGAGUGCUGCAUGGUAGGCAAAGGGCUGCCUCAUGUCCCUGCGUUAAUUG
UGGUGUGUUUcaaagcccgccgaaaggcgggcuuuucugu
(((.(((((((((((.(..(((((..(((((......)))))))))).))).))))))..))
).))))......(((((((((((....))))))))))).....
```

Consensus aptamer + sRNA70
```
.((((((((((((....((((((.((((((((.((..(((((((((((((((((((..(((
((((..............&)).))))).))).)))))...)))))))).)))...))))))).
.)))).))))))).....))))).)))))))).(((((((((....))))))))))......
```

hybridization free energy = -27.5 kcal/mol
Seed-based free energy = -7.2 kcal/mol

sRNA80
```
UUUCUGUGCAUCGUGCCCACUGGUACAUGCGAUACGUGGUAUUGGGCAAGAGGUUGUGGU
UGUCGGGCGCCGCCGUAAUUcaaagcccgccgaaaggcgggcuuuucugu
.......((((.(((((....))))))))))..(((((((..((.(((((..........)
))))).)).).)))).))))....(((((((((((....))))))))))).....
```

Consensus aptamer + sRNA80
```
.....((((((((...))))).)))...(((((((((((..((....)))..((((((((((.(
(((((.(((((((..((((.&.)))).).))))..((((.(((((((....))))))))))....)))))
)...))))))))....)))))))))))))..))).))))))))....((((((((((....))))
))))))).....
```

hybridization free energy = -20.1 kcal/mol
Seed-based free energy = -7.2 kcal/mol

sRNA110
```
UAUCGGGCCUGAUCGUUGCGUACGUAGAAGAGUGUGCCGGGUUUCGGGACCGAUGCCUCG
GUUAUGCAGGGCGACUUCAUCGUCGUCUCCCUAUAGCCAUUGUGAGUGUUcaaagcccgc
cgaaaggcgggcuuuucugu
(((((((.(((((.(.(.(((((((.......)))))).).).)..))))).))))))....(
((((((..((((((((......)))))))...))))))))...........((((((((
(....))))))))))).....
```

Consensus aptamer + sRNA110
```
...(((((((((((..(((((.(((((.((((((((...((((((((.(.(((.((.....(
(((((..............&..)))))).....)).)))).)..))))))))).))))))))).)
))))).)...))))).....(((((((..((((((((......)))))))...)))))))))..)
)))))))))).(((((((((....))))))))))).....
```

hybridization free energy = -22.1 kcal/mol
Seed-based free energy = -4.8 kcal/mol

b

Figure 6 | Generation and *in silico* validation of sRNAs designed using the consensus sequence and consensus structure of the purine-binding aptamers of *B. subtilis*.

(**a**) Generation of sRNA30, sRNA70, sRNA80, and sRNA110 using Ribomaker. These sRNAs were designed against the consensus sequence and structure generated in Figure 5d to target conserved nucleotides in the purine aptamers, and each sRNA is theoretically able to bind to all of the five purine aptamers in *B. subtilis*. The displayed output data show the primary sequence of each sRNA, its secondary structure, and the secondary structure of the complex sRNA-aptamer. sRNAs are named after their length in number of nucleotides excluding the terminator sequence shown in lower case letters. Secondary structures are represented by a string of points and brackets, such that base pairs are indicated by corresponding opening and closing brackets. Calculated free energies are displayed for the hybridization of the whole sRNA-aptamer complex as well as for the hybridization of the sRNA with the corresponding seed sequence part set in the 5′ end of aptamer sequences.

(**b**) Validation of sRNA-consensus aptamer interactions using the NUPACK software. Predicted MFE structures of sRNA-aptamer complexes at 30 °C are displayed with calculated free-energies of secondary structures. The color code shows the probability for each nucleotide to be in the predicted state at reaction equilibrium.

a

sRNA40$_{xpt}$
UGCUUCCUACGCGGACCAUAUCUACGCCUGUGUGAGUAGCcaaagcccgccgaaaggcgg
gcuuuucugu
.((((.(.(((((((((.............)))))))).).))).((((((((((((....)))))
)))))).....

xpt-pbuX aptamer
AGGAACACUCAUAUAAUCGCGUGGAUAUGGCACGCAAGUUUCUACCGGGCACCGUAAAUG
UCCGACUAUGGGUGAGCA
.....(((((((((......((((.......)))))..........(((((((.......))
))))..)))))))))....

xpt-pbuX aptamer + sRNA40$_{xpt}$
.((......(((((((((...(((((((((((((..(((..........(((((((.......))
))))....((((.(((((. & .)))))))).))).))))))))))))).)))))))....
)).(((((((((....))))))))))......

hybridization free energy = -20.9 kcal/mol
Seed-based free energy = -8.4 kcal/mol

sRNA30$_{pur}$
GAUUCAACAAUUAUGUGAGGUCGUGUUUCUcaaagcccgccgaaaggcgggcuuuucugu
(((((..(((....))).)))))..........((((((((((....))))))))))...

pur operon aptamer
ACACGACCUCAUAUAAUCUUGGGAAUAUGGCCCAUAAGUUUCUACCCGGCAACCGUAAAU
UGCCGGACUAUGCAGGAAAG
......(((((((......((((.......))))............(((((((.......)
))))))..)))).))).....

pur operon aptamer + sRNA30$_{pur}$
(((((((((((((((((..((((.......))))....(((((..(((((((.......)
))))))........))))).&..........))))))))))))))))).....(((((((
((....))))))))))).....

hybridization free energy = -25.8 kcal/mol
Seed-based free energy = -18.5 kcal/mol

sRNA71$_{pur}$
GUGCGAGCGGCAUAUGCUUGUGGGCCACUCCGCUUAACAUGGCGAAUGUUCCCGAGGCAU
UGUAUGAGGCUcaaagcccgccgaaaggcgggcuuuucugu
....((((..(((((((((.(((........(((......)))......)))))))))).
..)))))..)))).((((((((((....))))))))))......

pur operon aptamer + sRNA71$_{pur}$
......(((((((..)))).))))..(((((((((((..((....))..((((((((((((.(
(((((.((((((..(((((.. & .)))))..((((.(((((.(....))))))))))....)))
)))...))))))....))))))))))))..))).)))))))....(((((((((((....))
))))))))).....

hybridization free energy = -24.4 kcal/mol
Seed-based free energy = -8.4 kcal/mol

b

74

Figure 7 | Generation and *in silico* validation of sRNAs targeting specifically the *xpt-pbuX* aptamer or the *pur* operon aptamer of *B. subtilis*.

(**a**) Generation of sRNA40$_{xpt}$, sRNA30$_{pur}$, and sRNA71$_{pur}$ using Ribomaker. sRNA40$_{xpt}$ was designed against the sequence and confirmed structure of the *xpt-pbuX* aptamer. sRNA30$_{pur}$ and sRNA71$_{pur}$ were designed against the sequence of the *pur* operon aptamer and its optimized predicted structure. The figure shows the primary sequence of each sRNA and aptamer, their secondary structures, and the secondary structure of the complex sRNA-aptamer. The terminator sequence at the 3' ends of sRNAs is shown in lower case letters. Secondary structures are represented by a string of points and brackets, such that base pairs are indicated by corresponding opening and closing brackets. Calculated free energies are displayed for the hybridization of the whole sRNA-aptamer complex as well as for the hybridization of the sRNA with the corresponding seed sequence part set in the 5' end of aptamer sequences.

(**b**) Validation of sRNA interaction with its corresponding aptamer target using the NUPACK program. Predicted MFE structures of sRNA-aptamer complexes at 30 °C are displayed with calculated free-energies of secondary structures. The color code shows the probability for each nucleotide to be in the predicted state at reaction equilibrium.

4.1.4 Generation of perfectly complementary antisense RNAs

In addition to the partly complementary sRNAs designed using Ribomaker, two antisense RNAs (asRNAs) were synthesized with perfect complementarity to the *pur* operon aptamer (asRNA*pur*) or to the *nupG* aptamer (asRNA*nupG*) (**Fig. 8**), yet excluding the 3' end sequence that is involved in formation of the antiterminator as described in the general design strategy section above (section 4.1.1). Finally, a random RNA sequence (sRNA$_{random}$) was generated to be used as a general negative control.
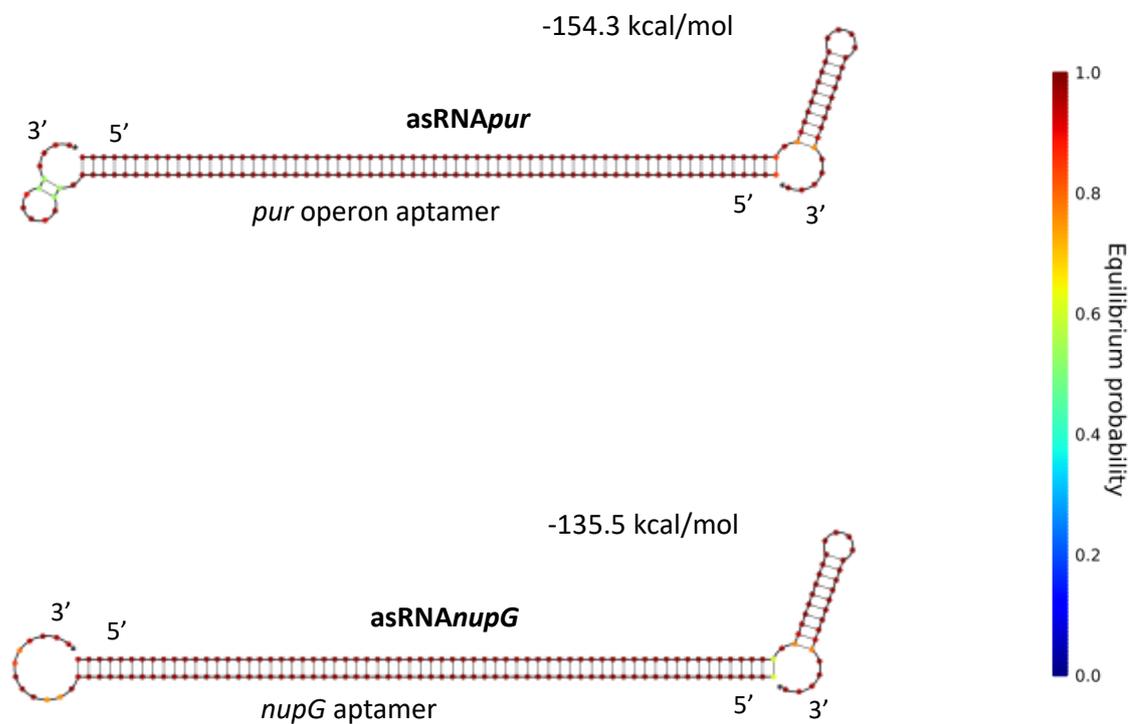
Figure 8 | Generation of antisense RNAs (asRNAs) perfectly complementary to the *pur* operon aptamer (upper panel) or to the *nupG* aptamer (lower panel) of *B. subtilis*. MFE structures of asRNA–aptamer complexes at 30 °C predicted using the NUPACK software, with calculated free-energies of secondary structures. The color code shows the probability for each nucleotide to be in the predicted state at reaction equilibrium.

4.2 Expression *in vitro* of designed small RNAs

DNA fragments corresponding to the designed sRNAs driven by the strong bacteriophage T7 promoter [84] and followed by the T7 terminator sequence were synthesized and inserted into the pIDTSMART plasmid (by Integrated DNA Technologies), generating a set of sRNA expression vectors (pT7sRNAx). An *in vitro* sRNA production test was performed for sRNA40$_{xpt}$, sRNA80, and a control sRNA using their corresponding plasmids pT7sRNA40$_{xpt}$, pT7sRNA80, and pT7sRNA$_{control}$, which were linearized using HindIII restriction site between the sRNA and the terminator sequence. The resulting linear plasmids were used in an *in vitro* run-off transcription assay employing the T7 RNA polymerase. sRNAs were successfully expressed (**Fig. 9**), and 0.7-1.3 µg RNA could be recovered from a 20 µL reaction after DNase I treatment and purification, allowing the direct use of purified sRNAs in further experiments *in vitro* for the evaluation of their activity.
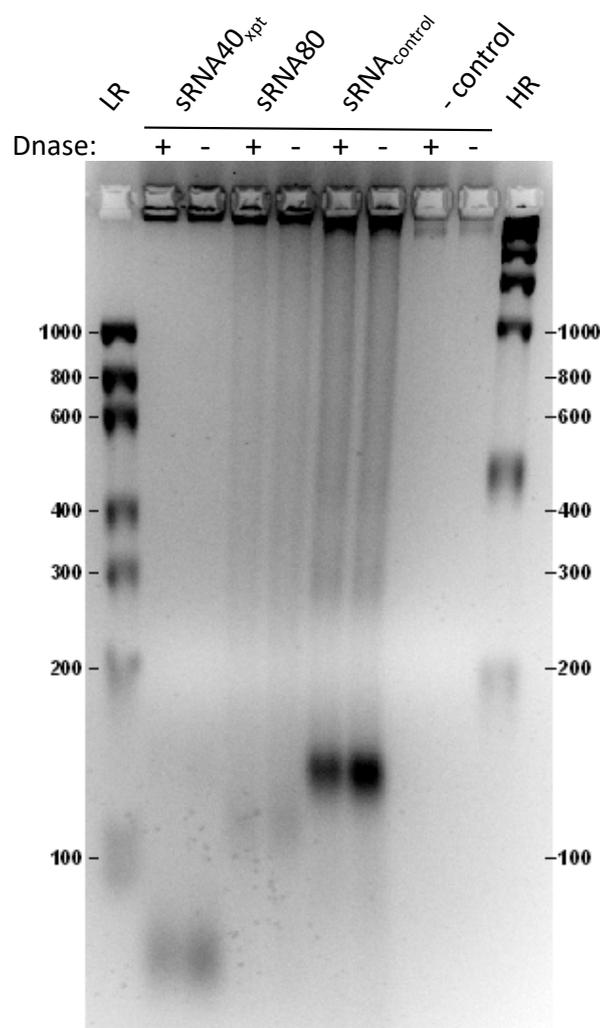
Figure 9 | Production *in vitro* of sRNAs using the bacteriophage T7 expression system. *In vitro* transcription reactions with (+) or without (-) subsequent DNase I treatment, analyzed using denaturing RNA gel electrophoresis stained with ethidium bromide. The negative control (- control) lanes correspond to the use of a DNA template lacking the T7 promoter. LR: low-range marker (Riboruler Low Range RNA Ladder). HR: high-range marker (Riboruler High Range RNA Ladder). The size is expressed in nucleotides. Expected sizes for sRNA40$_{xpt}$, sRNA80, and sRNA$_{control}$ are approximately 80, 120, and 158 nucleotides, respectively.

4.3 Development of an *in vitro* transcription assay for testing sRNA activity

To be able to directly monitor the regulatory function of purine riboswitches at the transcriptional level and to subsequently measure the effect of the designed sRNAs in the presence of purine ligand, a cell-free *in vitro* transcription assay system was developed employing the commercially available *Escherichia coli* RNA polymerase. A fusion construct was created by insertion of the *B. subtilis xpt-pbuX* riboswitch between the *E. coli ribB* promoter Prib [103] and the firefly luciferase gene *luc* in the plasmid pPrib-luc [25]. Thus, transcription of *luc* is modulated by the purine-sensing *xpt-pbuX* riboswitch in the resulting plasmid pPrib-RSxpt-luc.

4.3.1 *In vitro* transcripts were undetectable with gel electrophoresis using the promoter Prib in a linear template

The constructed vector pPrib-RSxpt-luc was linearized using the endonuclease SspI and used as a template in an *in vitro* transcription reaction in the absence of purine ligand. However, no detectable RNA fragment band was observed on an ethidium bromide-stained gel (**Fig. 10**). Undetectable transcripts indicate that the activity of the *E. coli* promoter Prib is too low to allow transcript detection without labelling and/or that Prib is a supercoiling sensitive promoter that is inactive in a linear DNA template.
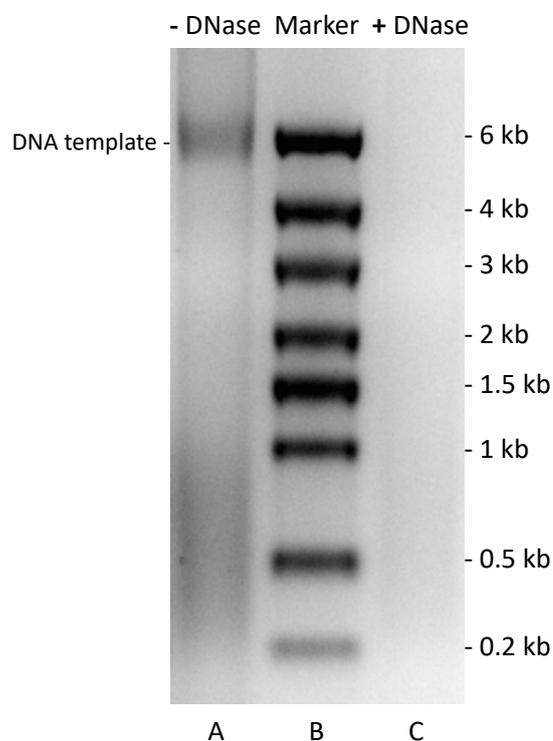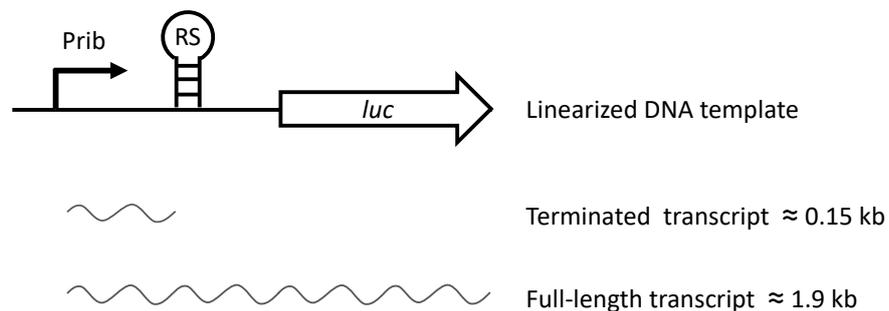
Figure 10 | *In vitro* **transcription using the** *E. coli* **promoter Prib in a linear DNA template does not result in detectable transcripts on a denaturing RNA gel stained with ethidium bromide.** Lane A: *in vitro* transcription reaction before treatment with DNase I. Lane B: RNA marker (Riboruler High Range RNA Ladder). Lane C: *in vitro* transcription reaction after treatment with DNase I. RS: *xpt-pbuX* riboswitch.

81

4.3.2 The terminated riboswitch is transcribed *in vitro* using a circular DNA template and detected with Northern blot

These suggestions were investigated further by combining *in vitro* transcription using either linear or circular (supercoiled) DNA template (**Fig. 11a**) with Northern blot analysis. Although the aborted transcript corresponding to the *xpt-pbuX* riboswitch could be detected in the circular but not in the linear template reaction (**Fig. 11b**. Lane A versus F), supporting that Prib is supercoiling sensitive, no full-length mRNA transcript (riboswitch−*luc*) could be detected despite the absence of ligand (**Fig. 11b**). This could be due to the low efficiency of *E. coli* RNA polymerase in transcribing *in vitro* the long 1.9 kilobase full-length mRNA and/or to the low efficiency of transcription termination using the circular plasmid *in vitro*. The reaction with a circular template was simultaneously performed under lower rNTPs concentrations (**Fig. 11b**. Lane B and C), as well as in the presence of sRNAs produced previously (see section 4.2) (**Fig. 11b**. Lane D and E). However, the absence of detectable full-length transcript prevents drawing further conclusions.

