

# **GlmY and GlmZ: a hierarchically acting regulatory cascade composed of two small RNAs**

**Dissertation**

for the award of the degree

"Doctor rerum naturalium" (Dr.rer.nat.)

Division of Mathematics and Natural Sciences  
of the Georg-August-Universität Göttingen

submitted by

**Yvonne Göpel**

from Eisenach

Göttingen 2013



**PD Dr. Boris Görke** (Supervisor and 1<sup>st</sup> Reviewer)

Max F. Perutz Laboratories, Department of Microbiology, Immunobiology and Genetics,  
University of Vienna

**Prof. Dr. Ralf Ficner** (2<sup>nd</sup> Reviewer)

Institute for Microbiology and Genetics, Department for Molecular Structural Biology,  
University of Göttingen

**Prof. Dr. Lutz Walter**

German Primate Center Göttingen, Department of Primate Genetics,  
University of Göttingen

Date of oral examination: October 2<sup>nd</sup>, 2013

Herewith I declare that the doctoral thesis entitled "**GlmY and GlmZ: a hierarchically acting regulatory cascade composed of two small RNAs**" was written independently and with no other aids and sources than quoted.

Göttingen,

Yvonne Göpel

## ACKNOWLEDGEMENTS

Zu allererst möchte ich mich ganz herzlich bei PD Dr. Boris Görke für die exzellente Betreuung bedanken und für die Möglichkeit an diesem äußerst spannenden Thema zu arbeiten. Seine offene Art und die Möglichkeit eigene Ergebnisse, Ideen und „phantastische“ Hypothesen offen und kritisch zu diskutieren, haben mich sehr beeindruckt und diese Arbeit immer wieder voran gebracht.

Weiterhin möchte ich Prof. Ralf Ficner für die Übernahme des Korreferates für diese Arbeit danken und Prof. Lutz Walter für seine Teilnahme an meinem Thesis Committee. Vielen Dank für die hilfreichen Tipps und kritischen Diskussionen während der Treffen und für das rege Interesse an meinem Thema! Prof. Stefanie Pöggeler, PD Dr. Michael Hoppert und Prof. Rolf Daniel möchte ich herzlich für Ihre Teilnahme an meiner Prüfungskommission danken. Auch dem Team der GGNB danke ich für Ihre Unterstützung.

Für die finanzielle Unterstützung danke ich dem Dorothea-Schlözer Programm der Universität Göttingen, sowie der GGNB - ohne diese Unterstützung wäre diese Arbeit nicht möglich gewesen.

Ein besonderer Dank geht an Sabine Lentes für ihren Einsatz und die kleinen Tipps und Kniffe und natürlich an den guten Geist der Abteilung: Bärbel Herbst. Bei meinen MasterstudentInnen und PraktikantInnen Tina Hollerbuhl, Tilmann Künzl, Lena Hoffmann, Muna Khan, Anna Kögler und Alexandra Juckert möchte ich mich für Ihr Interesse und die Beiträge zum Projekt bedanken. Weiterhin möchte ich der ganzen Abteilung danken und ganz besonders meinen lieben Kollegen der AG Görke: Denise, Jens, Chris und Muna. Es war immer lustig mit euch im Labor und auf Meetings und wir hatten viele spannende unterhaltungskulturelle Erkenntnisse... (Es gibt indische Zombie-Filme!) ☺ In diesem Sinne möchte ich auch meinem Austausch-Labor in Würzburg für ihre Hilfe im (Radioaktiv)Labor und die lustigen Abendaktivitäten in und um Würzburg danken: Movie Nights, Weinfeste, Kiliani und die illustren Abende am Mainufer haben mich in der Zeit wirklich aus dem Labortrübels gerettet. Ein großer Dank geht auch an Arne für den moralischen Zuspruch, die Gelassenheit und die Perspektive, die du dem Laborleben entgegen gebracht hast. Nicht zuletzt danke ich dem Taxi-Unternehmen Schmeisky, das mich sicher durch den einen oder anderen Winter gebracht hat... ☺

Last but not least, möchte ich mich bei meiner Familie für den bedingungslosen Rückhalt und das Vertrauen bedanken, dass Ihr mir entgegen gebracht habt. Mein aufrichtiger Dank geht auch an meine Freunde, die sich mit mir gefreut und mich bei den Talfahrten wieder aufgebaut haben. Rebecca, ich glaube hier muss ich keine Worte mehr verlieren, du weißt, wie sehr du mich

unterstützt hast. Stefan, du hast dir diesen Spaß nun schon ein zweites Mal angetan! Danke dafür. Und Danke, dass du immer ein offenes Ohr und eine Lösung (egal ob kurz- oder langfristig) parat hattest. Chris, die Arbeit mit dir werde ich sicher vermissen, es hat wirklich großen Spaß gemacht. Oliver, vielen, vielen Dank für deine Unterstützung, besonders in den schwierigen Phasen meiner Arbeit. Du hast mich überzeugt, dass alles gut werden wird...

*Ein Gelehrter in seinem Laboratorium ist nicht nur ein  
Techniker; er steht auch vor den Naturgesetzen wie ein Kind vor  
der Märchenwelt.*

*Marie Curie*

## LIST OF ABBREVIATIONS

|          |   |
|----------|---|
| 5' UTR   | 5' untranslated region                            |
| 3' UTR   | 3' untranslated region                            |
| ABS      | activator binding site                            |
| ATP      | adenosine triphosphate                            |
| bp       | base pairs  |
| DIG      | digoxigenin                                       |
| DNA      | deoxyribonucleic acid                             |
| DTT      | Dithiothreitol                                    |
| EMSA     | electro mobility shift analysis                   |
| Fig.     | Figure  |
| Nva-FMDP | N3-(4-methoxyfumaryl)-L-2,3-diaminopropanoic acid |
| Glc6P    | glucose-6-phosphate                               |
| GlcN     | glucosamine                                       |
| GlcNAc   | <i>N</i> -acetylglucosamine                       |
| GlcN1P   | glucosamine-1-phosphate                           |
| GlcN6P   | glucosamine-6-phosphate                           |
| GlcNAc6P | <i>N</i> -acetylglucosamine-6-phosphate           |
| GlmS     | glucosamine-6-phosphate synthase                  |
| GTP      | guanosine triphosphate                            |
| IHF      | integration host factor                           |
| IPTG     | isopropyl- $\beta$ -D-1-thiogalactopyranoside     |
| LB       | Luria Bertani                                     |

## List of abbreviations

---

|        |  |
|--------|--|
| LPS    | lipopolysaccharide                                       |
| mRNA   | messenger RNA  |
| nt     | nucleotide   |
| OD     | optical density  |
| PAP I  | poly(A) polymerase                                       |
| PCR    | polymerase chain reaction                                |
| PNPase | polynucleotide phosphorylase                             |
| psi    | pressure per square inch                                 |
| PTS    | phosphotransferase system                                |
| PVDF   | polyvinylidene difluoride membrane                       |
| RACE   | rapid amplification of cDNA ends                         |
| RBS    | ribosomal binding site                                   |
| RNA    | ribonucleic acid   |
| rpm    | revolutions per minute                                   |
| RppH   | pyrophosphohydrolase                                     |
| rRNA   | ribosomal RNA  |
| SD     | Shine-Dalgarno sequence                                  |
| SDS    | sodium dodecyl sulfate                                   |
| sRNA   | small RNA  |
| Tab.   | Table  |
| TAP    | tobacco acid pyrophosphatase                             |
| TCS    | two-component system                                     |
| UDP    | uridine diphosphate                                      |
| X-Gal  | 5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside |

---

**TABLE OF CONTENTS**

|   |            |
|---|------------|
| <b>ACKNOWLEDGEMENTS</b> .....   | <b>I</b>   |
| <b>LIST OF ABBREVIATIONS</b> .....  | <b>III</b> |
| <b>LIST OF PUBLICATIONS</b> .....   | <b>VI</b>  |
| <b>ABSTRACT</b> .....   | <b>1</b>   |
| <b>1. INTRODUCTION</b> .....  | <b>2</b>   |
| 1.1 Control of transcription initiation .....   | 2          |
| 1.2 Post-transcriptional gene regulation.....   | 5          |
| 1.3 RNA chaperon Hfq and its function.....  | 10         |
| 1.4 RNA degradation in <i>E. coli</i> .....   | 11         |
| 1.5 Regulation by protein modifications: new roles for acetylation in bacteria.....   | 15         |
| 1.6 Amino sugar metabolism in <i>E. coli</i> .....  | 17         |
| 1.7 The regulatory GlmYZ small RNA cascade.....   | 19         |
| 1.8 The GlrK/GlrR two-component system controls expression of small RNA GlmY in <i>E. coli</i> .....  | 20         |
| 1.9 Aim of this study.....  | 23         |
| <b>CHAPTER 2: Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in <i>Enterobacteriaceae</i>.</b> .....   | <b>24</b>  |
| Main manuscript.....  | 25         |
| Supplementary material .....  | 44         |
| <b>CHAPTER 3: The novel RNA binding protein RapZ (YhbJ) acts as a modulator affecting <math>\sigma^{54}</math>-dependent expression of sRNA GlmY independently of its function as an RNase adaptor protein.</b> ..... | <b>56</b>  |
| <b>CHAPTER 4: Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA</b> .....   | <b>83</b>  |
| Main manuscript.....  | 84         |
| Supporting Information .....  | 103        |
| <b>CHAPTER 5: Molecular requirements for Hfq binding and processing by RNase E within the regulatory GlmY/GlmZ sRNA cascade</b> .....   | <b>125</b> |
| <b>6. DISCUSSION</b> .....  | <b>152</b> |
| 6.1 Transcriptional control of small RNAs GlmY and GlmZ .....   | 152        |
| 6.2 Targeted decay of small RNA GlmZ by RNase adaptor RapZ and counteraction by sRNA GlmY .....   | 159        |
| 6.3 The GlmY/Z-cascade as a unique model system .....   | 166        |
| 6.4 Conclusion and perspectives .....   | 176        |
| <b>7. LIST OF REFERENCES</b> .....  | <b>178</b> |
| <b>CURRICULUM VITAE</b> .....   | <b>199</b> |

## LIST OF PUBLICATIONS

### research articles:

**Göpel, Y.**, Papenfort, K., Reichenbach, B., Vogel, J., and Görke, B. (2013) Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev.* **27**: 552-564

Resch M., **Göpel Y.**, Görke B., and Ficner R. (2013) Crystallization and preliminary X-ray diffraction analysis of YhbJ from *Escherichia coli*, a key protein involved in the GlmYZ sRNA regulatory cascade. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* **69**: 109-114

Lüttmann D., **Göpel Y.**, and Görke B. (2012) The phosphotransferase protein EIIA(Ntr) modulates the phosphate starvation response through interaction with histidine kinase PhoR in *Escherichia coli*. *Mol. Microbiol.* **86**: 96-110

**Göpel Y.**, Lüttmann D., Heroven AK., Reichenbach B., Dersch P., and Görke B. (2011) Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in *Enterobacteriaceae*. *Nucleic Acids Res.* **39**: 1294-1309

Reichenbach B., **Göpel Y.**, and Görke B. (2009) Dual control by perfectly overlapping sigma 54- and sigma 70- promoters adjusts small RNA GlmY expression to different environmental signals. *Mol. Microbiol.* **74**: 1054-1070

### review articles:

Göpel Y., Görke B. (2012) Rewiring two-component signal transduction with small RNAs. *Curr Opin Microbiol.* **15**: 132-139

Göpel Y., Görke B. (2012) Zusammenspiel von Zweikomponentensystemen und kleinen RNAs. *BIOspektrum.* **7.12**: 702-705

## ABSTRACT

In *Escherichia coli* and other enterobacteria, the homologous small RNAs (sRNAs) GlmY and GlmZ act in a hierarchical manner to feedback control expression of key enzyme glucosamine-6-phosphate (GlcN6P) synthase GlmS. Enzyme GlmS catalyzes formation of GlcN6P, which is the rate limiting reaction in the pathway of cell wall biosynthesis. Only sRNA GlmZ can activate the *glmS* mRNA by base-pairing, which releases the ribosomal binding site and allows synthesis of GlmS. The second sRNA GlmY acts indirectly to activate *glmS* by stabilizing GlmZ in a process that involves protein YhbJ. However, the molecular mechanism remained elusive. As sRNAs GlmY and GlmZ are crucial in maintaining essential cellular functions and are thus tightly controlled, we aimed to gain insight into the complex regulation of these sRNAs at the level of biosynthesis and decay.

First, we investigated control of *glmY* and *glmZ* transcription in several species using *Yersinia pseudotuberculosis*, *Salmonella thyphimurium* and *E. coli* as representatives. Three different promoter architectures were observed: (I) In *Y. pseudotuberculosis* expression of both sRNAs is driven solely from  $\sigma^{54}$ -promoters; (II) perfectly overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -dependent promoters control expression of *glmY* in *E. coli* and of both sRNA genes in *S. thyphimurium*; (III) in contrast, *glmZ* of *E. coli* is constitutively expressed from a  $\sigma^{70}$ -promoter. These results suggest that the *glmY/Z* system is in evolutionary transition from  $\sigma^{54}$ - to  $\sigma^{70}$ -dependency in a subset of species. Moreover, the  $\sigma^{54}$ -dependent promoters are activated by the two component system GlrK/GlrR and rely on integration host factor IHF for activity. Further, we found that acetylation of YhbJ is required for full activity of the  $\sigma^{54}$ -promoter of *glmY* in *E. coli* and that sirtuin deacetylase CobB drastically reduces promoter activity. In sum, GlmY and GlmZ seem to compose a regulon dependent on  $\sigma^{54}$ , GlrK/GlrR and IHF in the majority of *Enterobacteriaceae*.

Second, we clarified the molecular mechanism of signal transduction within the GlmYZ cascade. We demonstrated that YhbJ is a novel RNA-binding protein that binds GlmY and GlmZ with high affinity and switches its sRNA binding partners depending on the intracellular GlcN6P level. Under conditions of ample GlcN6P, YhbJ preferably binds GlmZ and recruits its processing machinery by protein-protein interaction with the major endoribonuclease RNase E. GlmZ is inactivated by RNase E and subsequently degraded. Upon GlcN6P depletion, GlmY accumulates and sequesters YhbJ thereby counteracting processing of GlmZ. In analogy to regulated proteolysis we renamed protein YhbJ to RapZ as acronym for RNase adaptor protein for sRNA GlmZ; thus, GlmY acts as an anti-adaptor decoy for protein RapZ. Even though GlmY and GlmZ are highly similar in sequence and structure, both sRNAs act by distinct mechanisms. While GlmZ is a base-pairing sRNA that depends on RNA chaperon Hfq for functionality and stability, GlmY acts solely by protein-binding and does not require Hfq. Moreover, GlmZ is processed by RNase E in a RapZ-dependent manner, whereas GlmY is not. Hence, we exploited GlmY and GlmZ as model system to study the molecular requirements for Hfq-binding and processing by RNase E. We found that the entire 3' end of GlmZ is required for high affinity binding by Hfq and Hfq-dependent stabilization *in vivo*. In contrast, the lateral bulge within the central stem loop of GlmZ is of prime importance for recognition by RNase E. In sum, our findings reveal an unprecedented mechanism controlling the activity of a small RNA at the level of its turnover. This mechanism involves the novel RNase adaptor protein RapZ, which might be the first of similar proteins conferring substrate specificity to a general ribonuclease.

## 1. INTRODUCTION

Differential gene expression is of utmost importance in all living organisms. Whereas only a subset of genes is essential under all conditions, most of the genetic repertoire of an organism is differentially regulated in response to intrinsic and extrinsic cues. In higher eukaryotes, differential gene regulation serves the differentiation of tissues and specialized cells. Thus, deviations in cellular development or abnormalities in differentiation are frequent causes for severe diseases and disabilities. In contrast, the predominant function of differential gene expression in prokaryotic cells is to ensure survival and prosperity by allowing for a quick response and elaborate adaptation to rapidly changing beneficial or malignant environmental conditions. While some environmental cues, such as heat or cold shock, require an extensive adaptive response, it might suffice to alter only a small subset of cellular activities to meet the requirements for adaptation to other stimuli. Hence, bacterial cells have evolved various intricate mechanisms allowing for modulation of cellular activities at all regulatory levels. The following chapter will focus on gene regulatory mechanisms and post-translational regulation of cellular activities in bacteria.

For instance, regulation of transcription rates can be achieved at the level of initiation, e.g. by DNA-binding transcriptional regulators or at the level of elongation, for example by attenuation or riboswitches. Regulation of gene expression at the post-transcriptional level involves the alteration of transcript stabilities or the efficiency of translation by either *cis*-acting RNA elements, such as riboswitches and ribozymes or by autonomously expressed small regulatory RNAs. Riboswitches may act by sensing metabolites, pH, temperature and other cues and adopt mutually exclusive secondary structures that either allow or prevent translation or alter the stability of the message. Another mechanism largely employed by riboswitches is the control transcription elongation by the formation of structures that may weaken RNA/DNA duplexes and their interaction with the RNA polymerase. Small RNAs also possess the ability to influence translation efficiency or turn-over rates of transcripts either directly by base-pairing with the mRNA or indirectly by altering the activity or availability of a regulator. Finally, RNA-binding proteins also possess the potential to affect transcription and translation rates and RNA stability.

### 1.1 Control of transcription initiation

Regulation of transcription can be achieved at the level of initiation, but also during elongation. The initiation of transcription might require the absence or presence of specific DNA-binding proteins, such as response regulators of two-component systems or other transcriptional regulators. These proteins either recruit the RNA polymerase holoenzyme, or mask the specific recognition site, thereby preventing association of the polymerase with the promoter of the regulated gene.

Furthermore, alternative sigma factors allow for control of large subsets of genes, as they serve as RNA polymerase subunits dedicated to promoter recognition.

### **Transcriptional control by alternative sigma factors**

Sigma factors serve as prokaryotic transcription initiation factors that interact with the RNA-polymerase core enzyme and convey the ability to recognize and interact with a specific promoter sequence. The Gram-negative model bacterium *Escherichia coli* possesses seven distinct sigma factors that recognize distinct promoter consensus sequences. Most sigma factors, including  $\sigma^{70}$ , recognize promoter consensus sequences containing a -10 signal and a -35 sequence; however a few exceptions apply (see below). Whereas the primary sigma factor  $\sigma^{70}$  (RpoD) ensures expression of genes essential to maintain expression of basic cellular activities, specialized sigma factors allow differential expression of genes belonging to appropriate regulons for the adaptation to more specific conditions (Helmann and Chamberlin, 1988; Gruber and Gross, 2003). Thus, the activities of the six alternative sigma factors are tightly regulated and they only initiate transcription under specific conditions.

For instance, the general stress responsive sigma factor  $\sigma^S$  (RpoS, or  $\sigma^{38}$ ) responds to nutrient starvation, oxidative and temperature stress and other cues and ensures expression of genes in stationary phase (Loewen *et al.*, 1998; Gaal *et al.*, 2001; Maciag *et al.*, 2011). RpoS recognizes similar consensus sequences as the primary sigma factor (Gaal *et al.*, 2001). Further, an extensive regulatory antagonism has been described for RpoS with at least two other sigma factors, the nitrogen-related sigma factor  $\sigma^{54}$  (RpoN) and the flagellar sigma factor  $\sigma^{28}$  (RpoF) (Dong *et al.*, 2011). As a consequence, expression and activity of RpoS are tightly regulated on DNA, RNA and protein level (Battesti *et al.*, 2011). Another stress responsive sigma factor is  $\sigma^E$  (RpoE, or  $\sigma^{24}$ ), which is activated during envelope stress response. Sigma factor  $\sigma^E$  is required to cope with extreme heat, changes in membrane structure and composition and miss-folded proteins of the outer membrane or periplasma (Ades *et al.*, 2003; Tam and Missiakas, 2005; Bury-Moné *et al.*, 2009). The heat shock sigma factor  $\sigma^{32}$  (RpoH) belongs to the RpoE regulon and is involved in adaptation to heat stress. Its major target genes encode chaperons, proteases and DNA repairing enzymes (Zhao *et al.*, 2005). Sigma factor  $\sigma^{FecI}$  ( $\sigma^{19}$ ) specifically regulates genes required for ferric acid transport (Angerer *et al.*, 1995). Many sigma factors have been shown to share extensive regulatory overlap, i.e. many genes are transcribed from more than one promoter. Thus, one of the functions of alternative sigma factors is thought to be the increase of transcription rates from genes that usually depend on  $\sigma^{70}$  (Wade *et al.*, 2006; Zhao *et al.*, 2010).

### **The *rpoN* gene encodes an unusual alternative sigma factor**

Sigma factor  $\sigma^{54}$  (RpoN) has previously been described to control genes that are related to nitrogen limitation, however, while this may be the major role of RpoN it is by far not its only function and several genes activated by RpoN serve other purposes (Reitzer and Schneider, 2001; Zhao *et al.*, 2010). With the exception of  $\sigma^{54}$ , all sigma factors of *E. coli* are homologues of the primary sigma factor  $\sigma^{70}$  and thus belong to the same class of sigma factors (Merrick, 1993). Additionally the  $\sigma^{54}$  RNA polymerase holoenzyme recognizes specific -24 and -12 sequence motifs, rather than the more common -35 and -10 recognition sequences utilized by most sigma factors of the  $\sigma^{70}$  family (Reitzer and Schneider, 2001). In contrast to other sigma factors,  $\sigma^{54}$  is able to bind its recognition sequences on the DNA without prior association with the RNA polymerase (Tintut *et al.*, 1995). However, the  $\sigma^{54}$ -RNA polymerase holoenzyme is still unable to catalyze open complex formation. It requires interaction with gene-specific enhancer-like activator proteins to initiate transcription (Hoover *et al.*, 1990; Wigneshweraraj *et al.*, 2008). Analogous to the eukaryotic transcription initiation process, these activator proteins usually bind as hexamers to specific sequences far upstream of the promoter. In case of  $\sigma^{54}$ -dependent promoters, the specific enhancer-like activator proteins hydrolyze ATP in order to facilitate open complex formation (Wigneshweraraj *et al.*, 2008). While many of these proteins are response regulators of two-component systems or DNA-binding proteins belonging to other families, some  $\sigma^{54}$ -dependent promoters can also be activated by ATPases that do not contain a DNA-binding domain (Beck *et al.*, 2007). Due to the far distance between the region of transcription initiation and the activator binding sites, proteins such as integration host factor (IHF) that aid in DNA-bending are often required to bring the  $\sigma^{54}$  RNA polymerase holoenzyme and the specific activator ATPases in close proximity (Wigneshweraraj *et al.*, 2008).

### **The role of two-component systems in regulation of gene expression**

Signal perception and transduction is a prerequisite for gene regulation and the response to changing environmental conditions. In bacteria, two-component and phosphorelay systems are widely distributed sensory systems that allow perception of certain stimuli, transmission within the cell and conversion into an adaptive response. The number of two-component systems varies greatly in different bacteria and approximately 30 two-component systems are currently known for *E. coli* (Jung *et al.*, 2012). A trans-membrane sensory histidine kinase and a cytosolic response regulator constitute the majority of two-component systems. Upon signal perception, the histidine kinase auto-phosphorylates with ATP at a conserved histidine residue within its transmitter domain. Subsequently, the phosphoryl-group is transferred to a conserved aspartate residue within the receiver domain of the cognate response regulator (Mitrophanov and Groisman, 2008; Jung *et al.*, 2012). In most cases, the response regulator itself acts as a transcriptional regulator and

phosphorylation alters its DNA-binding affinity. Thus, activation of a two-component system by its cognate stimulus alters gene expression to fit the requirements determined by intra- or extracellular changes (Szurmant *et al.*, 2007; Mitrophanov and Groisman, 2008). Often, sensor kinases also possess phosphatase activity towards the cognate response regulator and are thus able to terminate the response in absence of the adequate stimulus (Szurmant *et al.*, 2007).

Recently, it became apparent that many two-component systems are organized in extensive networks that are interconnected by various other regulatory elements, such as global transcriptional regulators, alternative sigma factors and small regulatory RNAs that act at the post-transcriptional level. The involvement of sRNAs in regulatory networks of two-component systems provides a missing link to previously unaccountable observations of regulatory effects of a two-component system on gene expression without any apparent response regulator binding (Göpel and Görke, 2012a; Mandin and Guillier, 2013). This not only allows to expand or revert the regulatory repertoire of two-component systems, but also provides the possibility to integrate various different stimuli. Gene expression may thus be fine-tuned to an adequate response that fits the unique requirements for various combinations of stimuli. This increase in complexity and flexibility allows to control and time developmental processes or complex social behavior, such as quorum sensing, luminescence or virulence (Göpel and Görke, 2012a).

## **1.2 Post-transcriptional gene regulation**

Regulation of gene expression at the post-transcriptional level modulates the translation rate and stability of pre-existing transcripts and in turn alters the final amounts of the encoded protein. Various different mechanisms that depend on RNA molecules as regulators have evolved to perform these functions. For instance, *cis*-acting riboswitches form mutually exclusive stem-loop structures within the mRNA leader region in response to the availability of metabolites, small molecules or physiological changes (Serganov and Nudler, 2013). Another class of RNA-based regulators are autonomously expressed small regulatory RNAs that are either encoded anti-sense to their target transcripts (*cis*-acting sRNAs, asRNAs) or in *trans* at a different locus on the chromosome (Storz *et al.*, 2011).

### **Riboswitches and Ribozymes**

Riboswitches respond to various ligands or changes in pH or temperature and adopt alternative mutually exclusive secondary structures. Apart from acting at the level of transcription elongation, riboswitches often control translation initiation by adopting a structure that masks the ribosome binding site. Upon binding of the adequate ligand, riboswitches may form an alternative structure

either releasing the ribosome binding site or sequestering it (Serganov and Nudler, 2013). Recently, riboswitches were shown to additionally control the stability of their target transcript by actively masking and releasing a processing site (Caron *et al.*, 2012). An exceptional ribozyme/riboswitch is the *glmS* ribozyme that regulates expression of glucosamine-6-phosphate synthase GlnS in *Bacilli* by catalyzing self-cleavage rather than adopting an alternative structure (Winkler *et al.*, 2004).

### **The *glmS* ribozyme mediates feedback-regulation of glucosamine-6-P synthase GlnS in *Bacilli***

In bacteria, glucosamine-6-phosphate synthase GlnS catalyzes the key reaction in biosynthesis of cell wall precursors, the formation of glucosamine-6-phosphate (GlcN6P) from fructose-6-phosphate and glutamine. In eukaryotes, the enzymatic activity of GlnS is feedback inhibited by UDP-*N*-acetylglucosamine, the final product of the amino sugar pathway initiated by GlnS (Milewski, 2002). In *Bacilli*, rather than directly inhibiting the enzymatic activity of GlnS, synthesis of the GlnS protein is feedback regulated in response to the concentration of its enzymatic product, GlcN6P (Winkler *et al.*, 2004). Interestingly, this is achieved by a post-transcriptional mechanism employing a regulatory ribozyme located in the 5' un-translated region (UTR) of the *glmS* mRNA. In its apo-state, the ribozyme is inactive and *glmS* is expressed. Upon binding of the catalytic co-factor GlcN6P, the *glmS* ribozyme is active and catalyzes self-cleavage of its mRNA (Winkler *et al.*, 2004; Collins *et al.*, 2007). Subsequently, the *glmS* transcript is degraded by RNase J1 leading to a lower level of GlnS (Collins *et al.*, 2007). As opposed to regulation of the stability of the *glmS* transcript by a ribozyme, a cascade of small regulatory sRNAs controls expression of *glmS* in Gram-negative *Enterobacteriaceae* (Görke and Vogel, 2008; Reichenbach *et al.*, 2008; Urban and Vogel, 2008).

### **Small regulatory RNAs**

Small non-coding RNAs are major post-transcriptional regulatory molecules that employ various mechanisms and are involved in practically all physiological processes in bacterial cells (Storz *et al.*, 2011). Next to classifying sRNAs by their genetic localization as *cis*- or antisense RNAs and *trans*-acting sRNAs, regulatory RNAs can be categorized by their mode of action as either base-pairing and/or protein-binding sRNAs.

*Cis*-regulatory small RNAs were first described as regulators of plasmid replication and distribution or as members of toxin/anti-toxin systems (Brantl, 2007; Brantl, 2009). Recently, some chromosomally encoded antisense RNAs have been shown to fulfill other regulatory functions as well (Georg and Hess, 2011). For example, asRNA GadY of *E. coli* stabilizes the *gadX* transcript by base-pairing within the 3' UTR. GadX is a transcriptional regulator that directs expression of genes required for glutamate-dependent acid resistance as well as genes associated with multidrug efflux systems (Opdyke *et al.*, 2004; Nishino *et al.*, 2008). In addition to stabilizing the *gadX* transcript, GadY also

directs processing of the *gadXW* mRNA (Opdyke *et al.*, 2004; Opdyke *et al.*, 2011). Another example is provided by an asRNA regulating expression of the *ureB* subunit of the urease in the human pathogen *Helicobacter pylori* (Wen *et al.*, 2011). Even though urease activity ensures survival of *H. pylori* in the acidic habitat of the stomach, its activity becomes lethal in the absence of acid, e.g. after ingestion. Under these conditions the expression of the asRNA is activated by the non-phosphorylated ArsR response regulator of the ArsS/ArsR two-component system. This leads to down regulation of *ureB* and thus, decrease of urease activity. Interestingly, the acid-responsive ArsS/ArsR system also regulates expression of the *ureAB* urease genes, but in this case activation of the sensor kinase and subsequent phosphorylation of ArsR are required for activation of transcription (Wen *et al.*, 2011).

Whereas antisense RNAs that act by base-pairing share full complementarity with their targets, *trans*-encoded sRNAs act by imperfect base-pairing and often regulate multiple target transcripts, as exemplified by sRNAs RyhB, a master regulator of iron homeostasis and GcvB, an sRNA that controls various amino acid transporters (Storz *et al.*, 2011; Salvail and Massé, 2012; Sharma *et al.*, 2011). As imperfect base-pairing with the target often relies on short non-consecutive stretches of complementarity, at least in Gram negative bacteria, most *trans*-acting base-pairing sRNAs require the Sm-like RNA chaperon Hfq for functionality (Urban and Vogel, 2007; Vogel and Luisi, 2011). Hfq has been shown to stabilize sRNAs and facilitate cognate sRNA/mRNA duplex formation *in vivo* (Link *et al.*, 2009; Sauer *et al.*, 2012; Vogel and Luisi, 2011). In contrast, many protein-binding sRNAs seem to be independent of Hfq and rather act by altering the activity of their cognate protein targets.

### **Mechanisms employed by *trans*-encoded small RNAs**

Analogous to microRNAs in eukaryotes, most *trans*-encoded sRNAs act by gene silencing (De Lay *et al.*, 2013). Small RNAs may base-pair in the vicinity of the Shine-Dalgarno region of their respective target transcripts and render this region inaccessible for ribosomes. As a consequence, processing sites are exposed and thus transcripts devoid of polysomes are often rapidly degraded in *E. coli* (Baker and Mackie, 2003). Examples for this mode of action include the inhibition of translation of *galK* translation by Spot42 (Møller *et al.*, 2002), the repression of *ptsG* translation by SgrS (Kawamoto *et al.*, 2006), and the inhibition of translation of outer membrane porins by their respective sRNA regulators (Vogel and Papenfort, 2006). Whereas the absence of translating ribosomes may be sufficient to destabilize some mRNA molecules, sRNAs have also been shown to promote degradation of their target transcripts actively by recruiting and activating the respective RNA degrading complexes (Bandyra *et al.*, 2012; Prévost *et al.*, 2011).

In contrast to microRNAs, bacterial small RNAs also possess the potential to directly activate their target transcripts, even though activation occurs less frequently. Here, the nascent mRNA forms inhibitory stem loop structures within the 5' UTR that mask the ribosome binding site and prevents translation of the message (Fröhlich and Vogel, 2009). Base-pairing with the sRNA activates translation by disrupting the inhibitory structure and releasing the Shine-Dalgarno sequence. Small RNA DsrA, for instance, base-pairs with the *rpoS* mRNA and allows for synthesis of sigma factor  $\sigma^S$  in response to cold shock (Lease and Belfort, 2000). Another intriguing example is the activation of *glmS* expression upon base-pairing with sRNA GlmZ under conditions of GlcN6P limitation (Kalamorz *et al.*, 2007; Urban and Vogel, 2008). This elaborate control mechanism mediates GlcN6P homeostasis and involves a second sRNA, GlmY, and an RNA-binding protein, RapZ (Göpel *et al.*, 2013; Görke and Vogel, 2008; Reichenbach *et al.*, 2008; Urban and Vogel, 2008; this work; see below for details). By binding protein RapZ, GlmY indirectly aids activation of *glmS* (Reichenbach *et al.*, 2008; Göpel *et al.*, 2013, this work). Thus, GlmY belongs to the class of sRNAs that bind and alter the activity of proteins.

#### **On the roles of protein-binding small RNA**

Even though most sRNAs act by base-pairing, some examples of regulatory RNAs are known that modulate protein activities and thus influence target transcripts indirectly (Storz *et al.*, 2011). In *E. coli* for instance, sRNAs CsrB and CsrC control the activity of the master carbon storage regulator CsrA. The CsrA RNA-binding protein interacts with its target transcripts encoding proteins involved in various pathways of the carbohydrate metabolism, motility, biofilm formation and virulence (Babitzke and Romeo, 2007). CsrA may activate or repress translation and was also shown to influence the stability of its target mRNAs. For instance, a recent study reported that CsrA activates expression of the master regulator for flagella biosynthesis and chemotaxis *flhDC* by protecting the mRNA from cleavage by RNase E (Yakhnin *et al.*, 2013). Small RNAs CsrB and CsrC mimic targets of the CsrA protein by displaying multiple copies of hairpins containing GGA-motifs. These motifs are preferentially bound by CsrA (Dubey *et al.*, 2005). By titrating up to 18 molecules of CsrA, the CsrB/C sRNAs indirectly regulate gene expression of various CsrA-target mRNAs (Liu *et al.*, 1997). As the two small RNAs are expressed under distinct conditions, the Csr-System differentially regulates competing metabolic pathways in response to changing environmental cues. Homologs of CsrA and the small RNAs are widely distributed among bacteria, with varying numbers of sRNA homologs; in *Pseudomonas* and *Legionella* species the Csr-system is termed Rsm for regulator of secondary metabolism (Sonnleitner and Haas, 2011).

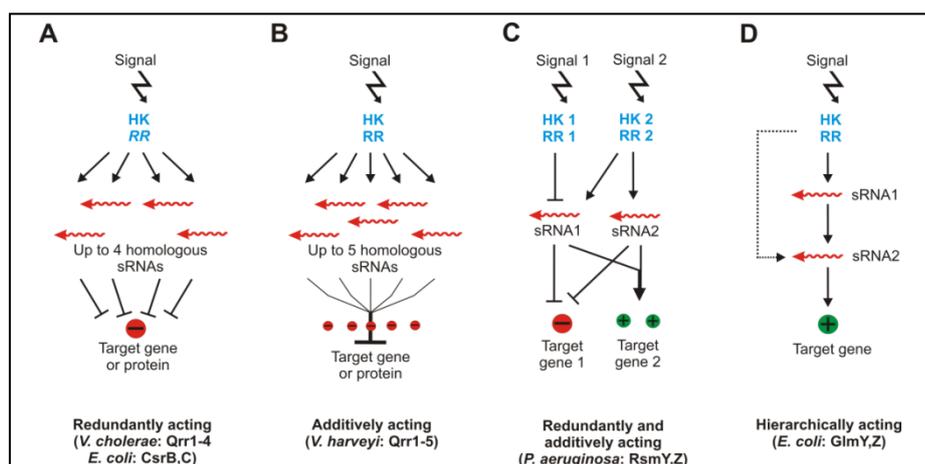
Another example for a protein-binding sRNA is the bacterial 6S RNA that binds and stores the  $\sigma^{70}$ -RNA polymerase holoenzyme, thus inhibiting transcription from  $\sigma^{70}$ -dependent promoters during

stationary growth. The structure of 6S RNA mimics an open complex at the promoter level. Transcription of a 6S-derived short product RNA (pRNA) releases the  $\sigma^{70}$ -RNA polymerase during outgrowth from stationary phase (Willkomm and Hartmann, 2005; Wassarman, 2007). Interestingly, *Bacillus subtilis* possesses two homologs of 6S RNA that differ in pRNA formation (Beckmann *et al.*, 2011).

### **Homologous sRNAs operate in four distinct modes of action**

As mentioned above, small RNAs are frequently found in regulatory networks of two-component systems and other global regulators. Peculiarly, several sRNAs possess multiple homologs involved in the same regulatory circuits. However, homologous sRNAs employ different modes of action to regulate their target transcripts (Göpel and Görke, 2012a). For instance, the aforementioned CsrB/C sRNAs or the quorum sensing related sRNAs Qrr1-4 of *Vibrio cholerae* are redundant sRNAs (Babitzke and Romeo, 2007; Lenz *et al.*, 2004). Thus, each single sRNA is sufficient for regulation of the respective targets and loss of one sRNA is redeemed by up-regulation of the remaining homologs. This is achieved by negative feedback loops between the respective regulator and the small RNAs (Fig. 1.1 A; Svenningsen *et al.*, 2009; Weilbacher *et al.*, 2003).

In contrast, the quorum sensing sRNAs Qrr1-5 of *Vibrio harveyi* act additively. Hence, full repression of the target transcript is only achieved by the combined action of all five sRNAs. This allows to generate an accurate gradient of target gene expression corresponding to the strength of the quorum sensing stimulus (Fig. 1.1 B; Tu and Bassler, 2007). Expanding the regulatory potential of homologous sRNAs, redundant and additive mechanisms may be combined to differentially regulate additional targets. For example, sRNAs RsmY and RsmZ of *Pseudomonas aeruginosa* are regulated by the GacS/GacA two-component system (Valverde *et al.*, 2003). Recently, it was demonstrated that sRNA RsmY but not RsmZ may additionally be repressed by a different two-component system. As a result, activation of GacS/GacA and the absence of the stimulus for the second two-component system are required for full activation of a specific target gene that relies on additive action of RsmY and RsmZ. In contrast, for regulation of other targets by a redundant mechanism, the activation of the GacS/GacA two-component systems is sufficient (Fig. 1.1 C; Bordi *et al.*, 2010). This allows to modify the strength of regulation on a subset of targets that are regulated additively, while another subset is constantly regulated by a redundant mechanism.



**Figure 1.1: Homologous sRNAs employ distinct modes of action (adapted from Göpel and Görke, 2012a).** **A.** For redundantly acting sRNAs, one copy of the sRNA is sufficient for full regulation of the respective target transcript or protein. Loss of sRNA copies can be compensated by up-regulation of the remaining sRNAs usually due to negative feedback loops between the regulator and the sRNAs. **B.** Each additively acting sRNA contributes to the regulation of the target genes and complete regulation is accomplished by the combined effect of all sRNAs. **C.** Redundantly acting sRNAs may influence a subset of targets additively. The homologs may be controlled by a common two-component system, while only one of the sRNAs is also subject to control by a second two-component system. Thus, activation of the two systems differentially influences the target transcripts or proteins. **D.** Homologous sRNAs may act in a hierarchical cascade to control expression of the target gene. One or both sRNAs may be controlled by a two-component system. While one sRNA acts directly on the target transcript, the second sRNA may act indirectly by stabilizing its homolog.

So far, only one case of hierarchically acting homologous sRNAs has been described (Fig. 1.1 D). In *E. coli* and other Gram negative bacteria, sRNAs GlmY and GlmZ act in a cascade to activate expression of *glmS*. Exclusively, GlmZ directly base-pairs with the *glmS* transcript, while GlmY acts indirectly by protecting GlmZ from degradation (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Görke and Vogel, 2008; Göpel *et al.*, 2013, this work; see below for details).

### **1.3 RNA chaperon Hfq and its function**

At least in Gram negative bacteria, many *trans*-encoded base-pairing small RNAs require the homo-hexameric, ring-shaped RNA chaperon Hfq for functionality and stability *in vivo* (Urban and Vogel, 2007; Link *et al.*, 2009; Vogel and Luisi, 2011). Hfq is believed to possess three distinct RNA binding regions: the distal site that was shown to specifically interact with ARN-repeats (R denotes a purine and N any nucleotide), the proximal face that preferentially binds to U-rich sequences often succeeding Rho-independent terminators of sRNAs, and the lateral surface (Link *et al.*, 2009; Sauer and Weichenrieder, 2011; Otaka *et al.*, 2011; Sauer *et al.*, 2012). Whereas mRNAs containing ARN-motif and poly-A sequences are preferentially bound the distal site, initial binding of sRNAs occurs at the proximal site. Recently, the rim or lateral face of Hfq was suggested to serve as a platform for extended Hfq/sRNA interaction thereby protecting entire sRNA molecules from cleavage by ribonucleases *in vivo* (Sauer *et al.*, 2012). Furthermore, Hfq simultaneously binds sRNA and mRNA molecules and/or induces structural changes within the RNA molecules and facilitates base-pairing between cognate sRNA/mRNA pairs (Vogel and Luisi, 2011). However, Hfq was shown to

differentially affect various sRNAs. Whereas some sRNAs require Hfq mainly for stability, other sRNAs strongly rely on Hfq to exert their function on target transcripts (Henderson *et al.*, 2013).

Hence, Hfq is of utmost importance for post-transcriptional gene regulation by small RNAs. Various mechanisms are known by which Hfq may act to induce sRNA-mediated gene silencing. For instance, Hfq was shown to directly interact with the RNA degradosome thereby recruiting the RNA processing machinery for coupled degradation of sRNA/mRNA duplexes (Massé *et al.*, 2003; Ikeda *et al.*, 2011; Prévost *et al.*, 2011). Since Hfq was suggested to replace the canonically associated DEAD-box helicase RhlB in the RNA degradosome (see below), this mechanism may lead to formation of an alternative degradosome reprogramed for degradation of sRNA/mRNA pairs (Ikeda *et al.*, 2011). In addition, Hfq may directly compete with 30S ribosomal subunits and hinder translation initiation. This non-canonical mechanism was suggested for the regulation of the *sdhC* transcript by sRNA Spot42 (Desnoyers and Massé, 2011). The *sdhC* transcript encodes a subunit of the iron-containing succinate dehydrogenase and is tightly repressed by various sRNAs during iron-starvation. Desnoyers and Massé could show that sRNA Spot 42 recruits Hfq to a precise A/U-rich sequence in close proximity to the translation initiation region and thus represses translation initiation indirectly. In *E. coli*, transcripts that are not associated with actively translating polysomes are rapidly degraded (Baker and Mackie, 2003). Therefore, while stabilizing unpaired sRNAs, Hfq may also actively and passively induce degradation of sRNA/mRNA pairs.

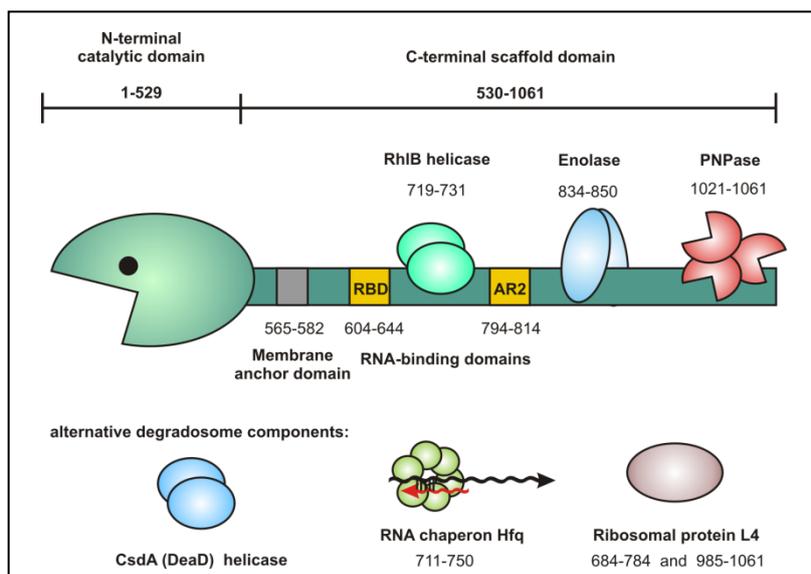
### **1.4 RNA degradation in *E. coli***

The mechanism of RNA degradation in bacteria as a consequence of cleavage reactions catalyzed by various ribonucleases is best characterized in *E. coli* (Mackie, 2013a; Górna *et al.*, 2012). The canonical pathway of RNA decay is initiated by endoribonucleolytic cleavage by RNase E, which cleaves within single-stranded A/U-rich regions or RNase III, which degrades double-stranded and highly structured RNA molecules (Kim *et al.*, 2004; Carpousis, 2007; Arraiano *et al.*, 2010). Due to its vital role in rRNA and tRNA maturation as well as in bulk RNA turn-over, the catalytic activity of RNase E is essential (Condon, 2007; Mackie, 2013a). RNase E has been shown to preferentially cleave 5' mono-phosphorylated RNA. Thus, in analogy to the process of decapping of transcripts in eukaryotes, bacterial transcripts are often marked for decay by pyrophosphohydrolase RppH that catalyzes the conversion of the 5' tri-phosphate to a mono-phosphate (Deana *et al.*, 2008). Following the endonucleolytic cleavage, exoribonucleases with a 3'→5' directionality, such as polynucleotide phosphorylase PNPase, RNase II and RNase R attack and degrade the RNA (Arraiano *et al.*, 2010; Górna *et al.*, 2012; Mackie, 2013a). Decay of structured RNA molecules often requires the aid of additional enzymes. RNA helicases remodel the structure of the RNA molecule and

poly(A)polymerase PAP-I adds 3' poly-(A)-sequences, which serve as binding platform for processive exonucleases (Mohanty and Kushner, 2006). Finally, the essential oligoribonuclease catalyzes the reduction of RNA oligomers to single nucleotides (Jain, 2002).

### RNase E and the degradosome

In many bacteria, the key enzymes for RNA turn-over assemble and form a multi-enzyme RNA degrading complex, the degradosome. In *E. coli* and other gamma-proteobacteria, the scaffold for the organization of this complex is provided by the C-terminal non-catalytic domain of RNase E (Ait-Bara and Carpousis, 2010; G3rna *et al.*, 2012). The canonical components include the ATP-dependent RNA helicase RhlB, the polynucleotide phosphorylase PNPase and the glycolytic enzyme enolase (Carpousis *et al.*, 1994; Miczak *et al.*, 1996; Vanzo *et al.*, 1998). As a function of the degradosome, RNA decay was suggested to be linked to the physiology and the metabolic state of the cell (Bernstein *et al.*, 2004; Del Favero *et al.*, 2008; Newman *et al.*, 2012). For example association with the cold-shock helicase CsdA (DeaD) was suggested to adapt RNA turn-over to cold shock conditions and interaction with the ribosomal protein L4 was proposed to stabilize stress-responsive transcripts (Fig. 1.2; Prud'homme-Genereux *et al.*, 2004; Kaberdin and Lin-Chao, 2009).



**Figure 1.2: The RNA degradosome of *E. coli* (modified from G3rna *et al.*, 2012).** RNase E consists of two distinct domains, the N-terminal catalytic domain and the C-terminal scaffolding domain. The scaffold provides an interaction platform for the protein components constituting the RNA degrading enzyme complex termed degradosome. Enolase, helicase RhlB and PNPase are canonical components. RNase E can also transiently associate with cold shock helicase CsdA (Prod'homme-Genereux *et al.*, 2004), RNA chaperon Hfq (Ikeda *et al.*, 2011) and ribosomal protein L4 (Singh *et al.*, 2009), which may lead to the formation of alternative degradosome complexes.

Whereas the catalytic function of RNase E is essential, the ability to assemble the degradosome is not (Ow *et al.*, 2000; Bandyra *et al.*, 2013). Nonetheless, the degradosome has been implicated in differential RNA turn-over. Various other proteins were shown to transiently interact with RNase E.

These interactions may trigger to the formation of alternative degradosomes programmed for the degradation of certain subsets of target transcripts. For instance, the repressor of RNase activity RraA globally alters RNA turn-over, while Hfq was shown to recruit the degradosome for coupled degradation of sRNA/mRNA duplexes (Lee *et al.*, 2003; Gao *et al.*, 2006; Ikeda *et al.*, 2011; Massé *et al.*, 2003). Hence, RNase E and the degradosome are key factors involved in post-transcriptional gene regulation processes that are mediated by small RNAs (Storz *et al.*, 2011; Luisi and Vogel 2011; Górna *et al.*, 2012).

### **Substrate recognition by RNase E**

Substrate recognition by RNase E is poorly understood as RNase E cleaves a multitude of RNA substrates within A/U-rich sequences, but does not possess a canonical cleavage sequence. So far, two distinct mechanisms are considered for substrate recognition by RNase E. The first mechanism is dependent on the 5' phosphorylation state of the targeted transcript. As mentioned above, RNase E preferentially cleaves 5' mono-phosphorylated RNA molecules (Kim *et al.*, 2004; Carpousis, 2007; Deana *et al.*, 2008). Thus, after generation of a 5' mono-phosphorylated RNA molecule by RppH, for instance, this terminus directly interacts with a phosphate 'sensory pocket' at the N-terminal catalytic domain and allosterically stimulates RNase E for cleavage (Callaghan *et al.*, 2005; Mackie, 1998; Mackie, 2013b). In contrast, another subset of target transcripts seems to be recognized by their specific fold regardless of the 5' phosphorylation state (Bouvier and Carpousis, 2011; Kime *et al.*, 2010; Mackie, 2013a). This mechanism has been termed 'direct entry' and may require certain stem loop structures in proximity of the cleavage site at least for some target transcripts (Schuck *et al.*, 2009; Kime *et al.* 2010). However, the structural requirements, if they apply, are so far unknown for most RNA substrates.

The recent discoveries of proteins mediating the turn-over of specific substrate RNAs that most likely belong to the class of substrates that are recognized by 'direct entry', provides an explanation of how substrate recognition might be achieved independently of the 5' phosphorylation state (Stoppel *et al.*, 2012; Suzuki *et al.*, 2006; Göpel *et al.*, 2013; this work). Examples for these proteins include the *E. coli* protein CsrD, which is involved in turn-over of sRNAs CsrB and CsrC, as well as the plastid protein RHON1 of *Arabidopsis thaliana* (Suzuki *et al.*, 2006; Stoppel *et al.*, 2012). These proteins may deliver RNA substrates to the membrane associated degradosome (Liou *et al.*, 2001; Stoppel *et al.*, 2012); remodel the RNA molecules for recognition by RNase E (Suzuki *et al.*, 2006), or even prime RNase E for cleavage, as proposed in this work. Hence, it is possible that various protein co-factors may contribute to substrate recognition by direct entry rather than the fold of the RNA molecules itself.

### **The impact of RNase E in riboregulation of gene expression**

Various sRNA/mRNA pairs have been shown to undergo coupled degradation after base-pairing (Massé *et al.*, 2003; Aiba, 2007; Caron *et al.*, 2010). In some cases, base-pairing of sRNAs with their target transcripts occurs in the vicinity of the Shine-Dalgarno sequence and thus prevents translation initiation. As a result, mRNAs that are devoid of translating polysomes are subject to rapid degradation (Baker and Mackie, 2003). Interestingly, while some mRNAs are degraded as a consequence of a block in translation, decay of other sRNA/mRNA complexes requires recruitment of the degradosome or even direct stimulation of the catalytic activity of RNase E (Ikeda *et al.*, 2011; Bandyra *et al.*, 2012, Mackie 2013b). Small RNA MicC for instance, was reported to directly stimulate decay of its target mRNA *ompD* by allosteric activation of RNase E activity with its 5' mono-phosphorylated terminus (Bandyra *et al.*, 2012). This process was suggested to rely on RNA chaperon Hfq for recruitment of the RNA degradosome. Later, it was shown that 5' mono-phosphorylated RNA and DNA oligonucleotides may act to *trans*-stimulate cleavage of complementary target RNAs by RNase E *in vitro*. Here, the catalytic domain of RNase E was sufficient for cleavage of the target and at least *in vitro* Hfq was not required for efficient recognition (Mackie, 2013b). Thus, sRNAs molecules themselves might act as allosteric activators of RNase E. Furthermore, RNase E was shown to be involved in lysine homeostasis by initiating decay of the *lysC* mRNA. Expression of the *lysC* transcript is regulated by a lysine-responsive riboswitch within its 5' UTR. Upon binding of the ligand, an alternative structure is formed sequestering the ribosome binding site and diminishing translation. Strikingly, this alternative structure also exposes two RNase E cleavage sites thereby inducing rapid decay of the *lysC* mRNA (Caron *et al.*, 2012). Hence, RNase E may also actively participate in riboregulation mediated by riboswitches that may expose or sequester cleavage sites depending on the availability of the corresponding ligand.

In contrast, some sRNAs stabilize their target transcripts and enhance gene expression rather than act as silencers (Fröhlich and Vogel, 2009; Zhao *et al.*, 2013). Many transcripts that are activated by small RNAs form inhibitory structures that mask the ribosome binding site and diminish translation. The relief of translational repression by disruption of these structures through base-pairing with the cognate sRNA leads to actively translating polysomes and concomitantly stabilizes the mRNA (Fröhlich and Vogel, 2009). However, sRNAs may also directly impair degradation of their target transcript or guide processing to alternative sites. For instance, sRNA SgrS was recently described to stabilize the *yigL* transcript encoding a sugar phosphatase by sequestering a RNase E processing site and thereby interfering with the decay of the mRNA (Papenfort *et al.*, 2013). This mechanism allows for expression of the stress induced YigL sugar phosphatase and subsequent detoxification of accumulated phosphosugars (Papenfort *et al.*, 2013). Interestingly, SgrS also mediates the response

to phosphosugar stress by gene silencing of sugar transporters (*ptsG* and *manX*) and encodes a small peptide SgrT. SgrT directly inactivates the glucose PTS permease PtsG by protein-protein interaction (Görke and Vogel, 2008; Vanderpool *et al.*, 2011). Gene silencing of both sugar transporters, *ptsG* and *manX*, was suggested to require Hfq for recruitment of the degradosome and subsequent induction of degradation by RNase E (Kawamoto *et al.*, 2005; Rice *et al.*, 2012).

In addition, processing by RNase E may generate mature sRNA species with an enhanced regulatory potential. For instance, sRNA MicX in *V. cholerae* was shown to be processed by RNase E in an Hfq-dependent manner. As a consequence, a shorter and more stable variant of MicX is generated that functions as a repressor of an outer membrane protein and the periplasmic subunit of a peptide ABC transporter (Davis and Waldor 2007). Moreover, processing by RNase E within an operon may generate shorter variants of the poly-cistronic transcript that may possess distinct half-lives. For example, RNase E cleaves within the *glmU* stop codon of the bi-cistronic *glmUS* message and leads to the formation of a fairly stable *glmU* mRNA and a rather unstable *glmS* transcript. Interestingly, the *glmS* transcript is subject to further riboregulation by two sRNAs, GlmY and GlmZ (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban and Vogel, 2008).

In sum, RNase E and the degradosome are key enzymes in riboregulation involved in gene silencing as well as stabilization and activation of transcripts.

### **1.5 Regulation by protein modifications: new roles for acetylation in bacteria**

Post-translational modifications target proteins and thus alter their activities. The by far most abundant modification in prokaryotes is protein phosphorylation. Specialized phosphorelays, transport and two-component systems rely on sequential phosphorylation events to transmit signals, generate adequate responses and transport preferred nutrients into the cell (Görke and Stülke, 2008; Jung *et al.*, 2012; Postma *et al.*, 1993). Phospho-enolpyruvate, ATP or acetyl-P serve as global phosphoryl-group donors and specific kinases catalyze phosphorylation events on histidine, aspartate, cysteine, threonine, serine and tyrosine residues (Soufi *et al.*, 2012). Phosphoryl-groups may be removed by phosphatases to reset signaling systems or alter the activity of enzymes and transcriptional regulators.

Modulation of protein activity by acetylation just recently emerged as a widespread post-translational modification in prokaryotes. In eukaryotic cells on the other hand, protein acetylation is thoroughly studied and predominantly serves to modify cellular activities at the epigenetic level, e.g. it influences gene expression levels and DNA repair through acetylation and de-acetylation of histones (Smith, 1991). This regulation is closely linked to the energetic status of the cell, as central

metabolic enzymes that generate acetyl-CoA as acetyl-group donor may localize to the nucleus and directly donate acetyl-CoA for histone modification (Hu *et al.*, 2010; Ladurner, 2009). Further, sirtuin deacetylases that function in histone deacetylation depend on NAD<sup>+</sup> and are thus highly sensitive to the metabolic state of the cell (Schwer and Verdin, 2008). Recent proteomic studies in *E. coli* and *Salmonella enterica* revealed a large number of acetylated proteins involved in various physiological processes (Wang *et al.*, 2010; Yu *et al.*, 2008; Zhang *et al.*, 2009). Thus, modification by protein acetylation fulfills diverse roles in bacteria: reversible acetylation may regulate the activity of central metabolic enzymes, such as acetyl-CoA synthase ACS or pyruvate dehydrogenase, influence chemotaxis by modification of response regulator CheY and enhance the activity of the RNA polymerase holoenzyme improving stress resistance (Lima *et al.*, 2011; Lima *et al.*, 2012; Starai *et al.*, 2002; Yan *et al.*, 2008).