a
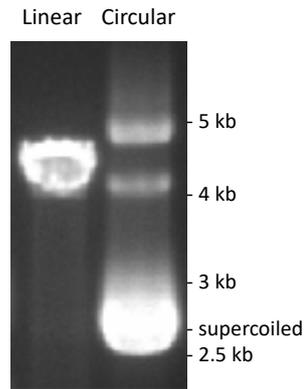
b
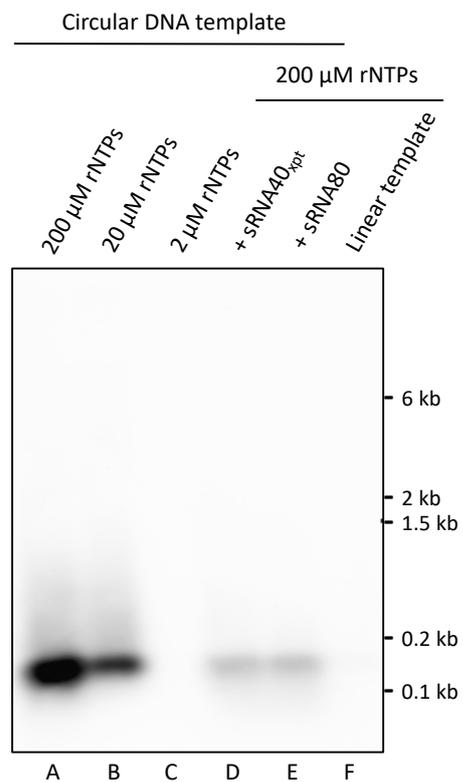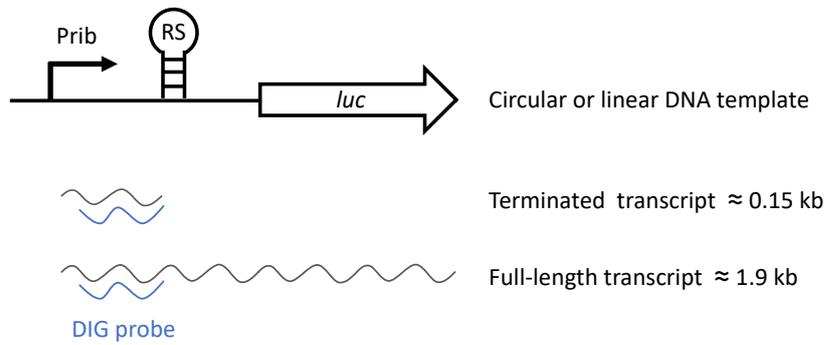
Figure 11 | The terminated transcript corresponding to the *xpt-pbuX* riboswitch is detected with Northern blot but not the full-length mRNA, and the *E. coli* promoter Prib is inactive when the DNA template is linearized.

(**a**) Agarose gel electrophoresis of the *in vitro* transcription template pPrib-RSxpt-luc in its linear and circular forms. The circular plasmid was isolated from *E. coli* cells and linearized using EcoRI restriction digest.

(**b**) Northern blot analysis of *in vitro* transcription reactions in the absence of purine ligand. The reactions were treated with DNase I before analysis. The DIG-labeled DNA probe binds to a sequence within the *xpt-pbuX* riboswitch, which is present in both terminated and full-length transcripts. The terminated transcript was detected using the circular DNA template at rNTPs concentrations of 20 and 200 μM, but it was not detected using the linearized template. The full-length transcript was not detected at all.

4.3.3 Optimization of the construct for transcription and detection of the unterminated transcript

To improve the assay, a new plasmid construct was designed, where Prib was replaced by the supercoiling-insensitive hybrid *E. coli* promoter PtacI [104]. In addition, the full-length *luc* gene was replaced by a truncated gene (*luc'*) and followed downstream by a XhoI restriction site. Therefore, the use of the resulting vector pPtac-RSxpt-luc', linearized with XhoI endonuclease, in an *in vitro* transcription run-off assay would lead to a shorter full-length transcript of 0.5 kb (**Fig. 12**) instead of the expected 1.9 kb using the former construct (see Fig. 11b).

	Linearized DNA template
	Terminated transcript ≈ 0.15 kb
	Full-length transcript = **0.5 kb**
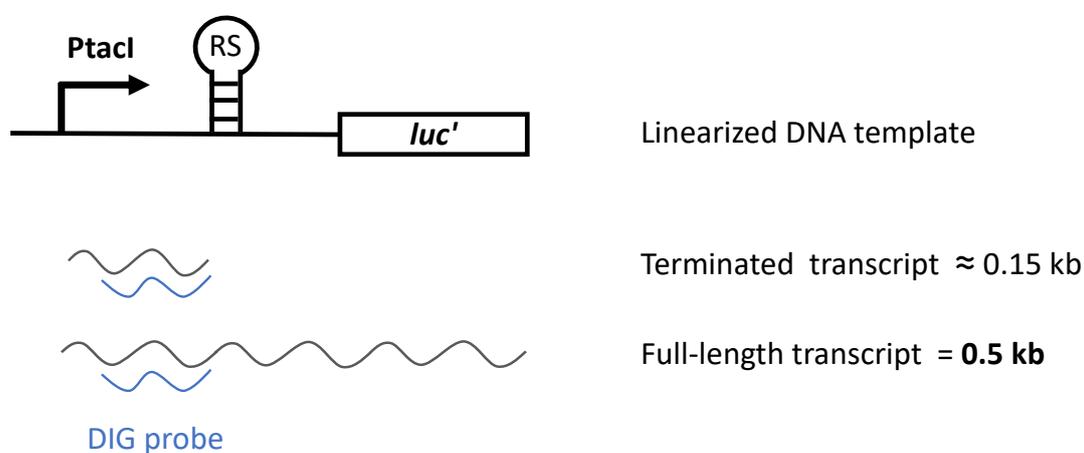
Figure 12 | **Optimization of the construct used in the *in vitro* transcription assay to perform a transcription run-off that generates a shorter unterminated transcript.** The promoter Prib was replaced by PtacI and the *luc* gene by a truncated *luc'* followed by a XhoI restriction site, which was used to linearize the plasmid. The DIG-labeled DNA probe binds to a sequence within the *xpt-pbuX* riboswitch, which is present in both terminated and unterminated transcripts. RS: *xpt-pbuX* riboswitch. *luc'*: truncated *luc* gene.

85

4.3.4 The *xpt-pbuX* riboswitch is unresponsive to xanthine *in vitro*

The new construct was tested in the absence of ligand and in the presence of 100 µM xanthine under two different experimental set up; either by directly adding the purified sRNAs produced previously (see section 4.2) to an *in vitro* transcription reaction based only on *E. coli* RNA polymerase (**Fig. 13**, left panel), or by adding the sRNA expression vectors pT7sRNAx prelinearized to a combined *E. coli*/T7 RNA polymerases-based reaction with mixed buffers (**Fig. 13**. Right panel). Using Northern blot analysis, the detection of both terminated (*xpt-pbuX* riboswitch) and unterminated run-off (*xpt-pbuX* riboswitch−*luc'*) transcripts was finally successful. However, the ratios unterminated/terminated transcripts calculated by quantification of band intensities show that addition of the ligand xanthine in the absence of sRNA did not trigger the expected downregulation of gene transcription (**Fig. 13**, lanes A vs B and F vs G). Consequently, no conclusions could be drawn about sRNA activity.
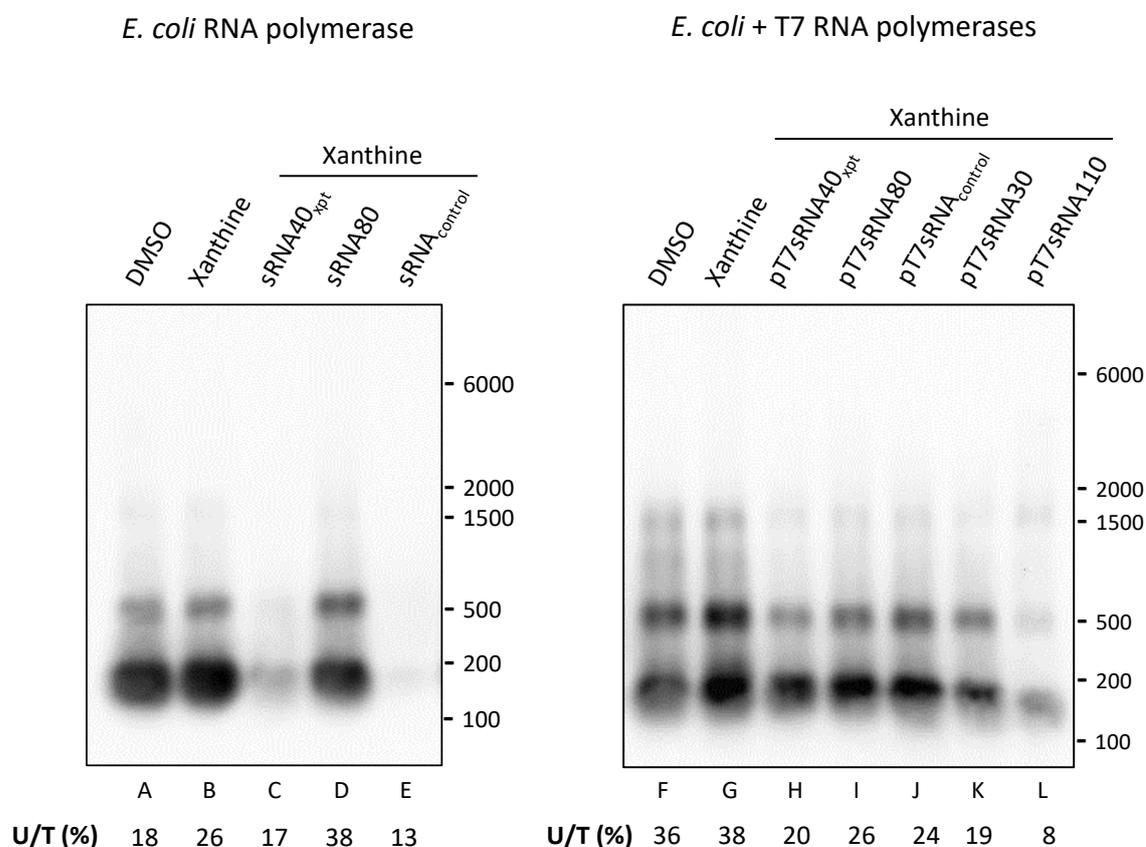
Figure 13 | The terminated *xpt-pbuX* riboswitch and the unterminated run-off transcript are both transcribed using PtacI promoter and detected with Northern blot, but the riboswitch is unresponsive to xanthine. The *in vitro* transcription reactions were performed in the presence of 100 µM xanthine, except for the DMSO control reactions (Lane A and F). The reactions were treated with DNase I before analysis. Left; reaction performed using *E. coli* RNA polymerase in its corresponding buffer for transcription of the riboswitch−*luc'* fusion, and purified sRNAs were directly added to the reaction. Right; reaction performed using *E. coli* and T7 RNA polymerases in a mixed buffer for transcription of the riboswitch−*luc'* fusion and the sRNA respectively. The ratio unterminated (U) / terminated (T) transcripts is expressed as a percentage for each well at the bottom of the figure.

The unresponsiveness of the riboswitch *in vitro* could be due to many factors, such as the composition of reaction buffers [105], the type of purine ligand, and the temperature. Furthermore, considering the kinetic regulation mechanism of transcriptional riboswitches [12-15], an optimal combination of the type of RNA polymerase, the strength of the promoter, and the concentration of rNTPs might be crucial for purine riboswitches to be functional *in vitro*. Considering all these uncertainties, optimization of the *in vitro* transcription assay would be highly challenging. Therefore, sRNA activity was directly evaluated *in vivo* where purine riboswitches are in their native functional environment.

4.4 Establishment of a reporter system *in vivo* to study purine riboswitch regulation and test sRNA activity

To be able to study the regulation of genes and operons controlled by the purine riboswitches in *B. subtilis* and to subsequently measure the effect of the designed sRNAs *in vivo*, five *B. subtilis* test strains corresponding to each purine riboswitch were constructed where the *E. coli* β-galactosidase reporter gene *lacZ* is modulated by a *B. subtilis* purine riboswitch directed upstream by its respective native promoter. Following characterization of the promoters and validation of the purine-triggered modulation, activity of designed sRNAs on purine riboswitches can be evaluated by co-expression of sRNAs under the control of the IPTG-inducible strong promoter Pgrac using the *B. subtilis* episomal expression vector pHT01 [79] (**Fig. 14**).
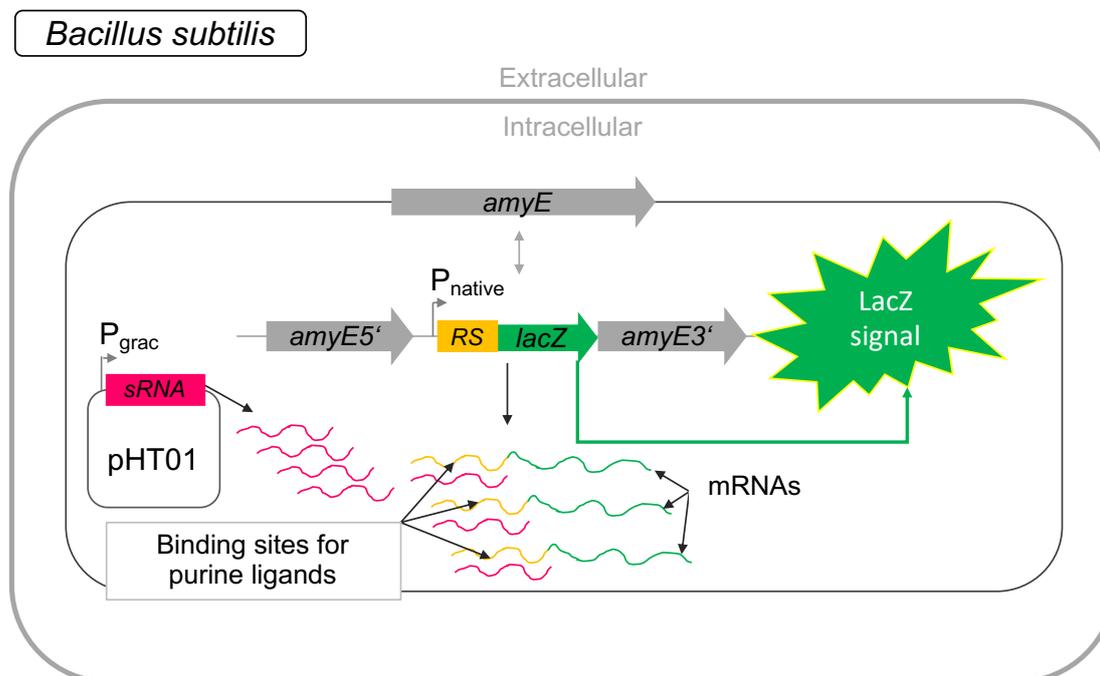
Figure 14 | Construction of a reporter system to characterize purine riboswitches and identify functional sRNAs in *B. subtilis*. The five purine riboswitches from *B. subtilis* coupled to *lacZ* are separately integrated into the *amyE* locus of a *B. subtilis* strain by homologous recombination. sRNAs are expressed from the episomal plasmid pHT01 under the control of the strong promoter Pgrac. Modulation of LacZ expression is monitored via the measurement of β-galactosidase enzymatic activity. RS: purine riboswitch. Pnative: the native promoter controlling the corresponding purine riboswitch's transcriptional unit (P*pur*, P*xpt-pbuX*, P*nupG*, P*pbuG*, or P*pbuE*).

First, a *B. subtilis* integrative vector carrying the *lacZ* gene was constructed. The resulting plasmid pBGAB was then used to generate five final vectors where *lacZ* expression is under the control of the regulatory region of each purine riboswitch-controlled transcriptional unit. This was carried out by inserting in a translational fusion setup the full-length regulatory sequence (containing the corresponding promoter, riboswitch, plus PurBoxes for *purEKBCSQLFMNHD, xpt-pbuX,* and *pbuG*) upstream of *lacZ* employing scarless type IIS endonuclease-based cloning [85]. The resulting integrative plasmids pBGABpur, pBGABxpt, pBGABnupG, pBGABpbuG, and pBGABpbuE carrying *lacZ* fused to the regulatory sequence of the *pur* operon, the *xpt-pbuX* operon, the *nupG* gene, the *pbuG* gene, and the *pbuE* gene respectively, were used to incorporate each final construct into the *amyE* chromosomal locus of a *B. subtilis* 168 strain by Campbell-type recombination. Cells with successful integration were selected using antibiotic resistance, and proper integration into the chromosomal *amyE* locus was confirmed with alpha-amylase activity assays, resulting into the following *B. subtilis* reporter strains BGABpur, BGABxpt, BGABnupG, BGABpbuG, and BGABpbuE, which were respectively named after the corresponding plasmid used to generate each strain.

4.4.1 Characterization of the promoters that control transcription of purine riboswitches-regulated genes in *B. subtilis*

The first step towards studying the regulation exerted by purine riboswitches *in vivo* was to experimentally characterize the promoters controlling their expression in *B. subtilis* using the β-galactosidase reporter system described above. Therefore, the five generated test strains were cultivated in a minimal medium without purine supplementation for the four strains containing repressive riboswitches BGABpur, BGABxpt, BGABnupG, and BGABpbuG, or in a minimal medium supplemented with 3.7 mM adenine for the strain BGABpbuE, which contains the activating adenine-binding *pbuE* riboswitch. Cell extracts were prepared and β-galactosidase activity was measured (**Fig. 15**). With a β-galactosidase signal detected in each one of the samples, the results of the assay validate experimentally the presence of an active promoter upstream of each purine riboswitch. Interestingly, the intensity of the signal can vary greatly from one strain to another, showing that the promoters that control different genes and operons involved in purine metabolism and transport in *B. subtilis* can have highly different strengths. P*pur* and P*pbuG* (controlling expression of the purine biosynthesis operon and the guanine/hypoxanthine permease, respectively) represent the strongest promoters and display approximately a 12-fold higher activity than the second strongest promoter P*xpt-pbuX* (controlling expression of the xanthine phosphoribosyltransferase *xpt* and the xanthine permease *pbuX*) and approximately a 90-fold higher activity than P*nupG* (controlling expression of the purine nucleoside importer), which represents the weakest promoter.