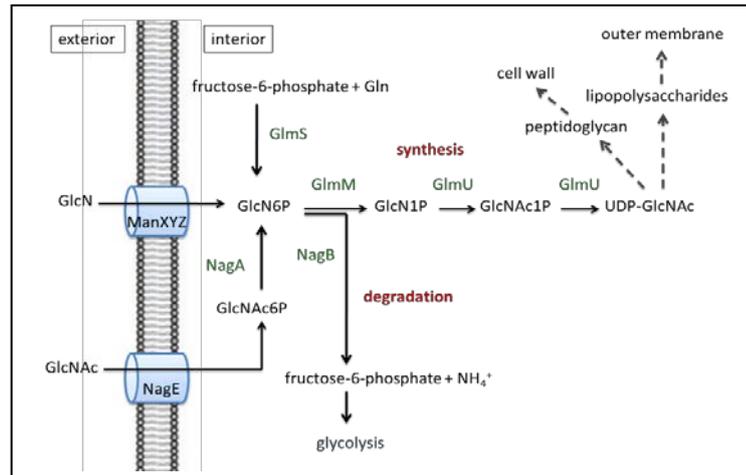
Homologs of the Gcn5-related acetyl-transferase YfiQ (Pat) are ubiquitously distributed among bacteria and were therefore suggested to be the major class of acetyl transferases (Hu *et al.*, 2010). These enzymes catalyze acetyl-CoA dependent acetylation of lysine-residues at the N<sup>ε</sup>-position, a process that can be reverted by deacetylases. In *E. coli* the sirtuin-related deacetylase CobB remains the only known deacetylase so far (Starai *et al.*, 2002; Weinert *et al.*, 2013). Reversible (de)acetylation by YfiQ (Pat) and CobB inversely regulates a number of target proteins involved in metabolism, chemotaxis, transcription initiation and even modulates activity of the core transcription machinery (Starai *et al.*, 2002; Hu *et al.*, 2010; Lima *et al.*, 2011). Interestingly, a recent study reported that the majority of acetylation observed *in vivo* may be attributed to non-enzymatic (auto-) acetylation of proteins depending on the high energy intermediate acetyl-phosphate as acetyl-group donor (Weinert *et al.*, 2013). This may also serve as a direct link between the acetylation state of various proteins and the metabolic fluxes within a bacterial cell. Whereas most acetylation events in bacteria seem to be independent of acetyl-transferase YfiQ, CobB is of major importance in deacetylation of chemically as well as enzymatically acetylated proteins (Weinert *et al.*, 2013).

Recently, it was reported that acetylation may also possess a role in the expression of sRNAs (Hu *et al.*, 2013). Expression of small RNA RprA, which is required for synthesis of the stress-responsive sigma factor RpoS upon osmotic shock, is dependent on the RcsC-RcsB phosphorelay system (Majdalani *et al.*, 2001; Majdalani *et al.*, 2002). Hu and colleagues demonstrated that response regulator RcsB is acetylated at multiple sites *in vivo* and suggested that acetylation at lysine residue 154 reduces DNA binding activity of RcsB. In turn, expression of *rprA* was observed to be diminished upon hyper-acetylation at this site. In contrast, deacetylase CobB was shown to positively influence *rprA* transcription, presumably by deacetylation of RcsB (Hu *et al.*, 2013). These observations by Hu

*et al.* are in concordance with previous *in vitro* studies demonstrating that deacetylated RcsB possesses a higher affinity for a DNA fragment encompassing the promoter sequence of the flagella biosynthesis master regulator *flhDC* as compared to acetylated RcsB (Thao *et al.*, 2010). Interestingly, a global proteomics study suggested protein RapZ (formerly YhbJ) to be acetylated *in vivo* (Zhang *et al.*, 2009). RapZ has been shown to mediate glucosamine-6-phosphate homeostasis (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Görke and Vogel, 2008; Göpel *et al.*, 2013; this work). In addition, RapZ was suggested to play an important role in regulation of the expression of sRNA GlmY (Reichenbach, 2009). In this work, we could demonstrate that RapZ indeed is acetylated *in vivo*. In contrast to regulation of *rprA* transcription by RcsB, acetylation of RapZ is required for activity of the  $\sigma^{54}$ -dependent *glmY* promoter and deacetylase CobB negatively regulates *glmY* expression (this work). Thus, acetylation of transcriptional regulators or other regulatory proteins may be more widespread in bacteria than previously thought and may also contribute to regulation of sRNA expression at the level of transcription.

### **1.6 Amino sugar metabolism in *E. coli***

In concordance with its significance for the bacterial cell, the amino sugar metabolism is subject to complex post-transcriptional regulation. Amino sugars are essential precursors for peptidoglycan and lipopolysaccharides, which are components of the bacterial cell wall and the outer membrane, respectively. Glucosamine-6-phosphate (GlcN6P) synthase GlmS is the key enzyme of the biosynthesis of activated amino sugar precursors and catalyzes the rate limiting step: *de novo* synthesis of GlcN6P and glutamate from fructose-6-phosphate and glutamine (Durand *et al.*, 2008; Milewski, 2002). GlcN6P is subsequently converted to glucosamine-1-phosphate (GlcN1P) by phosphoglucosamine mutase GlmM. The bi-functional enzyme GlmU possesses acetyl-transferase and uridyl-transferase activity and converts GlcN1P in a two-step reaction first into N-acetylglucosamine-1-phosphate (GlcNAc1P) and finally into the activated amino sugar UDP-N-acetylglucosamine (UDP-GlcNAc), which serves as dedicated precursor for peptidoglycan and lipopolysaccharide biosynthesis (Fig. 1.3; Mengin-Lecreux and van Heijenoort, 1993; Mengin-Lecreux and van Heijenoort, 1994; Mengin-Lecreux and van Heijenoort, 1996).



**Figure 1.3: UDP-GlcNAc biosynthesis pathway in *E. coli* (modified from Plumbridge, 1995; Plumbridge and Vimr, 1999).**

GlcN6P is an essential metabolite in the bacterial cell. *De novo* synthesis of GlcN6P is catalyzed by glucosamine-6-P synthase GlmS. Subsequently GlcN6P is converted into the activated amino sugar UDP-*N*-acetylglucosamine (UDP-GlcNAc) by enzymes GlmM and GlmU. Intermediates in these reactions are glucosamine-1-phosphate (GlcN1P) and *N*-acetylglucosamine-1-phosphate (GlcNAc1P). UDP-GlcNAc serves as essential precursor for biosynthesis of the bacterial cell wall and the lipopolysaccharides of the outer membrane. As opposed to *de novo* synthesis, external amino sugars can be internalized and utilized in this pathway, omitting the rate limiting step of GlcN6P synthesis. For example, glucosamine (GlcN) and *N*-acetylglucosamine (GlcNAc) are substrates of the PTS transporters ManXYZ and NagE, respectively (Postma *et al.*, 1993). Upon uptake the amino sugars are phosphorylated. GlcNAc6P is subsequently converted to GlcN6P by enzyme NagA. In addition to its role in biosynthesis of peptidoglycan and lipopolysaccharides, GlcN6P may serve as carbon- and nitrogen source. Enzyme NagB initiates degradation of GlcN6P by converting it to ammonium and fructose-6-phosphate, which is then channeled into glycolysis for further degradation.

In addition to *de novo* synthesis of amino sugars initiated by enzyme GlmS, bacterial cells possess a variety of uptake systems for various amino sugars. For instance, glucosamine (GlcN) and *N*-acetylglucosamine are substrates for the PTS-permeases ManXYZ and NagE, respectively (Fig. 1.3; (Postma *et al.*, 1993). Furthermore, chitobiose, *N*-acetylmannosamine and *N*-neuraminic acid can also be internalized by bacterial cells and converted to GlcN6P (Plumbridge, 1995; Plumbridge and Vimr, 1999; Keyhani *et al.*, 2000a; Keyhani *et al.*, 2000b). Aside from their role as essential precursors, amino sugars also serve as carbon and nitrogen sources, supplying the cell with energy. GlcN6P is an important metabolite interconnecting the anabolic and catabolic cycles of the metabolism of amino sugars. Degradation of GlcN6P is initiated by deaminase NagB, which converts GlcN6P to fructose-6-phosphate and ammonium. Fructose-6-phosphate is subsequently channeled into glycolysis for further degradation (Fig. 1.3).

Hence, in the presence of an ample supply of external amino sugars, the enzymatic activity of GlmS is dispensable. However, GlmS becomes essential in absence of other sources for amino sugars. In contrast, GlmM and GlmU are essential under all conditions, as they provide the cell with activated precursors for cell wall biosynthesis. In *E. coli*, *glmU* and *glmS* are encoded in the same operon. Differential expression of both genes is achieved by an elaborate post-transcriptional mechanism employing a cascade of two regulatory sRNAs, GlmY and GlmZ, and the RNA-binding protein RapZ (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Görke and Vogel, 2008; Göpel *et al.*, 2013; this work).

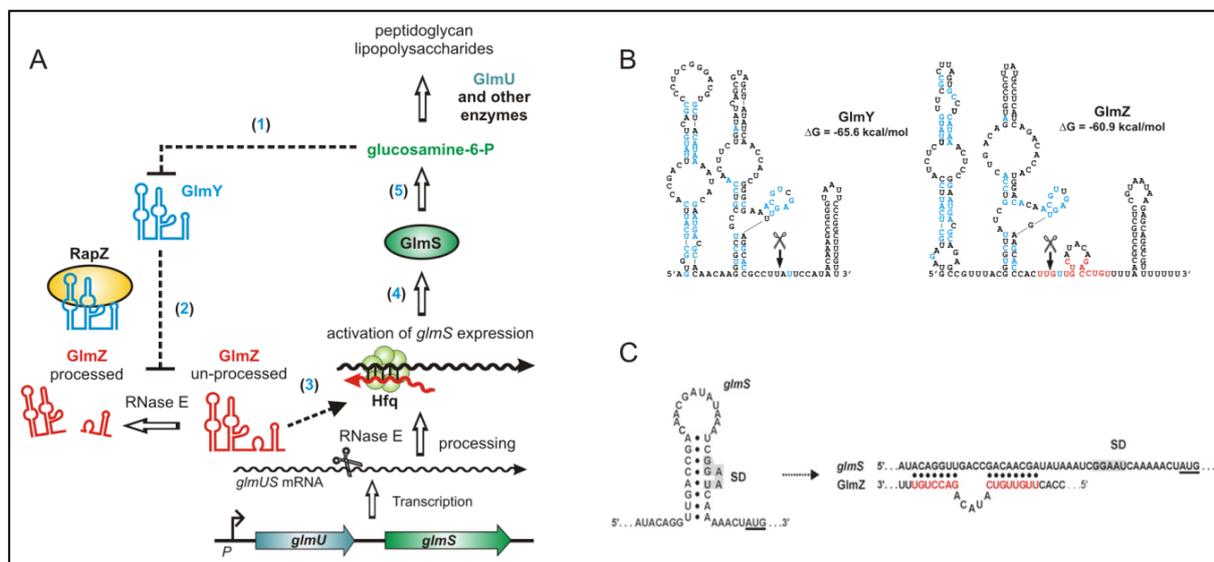
### **1.7 The regulatory GlmYZ small RNA cascade**

Prior to the discovery of regulatory small RNAs, it was a long-standing mystery how differential gene expression is achieved within the *glmUS* operon. After transcription, the bi-cistronic *glmUS* message is subject to processing by RNase E occurring within the *glmU* stop codon. Whereas the *glmU* transcript is fairly stable, the *glmS* mRNA harbors an inhibitory hairpin within the 5' UTR masking the ribosomal binding site and is subject to complex feedback regulation (Fig. 1.4 C; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Görke and Vogel, 2008; Fröhlich and Vogel, 2009). As a consequence of the inaccessibility of the Shine-Dalgarno sequence, the *glmS* mRNA is only sporadically translated. Therefore, the *glmS* transcript may presumably be prone to rapid degradation, as the absence of actively translating ribosomes was suggested to be a cause for decay of translationally blocked mRNAs in *E. coli* (Baker and Mackie, 2003). However, under GlcN6P limiting conditions, the *glmS* transcript is activated by base-pairing with small RNA GlmZ. This interaction requires RNA chaperon Hfq and leads to disruption of the inhibitory structure and release of the ribosome binding site (Fig. 1.4 C; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban and Vogel, 2008; Görke and Vogel, 2008). Subsequently, GlmS is synthesized and replenishes the GlcN6P pool in the cell.

Under conditions of ample GlcN6P, the ~207 nt-long sRNA GlmZ is subject to processing at position ~155, which removes most of the nucleotides required for base-pairing with *glmS* and initiates decay of GlmZ (Fig. 1.4 B; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Görke and Vogel, 2008). Processing of GlmZ is dependent on RNase E and requires protein RapZ (formerly YhbJ; Fig. 1.4 A; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Reichenbach, 2009). However, the molecular mechanism of GlmZ decay remained elusive and was just recently unraveled (Göpel *et al.*, 2013; this work).

A second small RNA, GlmY, was shown to indirectly activate the *glmS* transcript by stabilizing sRNA GlmZ (Reichenbach *et al.*, 2008; Urban *et al.*, 2007; Urban and Vogel, 2008). The homologous sRNAs GlmY and GlmZ are highly similar in sequence and structure, as they share a sequence identity of 63%. As opposed to GlmZ, the nucleotides required for base-pairing with *glmS* are lacking in GlmY (Fig. 1.4 B; Reichenbach *et al.*, 2008). GlmY is also subject to processing, however, the responsible enzyme has not been identified yet. In contrast to GlmZ, which is inactivated by processing, the processed variant of GlmY is the more abundant species and seems to be the active form of the sRNA *in vivo* (Reichenbach *et al.*, 2008; Reichenbach *et al.*, 2009; Göpel *et al.*, 2013). Upon depletion of intracellular GlcN6P, sRNA GlmY accumulates by a post-transcriptional mechanism and counteracts processing of GlmZ (Reichenbach *et al.*, 2008; Reichenbach *et al.*, 2009; Göpel *et al.*, 2013). This process involves the novel RNA-binding protein RapZ that is absolutely required for processing of GlmZ *in vivo* and *in vitro* (Fig. 1.4 A; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Göpel *et al.*,

2013; this work). As GlmZ is stabilized by the highly homologous sRNA GlmY and protein RapZ was shown to act downstream of GlmY and upstream of GlmZ, it was speculated that GlmY might act by titration of RapZ (Görke and Vogel, 2008), reminiscent of the mechanism underlying regulation within the Csr-system (Babitzke and Romeo, 2007). However, the molecular mechanism of signal transduction within the GlmYZ sRNA cascade was just recently clarified within the scope of this work (Göpel *et al.*, 2013).



**Figure 1.4: A cascade composed of two hierarchically acting sRNAs regulates GlcN6P homeostasis in *E. coli*.**

**A.** Regulation of *glmS* expression by the GlmYZ sRNA cascade. Gene *glmS* is encoded in one operon with the essential gene *glmU*. The *glmUS* co-transcript is subsequently processed by RNase E within the *glmU* stop codon. Processing generates an unstable *glmS* transcript carrying an inhibitory structure within its 5' UTR (Fig. 1.3 C). Upon depletion of intracellular GlcN6P, GlmY accumulates (1) and counteracts processing of sRNA GlmZ by RNase E in a process that involves protein RapZ (2). GlmZ is stabilized in its active form and can activate the *glmS* mRNA by base-pairing with the aid of RNA chaperon Hfq (3). GlmS is synthesized (4) and replenishes the GlcN6P pool of the cell (5). In turn, ample amounts of GlcN6P feedback inhibit accumulation of GlmY and down-regulate the cascade (Göpel and Görke, 2012b). **B.** Secondary structures of sRNAs GlmY and GlmZ. Fully conserved residues are highlighted in blue, residues involved in base-pairing with the *glmS* 5'UTR are indicated in red. Scissors mark the processing sites within GlmY and GlmZ (Fig. adapted from Göpel *et al.*, 2013). **C.** An inhibitory stem loop structure sequesters the Shine-Dalgarno (SD) sequence of the *glmS* mRNA. Upon base-pairing with sRNA GlmZ aided by Hfq, the inhibitory structure is disrupted and the Shine-Dalgarno region is released. Concomitantly, *glmS* translation is activated and the mRNA is stabilized. As the base-pairing sites in *glmS* and GlmZ are highly conserved in enterobacteria, it is conceivable that the regulatory mechanisms activating *glmS* expression is conserved among these bacteria (Fig. modified from Görke and Vogel, 2008).

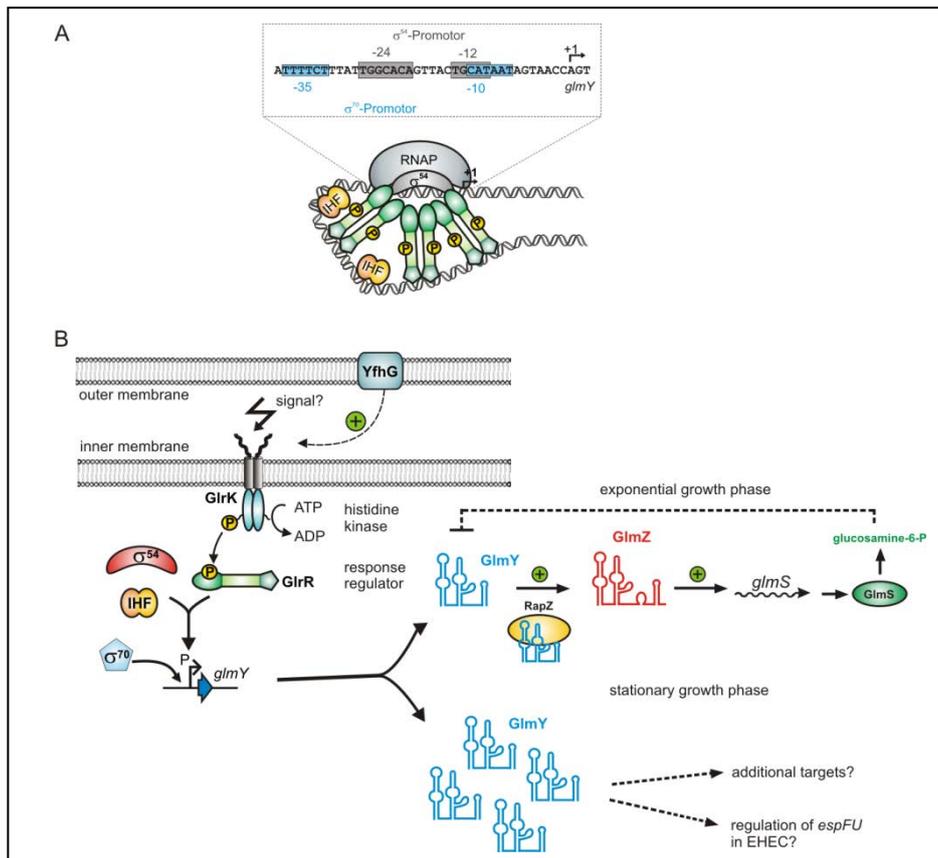
In this work, protein RapZ was identified as a novel RNA-binding protein that specifically binds sRNAs GlmY and GlmZ. Interestingly, RapZ switches its sRNA binding partners depending on the GlcN6P concentration *in vivo*. In addition, we clarified the role of RapZ in processing of sRNA GlmZ by RNase E and demonstrated that GlmY serves as sRNA mimicry, thereby counteracting processing of GlmZ by sequestering RapZ (Göpel *et al.*, 2013; this work).

### **1.8 The GlrK/GlrR two-component system controls expression of small RNA GlmY in *E. coli***

As small RNAs are involved in virtually every physiological process within the bacterial cell, the expression of sRNAs is tightly regulated at the level of biosynthesis and/or decay. In *E. coli*, expression of the gene encoding sRNA GlmY is driven by two perfectly overlapping promoters: a  $\sigma^{70}$ -

dependent promoter, which is active mainly during exponential growth, and a  $\sigma^{54}$ -dependent promoter that is strongly activated upon entry of stationary phase (Fig. 1.5<sup>°</sup>A; Reichenbach *et al.*, 2009). As mentioned above, transcription from  $\sigma^{54}$ -dependent promoters requires specific activator proteins for open complex formation and transcription initiation (Hoover *et al.*, 1990; Wigneshweraraj *et al.*, 2008). Response regulator GlrR of the GlrK/GlrR two-component system that is encoded directly downstream of the *glmY* gene, serves as the specific activator protein for the  $\sigma^{54}$ -dependent *glmY* promoter (Fig. 1.5; Reichenbach *et al.*, 2009). Upon signal perception, histidine kinase GlrK auto-phosphorylates with ATP. It was previously suggested that kinase GlrK (QseE) senses epinephrine in enterohemorrhagic *E. coli* (Reading *et al.*, 2009), however GlrK of *E. coli* K12 does not seem to respond to this compound (Reichenbach and Görke, unpublished). Thus, the signal sensed by GlrK in non-pathogenic *E. coli* remains to be identified. Following auto-phosphorylation of GlrK, the phosphoryl-group is transmitted to the conserved aspartate at position 56 in response regulator GlrR. In turn, phosphorylated GlrR binds to three conserved sequence motifs far upstream of the *glmY* promoter and activates transcription (Fig. 1.5; Reichenbach *et al.*, 2009; Göpel *et al.*, 2011; this work). In this work, we demonstrated that DNA bending by integration host factor IHF is required for activation of *glmY* transcription by the  $\sigma^{54}$ -dependent promoter. IHF presumably facilitates interaction between the GlrR hexamer and the  $\sigma^{54}$ -RNA polymerase holoenzyme (Fig. 1.5 A).

Whereas GlmY mediates GlcN6P homeostasis in exponential growth phase, it is speculated that GlmY serves additional functions in stationary phase (Göpel and Görke, 2012a). The GlrK/GlrR two-component system (alternative names QseE/QseF) has been implicated in virulence in enterohemorrhagic *E. coli* (EHEC) by somehow modulating expression of the *espFU* effector required for pedestal formation. However, there is no evidence of GlrR binding sites in proximity of the  $\sigma^{70}$ -dependent *espFU* gene, nor does GlrR bind to DNA-fragments encompassing the *espFU* promoter region (Reading *et al.* 2007). Thus, it is tempting to speculate that regulation of additional genes may occur indirectly via GlmY, since GlmY seems to be the only direct target of the GlrK/GlrR two-component system in *Escherichia* species as suggested by *in silico* analysis (Fig. 1.5; Reichenbach *et al.*, 2009; Göpel *et al.*, 2011; Göpel and Görke, 2012a; Göpel and Görke, unpublished).



**Figure 1.5: Regulation of *glmY* expression by two perfectly overlapping promoters and the GlrK/GlrR TCS in *E. coli*.**

**A.** A  $\sigma^{70}$ -dependent promoter and a perfectly overlapping  $\sigma^{54}$ -dependent promoter control expression of *glmY* in *E. coli*. During exponential phase, expression of *glmY* from its  $\sigma^{70}$ -dependent promoter is sufficient to ensure GlcN6P homeostasis mediated by GlmY within the GlmYZ regulatory sRNA cascade. The  $\sigma^{54}$ -dependent promoter requires response regulator GlrR and integration host factor IHF for activity (Fig. adapted from Göpel and Görke, 2012b). **B.** In non-pathogenic *E. coli*, the GlrK/GlrR two-component system is activated by a so far unknown signal and enhances *glmY* transcription from its  $\sigma^{54}$ -dependent promoter upon transition to stationary phase. Outer membrane protein YfhG (QseG) was suggested to be functionally linked with the GlrK/GlrR TCS as *yfhG* is encoded within the *glrK-yfhG-glrR* operon although the nature of this connection remained elusive (Reading *et al.*, 2009). The accumulation of GlmY in stationary phase may hint at further regulatory functions of this sRNA possibly by modifying expression of further targets. It is tempting to speculate that the link between the GlrK/GlrR TCS and the indirect target *espFU* is provided by sRNA GlmY (Reading *et al.*, 2009; Göpel and Görke, 2012a).

With this work, we are expanding the regulatory network controlling *glmY* expression by identifying sirtuin deacetylase CobB as an important regulator of *glmY* transcription. Furthermore, we will show in more detail that protein RapZ is absolutely required for the activity of the  $\sigma^{54}$ -dependent *glmY* promoter and that RapZ is the target of acetylation *in vivo*.

### The GlrK/GlrR two-component system is conserved among *Enterobacteriaceae*

Initial sequence analysis predicted conserved GlrR binding sites in the proximity of *glmY* genes of all enterobacterial species that possess genes encoding the sRNAs GlmY and GlmZ. In contrast to *E. coli* and its close relatives *Shigella* and *Klebsiella*, expression of the *glmZ* gene was also predicted to be regulated by the GlrK/GlrR two-component system in most *Enterobacteriaceae*. Interestingly, *in silico* analysis categorized the promoter architectures of all analyzed species into three groups: (I) both *glmY* and *glmZ* are solely transcribed from  $\sigma^{54}$ -dependent promoters and both promoters possess GlrR binding sites; (II) overlapping  $\sigma^{54}$ - and  $\sigma^{70}$ -dependent promoters drive transcription of *glmY* and

*glmZ* and the  $\sigma^{54}$ -dependent promoters seem to require GlrR; (III) only *glmY* is transcribed from overlapping  $\sigma^{54}$ - and  $\sigma^{70}$ -dependent promoters, while expression of *glmZ* is driven by a  $\sigma^{70}$ -dependent promoter. In the latter case, GlrR binding sites were predicted for the overlapping *glmY* promoter, but not for  $\sigma^{70}$ -promoter controlling *glmZ* transcription (Reichenbach *et al.*, 2009).

In this work, we verified the predictions by extending our investigation of transcriptional control of *glmY* and *glmZ* towards further species, i.e. *Yersina pseudotuberculosis* YPIII, as a representative for the first group, and *Salmonella enterica* serovar Thyphimurium LT2 as a representative of the second group. We also studied *glmZ* expression in *E. coli*, revealing that *Escherichia* species and their close relatives represent the last group.

### **1.9 Aim of this study**

In *Enterobacteriaceae*, small RNAs GlmY and GlmZ mediate GlcN6P homeostasis and thus provide an essential function ensuring survival and prosperity of the cell. The elaborate regulatory circuit composed of two homologous and hierarchically acting sRNAs and the RNA-binding protein RapZ (YhbJ) is unprecedented. The main focus of this work was to unravel the molecular mechanism underlying this unique cascade. Whereas it was previously predicted that RapZ may be a novel kind of RNA-binding protein (Kalamorz, 2008), it was unclear whether RapZ itself might be a nuclease or how RapZ might influence processing of GlmZ. In this respect, the role of RapZ and its mode of action were subject of our investigations within the scope of this work. In order to explore the function of RapZ and the molecular mechanism within the GlmYZ-cascade, a set of *in vivo* and *in vitro* techniques were established. An *in vivo* approach was established to investigate differential sRNA-binding upon ample GlcN6P supply and depletion using physiological amounts of epitope-tagged RapZ. Furthermore, the GlmYZ cascade was reconstituted *in vitro* to study the mechanism of signal transduction from GlmY to GlmZ in more detail. This system also provides the possibility to study the function of RapZ and the potential of possible additional effectors.

In addition to decay, small RNAs are elaborately regulated at the level of synthesis. It becomes more and more apparent that two-component systems and small RNAs are tightly interconnected within regulatory networks. The GlrK/GlrR two-component system was recently shown to direct expression of sRNA GlmY from its  $\sigma^{54}$ -dependent promoter (Reichenbach *et al.*, 2009). Interestingly, the small RNAs as well as the two-component system are highly conserved among *Enterobacteriaceae* (Reichenbach *et al.*, 2009). Thus, to expand the knowledge about the potential GlrK/GlrR regulon, the regulation of transcription studied for *glmY* and *glmZ* of representatives of other enterobacterial groups. Furthermore, additional factors involved in transcriptional control of the  $\sigma^{54}$ -dependent *glmY* promoter were identified in *E. coli* and their functions were analyzed.

## Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in *Enterobacteriaceae*.

*The results described in this chapter are published in:*

Göpel, Y., Lüttmann D., Heroven AK., Reichenbach B., Dersch P., and Görke B. (2011) Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in *Enterobacteriaceae*. *Nucleic Acids Res.* **39**: 1294-1309

*Author contributions:*

This study was designed by B.R., Y.G. and B.G.. B.R. and B.G. performed sequence alignment analysis identifying the *glmY* and *glmZ* promoters and gene synteny of the *glmY* and *glmZ* regions in different *enterobacteriaceae*. A.K.H. performed  $\beta$ -galactosidase measurements in *Y. pseudotuberculosis*. P.D. provided *Y. pseudotuberculosis* DNA, insightful discussion and advice for the manuscript. D.L. conducted  $\beta$ -galactosidase measurements of truncated *glmZ* promoter *lacZ*-fusions. Y.G. performed protein purifications, EMSAs and *in vitro* acetylation assays.  $\beta$ -galactosidase measurements of chromosomal sRNA-*lacZ* fusions were performed by Y.G. assisted by Sabine Lentjes or by intern students under the supervision of D.L., Y.G. and B.R.. Strains and plasmids were constructed by D.L., Y.G. or intern students under the supervision of D.L., Y.G. and B.R.. D.L., Y.G. and B.G. wrote the paper.

*The copy right lies with the authors.*

**ABSTRACT**

Small RNAs GlmY and GlmZ compose a cascade that feedback-regulates synthesis of enzyme GlmS in *Enterobacteriaceae*. Here, we analyzed the transcriptional regulation of *glmY/glmZ* from *Yersinia pseudotuberculosis*, *Salmonella typhimurium* and *Escherichia coli*, as representatives for other enterobacterial species, which exhibit similar promoter architectures. The GlmY and GlmZ sRNAs of *Y. pseudotuberculosis* are transcribed from  $\sigma^{54}$ -promoters that require activation by the response regulator GlrR through binding to three conserved sites located upstream of the promoters. This also applies to *glmY/glmZ* of *S. typhimurium* and *glmY* of *E. coli*, but as a difference additional  $\sigma^{70}$ -promoters overlap the  $\sigma^{54}$ -promoters and initiate transcription at the same site. In contrast, *E. coli glmZ* is transcribed from a single  $\sigma^{70}$ -promoter. Thus, transcription of *glmY* and *glmZ* is controlled by  $\sigma^{54}$  and the two-component system GlrR/GlrK (QseF/QseE) in *Y. pseudotuberculosis* and presumably in many other Enterobacteria. However, in a subset of species such as *E. coli* this relationship is partially lost in favor of  $\sigma^{70}$ -dependent transcription. In addition, we show that activity of the  $\sigma^{54}$ -promoter of *E. coli glmY* requires binding of the integration host factor to sites upstream of the promoter. Finally, evidence is provided that phosphorylation of GlrR increases its activity and thereby sRNA expression.

**INTRODUCTION**

Post-transcriptional gene regulation involving regulatory RNAs has emerged as a widespread principle occurring in all three domains of life. In bacteria, one important mode of riboregulation involves *trans*-encoded small RNAs (sRNAs), which appear to be involved in regulation of almost every important physiological function (Brantl, 2009; Görke and Vogel, 2008; Guillier and Gottesman, 2006; Guillier *et al.*, 2006; Repoila and Darfeuille, 2009; Vanderpool, 2007). The majority of sRNAs acts by base-pairing with target mRNAs usually in the vicinity of the ribosome binding site (Brantl, 2009; Waters and Storz, 2009). Most often, this interaction represses translation and/or stimulates mRNA degradation, although a few cases are known where sRNA-mRNA interaction increases gene expression (Fröhlich and Vogel, 2009). One example is provided by the sRNA GlmZ in *Escherichia coli*. Binding of GlmZ to its target mRNA *glmS* destroys an inhibitory stem loop that sequesters the Shine-Dalgarno sequence of *glmS*. GlmZ is also an unusual case, because it works in concert with a second homologous sRNA, GlmY (Görke and Vogel, 2008; Reichenbach *et al.*, 2008; Urban and Vogel, 2008). However, while other homologous sRNAs regulate their targets redundantly or additively (Waters and Storz, 2009), GlmY/GlmZ act hierarchically to activate expression of the *glmS* gene, which encodes glucosamine-6-phosphate (GlcN6P) synthase GlmS. GlmS catalyzes formation of GlcN6P, which initiates the pathway that generates precursors of cell wall synthesis. Of both sRNAs, only GlmZ is able to base-pair with *glmS* mRNA. However, ongoing processing removes most of the base-pairing residues and thereby inactivates GlmZ. Upon depletion of GlcN6P, the second sRNA GlmY accumulates and counteracts processing of GlmZ. This activates synthesis of GlmS, which re-synthesizes GlcN6P. Hence, both sRNAs work in a cascade to mediate feedback control of GlmS (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban *et al.*, 2007; Urban and Vogel, 2008).

To understand the impact of sRNAs on bacterial physiology, it is important to identify the signals and mechanisms that control expression of a particular sRNA. sRNA transcription is often controlled by transcriptional regulatory proteins similar to that of protein-coding genes (for an overview, see (Brantl, 2009)). Some sRNA genes are controlled by two-component systems (TCS) and/or alternative

sigma factors, which are the key devices for perception of environmental signals and their conversion into gene expression changes (Guillier *et al.*, 2006; Valverde and Haas, 2008; Vogel, 2009). Evidence is accumulating that sRNAs are also members of the modulon controlled by  $\sigma^{54}$  involving genes important for nitrogen and carbon-utilization, uptake of metal ions, stress responses and other apparently unrelated functions. Transcription of sRNA genes from  $\sigma^{54}$ -dependent promoters has been demonstrated in *Pseudomonas aeruginosa* and *Vibrio harveyi* (Sonnleitner *et al.*, 2009; Tu and Bassler, 2007). It is estimated that there are ~70  $\sigma^{54}$ -dependent promoters in *E. coli* (Reitzer and Schneider, 2001; Zhao *et al.*, 2010).  $\sigma^{54}$  is unique among  $\sigma$  factors since it is not related to other  $\sigma$  factors and recognizes a different sequence composed of -24/-12 motifs (Wigneshweraraj *et al.*, 2008). The  $\sigma^{54}$ -RNAP holo-enzyme is unable to catalyze formation of the open promoter complex. This reaction requires interaction with an activator protein that usually binds to activating binding sites (ABS) located far upstream of the promoter.

Despite the parallels in the transcriptional control of protein-coding and sRNA genes, there appears to be at least one difference: Many protein-coding genes are transcribed from multiple promoters that can be activated by different  $\sigma$  factors and use different transcriptional start sites (Mendoza-Vargas *et al.*, 2009; Wade *et al.*, 2006). While differing 5' sequences of mRNAs are without consequences for the nature of the encoded protein, they have functional consequences for sRNAs as shown for the *IstR-1* and *IstR-2* sRNAs, which are transcribed from consecutive promoters (Darfeuille *et al.*, 2007). To allow transcription of identical sRNA species from alternative promoters, these promoters must overlap to allow transcription initiation at the same nucleotide. Such an unorthodox arrangement has recently been identified for the *E. coli glmY* gene, where overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters start transcription at the same site (Reichenbach *et al.*, 2009). The  $\sigma^{54}$ -promoter requires activation by the TCS GlrR/GlrK (alternative names: QseF/QseE or YfhA/YfhK), which is encoded downstream of *glmY* and transcribed independently (Reichenbach *et al.*, 2009). The activator protein GlrR consists of an N-terminal response regulatory domain, a central  $\sigma^{54}$ -interaction module and a C-terminal helix-turn-helix DNA-binding motif. GlrR binds three TGTCN<sub>10</sub>GACA motifs located more than 100 bp upstream of *glmY* and thereby activates the  $\sigma^{54}$ -promoter, while activity of the  $\sigma^{70}$ -promoter is unaffected. Both promoters are moderately active during the exponential growth phase. Their activities interfere since binding of  $\sigma^{54}$  represses activity of the overlapping  $\sigma^{70}$ -promoter to some extent (Reichenbach *et al.*, 2009).

In this work, we analyzed the transcriptional regulation of *glmY* and *glmZ*. The TCS GlrR/GlrK as well as *glmY* and *glmZ* are conserved in *Enterobacteriaceae*. *In silico* analyses of the *glmY* and *glmZ* promoter sequences identified three groups within the enterobacterial species. *Yersinia pseudotuberculosis*, *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 and *E. coli* K12 are representatives of each group and were analyzed. We show that both, *glmY* and *glmZ*, are controlled by GlrR and  $\sigma^{54}$  in *Y. pseudotuberculosis*, *S. typhimurium* and presumably in many other species. In these species, both sRNAs are expressed from  $\sigma^{54}$ -dependent promoters that require activation by GlrR. However, overlapping  $\sigma^{70}$ -promoters additionally contribute to expression in *S. typhimurium*. In *E. coli*, *glmY* is transcribed from overlapping  $\sigma^{54}$ - and  $\sigma^{70}$ -promoters, while *glmZ* is expressed from a single  $\sigma^{70}$ -promoter that is constitutively active. In conclusion, *glmY* and *glmZ* appear to be strictly  $\sigma^{54}$ -dependent genes in one subgroup of *Enterobacteriaceae*, while  $\sigma^{54}$ -dependency is lost in favor of unregulated  $\sigma^{70}$ -promoters in a second subgroup. Furthermore, we show for *E. coli glmY* that activity of the  $\sigma^{54}$ -promoter requires the integration host factor IHF, which

presumably binds to two conserved sites flanking the proximal GlrR binding site. Finally, our data indicate that phosphorylation of GlrR increases its affinity for its target sites on the DNA.

## **MATERIALS AND METHODS**

### ***Growth conditions and strains***

LB was used as standard medium for cultivation of bacteria. *E. coli* and *S. typhimurium* LT2 were grown routinely under agitation (200 r.p.m.) at 37°C and *Y. pseudotuberculosis* YPIII was cultivated at 25°C. When necessary, antibiotics were added to the medium (ampicillin 100 µg/ml, kanamycin 30 µg/ml, chloramphenicol 15 µg/ml, spectinomycin 50 µg/ml). For induction of the *P<sub>Ara</sub>* promoter on pBAD plasmids, 0.2% L-arabinose was added. The *E. coli* strains are listed in Table 1, including a description of their relevant genotypes. The  $\Delta$ *ihfA::kan*, and  $\Delta$ *ihfB::kan* alleles were transduced to strains Z190 and Z197 using bacteriophage T4GT7 (Wilson *et al.*, 1979). Most of the *lacZ* reporter fusions used in this study were first established on plasmids and subsequently integrated into the  $\lambda$ *attB*-site on the *E. coli* chromosome by site-specific recombination yielding the strains as indicated in Table 2.1. Recombination was achieved using helper plasmid pLDR8 as described (Diederich *et al.*, 1992). Briefly, origin-less DNA-fragments encompassing the respective *lacZ* fusion, the *aadA* spectinomycin resistance gene and the  $\lambda$ *attP*-site were isolated by BamHI digestion and agarose gel-electrophoresis. The DNA-fragments were self-ligated and subsequently introduced into target strains carrying the temperature-sensitive  $\lambda$ -integrase expression plasmid pLDR8. Recombinants were obtained by selection on spectinomycin-plates at 42°C. Correct integration was verified by PCR using appropriate primers and loss of plasmid pLDR8 was confirmed by sensitivity to kanamycin.

### ***Construction and site-directed mutagenesis of plasmids***

DNA-cloning was carried out in *E. coli* strain DH5 $\alpha$  following standard procedures. The plasmids and oligonucleotides used in this study are listed in Tables S2.1 and S2.2, respectively (see “Supplementary data”). Plasmid constructions are also described under “Supplementary data”.

### ***Analysis of glmY and glmZ transcription ( $\beta$ -galactosidase assays)***

Overnight cultures of *E. coli* were inoculated into fresh LB medium to an OD<sub>600</sub> of 0.1 and grown to an OD<sub>600</sub> of 0.5-0.7. Subsequently, the cells were harvested and the  $\beta$ -galactosidase activities were determined as previously described (Miller, 1972).  $\beta$ -Galactosidases activities were determined from *Y. pseudotuberculosis* cells as described recently (Heroven *et al.*, 2008). The presented values are the average of at least three measurements using independent cultures.

### ***Protein purification***

C-terminally His-tagged *E. coli* and *Y. pseudotuberculosis* GlrR proteins were overproduced in *E. coli* DH5 $\alpha$  carrying plasmid pBGG219 or pBGG397, respectively. Cells were grown in 1 l LB-ampicillin to an OD<sub>600</sub> = 0.5-0.8. After addition of 1 mM IPTG for the induction of GlrR::His<sub>10</sub> synthesis, growth was continued for one additional hour. Cells were harvested and washed in ZAP-buffer (50 mM Tris-HCl, 200 mM NaCl, pH 7.5). The crude lysate was prepared using a one shot cell disrupter at 2600 psi (Constant systems Ltd.) and subsequently cleared by low speed centrifugation followed by ultracentrifugation. The cleared lysates were loaded onto pre-equilibrated Ni-NTA Superflow columns (Qiagen) and proteins were eluted with a gradient of imidazol solved in ZAP buffer. Samples of the different purification steps and elution fractions were analyzed by SDS-polyacrylamide gel electrophoresis and Coomassie blue staining. The 250 mM imidazol fractions contained the pure GlrR-His<sub>10</sub> proteins. These fractions were dialysed two times for 24 h against buffer (20 mM Tris-HCl pH 7.5, 100 mM KCl, 2 mM DTT). In the second dialysis step, the buffer additionally contained 25% (v/v) glycerol. The purified proteins were aliquoted and stored at -20°C until their use.

**Electrophoretic mobility shift assay (EMSA)**

EMSAs were carried out as described previously (Reichenbach *et al.*, 2009; Stratmann *et al.*, 2008). The DNA fragments tested in the EMSAs were amplified by PCR using the same oligonucleotides that were used for construction of the corresponding *glmY'*- and *glmZ'-lacZ* gene fusions (Table S1). The 200 bp and 400 bp *lacZ* promoter fragments, which were used as internal controls, were generated by PCR using primer pairs BG580/BG581 and BG578/BG579, respectively. DNA concentrations were determined with the NanoDrop Spectrometer ND-1000 (Peqlab). Binding assays were carried out in 10  $\mu$ l volume containing binding buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2 mM DTT, 10% glycerol), 30 ng of each DNA-fragment, and the protein concentrations as indicated in the Figures. The reactions were incubated at 30°C for 20 min and subsequently 6  $\mu$ l of the samples were separated at 4°C alongside with a DNA size marker on non-denaturing 8% acrylamide gels prepared in 0.5 $\times$ TBE. Gels were stained with ethidium bromide for visualization of the DNA. For testing the effect of acetyl phosphate, GlrR protein was incubated for 1 h at 37°C in binding buffer containing 50 mM acetyl phosphate and then used for the binding assays.

**Table 2.1.** *E. coli* strains used in this study

| Name         | Genotype   | Reference or construction                    |
|--------------|--|--|
| DH5 $\alpha$ | $\phi$ 80d <i>lacZ</i> $\Delta$ M15, <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> ( $r_{\kappa^-}$ , $m_{\kappa^+}$ ), <i>supE44</i> , <i>relA1</i> , <i>deoR</i> , $\Delta$ ( <i>lacZYA-argF</i> ) U169 | Laboratory collection                        |
| JW0895       | F-, $\Delta$ ( <i>araD-araB</i> )567, $\Delta$ <i>lacZ4787</i> :: <i>rrnB-3</i> , $\Delta$ <i>ihfB735</i> :: <i>kan</i> , <i>LAM-</i> , <i>rph-1</i> , $\Delta$ ( <i>rhaD-rhaB</i> )568, <i>hsdR514</i>                                | (Baba <i>et al.</i> , 2006)                  |
| JW1702       | $\Delta$ ( <i>araD-araB</i> )567, $\Delta$ <i>lacZ4787</i> :: <i>rrnB-3</i> , $\lambda$ -, $\Delta$ <i>ihfA786</i> :: <i>kan</i> , <i>rph-1</i> , $\Delta$ ( <i>rhaD-rhaB</i> )568, <i>hsdR514</i>                                     | (Baba <i>et al.</i> , 2006)                  |
| R1279        | CSH50 $\Delta$ ( <i>pho-bgl</i> )201 $\Delta$ ( <i>lac-pro</i> ) <i>ara thi</i>  | (Görke and Rak, 1999)                        |
| Z179         | As R1279, but $\Delta$ <i>glrR</i>   | (Reichenbach <i>et al.</i> , 2009)           |
| Z184         | As R1279, but $\Delta$ <i>rpoN</i>   | (Reichenbach <i>et al.</i> , 2009)           |
| Z190         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> , -10 mutated]   | (Reichenbach <i>et al.</i> , 2009)           |
| Z196         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> , -10 mutated]  | (Reichenbach <i>et al.</i> , 2009)           |
| Z197         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ]  | (Reichenbach <i>et al.</i> , 2009)           |
| Z206         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ]   | (Reichenbach <i>et al.</i> , 2009)           |
| Z227         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ]   | (Reichenbach <i>et al.</i> , 2009)           |
| Z360         | As R1279, but <i>attB</i> ::[ <i>E.c. glmZ</i> (-424 to +32)- <i>lacZ</i> ]  | pBGG59/BamHI $\rightarrow$ R1279, this work  |
| Z361         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>E.c. glmZ</i> (-424 to +32)- <i>lacZ</i> ]   | pBGG59/BamHI $\rightarrow$ Z179, this work   |
| Z362         | As R1279, but <i>attB</i> ::[ <i>Y.p. glmY</i> (-257 to +22)- <i>lacZ</i> ]  | pYG1/BamHI $\rightarrow$ R1279, this work    |
| Z363         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmY</i> (-257 to +22)- <i>lacZ</i> ]   | pYG1/BamHI $\rightarrow$ Z179, this work     |
| Z364         | As R1279, but <i>attB</i> ::[ <i>Y.p. glmZ</i> (-303 to +22)- <i>lacZ</i> ]  | pYG2/BamHI $\rightarrow$ R1279, this work    |
| Z365         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-303 to +22)- <i>lacZ</i> ]   | pYG2/BamHI $\rightarrow$ Z179, this work     |
| Z370         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ; IHF1 mutated]  | pBGG390/BamHI $\rightarrow$ R1279, this work |
| Z371         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ; IHF1 mutated; -10 mutated]   | pBGG391/BamHI $\rightarrow$ R1279, this work |
| Z372         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ; IHF1 mutated]   | pBGG390/BamHI $\rightarrow$ Z179, this work  |
| Z373         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ; IHF1 mutated; -10 mutated]  | pBGG391/BamHI $\rightarrow$ Z179, this work  |
| Z388         | As R1279, but <i>attB</i> ::[ <i>S.t. glmY</i> (-242 to +22)- <i>lacZ</i> ]  | pYG7/BamHI $\rightarrow$ R1279, this work    |
| Z389         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>S.t. glmY</i> (-242 to +22)- <i>lacZ</i> ]   | pYG7/BamHI $\rightarrow$ Z179, this work     |
| Z390         | As R1279, but <i>attB</i> ::[ <i>S.t. glmZ</i> (-242 to +22)- <i>lacZ</i> ]  | pYG8/BamHI $\rightarrow$ R1279, this work    |
| Z391         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>S.t. glmZ</i> (-242 to +22)- <i>lacZ</i> ]   | pYG8/BamHI $\rightarrow$ Z179, this work     |
| Z392         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> , -10 mutated], <i>ihfB</i> :: <i>kan</i>  | T4GT7 (JW0895) $\rightarrow$ Z190, this work |
| Z393         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>ihfB</i> :: <i>kan</i>   | T4GT7 (JW0895) $\rightarrow$ Z197, this work |
| Z394         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> , -10 mutated], <i>ihfA</i> :: <i>kan</i>  | T4GT7 (JW1702) $\rightarrow$ Z190, this work |
| Z395         | As R1279, but <i>attB</i> ::[ <i>E.c. glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>ihfA</i> :: <i>kan</i>   | T4GT7 (JW1702) $\rightarrow$ Z197, this work |
| Z397         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-292 to +22)- <i>lacZ</i> ; ABS1 mutated]   | pYG9/BamHI $\rightarrow$ Z179, this work     |
| Z398         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-303 to +22)- <i>lacZ</i> ; ABS2 mutated]   | pYG10/BamHI $\rightarrow$ Z179, this work    |
| Z399         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-303 to +22)- <i>lacZ</i> ; ABS3 mutated]   | pYG11/BamHI $\rightarrow$ Z179, this work    |
| Z400         | As R1279, but $\Delta$ <i>glrR</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-292 to +22)- <i>lacZ</i> ; ABS1,2,3 mutated]   | pYG12/BamHI $\rightarrow$ Z179, this work    |
| Z443         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>E.c. glmZ</i> (-424 to +32)- <i>lacZ</i> ]   | pBGG59/BamHI $\rightarrow$ Z184, this work   |
| Z444         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>Y.p. glmY</i> (-257 to +22)- <i>lacZ</i> ]   | pYG1/BamHI $\rightarrow$ Z184, this work     |
| Z445         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>Y.p. glmZ</i> (-303 to +22)- <i>lacZ</i> ]   | pYG2/BamHI $\rightarrow$ Z184, this work     |
| Z446         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>S.t. glmY</i> (-242 to +22)- <i>lacZ</i> ]   | pYG7/BamHI $\rightarrow$ Z184, this work     |
| Z447         | As R1279, but $\Delta$ <i>rpoN</i> , <i>attB</i> ::[ <i>S.t. glmZ</i> (-242 to +22)- <i>lacZ</i> ]   | pYG8/BamHI $\rightarrow$ Z184, this work     |

*E.c.*, *Escherichia coli*-K12; *Y.p.*, *Yersinia pseudotuberculosis* YPIII; *S.t.*, *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2

## RESULTS

### ***Conservation and gene synteny of glmY and glmZ in Enterobacteriaceae***

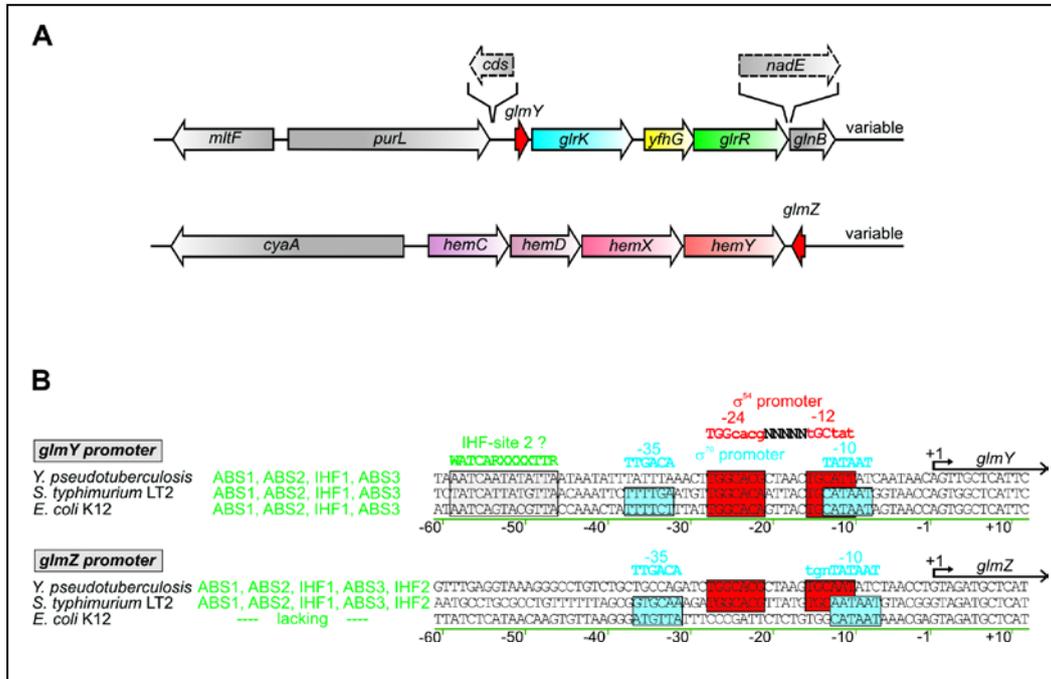
In *E. coli*, transcription of *glmY* is controlled by overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -dependent promoters. Activity of the  $\sigma^{54}$ -promoter is governed by the TCS GlrR/GlrK, which is encoded downstream of *glmY* (Reichenbach *et al.*, 2009). To investigate, whether this unusual promoter architecture is conserved in other bacteria and to increase our understanding of regulation of *glmZ* transcription, we compared the promoter sequences of *glmY* and *glmZ* from a comprehensive number of genomes. To retrieve these sequences, we used the sRNA sequences of *Escherichia coli* K12 (strain MG1655) as queries in NCBI Blast analyses. This search generated a list of species, all belonging to the *Enterobacteriaceae* family, which coincidentally contained both sRNA genes. Inspection of gene synteny using the MicrobesOnline tool (Dehal *et al.*, 2009) and the KEGG database (Kanehisa *et al.*, 2008) revealed conserved localization of *glmZ* downstream of the divergently orientated *hemCDXY* operon encoding enzymes involved in tetrapyrrole synthesis, whereas the region upstream of *glmZ* is variable and may carry insertion elements (Figs. 2.1 A and S2.1). Gene *glmY* is always located upstream of the gene cluster *glrK-yfhG-glrR* (Fig. 2.1 A and S2.2). Collectively, these observations suggest that sRNA genes *glmY* and *glmZ* are elements of the core genome conserved in *Enterobacteriaceae*. The conserved co-localization of *glmY* with the genes encoding the sensor kinase GlrK and the response regulator GlrR suggests that regulation of *glmY* expression by this TCS might be likewise conserved.

### ***Sequences for a $\sigma^{54}$ -promoter and for binding sites of the response regulator GlrR as well as of IHF are shared features of the glmY and glmZ promoter regions of many, but not all Enterobacteriaceae***

We performed sequence alignments of the promoter regions of the *glmY* as well as *glmZ* genes retrieved from 39 genome sequences representing the most important genera of *Enterobacteriaceae*. The  $\sigma^{54}$ -dependent promoter of *E. coli glmY* is conserved in all species (Fig. S2.3). The GlrR binding sites are likewise conserved although sequence deviations from the consensus TGTCN<sub>10</sub>GACA occur in a few cases, in particular in ABS 1 and 3. Two additional regions flanking ABS3 exhibit a higher degree of conservation and show similarity to binding sites of IHF, which are represented by the consensus WATCARXXXXTTR (Swinger and Rice, 2004). The previously characterized -10 sequence (CATAAT) of the  $\sigma^{70}$ -promoter, which overlaps with the -12 sequence of the  $\sigma^{54}$ -promoter of *glmY* in *E. coli*, is conserved only in a subset of genera, i.e. in *Escherichia* (which includes *Shigella* strains), *Klebsiella*, *Salmonella*, *Enterobacter*, *Citrobacter* and *Cronobacter*. Putative -35 sequences are also detectable at the appropriate positions. In contrast, in other genera, such as *Erwinia*, *Photobacterium*, *Serratia* and *Yersinia*, overlapping potential  $\sigma^{70}$ -promoter sequences are not detectable (Fig. S2.3).

The analysis of the promoter of the second sRNA gene *glmZ* revealed two groups of sequences, which exhibit no similarity and could not be aligned with each other (Fig. S2.4). In the group comprising the majority of sequences, the *glmZ* promoter region is strongly reminiscent of the organization of the *glmY* promoter. Sequence motifs of a  $\sigma^{54}$ -promoter, three GlrR binding sites and two IHF binding sites are detectable. The putative ABS1 and IHF-sites are less conserved in comparison to the *glmY* promoters (compare Figs. S2.3 and S2.4). In a subset of genera, i.e. *Cronobacter*, *Citrobacter*, *Enterobacter* and *Salmonella*, putative  $\sigma^{70}$ -promoters overlapping with the

$\sigma^{54}$ -promoters are also detectable upstream of *glmZ* (Fig. S2.4). Interestingly, these species also possess overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoter sequences upstream of *glmY* (Fig. S2.3). The second group comprised the genera *Klebsiella* and *Escherichia*. In these cases, sequence motifs for  $\sigma^{54}$ -promoters and for GlrR- and IHF-binding sites are lacking. Instead of that, putative  $\sigma^{70}$ -promoter sequences (ATGTTA-N<sub>15</sub>-tggCATAAT in *Escherichia sp.* and *Shigella* strains and ATGCAA-N<sub>15</sub>-tgcGATAAT in *Klebsiella pneumoniae*) are present at the appropriate positions.



**Figure 2.1: Organization of the *glmY* and *glmZ* genes in *Enterobacteriaceae*.** (A) Diagram illustrating gene synteny of the *glmY* and *glmZ* regions in *Enterobacteriaceae*. The gene cluster *glmY*-*glrK*-*yfhG*-*glrR*-*glnB* is conserved in *Enterobacteriaceae*, but in some species e.g. *Yersinia* and *Photobacterium*, gene *nadE* is inserted between *glrR* and *glnB*. Upstream of *glmY*, genes *mltF* and *purL* are present except for *Providencia sp.*. Small *orfs* of unknown function are interspersed between *purL* and *glmY* in *Yersinia*, *Photobacterium* and other species. Gene *glmZ* clusters with the downstream located and divergently orientated *hemCDXY* cluster, while the region upstream is variable. (B) Organization of enterobacterial *glmY* and *glmZ* promoters. Sequence alignments of the *glmY* and *glmZ* promoter regions from 39 enterobacterial genomes classified the species into three groups, for which *Y. pseudotuberculosis*, *S. typhimurium* and *E. coli* are representatively shown (for details, see Figs. S3 and S4). *Yersinia* possesses the sequences for a  $\sigma^{54}$ -promoter (labeled in red) and three GlrR binding sites upstream of both sRNA genes, while overlapping  $\sigma^{70}$ -promoters appear to be absent. GlrR binding sites and  $\sigma^{54}$ -promoters are also detectable upstream of both sRNA genes in *Salmonella*, but in addition putative  $\sigma^{70}$ -promoters (labeled in blue) that overlap the  $\sigma^{54}$ -promoters, are detectable. This arrangement is also found upstream of *E. coli glmY*. However, *E. coli glmZ* appears to be transcribed from a single  $\sigma^{70}$ -promoter. The sequence alignment also detected two putative IHF binding sites that coincide with the occurrence of  $\sigma^{54}$ -promoters.