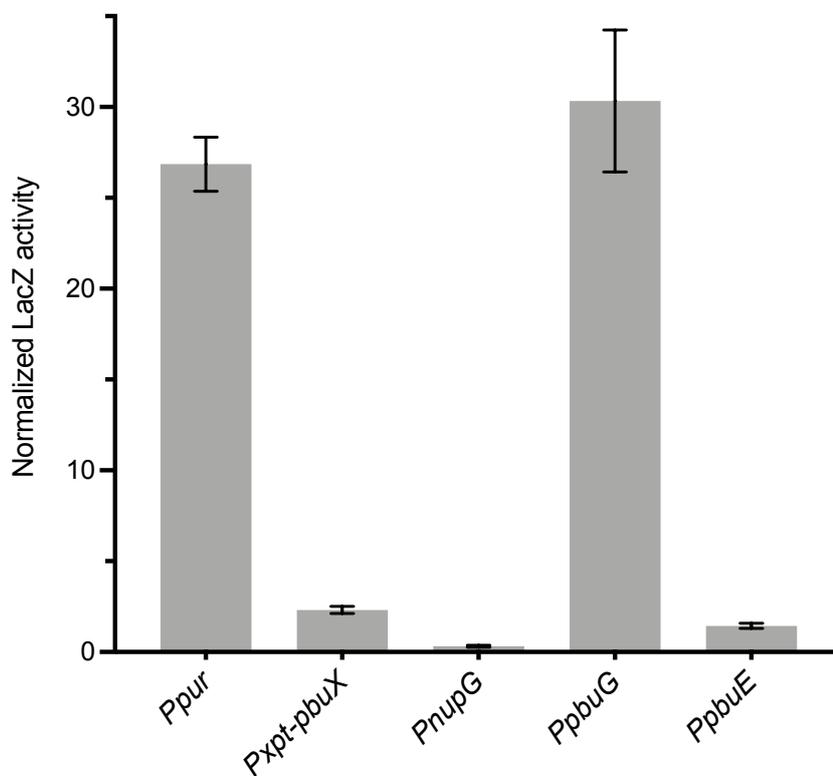
Figure 15 |Experimental validation and comparison of the activity of the promoters governing expression of the purine riboswitches-modulated genes and operons in *B. subtilis*. β-galactosidase activity measured in test strains containing the regulatory regions of purine riboswitch-regulated transcriptional units coupled to *lacZ* and integrated into the *amyE* chromosomal locus of *B. subtilis* 168. Cells were grown in a minimal medium without purine supplementation, except for the culture of the strain BGABpbuE (P*pbuE*), which was supplemented with 3.7 mM adenine to activate transcription. LacZ activity values expressed in enzyme unit per liter of culture were determined from cell extracts and normalized to the optical densities of corresponding cultures. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

4.4.2 **Expression of LacZ in the reporter strains is strongly modulated in response to guanosine or adenine supplementation**

To assess the effect of purine metabolites on the constructed reporter systems *in vivo*, the five reporter strains were cultivated in a minimal medium with and without purine supplementation. The supplemented medium contained 1 mM guanosine for the four strains BGABpur, BGABxpt, BGABnupG, and BGABpbuG containing guanine-sensing riboswitches, or 3.7 mM adenine for the strain BGABpbuE containing the adenine-sensing riboswitch. Measurements of β-galactosidase activity from cell extracts show that the constructed *in vivo* system is highly responsive to purine metabolites (**Fig. 16**). In line with the type of purine riboswitch and the function of regulated genes in *B. subtilis*, high guanosine levels lead to a downregulated expression of the *lacZ* genes fused to the repressive guanine-binding riboswitches (*pur* operon, *xpt-pbuX*, *nupG*, and *pbuG* riboswitches), which natively regulate genes involved in purine biosynthesis (the *pur* operon), salvage (the *xpt* gene*)* and uptake (the *pbuX*, *nupG*, and *pbuG* genes) in *B. subtilis*, whereas high adenine levels upregulate expression of the *lacZ* gene fused to the activating adenine-binding *pbuE* riboswitch, which natively regulate the *pbuE* gene involved in purine export.
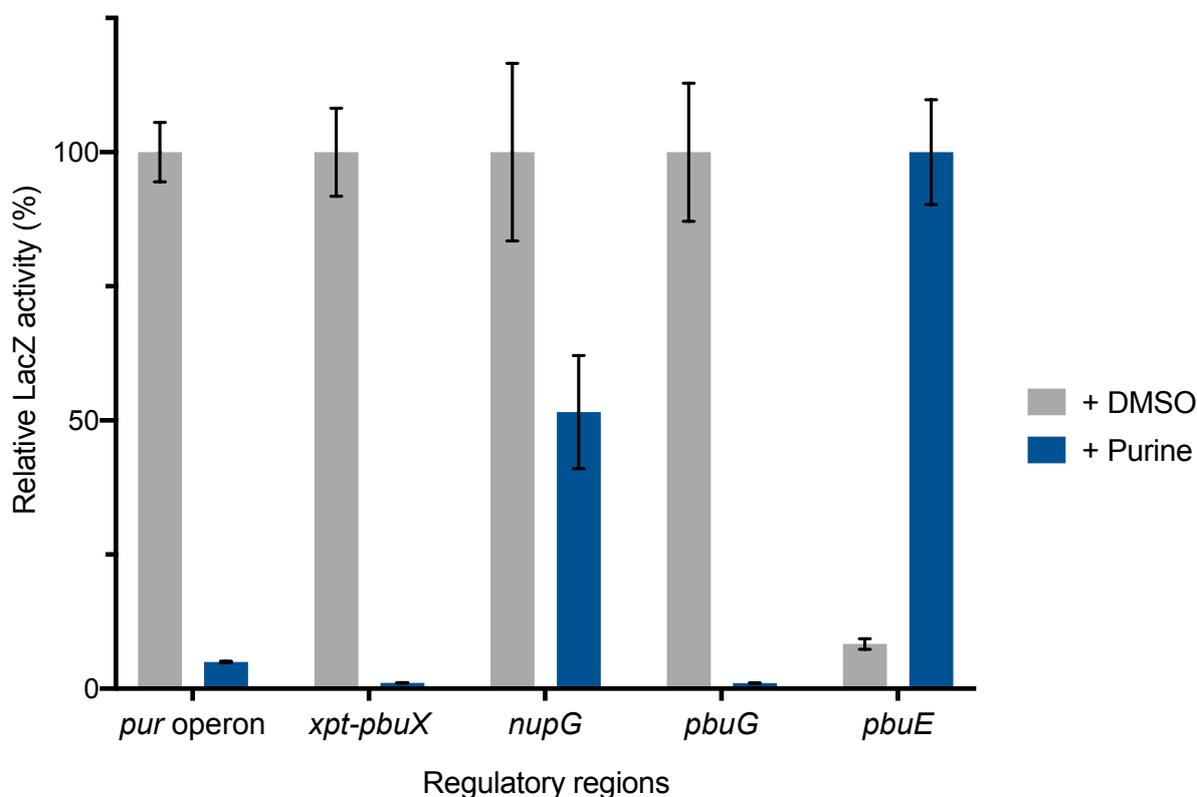
Figure 16 | LacZ expression in the generated reporter strains is highly responsive to purine metabolites supplementation. β-galactosidase activity measured in test strains containing the regulatory region of purine riboswitch-regulated transcriptional units coupled to *lacZ*. Cells grown in a minimal medium without purine supplementation (+ DMSO) or with 1 mM guanosine supplemented in the cultures of BGABpur, BGABxpt, BGABnpuG, and BGABpbuG, and 3.7 mM adenine supplemented in the culture of BGABpbuE strain. LacZ activity values, expressed as a percentage of the activated state for each construct, were determined from cell extracts and normalized to the optical densities of the cultures. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

With a minimum of 50% decrease in β-galactosidase activity in response to purine supplementation, or to its absence in the case of BGABpbuE, and over 91% decrease for four reporter strains out of five including BGABxpt and BGABpbuG that display up to 99% repression, the data show that purine biosynthesis, salvage, and transport genes in *B. subtilis* are highly responsive to purines and tightly regulated within a broad range of expression levels. Regulation of the *pur* operon is of the highest interest for the metabolic engineering of purines as it encodes the twelve enzymes responsible for purine *de novo* biosynthesis in *B. subtilis*. The guanosine-triggered 95% downregulation of LacZ expression in BGABpur indicates that a 20-fold increase in the expression of purine biosynthetic enzymes is potentially achievable in *B. subtilis*, showing how promising it is to successfully impair the negative feedback modulation exerted on the purine biosynthetic operon. Actually, the potential increase is likely even higher than 20-fold as even in the absence of purine supplementation, a certain level of *lacZ* repression should be exerted by the guanine (and potentially other purine metabolites [5]) synthesized *de novo* by the strain.

4.4.3 Mutations of PurR binding sites (PurBoxes) lead to an increase in LacZ expression and allow to study the activity of the *pur* operon riboswitch

The purine-mediated negative feedback repression of the *pur* operon, as well as the *xpt-pbuX* operon and the *pbuG* gene, is exerted by both of the PurR repressor and their respective riboswitch in *B. subtilis* (see Fig. 2a). To be able to exclusively measure the modulation exerted by the *pur* operon riboswitch, the two regulatory PurBoxes present in the original wild-type fusion in pBGABpur (see section 4.4) were mutated, leading to the plasmid pBGAB11. In this plasmid, the distal PurBox was deleted whereas the proximal PurBox was disrupted by introducing a point mutation at the level of a nucleotide that is conserved among all PurBoxes identified in *B. subtilis* (**Fig. 17**). This point mutation was shown in a previous study to abolish the activity of PurR on the *glyA* gene [21]. Integration of pBGAB11 into the *B. subtilis* 168 chromosome generated the strain BGAB11 (mutated PurBoxes in the fusion *pur* operon regulatory sequence−*lacZ*).

```
                Distal PurBox              Proximal PurBox

GATT AAATCCGTATGTTA - 16nt - AAATATTCGGATTT TGGGG-  8nt-ATG  purR

GAAT GGAAGCGAACGAAT - 17nt - TAATGTTCGGATTT ACAAT- 78nt-ATG  purA

TGTA AAACACGAACATTA - 16nt - TATCGTTCGATAAT ATCGT-274nt-ATG  purE
                                       G
CTTG AAATACGAATGATA - 16nt - TAAAGTTCGGGAAT TTTTA-230nt-ATG  xpt

TGTT TATTACGAACAAAA - 16nt - TATTGTTCGCTTTT TGTAT-270nt-ATG  yebB

AATA AATTCCGAACTTTA - 17nt - TAATATTCGTTTTT ACCAA-108nt-ATG  glyA

TATA TAAGGCGAACATTT - 16nt - AAATATTCGTTTTT AGGAG-  8nt-ATG  yumD

CGTG AAATCCGAATAATC - 16nt - AAATATTCGGTAAT AGGGT-100nt-ATG  yqhZ

CAAA TAAAACGAATAATA - 17nt - AAACGTTCGTAATT GGAGG-  5nt-ATG  ytiP

    -A---CGAA-----            -A---TTCG----T  conserved nucleotides
```

Figure 17 | Alignment of the tandem PurBox sequences identified upstream of promoters governing the expression of the nine PurR-regulated genes and operons in _B. subtilis_. To impair the binding of PurR to the PurBoxes upstream of the promoter P_pur_ (_purE_), the distal PurBox (shown in a red box) was fully deleted and a conserved nucleotide within the proximal PurBox was substituted (C−>G shown in red) in the _pur_ operon reporter plasmid pBGABpur, leading to pBGAB11. Boxed sequences are individual PurBox sequences. Shaded positions indicate nucleotides that diverge from the 5'-AWWWCCGAACWWTH-3' consensus sequence as defined by Kilstrup and coworkers [106]. Letters in the two bottom boxes show nucleotides that are conserved in the tandem PurBox motif. Lightface letters indicate nucleotides that are conserved in eight of the nine regulatory regions, and boldface letters indicate nucleotides that are conserved in all of them. Figure adapted from Saxild et al. (2001) [21].

LacZ activity was measured using cell extracts of BGAB11 (mutated PurBoxes in the fusion *pur* operon regulatory sequence−*lacZ*) cultures grown in a minimal medium without purine supplementation or supplemented with 1mM guanosine (**Fig. 18**). As a result of the impairment of PurR activity on the promoter P*pur*, an increase in LacZ expression is observed in BGAB11 when compared to BGABpur (wt *pur* operon regulatory region) both in the absence of purine supplementation (2.6-fold increase in LacZ activity) and in the presence of 1 mM supplemented guanosine (1.8-fold increase). The resulting response to guanosine supplementation in BGAB11, represented by 96.6% repression, corresponds to the modulation exerted by the *pur* operon riboswitch solely, allowing to test *in vivo* the activity of designed sRNAs on the riboswitch regulating purine *de novo* biosynthesis in *B. subtilis*.
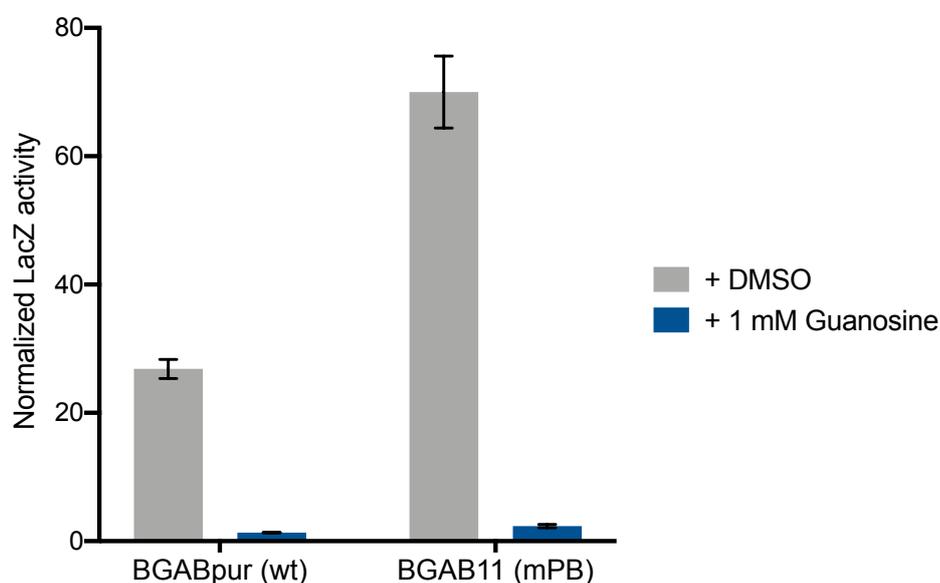
Figure 18 | Mutations of the *pur* operon PurBoxes lead to an increase in LacZ expression and reveal the modulation exerted by the *pur* operon riboswitch. β-galactosidase activity measured in test strains containing wild-type (BGABpur) or mutated (BGAB11) PurBoxes in the *pur* operon regulatory region coupled to *lacZ*. PurBox mutations consisted in the complete removal of the distal PurBox and the introduction of a point mutation in the proximal PurBox (see Fig. 17) to prevent PurR binding. Cells grown in a minimal medium without purine supplementation (+ DMSO) or with 1 mM supplemented guanosine. LacZ activity values, expressed in enzyme unit per liter of culture, were determined from cell extracts and normalized to the optical densities of corresponding cultures in the exponential phase of growth. mPB: mutated PurBoxes. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

4.5 Testing sRNAs activity on the *pur* operon riboswitch in *B. subtilis* reporter strains

To evaluate the functional activity of the designed small RNAs on the *pur* operon riboswitch in *B. subtilis*, DNA fragments corresponding to the small RNAs were inserted into the *B. subtilis* episomal vector pHT01 [79], generating a set of sRNA expression plasmids (pHTsRNAx). Thereafter, the reporter strain BGAB11, generated and characterized in the previous section, was transformed with the resulting plasmids, leading to a set of test strains (BGABsRNAx) co-expressing a small RNA and the *pur* riboswitch-modulated LacZ reporter (see Fig. 14).

4.5.1 Designed sRNAs are sustainably expressed in *B. subtilis,* but they have no activity on the *pur* operon riboswitch

First, proper expression of different sRNAs in their corresponding test strain was checked at different growth phases as follows. Total RNA was extracted at 4, 6, and 8 hours post-inoculation from the test strains BGABsRNA30, BGABsRNA70, BGABsRNA71$_{pur}$, and BGABsRNA110, which were cultivated in a minimal medium supplemented with 1 mM IPTG for the induction of sRNA synthesis. Thereafter, Northern blot analysis of total RNA was performed using labelled probes specifically targeting the sRNAs or the *lacZ* transcript as a control. The results show that the sRNAs are sustainably expressed in *B. subtilis*, with different expression levels depending on the sRNA and on the growth phase (**Fig. 19**).
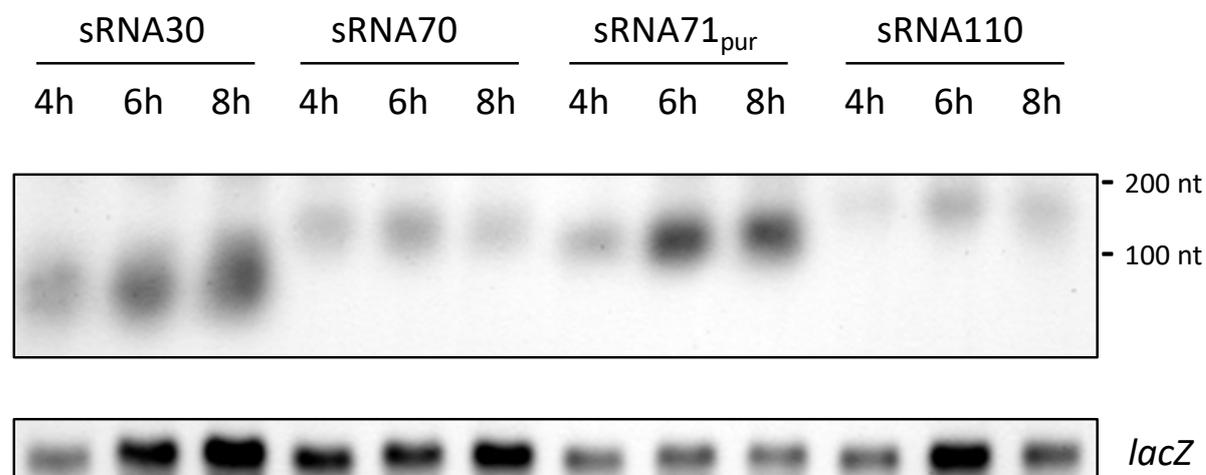
Figure 19 | Designed sRNAs are sustainably expressed in *B. subtilis*. Northern blot analysis of total RNA extracted at 4, 6, and 8 hours post-inoculation from *B. subtilis* test strains co-expressing the respective sRNA and a *pur* riboswitch-modulated LacZ. Cells were grown in a minimal medium with 1 mM IPTG for the induction of sRNA expression. Double digoxigenin-labeled DNA probes were used for detection of the sRNAs or the *lacZ* transcript. Expected sRNA sizes: approximately 90, 130, 130, and 170 nucleotides for sRNA30, sRNA70, sRNA71$_{pur}$, and sRNA110 respectively.

Following validation of the sRNA expression system, the activity of the designed sRNAs on the *pur* operon riboswitch was tested in the presence of high purine levels by cultivating the corresponding test strains (BGABsRNAx) in a minimal medium supplemented with guanosine and measuring the activity of LacZ, whose expression is modulated by the *pur* operon riboswitch. In addition to the control test strain BGABsRNA$_{random}$ expressing a random sequence, an additional control strain containing the empty pHT01 vector (BGABempty) was included in the experiment, and it was cultivated both with and without purine supplementation. Comparison of LacZ activity in the test strains expressing the sRNAs with the control strains shows that the designed sRNAs do not impair the repressive function exerted by the *pur* operon riboswitch in response to purine (**Fig. 20**). Indeed, the designed sRNAs have no significant effect on the purine-binding riboswitch, which still downregulates LacZ expression to similar levels in the absence and in the presence of the designed sRNAs. The absence of activity is observed for both types of sRNAs with regard to the design strategy; for sRNA30, sRNA70, and sRNA110, which were designed against purine aptamers consensus sequence and structure (see Fig. 5 and Fig. 6) as well as for sRNA30$_{pur}$ and sRNA71$_{pur}$, which were designed specifically against the sequence of the *pur* operon aptamer (see Fig. 7).
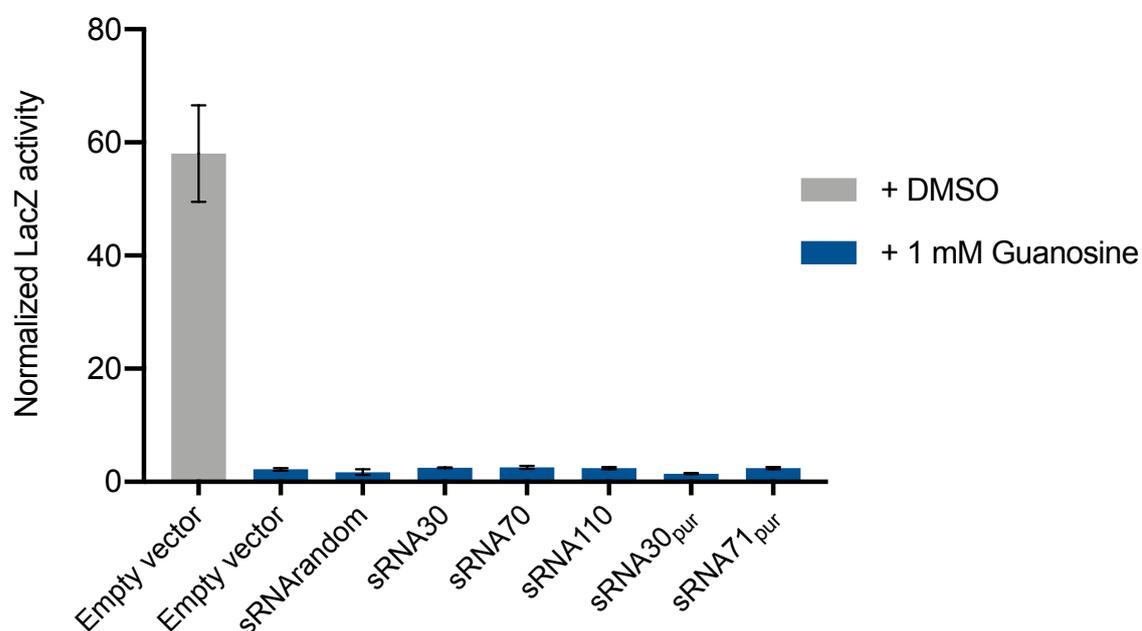
Figure 20 | Designed sRNAs do not impair the purine-triggered downregulation exerted by the *pur* operon riboswitch in *B. subtilis*. β-galactosidase activity measured in the test strains BGABsRNAx co-expressing the riboswitch-modulated LacZ reporters and the sRNAs that were designed using either the consensus sequence and structure of the five purine aptamers (sRNA30, sRNA70, and sRNA110) or using specifically the sequence of the *pur* operon aptamer (sRNA30pur and sRNA70pur). Cells grown in a minimal medium supplemented with 1 mM guanosine. The control strain transformed with the empty pHT01 vector was also grown without guanosine supplementation. LacZ activity values, expressed in enzyme unit per liter of culture, were determined from cell extracts and normalized to the optical densities of corresponding cultures. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

Although the small RNAs were designed to not bind to the aptamer's 3' end sequence that is involved in forming the antiterminator in the ON conformation, their binding to upstream nucleotides would affect the structure of the aptamer. Therefore, this strategy comes with a risk of affecting mRNA stability, for example by exposing an RNase site that would have otherwise been sterically inaccessible or protected via intra-aptamer base pairing. However, as the sRNA-aptamer interaction would still form complex secondary structures, and as different sRNAs target different regions within the aptamer, it is unlikely to observe mRNA degradation with all tested sRNAs. Another possible consequence of sRNA binding to the riboswitch could be the impairment of other potential interactions in *trans* with regulatory factors similarly to the characterized interaction of the aptamer of the *rib* operon riboswitch with the transcription activator RibR in *B. subtilis* [24, 25]. Still, there is no evidence for the presence of such a regulation at the level of the *pur* operon riboswitch targeted here. Additionally, no significant difference was measured in β-galactosidase activity between controls (empty vector and sRNA$_{random}$) and designed sRNA samples, indicating that the sRNAs might simply not be able to bind to the *pur* aptamer in the presence of purine.

4.5.2 The modulation exerted by the *pur* operon riboswitch is not impaired by the perfectly complementary antisense RNA neither

Activity of the antisense RNA asRNA*pur* on the *pur* operon riboswitch was tested in the LacZ reporter strain BGABasRNA*pur*. Unlike the sRNAs tested above, asRNA*pur* is perfectly complementary to the aptamer of the *pur* operon riboswitch (see Fig. 8). As a control, the activity of asRNA*nupG* was also parallelly tested on the *pur* operon riboswitch in the reporter strain BGABasRNA*nupG*; asRNA*nupG* being perfectly complementary to the aptamer of the *nupG* riboswitch but only partially complementary to the aptamer of the *pur* operon riboswitch. Both strains were cultivated, together with BGABsRNA$_{random}$ expressing a random small RNA sequence, in a minimal medium with or without guanosine supplementation. The resulting LacZ activity values measured using cell extracts show that both of the perfectly complementary asRNA*pur* and the partially complementary asRNA*nupG* do not affect the expression of the riboswitch-regulated LacZ, neither in the presence of supplemented guanosine nor in its absence (**Fig. 21**). The absence of effect of asRNA*pur* in the absence of purine supplementation demonstrates that despite being fully complementary to the *pur* aptamer, asRNA*pur* does not affect the stability of its target transcript. However, the absence of effect in the presence of guanosine supplementation shows that despite the perfect complementarity, asRNA*pur* is not functional on the *pur* operon riboswitch, similarly to the designed partially complementary sRNAs tested in the previous section.
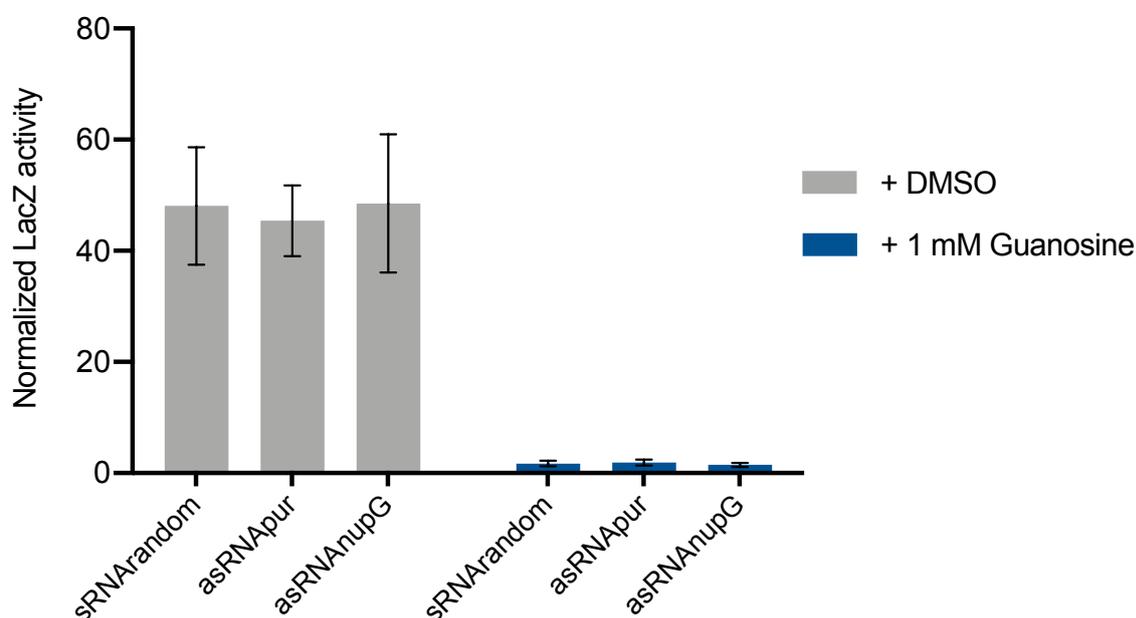
Figure 21 ǀ The antisense RNA perfectly complementary to the *pur* aptamer does not affect negatively the expression of target mRNA, but it does not impair the regulatory function exerted by the *pur* operon riboswitch. β-galactosidase activity measured in the test strains BGABasRNApur and BGABasRNAnupG co-expressing the *pur* operon riboswitch-modulated LacZ and asRNAs perfectly complementary to the *pur* aptamer (asRNApur) or to the *nupG* aptamer (asRNAnupG). The strain BGABsRNA_random expressing a random RNA sequence was used as a control. Cells grown in a minimal medium without purine supplementation (+ DMSO) or with 1 mM supplemented guanosine. LacZ activity values, expressed in enzyme unit per liter of culture, were determined from cell extracts and normalized to the optical densities of corresponding cultures. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

4.6 Kinetics of sRNA-aptamer interactions

Considering the kinetic regulation model of transcriptional riboswitches [12-15], kinetic requirements and limitations were evaluated for sRNAs to effectively compete with the ligand guanine and hybridize in a co-transcriptional manner with the nascent *pur* operon aptamer. In this model, a functional sRNA-aptamer interaction is limited to the time that the RNA polymerase (RNAP) takes to transcribe the riboswitch. During this process, two different windows of time could be distinguished: a first window during which only the sRNA could bind to the nascent aptamer; it starts when the seed sequence is transcribed up to the moment when the aptamer moiety is fully synthesized and folded. During the second window of time, the sRNA competes with the ligand; it starts when the aptamer domain is folded and lasts until RNAP reaches the regulatory decision point to continue or abort transcription.

4.6.1 Pre-folding sRNA−aptamer hybridization

In the present design strategy, sRNAs are able to start interacting with the aptamer as soon as the complementary seed sequence in its 5' end is transcribed. The time frame during which the ligand is not yet able to compete with the sRNA lasts for the time that RNAP spends to progress to the end of the aptamer domain, plus the time theoretically required for its folding. However, a recent study modeling the co-transcriptional folding kinetics of purine-binding aptamers showed that the ligand binding pocket structure is formed as soon as the aptamer domain is transcribed [16]. This limits the binding advantage that sRNAs have, to the time of aptamer's transcription. Different studies reported a rate of transcription by bacterial RNAP ranging

from 15 to 90 nucleotides per second [107-112]. If we base the model on the slowest reported speed, the maximum time required for transcription of purine-binding aptamers is approximately 6 seconds.

Based on a recent study in which hybridization and melting rate constants were measured at low concentrations for oligonucleotides of various lengths and under various experimental conditions [113], hybridization rate constants for sRNA-aptamer interaction at 30 °C are estimated to be in the 10^4 M^{-1} s^{-1} range at the highest.

Employing RNA-Seq transcriptome analysis, the expression level of the random sequence $sRNA_{random}$ expressed by $pHTsRNA_{random}$ in *B. subtilis* 168 was measured, and its proportion relatively to total mRNAs was quantified. The absolute concentration of total mRNAs was calculated using a total number of mRNAs of 10300 molecules per *B. subtilis* cell [114], and a cell volume of 1.13 μm^3 in the late exponential growth phase [115], resulting in an sRNA concentration of 1.7×10^{-6} M (**Table 16**).

Table 16 | Quantification of sRNA expression in *B. subtilis*[a] using RNA-Sequencing.

	Normalized reads (RPKM)	Ratio $sRNA_{random}$/mRNAs	Concentration (mol L^{-1})
$sRNA_{random}$	165599	0.11	**1.7×10^{-6}**
Total mRNAs	1498140	1	1.5×10^{-5}

[a]*B. subtilis 168* cultivated to the late exponential growth phase in LB medium.

Implementing the above values of hybridization rate constant (k_{hyb}) and sRNA concentration in the hybridization rate equation, the fraction of aptamer that hybridizes with the sRNA per second is obtained:

$$\text{Hybridization rate} = k_{hyb} \, [\text{sRNA}] \, [\text{aptamer}]$$

$$\text{Hybridization rate (M s}^{-1}) = 0.017 \, [\text{aptamer}]$$

Assuming that sRNA and aptamer concentrations are constants, during the 6 seconds timeframe of aptamer transcription:

$$[\text{sRNA-aptamer}] = 0.1 \, [\text{aptamer}]$$

Thereby, at the time when the aptamer becomes able to bind its ligand, the fraction hybridized with the sRNA is estimated to correspond to only 10% of total aptamer. The actual value should be even lower, as the slowest reported speed for bacterial RNAP was used, and it was assumed that the hybridization is irreversible and that reactants concentrations are constants in order to obtain the highest realistic approximation of the hybridized fraction.

4.6.2 Post-folding competitive binding

The time available for sRNAs to interact with the aptamer is typically longer after its folding because of RNAP pause sites present in the expression platform moiety of transcriptional riboswitches [13, 116-120]. The major limitation during this phase lies in the kinetic competition between the sRNA and the ligand for binding to the aptamer. Intracellular guanine concentration previously measured in exponentially growing *E. coli* is 1.9×10^{-4} M [121]. Assuming comparable guanine concentration in *B. subtilis*, it is two orders of

magnitude higher than the calculated sRNA concentration (see Table 16). Therefore, to be able to kinetically compete with guanine, sRNAs would require hybridization rate constants that are several orders of magnitude higher than the association rate constant of guanine. Such k_{hyb} values are likely unattainable for sRNA-aptamer hybridizations as the 10^4 M^{-1} s^{-1} range estimated in the previous section is even lower than the association rate constant of the guanine analog 7-deazaguanine binding to the *xpt-pbuX* riboswitch (2.1×10^5 M^{-1} s^{-1}) [14]. This indicates that once the aptamer is folded, it is quite unlikely for sRNAs to compete effectively with guanine, especially at physiologically high guanine intracellular concentrations, conditions under which sRNAs were expected to be functional.

5 Conclusion

The attempt to impair the regulation exerted by transcriptional riboswitches employing sRNAs was unsuccessful for the guanine-binding *pur* operon riboswitch in *Bacillus subtilis*. sRNA quantification and estimations of kinetic parameters from recent studies indicated that the reason could be a kinetic advantage, which guanine has over sRNAs. Similar results are anticipated for sRNAs targeting the FMN-binding *rib* operon riboswitch in *B. subtilis,* as both the concentration and the association rate constant values for FMN (5.4×10^{-5} M and 4.2×10^{5} M^{-1} s^{-1} respectively) [14, 121] are higher than the respective measured concentration and estimated k_{hyb} range for sRNA binding (see section 4.6.1). Nevertheless, the approach could be effective on other transcriptional riboswitches whose ligands are less abundant in the cell and/or that have lower association rate constants.

CHAPTER II. Rational engineering of transcriptional

riboswitches employing CRISPR-Cas9

6 Results and Discussions

6.1 Structural analysis of the *xpt-pbuX* riboswitch from *B. subtilis* and design of a rational engineering approach

The purine-binding riboswitch that controls transcription of the *xpt-pbuX* operon in *B. subtilis* is one of the most studied riboswitches at both structural and functional levels [5, 26, 99]. Thus, it served as a model in the present study to establish a rational editing approach for engineering transcriptional riboswitches (**Fig. 22**). The aim consists in keeping the riboswitch constitutively ON by exclusively preventing formation of the termination hairpin, while concurrently promoting formation of the antiterminator and preserving the structure. According to the structures of the ON and OFF conformations (**Fig. 22a**), this objective should be achievable via the sole deletion of the 3' strand of the stem within the termination hairpin (**Fig. 22a**, red-boxed sequence). Indeed, these key nucleotides are exclusively involved in forming a part of the terminator element. As a consequence of their deletion, the unpaired 5' strand of the former stem becomes free to interact with a complementary upstream sequence (**Fig. 22a**, sequences in blue) and to engage thereby into forming the antitermination fold within a largely intact riboswitch. To identify the target nucleotides to delete for engineering other

transcriptional riboswitches, secondary structures of terminator regions are easily predictable using RNA structure prediction programs because of the simple stem-loop structure of the terminator, and no further structural information is required.

The strategy was applied in separate sets of experiments to two riboswitches without available crystal structures and belonging to different classes; the guanine-binding riboswitch that controls transcription of the purine biosynthetic genes *purEKBCSQLFMNHD* (*pur* operon riboswitch) and the FMN-binding riboswitch that controls transcription of the riboflavin biosynthetic genes *ribDEAHT* (*rib* operon riboswitch) in *B. subtilis*. Structures of the terminator regions were predicted using the NUPACK program [102] (**Fig. 22b**). The identified target nucleotides are highlighted as red-boxed sequences within termination hairpins in Figure 22b. In the case of the *pur* operon riboswitch, 18 nucleotides were predicted to constitute the 3' strand of the terminator's stem (**Fig. 22b**, left panel), whereas in the case of the *rib* operon riboswitch, 14 nucleotides were identified (**Fig. 22b**, right panel).
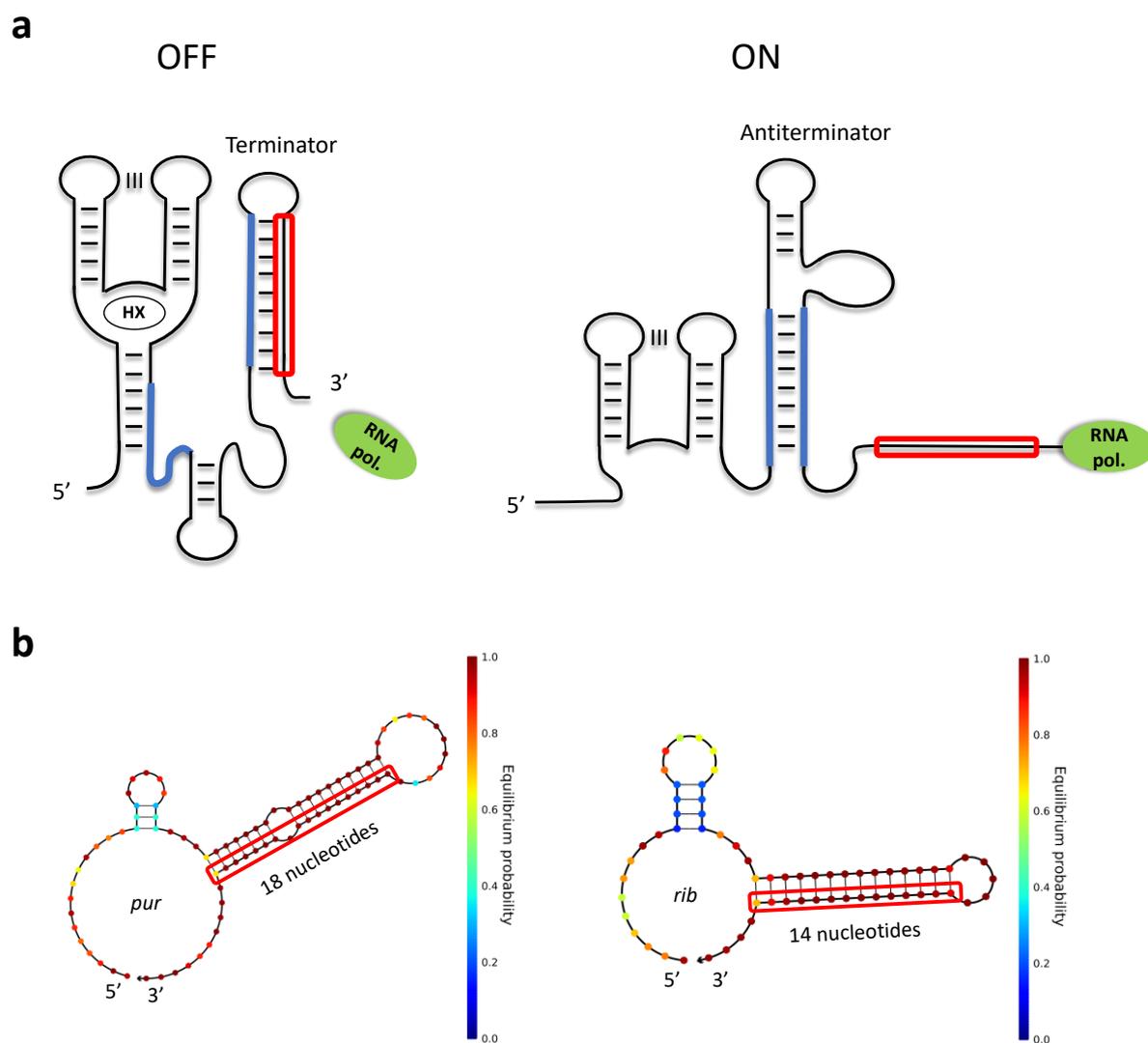
Figure 22 | A general approach for engineering transcriptional riboswitches to enhance gene expression.