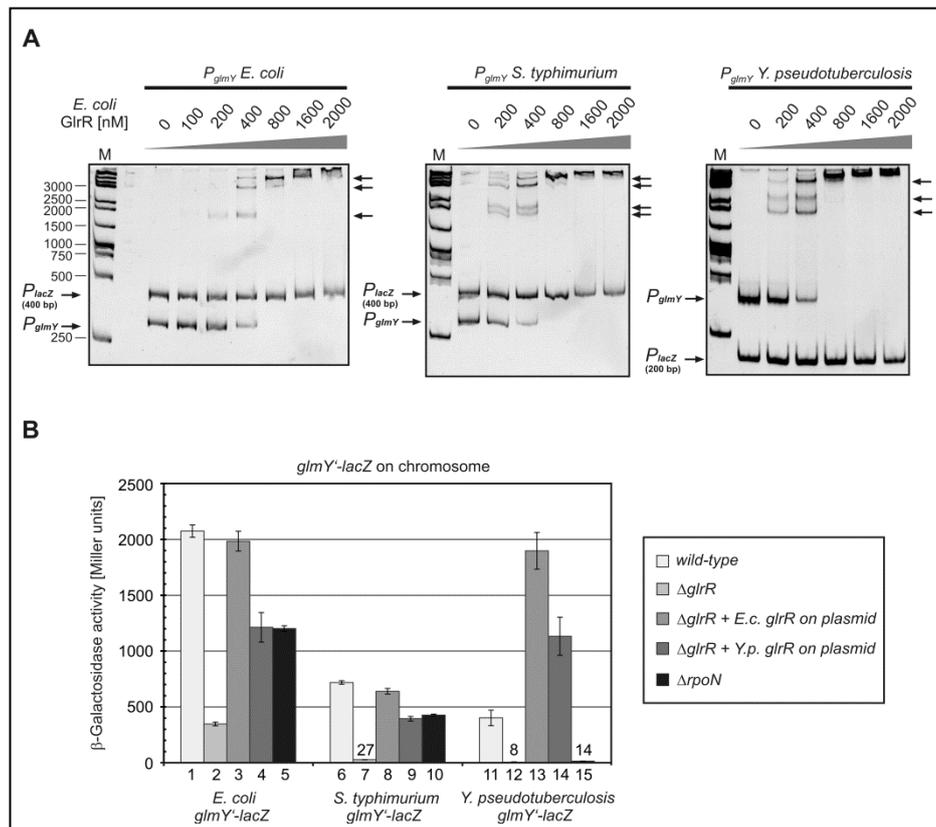
From these analyses we hypothesized that enterobacterial species can be classified into three groups in respect to control of *glmY* and *glmZ* expression (Fig. 2.1 B): (I) Species of the genera *Pantoea*, *Erwinia*, *Pectobacterium*, *Arsenophonus*, *Photobacterium*, *Serratia*, *Proteus*, *Yersinia* and *Dickeya* may transcribe both, *glmY* and *glmZ*, from  $\sigma^{54}$ -dependent promoters, which might be controlled by GlrR/GlrK. (II) This may also apply to species of the genera *Cronobacter*, *Citrobacter*, *Enterobacter* and *Salmonella*, but as a difference, additional overlapping  $\sigma^{70}$ -promoters are present, which may start transcription at the same site. (III) Overlapping  $\sigma^{54}$ - and  $\sigma^{70}$ -promoters also control expression of *glmY* in *Klebsiella* and *Escherichia* species. In contrast, transcription of *glmZ* is driven exclusively from  $\sigma^{70}$ -promoters.

Finally, IHF might be important for the activities of the  $\sigma^{54}$ -dependent *glmY* and *glmZ* promoters. To address these hypotheses, we selected one species per group to experimentally analyse the *glmY*

and *glmZ* promoters (Fig. 2.1 B). These were *Y. pseudotuberculosis* YPIII (group I), *S. enterica* subsp. *enterica* serovar *typhimurium* str. LT2 (group II) and *E. coli* K12 (group III).

**Response regulator GlrR binds to the *glmY* promoter regions of *S. typhimurium* and *Y. pseudotuberculosis***

First, we wanted to verify if the putative  $\sigma^{54}$ -dependent *glmY* promoters of *S. typhimurium* and *Y. pseudotuberculosis* are controlled by the response regulator GlrR. Therefore, we tested whether purified GlrR protein is able to bind to these promoters. Electrophoretic mobility shift assays (EMSA) were carried out using purified GlrR protein from *E. coli* and DNA fragments covering the *glmY* promoter regions of these species. For comparison, binding of GlrR to the corresponding DNA fragment of *E. coli* was tested. Different concentrations of purified His-tagged GlrR protein were incubated with the various *glmY* promoter fragments, respectively. In order to verify binding specificity, an additional DNA fragment, which covered the *lacZ* promoter and had a size of either 400 bp or 200 bp was simultaneously present in these assays. Protein/DNA-complexes and unbound DNA were separated by polyacrylamide gel electrophoresis (Fig. 2.2 A).



**Figure 2.2: Comparison of the roles of GlrR and  $\sigma^{54}$  for expression of *glmY* from *E. coli*, *S. typhimurium* and *Y. pseudotuberculosis*.** (A) EMSAs to test binding of *E. coli* GlrR protein to the *glmY* promoter regions of *E. coli* (-238 to +22), *S. typhimurium* (-242 to +22) and *Y. pseudotuberculosis* (-257 to +22). In addition to the *glmY* promoter fragments, 400 bp (panels 1 and 2) or 200 bp DNA fragments (panel 3) covering the *lacZ* promoter were present as internal controls. The sizes of the DNA size standard are given at the left. The apparent  $K_D$  values are 360 nM for the *E. coli* *glmY* promoter, 230 nM for the *Salmonella* *glmY* promoter and 290 nM for the *Y. pseudotuberculosis* *glmY* promoter. (B)  $\beta$ -Galactosidase activities of *E. coli* strains carrying fusions of *glmY* from *E. coli*, *S. typhimurium* and *Y. pseudotuberculosis* to the *lacZ* reporter gene. In addition, these strains had the genotypes indicated in the legend. The following strains and transformants were tested (corresponding to the columns from left to right): Z197, Z206+pBGG223, Z206+pYG6, Z227, Z388, Z389, Z389+pBGG223, Z389+pYG6, Z446, Z362, Z363, Z363+pBGG223, Z363+pYG6, Z444.

The *glmY* promoter fragments of all three species were shifted to distinct slower migrating bands indicating DNA/GlrR complexes, while the *lacZ* control fragments were not bound. Comparable protein concentrations were required to achieve binding, indicating that GlrR binds with similar affinities to all these *glmY* fragments. GlrR of *E. coli* shares 95% and 87% amino acid sequence identity with its homologs from *S. typhimurium* and *Y. pseudotuberculosis*, respectively. To confirm that the results obtained with the heterologous GlrR protein are valid, we additionally performed EMSAs using purified *Y. pseudotuberculosis* GlrR. This protein also bound the *glmY* promoter DNA fragments of both, *Y. pseudotuberculosis* and *E. coli*, with comparable affinities (Fig. S2.5). However, in comparison to GlrR from *E. coli* higher protein concentrations were required to achieve binding.

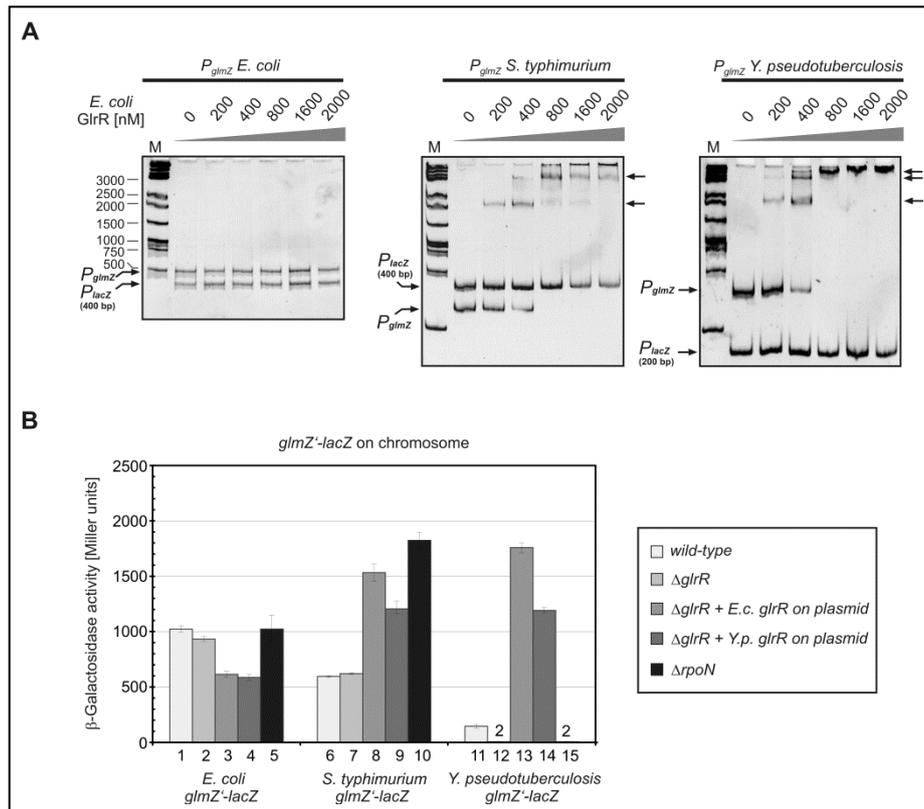
### **Analysis of *S. typhimurium* and *Y. pseudotuberculosis* *glmY* expression**

The EMSAs suggested that *glmY* expression is regulated by GlrR in all three species. To validate this conclusion and to determine whether single or overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters control expression of *glmY*, we constructed fusions of the *glmY* genes of all three species to the *lacZ* reporter gene. The fusions were integrated into the chromosome of *E. coli* wild-type and isogenic  $\Delta glrR$  and  $\Delta rpoN$  mutants (*rpoN* encodes  $\sigma^{54}$ ). The resulting strains were grown to exponential phase and the  $\beta$ -galactosidase activities were determined. The *E. coli glmY'*-*lacZ* fusion was readily expressed in the wild-type, while its expression was 6-fold lower in the  $\Delta glrR$  mutant reflecting the lack of activity of the  $\sigma^{54}$ -promoter (Fig. 2.2 B, columns 1 and 2). However, a certain level of expression was retained in the  $\Delta glrR$  mutant, which is due to the activity of the overlapping  $\sigma^{70}$ -promoter (Reichenbach *et al.*, 2009). Complementation of the  $\Delta glrR$  mutant with a plasmid carrying *E. coli glrR* under  $P_{Ara}$  promoter control restored expression of *glmY'*-*lacZ* to wild-type levels (Fig. 2.2 B, columns 1 and 3), while a somewhat lower activity was obtained when a plasmid carrying *glrR* from *Y. pseudotuberculosis* was used (Fig. 2.2 B, column 4). This effect was also seen in all subsequent complementation experiments suggesting that GlrR from *Y. pseudotuberculosis* is less active than the *E. coli* GlrR protein. In agreement with previous data (Reichenbach *et al.*, 2009), the *E. coli glmY'*-*lacZ* fusion was expressed at higher levels in the  $\Delta rpoN$  mutant in comparison to the  $\Delta glrR$  mutant (Fig. 2.2 B, columns 2 and 5). This difference results from repression of the  $\sigma^{70}$ -dependent promoter by binding of  $\sigma^{54}$ -RNAP to the overlapping  $\sigma^{54}$ -promoter in the  $\Delta glrR$  mutant (Reichenbach *et al.*, 2009).

Similar results were obtained using the *S. typhimurium glmY'*-*lacZ* fusion (Fig. 2.2 B, columns 6-10). However, expression of this fusion was almost completely abolished in the  $\Delta glrR$  mutant (Fig. 2.2 B, columns 2 and 7). A considerable level of expression was detectable in the  $\Delta rpoN$  mutant as it was also observed for the *E. coli glmY'*-*lacZ* fusion (Fig. 2.2 B, columns 5 and 10). Hence, the data are compatible with overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters, as predicted by the sequence alignment (Fig. 2.1 B; Fig. S2.3). The  $\sigma^{70}$ -promoter of *S. typhimurium glmY* appears to be completely repressed by binding of  $\sigma^{54}$  to the overlapping  $\sigma^{54}$ -promoter. The *Y. pseudotuberculosis glmY'*-*lacZ* fusion exhibited a different pattern of expression (Fig. 2.2 B, columns 11-15). This fusion was neither expressed in the  $\Delta glrR$  nor in the  $\Delta rpoN$  mutant. Complementation of the  $\Delta glrR$  mutant with plasmids encoding *glrR* either from *E. coli* or *Y. pseudotuberculosis* restored expression to higher levels than in the wild-type strain (Fig. 2.2 B, columns 11, 13, 14). Collectively, the data support the conclusions drawn from the sequence alignments: The *glmY* genes of all three species are transcribed from  $\sigma^{54}$ -dependent promoters that require activation by GlrR. An additional  $\sigma^{70}$ -promoter overlapping the  $\sigma^{54}$ -promoter exists in *E. coli* and *S. typhimurium*, but not in *Y. pseudotuberculosis*.

**The response regulator GlrR binds the *glmZ* promoter region of *S. typhimurium* and *Y. pseudotuberculosis*, while the *E. coli* *glmZ* promoter is not bound**

The sequence alignment analysis of the *glmZ* promoter regions had revealed putative  $\sigma^{54}$ -promoters and GlrR binding sites in *S. typhimurium* and *Y. pseudotuberculosis*, while these elements are missing upstream of *E. coli* *glmZ* (Fig. 2.1 B; Fig. S2.4). To determine whether GlrR is able to bind to these promoter regions, EMSAs were carried out using *E. coli* GlrR protein and DNA fragments encompassing the respective *glmZ* promoter regions. These experiments showed that GlrR binds the *glmZ* promoters of *S. typhimurium* and *Y. pseudotuberculosis* with comparable affinities, whereas the *E. coli* *glmZ* promoter is not bound (Fig. 2.3 A).



**Figure 2.3: Comparison of the roles of GlrR and  $\sigma^{54}$  for expression of *glmZ* from *E. coli*, *S. typhimurium* and *Y. pseudotuberculosis*. (A) EMSAs to test binding of *E. coli* GlrR protein to the *glmZ* promoter regions of *E. coli* (-424 to +32), *S. typhimurium* (-242 to +22) and *Y. pseudotuberculosis* (-303 to +22). The apparent  $K_D$  values for binding of GlrR to the *S. typhimurium* and *Y. pseudotuberculosis* *glmY* promoter fragments are 370 nM in both cases. (B)  $\beta$ -Galactosidase activities of *E. coli* strains carrying fusions of *glmZ*' from *E. coli*, *S. typhimurium* and *Y. pseudotuberculosis* to the *lacZ* reporter gene. In addition, these strains had the genotypes indicated in the legend. The following strains and transformants were tested (corresponding to the columns from left to right): Z360, Z361, Z361+pBGG223, Z361+pYG6, Z443, Z390, Z391, Z391+pBGG223, Z391+pYG6, Z447, Z364, Z365, Z365+pBGG223, Z365+pYG6, Z445.**

In addition, EMSAs were carried out using GlrR from *Y. pseudotuberculosis* (Fig. S2.6). Binding of the *glmZ* promoter fragment from *Y. pseudotuberculosis* was detectable, but four times higher protein concentrations were required in comparison to GlrR from *E. coli*, as already observed in the EMSAs using the *glmY* promoter fragments (Fig. S2.5). In contrast, the *E. coli* *glmZ* promoter fragment was not bound (Fig. S2.6). In conclusion, GlrR binds the *glmZ* promoters of *Y. pseudotuberculosis* and *S. typhimurium*, but not of *E. coli*.

**Analysis of *E. coli*, *S. typhimurium* and *Y. pseudotuberculosis* *glmZ* expression**

To obtain further evidence that  $\sigma^{54}$  and GlrR regulate the *glmZ* genes of *S. typhimurium* and *Y. pseudotuberculosis* and are not involved in *E. coli* *glmZ* regulation, *lacZ* fusions of the *glmZ* genes were constructed and integrated into the chromosome of *E. coli* wild-type,  $\Delta glrR$  and  $\Delta rpoN$  strains. Expression of the *E. coli* *glmZ'*-*lacZ* fusion was neither affected by the  $\Delta glrR$  nor by the  $\Delta rpoN$  mutation and expression of *glrR* from a plasmid had also no stimulatory effect (Fig. 2.3 B, columns 1-5). Hence, expression of *E. coli* *glmZ* is not controlled by GlrR or  $\sigma^{54}$ . Expression of the *S. typhimurium* *glmZ'*-*lacZ* fusion was also not decreased in the  $\Delta glrR$  mutant. In contrast to the *E. coli* *glmZ'*-*lacZ* fusion, expression was significantly increased when *glrR* was expressed from a plasmid (Fig. 2.3 B, compare columns 6-9 and 1-4). Interestingly, expression of this fusion was also strongly increased in the  $\Delta rpoN$  mutant (Fig. 2.3 B, columns 6 and 10). These results can be explained by the existence of overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters. The high levels of *glmZ* transcription detected in the  $\Delta glrR$  and  $\Delta rpoN$  mutants (Fig. 2.3 B, columns 7 and 10) suggest that this  $\sigma^{70}$ -promoter is stronger than the  $\sigma^{70}$ -promoter preceding the *glmY* gene in *S. typhimurium*.

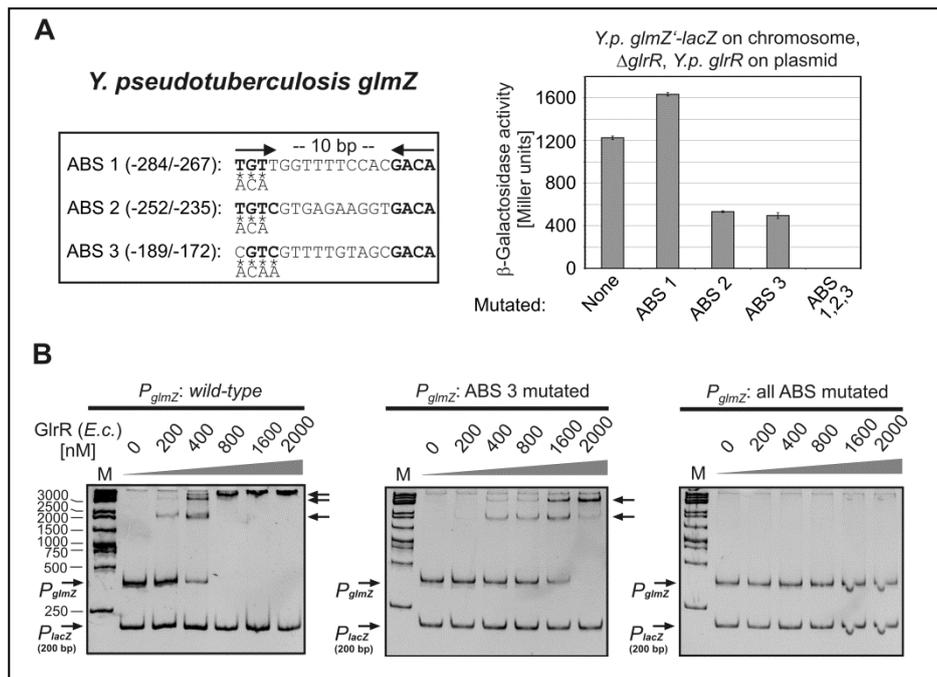
The *Y. pseudotuberculosis* *glmZ'*-*lacZ* fusion showed an expression pattern that was reminiscent of the results obtained with the cognate *glmY'*-*lacZ* fusion. Expression of both fusions was abolished in  $\Delta glrR$  as well as  $\Delta rpoN$  mutants (columns 12 and 15 in Figs. 2.2 B and 2.3 B, respectively). Complementation of the  $\Delta glrR$  mutant with plasmids carrying *glrR* either from *E. coli* or *Y. pseudotuberculosis* restored expression to levels that were even higher than in the wild-type strain (columns 11, 13 and 14 in Fig. 2.2 B and 2.3 B). In conclusion, *Y. pseudotuberculosis* *glmY* as well as *glmZ* appear to be expressed exclusively from  $\sigma^{54}$ -dependent promoters that require activation by GlrR. Apparently, overlapping  $\sigma^{70}$ -promoters do not exist in these cases.

**Expression of *glmY* and *glmZ* in *Y. pseudotuberculosis***

Among *Enterobacteriaceae*, *E. coli* and *Y. pseudotuberculosis* are distantly related (Paradis *et al.*, 2005). Although the transcriptional machinery and all elements involved in regulation of *glmY* and *glmZ* expression are conserved in both species, one might argue that the patterns of *Y. pseudotuberculosis* *glmY* and *glmZ* expression, as observed here in *E. coli*, do not appropriately reflect expression of these sRNAs in the authentic host. To address this possibility, we transformed *Y. pseudotuberculosis* with plasmids carrying either the *Y. pseudotuberculosis* *glmY'*-*lacZ* or the *glmZ'*-*lacZ* fusion or with the empty fusion vector. The cells carrying the *glmY'*-*lacZ* or the *glmZ'*-*lacZ* fusion displayed significantly higher  $\beta$ -galactosidase activities than the transformant carrying the empty *lacZ* fusion plasmid (Fig. S2.7 A). Thus, both fusions are expressed in *Y. pseudotuberculosis*. The *glmY'*-*lacZ* fusion was approximately two-fold higher expressed than the *glmZ'*-*lacZ* fusion. The same difference was observed in *E. coli* (compare columns 11 in Figs. 2.2 B and 2.3 B). Next, a second compatible plasmid carrying either *glrR* from *E. coli* or *Y. pseudotuberculosis* or no gene (empty vector) under control of the  $P_{Ara}$  promoter was introduced. Presence of the *glrR* expression plasmids strongly increased expression of the *lacZ* fusions (Fig. S2.7 B). Expression of *E. coli* *glrR* resulted in higher expression levels of the *lacZ* fusions in comparison to *Y. pseudotuberculosis* *glrR*. These differences were also detected in *E. coli* (Figs. 2.2 B and 2.3 B). Taken together, it appears justified to conclude that the data obtained with these *lacZ* fusions in *E. coli* reflect their expression in *Y. pseudotuberculosis*.

***E. coli glmZ* is exclusively transcribed from a  $\sigma^{70}$ -promoter, while *Y. pseudotuberculosis glmZ* transcription depends on  $\sigma^{54}$  and GlrR**

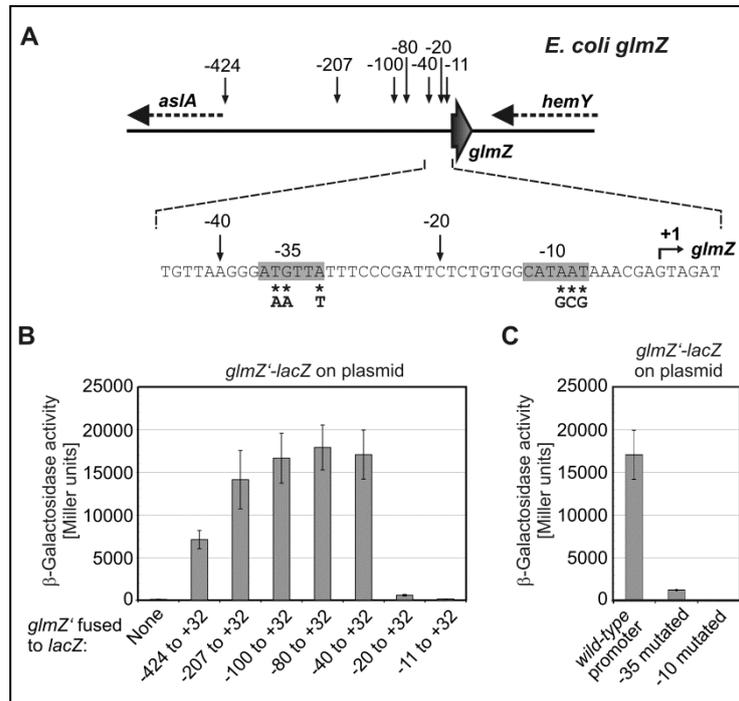
Our data suggested that *glmZ* of *Y. pseudotuberculosis* is transcribed from a single promoter that requires activation by  $\sigma^{54}$  and GlrR, whereas expression of *E. coli glmZ* is not affected by these factors. To confirm this conclusion, we mutated the left half-site of each of the three putative ABS of GlrR individually or in combination (Fig. 2.4 A, left). Fusions of *Y. pseudotuberculosis glmZ'* to *lacZ* carrying these mutations were integrated into the chromosome of the *E. coli  $\Delta$ glrR* mutant. These strains were subsequently complemented with the plasmid carrying *Y. pseudotuberculosis glrR* under *P<sub>Ara</sub>* promoter control and the  $\beta$ -galactosidase activities were determined (Fig. 2.4 A, right).



**Figure 2.4: Transcription of *Y. pseudotuberculosis glmZ* depends on binding of GlrR to its three target sites upstream of the promoter.** (A)  $\beta$ -Galactosidase activities of *E. coli* strains carrying mutated GlrR binding sites in the chromosomal *Y. pseudotuberculosis glmZ'-lacZ* fusion. In order to monitor activation by the cognate GlrR protein, *Y. pseudotuberculosis glrR* was expressed from plasmid pYG6, while the endogenous *glrR* gene was deleted. The nucleotide exchanges introduced into the ABS are depicted at the left. The following strains were employed (corresponding to the columns from left to right): Z365, Z397, Z398, Z399, Z400. (B) EMSAs to monitor binding of *E. coli* GlrR to DNA fragments covering the *Y. pseudotuberculosis glmZ* promoter and carrying mutations in the ABS as depicted in the Figure.

Mutation of ABS 1 had no negative impact on *Y. pseudotuberculosis glmZ'-lacZ* transcription, whereas mutation of ABS 2 or ABS 3 reduced expression more than two-fold. Expression was completely abolished, when all three ABS were simultaneously mutated. To corroborate these data, we performed EMSA experiments using *Y. pseudotuberculosis glmZ* promoter fragments carrying a mutation in ABS3 or simultaneously in all three ABS. These EMSAs were carried out using purified GlrR from *E. coli* (Fig. 2.4 B) or from *Y. pseudotuberculosis* (Fig. S2.8). In addition, a truncated *glmZ* promoter fragment lacking all three ABSs was tested in EMSA with *Y. pseudotuberculosis* GlrR (Fig. S2.8). The data show that mutation of ABS3 decreased binding efficiency significantly. Finally, binding of GlrR was completely prevented, when all three ABS were truncated or simultaneously mutated (Fig. 2.4 B, Fig. S2.8). These results show that *Y. pseudotuberculosis glmZ* is transcribed from a single  $\sigma^{54}$ -promoter, which requires activation by binding of GlrR to its upstream located ABS.

To further confirm that *E. coli glmZ* expression is independent of upstream activating sequences, a promoter deletion analysis was performed. For this purpose, DNA fragments carrying gradually 5'-truncated versions of the *aslA-glmZ* intergenic region were fused to *lacZ* (Fig. 2.5 A).



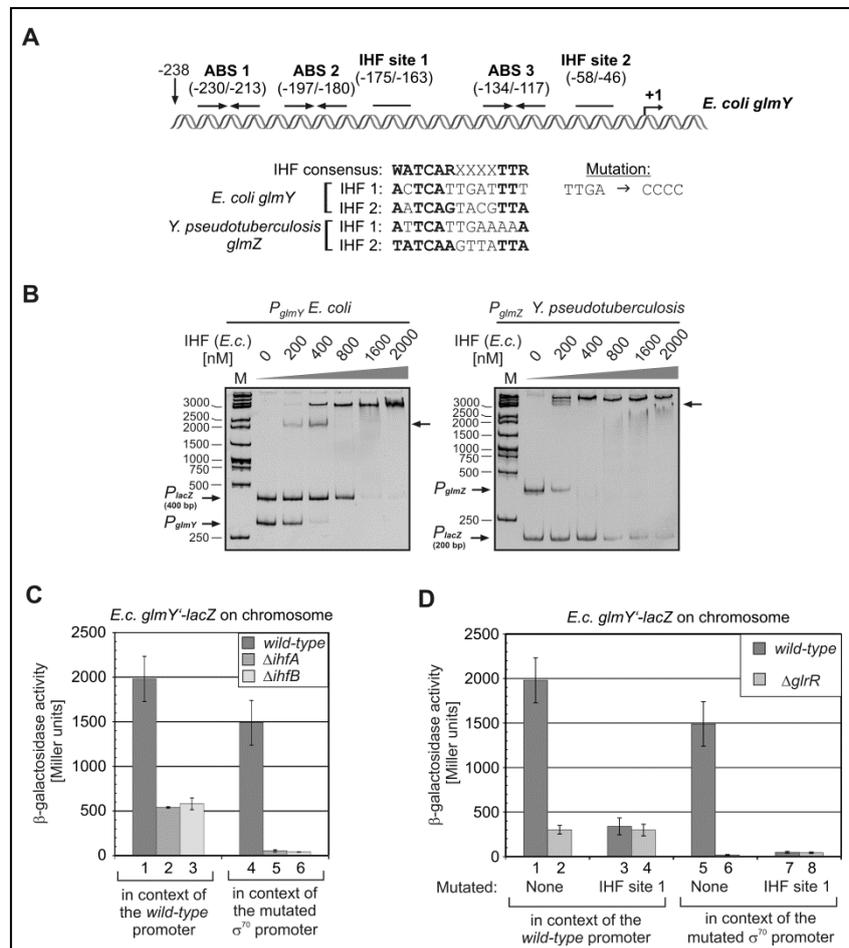
**Figure 2.5: Analysis of the *E. coli glmZ* promoter (A)** Schematic representation of the *aslA-hemY* intergenic region comprising the *E. coli glmZ* gene. DNA fragments extending until position +32 relative to the *glmZ* start site and with the 5' ends indicated by arrows were fused to *lacZ*. The sequence of the *glmZ* promoter region with the putative -35/-10 motifs of a  $\sigma^{70}$ -promoter is shown below. The nucleotide exchanges that were introduced into these motifs and tested in (C) are marked with asterisks. **(B)** 5'→3' deletion analysis of the *E. coli glmZ* upstream region.  $\beta$ -Galactosidase activities of *E. coli* wild-type strain R1279 carrying the gradually 5' truncated *glmZ*-*lacZ* fusions on plasmids. The following plasmids were tested (corresponding to the columns from left to right): pKEM04, pBGG59, pBGG111, pBGG112, pBGG113, pBGG114, pBGG170, pBGG135. **(C)** Mutational analysis of the *glmZ* promoter. The putative -35 and -10 sequences were mutated as indicated in (A) in the context of the *glmZ*'(-40 to +32)-*lacZ* fusion. Plasmids pBGG114, pBGG157 and pBGG171 (corresponding to the columns from left to right) were introduced into wild-type strain R1279 and the  $\beta$ -galactosidase activities were determined.

Plasmids carrying these various fusions were subsequently introduced into *E. coli* wild-type and the  $\beta$ -galactosidase activities were determined. The data show that the region upstream of position -40 relative to *glmZ* is dispensable for promoter activity (Fig. 2.5 B). Deletion of the sequences upstream position -20, which removes the -35 motif of the putative  $\sigma^{70}$ -promoter (Fig. 2.5 A and Fig. S2.4), abrogates expression. To verify if the assumed -35 and -10 sequences are indeed elements of a functional  $\sigma^{70}$ -promoter, these sequence elements were mutated. Mutation of the three bases matching the consensus sequence TTGACA within the putative -35 sequence (Fig. 2.5 A) reduced expression of the fusion drastically (Fig. 2.5 C). Mutation of the right half site of the putative -10 motif completely abolished expression (Fig. 2.5 C). These data confirm that *E. coli glmZ* is transcribed from a single  $\sigma^{70}$ -promoter, which is constitutively active and apparently unregulated, at least under the tested conditions.

#### Activity of the $\sigma^{54}$ -dependent *glmY* promoter requires binding of IHF

The sequence alignment analyses detected two additional sequence motifs with similarity to the binding site of the global transcriptional regulator IHF. These sequence elements were detectable in all species, except for the *glmZ* promoters of *Escherichia*, *Shigella* and *Klebsiella* (Figs. S2.3 and S2.4),

which according to all evidence are transcribed from single  $\sigma^{70}$ -promoters. This suggested a role of these sites for activities of the  $\sigma^{54}$ -promoters upstream of *glmY* and *glmZ* (Fig. 2.6 A). Therefore, we tested whether IHF is able to bind to the promoter fragments of *E. coli glmY* and *Y. pseudotuberculosis glmZ*.



**Figure 2.6: Role of IHF for expression of *glmY*.** (A) Schematic representation of the *E. coli glmY* promoter region and location of GlrR and putative IHF binding sites. The sequences of the putative IHF binding sites upstream of *E. coli glmY* and *Y. pseudotuberculosis glmZ* are shown and the nucleotide exchanges introduced in IHF site 1 of the *E. coli glmY* promoter are indicated. (B) EMSAs to test binding of purified IHF to the *glmY* and *glmZ* promoter regions of *E. coli* and *Y. pseudotuberculosis*, respectively. The DNA fragments were obtained by PCR making use of the primer pairs BG377/BG456 and BG700/BG701, respectively. As controls, DNA fragments encompassing the *lac* promoter were additionally present. (C) Expression of *E. coli glmY* in  $\Delta*ihfA*$  and  $\Delta*ihfB*$  mutants.  $\beta$ -Galactosidase activities of strains carrying the chromosomal *E. coli glmY-lacZ* fusion in the context of the wild-type promoter (columns 1-3) or in the context of the mutated  $\sigma^{70}$ -promoter leaving the  $\sigma^{54}$ -promoter as single active promoter (columns 4-6). Genes *ihfA* or *ihfB* were deleted as indicated in the legend. The following strains were tested (corresponding to the columns from left to right): Z197, Z395, Z393, Z190, Z394, Z392. (D) Mutational analysis of the putative IHF site 1 in the *E. coli glmY* promoter region.  $\beta$ -Galactosidase activities of wild-type and  $\Delta*glrR*$  *E. coli* strains carrying the wild-type or mutated alleles of the *E. coli glmY-lacZ* fusion. Mutations were either in the putative IHF-site 1 (columns 3, 4, 7, 8) as indicated in (A) or in the -10 sequence of the *glmY* promoter (columns 5-8) rendering *glmY-lacZ* expression fully dependent on  $\sigma^{54}$ . The following strains were employed (corresponding to the columns from left to right): Z197, Z206, Z370, Z372, Z190, Z196, Z371, Z373.

Both DNA fragments were bound by IHF protein (Fig. 2.6 B). The *lacZ* promoter fragments, which served as internal controls, were also bound, but at higher protein concentrations. The *lacZ* promoter is not known to contain any IHF site indicating unspecific binding. To confirm this conclusion, we repeated the experiments using a DNA fragment covering the *ptsG* promoter from *Bacillus subtilis* as internal control. *B. subtilis* does not possess IHF. Once more, efficient binding of the *glmY* and *glmZ* promoters could be observed, while the *ptsG* promoter was only bound at higher protein concentrations (Fig. S2.9). Hence, binding of IHF to the *lacZ* and *ptsG* promoters is unspecific, which

is in line with previous data reporting that IHF binds DNA with lower affinity also in sequence-independent manner (Swinger and Rice, 2004).

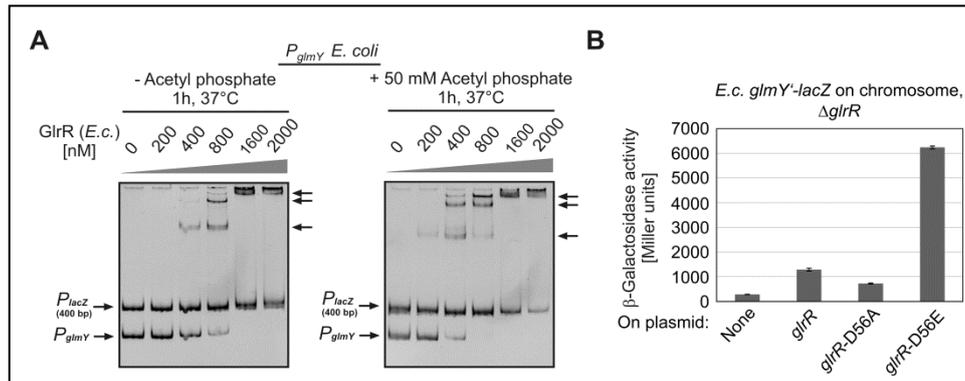
Next, we determined whether IHF is important for the activities of the  $\sigma^{54}$ -promoters. Therefore, we examined the role of IHF for expression of *E. coli glmY*. Expression of the chromosomally encoded *E. coli glmY'-lacZ* fusion was determined in mutants lacking *ihfA* or *ihfB*, which encode the subunits of IHF (Swinger and Rice, 2004). In both mutants, expression of the fusion was reduced ~four-fold (Fig. 2.6 C, columns 1-3). The remaining activities were comparable with the expression level of this fusion in the  $\Delta glrR$  mutant (Fig. 2.2 B, column 2), suggesting that it is caused by activity of the overlapping  $\sigma^{70}$ -promoter (Reichenbach *et al.*, 2009). Therefore, we repeated the experiment using a *glmY'-lacZ* fusion in which the -10 sequence of the  $\sigma^{70}$ -promoter is mutated, while the  $\sigma^{54}$ -promoter is unaffected (Reichenbach *et al.*, 2009). Expression of this fusion was abolished in the  $\Delta ihfA$  and  $\Delta ihfB$  mutants (Fig. 2.6 C, columns 4-6). This demonstrates that IHF is essential for activity of the  $\sigma^{54}$ -promoter of *glmY*.

To assess whether the two sequence elements resembling IHF binding sites are important for  $\sigma^{54}$ -promoter activity, we mutated the putative IHF-site 1 in the *E. coli glmY'-lacZ* fusion. Four highly conserved nucleotides (Fig. S2.3) were exchanged within the putative IHF site 1 (Fig. 2.6 A). This mutation yielded the same effects as the  $\Delta ihfA$  and  $\Delta ihfB$  mutations. Expression of the *glmY'-lacZ* fusion dropped five-fold and the remaining expression was comparable to the expression obtained in the  $\Delta glrR$  mutant, in which solely the  $\sigma^{70}$ -promoter is active (Fig. 2.6 D, columns 1-3). Mutation of the putative IHF-1 site had no further negative impact on the residual expression of the fusion in the  $\Delta glrR$  mutant (Fig. 2.6 D, columns 2 and 4) suggesting that activity of the  $\sigma^{70}$ -promoter is unaffected by this mutation. To verify the role of site 1 for activity of the  $\sigma^{54}$ -promoter, the experiments were repeated using the *glmY'-lacZ* fusion in which the  $\sigma^{70}$ -promoter had been mutated. Mutation of IHF site 1 abolished expression of this fusion and therefore had the same effect as a  $\Delta glrR$  or the  $\Delta ihf$  mutations (Fig. 2.6 D, columns 5-8; Fig. 2.6 C, columns 4-6). Hence, site 1 is essential for activity of the  $\sigma^{54}$ -promoter. Collectively, these data show that activity of the  $\sigma^{54}$ -promoter of *E. coli glmY* requires binding of IHF to the promoter region. The two sites identified by sequence alignment are likely candidates for these IHF binding sites. In contrast, activity of the overlapping  $\sigma^{70}$ -promoter appears to be unaffected by IHF.

### **Phosphorylated GlrR is active and stimulates sRNA expression**

GlrR contains a response regulatory domain including the conserved putative phosphorylation site aspartate 56 at its N-terminus. Phosphorylation of GlrR by its cognate kinase GlrK has been previously demonstrated *in vitro* (Yamamoto *et al.*, 2005). Furthermore, a  $\Delta glrK$  mutation was shown to abolish activity of the  $\sigma^{54}$ -promoter of *glmY* in *E. coli*, suggesting that GlrK controls activity of this promoter through modulation of the phosphorylation state of GlrR (Reichenbach *et al.*, 2009). In many TCS, the histidine kinase is capable of phosphorylating as well as dephosphorylating the response regulator. Phosphorylation of the response regulator results in structural changes, which in most cases activate the protein and stimulate interaction with the target DNA (Stock *et al.*, 2000). In a few cases the dephosphorylated protein was shown to be active (Zakikhany *et al.*, 2010). We wanted to discriminate, whether phosphorylated or dephosphorylated GlrR is active. Therefore, we exploited the fact that many response regulators can autophosphorylate *in vitro* using small

molecules such as acetyl phosphate as phosphoryl group donors (Scharf, 2010). Therefore, EMSAs were carried out using the *E. coli glmY* promoter fragment and the *E. coli* GlrR protein that was pre-incubated with 50 mM acetyl phosphate for 1 h at 37°C prior to EMSA. Since ongoing incubation of GlrR at 37°C resulted in increasing inactivation of the protein (compare left panels in Figs. 2.2 A and 2.7 A), a control experiment was performed in which GlrR was treated the same way but acetyl phosphate was omitted. These experiments revealed that binding affinity of GlrR was somewhat increased by the acetyl phosphate treatment relative to the control (compare panels in Fig. 2.7 A).



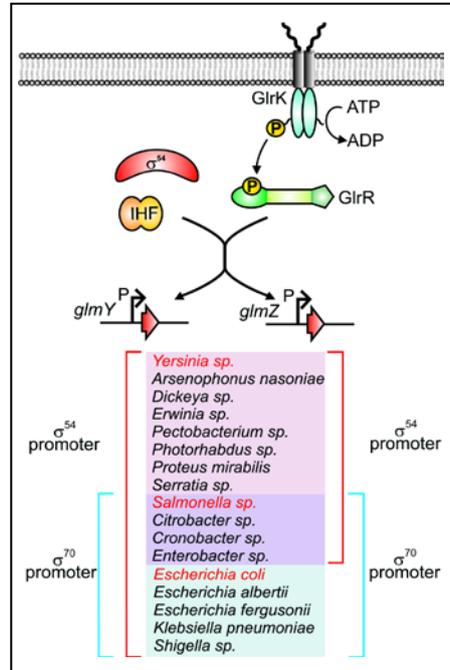
**Figure 2.7: Phosphorylation increases activity of response regulator GlrR.** (A) Effect of acetyl phosphate on the DNA binding activity of GlrR as revealed by EMSA. EMSAs were performed using purified *E. coli* GlrR and the *E. coli glmY* promoter fragment. To test the possible effect of phosphorylation on GlrR activity, the protein was pre-incubated at 37°C for 1 h in the absence (left panel) or presence (right panel) of 50 mM acetyl phosphate before continuing with the EMSA protocol. (B) A glutamate replacement of the phosphorylation site Asp56 in GlrR strongly up-regulates *glmY* expression. *E. coli* strain Z206 carrying a  $\Delta glrR$  mutation and the *E. coli glmY-lacZ* fusion on the chromosome was complemented with plasmids carrying *E. coli* wild-type *glrR* (pBGG389, column 2), *glrR-D56A* (pBGG398, column 3), *glrR-D56E* (pBGG399, column 4) or no gene (pBAD33, column 1) under  $P_{Ara}$  promoter control. Subsequently, the  $\beta$ -galactosidase activities were determined from these transformants.

To obtain *in vivo* evidence that phosphorylated rather than dephosphorylated GlrR is active, we replaced the phosphorylation site Asp 56 in GlrR with an alanine and a glutamate residue, respectively. An Ala replacement is reported to mimic the dephosphorylated form of a response regulator, while a Glu replacement is able to mimic the phosphorylated Asp in some response regulators resulting in kinase-independent activation (Scharf, 2010). Plasmids carrying the various *glrR* variants or no gene (empty vector control) under  $P_{Ara}$  promoter control were used to complement the  $\Delta glrR$  mutant that carries the *E. coli glmY-lacZ* fusion on the chromosome. Subsequently the  $\beta$ -galactosidase activities were determined from these transformants. Expression of the *glrR-D56A* allele resulted in ~two-fold lower activity when compared with wild-type *glrR* (Fig. 2.7 B, columns 2 and 3). In contrast, expression of *glrR-D56E* enhanced *glmY-lacZ* expression five-fold. Taken together, the data indicate that phosphorylation of GlrR increases its DNA binding activity and thereby expression of the sRNA.

## DISCUSSION

In this study we addressed the transcriptional regulation of two sRNA genes, *glmY* and *glmZ*, which are conserved in *Enterobacteriaceae*. Our analysis reveals three different scenarios of control of *glmY* and *glmZ* expression operative in enterobacterial species as described for *Y. pseudotuberculosis*, *S. typhimurium* and *E. coli*. Sequence alignment analyses (Figs. S2.3 and S2.4) suggest that these species are representatives for other species showing similar *glmY* and *glmZ* promoter architectures, respectively (Fig. 2.8). Most importantly, our results suggest that in most species expression of both

sRNAs is controlled by  $\sigma^{54}$  and the response regulator GlrR (Fig. 2.8). This adds two sRNA genes to the regulon governed by  $\sigma^{54}$  in *Enterobacteriaceae*. The *glmY* and *glmZ* genes of *Y. pseudotuberculosis* exhibit all features of canonical  $\sigma^{54}$ -dependent genes. Their expression depends on  $\sigma^{54}$  (Figs. 2.2 and 2.3) and on binding of the activator protein GlrR to ABS present upstream of the  $\sigma^{54}$ -promoter, as demonstrated for *Y. pseudotuberculosis glmZ* (Figs. 2.4 and S2.8).



**Figure 2.8: Model illustrating the roles of the TCS GlrK/GlrR,  $\sigma^{54}$  and IHF for transcription of sRNA genes *glmY* and *glmZ* in *Enterobacteriaceae*.** Histidine kinase GlrK phosphorylates response regulator GlrR, which stimulates binding of GlrR to its target sites on the DNA. GlrR binds to three activator binding sites present upstream of  $\sigma^{54}$ -dependent promoters that control the expression of sRNA genes *glmY* in all species and *glmZ* in a subset of species. GlrR, which contains a  $\sigma^{54}$  interaction domain, is absolutely required for activity of these  $\sigma^{54}$ -promoters. In addition, promoter activity depends on IHF, which might facilitate interaction of GlrR with the  $\sigma^{54}$ -RNA polymerase by binding-induced bending of the promoter DNA. In *Y. pseudotuberculosis*, transcription of *glmY* and *glmZ* is directed by single  $\sigma^{54}$ -promoters that require activation by GlrR. Hence, *glmY* and *glmZ* compose a regulon controlled by GlrR and  $\sigma^{54}$ . A similar arrangement is found in *S. typhimurium*, but  $\sigma^{70}$ -promoters that overlap the  $\sigma^{54}$ -promoters additionally contribute to *glmY* and *glmZ* expression. Overlapping  $\sigma^{54}$ - and  $\sigma^{70}$ -promoters also direct expression of the *E. coli glmY* gene, while expression of *glmZ* is achieved from a single constitutively active  $\sigma^{70}$ -promoter. Sequence alignment analyses suggest that these three different arrangements might also apply to other enterobacterial species as shown in the Figure.

In conclusion, transcription is initiated from single  $\sigma^{54}$ -promoters that require activation by GlrR and the same may also hold true for species of the genera *Arsenophonus*, *Dickeya*, *Erwinia*, *Pectobacterium*, *Photorhabdus*, *Proteus* and *Serratia* (Fig. 2.8). A somewhat different scenario is operative in the case of *S. typhimurium glmY* and *glmZ*. The corresponding promoter regions also contain three ABS and a  $\sigma^{54}$ -promoter. Accordingly, GlrR specifically binds to these regions and stimulates transcription (Figs. 2.2 and 2.3). However, both genes are still expressed in mutants lacking  $\sigma^{54}$ , which is at first glance incompatible with the properties of genuine  $\sigma^{54}$ -dependent genes. The expression in the absence of  $\sigma^{54}$  is explained by additional  $\sigma^{70}$ -promoters that overlap the  $\sigma^{54}$ -promoters and can potentially start transcription at the same site. According to the sequence alignment, such overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters may also exist in *Citrobacter*, *Cronobacter* and *Enterobacter* species (Fig. 2.8). We have recently shown that in *E. coli* transcription of *glmY* is controlled by a similar mechanism (Reichenbach *et al.*, 2009). In contrast, *E. coli glmZ* is not controlled by GlrR or  $\sigma^{54}$  and accordingly GlrR does not bind the *E. coli glmZ* promoter (Fig. 2.3). A single constitutively active  $\sigma^{70}$ -promoter directs expression of *glmZ* in *E. coli* (Fig. 2.5) and

presumably also in *Klebsiella* and other *Escherichia* species (including *Shigella*) (Fig. 2.8). In sum, our work suggests that *glmY* and *glmZ* transcription is controlled by  $\sigma^{54}$  and the TCS GlrR/GlrK in most Enterobacteria, but in a subset of species this relation is gradually lost in favor of unregulated  $\sigma^{70}$ -dependent transcription.

How did these different scenarios evolve? GlmY and GlmZ are homologous sRNAs (Reichenbach *et al.*, 2008; Urban and Vogel, 2008). A sequence alignment of the *glmY/glmZ* genes of several species reveals sequence elements that are conserved in both sRNAs, while the *glmS* binding site is exclusively present in GlmZ species (Fig. S2.10). A phylogenetic tree built from this sequence alignment clusters *glmZ* genes together, while the *glmY* genes form a distinct group (Fig. S2.11). A similar clustering can be observed when the sequences of the corresponding promoter regions are used for tree construction (Fig. S2.12). Accordingly, *glmY* and *glmZ* most likely originated from duplication of a single sRNA locus in an ancestor of *Enterobacteriaceae* and transcription of this ancient sRNA was presumably already controlled by  $\sigma^{54}$  and GlrR. Following duplication, divergence of the promoter regions by mutation might have generated the different promoter architectures detectable in recent bacteria.

What is the physiological meaning of regulation of *glmY/glmZ* transcription by GlrR/GlrK? In *E. coli*, GlmYZ feedback-regulate synthesis of the enzyme GlmS and are therefore crucial for maintaining the intracellular GlcN6P concentration required for undisturbed synthesis of the cell wall and the outer membrane (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008). This important role of GlmYZ may also apply to other *Enterobacteriaceae*, since the GlmZ/*glmS* base-pairing appears to be conserved (Fröhlich and Vogel, 2009). In *E. coli*, a decrease in the intracellular GlcN6P concentration induces accumulation of GlmY, which in turn increases concentration of the full-length form of GlmZ that is competent in *glmS* base-pairing (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008). Most likely, GlmY acts on GlmZ through sequestration of a protein that targets GlmZ to processing (Reichenbach *et al.*, 2008; Urban and Vogel, 2008), but it is unknown whether this mechanism is also operative in other species. In conclusion, up-regulation of the GlmYZ cascade in response to GlcN6P depletion occurs at the post-transcriptional level, and involves stabilization of the sRNAs rather than activation of their transcription in *E. coli* (Reichenbach *et al.*, 2009). Accordingly, the basal level of transcription of the sRNAs, as observed in the exponential growth phase, is sufficient for this function. However, GlrR/GlrK strongly up-regulate *glmY* expression through activation of the  $\sigma^{54}$ -promoter, when cells enter the stationary growth phase (Reichenbach *et al.*, 2009). In contrast, GlmZ levels decrease, i.e. stabilization of GlmZ as a consequence of accumulation of GlmY does not occur in this growth phase (Reichenbach *et al.*, 2008). Hence, GlmY accumulates in *E. coli* when growth ceases and ongoing cell wall synthesis and up-regulation of *glmS* are not required. This indicates a second function of GlmY, which requires a higher concentration of the sRNA and becomes relevant during transition to the stationary growth phase. We speculate that GlmY may have multiple functions and this may also hold for GlmZ in those species, which control expression of both sRNA through GlrR/GlrK: GlmYZ regulate *glmS* and thereby GlcN6P synthesis during the exponential growth phase and basal expression levels are sufficient for this purpose. In addition, they might have another function that requires further up-regulation of the sRNAs through the TCS GlrR/GlrK. What is this additional function? Interestingly, GlrR/GlrK have been implicated to play a role for virulence: Mutants of *Y. pseudotuberculosis* lacking GlrR exhibited reduced pathogenicity in mice (Flamez *et al.*, 2008). In enterohemorrhagic *E. coli* (EHEC) GlrR/GlrK (QseF/QseE) are required for transcription of *espFU*, which is an EHEC-specific gene and encodes an effector protein translocated to the host cell. Consequently, loss of GlrR/GlrK results

in the inability to form attaching and effacing lesions that are required for destruction of microvilli, pedestal formation and rearrangement of the cytoskeleton of host cells (Reading *et al.*, 2007; Reading *et al.*, 2009). In conclusion, GlrR/GlrK controls functions important for interaction with eukaryotic cells in at least two different bacteria. Whether this also holds for other *Enterobacteriaceae* and involves GlmY(Z) remains to be determined.

What is the reason for the existence of additional  $\sigma^{70}$ -promoters overlapping with the  $\sigma^{54}$ -dependent *glmY/glmZ* promoters in a subgroup of *Enterobacteriaceae*? They may allow better fine-tuning of the expression to meet the requirements of the multiple functions of these sRNAs, e.g. the  $\sigma^{70}$ -promoters ensure sRNA expression when the activating signal for GlrR/GlrK is absent and the  $\sigma^{54}$ -promoter is inactive. Alternatively, the  $\sigma^{70}$ -promoters could also be regulated and may allow regulation of the sRNAs in response to another yet unknown process. It is also possible, that the functional overlap of  $\sigma^{54}$ - and  $\sigma^{70}$ -dependent promoters is a more global phenomenon in certain species such as *E. coli*. Extensive functional overlap with  $\sigma^{70}$ -promoters has been observed for  $\sigma^{24}$ - and  $\sigma^{32}$ -dependent genes in *E. coli* (Wade *et al.*, 2006). Both,  $\sigma^{24}$  and  $\sigma^{32}$  recognize distinct promoter sequences. However, many of these promoters also contain matches to overlapping  $\sigma^{70}$ -promoters. Thus, the majority of the  $\sigma^{32}$ -promoters and about half of the  $\sigma^{24}$ -promoters are also recognized by  $\sigma^{70}$ -RNAP and transcription initiation at the same start site was demonstrated for some of these promoters (Wade *et al.*, 2006). This was interpreted to mean that the primary function of alternative  $\sigma$  factors is to increase transcription of  $\sigma^{70}$ -dependent genes. A recent study reported that 14% of the  $\sigma^{54}$ -dependent genes in *E. coli* can also be transcribed by  $\sigma^{70}$ -RNAP *in vitro* (Zhao *et al.*, 2010). Whether this occurs from overlapping or consecutive promoters is not known. However, our studies prove that arrangements of overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -promoters exist ((Reichenbach *et al.*, 2009); the present study). It remains to be elucidated whether functional overlap between  $\sigma^{70}$  and  $\sigma^{54}$  is a peculiarity of *E. coli* and its closest relatives or may apply to a wider range of bacterial species.

Activation of the  $\sigma^{54}$ -dependent *glmY* and *glmZ* promoters requires binding of GlrR to ABS located upstream of the promoter. However, the impact of each of the three ABS on the promoter activity appears to vary from case to case, e.g. ABS2 and ABS3 were shown to be essential for activity of the  $\sigma^{54}$ -promoter of *E. coli glmY* (Reichenbach *et al.*, 2009), while mutation of one of these sites upstream of *Y. pseudotuberculosis glmZ* reduced promoter activity only two-fold (Fig. 2.4 A). ABS1 appears to be dispensable for promoter activity in both cases, as reflected by its lower degree of conservation. Interaction of activator proteins with  $\sigma^{54}$ -RNAP requires bending of the DNA, which is usually induced by IHF (Swinger and Rice, 2004; Wigneshweraraj *et al.*, 2008). IHF might also be required for the activities of the  $\sigma^{54}$ -dependent *glmY* and *glmZ* promoters as demonstrated for the  $\sigma^{54}$ -promoter of *E. coli glmY* (Fig. 2.6). Two putative IHF binding sites were detected and we demonstrated an essential role for  $\sigma^{54}$ -promoter activity for the distal site (Fig. 2.6 D). We also provided evidence that phosphorylation of GlrR enhances *glmY* expression (Fig. 2.7). Substitution of the phosphorylation site Asp 56 with Ala reduced *glmY* expression two-fold, whereas a Glu exchange mimicking phosphorylation led to much stronger expression (Fig. 2.7 B). In addition, pre-incubation of GlrR with acetyl phosphate increased its binding affinity for the *glmY* promoter (Fig. 2.7 A). Taken together this indicates that the DNA-binding activity of GlrR is activated by its phosphorylation although it cannot be excluded yet that the mutations in GlrR affected the stability rather than activity of the protein. Our data indicate that just a minor fraction of GlrR is phosphorylated by GlrK

during exponential growth, which is in line with previous data suggesting that this TCS drastically increases *glmY* expression at the on-set of the stationary growth phase in *E. coli* (Reichenbach *et al.*, 2009). So far, *glmY* and *glmZ* are the only known direct targets of GlrR/GlrK suggesting that this TCS acts predominantly through these sRNAs.