(**a**) Riboswitch-mediated modulation of transcription: a schematic representation of structures adopted by the wild-type *xpt-pbuX* riboswitch from *B. subtilis* in its OFF and ON conformations. The scheme is drawn based on the published crystal structure [99]. Deletion of the red-boxed sequence disrupts the formation of the terminator solely and releases its stem's 5' strand, which is then free to engage into forming the antiterminator via the interaction with the complementary sequence shown in blue. Consequently, formation of the terminator is prevented and the ON conformation with intact structure is promoted. The key tertiary

interactions that form a pseudoknot are schematically shown as three vertical lines (III). HX: hypoxanthine. RNA pol.: RNA polymerase.

(**b**) Identification of target nucleotides to delete within terminators of transcriptional riboswitches. The secondary structures were predicted using the NUPACK program [102]. Left: predicted structure of the termination region of the guanine-binding *pur* operon riboswitch from *B. subtilis*. Right: predicted structure of the termination region of the FMN-binding *rib* operon riboswitch from *B. subtilis*. Red-boxed sequences represent the target nucleotides to be deleted. Bases are shown as dots whose colors stand for the probability to be in the displayed paired or unpaired state at equilibrium. Bars indicate hydrogen bonds.

6.2 Rational engineering of the *pur* operon riboswitch from *B. subtilis*, but not its deletion, alleviates gene repression in a reporter system

As a proof of concept with regard to the riboswitch engineering strategy described in the previous section, effects of riboswitch engineering and riboswitch deletion were measured employing β-galactosidase reporter systems based on the test strains BGABpur and BGAB11, which were previously generated to test sRNA activity (chapter I, section 4.4.3). BGABpur contains the wild-type regulatory region of the *pur* operon (containing the two regulatory PurBoxes, the endogenous promoter P*pur*, and the riboswitch, see Fig. 2a) translationally fused to the *lacZ* gene, whereas the PurBoxes are mutated in the reporter construct in BGAB11. Consequently, *lacZ* expression in BGABpur is controlled by both of the repressor PurR and the guanine-binding riboswitch, whereas in BGAB11, PurR activity on *lacZ* expression is impaired, and the sole activity of the riboswitch is measurable (see Fig. 18). To test the effect of the complete removal of the *pur* operon riboswitch on *lacZ* expression, the riboswitch in the plasmid pBGAB11 (used to generate the strain BGAB11) was deleted. Integration of the resulting construct pBGAB12 in the *amyE* locus of *B. subtilis* 168 generated the reporter strain BGAB12. A final plasmid was generated by deleting the key sequence identified in Figure 22b (left panel) from the riboswitch sequence in pBGAB11, resulting in the plasmid pBGAB13, which was then integrated in the chromosome of *B. subtilis* 168 to generate the reporter strain BGAB13. *lacZ* expression in BGAB13 is thus controlled by an engineered *pur* operon riboswitch.

The generated reporter strains BGAB12 (mutated PurBoxes, deleted riboswitch) and BGAB13 (mutated PurBoxes, engineered riboswitch) were cultivated in a minimal medium with and without 1 mM guanosine supplementation. Cell extracts were prepared, and β-galactosidase activity was determined and compared with the activities measured in BGABpur (wild-type regulatory region) and BGAB11 (mutated PurBoxes) (**Fig. 23**). LacZ expression decreased drastically in BGAB12 relatively to BGAB11, both in the absence (99.7% decrease in LacZ activity) and in the presence (94% decrease) of guanosine supplementation, to much lower levels than in BGABpur in which both the PurR repressor and the repressive riboswitch are active. These data show that the complete deletion of the *pur* operon riboswitch leads to drastically reduced levels of expression of the downstream gene, indicating that the presence of this riboswitch might be essential for maintaining physiological expression levels of the *pur* operon in *B. subtilis*. The effect of deletion of the *pur* operon riboswitch observed in the reporter strain BGAB12 is similar to the effects previously observed upon deletion of the *rib* operon riboswitch in a riboflavin overproducer *B. subtilis* strain [29]. In the presence of 1 mM supplemented guanosine, riboswitch engineering in BGAB13 led to a 14-fold increase in LacZ activity when compared to BGAB11 and to an overall increase of 25-fold when compared to BGABpur. Moreover, an increase in LacZ activity in BGAB13 compared to BGAB11 (1.6-fold) and to BGABpur (4.1-fold) is observed even in the absence of guanosine supplementation. These data indicate that the response of the engineered riboswitch to its ligand was effectively reduced and demonstrate the efficiency of the developed approach in activating expression of riboswitch-controlled genes.
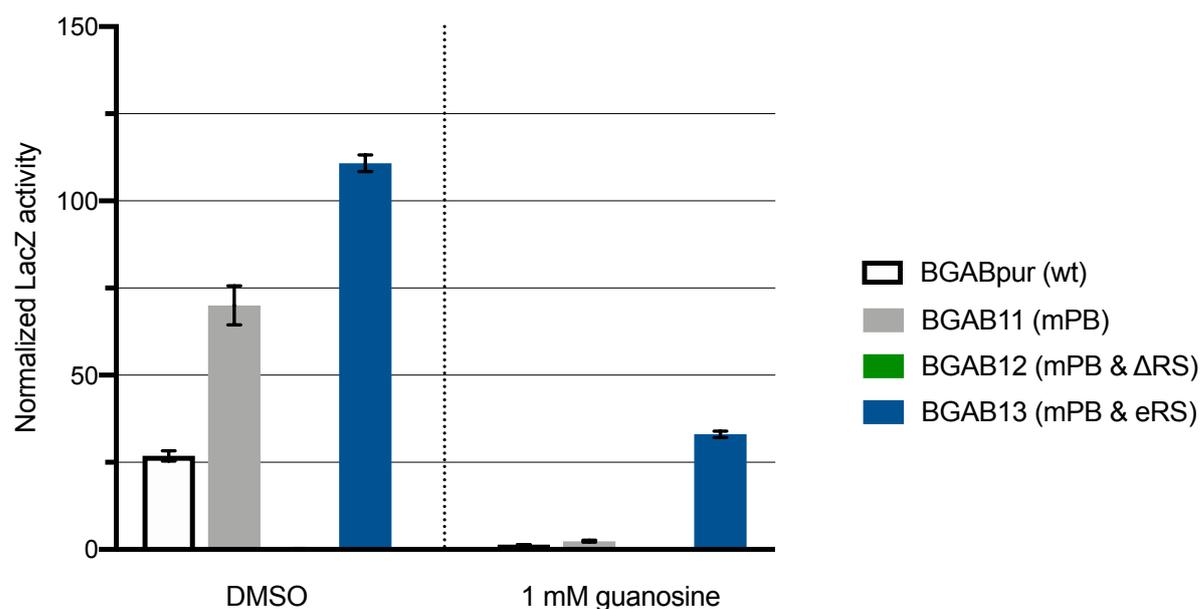
Figure 23 | Proof of concept: effect of *pur* operon riboswitch deletion or its rational engineering on gene expression in a *B. subtilis* reporter system. The reporter strains BGABpur (wt *pur* operon regulatory sequence–*lacZ* fusion), BGAB11 (mutated PurBoxes), BGAB12 (mutated PurBoxes and deleted riboswitch), and BGAB13 (mutated PurBoxes and engineered riboswitch) were cultivated in a minimal medium without purine supplementation or with 1 mM guanosine supplemented. Cell extracts were prepared in the exponential phase of growth, and LacZ activity values expressed in enzyme unit per liter of culture were determined and normalized to the optical densities of the cultures. The change in LacZ expression in BGABpur upon addition of 1 mM guanosine is governed by the repressor PurR and the purine riboswitch, whereas this change in BGAB11 displays the sole activity of the riboswitch, PurR binding to the PurBoxes being impaired. Gene expression collapsed upon removal of the riboswitch in BGAB12. Riboswitch engineering in BGAB13 increased gene expression in the absence of purine supplementation and reactivated expression in the presence of 1 mM guanosine to levels even higher than in the wild-type BGABpur in the absence of purine supplementation. Interestingly, LacZ expression is still affected to a certain degree by guanosine supplementation in BGAB13. mPB: mutated PurBoxes. ΔRS: deleted riboswitch. eRS: engineered riboswitch. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

Interestingly, comparison of LacZ activity in BGAB13 between the presence and the absence of guanosine supplementation shows that there is still a certain level of modulation of LacZ expression despite the mutation of PurR binding sites (distal PurBox deletion and proximal PurBox point mutation, see Fig. 17) and the engineering of the *pur* operon riboswitch, which is no more able to form the stem of the terminator in response to ligand binding. The same experiment performed in a *purR* deletion mutant showed similar results (**Fig. S1**), confirming that the remaining modulation of LacZ expression in BGAB13 does not result from PurR residual binding to the proximal PurBox. These data support the existence of another purine-responsive regulation, which is independent from PurR and from the formation of the known terminator of the riboswitch for the modulation of *pur* operon's expression in *B. subtilis*. Possible mechanisms include a ligand-influenced mRNA stability, metabolite-regulated interaction with transcription factors, or existence upstream of the targeted terminator of a second terminator that is active at high ligand concentrations as suggested for the *rib* operon riboswitch by Mironov et al. (2008) [30].

6.3   Genome editing in *Bacillus subtilis* employing the CRISPR-Cas9 system

The previous section shows that expression of a riboswitch-regulated reporter gene collapses upon riboswitch deletion, whereas the implementation of the riboswitch engineering approach developed here leads to the expected upregulation of expression of a reporter gene. The following experiments were carried out to confirm these effects via the deletion or engineering of native chromosomal riboswitches and measurement of the expression of their regulated genes. Furthermore, the consequent effects on bacterial physiology were investigated to figure out whether the riboswitch engineering approach developed here affects the levels of metabolites in a bacterial model and whether it is effective in improving performance of bacterial production strains. Two different *B. subtilis* riboswitches were either deleted or engineered following our strategy, the *pur* operon riboswitch regulating purine biosynthesis (see Fig. 2a) and the *rib* operon riboswitch regulating riboflavin biosynthesis (see Fig. 2b). In all engineered purine-related strains generated here, the repressor PurR was first deleted to prevent its interference. The modifications of the *B. subtilis* 168 chromosome were carried out employing the CRISPR-Cas9 genome editing technology.

The CRISPR-Cas9 constructs generated in the present study were based on the pJOE8999 vector that was developed by J. Altenbuchner [82] (**Fig. 24**). pJOE8999 is a 7.8 kb plasmid that can be propagated in *E. coli* via the pUC origin of replication and in *B. subtilis via* the thermo-sensitive replicon PE194. The vector encodes a kanamycin resistance gene functional in both bacteria.

The endonuclease Cas9 from *S. pyogenes* is driven by the P*manP* mannose-inducible promoter and the single guide RNA (sgRNA) is constitutively expressed *via* the semisynthetic promoter P*vanP**. Spacers that guide the complex to the chromosomal targets were cloned into the vector between BsaI sites as described by Altenbuchner (2016) [82]. For cloning the homology repair templates, the described sfiI-based ordered gene assembly in *B. subtilis* (OGAB) method did not work. Instead, a restriction site-independent method based on overlap extension PCR and referred to as circular polymerase extension cloning (CPEC) [86] was used (**Fig. 25**). Alternatively, when the overlaps contained strong secondary structures, classical cloning using the XbaI and SalI sites present in the vector was combined with the scarless Golden Gate assembly method [85]. Both strategies allowed the assembly and cloning of homology templates in a single step.
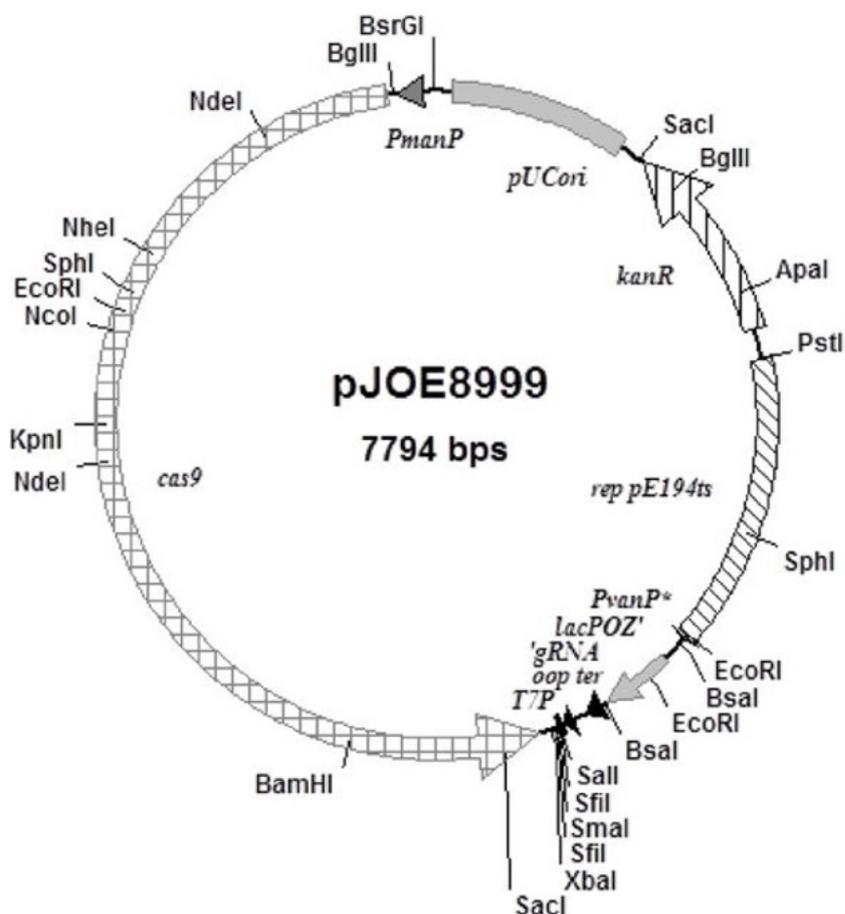
Figure 24 | CRISPR-Cas9 vector pJOE8999. Physical map of the vector pJOE8999 containing the pUC18 minimal origin, the temperature-sensitive replication origin from pE194ts, a kanamycin resistance gene (*kanR*), *cas9* under the control of the P*manP* promoter, the sgRNA transcribed from the semisynthetic promoter P*vanP** interrupted by the *lacZ α* fragment (l*acPOZ'*), the *λ oop* terminator, and the T7 promoter (T7P). Figure reproduced from Altenbuchner (2016) [82].
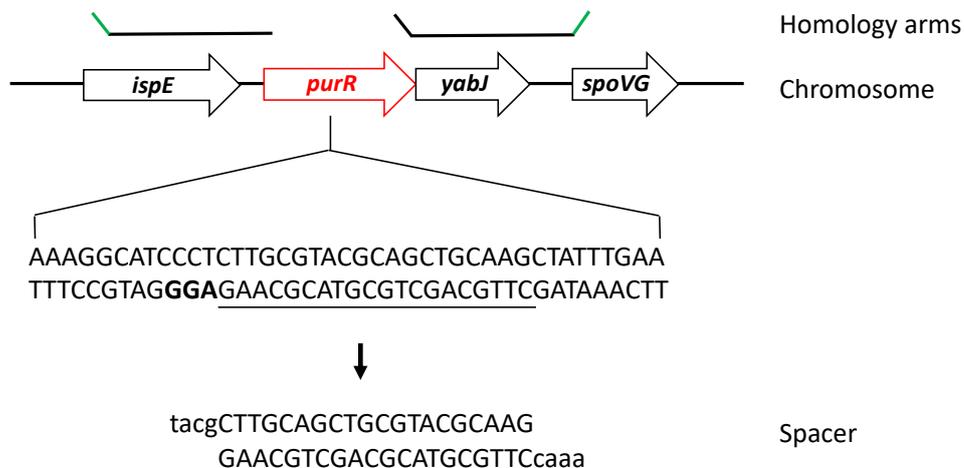
Figure 25 | Design and cloning strategy of the spacer and the homology repair template used for CRISPR/Cas9-based deletion of *purR* from the chromosome of *B. subtilis*. Schematic representation of *B. subtilis* chromosomal region carrying the *purR* gene. The PCR-amplified homology arms flanking the fragment to be deleted are depicted as horizontal lines in the top of the figure, with additional sequences representing overlaps (in green for overlaps with the ends of a SmaI-linearized pJOE8999, and in black for an inter-arms overlap) designed for CPEC assembly [86]. The sequence of the selected protospacer is underlined, and the protospacer adjacent motif (PAM) is depicted in bold. The double-stranded oligonucleotide fragment at the bottom shows the sequence of the spacer inserted into pJOE8999, with lowercase letters indicating the sticky ends fitting to the BsaI-digested sites in pJOE8999.