## **FUNDING**

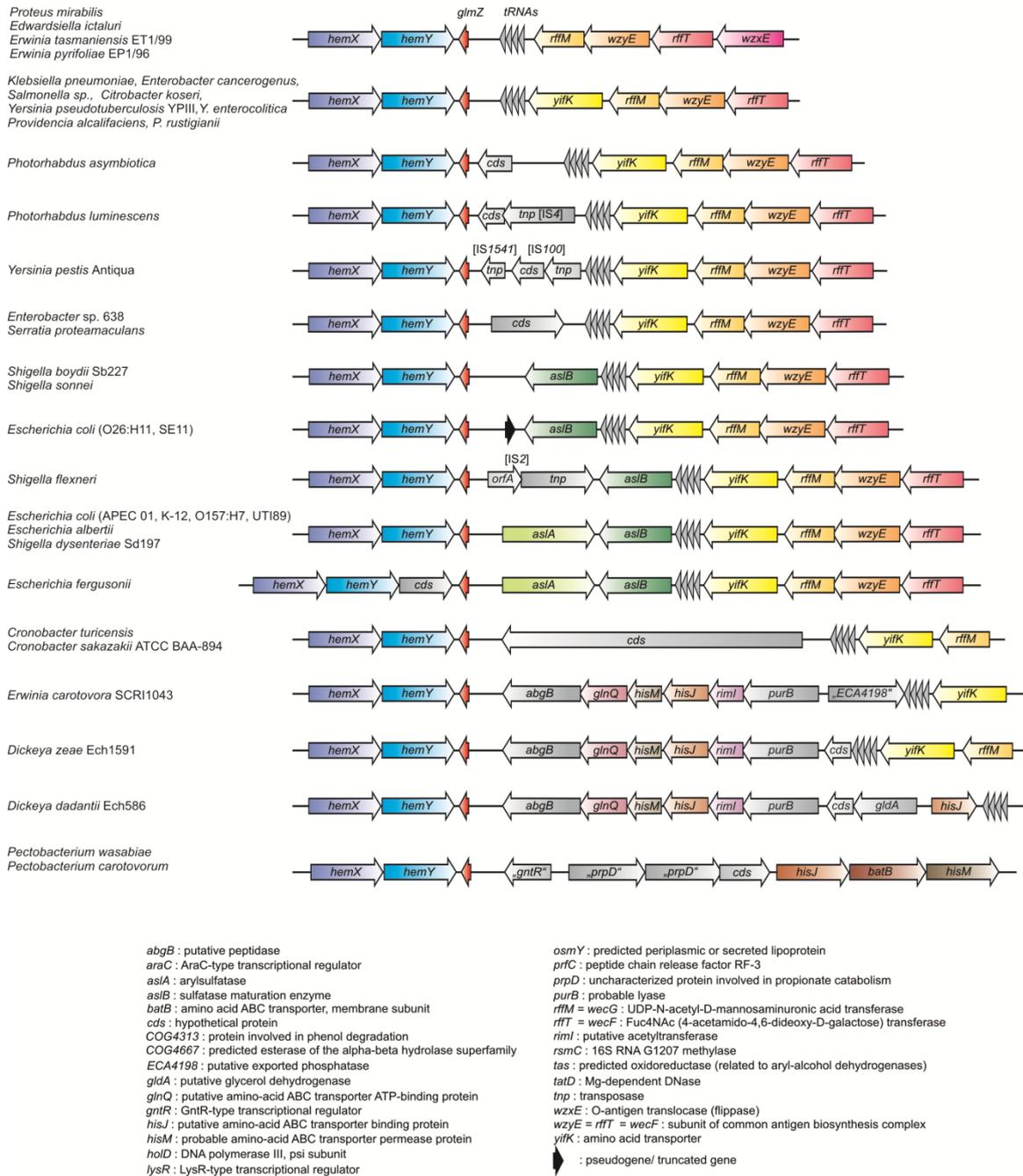
This work was supported by grants from the DFG Priority Program SPP1258 “Sensory and Regulatory RNAs in Prokaryotes” to B. G. and P. D..

## **ACKNOWLEDGEMENTS**

We thank Fabio Pisano, Karin Schnetz and Vanessa Sperandio for critical reading of the manuscript. We are grateful to Jacqueline Plumbridge for the gift of purified IHF protein and to Karin Schnetz for *S. typhimurium* LT2. We thank Sabine Lentjes for excellent technical assistance, Jörg Stülke for lab space and support and Konstantin Albrecht and Sabine Zeides for help with the construction of plasmids or strains.

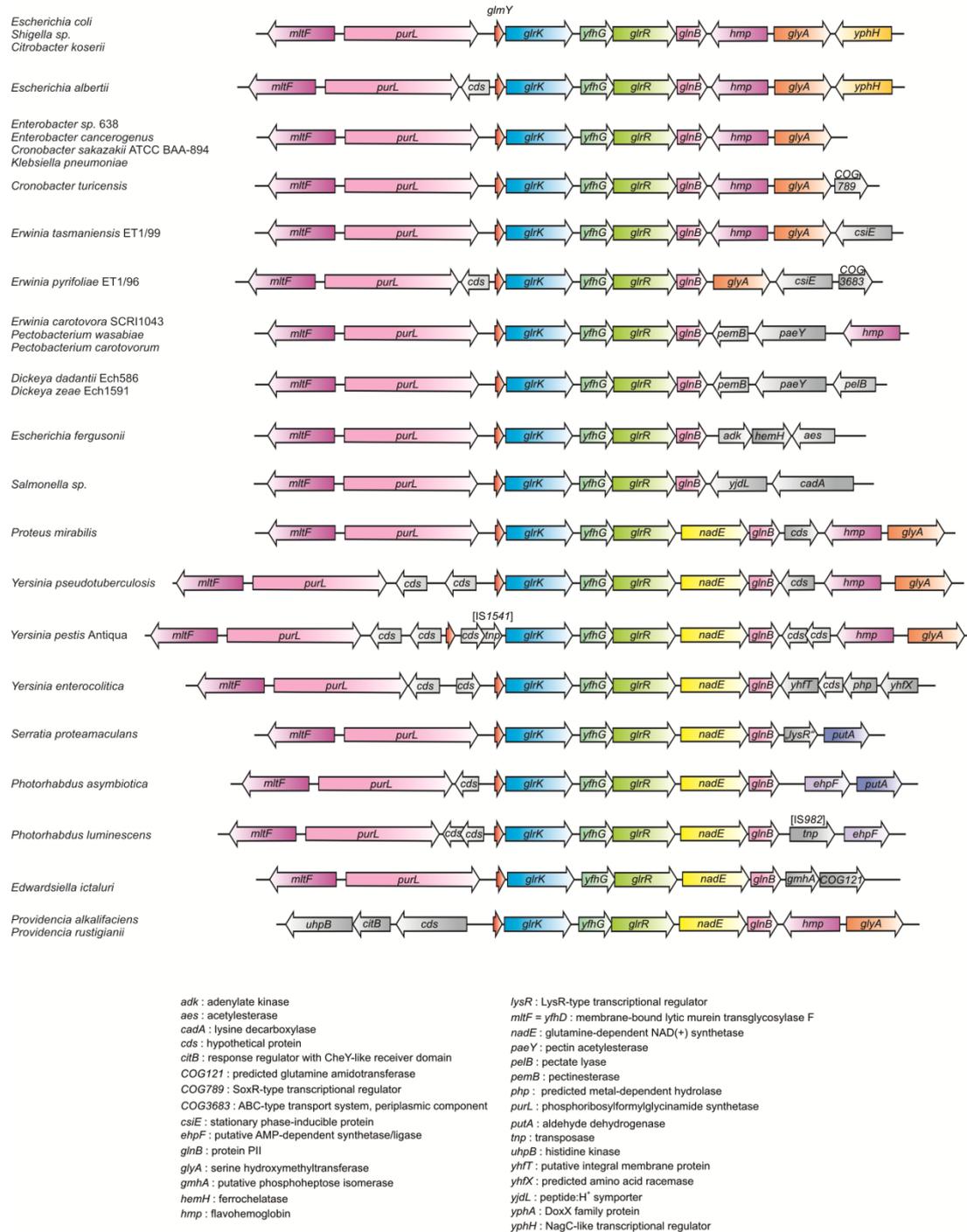
SUPPLEMENTARY MATERIAL

Fig. S2.1



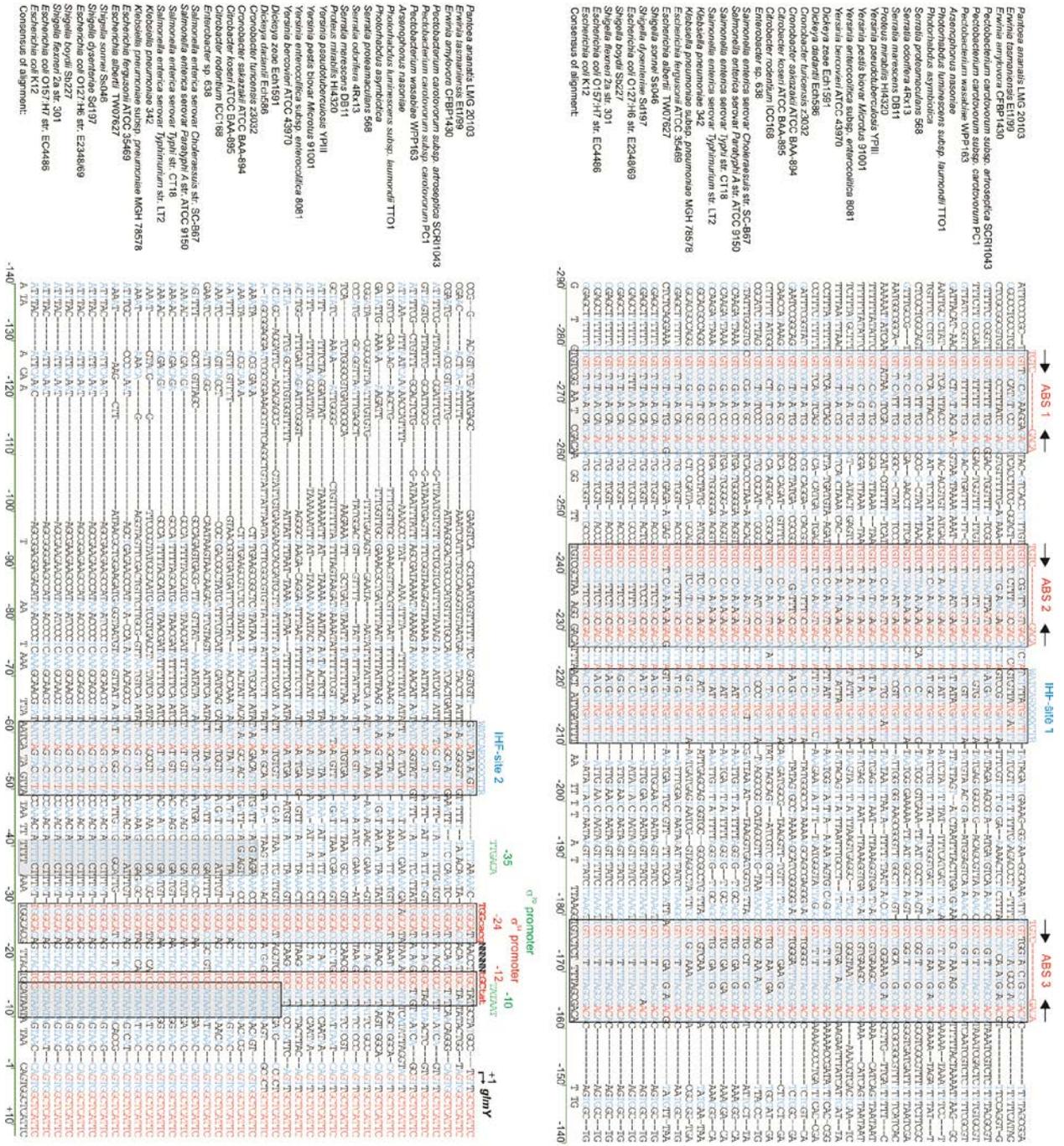
**Fig. S2.1. Gene synteny of the *glmZ* region in *Enterobacteriaceae*.** Information about gene co-localization and annotation was retrieved using the MicrobesOnline (Dehal *et al.*, 2009) and KEGG (Kanehisa *et al.*, 2008) databases. Genes are just approximately drawn to scale.

Fig. S2.2



**Fig. S2.2. Gene synteny of the *glmY* region in *Enterobacteriaceae*.** Information about gene co-localization and annotation was retrieved using the MicrobesOnline (Dehal *et al.*, 2009) and KEGG (Kanehisa *et al.*, 2008) databases. Genes are just approximately drawn to scale.

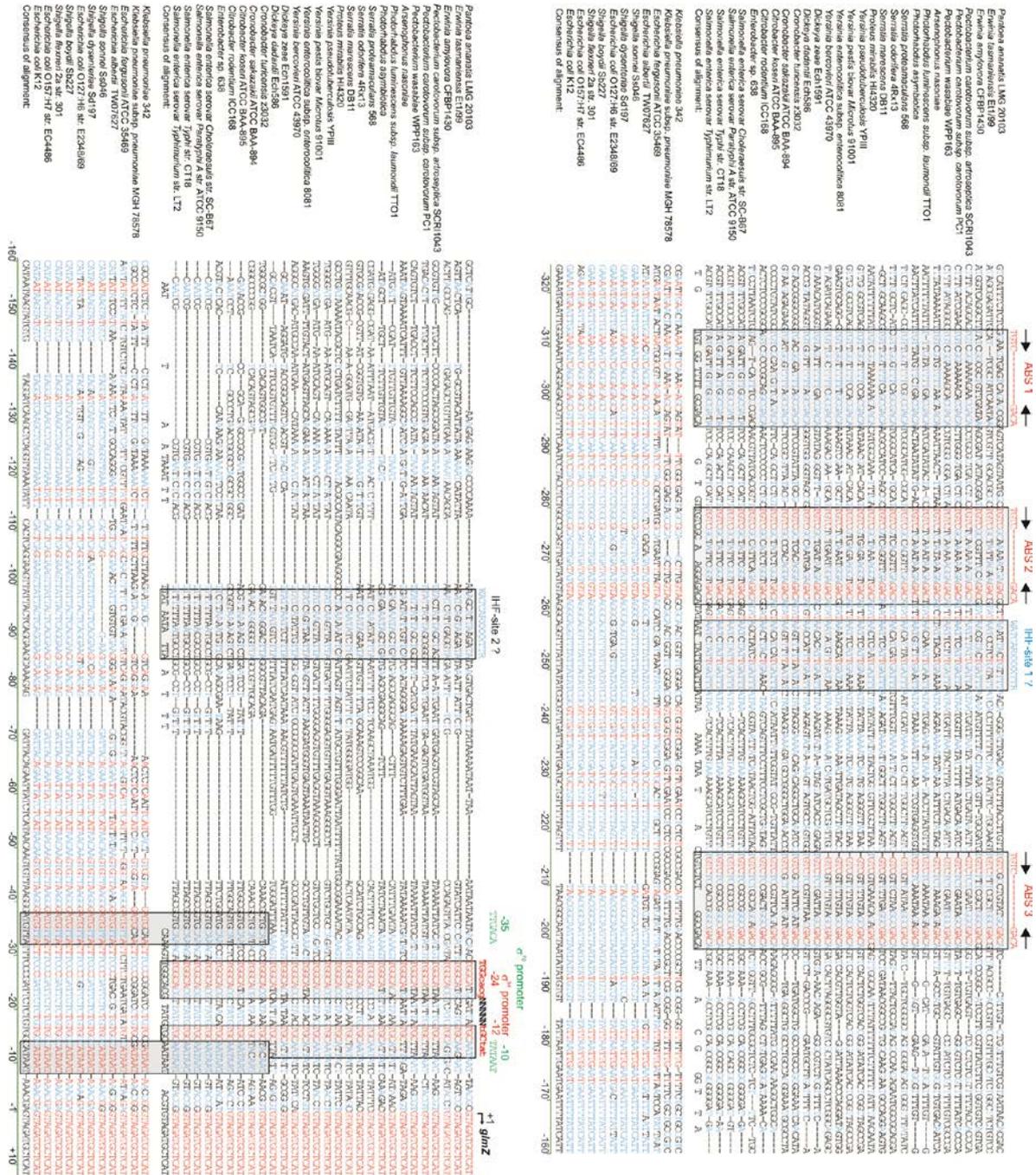
Fig. S2.3



**Fig. S2.3. Sequence alignment of the *glmY* promoter regions from 39 enterobacterial genomes.** Fully conserved nucleotide positions are highlighted in red, while residues conserved in the majority of these sequences are in blue. The transcriptional start site of *glmY* is marked with an arrow. The GlrR binding sites (ABS), putative IHF binding sites and the -24/-12 sequence motifs of  $\sigma^{54}$  promoters are boxed. The -35/-10 sequence motifs of overlapping  $\sigma^{70}$ -promoters are also boxed. The respective consensus sequences are shown above the alignment. Sequences were compiled from the following genomes (accession numbers are in parentheses): *Pantoea ananatis* LMG 20103 (NC\_013956.1), *Erwinia tasmaniensis* Et1/99 (NC\_010694.1), *Erwinia amylovora* CFBP 1430 (NC\_013961.1), *Pectobacterium carotovorum* subsp. *atroseptica* SCRI1043 (NC\_004547.2), *Pectobacterium carotovorum* subsp. *carotovorum* PC1 (NC\_012917.1), *Pectobacterium wasabiae* WPP163 (NC\_013421.1), *Arsenophonus nasoniae* (FN545167.1), *Photobacterium luminescens* subsp. *laumondii* TTO1 (NC\_005126.1), *Photobacterium asymbiotica* (NC\_012962.1), *Serratia proteamaculans* 568 (NC\_009832.1), *Serratia odorifera* 4Rx13 (NZ\_ADBX01000009.1), *Serratia marescens* Db11 [http://www.sanger.ac.uk], *Proteus mirabilis* HI4320 (NC\_010554.1), *Yersinia pseudotuberculosis* YPIII (NC\_010465.1), *Yersinia pestis* biovar *Microtus* str. 91001 (NC\_005810.1), *Yersinia enterocolitica* subsp. *enterocolitica* 8081 (NC\_008800.1), *Yersinia bercovieri* ATCC 43970 (NZ\_AALC02000019.1), *Dickeya zeae* Ech1591 (NC\_012912.1), *Dickeya dadantii* Ech586 (NC\_013592.1), *Cronobacter turicensis* z3032 (NC\_013282.1), *Cronobacter sakazakii* ATCC BAA-894 (NC\_009778.1), *Citrobacter koseri* ATCC BAA-895 (NC\_009792.1), *Citrobacter rodentium* ICC168 (NC\_013716.1), *Enterobacter* sp. 638 (NC\_009436.1), *Salmonella enterica* subsp. *enterica* serovar

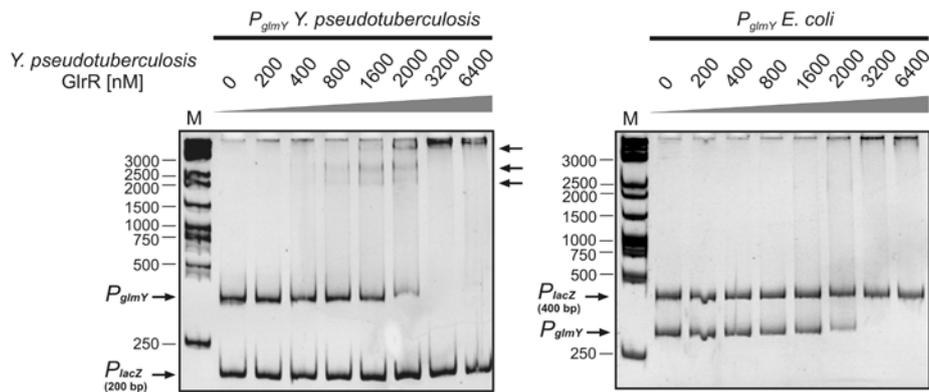
*Choleraesuis* str. SC-B67 (NC\_006905.1), *Salmonella enterica* subsp. *enterica* serovar Paratyphi A str. ATCC 9150 (NC\_006511.1), *Salmonella enterica* subsp. *enterica* serovar Typhi str. CT18 (NC\_003198.1), *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 (NC\_003197.1), *Klebsiella pneumoniae* 342 (NC\_011283.1), *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 (NC\_009648.1), *Escherichia fergusonii* ATCC 35469 (NC\_011740.1), *Escherichia albertii* TW07627 (NZ\_ABKX01000003.1), *Shigella sonnei* Ss046 (NC\_007384.1), *Shigella dysenteriae* Sd197 (NC\_007606.1), *Escherichia coli* O127:H6 str. E2348/69 (NC\_011601.1), *Shigella boydii* Sb227 (NC\_007613.1), *Shigella flexneri* 2a str. 301 (NC\_004337.1), *Escherichia coli* O157:H7 str. EC4486 (NZ\_ABHS01000009.1), *Escherichia coli* K12 str. MG1655 (U00096.2). The alignment was compiled using the AlignX tool of software Vector NTI Advance™ 9.0.

Fig. S2.4



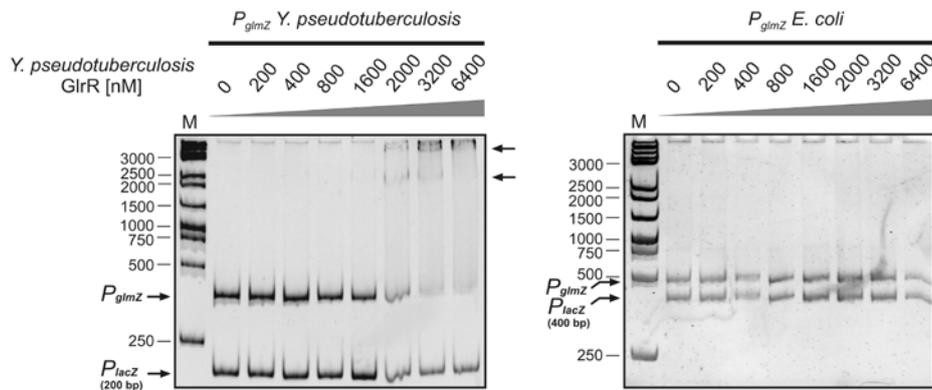
**Fig. S2.4. Sequence alignment of the *glmZ* promoter regions from 39 enterobacterial genomes.** The sequences classified into two groups, which exhibited no significant homologies to each other and are therefore shown in separate alignments. Fully conserved nucleotide positions within each group are highlighted in red, while residues conserved in the majority of sequences are in blue. Refer to legend to Fig. S2.3 for additional information. The same genome sequences as in Fig. S2.3 were used.

Fig. S2.5



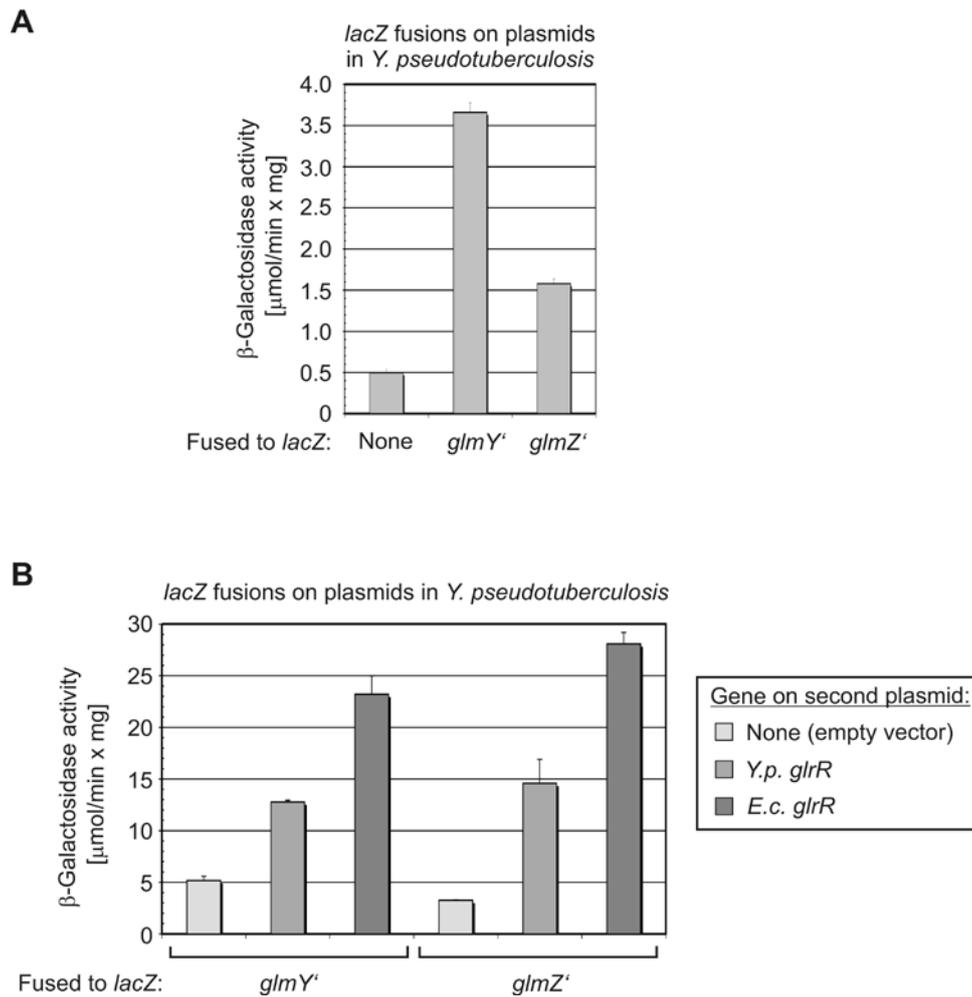
**Fig. S2.5.** *Y. pseudotuberculosis* GlrR binds the *glmY* promoters of *Y. pseudotuberculosis* and *E. coli*. EMSAs to test binding of *Y. pseudotuberculosis* GlrR protein to the *glmY* promoter regions of *Y. pseudotuberculosis* (-257 to +22) and *E. coli* (-238 to +22). The DNA fragments were obtained by PCR using the same primers as for construction of the corresponding *glmY*-*lacZ* fusions tested in Fig. 2.2 B. In addition to the *glmY* promoter fragments, 200 bp (panel 1) or 400 bp DNA fragments (panel 2) covering the *lacZ* promoter were present as internal controls. The sizes of the DNA size standard are given on the left.

Fig. S2.6



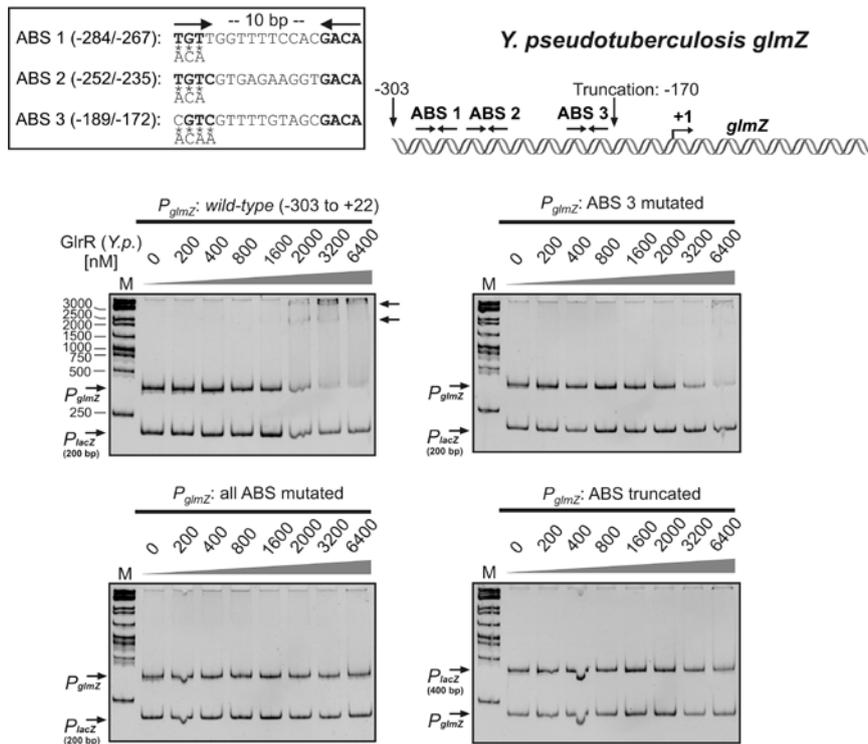
**Fig. S2.6.** *Y. pseudotuberculosis* GlrR binds the cognate *glmZ* promoter, while the *E. coli* *glmZ* promoter is not bound. EMSAs to test binding of *Y. pseudotuberculosis* GlrR protein to the *glmZ* promoter regions of *Y. pseudotuberculosis* (-303 to +22) and *E. coli* (-424 to +32). The DNA fragments were obtained by PCR using the same primers as for construction of the corresponding *glmZ*-*lacZ* fusions used in Fig. 2.3 B

Fig. S2.7



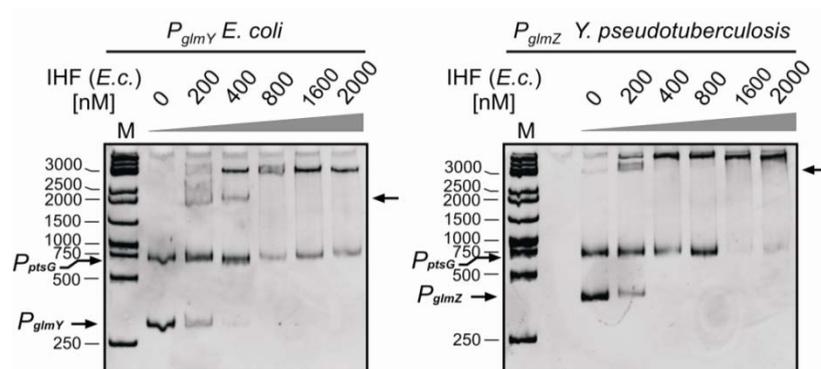
**Fig. S2.7. Expression of *glmY* and *glmZ* in *Y. pseudotuberculosis*.** (A) *glmY* and *glmZ* are expressed in *Y. pseudotuberculosis*.  $\beta$ -Galactosidase activities of *Y. pseudotuberculosis* carrying a *glmY*'-*lacZ* (column 2) or *glmZ*'-*lacZ* fusion (column 3) on a plasmid (pYG1 and pYG2, respectively). Cells carrying the empty *lacZ* fusion vector pKEM04 served as background control (column 1). (B) Stimulation of *glmY* and *glmZ* expression by GlrR in *Y. pseudotuberculosis*. Additional plasmids carrying either *glrR* from *Y. pseudotuberculosis* (plasmid pYG6, columns 2, 5) or *E. coli* (plasmid pBGG223, columns 3, 6) or no gene (empty vector pBAD18-cm, columns 1, 4) under  $P_{Ara}$  promoter control were introduced into *Y. pseudotuberculosis* carrying either the *glmY*'-*lacZ* fusion plasmid pYG1 (columns 1-3) or the *glmZ*'-*lacZ* fusion plasmid pYG2, respectively (columns 4-6). For the induction of *glrR* expression 0.2% arabinose was added and subsequently the  $\beta$ -galactosidase activities were determined.

Fig. S2.8



**Fig. S2.8. Binding of the *Y. pseudotuberculosis* GlrR protein to its cognate *glmZ* promoter requires three activator binding sites.** EMSAs to monitor binding of *Y. pseudotuberculosis* GlrR to DNA fragments covering the *Y. pseudotuberculosis* *glmZ* promoter (-303 to +22). DNA fragments were tested that carried mutations within ABS3 (top, right) or in all ABS simultaneously (bottom, left). These DNA fragments were obtained by PCR using the corresponding *glmZ*-*lacZ* fusions, presented in Fig. 2.4 A, as template and primers BG700/BG701. In addition, a truncated DNA fragment lacking all ABS was tested (-170 to +22; bottom, right). The DNA fragment was obtained using the primer pair BG755/BG701.

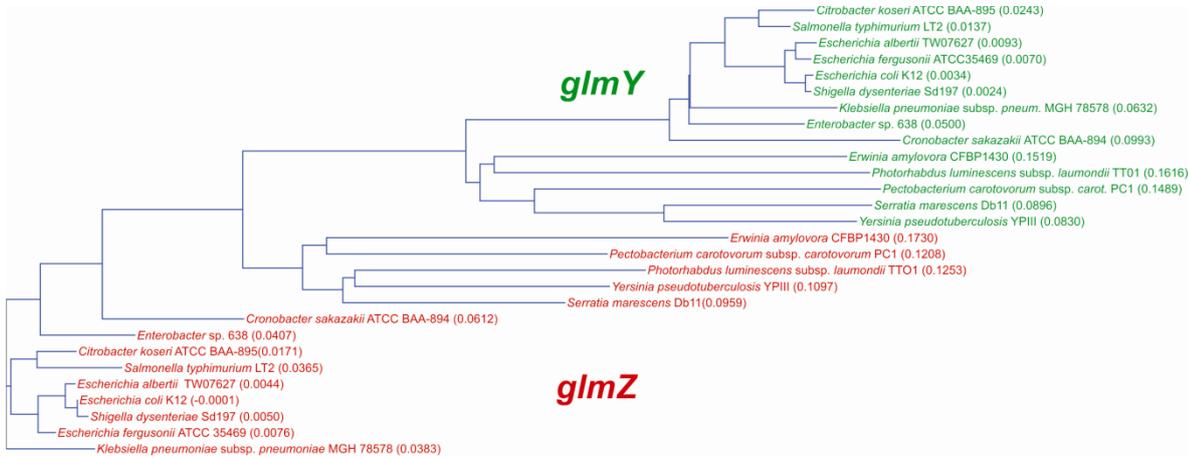
Fig. S2.9



**Fig. S2.9. Binding of IHF to the *E. coli* *glmY* and the *Y. pseudotuberculosis* *glmZ* promoter regions as revealed by EMSA.** The DNA fragments were obtained by PCR making use of the primer pairs BG377/BG456 and BG700/BG701, respectively. As a difference to the experiments shown in Fig. 2.6 B, a DNA fragment covering the *ptsG* promoter ( $P_{ptsG}$ ) from *Bacillus subtilis* was used as internal control rather than a *lacZ* promoter fragment. The  $P_{ptsG}$  fragment was amplified from the *B. subtilis* chromosome using primers IL5 (Langbein *et al.*, 1999) and JS11 (Stülke *et al.*, 1997).

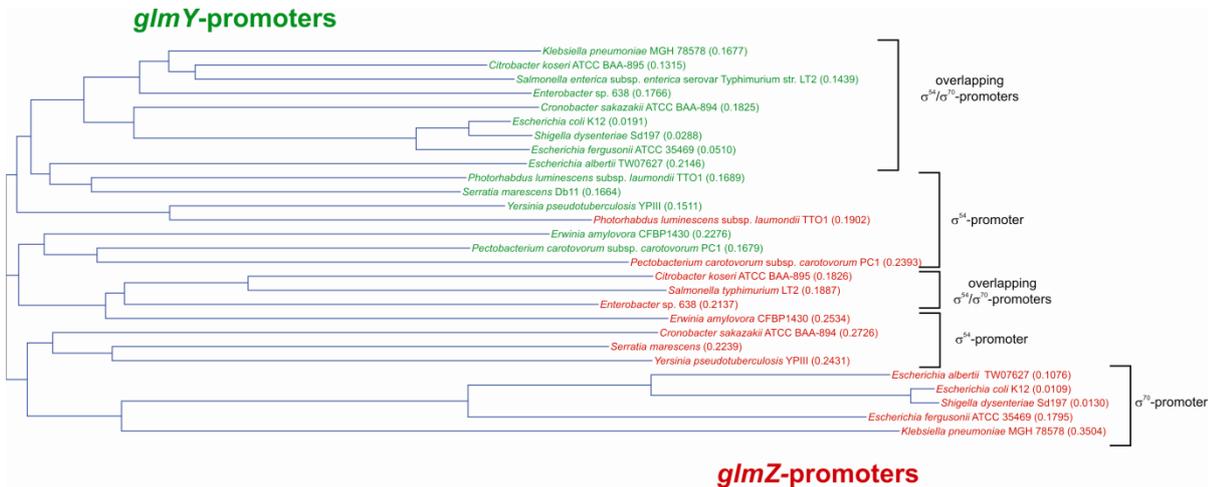


Fig. S2.11



**Fig. S2.11. Phylogenetic tree of *glmY* and *glmZ* genes of 14 enterobacterial species.** The tree was calculated from the sequence alignment shown in Fig. S2.10 using the AlignX tool of software Vector NTI Advance™ 9.0. The tree is built using the Neighbor joining method (Saitou and Nei, 1987), which works on a matrix of distances between all possible sequence pairs. The calculated distance values, which are related to the degree of divergence between the sequences, are shown in parentheses.

Fig. S2.12



**Fig. S2.12. Phylogenetic tree of *glmY*- and *glmZ*-promoter regions of 14 enterobacterial species.** The tree was calculated from a sequence alignment (data not shown) comprising the promoter sequences of the *glmY*- and *glmZ*-genes used for Fig. S2.11. The used sequences are shown in Figures S2.3 and S2.4, but the sequences downstream of the transcriptional start sites of the sRNAs were omitted. See Fig. S2.11 for additional information.

**SUPPLEMENTAL “MATERIALS AND METHODS”****Construction of plasmids**

For construction of the fusions of *Y. pseudotuberculosis glmY* (-257 to +22) and *glmZ* (-303 to +22) to *lacZ*, the corresponding *glmY*-5' and *glmZ*-5' regions were amplified from the *Y. pseudotuberculosis* chromosome using the primer combinations BG698/BG699 and BG700/BG701, respectively. The PCR fragments were subsequently used to replace the Sall-XbaI fragment in plasmid pKES15, which yielded plasmids pYG1 and pYG2, respectively. To obtain a *Y. pseudotuberculosis glmZ'*-*lacZ* fusion carrying a mutated ABS1, this mutation was introduced by PCR using forward primer BG747 together with BG701. Insertion of this fragment between the Sall/XbaI-sites of pKES15 resulted in plasmid pYG9. Mutations in ABS2 and ABS3 were introduced by multiple mutation reaction (MMR; (Hames *et al.*, 2005)). To this end, the 5'-phosphorylated oligonucleotides BG748 and/or BG754 carrying the mutations in ABS2 and ABS3, respectively, were used in addition to the forward primers BG700 or BG747 (mutation of ABS1) and reverse primer BG701 in PCRs. These PCRs contained thermo-stable Ampligase (Epicentre), which incorporated the mutagenesis primers during amplification. Insertion of the PCR fragments between the Sall/XbaI-sites of pKES15 yielded plasmids pYG10 (ABS2 mutated), pYG11 (ABS3 mutated) and pYG12 (all ABS mutated). The *glmY*-5' (-242 to +22) and *glmZ*-5' (-242 to +22) regions of *S. typhimurium* were amplified from chromosomal DNA using the primer combinations BG750/BG751 and BG752/BG753, respectively, and the PCR fragments were inserted between the Sall/XbaI sites of plasmid pKES15 to yield plasmids pYG7 and pYG8. Plasmids pBGG390 and pBGG391 are isogenic with plasmids pBGG201 and pBGG209, but carry mutations within the putative IHF1-site in the *E. coli glmY* upstream region. Plasmid pBGG390 was constructed by MMR using pBGG201 as template, BG377 and BG456 as external primers and the phosphorylated mutagenesis primer BG684. The resulting PCR fragment was cloned via Sall/XbaI into plasmid pKES15. To introduce the mutation in the -10 sequence, the AflIII-SacI fragment of pBGG390 was replaced by the corresponding fragment of pBGG209 resulting in plasmid pBGG391. The plasmids carrying the gradually 5'-truncated *E. coli glmZ'*-*lacZ* fusions were also constructed by cloning PCR fragments that were amplified from the *E. coli* chromosome between the Sall/XbaI sites of pKES15. The PCR fragments were obtained using reverse primer BG202 and one of the following forward primers resulting in the plasmid as indicated in parentheses: BG200 (pBGG111), BG333 (pBGG112), BG334 (pBGG113), BG335 (pBGG114), BG411 (pBGG170). Plasmid pBGG135 carrying the *glmZ'*(-11 to +32)-*lacZ* fusion was constructed by ligation of hybridized 5'-phosphorylated oligonucleotides BG347 and BG348 with the Sall/XbaI-digested vector pKES15. Hybridization was achieved by heating 150  $\mu$ l of a solution containing 20 pMol of each oligonucleotide, 10 mM Tris/HCl pH 7.5 and 1 M NaCl to 99°C followed by slow cooling and precipitation with ethanol. Plasmids pBGG157 and pBGG171 carry mutated -35 and -10 sequences in the context of the *glmZ'*(-40 to +32)-*lacZ* fusion, respectively. The mutations were introduced by forward primers BG388 and BG412, respectively, in PCRs using BG202 as reverse primer. The PCR fragments were subsequently inserted between the Sall/XbaI sites of plasmid pKES15. For construction of plasmid pBGG397 carrying *Y. pseudotuberculosis glrR::His10* under *tacOP* control, *glrR* was amplified from the *Y. pseudotuberculosis* chromosome using primers BG696 and BG697. Subsequently, the PCR product was inserted between the NdeI- and XbaI-sites on plasmid pKES170. For construction of plasmid pYG6 carrying *Y. pseudotuberculosis glrR* under *P<sub>Ara</sub>* promoter control, *glrR* was amplified using primers BG727/BG728 and cloned between the SacI and XbaI sites on plasmid pBAD18-cm. To construct plasmid pBGG389, which carries *E. coli glrR* under *P<sub>Ara</sub>* control, the SacI-HindIII fragment of plasmid pBGG223 encompassing the *glrR* gene was cloned

between these sites on plasmid pBAD33. To obtain the isogenic plasmids pBGG398 and pBGG399, which code for the *glrR*-D56A and *glrR*-D56E alleles, MMRs were carried out using pBGG223 as template, the external primers BG490/BG491 and the mutagenesis primers BG685 and BG686, respectively. The MMR fragments were subsequently cloned between the *SacI*/*XbaI* sites on plasmid pBAD33.

## SUPPLEMENTAL TABLES

**TABLE S2.1.** Plasmids used in this study

| Name      | Genotype or relevant structures <sup>a</sup>   | Reference or construction          |
|-----------|--|------------------------------------|
| pBAD18-cm | <i>P<sub>Ara</sub></i> , MCS 2, <i>cat</i> , ori pBR322  | (Guzman <i>et al.</i> , 1995)      |
| pBAD33    | <i>P<sub>Ara</sub></i> , MCS 2, <i>cat</i> , ori p15A  | (Guzman <i>et al.</i> , 1995)      |
| pBGG59    | Fusion of <i>E.c. glmZ'</i> (-424 to +32) to <i>lacZ</i>   | (Kalamorz <i>et al.</i> , 2007)    |
| pBGG111   | Fusion of <i>E.c. glmZ'</i> (-207 to +32) to <i>lacZ</i>   | this work                          |
| pBGG112   | Fusion of <i>E.c. glmZ'</i> (-100 to +32) to <i>lacZ</i>   | this work                          |
| pBGG113   | Fusion of <i>E.c. glmZ</i> (-80 to +32) to <i>lacZ</i>   | this work                          |
| pBGG114   | Fusion of <i>E.c. glmZ</i> (-40 to +32) to <i>lacZ</i>   | this work                          |
| pBGG135   | Fusion of <i>E.c. glmZ'</i> (-11 to +32) to <i>lacZ</i>  | this work                          |
| pBGG157   | Fusion of <i>E.c. glmZ</i> (-40 to +32) to <i>lacZ</i> , -35 region mutated  | this work                          |
| pBGG170   | Fusion of <i>E.c. glmZ'</i> (-20 to +32) to <i>lacZ</i>  | this work                          |
| pBGG171   | Fusion of <i>E.c. glmZ</i> (-40 to +32) to <i>lacZ</i> , -10 region mutated  | this work                          |
| pBGG201   | Fusion of <i>E.c. glmY'</i> (-238 to +22) to <i>lacZ</i>   | (Reichenbach <i>et al.</i> , 2009) |
| pBGG209   | Fusion of <i>E.c. glmY'</i> (-238 to +22) to <i>lacZ</i> , -10 region mutated  | (Reichenbach <i>et al.</i> , 2009) |
| pBGG219   | <i>E.c. glrR</i> :: His <sub>10</sub> in pKES170   | (Reichenbach <i>et al.</i> , 2009) |
| pBGG223   | <i>E.c. glrR</i> under <i>P<sub>Ara</sub></i> -control in pBAD18-cm  | (Reichenbach <i>et al.</i> , 2009) |
| pBGG389   | <i>E.c. glrR</i> under <i>P<sub>Ara</sub></i> -control in pBAD33   | this work                          |
| pBGG390   | Fusion of <i>E.c. glmY'</i> (-238 to +22) to <i>lacZ</i> , IHF1 mutated  | this work                          |
| pBGG391   | Fusion of <i>E.c. glmY'</i> (-238 to +22) to <i>lacZ</i> , -10 region and IHF1 mutated                                       | this work                          |
| pBGG397   | <i>Y.p. glrR</i> :: His <sub>10</sub> in pKES170   | this work                          |
| pBGG398   | <i>E.c. glrR</i> (D56A) under <i>P<sub>Ara</sub></i> -control in pBAD33  | this work                          |
| pBGG399   | <i>E.c. glrR</i> (D56E) under <i>P<sub>Ara</sub></i> -control in pBAD33  | this work                          |
| pKEM04    | Promoter-less <i>lacZ</i> , <i>kan</i> , <i>attP</i> , <i>aadA</i> , ori p15A  | (Nagarajavel <i>et al.</i> , 2007) |
| pKES15    | <i>bgI'</i> - <i>lacZ</i> , <i>kan</i> , <i>attP</i> , <i>aadA</i> , ori p15A  | (Nagarajavel <i>et al.</i> , 2007) |
| pKES170   | <i>lacI<sup>q</sup></i> , <i>Ptac</i> , T7gene10-RBS, <i>NdeI</i> , <i>XbaI</i> , <i>rrnBT1/T2</i> , <i>bla</i> , pBR322-ori | (Reichenbach <i>et al.</i> , 2009) |
| pLDR8     | $\lambda$ <i>int</i> under control of $\lambda P_R$ , $\lambda cl_{857}$ , <i>kan</i> , ori pSC101- <i>rep</i> <sup>TS</sup> | (Diederich <i>et al.</i> , 1992)   |
| pYG1      | Fusion of <i>Y.p. glmY'</i> (-257 to +22) to <i>lacZ</i>   | this work                          |
| pYG2      | Fusion of <i>Y.p. glmZ'</i> (-303 to +22) to <i>lacZ</i>   | this work                          |
| pYG6      | <i>Y.p. glrR</i> under <i>P<sub>Ara</sub></i> -control in pBAD18-cm  | this work                          |
| pYG7      | Fusion of <i>S.t. glmY'</i> (-242 to +22) to <i>lacZ</i>   | this work                          |
| pYG8      | Fusion of <i>S.t. glmZ'</i> (-242 to +22) to <i>lacZ</i>   | this work                          |
| pYG9      | Fusion of <i>Y.p. glmZ'</i> (-292 to +22) to <i>lacZ</i> , ABS1 mutated  | this work                          |
| pYG10     | Fusion of <i>Y.p. glmZ'</i> (-303 to +22) to <i>lacZ</i> , ABS2 mutated  | this work                          |
| pYG11     | Fusion of <i>Y.p. glmZ'</i> (-303 to +22) to <i>lacZ</i> , ABS3 mutated  | this work                          |
| pYG12     | Fusion of <i>Y.p. glmZ'</i> (-292 to +22) to <i>lacZ</i> , ABS1,2,3 mutated  | this work                          |

<sup>a</sup>Positions are relative to the first nucleotide of the respective gene. Gene names are according to <http://ecocyc.org/>

**TABLE S2.2.** Oligonucleotides used in this study

| Primer | Sequence <sup>a</sup>  | Res.<br>Sites | Position <sup>b</sup>           |
|--------|--|---------------|---------------------------------|
| BG200  | GCACGCGTCGACGATGCTGTTTTAGTTTTAACGGC                                | Sall          | <i>E.c. glmZ</i> -207 to -183   |
| BG202  | GCGTCTAGAGGGCAACATAAGAGATGGAATGAGC                                 | Xbal          | <i>E.c. glmZ</i> +32 to +6      |
| BG333  | GCACGCGTCGACTCAGGAAGTTATTACTCAGGAAG                                | Sall          | <i>E.c. glmZ</i> -100 to -76    |
| BG334  | GCACGCGTCGACAAGCAAAGAGGATTACAGAATTATC                              | Sall          | <i>E.c. glmZ</i> -80 to -55     |
| BG335  | GCACGCGTCGACAGGGATGTTATTTCCCGATTCTC                                | Sall          | <i>E.c. glmZ</i> -40 to -17     |
| BG347  | [P]-TCGACATAATAAACGAGTAGATGCTCATTCCATCTCTATTGT<br>TCGCCT           |               | <i>E.c. glmZ</i> -11 to +32     |
| BG348  | [P]-CTAGAGGCGAACATAAGAGATGGAATGAGCATCTACTCGTT<br>TATTATG           |               | <i>E.c. glmZ</i> +32 to -11     |
| BG377  | GCACGCGTCGACCTTTTTGTGTCTGTAATCACG                                  | Sall          | <i>E.c. glmY</i> -238 to -213   |
| BG388  | GCACGCGTCGACAGGGAAATTTTTCCCGATTCTCTGTG                             | Sall          | <i>E.c. glmZ</i> -40 to -13     |
| BG411  | GCACGCGTCGACTCTGTGGCATAATAAACGAG                                   | Sall          | <i>E.c. glmZ</i> -20 to -1      |
| BG412  | GCACGCGTCGACAGGGATGTTATTTCCCGATTCTCTGTGGCATGCGAAACGAGTAGA<br>TGCTC | Sall          | <i>E.c. glmZ</i> -40 to +10     |
| BG456  | GCTCTAGAATAAGTCGGTGAATGAGCCAC                                      | Xbal          | <i>E.c. glmY</i> +22 to +2      |
| BG490  | GCGAGCTCCCATCCACCCATGAGGTCAC                                       | SacI          | <i>E.c. glrR</i> -25 to -5      |
| BG491  | GGCTCTAGATCATTTCCTGAAATCGTTTGCATC                                  | Xbal          | <i>E.c. glrR</i> +1335 to +1311 |
| BG578  | CGGTGAAGGGCAATCAGCTG   |               | <i>E.c. lacZ</i> -271 to -252   |
| BG579  | GGCCTTCGCTATTACGCC   |               | <i>E.c. lacZ</i> +129 to +110   |
| BG580  | ATTAATGCAGCTGGCAGCACG  |               | <i>E.c. lacZ</i> -171 to -150   |
| BG581  | ACGGCCAGTGAATCCGTAATC  |               | <i>E.c. lacZ</i> +29 to +9      |
| BG684  | [P]-GCGACACTTAACTCACCCCTTTAATATTACTAATAAGTTTATC                    |               | <i>E.c. glmY</i> -185 to -141   |
| BG685  | [P]-GTAGATTTAGTCATCAGCGCTCTCGGGATGGATGAAATG                        |               | <i>E.c. glrR</i> +148 to +186   |
| BG686  | [P]-GTAGATTTAGTCATCAGCGAACTCGGGATGGATGAAATG                        |               | <i>E.c. glrR</i> +148 to +186   |
| BG696  | CTCGTACTCATATGACACCACGCAAACC                                       | NdeI          | <i>Y.p. glrR</i> +1 to +17      |
| BG697  | CGTCTCTAGACTCTTTAAATCGTTGGCATCC                                    | Xbal          | <i>Y.p. glrR</i> +1335 to +1314 |
| BG698  | GCACGCGTCGACTTTTTATATTCTGTGCGCAAG                                  | Sall          | <i>Y.p. glmY</i> -257 to -236   |
| BG699  | CGTCTCTAGACATAAAAAGGTGAATGAGCAAC                                   | Xbal          | <i>Y.p. glmY</i> +22 to +1      |
| BG700  | GCACGCGTCGACTTCGTTGTGTTGGGCGTCAG                                   | Sall          | <i>Y.p. glmZ</i> -303 to -284   |
| BG701  | CGTCTCTAGAAATAAGTGGGATGAGCATCTAC                                   | Xbal          | <i>Y.p. glmZ</i> +22 to +1      |
| BG727  | GCGAGCTCAAGGAATCTCATGACACCACG                                      | SacI          | <i>Y.p. glrR</i> -10 to +11     |
| BG728  | GGCTCTAGATTACTCTTTAAATCGTTGGCATC                                   | Xbal          | <i>Y.p. glrR</i> +1338 to +1315 |
| BG747  | GCACGCGTCGACGGGCGTCAGACATGGTTTTCCACGACAATAAACG                     | Sall          | <i>Y.p. glmZ</i> -292 to -259   |
| BG748  | [P]-TGTCACCTTCTCACGTGTATGTGATCGTTT                                 |               | <i>Y.p. glmZ</i> -234 to -263   |
| BG750  | GCACGCGTCGACCAAGATTAAGTGTGCGGAAATCC                                | Sall          | <i>S.t. glmY</i> -242 to -219   |
| BG751  | CGTCTCTAGACATAAAGAGGTGAATGAGCCAC                                   | Xbal          | <i>S.t. glmY</i> +22 to +1      |
| BG752  | GCACGCGTCGACGTGTTGCCATTATGATTTGTTGG                                | Sall          | <i>S.t. glmZ</i> -242 to -219   |
| BG753  | CGTCTCTAGATAAGAGATGGAATGAGCATCTAC                                  | Xbal          | <i>S.t. glmZ</i> +22 to +1      |
| BG754  | [P]-CAATGTAGGGTTATAACAAGTTTTGTAGCGACAG                             |               | <i>Y.p. glmZ</i> -204 to -170   |
| BG755  | GTTCACTCTGGTCACCGGG  |               | <i>Y.p. glmZ</i> -170 to -152   |

<sup>a</sup>Restriction sites are underlined; Nucleotide positions that differ from the wild-type sequence are in boldface; [P] indicates 5'-phosphorylation of the oligonucleotide. <sup>b</sup>Positions are relative to the first nucleotide of the respective gene. Gene names are according to <http://ecocyc.org/>

The novel RNA binding protein RapZ (YhbJ) acts as a modulator affecting  $\sigma^{54}$ -dependent expression of sRNA GlmY independently of its function as an RNase adaptor protein.

Göpel, Y., Kögler, A. C., and Görke, B.

*The results presented in this Chapter are unpublished.*

*Author contributions:*

This study was designed by Y.G. and B.G.. A.C.K. and Y.G. performed  $\beta$ -galactosidase activity measurements, construction of strains and plasmids with assistance by Sabine Lentjes or intern students under the supervision of Y.G. Plasmid pBGG225 was constructed by Birte Reichenbach. Western Blot experiments, protein purification and *in vitro* acetylation assays were performed by Y.G. B.G. and Y.G. compiled and analyzed sequence alignments. Y.G. wrote the manuscript, B.G. offered insightful discussion of the manuscript.

## ABSTRACT

In *Escherichia coli*, expression of key enzyme glucosamine-6-phosphate synthase (GlmS) is feedback-regulated by two homologous sRNAs, GlmY and GlmZ that function in an hierarchical cascade to activate the *glmS* transcript. Whereas sRNA GlmZ exerts its regulatory effect on *glmS* expression by direct base-pairing within the 5' un-translated region of the *glmS* mRNA, GlmY functions indirectly by stabilizing the functional form of sRNA GlmZ in a process that involves protein RapZ. Since this sRNA cascade provides a feedback mechanism to attune *glmS* expression to the intracellular amino sugar level, *glmY* expression must be tightly controlled. Transcription of *glmY* is driven by two perfectly overlapping promoters, a  $\sigma^{70}$ -dependent and a  $\sigma^{54}$ -dependent promoter. Whereas the  $\sigma^{70}$ -dependent *glmY*-promoter seems to be unregulated, the  $\sigma^{54}$ -dependent promoter is regulated by multiple factors including two-component system GlrK/GlrR and IHF. Here, we present another layer of regulation for the  $\sigma^{54}$ -dependent *glmY* promoter. We identified protein RapZ, which functions as mediator of signal transduction within the hierarchical GlmYZ-cascade, as indispensable for  $\sigma^{54}$ -driven *glmY* transcription. By a so far unknown mechanism, RapZ activates the *glmY* promoter in a process that presumably requires acetylation of a lysine residue by acetyl-phosphate generated by the AckA-Pta pathway. Accordingly, over-expression of deacetylase CobB abrogates  $\sigma^{54}$ -dependent transcription of *glmY*, presumably by deacetylation of RapZ. We therefore propose a second regulatory function for RapZ that seems to be independent of its role independent from its function as adaptor protein in the GlmY/Z cascade.

## INTRODUCTION

In recent years, bacterial small regulatory RNAs (sRNAs) emerged as major regulators of gene expression in various physiological processes within the bacterial cell. One predominant role of sRNAs is the fast response to changing environmental conditions, metabolite concentrations or adaptation to different stresses to ensure cell survival and prosperity (Gottesman and Storz, 2011; Richards and Vanderpool, 2011; Storz *et al.*, 2011). Thus, regulatory sRNAs must be strictly controlled, either at the level of degradation and/or at the transcriptional level. To date it is apparent that expression of many sRNA genes is extensively regulated at the level of transcription initiation and a number of regulators, such as alternative sigma factors, response regulators and transcriptional activators are known for several sRNA genes (Chao *et al.*, 2012; Göpel *et al.*, 2011; Göpel and Görke, 2012a). The fact that global regulators, such as two-component systems and alternative sigma factors, employ small RNAs in their regulons allows to fine-tune or to time the expression of a subset of genes. Further, such regulatory networks provide the possibility to integrate multiple stimuli into an accurately adjusted response (Beisel and Storz, 2010; Göpel and Görke, 2012a; Lee and Groisman, 2010). The use of alternative sigma factors enables the cell to

express or repress a large set of genes in response to specific stimuli, e.g. heat shock or envelope stress. In addition to the housekeeping sigma factor,  $\sigma^{70}$ , *Escherichia coli* possesses six alternative sigma factors (Reitzer and Schneider, 2001). Of these alternative sigma factors five are homologous to  $\sigma^{70}$  and DNA sequences that deviate from the  $\sigma^{70}$  consensus. In contrast, the unrelated  $\sigma^{54}$  (encoded by the *rpoN* gene) is a rather unusual sigma factor: it recognizes -12 and -24 consensus motifs and depends on specific activator proteins to form an open complex and initiate transcription (Reitzer and Schneider, 2001; Wigneshweraraj *et al.*, 2008). So far, it is unknown how the activity of  $\sigma^{54}$  itself is regulated. In *E. coli*,  $\sigma^{54}$  is encoded within the (*lptB*-) *rpoN-hpf-ptsN-rapZ-ptsO* operon, which encodes pleiotropic regulators. Enzyme IIA<sup>Ntr</sup>, encoded by *ptsN*, and its cognate phosphotransferase protein Npr, encoded by *ptsO*, form a regulatory phospho-relay system together with the Enzyme I<sup>Ntr</sup> component, encoded by *ptsP* (Rabus *et al.*, 1999; Zimmer *et al.*, 2008). This regulatory phosphotransferase system (PTS<sup>Ntr</sup>) modulates potassium transport and phosphate homeostasis by protein-protein interaction with sensor kinases KdpD and PhoR, respectively (Lüttmann *et al.*, 2009; Lüttmann *et al.*, 2012). The Hpf protein has been reported to regulate storage of ribosomes in stationary phase and thereby regulate translation (Ueta *et al.*, 2005; Ueta *et al.*, 2008). RNase adaptor protein RapZ was shown to mediate feedback regulation of *glmS* expression, in the hierarchically acting GlmY/Z sRNA cascade (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008). Gene *glmS* encodes for glucosamine-6-phosphate (GlcN6P) synthase GlmS, which catalyzes the key step in the cell wall biosynthesis pathway: synthesis of the essential precursor GlcN6P. While sRNA GlmZ base-pairs with the *glmS* 5' UTR, sRNA GlmY acts indirectly by stabilizing GlmZ in its active form in a process that requires sequestration of the adaptor RapZ (Göpel *et al.*, 2013).