Based on the described pJOE8999 vector, five constructs were generated to carry out the following chromosomal modifications: 1) Deletion of a 799 bp sequence representing most of the *purR* gene to abolish PurR repression of the promoter P*pur*. 2) Deletion of an 18 bp sequence corresponding to the rational editing of the *pur* operon riboswitch (see Fig. 22). 3) Deletion of the 198 bp fragment representing the full-length *pur* operon riboswitch. 4) Deletion of a 14 bp sequence corresponding to the rational editing of the *rib* operon riboswitch (see Fig. 22). 5) Deletion of the 258 bp fragment representing the full-length *rib* operon riboswitch. The two arms of each homology repair template (HT) were generated by amplification of 700-800 bp DNA fragments flanking each target sequence and assembled into pJOE8999, resulting respectively in the plasmids pPurRHT, peRSpurHT, pdRSpurHT, peRSribHT, and pdRSribHT. For deletion of *purR*, 5 spacer sequences were selected for their high ON-target activity out of the 67 possible spacers upstream of the protospacer adjacent motifs (PAM) NGG found within the target sequence. For the editing of the *pur* operon riboswitch, 2 PAM options are available within the targeted 18 bp sequence, and for the editing of the *rib* operon riboswitch, 3 PAM options are available (**Fig. 26**). The spacer candidates were discriminated according to their predicted target site specificity in *B. subtilis*, and the most promising spacer for each of the three targeted chromosomal regions was synthesized and inserted into the corresponding plasmid(s) pPurRHT, peRSpurHT, pdRSpurHT, peRSribHT, and pdRSribHT (see above). The resulting final genome editing plasmids were named pABpurR (*purR* knockout vector), pABeRSpur (*pur* operon riboswitch editing), pABdRSpur (*pur* operon riboswitch deletion), pABeRSrib (*rib* operon riboswitch editing), and

pABdRSrib (*rib* operon riboswitch deletion).

pur riboswitch   **rib riboswitch**

...GCTTCCTCATAAGTGCAATGCAGAG**CGG**GTATTTTTT...   ...AAAGCCCCGAATTTTTTATAAATTC**GGG**GCTTTTTTGACG...   Chromosome
...CGAAGGAGTATTCACGTTACGTCTCGCCCATAAAAAA...   ...TTTCGGGGCTTAAAAAATATTTAAGCCCCGAAAAAACTGC...

    |||||||||||||||||||| ||||||||||||||||||||

    CTCATAAGTGCAATGCAGAG   CCCGAATTTTTTATAAATTC   Spacer

    |||||||| | |||||||||||||| |

...GCTTCCTCATAAGTTTTTTATTTTCTGAAAA...   ...AAAGCCCCGAATTTTTTATTTTGACGGTAAA...   Homology
...CGAAGGAGTATTCAAAAATAAAAGACTTTT...   ...TTTCGGGGCTTAAAAAATAAAACTGCCATTT...   template

Figure 26 | Design of spacers used for CRISPR/Cas9-based editing of the *pur* operon riboswitch (left) and the *rib* operon riboswitch (right) in the chromosome of *B. subtilis*. Alignment of wild-type chromosomal sequences in the regions carrying the target nucleotides to be deleted (red-boxed sequences) with designed spacers and homology repair templates. Chosen PAM sequences (NGG) are depicted in bold.

Engineering of purine biosynthesis in *B. subtilis* 168 was performed in a stepwise manner; first, a *purR* knockout strain was generated employing pABpurR. Then, the *pur* operon riboswitch in the resulting BSpurR strain was either edited using pABeRSpur, leading to eBSpur2 strain, or deleted using pABdRSpur, leading to the strain BSdRSp. Engineering of riboflavin biosynthesis in the strain 168 was performed by editing the *rib* operon riboswitch employing pABeRSrib, resulting in the strain eBSflv. Deletion of the *rib* operon riboswitch in 168 was carried out using pABdRSrib, resulting in the

126

strain BSdRSr. The overall curing efficiency was 95% (out of 66 analyzed colonies, 3 of them still contained a plasmid following heat treatment). 57 transformants were analyzed with regard to their target loci using colony PCR followed by gel electrophoresis and DNA sequencing (**Fig. S2**). With 35 mutants having the expected chromosomal modification, the overall genome editing efficiency was 61%.

In J. Altenbuchner's study [82], the efficiencies in deleting a 25.1 kb and a 4.1 kb sequences using pJOE8999-based vectors were 89% and 97% respectively. The observed differences in the efficiency of genome editing could originate from many factors including the size of deleted sequences, gRNA activity, and loci accessibility. Indeed, despite an overall better chromosomal loci accessibility in bacteria than in eukaryotic cells, some regions are still more or less accessible than others for Cas9 within the bacterial genome [122].

6.4 Chromosomal deletions of the riboswitches that control expression of purine and riboflavin biosynthesis operons in *B. subtilis* lead to auxotrophic strains

Construction of the *B. subtilis* strains BSdRSp (*purR* knockout, *pur* operon riboswitch deleted) and BSdRSr (*rib* operon riboswitch deleted) is described in the previous section. Both strains were cultured in a minimal medium and growth was monitored by measurement of optical densities at 600 nm over time (**Fig. 27**). After 14 hours of incubation, no growth was observed for

BSdRSp and BSdRSr cultures. Growth could be recovered upon supplementation of the medium with 500 µM guanosine and 500 µM adenosine in the case of BSdRSp cultures (**Fig. 27a**), or 20 µM riboflavin in the case of BSdRSr cultures (**Fig. 27b**), showing that BSdRSp is auxotrophic for purines and BSdRSr is auxotrophic for riboflavin.

These results are perfectly in line with the collapse in reporter gene (*lacZ*) expression observed upon deletion of the *pur* operon riboswitch (Fig. 23) and with the decrease in *rib* genes expression and riboflavin yield upon deletion of the *rib* operon riboswitch in a riboflavin overproducer *B. subtilis* strain [29]. Furthermore, the present data show that riboswitches represent essential elements for *de novo* biosynthesis of purines and riboflavin in *B. subtilis*, indicating that these riboswitches might not only attenuate transcription but are possibly involved in additional processes hypothetically related to transcript stability [31-34] or to interactions with protein regulators [24, 25]. Finally, these results demonstrate that riboswitches controlling purine and riboflavin biosyntheses represent excellent targets for antibacterial therapy [123-126] as their deletion leads to auxotrophic strains.
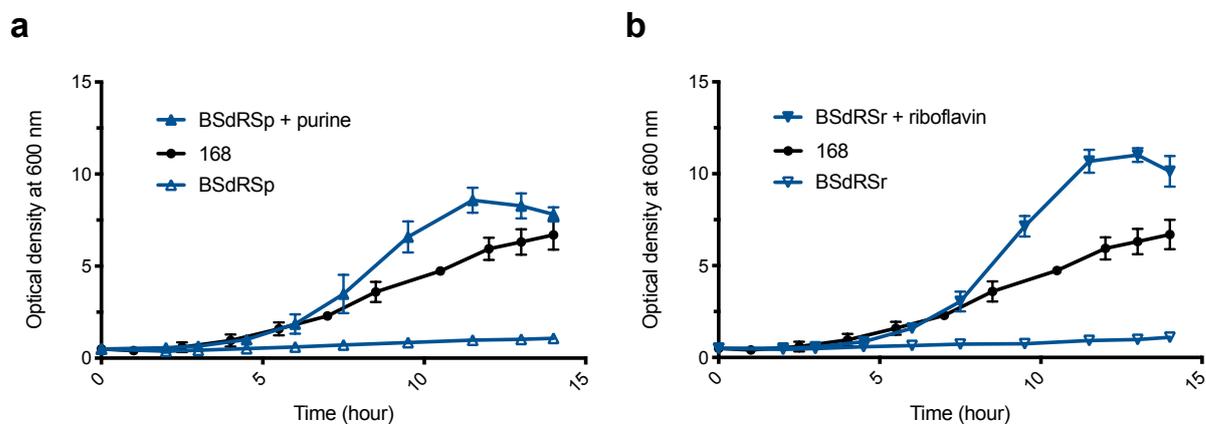
Figure 27 | Deletion of riboswitches controlling purine or riboflavin biosynthesis leads to auxotrophic strains.

(**a**) Growth of a *B. subtilis* strain lacking the *pur* operon riboswitch (BSdRSp) in a minimal medium, and effect of supplementation with purines (guanosine and adenosine).

(**b**) Growth of a *B. subtilis* strain lacking the *rib* operon riboswitch (BSdRSr) in a minimal medium, and effect of supplementation with riboflavin.

The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

6.5 Riboswitch editing activates transcription of purine and riboflavin biosynthetic genes and enhances metabolite production in *B. subtilis*

The following experiments were carried out to measure the effects of the riboswitch engineering strategy developed here on expression of regulated biosynthetic operons and on production of corresponding metabolites in *B. subtilis* strains edited with CRISPR-Cas9 (see section 6.3).

In a first analysis, relative transcript levels of the regulator gene *purR* and the first gene of the *pur* operon *purE* were measured by RT-qPCR in the strains BSpurR (*purR* knockout) and eBSpur2 (*purR* knockout, *pur* operon riboswitch edited) (**Fig. 28a**). As expected, there was no detectable *purR* transcript in both *purR* knockout strains, confirming the efficiency of CRISPR-Cas9 in deleting the *purR* gene (**Fig. 28a**, left). This deletion led to a 9.8-fold increase in transcript levels of *purE* in BSpurR relatively to the wild-type control strain 168 (**Fig. 28a**, right), revealing the extent to which the repressor PurR affects the activity of the promoter P*pur* in *B. subtilis*. Rational editing of the *pur* operon riboswitch led to an additional 9.2-fold increase in *purE* transcript levels in eBSpur2 when compared to BSpurR, resulting in a total enhancement of transcription by 90-fold in eBSpur2 relatively to the parent strain 168. These results demonstrate the efficacy of the implemented engineering strategy in activating transcription of riboswitch-controlled genes. To determine whether the observed increase in *purE* transcript levels was translated into an increase in the production of purine metabolites, the fusion construct *pur* operon regulatory sequence−*lacZ* in pBGAB11 (mutated PurBoxes, wild-type *pur*

130

operon riboswitch, see section 4.4.3) was employed as an intracellular genetic sensor that binds the purine nucleobase guanine and potentially xanthine and hypoxanthine [5]. Therefore, measured LacZ activity levels are inversely proportional to purine levels. The genetic sensor was integrated in the *amyE* locus of BSpurR and eBSpur2 and used to assess relative purine levels in the respective resulting strains BGBSpurR and BGeBSpur2 (**Fig. 28b**). In accordance with RT-qPCR data, the gradual decrease in LacZ activity observed respectively in BGBSpurR and BGeBSpur2 relatively to the unedited control strain BGAB11 (generated by integration of pBGAB11 into the chromosome of 168, see section 4.4.3) indicates that intracellular levels of purine increased upon *purR* deletion and upon riboswitch editing.

In similar experiments, relative transcript levels of *ribD*, the first gene of the *rib* operon, were determined by RT-qPCR in the strain eBSflv (*rib* operon riboswitch edited) and the wild-type control strain 168. The data show that the transcript level of *ribD* is 6.7-fold higher in eBSflv relatively to its level in 168 (**Fig. 28c**), showing that the rational editing of the *rib* operon riboswitch effectively activated transcription of the *rib* operon. Total production of riboflavin and its derivatives FMN and FAD was quantified in cultures of eBSflv and 168 strains using high performance liquid chromatography (HPLC) (**Fig. 28d**). The data show that the observed 6.7-fold increase in *ribD* transcript levels was translated into a remarkable 29-fold enhancement of riboflavin production and also into an increase in FMN levels. This second example further demonstrates the potency of the developed strategy and its applicability to different riboswitch classes for the metabolic engineering of various transcriptional riboswitch-controlled pathways.
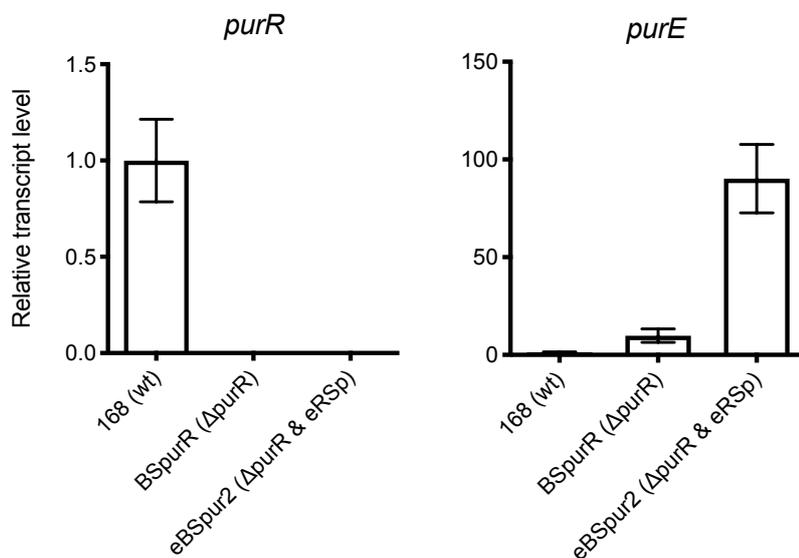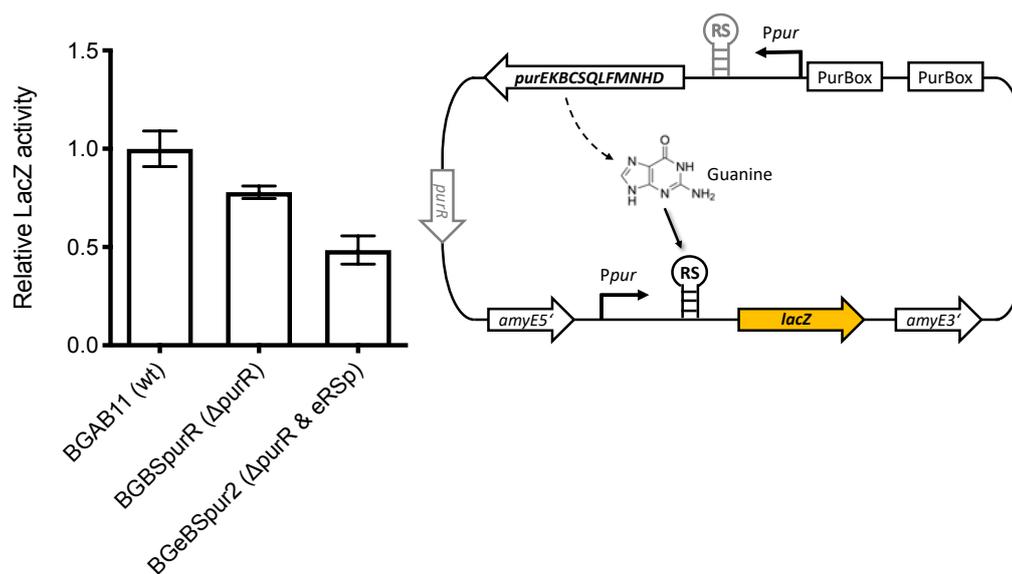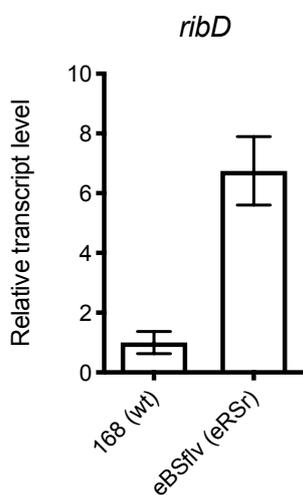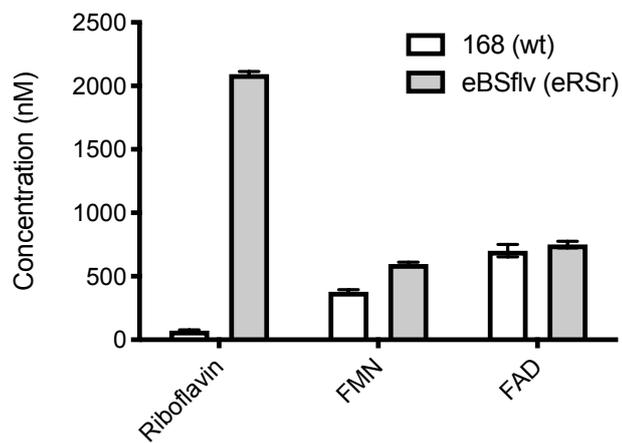
Figure 28 | Rational editing of transcriptional riboswitches leads to enhanced transcript and metabolite levels in *B. subtilis*. The edited strains were grown in minimal media, and transcript and metabolite levels were determined by RT-qPCR and a metabolite sensor or HPLC.

(**a**) Relative transcript levels of the repressor gene *purR* and the purine biosynthesis gene *purE* measured in the engineered strains BSpurR (*purR* knockout) and eBSpur2 (*purR* knockout, edited *pur* operon riboswitch "eRSp") in comparison to the *B. subtilis* 168 wild-type.

(**b**) Assessment of purine production in BGBSpurR and BGeBSpur2 relatively to the control strain BGAB11 using a purine-binding genetic sensor based on the *pur* operon riboswitch coupled to the reporter *lacZ* and integrated into the *amyE* chromosomal locus. Intracellular levels of guanine (and potentially xanthine and hypoxanthine) are inversely proportional to β-galactosidase activity. P*pur*: endogenous promoter controlling expression of the *pur* operon, RS: wt *pur* operon riboswitch. RS in grey: edited endogenous *pur* operon riboswitch. *purR* grey arrow: deleted regulatory gene.

(**c**) Relative transcript levels of the riboflavin biosynthesis gene *ribD* determined in the engineered *B. subtilis* strain eBSflv (edited *rib* operon riboswitch "eRSr") in comparison to the *B. subtilis* 168 wild-type.

(**d**) Total levels of riboflavin and its derivatives FMN and FAD in the wild-type strain and eBSflv. Metabolite concentrations were measured following cell lysis in the culture.

The values represent averages ± standard deviations of data obtained from a minimum of three independent cultures.

6.6 *purR* deletion and *pur* operon riboswitch editing lead to enhanced growth of *B. subtilis* 168 and of the riboflavin overproducer BSRF

To find out whether the genetic modifications and the resulting changes in metabolite levels negatively affect growth of *B. subtilis*, the edited strains were cultivated in a minimal medium and growth was monitored by measuring the OD at 600 nm (**Fig. 29a**). Interestingly, in comparison to the wild-type strain 168, higher cell densities were reached for both purine production strains eBSpurR (*purR* knockout) and eBSpur2 (*purR* knockout, *pur* operon riboswitch edited), as well as for the riboflavin production strain eBSflv (*rib* operon riboswitch edited), though to a lesser extent.

Considering the central role of purine nucleotides in paramount cellular processes such as DNA replication, transcription, translation, and energy metabolism, it is very likely that the observed influence of purine pathway engineering on growth resulted from the increased intracellular purine levels (see Fig. 28b). This is in line with the study performed by Samant et al. (2008) in which they showed that nucleotide biosynthesis represents the most critical function for bacterial proliferation in a medium with limited levels of nucleotide bases [127]. Therefore, the presented modifications of the purine pathway (deletion of *purR* and editing of the *pur* operon riboswitch) could be beneficial for industrial bioprocesses based on *B. subtilis.* Engineering of the riboswitch that modulates riboflavin biosynthesis also led to enhanced growth, albeit to a lower degree (**Fig. 29a**). Similarly to purine, it is not unlikely that this effect resulted from the increased flavin levels (see Fig. 28d) considering the central

role in energy metabolism of flavoproteins, which depend on the riboflavin-derived cofactors FMN and FAD. Nevertheless, existence of other mechanisms involving the riboswitch cannot be excluded.