While GlmZ is constitutively expressed and only regulated at the level of decay, sRNA GlmY is elaborately regulated at the level of transcription (Göpel *et al.*, 2011; Göpel *et al.*, 2013; Reichenbach *et al.*, 2009). Interestingly, in *E. coli*, *glmY* expression is driven by two perfectly overlapping promoters, a seemingly unregulated  $\sigma^{70}$ -dependent promoter that is active mainly during exponential growth, and a  $\sigma^{54}$ -dependent promoter (Reichenbach *et al.*, 2009). Upon entry of stationary phase, the  $\sigma^{54}$ -dependent *glmY* promoter is strongly activated by its cognate two-component system GlrK/GlrR that is encoded directly downstream of the *glmY* gene. After signal perception sensor kinase GlrK auto-phosphorylates and subsequently transfers the phosphoryl-group to aspartyl-residue 56 in response regulator GlrR. In turn, phosphorylated GlrR binds to three conserved sequence motifs far upstream of the *glmY* promoter and activates transcription together with integration host factor IHF (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011). However, the signal that is sensed by histidine kinase GlrK remains unknown. Here, we found that a third component YfhG that is encoded within the *glrK-yfhG-glrR* operon is involved in  $\sigma^{54}$ -driven *glmY* expression presumably by acting as an activator of sensor kinase GlrK.

In order to better understand the complex network of regulators controlling *glmY* expression, we investigated newly identified putative modulators of *glmY* promoter activity in more detail. In previous studies, we observed a reduction of transcription rate from the  $\sigma^{54}$ -dependent *glmY* promoter upon deletion of *rapZ* (Reichenbach, 2009). Here, we report that RapZ is indeed a regulator of  $\sigma^{54}$ -dependent *glmY* promoter activity and that this effect most likely depends on acetylation of a C-terminal lysine residue possibly with acetyl-phosphate. Further, we identified deacetylase CobB as an important regulator that acts upstream of RapZ and is likely to catalyze removal of the acetyl-group. Since the function of RapZ in *glmY* promoter control does not require RNA-binding activity and can be adopted by the *Bacillus subtilis* RapZ homologue YvcJ, we propose this to be a second function for RapZ that is independent of its role in signal transduction within the GlmY/GlmZ cascade.

## MATERIALS AND METHODS

### *Growth conditions and construction of strains*

LB was used as standard complex medium for cultivation of *E. coli* alternatively, M9 medium containing 1% glycerol supplemented with L-proline (40  $\mu\text{g/ml}$ ) and thiamine (1  $\mu\text{g/ml}$ ) or MOPS-medium (Neidhardt *et al.*, 1974) with glutamine or ammonium served as defined minimal media. Cells were grown routinely under agitation (200 r.p.m.) at 37°C. When required for selection, antibiotics were added to the medium in the following concentrations: ampicillin 100  $\mu\text{g/ml}$ , kanamycin 30  $\mu\text{g/ml}$ , chloramphenicol 15  $\mu\text{g/ml}$ , spectinomycin 75  $\mu\text{g/ml}$  and 1 mM IPTG was used to induce expression of genes placed under the control of the  $P_{tac}$  promoter. For induction of *zraP'*-*lacZ* expression 1 mM  $\text{ZnCl}_2$  was added to the medium, and induction of *glnA'*-*lacZ* was achieved by replacing  $\text{NH}_4$  with glutamine as nitrogen source in MOPS minimal medium.

*E. coli* strains used in this study and their respective genotypes are shown in Table 3.1. Previously established alleles tagged with an antibiotic resistance marker were moved between strains by general transduction using phage T4GT7 (Wilson *et al.*, 1979). Reporter gene fusions with *lacZ* were established on plasmids (see Tables 3.1 and 3.2) subsequently integrated into the phage  $\lambda attB$ -site on the chromosome of strains R1279 (CSH50 wild type), Z37 (CSH50,  $\Delta rapZ$ ), R2413 (CSH50,  $\Delta ptsN-O$ ), S162 (CSH50, *phoU*<sup>+</sup>, wild type), Z845 (CSH50, *phoU*<sup>+</sup>,  $\Delta rapZ$ ), and S4197 (MG1655, *ilvG*<sup>+</sup>, wild type) by site-specific recombination using helper plasmid pLDR8 as previously described (Diederich *et al.*, 1992). Briefly, origin-less DNA-fragments encompassing the respective *lacZ* fusion alleles, the *aadA* spectinomycin resistance gene and the  $\lambda attP$ -site were isolated by *Bam*HI digest. DNA-fragments were self-ligated and subsequently introduced into the respective strains carrying the temperature-sensitive  $\lambda$ -integrase expression plasmid pLDR8. Recombinants were obtained upon selection on spectinomycin-plates at 42°C. FLAG-tagging of *rpoN* at its authentic locus was performed by epitope tagging using a PCR-fragment obtained with primers BG1089/1090 from template plasmid pSUB11 as described before (Uzzau *et al.*, 2001). Deletion of *yfhG* was performed by the marker-less deletion method as previously described using a PCR-fragment obtained with primers BG767/BG768 (Datsenko and Wanner, 2000). Antibiotic resistance genes encompassed by FLP recognition sites were removed using the temperature-sensitive FLP recombinase expression plasmid pCP20 as described previously (Datsenko and Wanner, 2000). All strain constructions were verified by diagnostic PCRs and/or sequencing.

### **Construction of plasmids**

DNA-cloning was performed in *E. coli* strain DH5 $\alpha$  following standard procedures. All plasmids and oligonucleotides used in this study are listed in Tables 3.2 and 3.3, respectively.

Plasmid pDL43 encoding a *glnA'* (-372 to + 50)-*lacZ* fusion was constructed by PCR-amplification of *glnA'* using primers BG934/BG935 and *E. coli* w3110 chromosomal DNA. The fragment was ligated between the *SacI/XbaI*-sites of plasmid pKES15 thereby replacing the *bgI'*-fragment. Plasmid pBGG225 was constructed by ligation of the PCR-derived fragment encompassing the *yfhG* gene (BG494/BG495) into pBAD18-cm via *SacI/XbaI*. For construction of plasmid pYG38, *cobB* was amplified from *E. coli* chromosomal DNA using primers BG940/BG941 and cloned into the *EcoRI/BamHI* sites in pKESK22. Plasmid pYG82 encoding RapZ with K270A-K281A-R282A-K283A mutations in its 3' end was constructed by sub-cloning of the *rapZ* 3' *BamHI-XbaI* fragment from pYG29 into pFDX4324, which replaced the wild type 3' end with the mutated fragment. For construction of pYG89, encoding *glrR*, and pYG90, encoding *glrR*<sub>D56E</sub> under *P<sub>tac</sub>* promoter control, the respective alleles were separated from plasmids pBGG389 and pBGG399 via partial *EcoRI/XbaI* digestion. The resulting fragments were inserted into pKESK23 between its *EcoRI/XbaI* sites. Plasmids pKESK22 (Stratmann *et al.*, 2008) and pKESK23 were a gift from Karin Schnetz (University of Cologne, Germany). Plasmid constructions were verified by analytical digestion and DNA sequencing.

### **Analysis of *glmY* and *glmZ* transcription ( $\beta$ -galactosidase assays)**

Overnight cultures of *E. coli* were inoculated into fresh LB medium containing supplements as required to an OD<sub>600</sub> of 0.1 and grown to an OD<sub>600</sub> of 0.5-0.8 for exponential growth phase, OD<sub>600</sub> of 1.7- 2.0 for stationary growth phase or over the course of 10 hours. Subsequently, the cells were harvested and the  $\beta$ -galactosidase activities were determined as previously described (Miller, 1972).

### **Western blotting and dot-blot far-Western analysis**

Western blotting analysis were carried out as described previously (Lüttmann *et al.*, 2012). Polyclonal rabbit antisera were used at dilutions of 1:10000 (anti-FLAG, Sigma-Aldrich), 1:10000 (anti-GlmS), 1:5000 (anti-Hfq) and 1:1000 (anti-acetolysine, abcam). The antibodies were visualized with alkaline phosphatase conjugated goat anti-rabbit IgG secondary antibodies (Promega), diluted 1:100000, and the CDP\* detection system (Roche Diagnostics, Germany).

### **Protein purification**

Strep-tagged recombinant proteins were purified as described previously (Lüttmann *et al.*, 2012) Plasmids pBGG164 and pBGG217 were used for overproduction of Strep-RapZ and Strep-PtsN in strains Z741, Z762, Z762/pYG38 or DH5 $\alpha$ , respectively. The transformants were grown in 500 ml LB medium to an OD<sub>600</sub> = 0.5 to 0.8. Synthesis of the recombinant proteins was induced by addition of 1 mM IPTG and growth was subsequently continued for 1 h. Cells were harvested and washed in buffer W (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0) and lysed using a French pressure cell at 18.000 psi. Lysates were cleared by low speed centrifugation (10 min, 4°C, 8000 r.p.m.) followed by ultracentrifugation (1 h, 6°C, 35000 r.p.m.) and loaded onto columns containing 0.5 ml Strep-Tactin matrix (IBA, Germany) pre-equilibrated with buffer W. Columns were washed four times with 5 ml buffer W and Strep-tagged proteins were eluted in three steps using 3 $\times$ 0.5 ml buffer W with 2.5 mM desthiobiotin. Elution fractions were dialyzed two times for 16 h against dialysis buffer

(10 mM Tris-HCl, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 2 mM DTT, pH 7.0). In the second dialysis step, 25% (v/v) glycerol was added to the buffer. The purified proteins were stored at -20°C until further use.

### In vitro acetylation assay

Strep-RapZ was purified from cells carrying the *strep-rapZ* over-expression vector pBGG164 and the *cobB* over-expression plasmid pYG38 to ensure that RapZ was not completely acetylated after purification. In vitro acetylation assays were performed using 10 µg Strep-RapZ protein in 1x structure buffer [10 mM Tris pH 7, 100 mM KCl, 10 mM MgCl<sub>2</sub>] containing 20 mM nicotinamide and 50 mM acetyl-phosphate in a total volume of 30 µl. As a control acetyl-phosphate was omitted and 10 µg Strep-RapZ was incubated with buffer containing 20 mM nicotinamide to inhibit residual CobB activity. Reactions were incubated at 30°C for 2 hours. 5µl samples were taken after 5, 10, 30, 60 and 120 minutes and 5 µl 2x Laemmli buffer [125 mM Tris/HCl pH 6.8, 10% (v/v) β-mercaptoethanol, 4% (w/v) SDS, 20% (v/v) glycerol, 0.1% (w/v) bromophenol-blue] was added. Samples were loaded onto a 12.5% SDS gel following heat denaturing at 65°C for 5 minutes. SDS gels were electroblotted onto a PVDF membrane (BioRad) for 1 hour at 0.8 mA cm<sup>-2</sup>. For detection, a polyclonal rabbit antiserum directed against acetolysine (abcam) was used in a 1:1000 dilution following western blot procedures.

**Table 3.1.** *E. coli* strains used in this study

| Name         | Genotype  | Reference or construction   |
|--------------|---|---|
| AJW1939      | MC4100 $\Delta(\text{ackA-pta})::\text{cm}$   | (Wolfe <i>et al.</i> , 2008)  |
| DH5 $\alpha$ | $\phi$ 80d <i>lacZ</i> $\Delta$ M15, <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> ( $r_{K^-}$ , $m_K^+$ ), <i>supE44</i> , <i>relA1</i> , <i>deoR</i> , $\Delta(\text{lacZYA-argF})$ U169 | Laboratory collection   |
| JW1106       | F <sup>-</sup> , $\Delta(\text{araD-araB})567$ , $\Delta(\text{lacZ4787}::\text{rrmB-3})$ , $\lambda$ , $\Delta(\text{cobB779}::\text{kan}$ , <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$ , <i>hsdR514</i>             | (Baba <i>et al.</i> , 2006)   |
| JW2568       | F <sup>-</sup> , $\Delta(\text{araD-araB})567$ , $\Delta(\text{lacZ4787}::\text{rrmB-3})$ , $\lambda$ , $\Delta(\text{yfiQ752}::\text{kan}$ , <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$ , <i>hsdR514</i>             | (Baba <i>et al.</i> , 2006)   |
| JW4130       | $\Delta(\text{araD-araB})567$ , $\Delta(\text{lacZ4787}::\text{rrmB-3})$ , <i>LAM-</i> , <i>rph-1</i> , $\Delta(\text{rhaD-rhaB})568$ , $\Delta(\text{hfq-722}::\text{kan}$ , <i>hsdR514</i>                            | (Baba <i>et al.</i> , 2006)   |
| MG1655       | F <sup>-</sup> $\lambda$ <i>ilvG</i> <i>rfb</i> <i>rph</i>  | (Blattner <i>et al.</i> , 1997)   |
| R1279        | CSH50 $\Delta(\text{pho-bgl})201$ $\Delta(\text{lac-pro})$ <i>ara thi</i>   | (Schnetz <i>et al.</i> , 1996)  |
| R2413        | As R1279, but $\Delta(\text{ptsN-ptsO})$  | (Kalamorz <i>et al.</i> , 2007)   |
| S162         | CSH50 $\Delta(\text{bgl})$ $\Delta(\text{lac-pro})$ <i>ara thi</i>  | (Caramel and Schnetz, 1998)   |
| S4197        | As MG1655, but <i>rph</i> <sup>+</sup> <i>ilvG</i> <sup>+</sup> $\Delta(\text{lacZ-pFDY217})$   | (Venkatesh <i>et al.</i> , 2010)  |
| Z24          | As R1279, $\Delta(\text{rapZ}::\text{cm})$  | (Kalamorz <i>et al.</i> , 2007)   |
| Z37          | As R1279, $\Delta(\text{rapZ})$   | (Kalamorz <i>et al.</i> , 2007)   |
| Z190         | As R1279, <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> , -10 region mutated] <i>aadA</i>   | (Reichenbach <i>et al.</i> , 2009)  |
| Z197         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i>  | (Reichenbach <i>et al.</i> , 2009)  |
| Z201         | As R1279, <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> , -24 region mutated] <i>aadA</i>   | (Reichenbach <i>et al.</i> , 2009)  |
| Z206         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{glrR})$  | (Reichenbach <i>et al.</i> , 2009)  |
| Z225         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{rapZ})$  | (Reichenbach, 2009)   |
| Z274         | As R1279, but <i>attB</i> ::[ <i>zraP</i> <sup>-</sup> (-289 to +41)- <i>lacZ</i> ], <i>aadA</i>  | (Reichenbach <i>et al.</i> , 2009)  |
| Z275         | As R1279, but <i>attB</i> ::[ <i>zraP</i> <sup>-</sup> (-289 to +41)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{rapZ})$  | This work, pBGG324/BamHI→Z37  |
| Z449         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> , -10 region mutated], <i>aadA</i> $\Delta(\text{yfhG})$  | This work, $\Delta(\text{yfhG}::\text{cm})$ via primers BG767/768 from pKD3 into Z190, cured from <i>cm</i> |
| Z477         | As R1279, <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{yfhG})$  | This work, $\Delta(\text{yfhG}::\text{cm})$ via primers BG767/768 from pKD3 into Z197, cured from <i>cm</i> |
| Z492         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> , -24 region mutated], <i>aadA</i> $\Delta(\text{yfhG})$  | This work, $\Delta(\text{yfhG}::\text{cm})$ via primers BG767/768 from pKD3 into Z201, cured from <i>cm</i> |
| Z716         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{ptsN-ptsO})$   | This work, pBGG201/BamHI→R2413  |
| Z741         | As S4197, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i>  | This work, pBGG201/BamHI→S4197  |
| Z743         | As S4197, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> , -10 region mutated] <i>aadA</i>   | This work, pBGG209/BamHI→S4197  |
| Z747         | S4197, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{cobB}::\text{kan})$   | This work, Z741 $\Delta(\text{cobB}::\text{kan})$ transduced from JW1106 to Z741                            |
| Z762         | S4197, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{cobB})$   | This work, Z747 cured from <i>kan</i>   |
| Z780         | As R1279, but <i>attB</i> ::[ <i>zraP</i> <sup>-</sup> (-289 to +41)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{ptsN-O})$  | This work, pBGG324/BamHI→R2413  |
| Z787         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{ackA-pta})::\text{cm}$   | This work, $\Delta(\text{ackA-pta})::\text{cm}$ transduced from AJW1939 to Z197                             |
| Z788         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{rapZ})$ , $\Delta(\text{ackA-pta})::\text{cm}$   | This work, $\Delta(\text{ackA-pta})::\text{cm}$ transduced from AJW1939 to Z225                             |
| Z791         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{ackA-pta})$  | This work, Z787 cured from <i>cm</i>  |
| Z792         | As R1279, but <i>attB</i> ::[ <i>glmY</i> <sup>-</sup> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta(\text{rapZ})$ , $\Delta(\text{ackA-pta})$  | This work, Z788 cured from <i>cm</i>  |

| Name | Genotype   | Reference or construction  |
|------|--|--|
| Z806 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , <i>rpoN</i> -3xFLAG- <i>kan</i>  | This work, 3x-FLAG via primers BG1089/1090 from pSUB11 into Z197             |
| Z807 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , <i>rpoN</i> -3xFLAG  | This work, Z806 cured from <i>kan</i>  |
| Z808 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ ( <i>ptsN</i> - <i>ptsO</i> ) <i>rpoN</i> -3xFLAG- <i>kan</i> | This work, 3x-FLAG via primers BG1089/1090 from pSUB11 into Z716             |
| Z809 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ ( <i>ptsN</i> - <i>ptsO</i> ) <i>rpoN</i> -3xFLAG             | This work, Z808 cured from <i>kan</i>  |
| Z820 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>hfg</i> :: <i>kan</i>                                      | This work, $\Delta$ <i>hfg</i> :: <i>kan</i> transduced from JW4130 to Z197  |
| Z821 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i> , $\Delta$ <i>hfg</i> :: <i>kan</i>               | This work, $\Delta$ <i>hfg</i> :: <i>kan</i> transduced from JW4130 to Z225  |
| Z824 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>hfg</i>  | This work, Z820 cured from <i>kan</i>  |
| Z825 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i> , $\Delta$ <i>hfg</i>                             | This work, Z821 cured from <i>kan</i>  |
| Z830 | S4197, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i> :: <i>cm</i>   | This work, $\Delta$ <i>rapZ</i> :: <i>cm</i> transduced from Z24 to Z741     |
| Z833 | S4197, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i>  | This work, Z830 cured from <i>cm</i>   |
| Z841 | S4197, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> , -10 region mutated], <i>aadA</i> , $\Delta$ <i>rapZ</i> :: <i>cm</i>                     | This work, $\Delta$ <i>rapZ</i> :: <i>cm</i> transduced from JW1106 to Z197  |
| Z842 | As S162, but $\Delta$ <i>rapZ</i> :: <i>cm</i>   | This work, $\Delta$ <i>rapZ</i> :: <i>cm</i> transduced from Z24 to S162     |
| Z843 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>cobB</i> :: <i>kan</i>                                     | This work, $\Delta$ <i>cobB</i> :: <i>kan</i> transduced from Z24 to Z743    |
| Z844 | S4197, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> , -10 region mutated], <i>aadA</i> , $\Delta$ <i>rapZ</i>                                  | This work, Z841 cured from <i>cm</i>   |
| Z845 | As S162, but $\Delta$ <i>rapZ</i>  | This work, Z842 cured from <i>cm</i>   |
| Z846 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>cobB</i>   | This work, Z843 cured from <i>kan</i>  |
| Z847 | As S162, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i>  | This work, pBGG201/BamHI→Z845  |
| Z849 | As R1279, but <i>attB</i> ::[ <i>glnA</i> '(-372 to +50)- <i>lacZ</i> ], <i>aadA</i>   | This work, pDL43/BamHI→R1279   |
| Z850 | As R1279, but <i>attB</i> ::[ <i>glnA</i> '(-372 to +50)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>rapZ</i>  | This work, pDL43/BamHI→Z37   |
| Z851 | As S162, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i>   | This work, pBGG201/BamHI→S162  |
| Z852 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>cobB</i> , $\Delta$ <i>rapZ</i> :: <i>cm</i>               | This work, $\Delta$ <i>rapZ</i> :: <i>cm</i> transduced from Z24 to Z846     |
| Z853 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>cobB</i> , $\Delta$ <i>rapZ</i>                            | This work, Z852 cured from <i>cm</i>   |
| Z866 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>yfiQ</i> :: <i>kan</i>                                     | This work, $\Delta$ <i>yfiQ</i> :: <i>kan</i> transduced from JW2568 to Z197 |
| Z867 | As R1279, but <i>attB</i> ::[ <i>glmY</i> (-238 to +22)- <i>lacZ</i> ], <i>aadA</i> , $\Delta$ <i>yfiQ</i>   | This work, Z866 cured from <i>kan</i>  |

**Table 3.2.** Plasmids used in this study

| Name        | Relevant structure  | Reference or construction          |
|-------------|---|------------------------------------|
| pBAD18-cm   | <i>P</i> <sub>Ara</sub> , MCS 2, <i>cat</i> , ori pBR322  | (Guzman <i>et al.</i> , 1995)      |
| pBGG164     | <i>strep-rapZ</i> under <i>P</i> <sub>tac</sub> control, <i>bla</i> ,   |                                    |
| pBGG201     | Fusion of <i>glmY</i> '(-238 to +22) <sup>a</sup> to <i>lacZ</i>  | (Reichenbach <i>et al.</i> , 2009) |
| pBGG209     | Fusion of <i>glmY</i> '(-238 to +22) to <i>lacZ</i> , -10 region mutated  | (Reichenbach <i>et al.</i> , 2009) |
| pBGG225     | <i>yfhG</i> under <i>P</i> <sub>Ara</sub> -control in pBAD18-cm   | This work                          |
| pBGG324     | Fusion of <i>zraP</i> '(-289 to +41) to <i>lacZ</i>   | (Reichenbach <i>et al.</i> , 2009) |
| pBGG389     | <i>glrR</i> under <i>P</i> <sub>Ara</sub> -control  | (Göpel <i>et al.</i> , 2011)       |
| pBGG399     | as pBGG389, but <i>glrR</i> with D56E mutation  | (Göpel <i>et al.</i> , 2011)       |
| pBGM53      | encodes for <i>yvcJ</i> under <i>P</i> <sub>Spac</sub> -control in pDG148-Stul  | (Görke <i>et al.</i> , 2005)       |
| pDG148-Stul | <i>E. coli</i> - <i>B. subtilis</i> shuttle vector, pUB110 origin, <i>kan</i> , <i>ble</i> , <i>bla</i> , <i>P</i> <sub>spac</sub> promoter, Stul site, <i>lacI</i> | (Joseph <i>et al.</i> , 2001)      |
| pDL43       | Fusion of <i>glnA</i> '(-372 to +50) to <i>lacZ</i>   | (Lüttmann, 2011)                   |
| pFDX4291    | pSC101-ori, <i>cat</i> , operator-less <i>P</i> <sub>tac</sub> , BglII, <i>sacB</i> -RBS, NdeI, XbaI, HincII  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4292    | encodes for <i>ptsO</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4294    | encodes for <i>ptsN</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4296    | encodes for <i>ptsN</i> , <i>rapZ</i> and <i>ptsO</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4320    | encodes for <i>rapZ</i> and <i>ptsO</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4322    | encodes for <i>ptsN</i> and <i>ptsO</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pFDX4324    | encodes for <i>rapZ</i> under <i>P</i> <sub>tac</sub> -control in pFDX4291  | (Kalamorz <i>et al.</i> 2007)      |
| pKES15      | <i>bgI</i> '- <i>lacZ</i> , <i>kan</i> , <i>attP</i> , <i>aadA</i> , ori pACYC177   | (Nagarajavel <i>et al.</i> , 2007) |
| pKESK22     | ori p15A, <i>neo</i> , <i>lacI</i> <sup>q</sup> , <i>P</i> <sub>tac</sub> , MCS   | (Stratmann <i>et al.</i> , 2008)   |
| pKESK23     | ori p15A, <i>neo</i> , <i>attP</i> , <i>aadA</i> , <i>T4gene32</i> -terminator, <i>lacI</i> <sup>f</sup> ; <i>P</i> <sub>tac</sub> , MCS.                           | Schnetz unpublished                |
| pYG29       | encodes for <i>strep-rapZ</i> under <i>P</i> <sub>tac</sub> control, but <i>rapZ</i> with K270A-K281A-R282A-K283A mutations   | (Göpel <i>et al.</i> , 2013)       |
| pYG38       | encodes for <i>cobB</i> under <i>P</i> <sub>tac</sub> -control in pKESK22   | This work                          |
| pYG82       | as pFDX4324, but <i>rapZ</i> with K270A-K281A-R282A-K283A mutations   | This work                          |
| pYG89       | encodes for <i>glrR</i> under <i>P</i> <sub>tac</sub> -control in pKESK23   | This work                          |
| pYG90       | as pYG89, but <i>glrR</i> with D56E mutation  | This work                          |

Ori, origin of replication, MCS, multiple cloning site, <sup>a</sup> Positions given are relative to the first nucleotide of the respective gene. Gene names according to <http://ecocyc.org/>

**Table 3.3.** Oligonucleotides used in this study

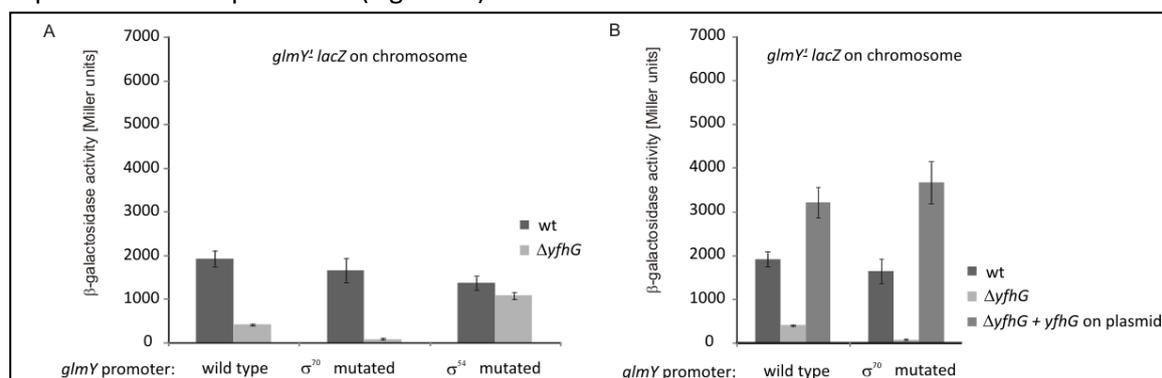
| Primer | Sequence <sup>a</sup>   | Res.Sites    | Position <sup>b</sup>      |
|--------|---|--------------|----------------------------|
| BG494  | <u>GCGAGCTCGCCCACTGGGATAGGCTTTAAG</u>   | <u>SacI</u>  | <i>yfhG</i> -32 to -11     |
| BG495  | <u>GGCTCTAGATTATGGCTCATCAGGAGTGACC</u>  | <u>XbaI</u>  | <i>yfhG</i> +714 to +693   |
| BG767  | AACAATTACGCCCACTGGGATAGGCTTTAAGTCTGGTGAATATG <i>TGTAGGC</i><br><i>TGGAGCTGCTTCG</i>     |              | <i>yfhG</i> -42 to +3      |
| BG768  | TCCCGATCGTCATCGACCAATAATAAATGCGCAGGTTTATGGCTCATATGA<br><i>ATATCCTCCTTAGTTCTATTCC</i>    |              | <i>glrR</i> +1 to +48      |
| BG934  | GCACGCTCGACAACCTTTCCTCAGGCATTAG   | <u>SalI</u>  | <i>glnA</i> -372 to -353   |
| BG935  | <u>GGCTCTAGAACTTCACTTCGTGCTCG</u>   | <u>XbaI</u>  | <i>glnA</i> +50 to +33     |
| BG940  | <u>CCGGAATTCCTAACCGATTAACAACAGAGG</u>   | <u>EcoRI</u> | <i>cobB</i> -26 to -6      |
| BG941  | <u>GCGGGATCCTCAGGCAATGCTTCCCGCTT</u>  | <u>BamHI</u> | <i>cobB</i> +843 to +823   |
| BG1089 | GAGAGTCTTTATCCATTCCGCCGTCAAACCAGCGTAAACAACCTCGTTGACTA<br><i>CAAAGACCATGACGG</i>         |              | <i>rpoN</i> +1385 to +1431 |
| BG1090 | CGCAGTGCCTCGGTGATCTCGACGTTATTTCCGGTAATGTTGAGCTGCATA<br><i>TGAATATCCTCCTTAGTTCTATTCC</i> |              | <i>hpf</i> +1 to +50       |

<sup>a</sup>Restriction sites are underlined, italic sequences are complementary to template pKD3 for deletions and pSUB11 for FLAG-tag fusions; <sup>b</sup>Positions are relative to the first nucleotide of the respective gene.

## RESULTS

### *Outer membrane protein YfhG is a putative activator of the GlrK/R two component system and is required for $\sigma^{54}$ -driven expression of glmY*

Interestingly, the *glrK* operon encodes a third gene, *yfhG*, which is localized in between *glrK* and *glrR*. Since genes that encode proteins with related functions often co-localize, we asked whether YfhG is involved in GlrK/GlrR signal transduction and therefore in regulation of *glmY* expression. To answer this question, we constructed *yfhG* deletion mutant and performed  $\beta$ -galactosidase activity measurements addressing the activity of the *glmY* promoter by means of a *glmY'*-*lacZ* reporter fusion (Fig. 3.1). A previously obtained commercially available *yfhG* mutant could not be complemented by expression of *yfhG* *in trans* but required additional expression of *glrR* (data not shown). Since the *yfhG* and *glrR* genes partially overlap, we constructed a non-polar *yfhG* deletion mutant by keeping the shared sequence intact. In addition to analysis of the deletion mutant (Fig. 3.1 A), we performed complementation experiments (Fig. 3.1 B).



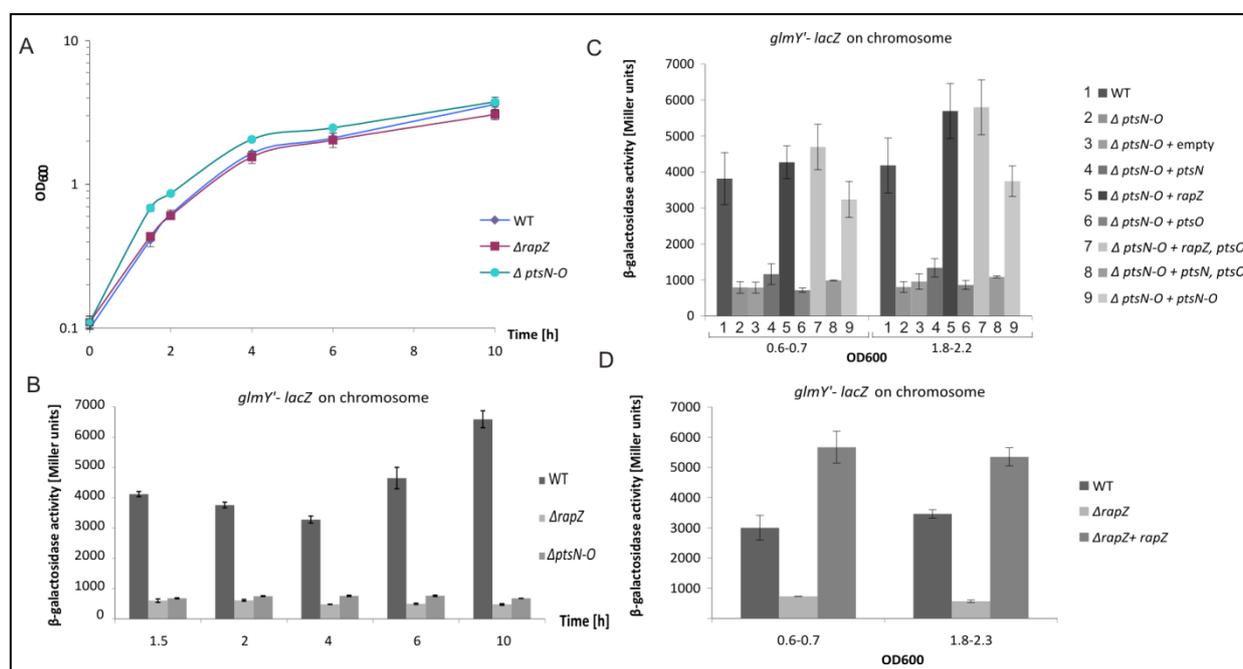
**Figure 3.1: YfhG is an activator of the  $\sigma^{54}$ -dependent *glmY* promoter.** A.  $\beta$ -galactosidase measurements of *glmY'*-*lacZ* expression in strains Z197 (wt), Z477 ( $\Delta yfhG$ ), Z190 (wt,  $\sigma^{70}$ -*glmY* promoter mutated), Z449 ( $\Delta yfhG$ ,  $\sigma^{70}$ -*glmY* promoter mutated), Z201 (wt,  $\sigma^{54}$ -*glmY* promoter mutated) and Z492 ( $\Delta yfhG$ ,  $\sigma^{70}$ -*glmY* promoter mutated). Cells were grown in LB and samples were harvested at OD<sub>600</sub> 0.5-0.8 and  $\beta$ -galactosidase activity was determined. B. Complementation analysis re-introducing *yfhG* on plasmid. Strains Z197, Z477, Z477/pBGG225 (*yfhG*), Z190, Z449 and Z449/pBGG225 (*yfhG*) were grown in LB to an OD<sub>600</sub> 0.5-0.8, samples were harvested and  $\beta$ -galactosidase activity measurements were performed.

Upon deletion of *yfhG*, the activity of the *glmY*-promoter is reduced ~four to five-fold indicating that YfhG indeed is involved in regulation of *glmY* expression (Fig. 3.1 A, first 2 columns). We then used *glmY'*-*lacZ* fusions with either a mutated -10 sequence, so that only the  $\sigma^{54}$ -dependent promoter is active (Fig. 3.1 A, columns 3, 4) or with a mutated -24 signal sequence which leaves only the  $\sigma^{70}$ -dependent promoter active (Fig. 3.1 A, last 2 columns). While expression from the  $\sigma^{70}$ -*glmY* promoter is unaltered in a strain lacking *yfhG*,  $\sigma^{54}$ -driven transcription of *glmY* is almost completely abolished (Fig. 3.1 A, columns 3, 4). Therefore, YfhG is involved in regulation of  $\sigma^{54}$ -dependent expression of *glmY*. Further, complementation of the *yfhG* mutant with a plasmid encoding for YfhG enhanced expression levels of the *glmY'*-*lacZ* reporter gene fusion ~two-fold as compared to wild type expression levels (Fig. 3.1 B, compare columns 1 and 4 to columns 3 and 6). Thus, we propose that outer membrane protein YfhG has a functional connection with the GlrK/GlrR two-component system acting as an activator presumably by directly or indirectly stimulating sensor kinase GlrK.

***RapZ is required for  $\sigma^{54}$ -dependent expression of *glmY* and up-regulation of promoter activity during transition to stationary growth***

Previous Northern blot experiments determined opposite effects of a deletion of the gene encoding RNA-binding protein RapZ on the amounts of sRNAs GlmY and GlmZ in the cell. GlmZ strongly accumulates in its unprocessed form, due to the absence of the RNase E recruiting factor RapZ, whereas GlmY amounts are decreased in cells lacking RapZ (Reichenbach *et al.*, 2008, Reichenbach, 2009; Göpel *et al.*, 2013). In contrast to GlmZ, which is regulated at the level of decay, transcription of *glmY* is extensively controlled (Reichenbach *et al.*, 2009, Göpel *et al.*, 2011). To study the effect of RapZ on GlmY abundance in more detail, we performed  $\beta$ -galactosidase measurements of chromosomally encoded *glmY'*-*lacZ* fusions at different growth stages in the wild type as well as mutant strains deficient in either *rapZ* alone or in mutants carrying a combined *ptsN-rapZ-ptsO* (*ptsN-O*) deletion (Fig. 3.2 B). Gene *ptsN*, encoding the EIIA<sup>Ntr</sup> component of the regulatory nitrogen PTS<sup>Ntr</sup>, is located directly upstream of the *rapZ* gene, the cognate phosphotransferase protein NPr is encoded by *ptsO* and partially overlaps with the 3' end of the *rapZ* gene. To address whether the neighboring genes also affect the activity of the *glmY* promoter we determined expression of *glmY'*-*lacZ* in the combined *ptsN-O* mutant (Fig. 3.2 B). Deletion of the genes encoding EIIA<sup>Ntr</sup>, RapZ and NPr (*ptsN-O*) led to an eight to 12-fold reduction of *glmY* promoter ( $P_{glmY}$ ) activity as compared to the wild type strain when determined for five different time intervals (Fig. 3.2 B). Furthermore, deletion of these genes prevented up-regulation of the  $P_{glmY}$  activity upon transition to stationary growth. This increased activity was shown to be dependent on the  $\sigma^{54}$ -*glmY* promoter and the GlrK/R two-component system that activates  $\sigma^{54}$ - driven *glmY* expression (Reichenbach *et al.*, 2009). The  $\beta$ -galactosidase levels in the  $\Delta ptsN-O$  mutant were similarly low as those determined in the absence of

the  $P_{glmY}$  specific  $\sigma^{54}$ -activator protein GlrR. This reflects the basal activity of the  $\sigma^{70}$ -dependent promoter under conditions where RpoN ( $\sigma^{54}$ ) is actively binding its recognition sequence but is unable to initiate transcription (Reichenbach *et al.*, 2009, Reichenbach 2009, Göpel *et al.*, 2011). Very similar results were obtained, when only *rapZ* was deleted, indicating that the absence of RapZ accounts for the effect observed on the expression of the *glmY-lacZ* fusion. Complementation experiments, re-introducing the missing genes, confirmed that expression of *rapZ* alone is essential and sufficient to restore high  $\beta$ -galactosidase levels reflecting strong  $P_{glmY}$  activity (Fig. 3.2 C column 5, Fig. 3.2 D columns 3 and 6), whereas expression of *ptsN* and *ptsO* were without effect (Fig. 3.2 C columns 4,6 and 8).



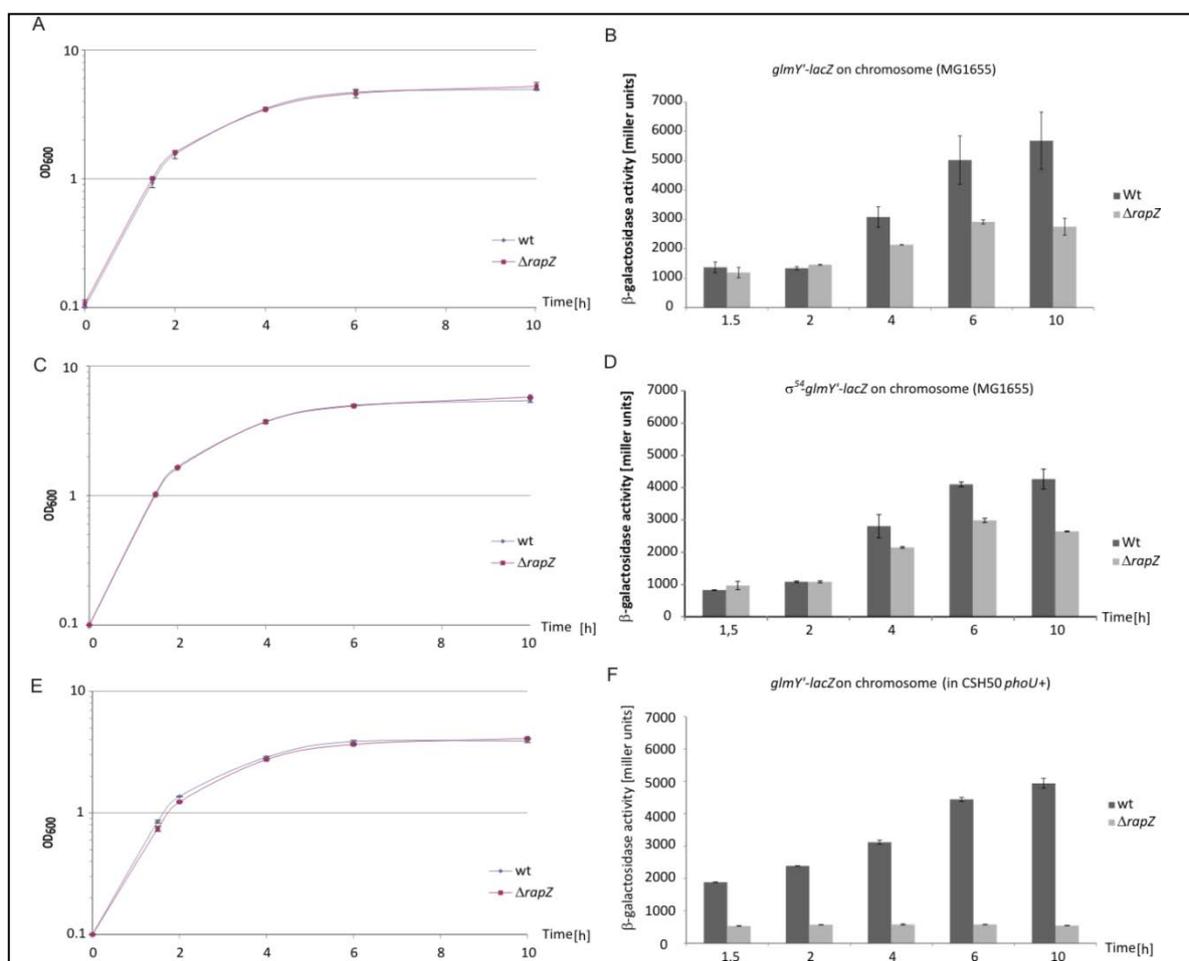
**Figure 3.2: RapZ is required for *glmY* promoter activity and up-regulation in stationary growth phase.** **A.** Strains Z197 (wild type), Z225 ( $\Delta rapZ$ ) and Z716 ( $\Delta ptsN-O$ ) were grown in LB over the course of 10 hours. **B.**  $\beta$ -galactosidase measurements of *glmY-lacZ* expression in strains Z197, Z225 and Z716 over 10 h. Samples were harvested at the indicated time intervals and  $\beta$ -galactosidase activity was determined. **C.** Complementation analysis of the  $\Delta ptsN-O$  mutant using several combinations of the previously deleted genes. Cells were grown in LB to the indicated OD<sub>600</sub> values and  $\beta$ -galactosidase activity was assessed in Z197, Z716, Z716/pFDX4291 (empty), Z716/pFDX4294 (*ptsN*), Z716/pFDX4324 (*rapZ*), Z716/pFDX4292 (*ptsO*), Z716/pFDX4320 (*rapZ, ptsO*), Z716/pFDX4322 (*ptsN, ptsO*) and Z716/pFDX4296 (*ptsN, rapZ, ptsO*). **D.** Complementation analysis of the  $\Delta rapZ$  single mutant Z225 using plasmid-encoded *rapZ* (pFDX4324). Cells were grown in LB, harvested at the indicated OD<sub>600</sub> values and  $\beta$ -galactosidase measurements were performed using samples derived from Z197 (wild type), Z225 ( $\Delta rapZ$ ) and Z225/pFDX4324 (*rapZ*).

Next, we addressed the influence of RapZ on  $P_{glmY}$  activity in derivatives of MG1655 *ilvG*<sup>+</sup> strain S4197 in order to investigate, whether the observed effect of RapZ is strain specific, i.e. if it only occurs in CSH50 derivative R1279 or if the effect is independent of strain specific traits and can be observed in other *E. coli* K12 wild type strains as well. Again, we performed growth experiments coupled to  $\beta$ -galactosidase measurements over the course of 10 hours (Fig. 3.3).

Indeed, we could observe an effect of a *rapZ* deletion on the activity of the *glmY* promoter (Fig. 3.3 B). However, the effect is not detectable during exponential growth, but becomes apparent as a two-fold reduction of LacZ activity in stationary growth phase. Even though the effect is weaker in *E. coli*

MG1655 derivatives as compared to CSH50 derivatives (a two-fold reduction of  $P_{glmY}$  activity vs. a 12-fold reduction during stationary growth, respectively), it is significant. As reported previously, phenotypes of deletion mutants might vary between CSH50 and MG1655 derivatives, i.e., effects of deleted genes on the expression of *lacZ* fusions may be lower in MG1655 derivatives possibly due to additional factors involved in the regulation or a higher fitness of this strain (Lüttmann *et al.*, 2012).

The observed effect of RapZ on  $P_{glmY}$  activity is independent of the  $\sigma^{70}$ -dependent *glmY* promoter, but has been associated with the  $\sigma^{54}$ -*glmY* promoter at least in CSH50 derivatives (Reichenbach, 2009). To address this question in derivatives of strain MG1655,  $\beta$ -galactosidase activity was assessed in a strain carrying a *glmY'*-*lacZ* fusion with a mutation within the -35 sequence of the  $\sigma^{70}$ -*glmY* promoter, leaving only the  $\sigma^{54}$ -dependent promoter intact (Fig. 3.2 D, Reichenbach *et al.*, 2009). Again, the activity of  $P_{glmY}$  was found to be reduced in this strain, however to a significantly lower extent than observed for the respective CSH50 derivative (Z226; Reichenbach, 2009).



**Figure 3.3: Activity of the  $\sigma^{54}$ -dependent *glmY* promoter in *E. coli* MG1655 *ilvG*<sup>+</sup> is also affected by RapZ.** A. Growth of MG1655 *ilvG*<sup>+</sup> (Z741) and the isogenic *rapZ* mutant (Z833) in LB was monitored over 10 hours. B. Influence of a *rapZ* deletion on *glmY'*-*lacZ* expression in *E. coli* K-12 MG1655 *ilvG*<sup>+</sup> was determined by  $\beta$ -galactosidase activity measurements from samples harvested at the indicated time intervals. C. MG1655 *ilvG*<sup>+</sup> strain Z743 carrying a mutation in the  $\sigma^{70}$ -dependent *glmY* promoter and the isogenic *rapZ* mutant (Z844) were grown in LB over the course of 10 hours. D.  $\beta$ -galactosidase activity measurements were performed using samples harvested at 1.5, 2, 4, 6 and 10 hours. E. Strains Z851 (wild type) and Z847 (*rapZ*) were grown in LB over the course of 10 hours. F.  $\beta$ -galactosidase measurements of *glmY'*-*lacZ* expression in strains Z851 and Z847 as determined along the growth curve. Samples were harvested at the indicated time intervals and  $\beta$ -galactosidase activity was measured.

One prominent mutation in CSH50 derivative R1279 and its descendants is a  $\Delta(\textit{phoU}\textit{-bgl})$  deletion. The *bgl* operon encodes genes required for uptake and utilization of  $\beta$ -glucosides, such as arbutin and salicin. Gene *phoU* encodes the repressor of the PhoB/R two-component system and is located directly upstream of the first gene in the *bgl* operon, *bglG*. Therefore, cells lacking *phoU* constitutively express the *pho* regulon regardless of the phosphate concentration leading to increased phosphate levels (Hsieh and Wanner, 2010; Lüttmann *et al.*, 2012). To exclude that elevated intracellular phosphate levels account for the differences in *glmY'*-*lacZ* expression observed in CSH50 and MG1655 derived *rapZ* deletion strains,  $\beta$ -galactosidase measurements were repeated in *phoU*<sup>+</sup> derivative of CSH50 (strain Z851) and the isogenic *rapZ* deletion strain Z847 (Fig. 3.3 E, F).

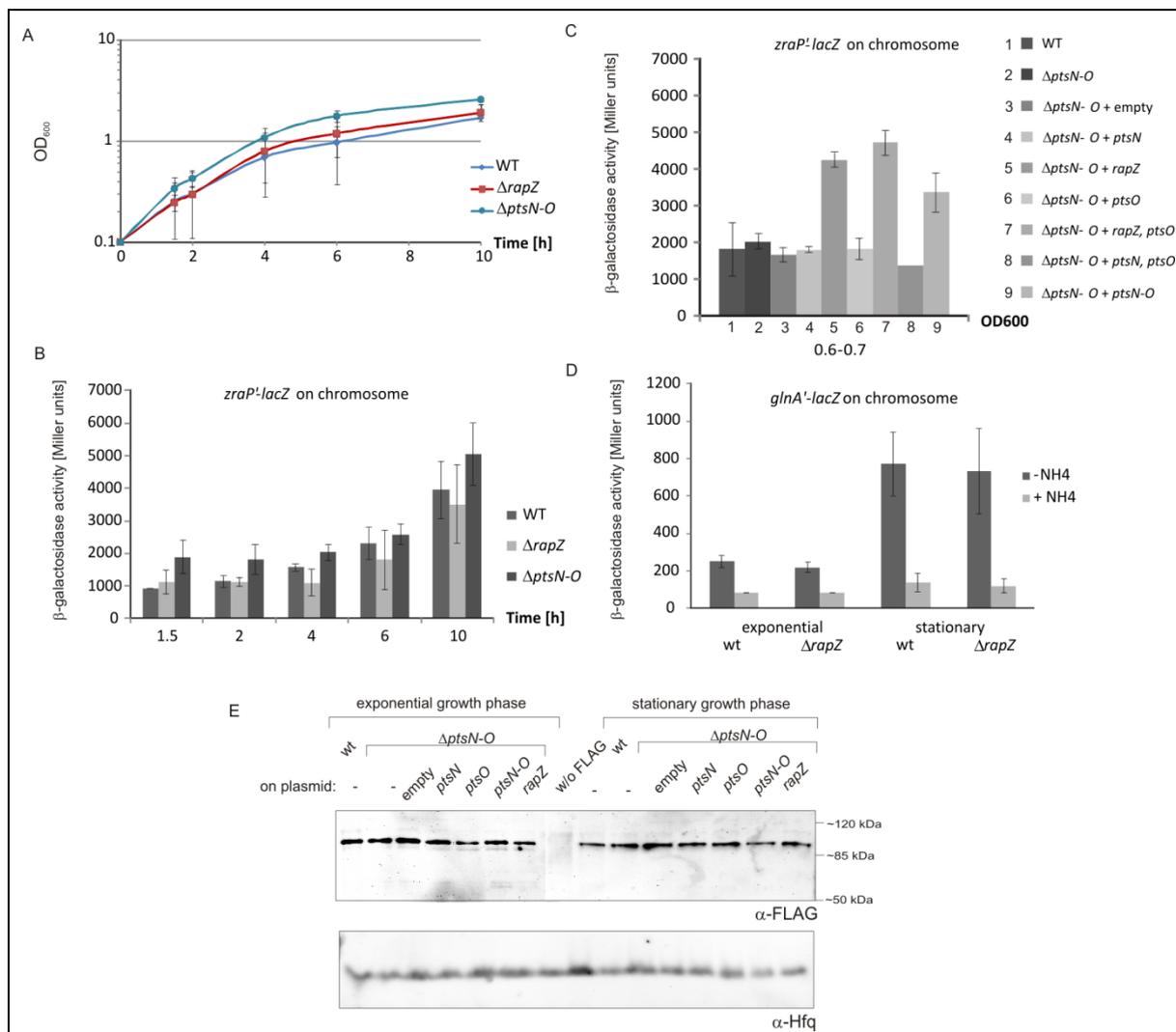
As the drastic reduction of  $P_{\textit{glmY}}$  activity of five to 10-fold persists in a *phoU*<sup>+</sup> genetic background (compare Figs. 3.2 B and 3.3 F), it can be excluded that differences in the intracellular phosphate levels are responsible for the observed variations in the phenotypes of *rapZ* deletion strains in R1279 (CSH50) descendants and MG1655 derivatives. Considering the strong regulatory phenotype, the following experiments were conducted in CSH50 derivatives.

***RapZ is specifically required for high  $P_{\textit{glmY}}$  activity, but over-expression of rapZ also enhances the activity of other  $\sigma^{54}$  dependent promoters***

Since *rapZ* is encoded within the *rpoN* operon (*rpoN* encodes for  $\sigma^{54}$ ) we reasoned that *rapZ* could be a modulator of RpoN activity or influence *rpoN* expression. To investigate whether the absence of *rapZ* also affects other  $\sigma^{54}$ -promoters we tested expression of a *zraP'*-*lacZ* fusion over the course 10 hours in the wild type as compared to  $\Delta\textit{ptsN-O}$  and  $\Delta\textit{rapZ}$  mutant strains (Fig. 3.4 B). Gene *zraP* encodes for a periplasmic zinc binding protein that is involved in zinc homeostasis. Expression of this gene is controlled by the ZraS/R two-component system that is induced upon high concentrations of zinc in the medium (Leonhartsberger *et al.*, 2001). To ensure expression of the reporter fusion, all experiments were carried out using cultures grown in the presence of 1 mM ZnCl<sub>2</sub>.

Neither deletion of *rapZ* alone, nor a triple deletion of *ptsN*, *rapZ* and *ptsO* had any significant effect on the expression of *zraP'*-*lacZ* under inducing conditions. The activity of the  $\sigma^{54}$ -dependent *zraP* promoter ( $P_{\textit{zraP}}$ ) increased over time in all three tested strains (wild type,  $\Delta\textit{rapZ}$ ,  $\Delta\textit{ptsN-O}$ ; Fig. 3.4 A and B). To further exclude a stimulatory effect of RapZ on  $P_{\textit{zraP}}$  activity, we tested whether increased levels of *rapZ* or its neighboring genes *ptsN* and *ptsO* had any effect on the expression of the *zraP'*-*lacZ* fusion by re-introducing the missing genes in several combinations. Unexpectedly,  $P_{\textit{zraP}}$  activity showed a ~two-fold increase upon addition of plasmids carrying *rapZ* either alone or in combination with *ptsO* or *ptsN* and *ptsO* (Fig 3.4 C). Thus, over-expression of *rapZ* led to mildly enhanced promoter activity of  $P_{\textit{zraP}}$  as previously observed for  $P_{\textit{glmY}}$  (compare Fig. 3.2 C, D and 3.4 C).

Therefore,  $P_{zraP}$  does not strictly require RapZ for activity, as was detected for  $P_{glmY}$ , but expression of the *zraP'*-*lacZ* can be enhanced by increased RapZ levels (Fig. 3.4 C, columns 5, 7 and 9).



**Figure 3.4: The absence of RapZ does not affect expression of *zraP* or *glnA*.** **A.** Strain Z274 (wild type), Z275 ( $\Delta rapZ$ ) and Z780 ( $\Delta ptsN-O$ ) were grown in LB with 1 mM ZnCl<sub>2</sub> over the course of 10 hours, samples were harvested at the indicated times. **B.**  $\beta$ -galactosidase activity measurements of samples derived from A at 1.5, 2, 4, 6 and 10 hours of growth. **C.** Complementation experiment re-introducing the missing genes in several combinations: Z274, Z780, Z780/pFDX4291 (empty), Z780/pFDX4294 (*ptsN*), Z780/pFDX4324 (*rapZ*), Z780/pFDX4292 (*ptsO*), Z780/pFDX4320 (*rapZ*, *ptsO*), Z780/pFDX4322 (*ptsN*, *ptsO*) and Z780/pFDX4296 (*ptsN*, *rapZ*, *ptsO*).  $\beta$ -galactosidase activities were determined at the indicated OD<sub>600</sub>. **D.** Strains Z849 (wild type) and Z850 ( $\Delta rapZ$ ) were grown in MOPS minimal medium (Neidhardt *et al.*, 1974) with and without ammonia and  $\beta$ -galactosidase measurements assessing the influence of the absence of *rapZ* on a chromosomal *glnA'*-*lacZ* fusion were performed from cells harvested in exponential (OD<sub>600</sub> 0.65-0.8) and stationary phase (OD<sub>600</sub> 1.4-1.9). **E.** Western Blot analyzing RpoN-3xFLAG levels in exponential and stationary growth phase. Strains Z807 (wild type, *rpoN*-3x-FLAG), Z809 ( $\Delta ptsN-O$ , *rpoN*-3x-FLAG) and transformants of Z809 over-expressing the indicated genes or the empty expression vector (Z809/pFDX4291 (empty), Z809/pFDX4294 (*ptsN*), Z809/pFDX4292 (*ptsO*), Z809/pFDX4296 (*ptsN*, *rapZ*, *ptsO*) and Z809/pFDX4324 (*rapZ*)) were grown in LB medium and harvested at OD<sub>600</sub> 0.5-0.8 or OD<sub>600</sub> 1.5-2.0. As a control strain Z197 (wild type, w/o FLAG) was used. Whole cell extracts were subjected to Western Blot analysis. FLAG-tagged RpoN was detected using an  $\alpha$ -FLAG antibody (upper panel), alternatively Hfq was detected by a specific antibody as loading control (lower panel).