   With the aim of exploiting further the remarkable growth-promoting feature of the presented engineering of the purine pathway, *purR* was deleted and the *pur* operon riboswitch was edited in the industrial strain BSRF employing the CRISPR-Cas9 vectors pABpurR and pABeRSpur (see section 6.3) in a stepwise manner, leading to the strain eBSRFpur2. BSRF is a recombinant *B. subtilis* 168 derivative that was engineered to overproduce riboflavin. In BSRF, the endogenous promoter P*rib* controlling expression of the *rib* operon is replaced by a strong promoter, and mutations are introduced in the *rib* operon riboswitch, in *ribC* (encoding the FMN producing flavokinase) and in *tkt* (encoding a transketolase of the hexose monophosphate shunt). As a side effect of these modifications, growth of BSRF is relatively impaired. Growth of the original strain BSRF, the modified strain eBSRFpur2, as well as the wild-type strain 168 were monitored by measurement of optical densities at 600 nm over time (**Fig. 29b**). Growth curves show that the introduced modifications (stepwise deletion of *purR* and editing of the *pur* operon riboswitch) boosted growth by 53% after 16 hours of incubation, allowing eBSRFpur2 cultures to reach even higher cell densities than the cultures of the wild-type strain 168.
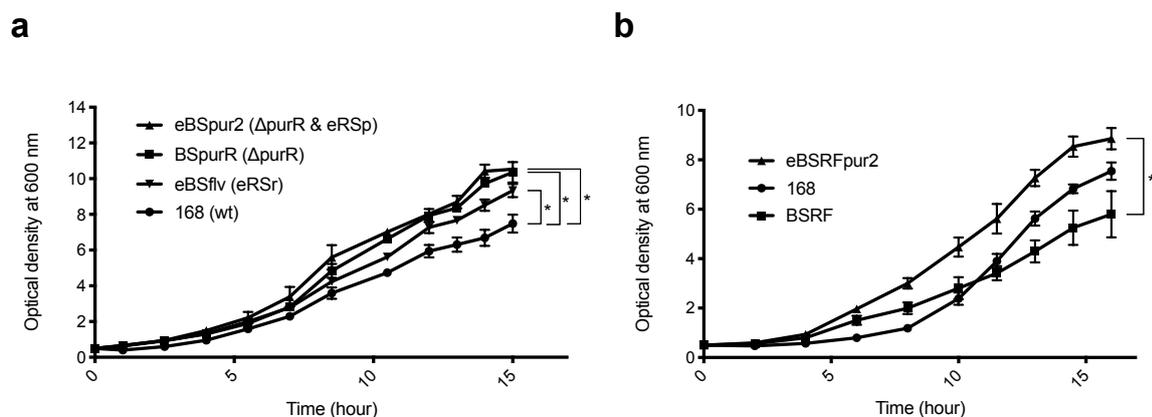
Figure 29 | *purR* deletion and *pur* operon riboswitch editing lead to enhanced growth in *B. subtilis*. The generated purine and riboflavin producer strains were grown in a minimal medium, and growth was monitored by OD measurement at 600 nm.

(**a**) Growth of the edited strains BSpurR (*purR* knockout), eBSpur2 (*purR* knockout, edited *pur* operon riboswitch "eRSp"), and eBSflv (edited *rib* operon riboswitch "eRSr") compared to the growth of the wild-type strain *B. subtilis* 168. The edited strains grow to a higher cell density.

(**b**) Influence of the double genetic modification *purR* deletion and *pur* operon riboswitch editing on growth of the industrial riboflavin overproducing strain BSRF. The edited strain eBSRFpur2 grows to a higher cell density than BSRF, reaching even higher densities than the original 168 strain.

The values represent averages ± standard errors of the means of data obtained from three independent cultures. *p < 0.05, Student's t test.

To determine whether the improved growth of eBSRFpur2 compared to BSRF was translated into an increase in riboflavin levels in the culture, total production of riboflavin was measured by HPLC (**Fig. 30**). With 44% increase in riboflavin levels in eBSRFpur2 cultures when compared to the cultures of BSRF, the increase in cell density resulted indeed in an increase in riboflavin production in the culture.
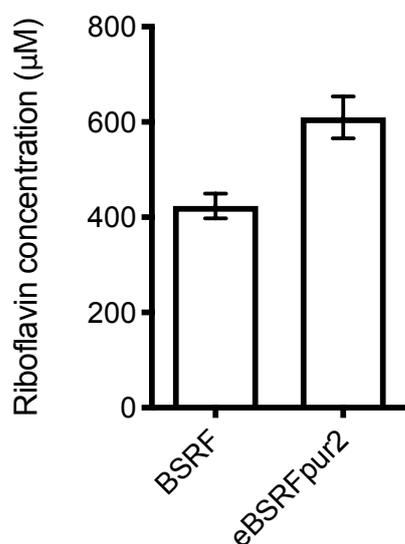
Figure 30 | **Effect of the enhanced growth of BSRF on total production of riboflavin.** Riboflavin concentration was measured following cell lysis in the culture. The values represent averages ± standard deviations of data obtained from a minimum of two independent cultures.

7 Conclusions

Despite the characterized role of riboswitches in preventing constitutive expression of regulated genes and operons, the present study shows for two transcriptional riboswitches controlling major biosynthetic pathways in *Bacillus subtilis* that their removal from the chromosome does not result in transcription activation, but rather to a drastic decrease in expression that led to auxotrophic strains. An alternative approach based on the rational engineering of riboswitches was developed to alleviate transcription attenuation without affecting other potential structure-related functions.

In contrast to many other industrial strains, the strains generated from *B. subtilis* 168 wild-type have no further genetic modifications and thus can be employed as a starting point for further development with regard to production of different purine and flavin-derived compounds. Alternatively, the generated CRISPR-Cas9 constructs could be directly used to implement the engineering conducted here into existing *B. subtilis* industrial strains with the aim of increasing metabolite production and/or enhancing growth.

With thousands of novel riboswitch classes expected to be discovered in the future [4], the versatility and straightforwardness of the developed approach in combination with CRISPR-Cas9 genome editing make the presented

procedure a promising tool to rapidly engineer newly discovered transcriptional riboswitches in *B. subtilis.* Furthermore, as transcriptional riboswitches have common mode of actions among different species [7], the present engineering strategy can be implemented in other organisms.

Finally, considering the role that riboswitches can play in stabilizing their transcripts [31-34], it might be worthwhile assessing the effect of rationally edited riboswitches as potential stabilizing elements when introduced in the 5' ends of unstable mRNAs. Such a use would have a multitude of applications in biotechnology and synthetic biology.

8 Supplementary Figures

Figure S1. The residual responsiveness of the reporter system after PurBoxes mutations and *pur* operon riboswitch engineering (BGAB13, Fig. 23) is not caused by PurR residual activity. The *B. subtilis* mutant BSpurR (*purR* knockout) generated using CRISPR-Cas9 (section 6.3) was used to generate new reporter strains where it is sure that PurR activity is totally abolished unlike in the wild-type-based reporter strains BGAB11 and BGAB13 previously used in Figure 23 (where the PurBoxes were mutated but PurR was present in the cell and the possibility of its residual binding to the point-mutated proximal PurBox could not be excluded). Integration of the reporter plasmids pBGAB11 (mutated PurBoxes, wild-type *pur* operon riboswitch) and pBGAB13 (mutated PurBoxes, engineered *pur* operon riboswitch) in the chromosome of BSpurR resulted into BGAB11Δ*purR* and BGAB13Δ*purR* respectively. The resulting strains were cultivated in a minimal medium without purine supplementation or with 1 mM guanosine supplemented. Cell extracts were prepared in the exponential phase of growth, and LacZ activity values expressed in enzyme unit per liter of culture were determined and normalized to the optical densities of the cultures. A 16-fold increase in LacZ expression was measured upon the engineering of the riboswitch (LacZ activity in BGAB13Δ*purR* compared to the activity in BGAB11Δ*purR*) in the presence

of supplemented guanosine and a 1.8-fold increase was measured in the absence of supplementation. These results are similar to the previous data measured using BGAB11 and BAGB13 (14-fold and 1.6-fold increase with and without guanosine supplementation respectively, Fig. 23). The residual responsiveness to guanosine supplementation previously observed in BGAB13 despite PurBoxes mutations and riboswitch engineering is still present in BGAB13Δ*purR* despite the additional deletion of *purR*. This result shows that PurR was not responsible for the residual downregulation of LacZ expression in response to guanosine supplementation in BGAB13.

mPB: mutated PurBoxes. eRS: engineered riboswitch. The values represent averages ± standard deviations of data obtained from three independent cultures.
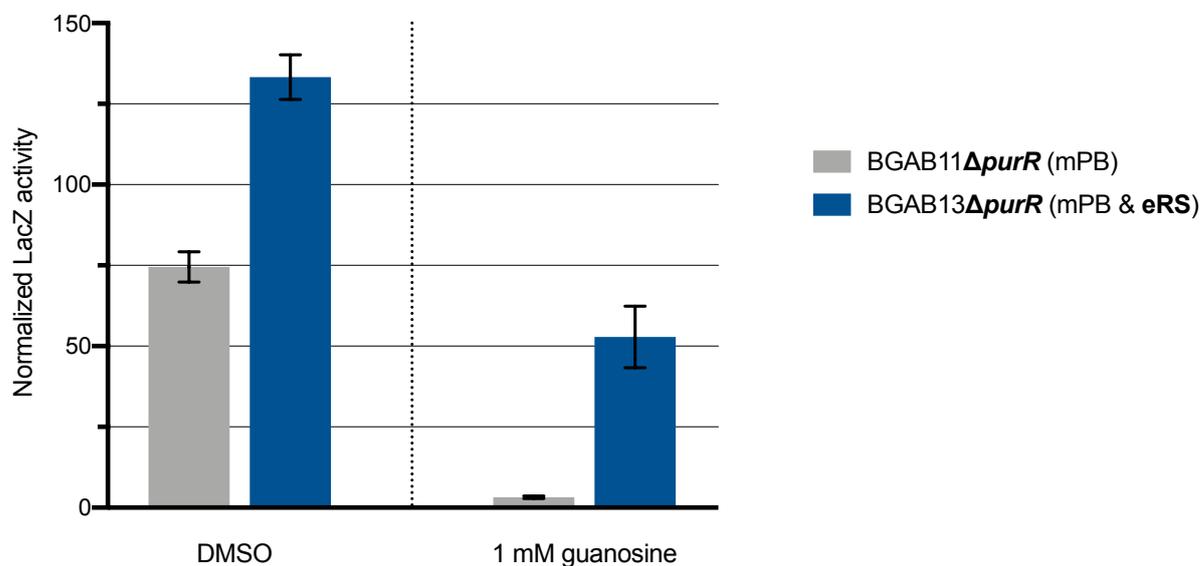
Figure S2. Analysis of *B. subtilis* transformants generated employing the CRISPR-Cas9 system.

(**a**) Confirmation of *purR* deletion using agarose gel electrophoresis.

(**b**) Confirmation of *pur* operon (upper panel) and *rib* operon (lower panel) riboswitches deletions by DNA sequencing of riboswitch regions.

(**c**) Confirmation of *pur* operon (upper panel) and *rib* operon (lower panel) riboswitches editing by DNA sequencing of target regions.

Chromosomal regions around *purR*, *pur* operon riboswitch, and *rib* operon riboswitch were amplified in colony PCR reactions using primers that bind outside the homology template sequences.

a

L   Δ*purR*   *wt*

(kb)
10
8
6
5
4
3.5
3
2.5
2
1.5

1

0.75

0.5

0.25

Δ 799 bp

b

168 genome TGTCCGTTGTTAAGATAAACATGAAATCAAAACACGACCTCATATAATCTTGGGAATATGGCCCATAAGTTTCTACCCGGCAACCGTAAATTGCCGGACTATGCAGGAAAGTGATCGAT

dRSpur TGTCCGTTGTTAAGATAAAC

Trace data

AAAACTGACATGGATATATCGCAGAAGCGAACGACTGACGATACATGTACCATGCCCGGTTTGTATTGCTTCCTCATAAGTGCAATGCAGAGCGGGTATTTTTTATTTTCTGAAAACAA

TTTTTATTTTCTGAAAACAA

168 genome GGATCGCTATAATAACCAATAAGGACAAATGAATAAAGATTGTATCCTTCGGGGCAGGGTGGAAATCCCGACCGGCGGTAGTAAAGCACATTTGCTTTAGAGCCCGTGACCCGTGT

dRSrib GGATCGCTATAATAACCAAT

Trace data

GCATAAGCACGCGGTGGATTCAGTTTAAGCTGAAGCCGACAGTGAAAGTCTGGATGGGAGAAGGATGATGAGCCGCTATGCAAAATGTTTAAAAATGCATAGTGTTATTTCCTATT

GCGTAAAATACCTAAAGCCCCGAATTTTTTATAAATTCGGGGCTTTTTTGACGGTAAATAACAAAA

TTTGACGGTAAATAACAAAA

c

Spacer   Target sequence   PAM   QU35_RS03570

B.s 168 genome ACGACTGACGATACATGTACCATGCCCGGTTTGTATTGCTTCCTCATAAGTGCAATGCAGAGCGGGTATTTTTTATTTTCTGAAAACAAAAGCATTAGAAGGTGGGGAACAGAATGCAGCCGCTAGTAGC

eRSpur ACGACTGACGATACATGTACCATGCCCGGTTTGTATTGCTTCCTCATAAGT GTTTTTATTTTCTGAAAACAAAAGCATTAGAAGGTGGGGAACAGAATGCAGCCGCTAGTAGC

Trace data

QU35_RS12140   Target sequence   PAM   Spacer

B.s 168 genome ATACTCTTCCATTTGTTTCCCTCCCCTCTTTTGTTATTTACCGTCAAAAAAGCCCCGAATTTATAAAAAATTCGGGGCTTTAGGTATTTTTACGCAATAGGAAATAACACTATGCATTTTTAAACATTTT

eRSrib ATACTCTTCCATTTGTTTCCCTCCCCTCTTTTGTTATTTACCGTCAAAA TAAAAAATTCGGGGCTTTAGGTATTTTTACGCAATAGGAAATAACACTATGCATTTTTAAACATTTT

Trace data

143

Abbreviations

5'-UTR	5' untranslated region
Amp	Ampicillin
asRNA	Antisense RNA
B. subtilis	*Bacillus subtilis*
Cas	CRISPR associated protein
Cm	Chloramphenicol
CMM	C minimal medium
CPEC	Circular Polymerase Extension Cloning
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
tracrRNA	*trans*-activating CRISPR RNA
DIG	Digoxigenin
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid

Abbreviations

eRS*pur*	Edited *pur* operon riboswitch
eRS*rib*	Edited *rib* operon riboswitch
FAD	Flavin adenine dinucleotide
FMN	Flavin mononucleotide
GRAS	Generally recognized as safe
gRNA	Guide RNA
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-d-1-thiogalactopyranoside
Kan	Kanamycin
LB	Lysogeny broth
MFE	Minimum free energy
MOPS	3-(N-morpholino)propanesulfonic acid
mPB	Mutated PurBoxes
mRNA	Messenger RNA
OD	Optical density
ONPG	Ortho-Nitrophenyl-β-galactoside
PAM	Protospacer Adjacent Motif
PCR	Polymerase chain reaction
RBS	Ribosomal binding site

RNA   Ribonucleic acid

RNAP   RNA polymerase

rNTP   Ribonucleoside tri-phosphate

RPKM   Reads Per Kilobase of transcript, per Million mapped reads

rRNA   Ribosomal RNA

RT   Room temperature

RT-qPCR   Quantitative reverse transcription PCR

SD   Shine-Dalgarno

sgRNA   Single guide RNA

Spc   Spectinomycin

sRNA   Small RNA

wt   Wild-type

Acknowledgements

To our beloved friend Mario Lederle who unexpectedly left us before defending his doctoral thesis. Rest in peace my friend, you are and will always be in our hearts.

The number of persons who participated directly or indirectly to this work are uncountable. I would like to express my gratitude to every single person who played a positive role during my education and especially during the challenging period of my doctoral thesis.

First and foremost, I would like to thank my beloved family from whom I was far away during this period and thank them for their understanding for not being able to visit them often. I thank my inspiring father, mother, grand-mother, sisters, brothers-in-law, and nephews and nieces for their unconditional love and support.

Special thanks go to Prof. Matthias Mack for offering me the opportunity to pursue research projects in his lab, for the support, the trust, and the freedom that he granted me to implement my own ideas, for his kindness, and for being a great source of knowledge to me who knew little about the Microbiology field when I started my PhD.

I am very grateful to Prof. Georg Stöcklin for accepting to be my first referee, for his guidance, extremely helpful discussions, and genuine interest into my projects despite the difference in the fields. I would also like to thank him and his former doctoral student Dr. Marius Bruer for welcoming me in their lab to perform the RT-qPCR experiments and gladly offering me material and great

help.

I am very grateful to Prof. Andres Jäschke for accepting to take part of my Thesis Advisory Committee, for his precious input and for the easy and great human contact with him.

I am also grateful to Dr. Ilka Bischofs for helpful and stimulating discussions in the early stage of my PhD and for kindly accepting to join the examination committee.

Many thanks go to every member of the Mack lab including former and current doctoral students, postdocs, technical assistants, and undergraduate students. I would like to thank specifically my colleagues and friends Dr. Rodrigo Mora Lugo for providing *B. subtilis* genomic DNA and for being such an amazing listener and compassionate person and Benjamin Kachel for translating the abstract to the German language, for his trust, and for the exciting deep discussions we have about science and beyond.

I would like to thank my Graduate School HBIGS, especially Dr. Rolf Lutz, Martina Galvan, and Sandra Martini for the invaluable support that I received from them since the first day of my doctoral studies and for the great program, events, and courses they offer to us.

Last but not least, I am grateful to Prof. Jörn Kalinowski and Dr. Tobias Busche for the RNA-Seq collaboration, to Dr. Josef Altenbuchner for providing the plasmid pJOE8999 and *E. coli* strain NM538, to Dr. Danielle Pedrolli for providing the sequence of sRNA30$_{pur}$, and to the "Landesstiftung Baden-Württemberg" for supporting my work.

Bibliography

[1] Roth A, Breaker RR. The structural and functional diversity of metabolite-binding riboswitches. Annu Rev Biochem. 2009;78:305-34.

[2] Serganov A, Nudler E. A decade of riboswitches. Cell. 2013;152:17-24.

[3] Sherwood AV, Henkin TM. Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. Annu Rev Microbiol. 2016;70:361-74.

[4] McCown PJ, Corbino KA, Stav S, Sherlock ME, Breaker RR. Riboswitch diversity and distribution. RNA (New York, NY). 2017;23:995-1011.

[5] Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. Riboswitches Control Fundamental Biochemical Pathways in Bacillus subtilis and Other Bacteria. Cell. 2003;113:577-86.

[6] Weinberg Z, Barrick JE, Yao Z, Roth A, Kim JN, Gore J, et al. Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. Nucleic Acids Res. 2007;35:4809-19.

[7] Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, et al. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. Proc Natl Acad Sci U S A. 2004;101:6421-6.

[8] Nudler E, Mironov AS. The riboswitch control of bacterial metabolism. Trends in Biochemical Sciences. 2004;29:11-7.

[9] Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, Kreneva RA, et al. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. Cell. 2002;111:747-56.

[10] Winkler WC, Cohen-Chalamish S, Breaker RR. An mRNA structure that controls gene expression by binding FMN. Proc Natl Acad Sci U S A. 2002;99:15908-13.

[11] Barrick JE, Breaker RR. The distributions, mechanisms, and structures of metabolite-binding riboswitches. Genome Biol. 2007;8:R239.

[12] Wickiser JK, Cheah MT, Breaker RR, Crothers DM. The kinetics of ligand binding by an adenine-sensing riboswitch. Biochemistry. 2005;44:13404-14.

[13] Wickiser JK, Winkler WC, Breaker RR, Crothers DM. The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. Mol Cell. 2005;18:49-60.

[14] Gilbert SD, Stoddard CD, Wise SJ, Batey RT. Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain. Journal of molecular biology. 2006;359:754-68.

[15] Lemay J-F, Desnoyers G, Blouin S, Heppell B, Bastet L, St-Pierre P, et al. Comparative Study between Transcriptionally- and Translationally-Acting Adenine Riboswitches Reveals Key Differences in Riboswitch Regulatory Mechanisms. PLoS Genetics. 2011;7:e1001278.

[16] Gong S, Wang Y, Wang Z, Sun Y, Zhang W. Folding behaviors of purine riboswitch aptamers. Wuhan University Journal of Natural Sciences. 2018;23:43-50.

[17] Baird NJ, Kulshina N, Ferre-D'Amare AR. Riboswitch function: flipping the switch or tuning the dimmer? RNA Biol. 2010;7:328-32.

[18] Hohmann HP, Dijl JMv, Krishnappa L, Prágai Z. Host Organisms: Bacillus subtilis. Industrial Biotechnology2016. p. 221-97.

[19] Ebbole DJ, Zalkin H. Interaction of a putative repressor protein with an extended control region of the Bacillus subtilis pur operon. J Biol Chem. 1989;264:3553-61.

[20] Weng M, Zalkin H. Mutations in the Bacillus subtilis purine repressor that perturb PRPP effector function in vitro and in vivo. Curr Microbiol. 2000;41:56-9.

[21] Saxild HH, Brunstedt K, Nielsen KI, Jarmer H, Nygaard P. Definition of the Bacillus subtilis PurR operator using genetic and bioinformatic tools and expansion of the PurR regulon with glyA, guaC, pbuG, xpt-pbuX, yqhZ-folD, and pbuO. J Bacteriol. 2001;183:6175-83.

[22] Bera AK, Zhu J, Zalkin H, Smith JL. Functional dissection of the Bacillus subtilis pur operator site. J Bacteriol. 2003;185:4099-109.

[23] Zakataeva NP, Romanenkov DV, Skripnikova VS, Vitushkina MV, Livshits VA, Kivero AD, et al. Wild-type and feedback-resistant phosphoribosyl pyrophosphate synthetases from Bacillus amyloliquefaciens: purification, characterization, and application to increase purine nucleoside production. Appl Microbiol Biotechnol. 2012;93:2023-33.

[24] Higashitsuji Y, Angerer A, Berghaus S, Hobl B, Mack M. RibR, a possible regulator of the Bacillus subtilis riboflavin biosynthetic operon, in vivo interacts with the 5'-untranslated leader of rib mRNA. FEMS microbiology letters. 2007;274:48-54.

[25] Pedrolli DB, Kuhm C, Sevin DC, Vockenhuber MP, Sauer U, Suess B, et al. A dual control mechanism synchronizes riboflavin and sulphur metabolism in Bacillus subtilis. Proc Natl Acad Sci U S A. 2015;112:14054-9.

[26] Johansen LE, Nygaard P, Lassen C, Agerso Y, Saxild HH. Definition of a second Bacillus subtilis pur regulon comprising the pur and xpt-pbuX operons plus pbuG, nupG (yxjA), and pbuE (ydhL). J Bacteriol. 2003;185:5200-9.

[27] Mandal M, Breaker RR. Adenine riboswitches and gene activation by disruption of a transcription terminator. Nature structural & molecular biology. 2004;11:29-35.

[28] Nygaard P, Saxild HH. The purine efflux pump PbuE in Bacillus subtilis modulates expression of the PurR and G-box (XptR) regulons by adjusting the purine base pool size. J Bacteriol. 2005;187:791-4.

[29] Shi T, Wang Y, Wang Z, Wang G, Liu D, Fu J, et al. Deregulation of purine pathway in Bacillus subtilis and its use in riboflavin biosynthesis. Microbial Cell Factories. 2014;13:101.

[30] Mironov AS, Karelov DV, Solov'eva IM, Eremina S, Errais-Lopes L, Kreneva RA, et al. [Relationship between the secondary structure and the regulatory activity of the leader region of the riboflavin biosynthesis operon in Bacillus subtilis]. Genetika. 2008;44:467-73.

[31] Soukup GA, Breaker RR. Relationship between internucleotide linkage geometry and the stability of RNA. RNA (New York, NY). 1999;5:1308-25.

[32] Shahbabian K, Jamalli A, Zig L, Putzer H. RNase Y, a novel endoribonuclease, initiates riboswitch turnover in Bacillus subtilis. The EMBO Journal. 2009;28:3523-33.

[33] Richards J, Liu Q, Pellegrini O, Celesnik H, Yao S, Bechhofer DH, et al. An RNA pyrophosphohydrolase triggers 5'-exonucleolytic degradation of mRNA in Bacillus subtilis. Mol Cell. 2011;43:940-9.

[34] DeLoughery A, Lalanne JB, Losick R, Li GW. Maturation of polycistronic mRNAs by the endoribonuclease RNase Y and its associated Y-complex in Bacillus subtilis. Proc Natl Acad Sci U S A. 2018;115:E5585-e94.

[35] Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, Mandin P, et al. A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. Cell. 2009;139:770-9.

[36] Kukanova A, Zhdanov VG, Stepanov AI. [Bacillus subtilis mutants resistant to roseoflavin]. Genetika. 1982;18:319-21.

[37] Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009;136:615-28.

[38] Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2011;3.

[39] Morfeldt E, Taylor D, von Gabain A, Arvidson S. Activation of alpha-toxin translation in Staphylococcus aureus by the trans-encoded antisense RNA, RNAIII. Embo j. 1995;14:4569-77.

[40] Sharma CM, Darfeuille F, Plantinga TH, Vogel J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev. 2007;21:2804-17.

[41] Prevost K, Salvail H, Desnoyers G, Jacques JF, Phaneuf E, Masse E. The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. Molecular microbiology. 2007;64:1260-73.

[42] Opdyke JA, Kang JG, Storz G. GadY, a small-RNA regulator of acid response genes in Escherichia coli. J Bacteriol. 2004;186:6698-705.

[43] Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. Nature structural & molecular biology. 2009;16:840-6.

[44] Desnoyers G, Morissette A, Prevost K, Masse E. Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. Embo j. 2009;28:1551-61.

[45] Georg J, Voss B, Scholz I, Mitschke J, Wilde A, Hess WR. Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. Mol Syst Biol. 2009;5:305.

[46] Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, Balestrino D, et al. The Listeria transcriptional landscape from saprophytism to virulence. Nature. 2009;459:950-6.

[47] Guell M, van Noort V, Yus E, Chen WH, Leigh-Bell J, Michalodimitrakis K, et al. Transcriptome complexity in a genome-reduced bacterium. Science. 2009;326:1268-71.

[48] Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature. 2010;464:250-5.

[49] Gerdes K, Wagner EG. RNA antitoxins. Curr Opin Microbiol. 2007;10:117-24.

[50] Fozo EM, Hemm MR, Storz G. Small toxic proteins and the antisense RNAs that repress them. Microbiol Mol Biol Rev. 2008;72:579-89, Table of Contents.

[51] Tramonti A, De Canio M, De Biase D. GadX/GadW-dependent regulation of the Escherichia coli acid fitness island: transcriptional control at the gadY-gadW divergent promoters and identification of four novel 42 bp GadX/GadW-specific binding sites. Molecular microbiology. 2008;70:965-82.

[52] Huntzinger E, Boisset S, Saveanu C, Benito Y, Geissmann T, Namane A, et al. Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. Embo j. 2005;24:824-35.

[53] Lease RA, Cusick ME, Belfort M. Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. Proc Natl Acad Sci U S A. 1998;95:12456-61.

[54] Valentin-Hansen P, Eriksen M, Udesen C. The bacterial Sm-like protein Hfq: a key player in RNA transactions. Molecular microbiology. 2004;51:1525-33.

[55] Brennan RG, Link TM. Hfq structure, function and ligand binding. Curr Opin Microbiol. 2007;10:125-33.

[56] Vogel J, Luisi BF. Hfq and its constellation of RNA. Nat Rev Microbiol. 2011;9:578-89.

[57] Rochat T, Delumeau O, Figueroa-Bossi N, Noirot P, Bossi L, Dervyn E, et al. Tracking the Elusive Function of Bacillus subtilis Hfq. PLoS One. 2015;10:e0124977.

[58] Silvaggi JM, Perkins JB, Losick R. Small untranslated RNA antitoxin in Bacillus subtilis. J Bacteriol. 2005;187:6641-50.

[59] Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, Possedko M, et al. Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes Dev. 2007;21:1353-66.

[60] Heidrich N, Moll I, Brantl S. In vitro analysis of the interaction between the small RNA SR1 and its primary target ahrC mRNA. Nucleic Acids Res. 2007;35:4331-46.

[61] Sharma V, Yamamura A, Yokobayashi Y. Engineering artificial small RNAs for conditional gene silencing in Escherichia coli. ACS Synth Biol. 2012;1:6-13.

[62] Isaacs FJ, Dwyer DJ, Ding C, Pervouchine DD, Cantor CR, Collins JJ. Engineered riboregulators enable post-transcriptional control of gene expression. Nature biotechnology. 2004;22:841-7.

[63] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315:1709-12.

[64] Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol. 2011;9:467-77.

[65] Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology. 2009;155:733-40.

[66] Karvelis T, Gasiunas G, Miksys A, Barrangou R, Horvath P, Siksnys V. crRNA and tracrRNA guide Cas9-mediated DNA interference in Streptococcus thermophilus. RNA Biol. 2013;10:841-51.

[67] Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468:67-71.

[68] Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A. 2012;109:E2579-86.

[69] Selle K, Barrangou R. Harnessing CRISPR-Cas systems for bacterial genome editing. Trends Microbiol. 2015;23:225-32.

[70] Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157:1262-78.

[71] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-Guided Human Genome Engineering via Cas9. Science. 2013;339:823-6.

[72] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816-21.

[73] Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology. 2013;31:233-9.

[74] Oh JH, van Pijkeren JP. CRISPR-Cas9-assisted recombineering in Lactobacillus reuteri. Nucleic Acids Res. 2014;42:e131.

[75] Dong H, Zhang D. Current development in genetic engineering strategies of Bacillus species. Microb Cell Fact. 2014;13:63.

[76] Hong KQ, Liu DY, Chen T, Wang ZW. Recent advances in CRISPR/Cas9 mediated genome editing in Bacillus subtilis. World J Microbiol Biotechnol. 2018;34:153.

[77] Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual: Cold Spring Harbor Laboratory Press; 2001.

[78] Commichau FM, Gunka K, Landmann JJ, Stulke J. Glutamate metabolism in Bacillus subtilis: gene expression and enzyme activities evolved to avoid futile cycles and to allow rapid responses to perturbations of the system. J Bacteriol. 2008;190:3557-64.

[79] Nguyen H, Phan T, Schumann W. Expression Vectors for the Rapid Purification of Recombinant Proteins in Bacillus subtilis. Current Microbiology. 2007;55:89-93.

[80] Antoniewski C, Savelli B, Stragier P. The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. J Bacteriol. 1990;172:86-93.

[81] Lewis PJ, Marston AL. GFP vectors for controlled expression and dual labelling of protein fusions in Bacillus subtilis. Gene. 1999;227:101-9.

[82] Altenbuchner J. Editing of the *Bacillus subtilis* Genome by the CRISPR-Cas9 System. Appl Environ Microbiol. 2016;82:5421-7.

[83] Harwood CR, Cutting SM. Molecular biological methods for Bacillus. Chichester; New York: Wiley; 1990.

[84] Tabor S. Expression using the T7 RNA polymerase/promoter system. Curr Protoc Mol Biol. 2001;Chapter 16:Unit16.2.

[85] Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. PLoS One. 2008;3:e3647.

[86] Quan J, Tian J. Circular polymerase extension cloning of complex gene libraries and pathways. PLoS One. 2009;4:e6441.

[87] Schaffert L, Schneiker-Bekel S, Dymek S, Droste J, Persicke M, Busche T, et al. Essentiality of the Maltase AmlE in Maltose Utilization and Its Transcriptional Regulation by the Repressor AmlR in the Acarbose-Producing Bacterium Actinoplanes sp. SE50/110. Front Microbiol. 2019;10:2448.

[88] Pfeifer-Sancar K, Mentz A, Ruckert C, Kalinowski J. Comprehensive analysis of the Corynebacterium glutamicum transcriptome using an improved RNAseq technique. BMC Genomics. 2013;14:888.

[89] Irla M, Neshat A, Brautaset T, Ruckert C, Kalinowski J, Wendisch VF. Transcriptome analysis of thermophilic methylotrophic Bacillus methanolicus MGA3 using RNA-sequencing provides detailed insights into its previously uncharted transcriptional landscape. BMC Genomics. 2015;16:73.

[90] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-20.

[91] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357-9.

[92] O'Kane C, Stephens MA, McConnell D. Integrable alpha-amylase plasmid for generating random transcriptional fusions in Bacillus subtilis. J Bacteriol. 1986;168:973-81.

[93] Frischauf AM, Lehrach H, Poustka A, Murray N. Lambda replacement vectors carrying polylinker sequences. Journal of molecular biology. 1983;170:827-42.

[94] Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature biotechnology. 2016;34:184-91.

[95] Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics. 2014;30:1473-5.

[96] Rodrigo G, Jaramillo A. RiboMaker: computational design of conformation-based riboregulation. Bioinformatics. 2014;30:2508-10.

[97] Will S, Joshi T, Hofacker IL, Stadler PF, Backofen R. LocARNA-P: accurate boundary prediction and improved detection of structural RNAs. RNA (New York, NY). 2012;18:900-14.

[98] Will S, Reiche K, Hofacker IL, Stadler PF, Backofen R. Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. PLoS Comput Biol. 2007;3:e65.

[99] Batey RT, Gilbert SD, Montange RK. Structure of a natural guanine-responsive riboswitch complexed with the metabolite hypoxanthine. Nature. 2004;432:411-5.

[100] Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF. RNAalifold: improved consensus structure prediction for RNA alignments. BMC Bioinformatics. 2008;9:474.

[101] Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL. The Vienna RNA websuite. Nucleic Acids Res. 2008;36:W70-4.

[102] Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, et al. NUPACK: Analysis and design of nucleic acid systems. Journal of Computational Chemistry. 2011;32:170-3.

[103] Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, et al. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. PLoS One. 2009;4:e7526.

[104] de Boer HA, Comstock LJ, Vasser M. The tac promoter: a functional hybrid derived from the trp and lac promoters. Proc Natl Acad Sci U S A. 1983;80:21-5.

[105] Buck J, Noeske J, Wohnert J, Schwalbe H. Dissecting the influence of Mg2+ on 3D architecture and ligand-binding of the guanine-sensing riboswitch aptamer domain. Nucleic Acids Res. 2010;38:4143-53.

[106] Kilstrup M, Jessing SG, Wichmand-Jorgensen SB, Madsen M, Nilsson D. Activation control of pur gene expression in Lactococcus lactis: proposal for a consensus activator binding sequence based on deletion analysis and site-directed mutagenesis of purC and purD promoter regions. J Bacteriol. 1998;180:3900-6.

[107] Adelman K, La Porta A, Santangelo TJ, Lis JT, Roberts JW, Wang MD. Single molecule analysis of RNA polymerase elongation reveals uniform kinetic behavior. Proceedings of the National Academy of Sciences. 2002;99:13538-43.

[108] Wang MD, Schnitzer MJ, Yin H, Landick R, Gelles J, Block SM. Force and velocity measured for single molecules of RNA polymerase. Science. 1998;282:902-7.

[109] Tolic-Norrelykke SF, Engh AM, Landick R, Gelles J. Diversity in the rates of transcript elongation by single RNA polymerase molecules. J Biol Chem. 2004;279:3292-9.

[110] Abbondanzieri EA, Greenleaf WJ, Shaevitz JW, Landick R, Block SM. Direct observation of base-pair stepping by RNA polymerase. Nature. 2005;438:460-5.

[111] Klumpp S, Hwa T. Growth-rate-dependent partitioning of RNA polymerases in bacteria. Proc Natl Acad Sci U S A. 2008;105:20245-50.

[112] Proshkin S, Rahmouni AR, Mironov A, Nudler E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science. 2010;328:504-8.

[113] Bielec K, Sozanski K, Seynen M, Dziekan Z, Ten Wolde PR, Holyst R. Kinetics and equilibrium constants of oligonucleotides at low concentrations. Hybridization and melting study. Phys Chem Chem Phys. 2019;21:10798-807.

[114] Lynch M, Marinov GK. The bioenergetic costs of a gene. Proc Natl Acad Sci U S A. 2015;112:15690-5.

[115] Maass S, Sievers S, Zuhlke D, Kuzinski J, Sappa PK, Muntel J, et al. Efficient, global-scale quantification of absolute protein amounts by integration of targeted mass spectrometry and two-dimensional gel-based proteomics. Anal Chem. 2011;83:2677-84.

[116] Artsimovitch I, Landick R. Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc Natl Acad Sci U S A. 2000;97:7090-5.

[117] Shundrovsky A, Santangelo TJ, Roberts JW, Wang MD. A single-molecule technique to study sequence-dependent transcription pausing. Biophys J. 2004;87:3945-53.

[118] Winkler ME, Yanofsky C. Pausing of RNA polymerase during in vitro transcription of the tryptophan operon leader region. Biochemistry. 1981;20:3738-44.

[119] Zhang H, Switzer RL. Transcriptional pausing in the Bacillus subtilis pyr operon in vitro: a role in transcriptional attenuation? J Bacteriol. 2003;185:4764-71.

[120] Steinert H, Sochor F, Wacker A, Buck J, Helmling C, Hiller F, et al. Pausing guides RNA folding to populate transiently stable RNA structures for riboswitch-based transcription regulation. eLife. 2017;6.

[121] Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. Nature chemical biology. 2009;5:593-9.

[122] Guo J, Wang T, Guan C, Liu B, Luo C, Xie Z, et al. Improved sgRNA design in bacteria via genome-wide activity profiling. Nucleic Acids Res. 2018;46:7052-69.

[123] Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. Nat Biotechnol. 2006;24:1558-64.

[124] Ott E, Stolz J, Lehmann M, Mack M. The RFN riboswitch of Bacillus subtilis is a target for the antibiotic roseoflavin produced by Streptomyces davawensis. RNA Biol. 2009;6:276-80.

[125] Lee ER, Blount KF, Breaker RR. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. RNA Biol. 2009;6:187-94.

[126] Pedrolli DB, Matern A, Wang J, Ester M, Siedler K, Breaker R, et al. A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to Streptomyces davawensis. Nucleic Acids Res. 2012;40:8662-73.

[127] Samant S, Lee H, Ghassemi M, Chen J, Cook JL, Mankin AS, et al. Nucleotide biosynthesis is critical for growth of bacteria in human blood. PLoS pathogens. 2008;4:e37.