To validate the role of RapZ for activity of yet another  $\sigma^{54}$ -dependent promoter, a *glnA'*-*lacZ* fusion was used as a further control. Gene *glnA* encodes for a glutamine synthase that converts glutamate and ammonia to glutamine. Expression of *glnA* is driven by a  $\sigma^{70}$ -dependent promoter ( $P_{glnA1}$ ) in rich media, whereas upon nitrogen limitation the  $\sigma^{54}$ -dependent  $P_{glnA2}$  promoter is activated by the two-component system NtrB/C (GlnL/G). The *glnA'*-*lacZ* fusion used in our experiments encompasses

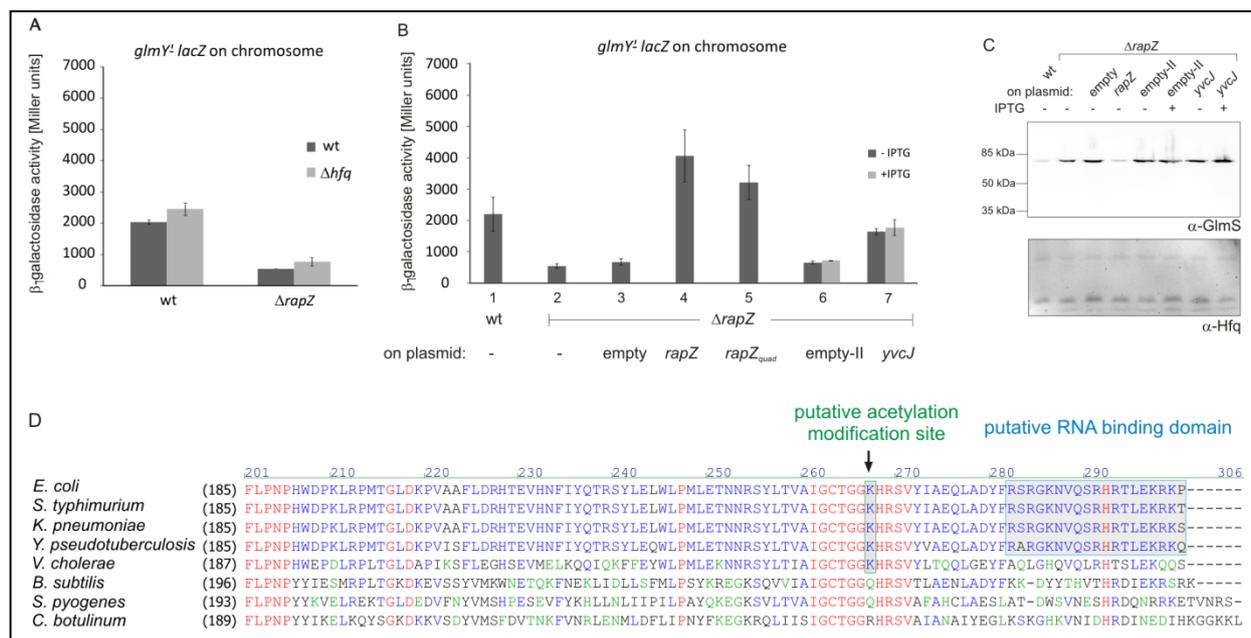
both promoters. Therefore, the basal expression obtained in MOPS minimal medium containing ammonia can be attributed to  $\sigma^{70}$ - $P_{glnA1}$ . Upon nitrogen limitation the  $\sigma^{54}$ -dependent  $P_{glnA2}$  promoter is activated by NtrB/C and RpoN, thus expression of the *glnA'*-*lacZ* reporter fusion should increase (Magasanik, 1993). In our experiments, nitrogen limitation led to a ~five-fold higher expression of the *glnA'*-*lacZ* fusion in exponential growth phase and an ~eight-fold increase in stationary phase as judged from  $\beta$ -galactosidase activity measurements (Fig. 3.4 D). Under all tested conditions, deletion of *rapZ* did not affect *glnA'*-*lacZ* expression levels (Fig. 3.4 D columns 3,4,7 and 8). These experiments were repeated in MG1655 *ilvG*<sup>+</sup> derivatives with similar results (data not shown). Analogous to  $P_{zraP}$ , activity of  $\sigma^{54}$ -dependent  $P_{glnA2}$  promoter does not strictly require presence of RapZ. Therefore, we conclude that RapZ is a modulator strictly required for the activity of the  $\sigma^{54}$ -dependent *glmY* promoter, but over-expression of *rapZ* mildly enhances at least one other  $\sigma^{54}$ -dependent promoter, namely  $P_{zraP}$ .

We performed western blot experiments to detect RpoN under various conditions to investigate whether changes in the *rpoN* expression level might be caused by deletion of *rapZ* or its over-expression and thus could account for the observed changes in activity of  $\sigma^{54}$ -dependent promoters (Fig. 3.4 E). In order to examine RpoN levels from whole cell extracts, the sequence coding for a 3' triple FLAG tag was fused to the *rpoN* gene at its authentic locus in the wild type as well as the  $\Delta$ *ptsN-O* triple mutant. Furthermore, we re-introduced the missing genes alone or in combination into the  $\Delta$ *ptsN-O* deletion mutant and analyzed RpoN-3x-FLAG levels in exponential and stationary growth phase. Interestingly, amounts of RpoN-3x-FLAG remained constant under all tested conditions. Therefore, RapZ and the other tested proteins, EIIA<sup>Ntr</sup> and NPr, encoded in the *rpoN* operon were without effect on *rpoN* abundance.

### ***The RNA binding function of RapZ is not required for regulation of $P_{glmY}$ activity***

Since RapZ was found to alter transcription of *glmY* as well as mediate feedback regulation of *glmS* expression via the GlmY/RapZ/GlmZ cascade, we next addressed the question whether these two functions are linked. It was shown previously, that regulation of *glmY* transcription is independent of sRNAs GlmY and GlmZ itself (Reichenbach, 2009). Thus, regulation of  $P_{glmY}$  activity by RapZ is unlikely to contribute to an auto-regulatory feedback loop, but rather serves a different function. However, it could not be excluded if other small RNAs are involved in this regulatory circuit or if the RNA binding function of RapZ is required for its stimulatory effect on the *glmY* promoter. To address this question, we once again used the *glmY'*-*lacZ* reporter gene fusion and determined  $\beta$ -galactosidase activities in the wild type and  $\Delta$ *rapZ* deletion mutant in absence and presence of RNA chaperon Hfq (Fig. 3.5). At least in Gram-negative bacteria, most base-pairing sRNAs require Hfq for functionality and stability (Urban and Vogel, 2007; Vogel and Luisi, 2011). Therefore, if a base-pairing sRNA would be involved

in regulation of *glmY* expression via RapZ, the reporter gene fusion should be differentially expressed in absence and presence of Hfq.



**Figure 3.5: RNA binding properties of RapZ are not essential for activity of the *glmY* promoter.** **A** Strain Z197 (wild type), strain Z824 (wild type,  $\Delta hfq$ ), strain Z225 ( $\Delta rapZ$ ) and strain Z825 ( $\Delta rapZ$ ,  $\Delta hfq$ ) were grown in LB to an OD<sub>600</sub> of 0.5-0.8 and  $\beta$ -galactosidase activity was determined. **B** Strains Z197 (wild type) and Z225 ( $\Delta rapZ$ ) as well as the following transformants: Z225/pFDX4291 (empty), Z225/pFDX4324 (*rapZ*), Z225/pYG82 (*rapZ<sub>quad</sub>*), Z225/pDG148-Stul (empty-II) and Z225/pBGM53 (*yvcJ*) were grown in LB to an OD<sub>600</sub> of 0.5-0.8, Z225/pDG148-Stul and Z225/pBGM53 were grown in the absence or presence of 1mM IPTG to induce gene expression and  $\beta$ -galactosidase activity was measured. **C** Western Blot analysis of GlmS levels in strains Z197 (wild type), Z225 ( $\Delta rapZ$ ) and transformants Z225/pFDX4291, Z225/pFDX4324, Z225/pDG148-Stul and Z225/pBGM53 grown in LB and harvested in exponential growth phase. 1 mM IPTG was added to the cultures as indicated. Blots were detected using an antibody directed against GlmS or against Hfq as loading control. **D** Sequence alignment of the C-Termini of RapZ proteins from *E. coli* K12 str. MG1655 (U00096.2), *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 (NC\_003197.1), *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 (NC\_009648.1), *Yersinia pseudotuberculosis* YPIII (NC\_010465.1), *Vibrio cholerae* O1 biovar El Tor str. N16961 (NC\_002505.1), *Bacillus subtilis* subsp. *subtilis* str. 168 (NC\_000964.3), *Streptococcus pyogenes* MGAS1882 (CP003121.1), and *Clostridium botulinum* A str. ATCC3502 (NC\_009495). Conserved amino acids are marked in red. Residues conserved in the majority of sequences are highlighted in blue and residues weakly similar are in green. The putative RNA binding domain as determined in Göpel *et al.*, 2013 is boxed in blue, a putative acetylation modification site as determined in Zhang *et al.*, 2009 is boxed in green.

As  $P_{glmY}$  activity is independent of the absence or presence of Hfq and as the effect of the  $\Delta rapZ$  deletion persists also in the  $\Delta hfq$  background (Fig. 3.5 A), it was concluded that the regulation of  $P_{glmY}$  activity by RapZ is most likely independent of other base-pairing sRNAs in *E. coli*.

To verify whether the regulation of *glmY* expression is indeed a second function of RapZ and independent of its RNA binding properties, we performed complementation experiments using a RapZ mutant that carries four amino acid exchanges in its putative C-terminal RNA-binding domain (designated RapZ<sub>quad</sub>). This mutant was shown to be incapable to bind sRNAs GlmY and GlmZ (Göpel *et al.*, 2013; Fig. 3.5 B). Interestingly, this mutant still complemented the  $\Delta rapZ$  deletion and restored high  $\beta$ -galactosidase activities (Fig. 3.5 B column 5) suggesting that RapZ does not require the RNA binding function for stimulation of *glmY* expression.

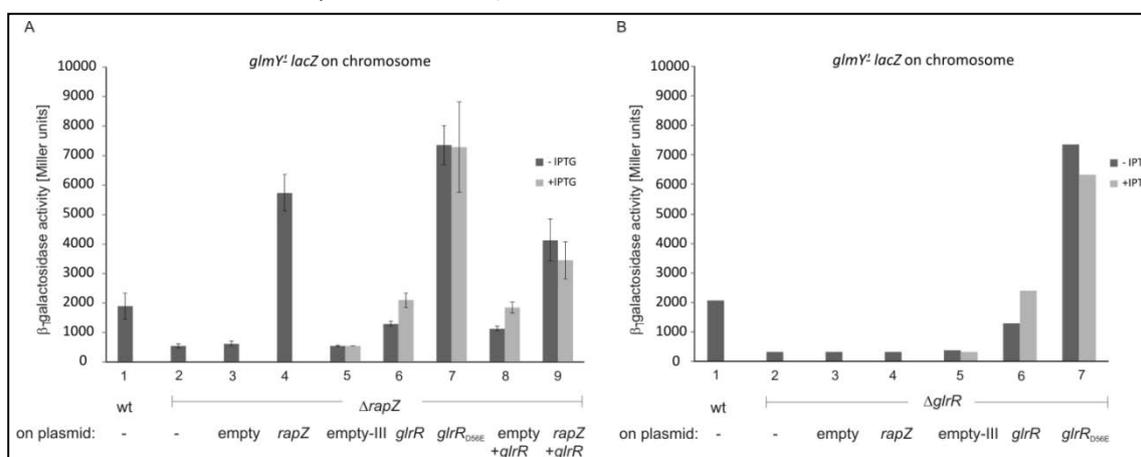
To further validate this finding, we conducted complementation experiments over-expressing the *Bacillus subtilis* 168 *rapZ* homolog *yvcJ*. As expression of *glmS* is controlled by the well-studied *glmS*-ribozyme/riboswitch in *Bacilli*, sRNAs GlmY and GlmZ are not present in *Bacillus* species (Göpel *et al.*, 2011; McCown *et al.*, 2012). Furthermore, while certain amino acid motifs in sequences of RapZ homologs are highly or totally conserved, e.g. the N-terminal Walker A and Walker B motifs, the putative RNA-binding domain is only conserved in *Enterobacteriaceae* that contain GlmY and GlmZ, but not in homologs from species like *B. subtilis* (Göpel *et al.*, 2013, Fig. 3.5 D). Thus, the latter homologs presumably lack the ability to bind RNA. Interestingly, introduction of a plasmid over-expressing *yvcJ* restored  $\beta$ -galactosidase activity of the *glmY'*-*lacZ* fusion almost to wild type levels (Fig. 3.5 B column 7) but did not further enhance  $P_{glmY}$  activity as observed when *rapZ* or *rapZ<sub>quad</sub>* were over-expressed (Fig. 3.5 B columns 4 and 5).

Nevertheless, our data suggests that the protein domains responsible for regulation of  $P_{glmY}$  activity are distinct from the RNA binding domain of enterobacterial RapZ homologs, which functions in regulating *glmS* expression. To further validate this hypothesis, we performed Western blot experiments addressing the cellular amounts of GlmS in the *rapZ* mutant and the complemented mutant as compared to the wild type (Fig. 3.5 C). Indeed, although YvcJ complemented a *rapZ* deletion in respect to *glmY* expression levels, it failed to counteract the constitutive over-expression of *glmS* in *rapZ* mutants (Fig. 3.5 C). In the absence of RapZ, RNase E is unable to process GlmZ, which leads to constitutive activation of the *glmS* mRNA and thus to elevated GlmS levels (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Göpel *et al.*, 2013). Western blot experiments addressing the cellular amounts of GlmS in the wild type and  $\Delta rapZ$  mutant show a considerable increase of GlmS upon deletion of *rapZ* (Fig. 3.5 C lane 1 and 2). Complementation of the deletion mutant with a plasmid expressing *rapZ* strongly reduced GlmS amounts to nearly wild type levels (Fig. 3.5 C lane 4). In contrast, introduction of a plasmid expressing *yvcJ* did not lead to reduction of GlmS levels (Fig. 3.5 C lane 7 and 8). Thus, RapZ fulfills two distinct functions: modulation of the *glmY* promoter and feedback regulation of *glmS* expression within the GlmYZ cascade; and our data suggest that these functions are likely independent of each other.

### ***RapZ requires $\sigma^{54}$ - specific activator protein GlrR to exert its stimulatory effect on $P_{glmY}$ promoter activity***

In theory, it is likely that RapZ might either act by modulation of the activity of the sigma factor itself, or by modifying the activity of another factor involved in control of *glmY* transcription. To address the question whether RapZ acts by modulation of the GlrK/GlrR two-component system on the *glmY* expression, we investigated the effect of a *glrR* over-expression in cells lacking *rapZ* (Fig. 3.6 A). As

mentioned above, in *E. coli*, response regulator GlrR is a  $P_{glmY}$  specific  $\sigma^{54}$ -activator protein and activity of  $\sigma^{54}$ -dependent  $P_{glmY}$  promoter strictly depends on phosphorylated GlrR (Fig. 3.6 B; Reichenbach *et al.*, 2009; Göpel *et al.*, 2011).



**Figure 3.6: Stimulation of the *glmY* promoter by RapZ depends on GlrR.** **A.** Strains Z197 (wild type), Z225 ( $\Delta rapZ$ ) and transformants Z225/pFDX4291 (empty), Z225/pFDX4324 (*rapZ*), Z225/pKESK-23 (empty-III), Z225/pYG89 (*glrR*), Z225/pYG90 (*glrR*<sub>D56E</sub>), Z225/pFDX4291+pYG89 and Z225/pFDX4324+pYG89 were grown in LB to an OD<sub>600</sub> of 0.5-0.8, 1 mM IPTG was added to the indicated transformants and  $\beta$ -galactosidase activity was determined. **B.** Strains Z197 (wild type), Z206 ( $\Delta glrR$ ) and transformants Z206/pFDX4291 (empty), Z206/pFDX4324 (*rapZ*), Z206/pKESK-23 (empty-III), Z206/pYG89 (*glrR*), Z206/pYG90 (*glrR*<sub>D56E</sub>), were grown in LB to an OD<sub>600</sub> of 0.5-0.8, 1 mM IPTG was added to the indicated transformants and  $\beta$ -galactosidase activity was determined. The graph depicted in B shows  $\beta$ -galactosidase values of one experiment.

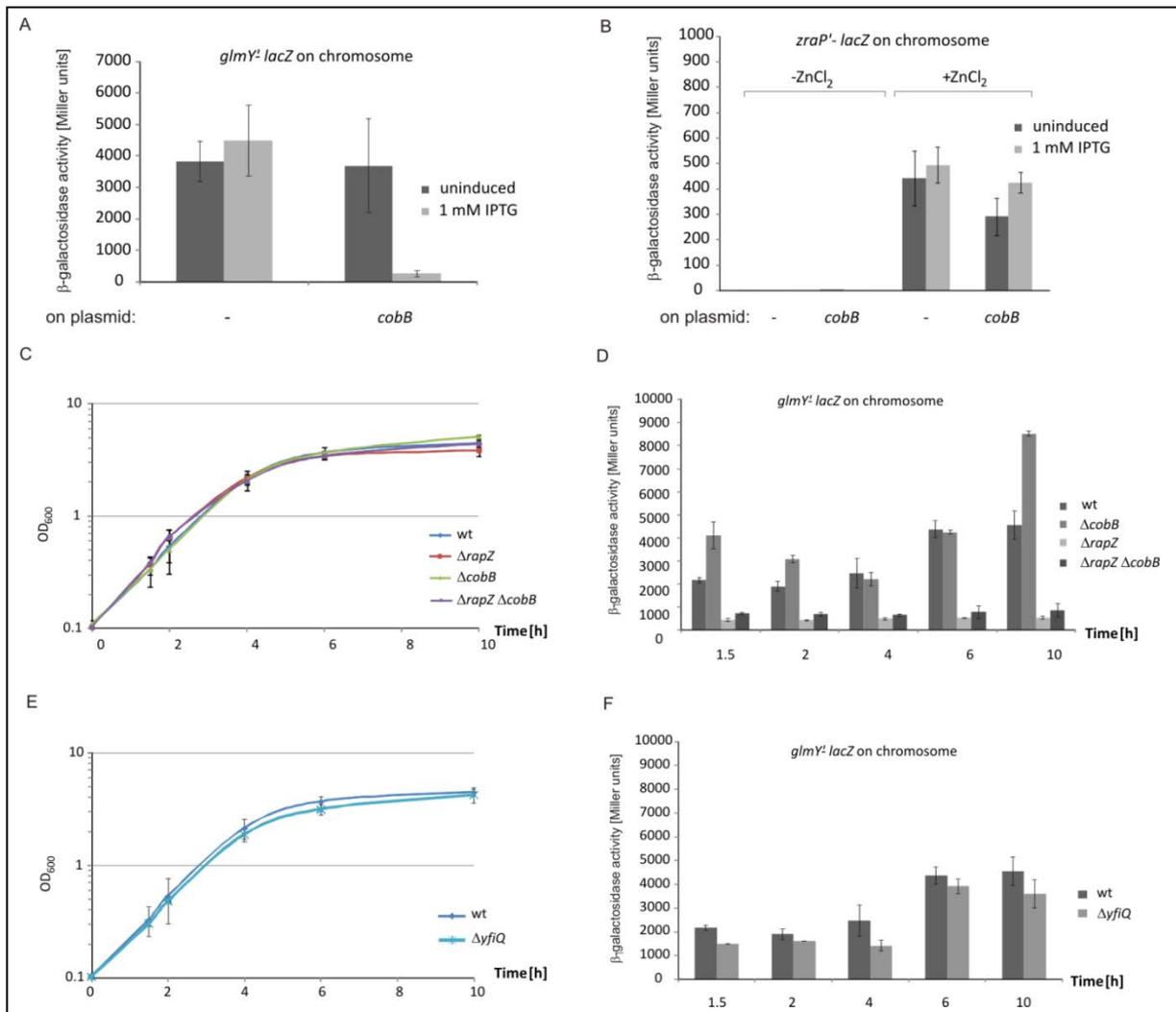
Upon introduction of a plasmid expressing *glrR*, in addition to the endogenous *glrR*, *glmY'*-*lacZ* expression was restored to wild type levels even in a *rapZ* deletion mutant (Fig. 3.6 A, column 6). When using a phospho-mimetic *glrR*<sub>D56E</sub> mutant, *glmY'*-*lacZ* expression levels were increased ~four-fold while over-expression of *rapZ* lead to a higher *glmY'*-*lacZ* expression as compared to wild type levels (Fig. 3.6 A, columns 7 and 4). Simultaneous over-expression of *glrR* and *rapZ* did not further increase  $P_{glmY}$  activity. Rather, promoter activity resembled the observed ~two to three-fold increase of the reporter fusion upon *rapZ* over-expression (Fig. 3.6 A column 9, compare Fig. 3.6 A column 4 and Fig. 3.5 B column 4). In contrast, over-expression of *rapZ* is insufficient to complement a *glrR* deletion mutant (Fig. 3.6 B, column 4). Only, complementation with *glrR* or the *glrR*<sub>D56E</sub> variant restored high  $\beta$ -galactosidase activities. Again, while expression of *glrR* restored wild type *glmY'*-*lacZ* levels, over-expression of *glrR*<sub>D56E</sub> resulted in a three to four-fold activation of the *glmY* promoter (Fig. 3.6 B, columns 6 and 7). Thus, over-expression of *glrR* rescued a *rapZ* deletion and restored wild type levels of *glmY'*-*lacZ* expression, whereas *rapZ* requires the presence of GlrR for its stimulatory effect. The phospho-mimetic *glrR*<sub>D56E</sub> variant bypasses the need of RapZ for full activation of the *glmY* promoter under the conditions tested.

***Sirtuin deacetylase CobB is another potent regulator that acts indirectly to modulate activity of the  $\sigma^{54}$ -dependent *glmY* promoter***

Recently, we performed a gene bank screen in order to identify unknown factors involved in the modulation of *glmY* expression from its  $\sigma^{54}$ -dependent promoter (Künzl, Göpel and Görke, unpublished). In this screen an *E. coli* K-12 derived genomic library was introduced into a strain carrying a chromosomal *glmY'*-*lacZ* fusion with a mutated -10 sequence. Thus, only the  $\sigma^{54}$ -dependent  $P_{glmY}$  promoter is active. The obtained transformants carrying plasmids over-expressing parts of the *E. coli* genome were selected on M9-glycerol-XGal plates for altered *glmY'*-*lacZ* expression levels by blue-white screening. This approach identified deacetylase CobB as putative regulator of  $\sigma^{54}$ - $P_{glmY}$  activity (Künzl, Göpel and Görke, unpublished). To validate the effect of *cobB* over-expression on the expression of the *glmY'*-*lacZ* fusion, the *cobB* gene was subcloned under control of an IPTG-inducible promoter. The resulting plasmid was introduced into the *glmY'*-*lacZ* reporter strain and  $\beta$ -galactosidase measurements were performed (Fig. 3.7A).

Upon induction with IPTG, over-expression of *cobB* drastically reduced  $\sigma^{54}$ -dependent *glmY'*-*lacZ* expression (~17-fold, Fig. 3.7 A columns 2 and 4). However, no significant effect was observed, when the *zraP'*-*lacZ* fusion was used as a control (Fig. 3.7 B). Note that expression levels of *zraP'*-*lacZ* as determined in exponential growth phase in M9-glycerol medium containing 1mM ZnCl<sub>2</sub> are about ~two-fold lower compared to *zraP'*-*lacZ* expression in LB supplemented with 1mM ZnCl<sub>2</sub> (compare Fig. 3.7 B and 3.4 B, C first columns). Since over-expression of *cobB* only decreased activity of  $\sigma^{54}$ - $P_{glmY}$  but not of  $P_{zraP}$ , deacetylase CobB and thus acetylation seems to play an important role for regulation of the  $\sigma^{54}$ -dependent *glmY* promoter, but not for  $\sigma^{54}$ -dependent promoters in general.

Next, we checked whether a deletion of *cobB* also effects  $P_{glmY}$  activity as determined over the course of 10 hours by monitoring expression of the *glmY'*-*lacZ* reporter fusion. Here, the *cobB* deletion mutant showed a ~two-fold increase in *glmY'*-*lacZ* expression in stationary phase as compared to the wild type (Fig. 3.7 D, second bar). Upon introduction of an additional *rapZ* mutation, *glmY'*-*lacZ* expression was reduced ~five to 12-fold as compared to wild type levels. Note that in the double mutant *glmY'*-*lacZ* expression was at all times approximately 1.3× higher than in the *rapZ* single mutant (Fig. 3.7 D, third and fourth bar). However, *glmY'*-*lacZ* expression remained constant and did not further increase in stationary phase. These data suggest that *rapZ* is epistatic over *cobB* and CobB might act via RapZ on the activity of the *glmY* promoter.



**Figure 3.7: Deacetylase CobB is a potent regulator of *glmY* promoter activity.** **A** Strain Z190 carrying a *glmY*<sup>+</sup>-*lacZ* fusion with a mutated -10 sequence on chromosome was transformed with the empty expression vector pKESK-22 (-) or the *cobB* over-expression vector pYG38, respectively. Transformants were grown in M9-Glycerin (1%) to an OD<sub>600</sub> of 0.3 to 0.6 in the absence and presence of 1mM IPTG and  $\beta$ -galactosidase activity was determined. **B** The experiment described in A was repeated using strain Z274 carrying a chromosomal *zraP'*-*lacZ* fusion. In addition, expression levels were determined in absence and presence of ZnCl<sub>2</sub>, the inducer of the ZraS/R two-component system that controls *zraP* expression. **C** Strains Z197 (wild type), Z225 ( $\Delta rapZ$ ), Z846 ( $\Delta cobB$ ) and Z853 ( $\Delta rapZ \Delta cobB$ ) were grown in LB over the course of 10 hours and samples were harvested at the indicated times. **D**  $\beta$ -galactosidase activity measurements of the samples derived from the growth experiment depicted in C. **E** Strains Z197 (wild type) and Z867 ( $\Delta yfiQ$ ) were grown in LB for 10 hours and samples were harvested at 1.5, 2, 4, 6 and 10 hours of growth. **F**  $\beta$ -galactosidase activities of the samples derived from the growth experiment shown in E.

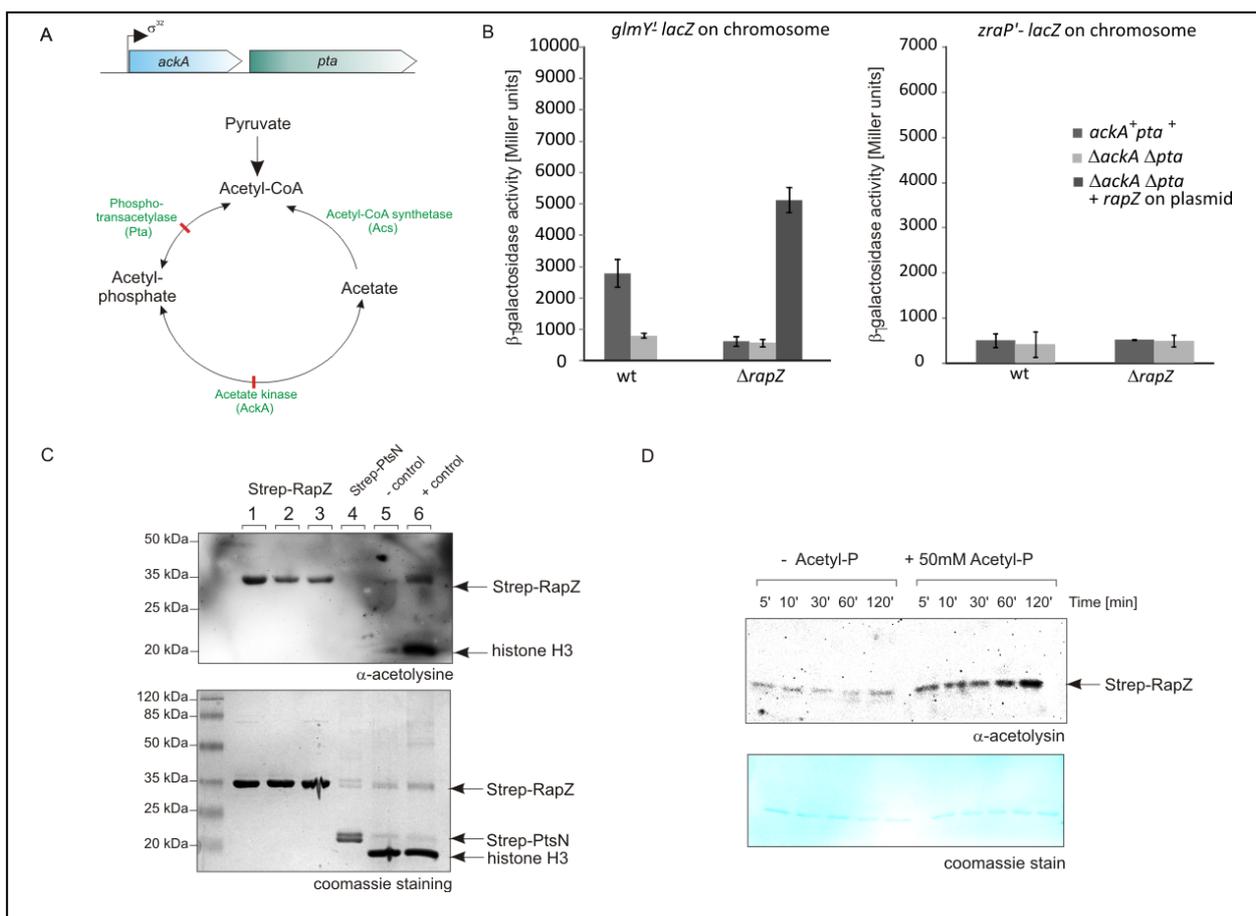
Since deacetylase CobB and its antagonist acetyl transferase YfiQ (Pka) provide a regulatory circuit that modulates activity of acetyl-CoA synthase Acs, it is likely that the activities of other acetylation targets are inversely regulated by CobB/YfiQ as well (Hu *et al.*, 2010; Starai *et al.*, 2002). We therefore tested whether acetyl transferase YfiQ is also involved in control of *P<sub>glmY</sub>* promoter activity in different growth phases. However, we could not detect any significant effect of a *yfiQ* deletion on *glmY* expression as monitored by the *lacZ* reporter fusion over the course of 10 hours (Fig. 3.7 E and F). Thus, it is likely that either another unknown acetyl transferase is required for acetylation or acetylation might occur independently of an acetyl transferase, e.g. by auto-acetylation using acetyl-phosphate as a donor.

***Acetyl-phosphate generated by the AckA/Pta pathway is a prerequisite for activity of the  $\sigma^{54}$ -dependent  $P_{glmY}$  promoter***

We reasoned that, if a so far unknown acetyl transferase is involved in regulation of the *glmY* promoter, elevated levels of acetyl-Coenzyme A (acetyl-CoA) that serves as acetyl-group donor would enhance the activity of  $P_{glmY}$ . To investigate the role of acetyl-CoA we performed  $\beta$ -galactosidase measurements in strains deficient in the Pta/AckA pathway (Fig. 3.8 B). This pathway converts acetyl-CoA to acetate and vice versa. Disruption of these reactions enhances the intracellular acetyl-CoA pool, so that more acetyl-CoA is available for anabolic processes and acetylation events (Fig. 3.8 A).

Surprisingly, introduction of a  $\Delta ackA \Delta pta$  double mutation in the *glmY'-lacZ* reporter strain resulted in a ~four-fold decrease of  $\beta$ -galactosidase activity (Fig. 3.8 B second column), whereas expression of a *zraP'-lacZ* fusion was not altered (Fig. 3.8 B second graph). However, in a *rapZ* deletion strain, which shows a ~four-fold reduction of  $P_{glmY}$  activity due to lack of *rapZ*, an additional  $\Delta ackA \Delta pta$  double mutation was without effect and did not further decrease promoter activity (Fig. 3.8 B column 3 and 4). Introduction of a plasmid expressing *rapZ* restored high  $\beta$ -galactosidase activity and even abolished the negative effect of the additional  $\Delta ackA \Delta pta$  deletion (Fig. 3.8 B column 5) indicating that the Pta/AckA pathway functions upstream of RapZ, for example by providing an acetyl-group donor.

Acetyl-P is the high energy intermediate of this pathway that can either be generated from acetate by acetate kinase AckA or from acetyl-CoA by phospho-transacetylase Pta (Fig. 3.8 A). Acetyl-P levels depend on the availability of intracellular acetyl-CoA and extracellular acetate (Wolfe, 2005). Further, acetyl-P is known to function as phosphoryl-group donor for response regulators *in vitro* and also possesses some *in vivo* activity (Lukat *et al.*, 1992; Scharf, 2010). Recently, *glmY*-specific response regulator GlrR was shown to exhibit increased DNA-binding affinity when incubated with acetyl-P *in vitro* (Göpel *et al.*, 2011).



**Figure 3.8: Acetyl-P can act as acetyl-group donor for RapZ *in vitro*.** **A** schematic overview of the main acetyl-P generating pathway, modified from (Wolfe, 2005). Top: genetic organization of the *ackA-pta* operon. Bottom: Phospho-transacetylase Pta converts acetyl-CoA to acetyl-P in a reversible reaction. Acetyl-P can also be generated from acetate by acetate kinase AckA; this reaction is reversible. Further, acetyl-CoA synthetase (Acs) can refill the acetyl-CoA pool from acetate. Blocking both Pta and AckA reactions increases the acetyl-CoA pool in the cell. **B**. Strains Z197 (wild type), Z791 ( $\Delta ackA-pta$ ), Z225 ( $\Delta rapZ$ ), Z792 ( $\Delta rapZ, \Delta ackA-pta$ ) and transformant Z792/pFDX4324 (*rapZ*) were grown in LB to an  $OD_{600}$  of 0.6-0.8 and  $\beta$ -galactosidase measurements were performed. **C** Top panel: western blot analyzing the acetylation state of Strep-RapZ purified from Z741 (wild type), Z762 ( $\Delta cobB$ ) and Z762/pYG38 (*cobB*). Purified Strep-PtsN and *Xenopus laevis* histone H3 served as negative control, *X. laevis* lysine acetylated histone H3 K56Ac served as positive control. The acetylation state of the proteins was detected using an  $\alpha$ -acetyllysine antibody. Lower panel: SDS-Gel of the samples analyzed by western blot serves as loading control. **D**. *in vitro* acetylation assay using purified Strep-RapZ from Z762/pYG38 (*cobB*). Upper panel: 10  $\mu$ g of purified Strep-RapZ was incubated for 2 h at 30°C with 20 mM nicotinamid in 1x structure buffer in the absence and presence of 50 mM acetyl-P and 2  $\mu$ g samples were collected at the indicated times. The acetylation state of RapZ over the course of 2 h was addressed by western blotting using an  $\alpha$ -acetyllysine antibody. Lower panel: Coomassie stain of the membrane after blotting serves as loading control.

So far, it is not clear if the auto-phosphorylation of GlrR is relevant *in vivo*, but it could, at least in part account for the observed reduction of *glmY* promoter activity upon deletion of *ackA* and *pta*. However, this does not provide an explanation for the observation that over-expression of *rapZ* again led to two to three-fold enhanced promoter activity even in the absence of *ackA* and *pta* (Fig. 3.8 B, fifth column), suggesting that RapZ might be the limiting factor.

### **RapZ is acetylated *in vivo* and is capable to auto-acetylate with acetyl-phosphate *in vitro***

Interestingly, RapZ possesses a putative acetylation site at lysine residue K251 and proteome data suggests that RapZ might be acetylated *in vivo* (Zhang *et al.*, 2009). To address the question if acetylation of RapZ could in part be responsible for the observed effect of the  $\Delta ackA \Delta pta$  mutation

on  $P_{glmY}$  activity, we first conducted western blot experiments addressing the acetylation state of RapZ (Fig. 3.8 C). For western blot analysis, Strep-RapZ was purified from wild type cells (Z741), cells deficient in *cobB* (Z762) and cells over-expressing *cobB* (Z762/pYG38). Purified Strep-PtsN and histone H3 from *Xenopus laevis* served as negative controls, pre-acetylated *X. laevis* histone H3<sub>K56Ac</sub> served as positive control.

Indeed, Strep-RapZ could be detected using an antibody directed against acetylated lysine residues, although signal strength did not vary significantly between Strep-RapZ derived from wild type, *cobB* deficient or *cobB* over-producing cells (Fig. 3.8 C, lanes 1-3). Further, signal strength of purified Strep-RapZ resembled in intensity the signal for the positive control, *in vitro* acetylated histone H3, whereas the negative controls could not be detected. These data are in agreement with a previous study identifying a RapZ derived peptide as acetylated in a mass spectrometry approach (Zhang *et al.*, 2009). However, a corresponding acetyl transferase is currently unknown. Taking into account that a *yfiQ* (*pat*) deletion did not have an effect on the *glmY*'-lacZ reporter fusion and disruption of the Pta/AckA pathway lead to reduced  $P_{glmY}$  activity, we hypothesized that acetyl-P rather than acetyl-CoA might possibly be the acetyl-group donor for RapZ. Consistently, the effect of the  $\Delta$ *ackA*  $\Delta$ *pta* deletion on  $P_{glmY}$  activity could be explained by the fact that in strains lacking *ackA* and *pta* the acetyl-P levels are drastically reduced (Mizrahi *et al.*, 2006).

To corroborate that acetyl-P could indeed serve as acetyl-group donor we performed *in vitro* acetylation assays incubating purified Strep-RapZ with 50 mM acetyl-P for two hours (Fig. 3.8 D). Samples of 2  $\mu$ g protein were collected at the indicated time intervals and the acetylation state of Strep-RapZ was monitored by western blotting using the  $\alpha$ -acetyllysine antibody. While samples derived from mock treatment (without acetyl-P) exhibited constantly low acetylation levels, signal strength increased over time for samples treated with acetyl-P (Fig. 3.8. D, upper panel). These observations indicate an increase in acetylation upon incubation with acetyl-P. Thus, we conclude that at least *in vitro* RapZ is able to auto-acetylate using acetyl-P as donor. However, it remains to be clarified if acetylation with acetyl-P also plays a role *in vivo*.

## DISCUSSION

Extensive control of the expression of sRNA regulators on the transcriptional level emerges as a widespread regulatory principle (Beisel and Storz, 2010; Göpel *et al.* 2011; Chao *et al.* 2012; Göpel and Görke 2012a). The need for such elaborate control is founded in the global role of sRNA regulators that are involved in virtually every physiological process in the cell (Gottesman and Storz, 2011).

In this work, we add to the repertoire of regulators, identifying outer membrane protein YfhG as an activator for  $\sigma^{54}$ -dependent *glmY* expression that acts presumably via GlrK/GlrR. Furthermore, we could demonstrate a regulatory function of RNA-binding protein RapZ in the control of  $\sigma^{54}$ -dependent expression of small RNA GlmY that seems to be dependent on acetylation of RapZ. In this respect, we found that deacetylase CobB can inhibit *glmY* transcription from the  $\sigma^{54}$ -dependent promoter possibly by deacetylation of RapZ. Interestingly, whereas acetylation of RapZ seems to be independent of the major acetyl-transferase YfiQ (Pat), we propose a role for acetyl-P as acetyl-group donor. At least *in vitro*, the acetylation state of RapZ can be increased by incubation with acetyl-P.

Importantly, the function of RapZ in *glmY* promoter control appears to be independent of its recently described function as an RNase E adaptor protein promoting turn-over of sRNA GlmZ. This conclusion is supported by the observation that a RapZ quadruple mutant unable to bind sRNA GlmY and GlmZ as well as the *Bacillus* RapZ homologue YvcJ can complement a *rapZ* deletion in respect to *glmY* promoter activity, but not in respect to regulation of *glmS* expression (Fig. 3.5 B and C; Göpel *et al.*, 2013). Since *rapZ* is an “evolutionary old” gene, it is conserved in number of bacterial phyla and might have been present in the last common ancestor (Pompeo *et al.*, 2011). Thus, it is likely that RapZ and its homologs serve diverse functions in different bacteria. However, they might possibly share a more wide spread conserved function.

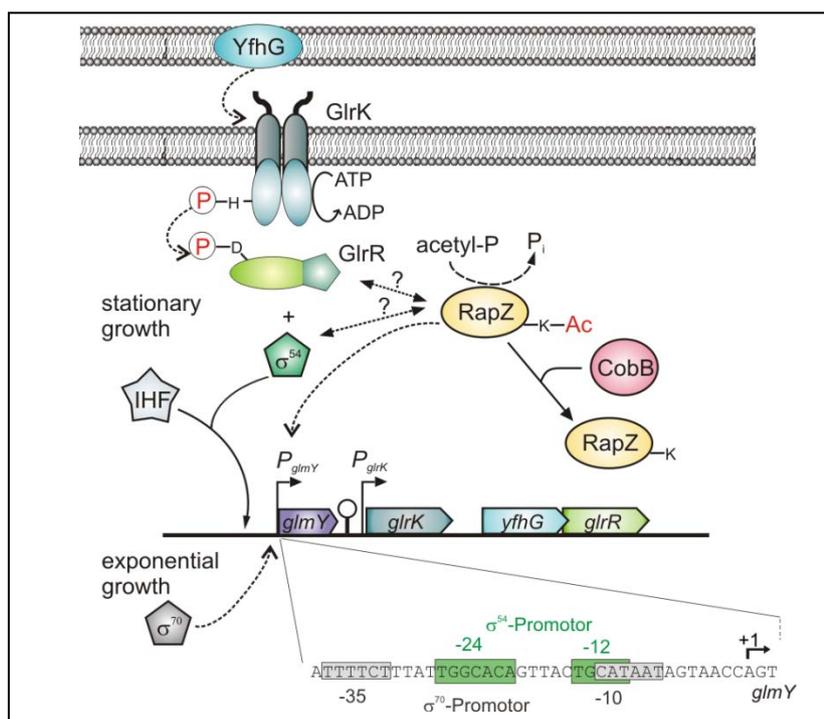
In *E. coli*, we identified such a second function for RapZ, namely regulation of *glmY* promoter activity. As RapZ strictly requires the *glmY*-specific  $\sigma^{54}$ -activator protein GlrR for its role in regulation of *glmY* expression, it is likely that RapZ acts indirectly. However, the molecular mechanism underlying this regulation remains to be clarified. In principle, two general modes of action are plausible for the RapZ-mediated control of the activity of the  $\sigma^{54}$ -dependent *glmY* promoter. Either RapZ might function by modifying the activity of a sigma factor, in this case  $\sigma^{54}$ , or it might act as a modulator of the activity of a transcriptional regulator, for which GlrR would be a likely candidate.

Alternative sigma factors need to be tightly regulated since they control large sets of genes. Because of its role as a regulatory hub for integration of a multitude of stress-related stimuli, sigma factor  $\sigma^S$  (encoded by *rpoS*) for example, is subject to complex control on virtually every regulatory level. At the transcriptional level *rpoS* expression is negatively controlled by CRP-cAMP and the ArcB/A two-component system. While response regulator ArcA directly represses *rpoS*, histidine kinase ArcB additionally phosphorylates adaptor protein RssB that then targets the RpoS protein for proteolysis by ClpXP (Becker *et al.*, 1999; Mika and Hengge, 2005). RpoS degradation is controlled by a complex cascade composed of RssB and several anti-adaptors, e.g. IraP, IraM and IraD (Bougdour *et al.*, 2008). Three sRNAs, RprA, DsrA and OxyS, control *rpoS* mRNA stability and translation initiation (Lease and Belfort, 2000; Majdalani *et al.*, 2001; Majdalani *et al.*, 2002; Zhang *et al.*, 1998).

Interestingly, no such regulation is currently known for  $\sigma^{54}$  and we could demonstrate that it seems to be constitutively expressed during exponential and stationary growth in complex medium (Fig. 3.4 E). RpoN expression levels have been estimated to 1/10 of the amount of  $\sigma^{70}$  (RpoD) present in cells (Jishage *et al.*, 1996). Considering that  $\sim 70$  transcription units are regulated by  $\sigma^{54}$  jointly with specific activator proteins, it seems to be presumptuous that *rpoN* itself should not be regulated at some level (Zhao *et al.*, 2010). From this study, it could not be excluded that RapZ acts as a modulator of RpoN activity since it had a strong stimulatory effect on the  $\sigma^{54}$ -dependent *glmY* promoter (Figs. 3.2 - 3.3) and a slight effect on  $\sigma^{54}$ -dependent *zraP* promoter (Fig. 3.4 C), while RpoN expression levels remained constant under all tested conditions (Fig. 3.4 E). If RapZ should indeed have an influence on RpoN activity, it is tempting to speculate that that this regulation might require protein-protein interaction and maybe even transient modification of RpoN, e.g. by phosphorylation or acetylation.

Further, we could demonstrate that RapZ acts upstream of response regulator GlrR and therefore RpoN since over-expression of *glrR* could compensate for the deletion of *rapZ* (Fig. 3.6 A). In contrast, RapZ is strictly dependent on GlrR to exert its stimulatory role on the *glmY* promoter (Fig. 3.6 B). These observations are in agreement with previous gel shift experiments showing that RapZ alone does not bind to the *glmY* promoter fragment that is bound by GlrR (Reichenbach *et al.*, 2009; Reichenbach, 2009). It could also be imagined that RapZ acts by modifying activity of GlrR instead of RpoN. For example, RapZ might interact with GlrR and enhance its DNA-binding affinity by conferring structural changes or even modify the protein.

RapZ and its homologs contain highly conserved Walker A and Walker B motifs. The *B. subtilis* homolog YvcJ and RapZ were shown to bind and hydrolyze ATP and GTP. Furthermore, deletion of *yvcJ* reduced expression of competence genes two to four-fold in a process that presumably involves ATP-dependent phosphorylation of a so far unidentified cellular component by YvcJ (Luciano *et al.*, 2009). Thus, at least in *B. subtilis* a RapZ homolog might possess kinase activity. Interestingly, processing of GlmZ by RNase E and RapZ did not depend on ATP or GTP *in vitro* (Göpel, Vogel and Görke, unpublished). Therefore, it seems likely that hydrolysis of these nucleotides and/or possibly kinase activity might be required for the role of RapZ in promoter control. Thus, we propose a model in which RapZ indirectly activates  $\sigma^{54}$ -dependent *glmY* transcription in concert with the GlrK/GlrR two-component system either by modulating activity of response regulator GlrR or  $\sigma^{54}$  itself (Fig. 3.9). However, the molecular mechanism by which RapZ activates the  $\sigma^{54}$ -dependent *glmY* promoter remains elusive and is subject of ongoing investigation.



**Figure 3.9: Acetylated RapZ enhances the activity of the  $\sigma^{54}$ -dependent *glmY* promoter.** Proposed working model for the regulatory circuit involved in control of the *glmY* promoter. During exponential growth, the  $\sigma^{70}$ -dependent *glmY* promoter ensures that sufficient amounts of GlmY are present to mediate glucosamine-6-phosphate homeostasis by feed-back regulation of *glmS* translation via the GlmY/RapZ/GlmZ cascade. During transition to stationary phase, the  $\sigma^{54}$ -dependent *glmY* promoter is strongly activated due to activation of the GlrK/GlrR two-component system. The signal that is sensed by histidine kinase GlrK, however, remains elusive. Outer membrane protein YfhG presumably acts as an activator of the GlrK/GlrR system and might act in signal perception/transition. Upon activation, GlrK auto-phosphorylates and the phosphoryl-group is transferred to Asp56 residue in GlrR. GlrR in turn is activated by phosphorylation and binds to three GlrR recognition motives far upstream of the *glmY* promoter. IHF aids in DNA bending and thus ensures interaction between the GlrR dimers and the  $\sigma^{54}$ -RNA-polymerase holoenzyme. This process further involves RapZ by a so far unknown mechanism. RapZ auto-acetylates presumably with acetyl-P and thus is able to activate the  $\sigma^{54}$ -dependent *glmY* promoter. This regulation is likely to be indirect, possibly by interaction of RapZ with either response regulator GlrR or sigma factor the  $\sigma^{54}$ . Acetylation of RapZ is likely to be reversed by deacetylase CobB.

Recently, acetylation of the RNA polymerase holoenzyme has emerged as a regulatory mechanism controlling activity of this important enzyme, e.g. acetylation of lysine residues within the  $\alpha$ -C-terminal domain of RNA polymerase is important for CpxA-independent activation of transcription of stress resistance factor *cpxP* (Lima *et al.*, 2011; Lima *et al.*, 2012). This acetylation is facilitated by different carbon sources that generate high amounts of acetyl-CoA upon their degradation. However, we hypothesize that this carbon source dependent acetylation of the RNA polymerase or one of the sigma factors does not play a role in regulation of *glmY* expression. First, no major differences in *glmY'*-*lacZ* expression levels were observed upon growth in minimal medium with glycerol, glucose or succinate (Hoffmann, Göpel and Görke, unpublished). And second, artificial elevation of the acetyl-CoA levels by disrupting the AckA/Pta pathway did not increase *glmY'*-*lacZ* expression but led to a ~four-fold decrease in *glmY* promoter activity (Fig. 3.8 B). Furthermore, over-expression of *cobB* affects the  $\sigma^{54}$ -dependent *glmY* promoter, whereas the *zraP* promoter is not affected (Fig. 3.6 A and B). Thus, we hypothesized that a regulatory factor unique for the *glmY* promoter is acetylated and thereby enhances  $P_{glmY}$  promoter activity.

Interestingly, RapZ was recently suggested to be acetylated in a proteomics study (Zhang *et al.*, 2009). The acetylated residue was proposed to be K251, a lysine residue in close proximity to the predicted C-terminal RNA-binding domain in RapZ of *Enterobacteriaceae* (Fig. 3.5, Göpel *et al.*, 2013; Zhang *et al.*, 2009). Indeed, by western blot, we could demonstrate that RapZ is acetylated *in vivo* when grown in complex medium (Fig. 3.8 C). However, acetyl transferase YfiQ (Pat) is not involved in the acetylation process (Fig. 3.7 F). In contrast, acetylation of RapZ and thus activity of the *glmY* promoter seems to be dependent on acetyl-P generated by the AckA-Pta pathway (Fig. 3.8 A and B). Response regulator GlrR was suggested to auto-phosphorylate with acetyl-P *in vitro*, since treatment with acetyl-P enhanced binding of a DNA fragment encompassing the *glmY* promoter (Göpel *et al.*, 2011). Thus, the inhibitory effect of disruption of the AckA-Pta pathway on  $P_{glmY}$  promoter activity might in part be due to reduced auto-phosphorylation of GlrR with acetyl-P. So far, there is no evidence that GlrR might be acetylated *in vivo*. Western blot experiments assessing the acetylation state of factors involved in regulation of *glmY* transcription, i.e. integration host factor, GlrR, GlrK and RapZ identified only RapZ as acetylated, whereas no signals were obtained for any of the other factors (Hoffmann and Görke; Göpel and Görke, unpublished).

Nevertheless, over-expression of *rapZ* restored high  $P_{glmY}$  activity as monitored by a *glmY'-lacZ* reporter gene fusion even in cells lacking *ackA* and *pta* suggesting that RapZ is the limiting factor (Fig. 3.8 B). We therefore tested whether RapZ is able to catalyze its auto-acetylation with acetyl-P *in vitro*. Indeed, we found that the acetylation level of RapZ increased over time when incubated with acetyl-P as compared to a mock treated control (Fig. 3.8 D). Recently, a study by Weinert and colleagues reported extensive chemical acetylation by acetyl-P in growing cells and accumulation of acetylated proteins stationary phase (Weinert *et al.*, 2013). This process was further shown to be independent of acetyl-transferase YfiQ. Our observation that RapZ is chemically acetylated by acetyl-P *in vitro* is thus in complete agreement with the notion that the majority of acetylation *in vivo* is dependent on acetyl-P rather than acetyl-CoA (Weinert *et al.*, 2013). Thus chemical acetylation may add a novel mode of action to our understanding of protein modifications in bacteria. However, the impact of acetylation with acetyl-P *in vivo* remains to be determined in future studies.

Clearly, stationary phase induction of GlmY by such an elaborate mechanism serves a regulatory purpose. However, we can only speculate what additional function(s) GlmY might have in stationary phase. Since *glmY* is the only gene regulated by the GlrK/GlrR two-component system in *E. coli*, it is likely that additional functions of GlmY are directly related to stimuli sensed by GlrK. So far, the signal sensed by GlrK is unknown, but the GlrK/GlrR two-component system seems to be involved in virulence in some pathogenic bacteria. For example, *Y. pseudotuberculosis* strains deficient in GlrR exhibit a reduced pathogenicity in mice (Flamez *et al.*, 2008) and in enterohemorrhagic *E. coli* (EHEC)

GlrK and GlrR (alternative names: QseF/QseE) indirectly activate transcription of the gene encoding the EspFU effector protein that is injected into the host cell. Hence, loss of the two-component system results in an inability to form attaching and effacing lesions and greatly reduces virulence of EHEC (Reading *et al.*, 2007; Reading *et al.*, 2009). Thus, in pathogenic bacteria, GlrK and GlrR seem to be involved in regulation of cell contact. However, these effects might be due to a more general function of this system. In agreement, *glrK* and *glrR* are conserved in a variety of species; most of them are non-pathogenic enterobacteria (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011). Therefore, the stimulus sensed by GlrK is likely to be more general than a virulence signal. It is tempting to speculate that the system senses membrane integrity possibly indirectly via outer membrane protein YfhG and thus might have a role in membrane remodeling and homeostasis. Interestingly, we found that YfhG is strictly required for the activity of the GlrK/GlrR two-component system since deletion of *yfhG* almost completely abolished transcription from the  $\sigma^{54}$ -dependent *glmY* promoter (Fig. 3.1 A). Consistently, complementation with *yfhG* on a plasmid enhanced  $\sigma^{54}$ -driven *glmY* expression by two-fold (Fig. 3.1 B). Further, a recently conducted study found that cells lacking *glmY* showed a 40% increase in sensitivity towards cell envelope stress (Hobbs *et al.*, 2009). This might possibly explain the drastic effect of *glrR*/*glrK* deletions in pathogens that rely on cell contact with their host cells and underlines the importance of sRNA(s) GlmY (and GlmZ) for these processes. Thus the next step is to identify the signal sensed by GlrK in order to draw conclusions as to further functions of GlmY and thoroughly investigate the implication of this system membrane homeostasis and in regard to virulence aspects involving host cell contact. Further studies dealing with these questions are currently subject of our investigations.

#### ACKNOWLEDGEMENTS

Sabine Lentjes is thanked for excellent technical assistance and Jörg Stülke for laboratory space. We are thankful to Jörg Vogel for the gift of Hfq antiserum and Heinz Neumann for the gift of *X. laevis* histone H3 and histone H3 K56Ac. Denise Lüttmann is thanked for construction of plasmid pDL43 and Tina Hollerbuhl is acknowledged for construction of the non-polar *yfhG* deletion mutant. Alan J. Wolfe is acknowledged for providing strains. Y.G. received a fellowship of the Dorothea Schlözer Program of the Göttingen University and AC.K. was supported by a stipend of the Studienstiftung des Deutschen Volkes. This work was supported by grants of the DFG priority program SPP1258 "Sensory and regulatory RNAs in prokaryotes" to B.G..

## Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA

*The results described in this chapter are published in:*

Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J., and Görke, B. (2013) Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev.* **27**: 552-564

### *Author contributions:*

This study was designed by Y.G. and B.G.. B.R. performed the Northern Blot in RNase E temperature sensitive strains. Y.G. performed all Northern Blots except for the Blots from RNase E temperature sensitive strains, RNase E cleavage assays, Gelretardation experiments, growth experiments and  $\beta$ -galactosidase measurements, structure probing and western blot analysis. Far-western experiments and BACTH assays were performed by Y.G.. B.G. and Y.G. conducted sequence alignment analysis. K.P. and J.V. provided insightful discussion and advice for preparation of the manuscript. Deepsequencing analyses were performed by Cynthia M. Sharma in the framework of the DFG priority program SPP 1258. Strains were constructed by B.R. and Y.G., plasmids were constructed by B.R. and Y.G. with assistance by Sabine Lentjes or intern students under the supervision of B.R. or Y.G.. Y.G. and B.G. wrote the paper.

## ABSTRACT

Bacterial small RNAs (sRNAs) are well-established to regulate diverse cellular processes but how they themselves are regulated is less understood. Recently, we identified a regulatory circuit wherein the GlmY and GlmZ sRNAs of *Escherichia coli* act hierarchically to activate mRNA *glmS* which encodes glucosamine-6-phosphate (GlcN6P) synthase. Although the two sRNAs are highly similar, only GlmZ is a direct activator that base-pairs with the *glmS* mRNA aided by protein Hfq. GlmY, however, does not bind Hfq and activates *glmS* indirectly by protecting GlmZ from RNA cleavage. This complex regulation feedback-controls the levels of GlmS protein in response to its product GlcN6P, a key metabolite in cell wall biosynthesis. Here, we reveal the molecular basis for the regulated turnover of GlmZ, identifying RapZ (formerly YhbJ) as a novel type of RNA-binding protein that recruits the major endoribonuclease RNase E to GlmZ. This involves direct interaction of RapZ with the catalytic domain of RNase E. GlmY binds RapZ through a secondary structure shared by both sRNAs and therefore acts by molecular mimicry as a specific decoy for RapZ. Thus, in analogy to regulated proteolysis, RapZ is an adaptor and GlmY an anti-adaptor in regulated turnover of a regulatory small RNA.

## INTRODUCTION

Bacterial small non-coding RNAs (sRNAs) function as regulators of gene expression in numerous diverse physiological circuits in response to changing internal or environmental cues such as metabolite concentrations or stresses (Gottesman and Storz, 2011; Richards and Vanderpool, 2011; Storz *et al.*, 2011). Therefore, the activities of sRNAs must be tightly controlled, which can occur at the level of biogenesis or decay, or both. In the past decade it became evident that sRNA expression is elaborately regulated at the transcriptional level and cognate transcriptional regulators have been identified for an increasing number of sRNA genes (Chao *et al.*, 2012; Göpel *et al.*, 2011; Holmqvist *et al.*, 2012). Many global regulators including alternative sigma factors and two-component systems employ sRNAs to control target gene expression indirectly, providing further flexibility and fine-tuning in regulation (Gogol *et al.*, 2011; Göpel and Görke, 2012a). By contrast, it is much less understood how sRNAs are regulated at the level of decay.

Key factors in sRNA degradation are the double-strand specific endoribonuclease RNase III, the single-strand specific endoribonuclease RNase E and the 3'→5' exoribonuclease polynucleotide phosphorylase (PNPase) (Arraiano *et al.*, 2010; Storz *et al.*, 2011; Vogel and Luisi, 2011). RNase E provides the scaffolding core of the RNA degradosome protein complex, which also contains helicase RhlB, enolase and PNPase. The degradosome promotes bulk RNA turn-over (Górna *et al.*, 2012). RNase E, and less frequently RNase III, are also responsible for sRNA-mediated gene silencing through initiation of target mRNA degradation induced by base-pairing with the sRNA, which is often co-degraded (Bandyra *et al.*, 2012; Caron *et al.*, 2010; Figueroa-Bossi *et al.*, 2009; Massé *et al.*, 2003).

Several sRNAs are cleaved by RNase E or RNase III independent of their targets. Initial processing may facilitate further degradation of the sRNA by enzymes such as PNPase (Andrade *et al.*, 2012) or generate mature sRNA species with higher regulatory activity either as a consequence of increased stability or increased affinity for the target mRNA (Davis and Waldor, 2007; Papenfort *et al.*, 2009; Soper *et al.*, 2010). For micro-RNAs, the eukaryotic counterparts of bacterial sRNAs, protein factors have been implicated in the targeted processing and degradation of selected microRNAs (Suzuki and Miyazono, 2011; Xhemalce *et al.*, 2012; Zisoulis *et al.*, 2012). Whether there are also specific protein co-factors for the degradation of individual Hfq-associated sRNAs by core endoribonucleases such as RNase E, is unknown. So far, only few specific sRNA binding proteins have been identified in bacteria (Pichon and Felden, 2007).

An intriguing example of a complex sRNA-regulated circuit with many input functions is the GlmY/GlmZ sRNA cascade of *E. coli*. GlmY and GlmZ are homologous sRNAs that jointly control synthesis of the metabolic enzyme glucosamine-6-phosphate (GlcN6P) synthase GlmS (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban *et al.*, 2007; Urban and Vogel, 2008). GlmS catalyzes synthesis of GlcN6P, which is the key reaction of the amino sugar pathway providing precursors for assembly of the cell wall and the outer membrane. The ~207 nt long sRNA GlmZ base-pairs with an anti-Shine-Dalgarno sequence in the *glmS* mRNA, thereby promoting translation and concomitant stabilization of the mRNA. GlmZ is inactivated by processing to a shorter variant of 151 nt which lacks the *glmS* interaction site. The related 184 nt long GlmY sRNA is also subject to processing at its 3' end by a so far unknown enzyme resulting in a variant of 147 nt. Processed GlmY accumulates upon decreasing intracellular concentrations of GlcN6P and inhibits processing of GlmZ thus activating *glmS* indirectly. Thus, GlmY and GlmZ operate hierarchically to attune *glmS* expression thereby mediating GlcN6P homeostasis.

Genetic analysis identified YhbJ, a protein of unknown function, as an additional factor in this regulatory cascade (Kalamorz *et al.*, 2007; Urban and Vogel, 2008). YhbJ positively controls processing of GlmZ. In the absence of YhbJ, GlmZ is not processed, whereas overproduction of YhbJ increases the fraction of processed GlmZ, suggesting that YhbJ is limiting for GlmZ processing (Fig. S4.1 A). GlmY has no impact on GlmZ and *glmS* in a *yhbJ* mutant suggesting that GlmY acts through YhbJ (Fig. S4.1 B, (Reichenbach *et al.*, 2008; Urban and Vogel, 2008)). The high similarity of both sRNAs in sequence and structure suggested that GlmY possibly acts by a mimicry mechanism on GlmZ, perhaps through targeting protein YhbJ.

In this work we provide molecular evidence that YhbJ acts as an adaptor protein guiding processing of GlmZ. YhbJ specifically binds both sRNAs at a conserved central stem loop. However, the ribonuclease responsible for GlmZ processing is RNase E rather than YhbJ itself. We find that YhbJ directly interacts with the catalytic domain of RNase E. *In vitro*, RNase E requires YhbJ for specific and

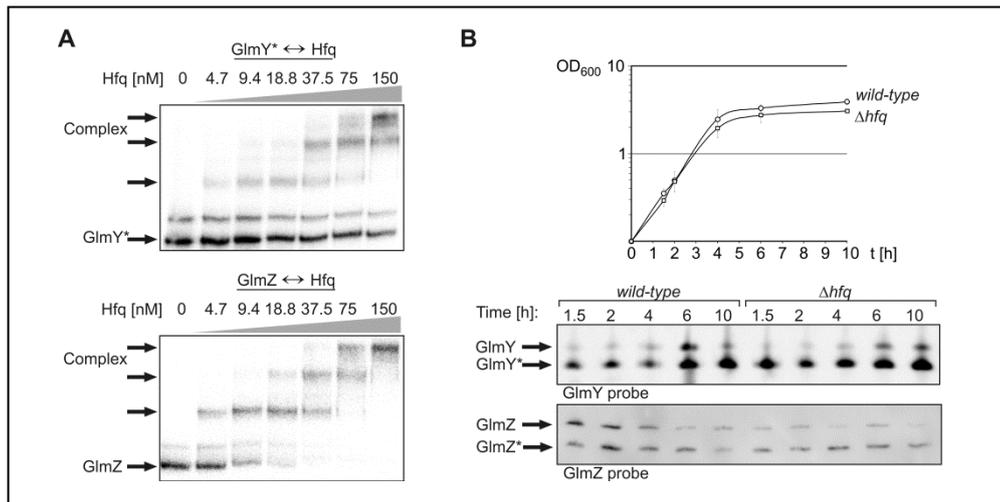
rapid processing of GlmZ, while GlmY inhibits this reaction through sequestration of YhbJ. Thus, YhbJ is a novel type of sRNA binding protein that targets GlmZ for degradation by RNase E. This process is counteracted by the decoy sRNA GlmY mediating hierarchy within the GlmY/GlmZ cascade. By analogy, this regulatory circuit bears strong resemblance to the mechanism of controlled proteolysis, in which a regulatory protein is delivered to its degrading protease by an adaptor protein, while anti-adaptor proteins counteract this process through sequestration of the adaptor (Bougdour *et al.*, 2008). In this scenario, YhbJ serves as an adaptor recruiting GlmZ to RNase E and sRNA GlmY acts as an anti-adaptor. Thus, we propose renaming of YhbJ as RapZ (RNase adaptor protein for sRNA GlmZ).

## RESULTS

### ***Unlike GlmZ, GlmY is not an Hfq-dependent small RNA and presumably acts by protein binding***

Although GlmY and GlmZ are very similar sRNAs, they use different mechanisms to regulate *glmS*. GlmZ activates *glmS* through base-pairing. Many base-pairing sRNAs require the hexameric RNA chaperone Hfq for functionality, at least in Gram-negative bacteria (Vogel and Luisi, 2011). Hfq facilitates annealing of cognate sRNA/mRNA pairs and stabilizes the resulting duplexes. Moreover, base-pairing sRNAs are often destabilized in *hfq* mutants, while protein binding sRNAs do not show these properties (Urban and Vogel, 2007). Recent coIP experiments identified GlmZ as a sRNA associating with Hfq, whereas GlmY could not be detected in these Hfq pull-down assays (Chao *et al.*, 2012). GlmY regulates *glmS* indirectly in a process that is likely to involve RapZ. However, GlmY does not control synthesis of RapZ since neither absence nor overexpression of GlmY had any significant effect on the amount of RapZ (Fig. S4.2). Collectively, these observations suggest that GlmY may not be an Hfq-binding base-pairing sRNA but uses a distinct molecular mechanism.

To corroborate this conclusion, we determined Hfq binding affinities for GlmY and GlmZ *in vitro* and examined GlmY and GlmZ stabilities in an *hfq* mutant. *Trans*-encoded base-pairing sRNAs display high binding affinities for Hfq exhibiting  $K_d$ s in the range of 0.1-10 nM, whereas non-specific RNAs have significantly lower affinities (Olejniczak, 2011; Panja and Woodson, 2012). Indeed, GlmZ was efficiently bound by purified Hfq *in vitro* (apparent  $K_d$ ~10 nM), while ~15-fold higher Hfq concentrations were required for binding of GlmY indicating unspecific binding (Fig. 4.1 A).



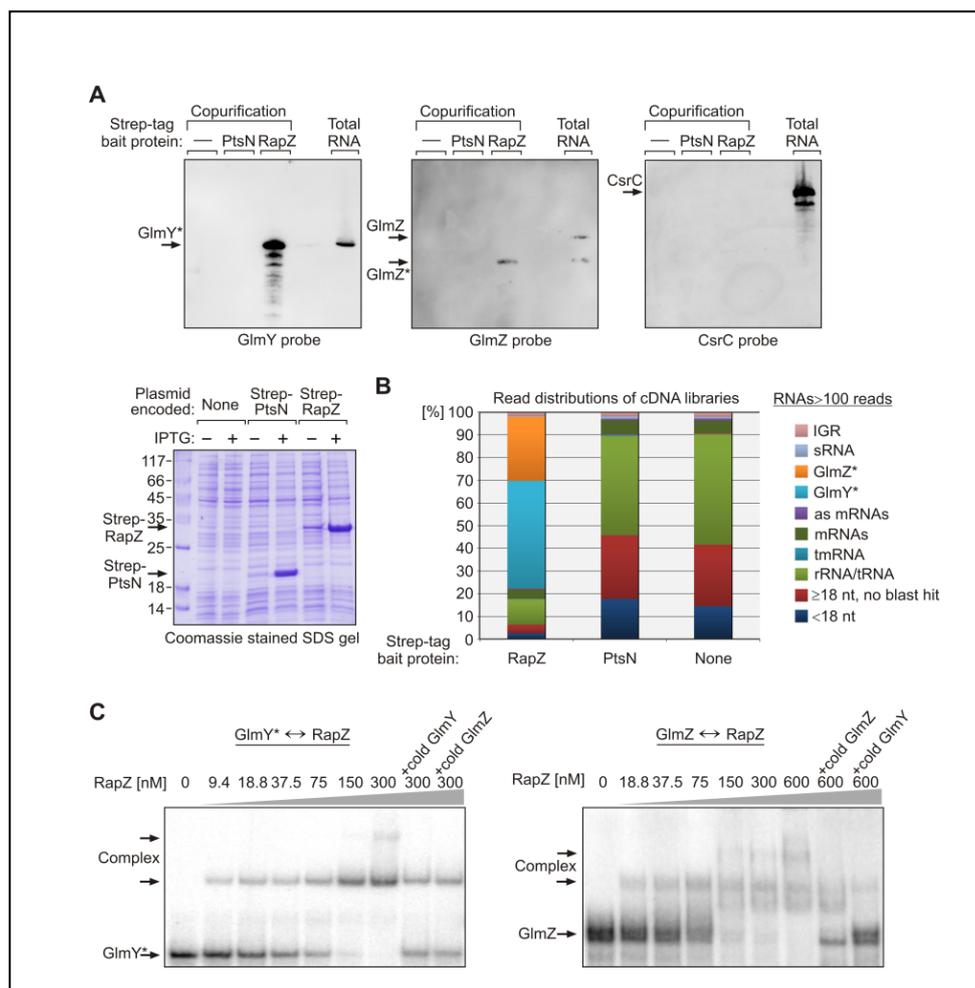
**Figure 4.1. GlmY is not an Hfq-binding sRNA.** (A) Hfq binds GlmZ, but not GlmY, with high affinity *in vitro*. EMSA using  $\alpha$ -<sup>32</sup>P-UTP labeled GlmY (top) or GlmZ (bottom) sRNAs and various concentrations of purified Hfq as indicated. (B) GlmZ, but not GlmY, is destabilized in an *hfq* mutant. The processed forms of the sRNAs are marked with an asterisk in this and all other figures. Strain R1279 (wild-type) and the  $\Delta hfq$  mutant Z664 were grown in LB (top). Total RNA was isolated from samples harvested at various time intervals and analyzed by Northern blotting using the indicated probes. As loading controls, the membranes were re-probed against 5S rRNA (Fig. S4.3).

Consistently, the amount of the full-length (i.e. base-pairing) variant of GlmZ was significantly reduced in the *hfq* mutant, whereas GlmY was not affected (Fig. 4.1 B). These results indicate that GlmY does not act by Hfq-assisted base-pairing to prevent processing of GlmZ, but might act by targeting of a protein for which the most likely candidate is RapZ.

#### **RapZ specifically binds small RNAs GlmY and GlmZ**

To explore whether RapZ binds GlmY and perhaps also GlmZ, copurification experiments were performed. Recombinant RapZ carrying the Strep-tag epitope at its N-terminus was overexpressed in wild-type cells (Fig. 4.2 A, lower panel). Cells carrying the empty bait vector or a bait vector expressing the Strep-tagged phosphotransferase protein PtsN served as negative controls. Cell extracts were subjected to StrepTactin affinity chromatography and recombinant proteins were eluted (Fig. S4.4). Copurifying RNA was extracted from the elution fractions and analyzed by Northern blot (Fig. 4.2 A, top panels). Notably, GlmY as well as GlmZ, both in their processed forms, were recovered with Strep-RapZ, but undetectable in the control samples. The abundant protein-binding sRNA CsrC, could not be detected in any of the analyzed samples, suggesting that RapZ does not bind every sRNA. To corroborate that RapZ binds GlmY and GlmZ specifically, the copurified RNAs were analyzed by deep sequencing (Fig. 4.2 B). Blast analysis of the sequences of the RapZ-derived cDNA library yielded 43644 mapped reads, 50% of which corresponded to GlmY and 30% to GlmZ (Fig. 4.2 B; Fig. S4.5, Excel file S4.1). Thus, GlmY and GlmZ were approximately 1000-fold enriched by Strep-RapZ copurification as compared to the negative controls. These results demonstrated that

GlmY and GlmZ are the main, if not the sole targets of RapZ. Finally, we tested binding of GlmY and GlmZ by purified RapZ *in vitro* by gel retardation analysis.



**Figure 4.2. RapZ binds GlmY as well as GlmZ *in vivo* and *in vitro*.** (A) GlmY and GlmZ copurify with Strep-RapZ. Transformants of strain R1279 carrying either plasmid pBGG237 (empty vector), plasmid pBGG217 encoding Strep-PtsN (MW = 19.29 kDa) or plasmid pBGG164 encoding Strep-RapZ (MW = 34.04 kDa) were grown in LB and IPTG was added to induce synthesis of recombinant proteins. To verify overproduction of Strep-PtsN and Strep-RapZ, total protein extracts of samples collected before and 1 h after addition of IPTG were separated on 12.5 % SDS-polyacrylamide gels and gels were analyzed by staining with Coomassie blue (lower panel). The cell extracts were passed through StrepTactin columns resulting in purification of the Strep-tagged proteins (Fig. S4.4). Co-eluting RNA was isolated and subjected to Northern analysis using probes specific for GlmY (left), GlmZ (middle) and CsrC (right). 5  $\mu$ g total RNA of strain R1279 served as positive control (last lane, respectively). (B) The majority of RNA molecules copurifying with Strep-RapZ correspond to GlmY and GlmZ. The RNA preparations isolated in the copurification experiments described in (A) were converted to cDNA and analyzed by 454 pyrosequencing. The relative distribution of reads mapping to the different RNA categories is shown in 100% stacked column charts. (C) RapZ binds GlmY and GlmZ *in vitro*. EMSAs using  $\alpha$ - $^{32}$ P-UTP labeled GlmY (left) or GlmZ (right) sRNAs and various concentrations of purified Strep-RapZ as indicated. In the last two lanes 120 nM unlabeled GlmY or GlmZ corresponding to a 30-fold excess over the labeled sRNAs was added as a competitor.

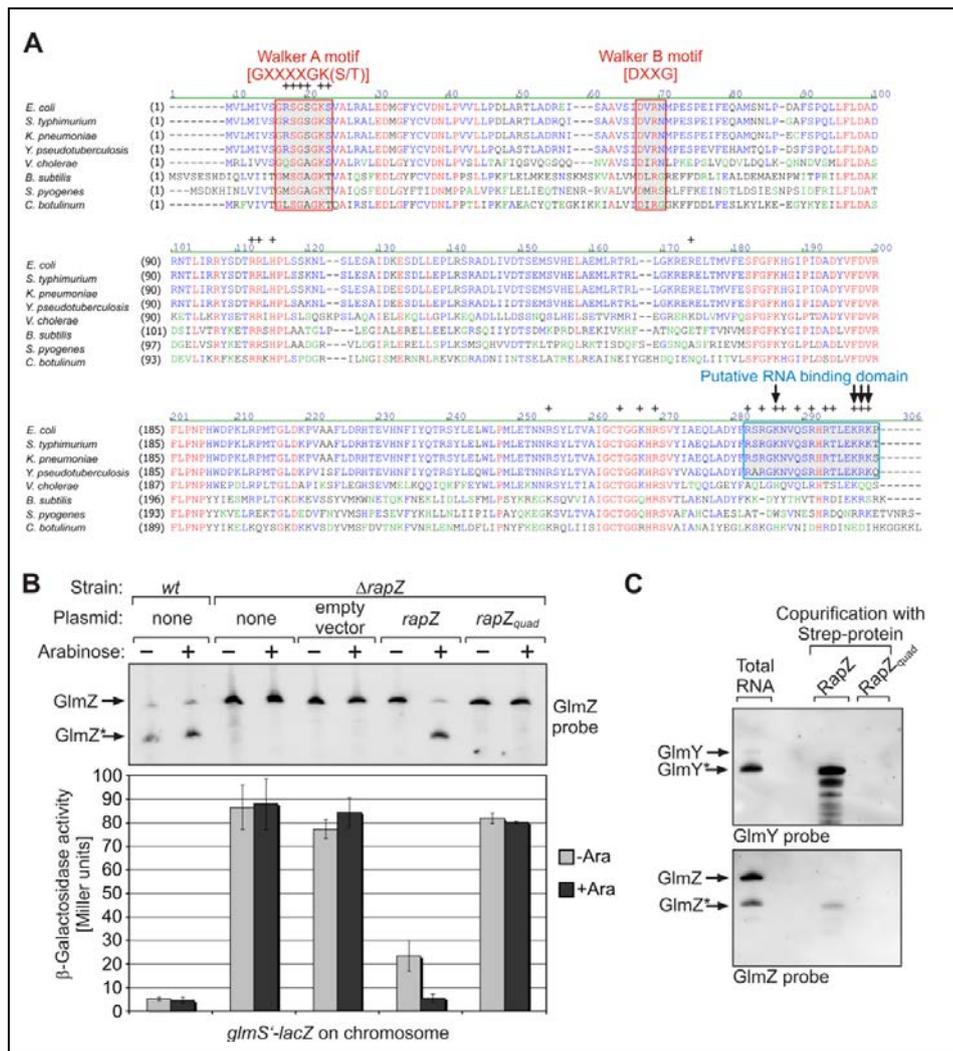
Both sRNAs were efficiently bound by RapZ with an apparent  $K_d$  of  $\sim$ 30 nM for GlmY and  $\sim$ 75 nM for GlmZ (Fig. 4.2 C). Addition of un-labeled GlmY or GlmZ competitor RNA decreased complex

formation. RapZ bound the full-length and the processed form of each sRNA with similar affinity (Fig. S4.6), suggesting that the 3' ends of the sRNAs are dispensable for efficient binding. A control experiment using the unrelated sRNA MicF revealed complex formation with RapZ only at the highest protein concentrations reflecting unspecific binding (Fig. S4.7). Collectively, the results show that RapZ specifically binds GlmY and GlmZ *in vivo* and *in vitro* with high affinity. Moreover, RapZ appears to have a higher affinity for GlmY than for GlmZ.

***RapZ interacts with the central stem loops of GlmY and GlmZ presumably via its C-terminal end***

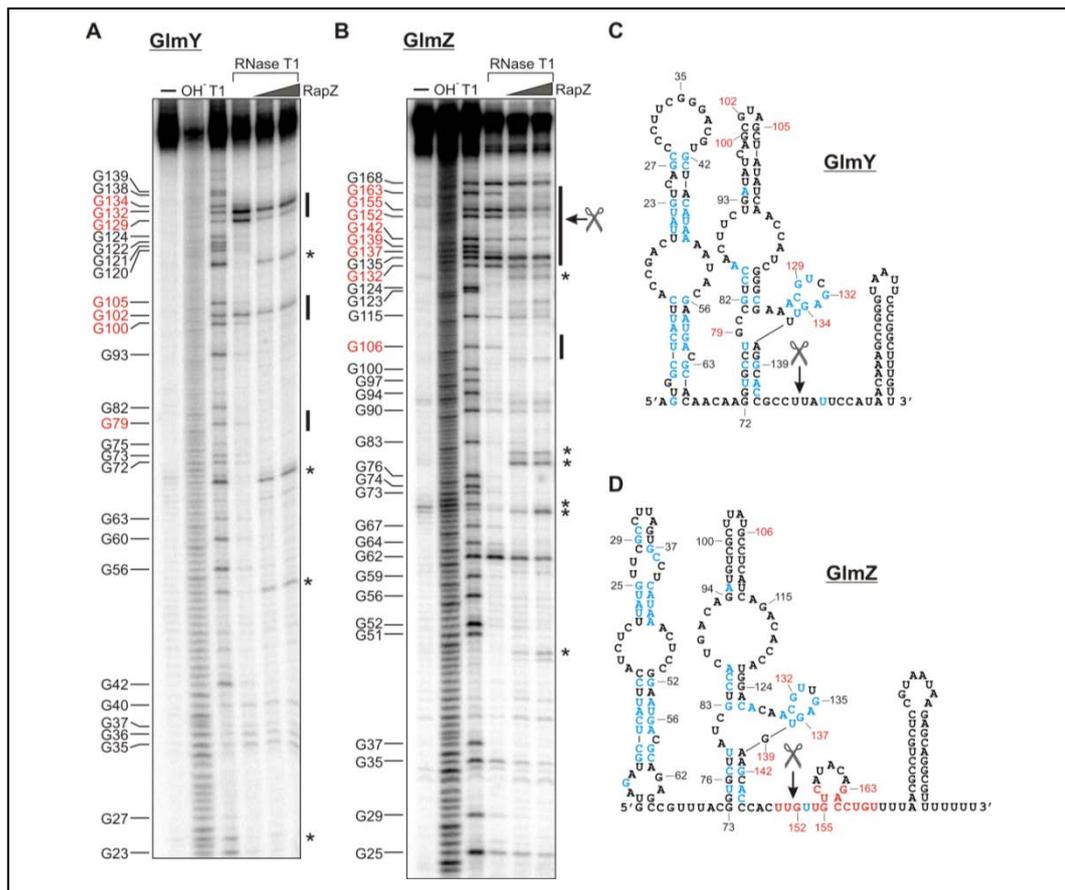
RapZ is an uncharacterized protein that shows no extended homology to other proteins; the only discernable motifs are a Walker A and a Walker B motif, which allow binding of GTP or ATP (Luciano *et al.*, 2009). We used the software tool BindN to predict potentially RNA-binding residues in RapZ. The analysis revealed clustering of candidate residues in the C-terminus of RapZ (Fig. 4.3 A; supplemental experimental procedures). Sequence alignment analysis of RapZ homologs from various bacteria showed that the sequence of this region is highly conserved in *Enterobacteriaceae*, which possess the GlmY/GlmZ sRNAs, but deviates in those bacteria that lack these sRNAs (Fig. 4.3 A; Fig. S4.8; (Göpel *et al.*, 2011)).

To determine whether the C-terminal region of RapZ is required for binding of the sRNAs we substituted the residues K270, K281, R282 and K283 with alanine, resulting in a quadruple mutant (subsequently designated RapZ<sub>quad</sub>; Fig. 4.3 A). Complementation analysis revealed that these mutations abrogated the function of RapZ *in vivo*. In a  $\Delta rapZ$  mutant unprocessed GlmZ accumulates and up-regulates *glmS* expression as monitored by a *glmS'-lacZ* reporter fusion (Fig 4.3 B) and by Northern blotting (Fig. S4.10 A). Introduction of a plasmid carrying the wild-type *rapZ* gene under control of an arabinose-inducible promoter complemented the  $\Delta rapZ$  mutant, while this was not the case when using an isogenic construct expressing RapZ<sub>quad</sub> (Fig. 4.3 B). A Western blot confirmed that the RapZ<sub>quad</sub> protein is expressed and stable (Fig. S4.10 B). Next, copurification experiments were carried out to determine whether the inactivity of RapZ<sub>quad</sub> *in vivo* is caused by an inability to bind the sRNAs. Strep-tagged RapZ<sub>quad</sub> and wild-type Strep-RapZ were purified by StrepTactin affinity chromatography (Fig. S4.11) and copurifying RNAs were analyzed by Northern blot (Fig. 4.3 C). Once again, GlmY and GlmZ efficiently copurified with the wild-type protein, but not with the quadruple mutant. Finally, purified RapZ<sub>quad</sub> also failed to bind GlmY and GlmZ in gel retardation experiments (Fig. S4.12). Therefore, the RapZ C-terminus is required for sRNA binding suggesting that this region is involved in RNA-binding.



**Figure 4.3.** RapZ variant carrying the quadruple mutation K270A-K281A-R282A-K283A in the supposed RNA-binding domain fails to bind GlmY and GlmZ. (A) Alignment of RapZ homologs. Putative RNA-binding amino acid residues (marked with "+") predicted by BindN, cluster in the C-termini of enterobacterial RapZ proteins (boxed in blue). The positions mutated in the quadruple mutant RapZ<sub>quad</sub> are indicated by arrows. Nucleotide binding Walker A and B motifs are boxed in red. See legend to Fig. S4.8 for further information (B) The RapZ quadruple mutant (RapZ<sub>quad</sub>) does not complement a  $\Delta rapZ$  mutation, as monitored by Northern blot (top panel) and by the expression of a *glmS'-lacZ* fusion (histogram). Strains Z8 (wild-type) and Z28 ( $\Delta rapZ$ ) were employed, which carry a *glmS'-lacZ* reporter fusion on the chromosome. Additionally, transformants of strain Z28 were tested, which contained plasmids carrying either wild-type *rapZ* (plasmid pBGG61) or the mutant *rapZ<sub>quad</sub>* gene (plasmid pYG30) or no insert (empty vector pBAD33) under control of an arabinose-inducible promoter. Cells were grown in LB in the absence or presence of arabinose as indicated and the  $\beta$ -galactosidase activities were determined. In addition, total RNAs were isolated and subjected to Northern blotting using a probe directed against GlmZ. The loading controls are provided in Fig. S4.9. (C) GlmY and GlmZ do not copurify with RapZ<sub>quad</sub>. A copurification experiment as shown in Fig. 4.2 A was carried out. Transformants of strain R1279 overproducing either Strep-RapZ<sub>quad</sub> (encoded on plasmid pYG29) or Strep-RapZ (encoded on plasmid pBGG164; positive control) were tested. The copurifying RNA was analyzed by Northern blot using probes specific for GlmY (top) or GlmZ (bottom). Total RNA of strain R1279 served as positive control (first lane, respectively).

To determine which regions in the GlmY and GlmZ sRNAs are bound by RapZ, we performed structural probing of the sRNAs by limited RNase T1 digestion in the absence and presence of RapZ. RNase T1 cleaves single-stranded RNA downstream of guanosine residues. The sRNAs, after 5' end labeling with <sup>32</sup>P, were partially digested by RNase T1 in their denatured (Fig.4.4 A and B, lanes 3) as well as in their native forms (lanes 4).



**Figure 4.4. Identification of RapZ binding sites in GlmY and GlmZ by RNase T1 protection assay.** 5'-end labeled GlmY (A) and GlmZ (B) sRNAs were subjected to partial RNase T1 cleavage in the absence (lanes 4) and presence of 80 nM (lanes 5) and 300 nM RapZ (lanes 6). The untreated RNAs were loaded in lanes 1. The ladders obtained by alkali treatment were separated in lanes 2. The RNase T1 ladders of the denatured sRNAs were separated in lanes 3. The positions of the cleaved G residues are given at the left of the gels. The positions of G residues that become protected from RNase T1 cleavage by RapZ are indicated in red. Asterisks indicate unspecific cleavage. In (C) and (D) the positions of the G residues that are protected by RapZ are depicted in the structures of GlmY and GlmZ (labeled in red). Residues labeled in blue are conserved in both GlmY and GlmZ of various enterobacterial species (Göpel *et al.*, 2011). Processing sites are indicated by pairs of scissors.

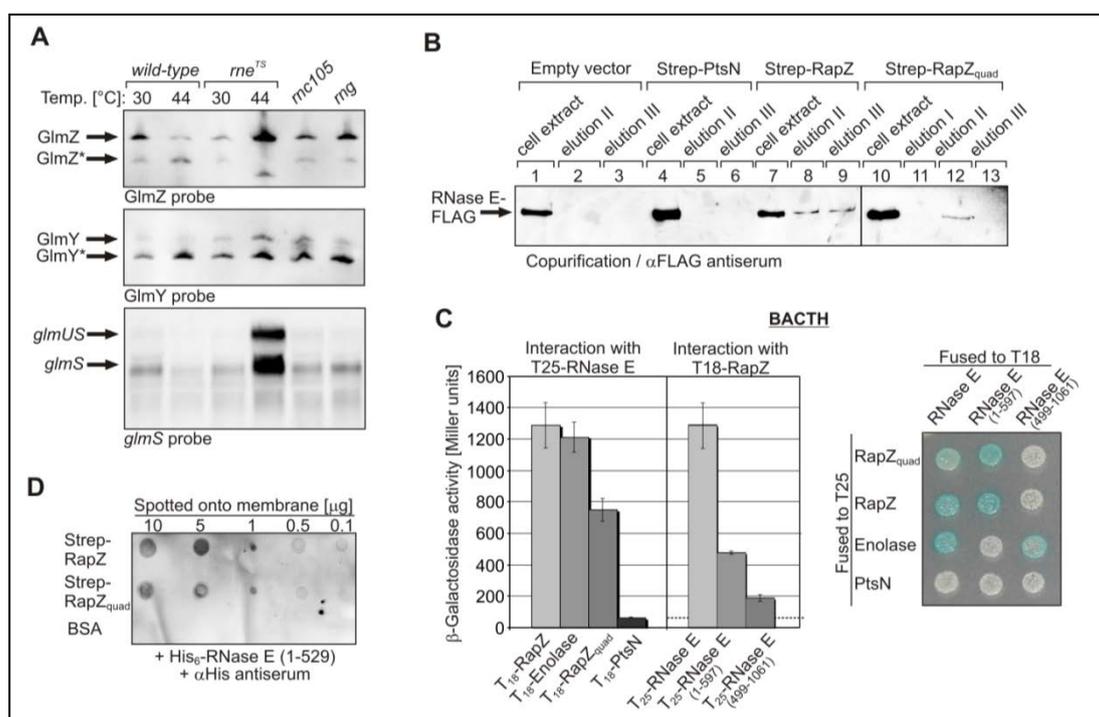
Comparison of the cleavage patterns generated by RNase T1 generally confirmed our previous structure predictions for GlmY and GlmZ (Fig. 4.4 C and D; (Reichenbach *et al.*, 2008; Urban and Vogel, 2008)). In the presence of 40 nM or 300 nM RapZ (Fig. 4.4 A and B, lanes 5 and 6, respectively) several residues in GlmY and GlmZ became partially or totally protected from cleavage. These residues were located in the central stem loop structures, in particular in the conserved lateral bulges of the sRNAs. In addition, residues in the vicinity of the processing site became protected in GlmZ. These data suggest that RapZ binds both sRNAs at similar structural motifs.

#### **RapZ physically interacts with RNase E, which is responsible for processing of GlmZ**

Previous *in vivo* evidence suggested that RapZ somehow triggers processing and thereby inactivation of GlmZ (Kalamorz *et al.*, 2007; Urban and Vogel, 2008). However, the *in vitro* experiments above did not yield any evidence that RapZ itself could be this processing enzyme (Fig. 4.2 C and Fig. 4.4 B).

Therefore, we searched for ribonucleases required for processing of GlmZ. To this end, we investigated the fates of GlmY, GlmZ and *glmS* transcripts in mutants defective for the major endoribonucleases RNase E, RNase III and RNase G by Northern blot analysis (Fig. 4.5 A). Since RNase E is essential, a temperature-sensitive mutant was employed, which becomes non-permissive at 44°C. While none of the RNase mutations substantially affected GlmY, processing of GlmZ was specifically inhibited upon inactivation of RNase E. As expected, this resulted in concomitant up-regulation of the *glmS* mRNA and the *glmUS* co-transcript, which undergoes processing by RNase E at the *glmU* stop codon (Joanny *et al.*, 2007; Kalamorz *et al.*, 2007). In conclusion, RNase E, and not RNase III as suggested previously (Argaman *et al.*, 2001), is required for processing of GlmZ *in vivo*.

Next, we asked whether the stimulatory effect of RapZ on RNase E-mediated processing of GlmZ might involve a physical interaction of RapZ and RNase E. Therefore, we tested whether a chromosomally encoded RNase E-FLAG variant copurified together with Strep-RapZ upon StrepTactin affinity chromatography. RNase E-FLAG was indeed detectable in the elution fractions of the Strep-RapZ purification, but absent from control elutions carried out with extracts of cells that expressed the Strep-PtsN protein or carried the empty bait vector (Fig. 4.5 B and Fig. S4.14). To confirm these results, bacterial two-hybrid assays were performed using the BACTH system, which relies on reconstitution of the activity of split adenylate cyclase CyaA in *E. coli* (Karimova *et al.*, 1998). Indeed, high  $\beta$ -galactosidase activities, reflecting successful restoration of CyaA activity and therefore protein-protein interaction, were observed when RNase E and RapZ were fused to the C-termini of the T25- and T18-fragments of CyaA, respectively (Fig. 4.5 C, column 1; Fig. S4.15 A). Comparable enzyme activities were obtained when the already established interaction of enolase and RNase E was tested as a positive control. In contrast, interaction between PtsN and RNase E (negative control) could not be detected by the BACTH system, consistent with the results of the copurification approach (Fig. 4.5 C, columns 2, 4).



**Figure 4.5. RapZ interacts with RNase E, which is essential for processing of GlmZ, and this interaction is independent of the RNA-binding function of RapZ.** (A) Inactivation of RNase E abrogates processing of GlmZ in vivo. Northern blot analysis of total RNA isolated from strains N3433 (wild-type), N3431 (*rne<sup>TS</sup>*), IBPC633 (*rnc*) and IBPC935 (*rng*). Strain N3431 carries a temperature sensitive RNase E variant, which becomes inactivated upon a temperature shift from 30°C to 44°C. Northern blot analysis was performed using the indicated probes. The loading controls are provided in Fig. S4.13. (B) RNase E copurifies with RapZ independently of RNA-binding by RapZ. Strain Z64, which carries a FLAG-tagged *rne* gene (encoding RNase E-FLAG) on the chromosome, and additionally contained either plasmid pBGG237 (empty vector; lanes 1-3), plasmid pBGG217 encoding Strep-PtsN (lanes 4-6), plasmid pBGG164 encoding Strep-RapZ (lanes 7-9) or plasmid pYG29 encoding Strep-RapZ<sub>quad</sub> (lanes 10-13) was grown in LB containing 1 mM IPTG for induction of synthesis of the Strep-tagged proteins. The cell extracts were subjected to the copurification protocol using StrepTactin affinity chromatography (Fig S4.11). Presence of RNase E-FLAG in the eluates was tested by Western blotting using anti-FLAG antiserum. (C) Bacterial two-hybrid (BACTH) assays indicating interaction of the N-terminal part of RNase E with RapZ independent of its RNA-binding function. High β-galactosidase activities, which were monitored either quantitatively (left) or phenotypically on X-Gal agarose plates (right), reflect reconstitution of split adenylate cyclase CyaA activity through interaction of the proteins that are fused to its separately encoded T25- and T18-domains. The various plasmids that were tested in strain BTH101 are listed in Table S2. (D) Dot-blot far-Western indicating interaction of RapZ with the catalytic domain of RNase E *in vitro*. Various amounts of purified Strep-RapZ, Strep-RapZ<sub>quad</sub> and BSA (negative control) were spotted onto a membrane and subsequently incubated in 50 nM of the His<sub>6</sub>-tagged catalytic domain of RNase E. Interaction was visualized with an antiserum directed against the His-tag.

To narrow down the RapZ interaction site in RNase E, N- and C-terminally truncated RNase E variants were tested in the BACTH assay (Fig. 4.5 C, columns 5-7; Fig. S4.15 A): RapZ preferentially interacts with the N-terminal part of RNase E (residues 1-597), which comprises its catalytic domain. However, the less efficient interaction of RapZ with the N-terminal fragment of RNase E as compared to the full-length protein might indicate that residues in the RNase E C-terminus also contribute to this interaction. Control experiments demonstrated the expected interaction of enolase with the C-terminal but not with the N-terminal fragment of RNase E, thus verifying functionality of the T25-RNase E (499-1061) fusion protein (Fig. S4.15 B). In additional experiments all interactions were tested in the opposite orientation, i.e. the various RNase E variants were fused to the CyaA-T18

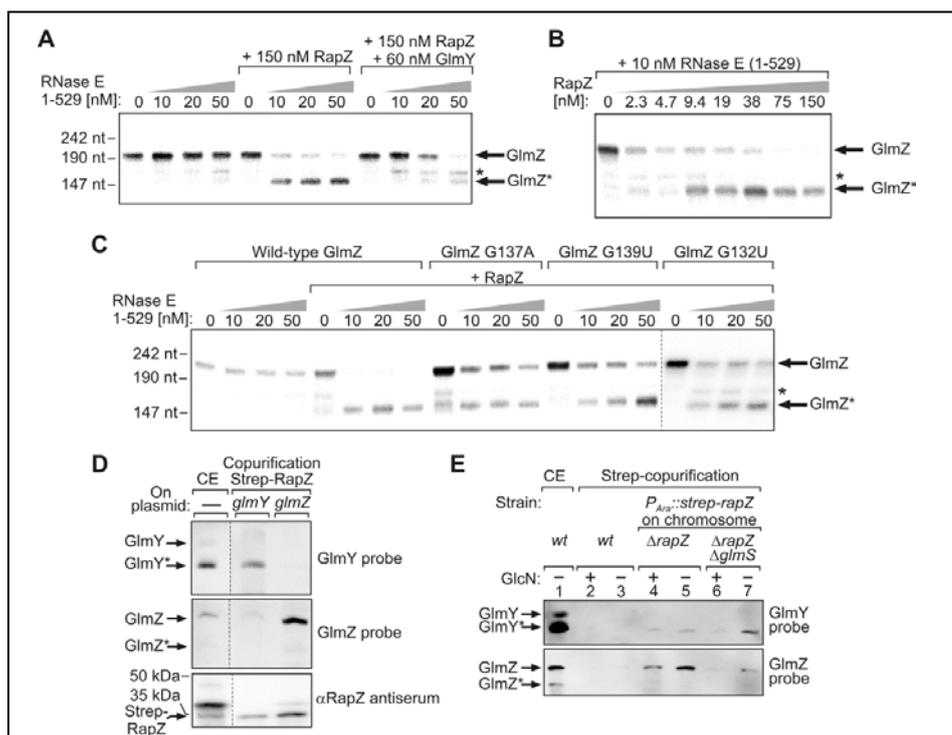
domain whereas its potential interaction partners were fused to CyaA-T25. Phenotypic monitoring of the  $\beta$ -galactosidase activities of the various transformants confirmed the results (Fig. 4.5 C, right): RapZ interacts with the N-terminal part of RNase E, while enolase binds to the RNase E C-terminus. To further corroborate these results, we tested interaction *in vitro* by dot-blot far-Western analysis using purified proteins (Fig. 4.5 D). Strep-RapZ was spotted in serial dilutions on a membrane, which was subsequently incubated in a solution containing the 6 $\times$ His-tagged catalytic domain of RNase E (residues 1-529). Incubation with  $\alpha$ His antiserum revealed interaction of the truncated RNase E with RapZ, but not with the negative controls BSA, Strep-PtsN or RhlB (Fig. 4.5 D and Fig. S4.16). RhlB is known to interact with the C-terminal part of RNase E (Górna *et al.*, 2012). Essentially, the same result was obtained when the reverse experiment was carried out, i.e. when membrane-immobilized His<sub>6</sub>-RNase E (1-529) was treated with a solution containing Strep-RapZ and complexes were detected with  $\alpha$ Strep antiserum (Fig. S4.17 A).

Next we analyzed whether the observed interaction between RapZ and the catalytic domain of RNase E was direct or whether it resulted from the simultaneous interaction of both proteins with GlmZ and thus required RNA. Therefore, we tested the RapZ<sub>quad</sub> mutant, which lacks sRNA binding activity (Fig. 4.3), in the various protein-protein interaction assays (Fig. 4.5 B-D, Fig. S4.15-S4.17). The data revealed that the RapZ<sub>quad</sub> protein still interacted with the catalytic domain of RNase E, although with a slightly decreased affinity as compared to wild-type RapZ. Together, our data suggest that RapZ directly interacts with the catalytic domain of RNase E in a RNA-independent manner.

### ***RapZ triggers processing of GlmZ by RNase E in vitro***

Our observations that RapZ interacts with GlmZ as well as with RNase E, and that both proteins are required for processing of GlmZ *in vivo*, raised the possibility that RapZ recruits GlmZ to RNase E for subsequent processing. To test this idea, processing of GlmZ by RNase E was studied *in vitro*. 40 nM of radio-labeled GlmZ were incubated with increasing concentrations of the purified catalytic domain of RNase E (Fig. 4.6 A). This procedure did not yield significant amounts of processed GlmZ species that matched the 151 nt fragment observed *in vivo* (Fig.4.6 A-C), suggesting that RNase E alone is not sufficient to mediate efficient GlmZ processing. Importantly, the presence of 150 nM purified RapZ in this assay promoted efficient processing of GlmZ to the product of expected size (Fig. 4.6 A). Additional assays were carried out using a fixed concentration of 10 nM RNase E and varying concentrations of RapZ (Fig. 4.6 B). These data revealed that RapZ triggered correct processing of GlmZ in a concentration dependent manner. An  $\sim$ 3-fold excess of RapZ over RNase E was required to obtain complete processing of GlmZ within the reaction time. Interestingly, Yhbj forms a homotrimer in solution (Resch *et al.*, 2013) suggesting a 3:1 stoichiometry of RapZ and RNase E in the complex. Moreover, nonspecific cleavage was suppressed towards higher concentrations of RapZ.

Our experiments suggested that RapZ makes contacts with the lateral bulge in the central stem loop of GlmZ (Fig. 4.4 B, D). Residues G132, G137 and G139 located in this bulge became protected from RNase T1 digestion in the presence of RapZ (Fig. 4.4 B). To determine, whether these residues are important for binding of GlmZ by RapZ and/or its processing by RNase E, we studied the effects of their individual mutation. Substitution of wild-type GlmZ with the corresponding GlmZ mutants caused strong up-regulation of *glmS* expression as inferred from expression of a *glmS'*-*lacZ* reporter gene fusion and Northern blot analysis (Fig. S4.18 A, B).



**Figure 4.6.** RapZ recruits GlmZ to processing by RNase E and GlmY counteracts this reaction through sequestration of RapZ. (A) RapZ promotes cleavage of GlmZ by RNase E *in vitro* and GlmY inhibits this process. *In vitro* cleavage assay of  $\alpha$ - $^{32}$ P-UTP labeled GlmZ using varying concentrations of the catalytic domain of RNase E in the absence (lanes 1-4) or presence of 150 nM Strep-RapZ (lanes 5-12). In lanes 9-12 unlabeled GlmY (60 nM) was additionally present. The asterisk indicates a non-specific cleavage product of GlmZ. (B) RapZ triggers specific processing of GlmZ. *In vitro* cleavage assay of  $\alpha$ - $^{32}$ P-UTP labeled GlmZ using 10 nM of the catalytic domain of RNase E and varying concentrations of Strep-RapZ. (C) Mutations in the lateral bulge inhibit processing of GlmZ by RNase E *in vitro*. *In vitro* cleavage assay of  $\alpha$ - $^{32}$ P-UTP labeled GlmZ variants carrying the indicated mutations. Varying concentrations of the catalytic domain of RNase E and 150 nM Strep-RapZ were added as indicated. (D) GlmY and GlmZ compete for binding to RapZ *in vivo*. The effects of plasmid-driven overexpression of GlmY and GlmZ on sRNA copurification with Strep-RapZ were addressed. Strain Z479, which carried the arabinose-inducible *strep-rapZ* gene on the chromosome, and either plasmid pYG23 for overexpression of *glmY* or plasmid pYG24 for overexpression of *glmZ* was grown in LB containing arabinose. The cell extracts were subjected to the copurification protocol using StrepTactin affinity chromatography resulting in purification of Strep-RapZ (Western blot, bottom panel). Co-eluting RNA was subjected to Northern analysis (top and medium panels). 5  $\mu$ g total RNA of strain Z479 served as positive control (first lane, respectively). The dotted line indicates cropping of lanes from the original blot. (E) Intracellular GlcN6P depletion in a *glmS* mutant shifts the proportion of the two sRNA that are bound to Strep-RapZ towards GlmY. Strains Z479 (lanes 4, 5) and the isogenic  $\Delta$ *glmS* mutant Z555 (lanes 6, 7) lack the authentic *rapZ* gene, but carry the arabinose-inducible *strep-rapZ* gene at an ectopic site. The strains were grown in LB containing GlcN and arabinose until an  $OD_{600} = 0.3$ . Subsequently, the cultures were split and growth was continued in the absence or presence of GlcN (Fig. S4.21 B). After 2 h the cultures were subjected to the copurification protocol, resulting in purification of Strep-RapZ (Fig. S4.23). The copurifying sRNAs were analyzed by Northern blotting. Wild-type strain R1279 carrying no *strep-rapZ* served as negative control (lanes 2, 3). 5  $\mu$ g total RNA of strain R1279 served as positive control (lane 1).

Gel retardation assays showed that the GlmZ mutants were still efficiently bound by RapZ similar to wild-type GlmZ (Fig. S4.18 C). Notably, when tested in the *in vitro* cleavage assay, processing of the GlmZ mutants by RNase E was strongly inhibited as compared to wild-type GlmZ, which was tested in parallel (Fig. 4.6 C). In conclusion, single mutations in the lateral bulge of GlmZ are not sufficient to abolish binding by RapZ, but they strongly impair processing by RNase E.

#### ***GlmY counteracts processing of GlmZ by acting as decoy for RapZ***

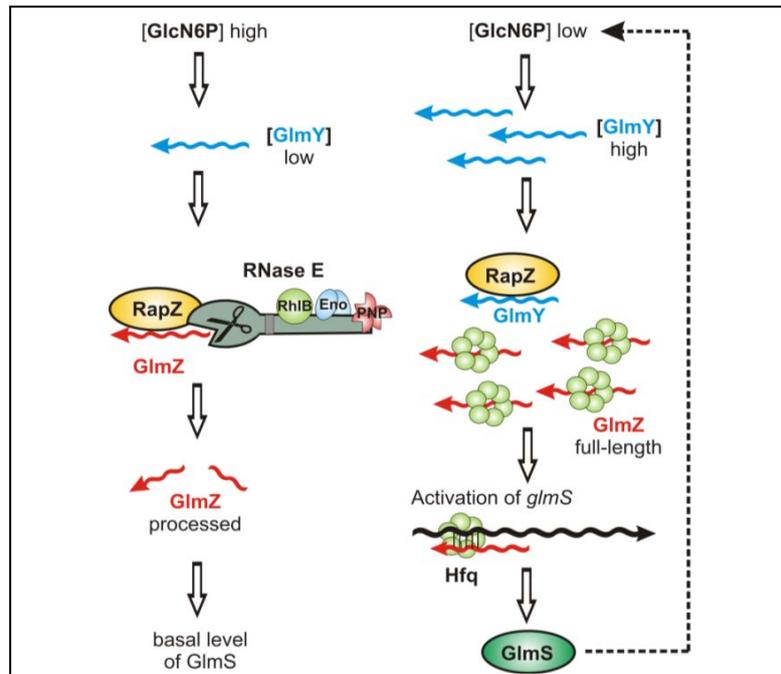
*In vivo*, GlmY counteracts processing of GlmZ. To reconstitute this antagonism with a minimal system, we tested the effect of GlmY on processing of GlmZ by RNase E *in vitro*. To this end 60 nM cold GlmY was added to the *in vitro* cleavage system containing radio-labeled GlmZ, RapZ and varying amounts of the catalytic domain of RNase E (Fig. 4.6 A). Intriguingly, the presence of GlmY strongly inhibited processing of GlmZ. High RNase E concentrations led to some processing but the majority of the emerging products corresponded to unspecific cleavage similar to the one observed when RapZ concentrations were limiting. Both, full-length and processed GlmY, inhibited cleavage of GlmZ in these *in vitro* assays with comparable efficiency (compare Fig. 4.6A and Fig. S4.19). This is consistent with our observation that RapZ binds both forms of GlmY with similar affinities *in vitro* (Fig. S4.6 A). Addition of 60 nM of the non-cognate sRNA SraC to the *in vitro* cleavage system did not impair processing of GlmZ demonstrating that this is a specific function of GlmY (Fig. S4.20). These results, together with the observation that GlmY binds RapZ with high affinity (Fig. 4.2), indicated that GlmY counteracts processing of GlmZ through sequestration of its processing co-factor RapZ.

To obtain *in vivo* evidence that GlmY and GlmZ compete for binding of RapZ, we carried out copurification experiments using Strep-RapZ as bait. To obtain physiological concentrations of RapZ, a strain was used in which the native *rapZ* gene was replaced by *strep-rapZ*, which was expressed from an arabinose-inducible promoter at the ectopic chromosomal *λattB*-site (Fig. S4.21 A). Control experiments verified that upon induction, Strep-RapZ triggered processing of GlmZ with the same rate as observed in the wild-type strain (Fig. S4.21 A). To investigate whether high concentrations of one sRNA may prevent binding of the other sRNA to RapZ, plasmids over-expressing either GlmY or GlmZ were introduced into this strain in addition to the chromosomally encoded sRNAs. Over-expression of GlmY, and to a marginally lower extent also of GlmZ, led to up-regulation of the *glmS* transcript (Fig. S4.22), which is in agreement with previous data (Reichenbach *et al.*, 2008; Urban and Vogel, 2008). From these strains Strep-RapZ was purified by affinity chromatography and copurifying sRNAs were analyzed by Northern blotting (Fig. 4.6 D). Only GlmY, not GlmZ copurified with Strep-RapZ when the strain over-expressing *glmY* was examined. The opposite result was obtained when the strain over-expressing *glmZ* was subjected to the copurification experiment. These results showed that high concentrations of GlmY displace GlmZ from RapZ, and vice versa.

In wild-type cells GlmY accumulates upon depletion of the cellular GlcN6P pool, which leads to the inhibition of GlmZ processing (Reichenbach *et al.*, 2008). Therefore, we predicted that under conditions of low GlcN6P concentrations, the ratio of GlmY/GlmZ bound to RapZ should shift in favor of GlmY. To achieve GlcN6P depletion, a  $\Delta glmS$  mutant was employed. This mutant requires exogenous supplementation with amino sugars such as glucosamine (GlcN) for viability. Upon uptake, GlcN is converted to GlcN6P thereby bypassing the need for GlmS (Plumbridge and Vimr, 1999). Withdrawal of GlcN from the culture results in cessation of growth and finally cell death (Fig. S4.21 B). Concomitantly, both sRNAs strongly accumulate, i.e. cells try to compensate the GlcN6P downshift through activation of the GlmYZ cascade (Fig. S4.21 B; (Reichenbach *et al.*, 2008)). Ultimately, this leads to up-regulation of a plasmid encoded *glmS<sup>+</sup>-lacZ* fusion when present in these strains (Fig. S4.21 C), demonstrating induction of the complete GmY/GlmZ/*glmS* regulatory cascade upon GlcN6P depletion. The GlcN6P downshift had no significant effect on cellular RapZ levels (Fig. S4.21 D). Two hours after amino sugar withdrawal (Fig. S4.21 B), cells were harvested and the sRNAs copurifying with the chromosomally encoded Strep-RapZ protein (Fig. S4.23) were isolated and analyzed by Northern blot (Fig. 4.6 E). Indeed, more GlmY was bound to RapZ in the GlcN-starved *glmS* mutant as compared to the isogenic *glmS<sup>+</sup>* strain (Fig. 4.6 E upper panel, lanes 5 and 7). Intriguingly, the opposite pattern was observed for GlmZ (Fig. 4.6 E lower panel, lanes 5 and 7). Collectively the data suggest that accumulation of GlmY, either as consequence of its artificial overexpression or of GlcN6P depletion, sequesters RapZ leading to the displacement of GlmZ. Consequently, GlmZ is not recruited to processing by RNase E and accumulates to activate *glmS* expression (Fig. 4.7).

## DISCUSSION

In this study we provide mechanistic insight into the GlmY/GlmZ sRNA cascade regulating synthesis of the metabolic enzyme GlmS. We show that protein RapZ is a novel type of an RNA-binding adaptor protein that binds sRNA GlmZ and targets it to processing by RNase E. Notably, this process involves direct interaction of RapZ with the N-terminal catalytic domain of RNase E. Cleavage of GlmZ by RNase E removes the base-pairing nucleotides and thereby inactivates the sRNA (Fig. 4.7, left). The homologous sRNA GlmY counteracts this process by acting as decoy sRNA for RapZ and thus functions as an anti-adaptor. GlmY accumulates upon GlcN6P depletion and sequesters RapZ. As a result, full-length GlmZ is stabilized and activates GlmS synthesis through Hfq-assisted base-pairing with its mRNA (Fig. 4.7, right). Thus, *E. coli* uses a regulatory circuit composed of a base-pairing sRNA, an adaptor protein and an anti-adaptor sRNA to achieve GlcN6P homeostasis. This mechanism likely applies to all species of *Enterobacteriaceae*, in which all components of the circuit are conserved (Fig. S4.8; (Göpel *et al.*, 2011)).



**Figure 4.7. Model for the control of the regulatory GlmY/GlmZ cascade by RapZ.** When GlcN6P concentrations are high in the cell, GlmY is present in low amounts. Under these conditions, RapZ is free to bind GlmZ and to recruit it to processing by RNase E through protein-protein interaction. Consequently, GlmZ is inactivated and unable to activate GlmS synthesis. When GlcN6P levels decrease, processed GlmY accumulates and binds and sequesters RapZ. Thereby, GlmZ remains unbound and cannot be processed by RNase E. As a result, GlmZ base-pairs with *glmS* in an Hfq-dependent manner and activates synthesis of GlmS, which re-synthesizes GlcN6P.

Our work has two major implications for bacterial sRNA-mediated regulation and control of RNA turnover in general. Firstly, proteins that selectively associate with particular sRNAs to alter their functionality might be more common in bacteria than previously thought. Secondly, proteins can play a major role in programming a particular RNA to processing or degradation, thus explaining how specificity in substrate recognition by RNases can be achieved.

Reconstitution of specific processing of GlmZ by RNase E *in vitro* required protein RapZ (Fig. 4.6 A, B). Thus, GlmZ per se is not an appropriate substrate for RNase E. How RNase E recognizes its various substrates is as yet to be understood. In functional analogy to decapping of eukaryotic mRNAs, many RNAs require dephosphorylation at their 5' end to initiate RNase E mediated decay (Deana *et al.*, 2008). Interaction with the 5' mono-phosphate end of the RNA was shown to activate the catalytic function of RNase E (Callaghan *et al.*, 2005). In the *in vitro* cleavage experiments (Fig. 4.6), GlmZ carried a tri-phosphate at its 5' end. Therefore, it seems unlikely that RapZ acts by providing access for RNase E to the 5' end of GlmZ, which is sequestered in a stable stem loop structure (Fig. 4.4). This is further supported by data indicating that RapZ contacts the lateral bulge in the central stem loop in GlmZ (Fig. 4.4). However, mutations in this structure were not sufficient to abolish RapZ binding (Fig. S4.18), instead they strongly impaired processing by RNase E (Fig. 4.6 C). Thus, GlmZ might belong to the class of substrates, which are recognized by RNase E through their fold and are cleaved regardless of their 5' phosphorylation status (Bouvier and Carpousis, 2011; Kime *et al.*, 2010). It is

worth noting that only processed GlmZ copurified with RapZ when RapZ was overexpressed (Fig. 4.2 A, B). In contrast, only full-length GlmZ copurified when RapZ was expressed at physiological concentrations (Fig.4.6 D, E). *In vitro* RapZ binds both forms of GlmZ with similar affinities (Fig. S4.6 B). *In vivo*, RapZ might preferably bind the more abundant of both forms. In principle, this could provide the basis of a feedback mechanism inhibiting processing of GlmZ when the level of processed GlmZ exceeds a threshold, which remains to be determined.

Different experimental approaches demonstrated interaction of RapZ with the N-terminal domain of RNase E, and this interaction is most likely direct (Fig. 4.5 B-D, Fig. S4.15-S4.17). Thus, two different scenarios can be envisioned for the mode of action of RapZ, which remains to be explored. In the first model, RapZ might allosterically activate RNase E upon interaction, and binding of GlmZ by RapZ would serve to deliver GlmZ to RNase E located at the membrane. Upon contact, GlmZ would be released from RapZ and transferred to RNase E. In the second model, binding of RapZ would induce and stabilize structural rearrangements in GlmZ converting it to a substrate that is recognized by RNase E. There is accumulating evidence that RNase E undergoes multiple, possibly dynamic interactions with other proteins beyond its role in the canonical RNA degradosome. For instance, an alternative degradosome containing Hfq rather than RhlB appears to degrade certain sRNAs/target mRNA duplexes (Ikeda *et al.*, 2011). Protein CsrD was shown to selectively bind sRNAs CsrB and CsrC to promote their degradation by RNase E, but whether this involves interaction with RNase E is unknown (Suzuki *et al.*, 2006). Recently, the RNA-binding protein RHON1 was shown to target certain plastid transcripts to processing by RNase E in *Arabidopsis* (Stoppel *et al.*, 2012). As observed here for RapZ, this process likely involves interaction of RHON1 with RNase E. Therefore, it is tempting to speculate that RNA-adaptor proteins are ubiquitously used to program ribonucleases for target specificity.

We also clarified how the sRNA GlmY counteracts processing of its homolog GlmZ and thereby controls *glmS* expression indirectly. In contrast to GlmZ, GlmY appears not to be a base-pairing sRNA as it is stable in an *hfq* mutant and is bound by Hfq only with low affinity (Fig. 4.1). Given the high similarity of both sRNAs, this difference is remarkable. According to recent reports, an accessible poly(U)-tail appears to be critical for sRNA-binding to Hfq (Otaka *et al.*, 2011; Sauer and Weichenrieder, 2011). The poly(U)-tail of GlmY, in contrast to that of GlmZ, is buried in a stem loop structure, which may explain the differences in Hfq binding. Consistently, GlmY acts on *glmS* not by base-pairing, but by targeting protein RapZ. RapZ appears to recognize similar structures in both sRNAs, i.e. the central stem loops including the lateral bulges (Fig. 4.4), presumably through a domain located at its C-terminus (Fig. 4.3). Both sRNAs compete for binding to RapZ (Fig.4.2, Fig. 4.6 D) and GlmY out-competes GlmZ, when it accumulates as consequence of intracellular GlcN6P

depletion (Fig. 4.6 E; Fig. S4.21 B). Finally, GlmY prevents the RapZ-mediated processing of GlmZ by RNase E *in vitro* (Fig. 4.6 A), reflecting the *in vivo* scenario (Reichenbach *et al.*, 2008; Urban and Vogel, 2008). *In vitro*, processed as well as unprocessed GlmY are able to counteract processing of GlmZ (Fig. 4.6 A and Fig. S4.19). However, *in vivo* it is processed GlmY that accumulates upon GlcN6P depletion and is responsible for inhibition of GlmZ cleavage (Fig. S4.21 B). In conclusion, a small RNA acts as decoy to inhibit degradation of a second sRNA through sequestration of an RNase recruiting protein.

The regulatory circuit investigated here strongly resembles the principle of controlled proteolysis of regulatory proteins in bacteria (Bougdour *et al.*, 2008). In this process a regulatory protein is recruited to the degrading protease complex through interaction with an adaptor protein. This can be counteracted by an anti-adaptor protein, which binds and sequesters the adaptor leading to stabilization of the regulator. The GlmYZ system works similarly, but here the adaptor protein targets a sRNA regulator to programmed decay and is antagonized by a sRNA anti-adaptor. Such a sophisticated regulatory cascade provides multiple points of entry and exit for interaction and communication with additional molecules. For instance, usage of multiple anti-adaptor proteins allows the cell to activate the regulatory protein in response to different environmental cues (Bougdour *et al.*, 2008). Whether similar scenarios hold true for the GlmYZ system remains to be explored.

## **MATERIALS AND METHODS**

### ***Growth conditions, plasmids and strains***

Plasmids, strains and culture conditions are described under "Supplemental material"

### ***Western blotting and dot-blot far-Western analysis***

Western blotting and dot-blot far-Western analysis were carried out as described previously (Lüttmann *et al.*, 2012). For dot-blot far-Western analysis Strep-RapZ and Strep-RapZ<sub>quad</sub> were purified from strain Z106, which lacks GlmY and GlmZ. Polyclonal rabbit antisera were used at dilutions of 1:5000 (anti-RapZ, SeqLab), 1:10000 (anti-FLAG, Sigma-Aldrich), 1:20000 (anti-Strep, PromoKine and anti-His, Antibodies online). The antibodies were visualized with alkaline phosphatase conjugated goat anti-rabbit IgG secondary antibodies (Promega), diluted 1:100000, and the CDP\* detection system (Roche Diagnostics, Germany).

### ***Isolation of total RNA and Northern analysis***

Purification of total RNA and Northern blotting was performed as described previously (Reichenbach *et al.*, 2008; Reichenbach *et al.*, 2009). Digoxigenin-labeled RNA probes were produced by *in vitro* transcription of PCR products. Oligonucleotides BG709 and BG710 were used to generate the *csrC*-specific PCR product.

### **Protein purification**

The purification of recombinant Strep- and His<sub>6</sub>-tagged proteins is described under "Supplemental material".

### **Electrophoretic mobility shift assays**

EMSA experiments were performed essentially as described previously (Sittka *et al.*, 2007) using  $\alpha$ -<sup>32</sup>P-UTP-labeled RNAs, which were generated as described under "Supplemental material". Briefly, ~4 nM of the heat-denatured labeled sRNA, 1  $\mu$ g yeast tRNA and various amounts of Hfq or Strep-RapZ were incubated in 1 $\times$  structure buffer (10 mM Tris pH 7, 100 mM KCl, 10 mM MgCl<sub>2</sub>) in a final volume of 10  $\mu$ l for 30 min at 30°C. Protein dilutions were prepared in 1 $\times$  structure buffer and protein concentrations were calculated for RapZ monomers and Hfq hexamers, respectively. Prior to loading 2  $\mu$ l of loading buffer (50 % glycerol, 0.5 $\times$  TBE, 0.2% bromophenol blue) was added and samples were subsequently separated by gel electrophoresis (8 % polyacrylamide, 1 $\times$  TBE) in 0.5 $\times$  TBE buffer at 300 V for 3 h at 4°C. Dried gels were analyzed by phospho-imaging.

### **RNA and protein copurification**

Copurification experiments were carried out as described previously (Lüttmann *et al.*, 2012). Briefly, Strep-tagged bait proteins were purified from *E. coli* cells by StrepTactin affinity chromatography. Successful purification of the bait proteins was confirmed by SDS-PAGE and subsequent analysis of the gels by staining with Coomassie blue (Figs. S4.4, S4.8, S4.11) or by Western blotting using anti-RapZ antiserum (Fig. 4.6 D, S4.15). Copurifying RNase E-3 $\times$ FLAG was detected by Western blotting using anti-FLAG antiserum. For isolation of copurifying RNAs, 300  $\mu$ l of the elution fractions were extracted with phenol:chloroform:isoamyl-alcohol (25:24:1) and the RNA was precipitated with EtOH:LiCl (30:1) and resolved in 30  $\mu$ l RNase-free water. 2.5  $\mu$ l of these samples were subjected to Northern blotting and 15  $\mu$ l were used for conversion to cDNA.

### **Preparation of cDNA and 454 pyrosequencing**

cDNA library construction and 454 pyrosequencing was performed as described for the identification of eukaryotic microRNA (Berezikov *et al.*, 2006), but omitting size-fractionation of RNA prior to cDNA synthesis (Sittka *et al.*, 2008). The cDNAs were PCR-amplified to 20-30 ng/ $\mu$ l. The resulting cDNA libraries were sequenced on a Roche FLX Titanium 454 sequencer. The sequences of cDNA inserts  $\geq$  18 nt were blasted against the *E. coli* K12 genome (NC\_000913). The Integrated Genome Browser software from Affymetrix was used for visualization of the location of blast hits and calculation of mapped reads per nucleotide.

### **$\beta$ -Galactosidase assays**

Overnight cultures of *E. coli* were inoculated into fresh LB medium to an OD<sub>600</sub> of 0.1 and grown to an OD<sub>600</sub> of 0.5-0.8. Subsequently, the cells were harvested and the  $\beta$ -galactosidase activities were determined as described previously (Miller, 1972). The presented values are the average of at least three measurements using independent cultures from at least two independent transformants.

### **RNase T1 protection assays**

RNase T1 protection assays were carried out in 10  $\mu$ l reactions as described previously (Sharma *et al.*, 2007). Briefly, 0.4 pmol of the sRNA, which was 5'-end labeled as described under "Supporting information", was denatured (95°C, 1 min) and chilled on ice for 5 min. Subsequently, 10 $\times$  structure

buffer (100 mM Tris-HCl, 1 M KCl, 100 mM MgCl<sub>2</sub>, pH 7.0), 1 µg yeast tRNA and optionally 80 nM or 300 nM Strep-RapZ were added. After incubation for 10 min at 30°C, the reaction was started by addition of 0.1 u RNase T1 (Ambion). Following 2 min incubation at 37°C, the reaction was stopped by addition of 100 µl stop solution (50 mM Tris, 10 mM EDTA, 0.1% SDS, pH 8.0). The solution was extracted with phenol:chloroform:isoamyl-alcohol (25:24:1) and the RNA was precipitated with ethanol:sodium acetate pH 5.2 (30:1) and finally dissolved in 22 µl 2x RNA loading buffer (95% Formamide; 18 mM EDTA; 0.025% SDS, 0.01% Xylene cyanol, 0.01% Bromophenol blue). OH ladders were obtained by incubation of 0.8 pmol 5'-end labeled sRNA for 5 min in alkaline hydrolysis buffer (Ambion) at 95°C. RNase T1 ladders were generated by incubating 0.8 pmol 5'-end labeled sRNA in 1x sequencing buffer (Ambion) for 1 min at 95°C and subsequent addition of 0.1 u RNase T1 and incubation for 5 min at 37°C. Reactions were stopped with 12 µL of 2x RNA loading dye. All samples were denatured for 1 min at 95°C and chilled on ice prior to loading on 6% polyacrylamide/7 M urea sequencing gels. Electrophoresis was carried out in 1x TBE at 40 W for ~2 h. The gels were dried and analyzed by phospho-imaging.

### ***Bacterial two-hybrid (BACTH) assays***

To assess protein-protein interaction in living *E. coli* cells, the BACTH system was used as described previously (Karimova *et al.*, 1998). Details are provided under "Supplemental information".

### ***RNase E in vitro cleavage assay***

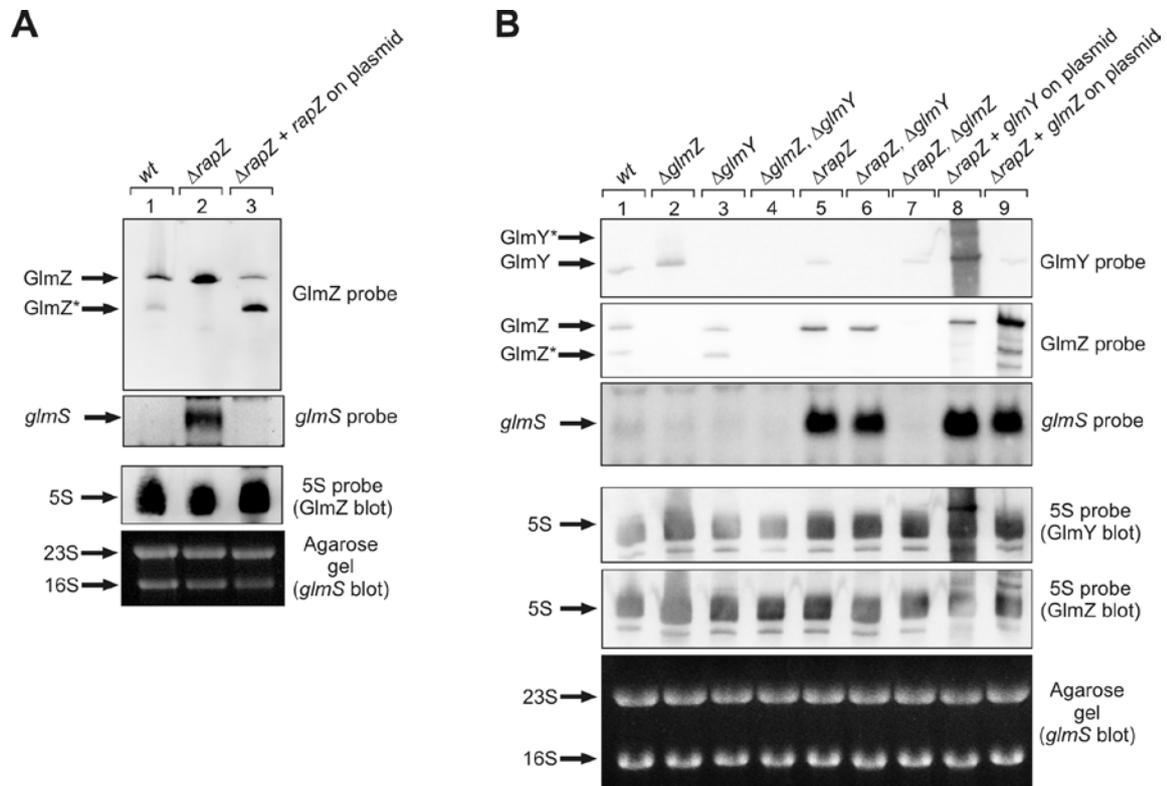
40 nM α-<sup>32</sup>P-UTP labeled GlmZ RNA was denatured by incubation at 70°C for 2 min, followed by incubation on ice for 5 min. The RNA was renatured by incubation at 30°C for 5 min in 1x reaction buffer (25 mM Tris-HCl, 50 mM NaCl, 50 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM DDT, pH 7.5) containing 0.1 mg/ml yeast tRNA (Ambion) in a volume of 8 µl. Subsequently, 1 µl 1x reaction buffer containing the indicated concentration of Strep-RapZ was added and incubation was continued for 10 min. To determine the impact of GlmY on the cleavage assay, 60 nM denatured unlabeled GlmY was added prior to renaturing of GlmZ. Cleavage was started by addition of 1 µl 1x reaction buffer containing the indicated concentration of the catalytic domain of RNase E, followed by incubation at 30°C for 20 min. Reactions were stopped by addition of 4 u Proteinase K (Fermentas) and PK-buffer (100 mM Tris/HCl, 12.5 mM EDTA, 150 mM NaCl, 1% SDS, pH 7.5) and the samples were incubated at 50°C for 10 min to ensure degradation of RNase E. 2xRNA loading buffer was added and samples were separated on 7M urea/TBE/8% polyacrylamide gels. The gels were dried and analyzed by phospho-imaging.

### **ACKNOWLEDGMENTS**

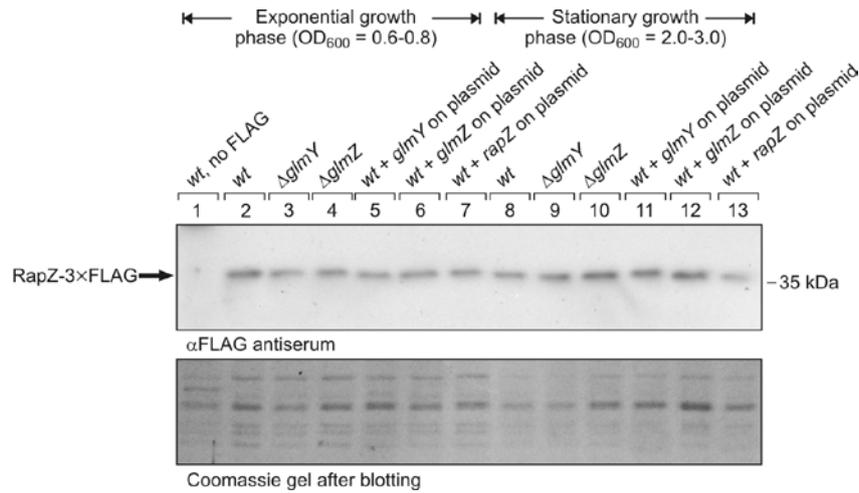
We thank Sabine Lentjes for excellent technical assistance and Jörg Stülke for laboratory space. Cynthia Sharma is acknowledged for deep sequencing analysis. We are grateful to Sue Lin-Chao, Ben F. Luisi, Eliane Hajnsdorf, and Jacqueline Plumbridge for providing plasmids and strains and Ben F. Luisi for the gift of purified RhlB protein. Falk Kalamorz is thanked for help with the preparation of RapZ antiserum and Barbara Waldmann for construction of plasmid pBGG403. Y.G. received a fellowship of the Dorothea Schlözer Program of the Göttingen University and B.R. was supported by a stipend of the Studienstiftung des Deutschen Volkes. This work was supported by grants of the DFG priority program SPP1258 "Sensory and regulatory RNAs in prokaryotes" to B.G. and J.V.

## SUPPORTING INFORMATION

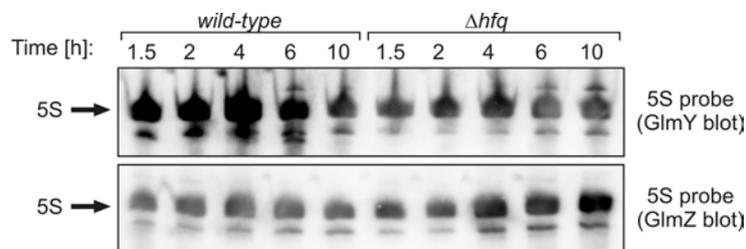
## SUPPLEMENTAL FIGURES



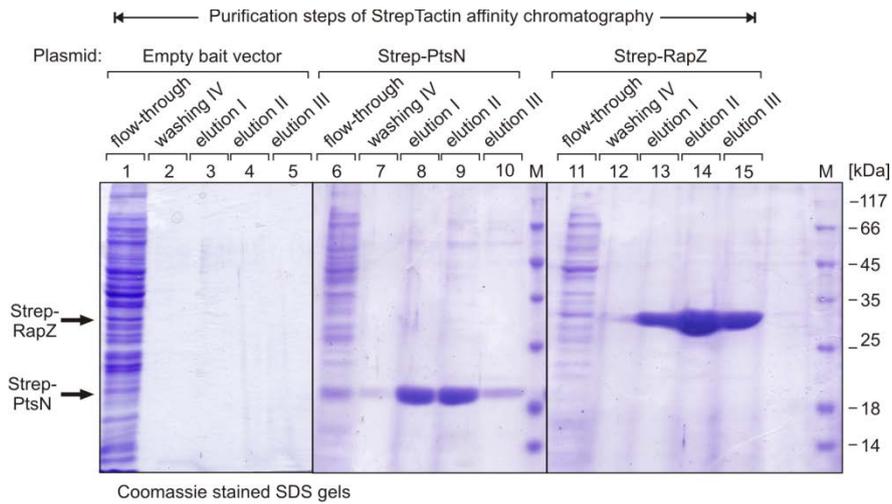
**Fig. S4.1. RapZ (YhbJ) acts upstream of GlmZ, but downstream of GlmY. (A)** Northern blot experiments addressing the role of RapZ (YhbJ) for processing of GlmZ. The fates of GlmZ and *glmS* were studied in strain R1279 (*wild-type*, lane 1), strain Z37 ( $\Delta rapZ$ , lane 2) and in strain Z37 carrying plasmid pBGG164 for overproduction of YhbJ (lane 3). IPTG was added to the culture for induction of *yhbJ* expression. The bacteria were grown in LB to exponential phase and total RNAs were isolated and subjected to Northern analysis using the indicated probes. **(B)** Northern blot experiments addressing the fates of the GlmY, GlmZ and *glmS* RNAs in various genetic backgrounds. The following strains and transformants were employed: R1279 (*wild-type*, lane 1), Z45 ( $\Delta glmZ$ , lane 2), Z96 ( $\Delta glmY$ , lane 3), Z106 ( $\Delta glmZ \Delta glmY$ , lane 4), Z37 ( $\Delta rapZ$ , lane 5), Z115 ( $\Delta rapZ \Delta glmY$ , lane 6), Z116 ( $\Delta rapZ \Delta glmZ$ , lane 7), Z37 carrying *glmY* on plasmid pYG23 (lane 8), Z37 carrying *glmZ* on plasmid pYG24 (lane 9), Z37 carrying *strep-rapZ* on plasmid pBGG164 (lane 10).



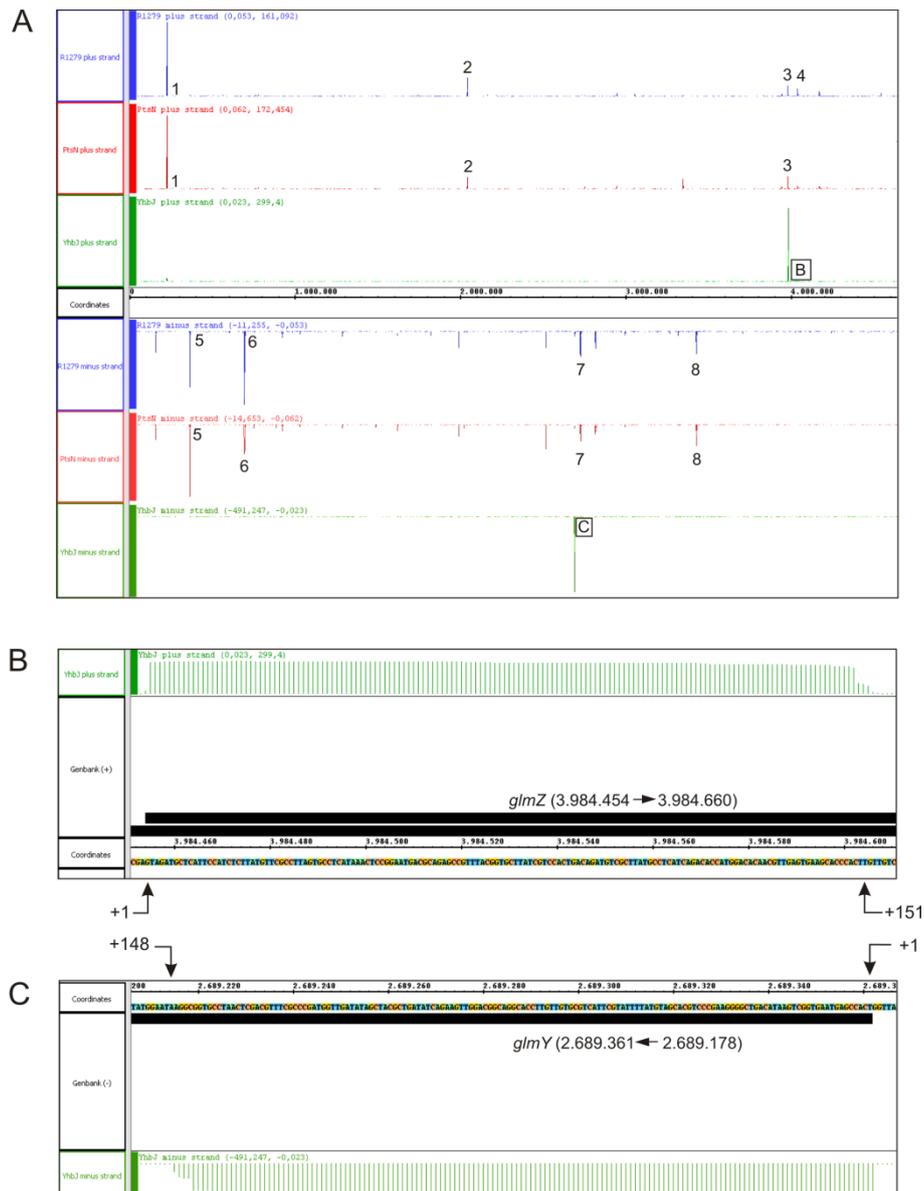
**Fig. S4.2. Absence or over-expression of GlmY or GlmZ has no effect on cellular amounts of RapZ.** Western blot analysis of strains carrying a *rapZ-3xFLAG* fusion gene at the authentic *rapZ* locus (except for lane 1: *wild-type* strain R1279). In addition, the strains carried either deletions of *glmY* or *glmZ* or plasmids over-expressing *glmY* (plasmid pYG23; lanes 5, 11), *glmZ* (plasmid pYG24; lanes 6, 12) or *rapZ* (plasmid pBGG61; lanes 7, 13) as indicated. The following strains were used: Z592 (lanes 2, 5-7, 8 and 11-13), Z610 (lanes 3, 9) and Z612 (lanes 4, 10). The strains and transformants were grown in LB to the  $OD_{600}$  as indicated. For the induction of expression of *rapZ*, arabinose was added to the cultures tested in lanes 7 and 13. Cell extracts were prepared and subsequently analyzed by Western blotting using  $\alpha$ FLAG antiserum (top). The polyacrylamide gel, which was stained with Coomassie blue after blotting, is provided as loading control (bottom). Note that over-expression of *rapZ* from plasmid pBGG61 had also no notable effect on the detected amount of RapZ-3xFLAG, making auto-regulation of synthesis of RapZ unlikely.



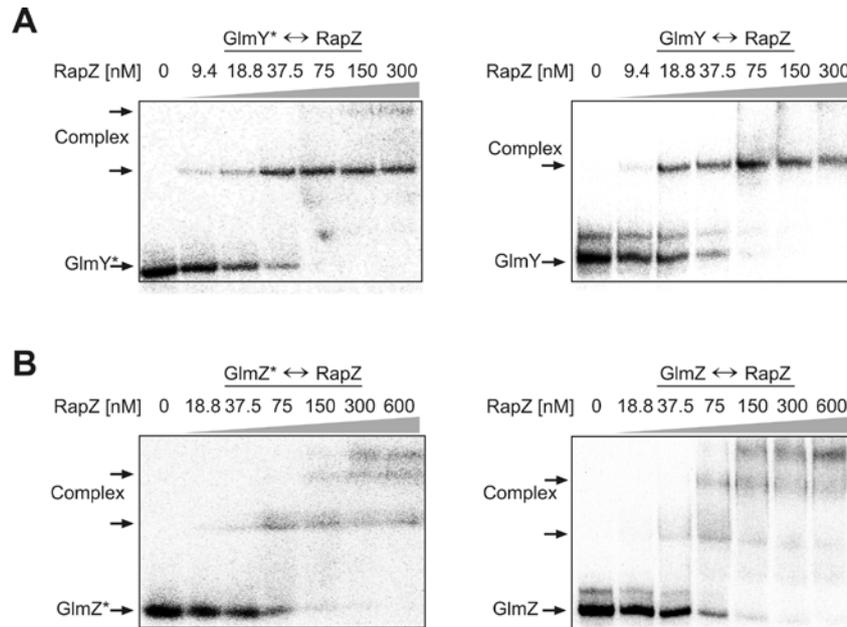
**Fig. S4.3. Loading controls for the Northern blot experiments shown in Fig. 4.1 B.** The nylon membranes used for the Northern blot experiments in Fig. 4.1 were re-probed using a probe directed against 5S rRNA.



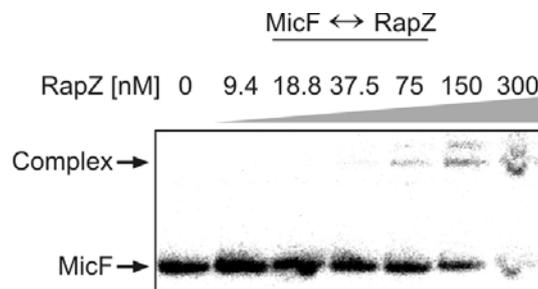
**Fig. S4.4. Analysis of the purification steps of the StrepTactin affinity chromatography of the Protein/RNA copurification experiment (Fig. 4.2) by SDS-PAGE and Coomassie blue staining.** The IPTG-treated cultures of the transformants described in Fig. 4.2 A were subjected to Streptactin affinity chromatography. Aliquots of the column flow-through, the fourth washing step and of the three elution fractions were separated on 12.5 % SDS-polyacrylamide gels and gels were analyzed by staining with Coomassie blue. The RNA samples analyzed in Fig. 4.2 were extracted from the elution fractions II (lanes 4, 9, 14).



**Fig. S4.5. GlmY and GlmZ copurify with RapZ as revealed by RNAseq. A)** Screenshot of IGB-files (Integrated Genome Browser 6.4.1, Oracle) visualizing the cDNA libraries derived from the copurification experiments shown in Fig. 4.2 A. The library depicted in blue was obtained from strain R1279 harboring the empty bait plasmid pBGG237 (negative control). The cDNA library shown in red was derived from copurification with protein Strep-PtsN from strain R1279 (second negative control). The cDNA library shown in green was generated from the copurification with Strep-RapZ (Strain R1279 harboring plasmid pBGG164). 454-pyrosequencing reads of the three cDNA libraries were mapped to the genome of *E. coli* K-12 strain MG1655 (U00096.2). Whereas the majority of cDNA reads obtained from the negative controls corresponded to rRNA (peaks 1, 4, 7, 8), tRNA (peaks 2, 3), some intergenic regions (*lacI-lacZ*, peak 5) or messenger RNAs (*glnX-glnW*, peak 6), these RNAs were underrepresented in cDNA libraries obtained from copurification with Strep-RapZ. Here, GlmY (C) and GlmZ (B) were predominantly detected. **B)** Detailed view of the *glmZ* region on the plus strand and the corresponding cDNA reads. Note that only the processed form of GlmZ (+1 to +151) was detectable in libraries obtained from RapZ copurification. **C)** Detailed view of the *glmY* region on the minus strand and the corresponding cDNA reads. Again, only processed GlmY (+1 to +148) was shown to copurify with RapZ (for quantification of cDNA reads see Fig. 4.2 B and Excel file S4.1).



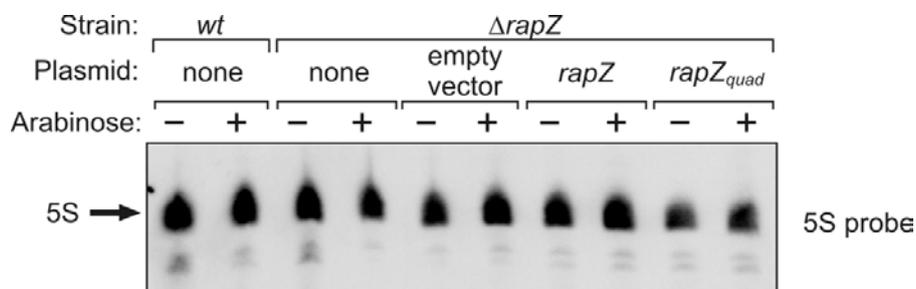
**Fig. S4.6. RapZ efficiently binds GlmY (A) and GlmZ (B) *in vitro* regardless of the processing state of these sRNAs. (A)** EMSAs using  $\alpha$ - $^{32}$ P-UTP labeled sRNAs corresponding to the processed form of GlmY (*GlmY\**; left) or full-length GlmY (right) and various concentrations of purified Strep-RapZ as indicated. **(B)** Same as (A) but processed GlmZ (*GlmZ\**; left) and full-length GlmZ (right) were employed.



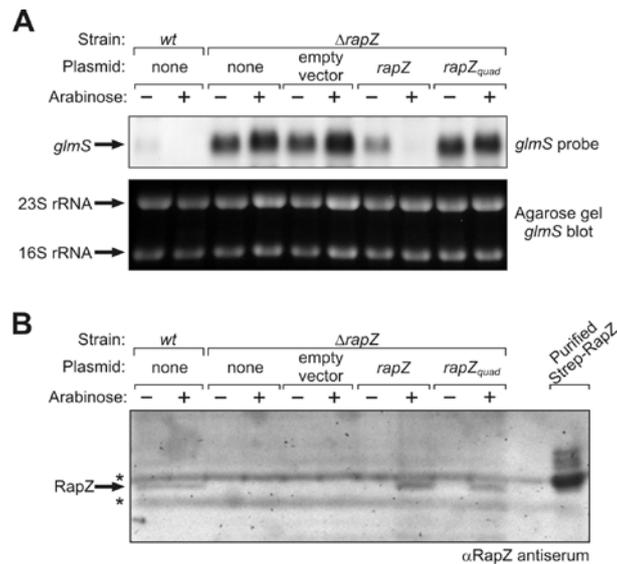
**Fig. S4.7. RapZ binds sRNA MicF only at high protein concentrations indicating unspecific binding.** EMSA using 4 nM of  $\gamma$ - $^{32}$ P-ATP 5' end-labeled sRNA MicF and various concentrations of purified Strep-RapZ as indicated. In vitro transcription and labeling of MicF was performed as previously described (Corcoran *et al.*, 2012).



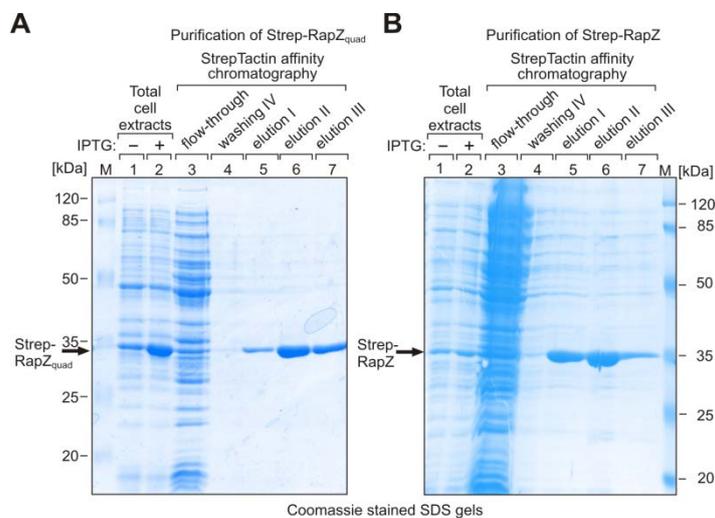
**Fig. S4.8. Alignment of amino acid sequences of RapZ homologs from 29 bacteria.** Fully conserved amino acid positions are highlighted in red. Residues conserved in the majority of sequences are in blue and residues weakly similar to the consensus residue are in green. The nucleotide binding Walker A and B motifs (Luciano *et al.*, 2009) are boxed in red. The putative RNA binding domain predicted by the BindN software is boxed in blue. Note that the sequence of this region is highly conserved in the top 14 sequences, which derive from species belonging to the *Enterobacteriaceae*. In bacteria belonging to other families, this sequence strongly deviates (bottom 15 sequences). Thus, the putative RNA-binding domain is exclusively conserved in species that possess the GlmY and GlmZ sRNAs (for comparison, see: (Göpel *et al.*, 2011)). The positions K270, K281, R282 and K283, which were mutated in *E. coli* RapZ<sub>quad</sub> (Fig. 4.3) are indicated by arrows. The alignment was compiled using the AlignX tool of software Vector NTI Advance™ 10.0. Sequences were compiled from the following genomes (accession numbers are in parentheses): *Escherichia coli* K12 str. MG1655 (U00096.2), *Shigella dysenteriae* Sd197 (NC\_007606.1), *Escherichia albertii* TW07627 (NZ\_ABKX01000003.1), *Escherichia fergusonii* ATCC 35469 (NC\_011740.1), *Citrobacter koseri* ATCC BAA-895 (NC\_009792.1), *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. LT2 (NC\_003197.1), *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 (NC\_009648.1), *Cronobacter sakazakii* ATCC BAA-894 (NC\_009778.1), *Erwinia amylovora* CFBP 1430 (NC\_013961.1), *Photobacterium luminescens* subsp. *laumondii* TTO1 (NC\_005126.1), *Pectobacterium carotovorum* subsp. *carotovorum* PC1 (NC\_012917.1), *Serratia marescens* Db11 [http://www.sanger.ac.uk], *Yersinia pseudotuberculosis* YPIII (NC\_010465.1), *Enterobacter* sp. 638 (NC\_009436.1), *Actinobacillus succinogenes* 130Z (CP000746.1), *Aeromonas hydrophila* subsp. *hydrophila* ATCC7966 (CP000462.1), *Bacillus licheniformis* ATCC14580 (NC\_006270.3), *Bacillus subtilis* subsp. *subtilis* str. 168 (NC\_000964.3), *Enterococcus faecalis* TX0645 (AECE01000054.1), *Lactobacillus casei* ATCC 334 (NC\_008526.1), *Streptococcus mutans* UA159 (NC\_004350.2), *Streptococcus pyogenes* MGAS1882 (CP003121.1), *Staphylococcus aureus* subsp. *aureus* NCTC 8325 (NC\_007795.1), *Clostridium botulinum* A str. ATCC3502 (NC\_009495), *Corynebacterium diphtheriae* HC02 (NC\_016802.1), *Burkholderia mallei* ATCC 10399 (CH899684.1), *Marinobacter aquaeolei* VT8 (NC\_008740.1), *Shewanella baltica* OS183 (NZ\_AECY020005.1), *Vibrio cholerae* O1 biovar El Tor str. N16961 (NC\_002505.1).



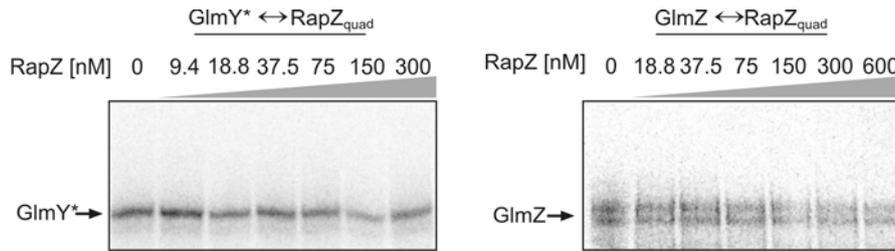
**Fig. S4.9. Loading controls for the Northern blotting experiments shown in Fig. 4.3 B.** The nylon membranes used for the Northern blotting experiments in Fig. 4.3 B were re-probed using a probe directed against 5S rRNA.



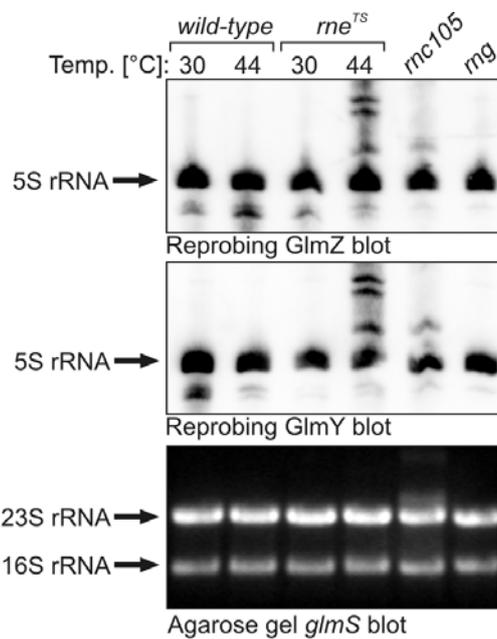
**Fig. S4.10. The RapZ quadruple mutant (RapZ<sub>quad</sub>) protein, although expressed and stable, does not complement a  $\Delta rapZ$  mutation. (A)** Northern blot analysis for the detection of glmS mRNA in total RNA isolated from the strains and transformants tested in Fig. 4.3 B. A probe directed against glmS was used. The ethidium bromide stained agarose gel is shown as a loading control below the blot. **(B)** Western blotting analysis of RapZ and RapZ<sub>quad</sub> proteins. Cell extracts of the strains and transformants tested in Fig. 4.3 B were subjected to Western blotting analysis using a polyclonal antiserum directed against RapZ. In the last lane 0.5  $\mu$ g of purified Strep-RapZ, which exceeds the molecular weight of native RapZ by 1.55 kDa, was loaded as a size marker. Note that the RapZ antiserum also detects several proteins in the extracts unspecifically (marked with asterisks).



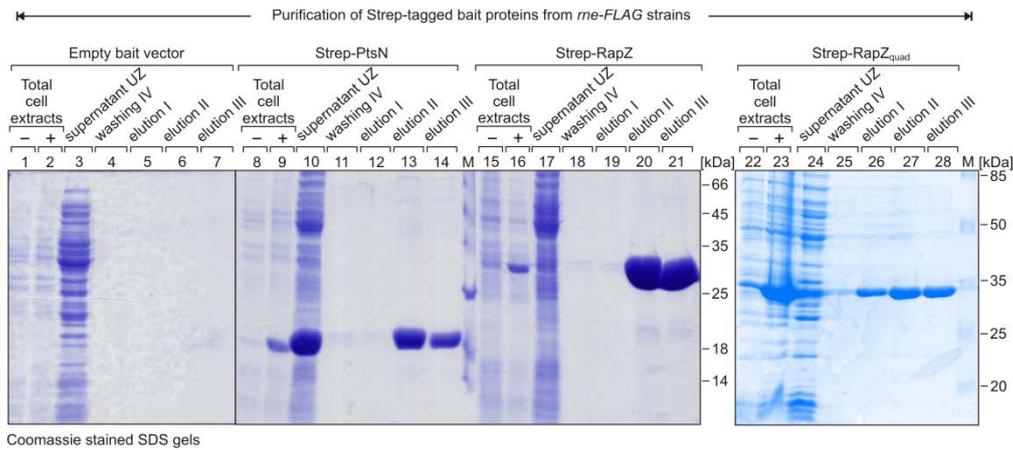
**Fig. S4.11. Analysis of samples collected during the RNA copurification experiment shown in Fig. 4.3 C by SDS-PAGE and Coomassie blue staining.** Total protein extracts were analyzed in lanes 1 and 2 to confirm overproduction of Strep-RapZ<sub>quad</sub> **(A)** and Strep-RapZ **(B)**. Strain R1279 carrying either plasmid pYG29 **(A)** or pBGG164 **(B)** was grown in LB until an OD<sub>600</sub> = 0.8. Subsequently, 1 mM IPTG was added to induce synthesis of the Strep-tagged proteins. The total protein extracts of samples collected before addition of IPTG (lanes 1) and 1 h after addition (lanes 2) were analyzed. The IPTG-treated cultures (lanes 2) were subjected to Streptactin affinity chromatography. Aliquots of the flow-through, the fourth washing step and of the three elution fractions were analyzed. Samples were separated on 12.5 % SDS-polyacrylamide gels and gels were stained with Coomassie blue. The RNA samples analyzed in Fig. 4.3 C were extracted from the elution fractions II.



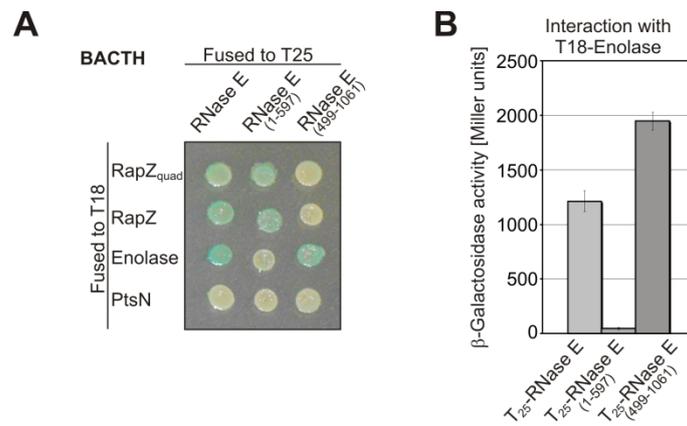
**Fig. S4.12. RapZ<sub>quad</sub> fails to bind GlmY and GlmZ *in vitro*.** EMSAs using 4 nM of  $\alpha$ -<sup>32</sup>P-UTP labeled GlmY (left) or GlmZ (right) sRNAs and various concentrations of purified Strep-RapZ<sub>quad</sub> as indicated.



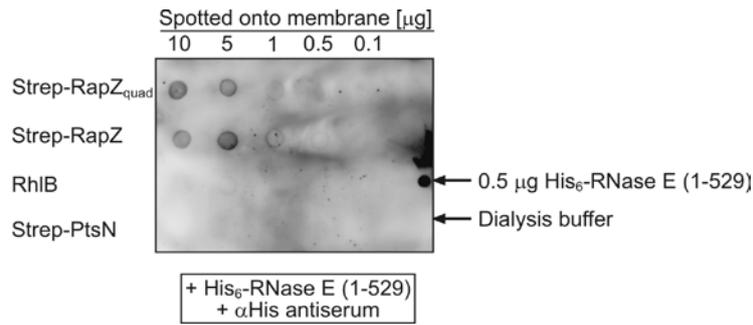
**Fig. S4.13. Loading controls for the Northern blotting experiments shown in Fig. 4.5 A.** The nylon membranes used to detect GlmZ and GlmY by Northern analysis were re-probed using a probe directed against 5S rRNA, respectively (top and middle). In addition, the ethidium-bromide-stained agarose gel used for the *glmS* blot is shown (bottom).



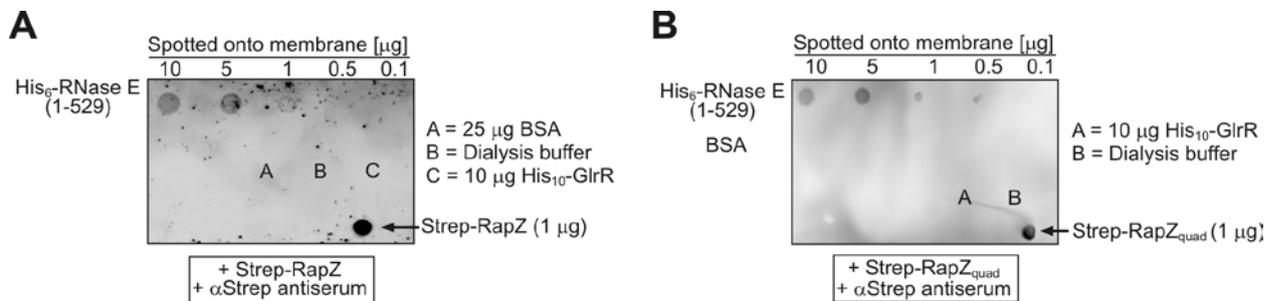
**Fig. S4.14. Analysis of samples collected during the copurification experiments shown in Fig. 4.5 B by SDS-PAGE and Coomassie blue staining.** In these experiments strain Z64 was used, which encodes an RNase E-FLAG fusion protein on the chromosome. Additionally, strain Z64 carried either plasmid pBGG237 (empty vector; lanes 1-7), plasmid pBGG217 encoding Strep-PtsN (lanes 8-14), plasmid pBGG164 encoding Strep-RapZ (lanes 15-21) or plasmid pYG29 encoding Strep-RapZ<sub>quad</sub> (lanes 22-28). Total protein extracts of these transformants were analyzed in lanes 1, 2, 8, 9, 15, 16, 22 and 23 to confirm overproduction of Strep-PtsN, Strep-RapZ and Strep-RapZ<sub>quad</sub>, respectively. Samples were collected before ("-") and 1 h after addition of IPTG ("+"). The IPTG-treated cultures were subjected to Streptactin affinity chromatography. Aliquots of the supernatants after ultra-centrifugation, the fourth washing steps and of the three elution fractions were analyzed. Samples were separated on 12.5 % SDS-polyacrylamide gels and gels were stained with Coomassie blue.



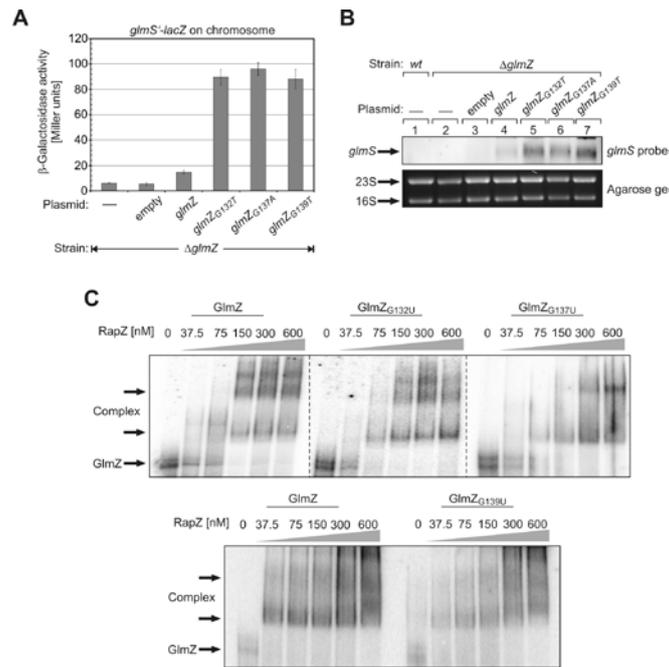
**Fig. S4.15. Bacterial two-hybrid (BACTH) assays addressing interaction of RapZ and enolase with RNase E.** (A) Phenotypic BACTH assay demonstrating that RapZ and the RapZ<sub>quad</sub> mutant interact with the N-terminal fragment of RNase E, while enolase binds to the RNase E C-terminal fragment. The following plasmid combinations were tested in strain BTH101 (from top to bottom and from left to right): pT25-Rne/pYG42, pT25-Rne/pBGG349, pT25-Rne/pYG45, pT25-Rne/pYG46, pT25-RE1/pYG42, pT25-RE1/pBGG349, pT25-RE1/pYG45, pT25-RE1/pYG46, pT25-RE3/pYG42, pT25-RE3/pBGG349, pT25-RE3/pYG45, pT25-RE3/pYG46. The co-transformants were spotted on LB-plates containing X-Gal for the phenotypic detection of β-galactosidase activity and incubated at 30°C. A blue color indicates reconstitution of split adenylate cyclase CyaA activity through interaction of the proteins that are fused to its separately encoded T25- and T18-domains. (B) Control experiment for Fig. 4.5 C demonstrating interaction of enolase with the C-terminal fragment of RNase E. Transformants of strain BTH101 containing plasmid pYG45 and either pT25-Rne, pT25-RE1 or pT25-RE3 were grown to stationary growth phase and the β-galactosidase activities were determined.



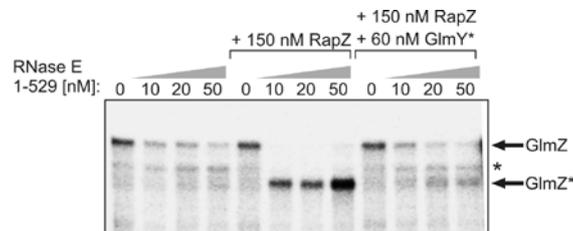
**Fig. S4.16.** Dot-blot far-Western indicating that RapZ and the RapZ<sub>quad</sub> mutant but not RhlB and Strep-PtsN interact with the catalytic domain of RNase E *in vitro*. Various amounts of purified Strep-RapZ<sub>quad</sub> and Strep-RapZ were spotted onto a nylon membrane. RhlB, which interacts with the C-terminus of RNase E and Strep-PtsN served as negative controls. As a positive control 0.5  $\mu\text{g}$  His<sub>6</sub>-RNase E was also spotted on the membrane (arrow). The membrane was subsequently incubated in 50 nM of the His<sub>6</sub>-tagged catalytic domain of RNase E and interaction was visualized with an antiserum directed against the His-tag.



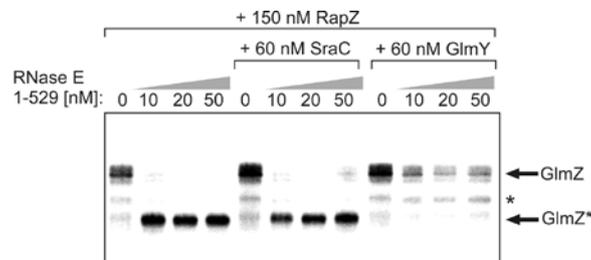
**Fig. S4.17.** Dot-blot far-Western indicating interaction of Strep-RapZ (A) and the Strep-RapZ<sub>quad</sub> mutant protein (B) with the catalytic domain of RNase E *in vitro*. Various amounts of the purified His-tagged catalytic domain of RNase E (His<sub>6</sub>-RNase E 1-529) were spotted onto a membrane. Spots containing 25  $\mu\text{g}$  BSA or dialysis buffer or 10  $\mu\text{g}$  His<sub>10</sub>-GlrR, which was purified as described previously (Göpel *et al.*, 2011), served as negative controls. As positive controls, 1  $\mu\text{g}$  Strep-RapZ (A) or Strep-RapZ<sub>quad</sub> (B) were applied. The membrane was incubated in solutions containing 50 nM Strep-RapZ (A) or Strep-RapZ<sub>quad</sub> (B). Interaction was visualized with an antiserum directed against the Strep-tag.



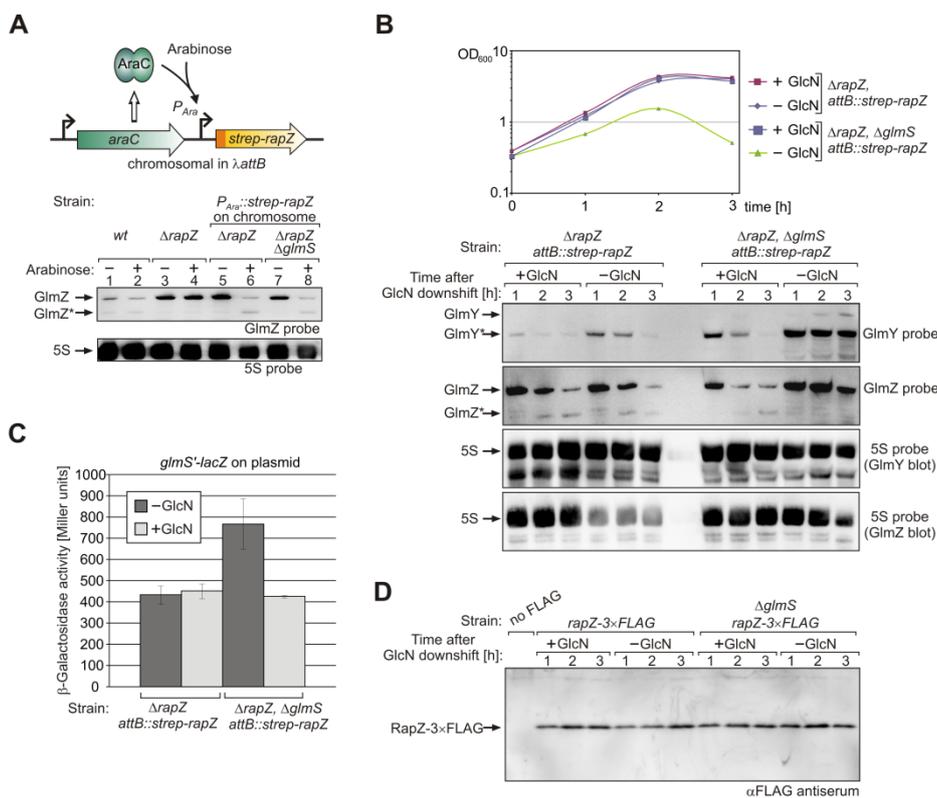
**Fig. S4.18. Mutations in the lateral bulge of GlmZ cause chronic activation of *glmS* expression, although binding of the GlmZ variants by RapZ is not affected.** (A) Complementation of a *ΔglmZ* mutant carrying a *glmS'-lacZ* fusion on the chromosome (strain Z38) with various *glmZ* alleles expressed from plasmids. The following plasmids were tested: pBAD30 (column 2), pBGG84 (column 3), pBGG403 (column 4), pYG51 (column 5), pYG52 (column 6). The transformants were grown in LB-arabinose to exponential phase and the  $\beta$ -galactosidase activities were determined. (B) Northern blot analysis of total RNA isolated from the transformants tested in (A). In addition wild-type strain R1279 (lane 1) was employed. (C) EMSAs using various concentrations of purified Strep-RapZ and 4 nM of  $\alpha$ -<sup>32</sup>P-UTP labeled GlmZ or of its mutant variants carrying the indicated base exchanges.



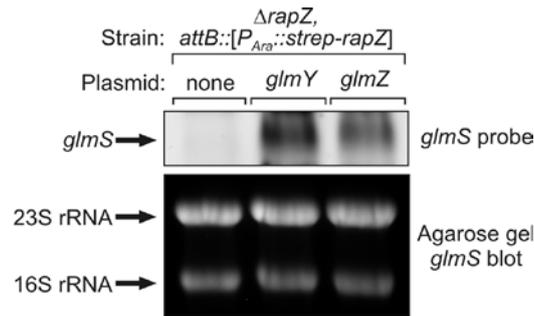
**Fig. S4.19. Processed GlmY (GlmY\*) inhibits the RapZ mediated cleavage of GlmZ by RNase E *in vitro*.** *In vitro* cleavage assay of  $\alpha$ -<sup>32</sup>P-UTP labeled GlmZ using varying concentrations of the catalytic domain of RNase E and 150 nM Strep-RapZ. In addition, 60 nM of unlabeled RNA corresponding to the processed form of GlmY (GlmY\*) were present in lanes 9-12. The asterisk indicates a non-specific cleavage product of GlmZ.



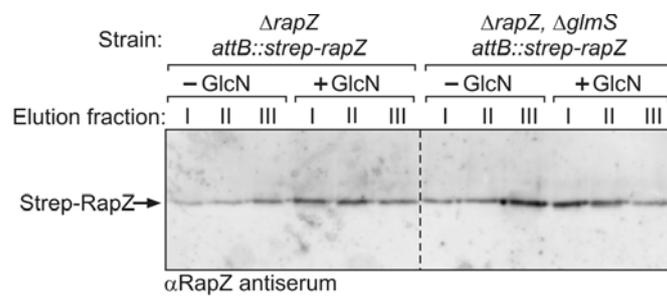
**Fig. S4.20. The RapZ mediated cleavage of GlmZ by RNase E *in vitro* can be inhibited by GlmY but not by the non-cognate sRNA SraC.** *In vitro* cleavage assay of  $\alpha$ -<sup>32</sup>P-UTP labeled GlmZ using varying concentrations of the catalytic domain of RNase E and 150 nM Strep-RapZ. In addition, 60 nM of the unlabeled sRNAs SraC or GlmY were present where indicated. The asterisk indicates a non-specific cleavage product of GlmZ.



**Fig. S4.21. Experimental system for studying the effects of GlcN6P depletion on the GlmYZ cascade. (A)** Functionality test of the chromosomally encoded Strep-RapZ protein. Strains Z479 (lanes 5, 6) and Z555 (lanes 7, 8) lack the authentic *rapZ* gene, but carry the recombinant *strep-rapZ* gene under control of the arabinose-inducible  $P_{Ara}$  promoter in the chromosomal *attB*-site (top). Strain Z555 additionally lacks the *glmS* gene. The strains were grown in LB medium in the absence and presence of arabinose. In addition, 0.2% glucosamine (GlcN) was added to the cultures of strain Z555 to compensate for its inability to synthesize GlcN6P. Total RNA was isolated from exponentially grown cells and analyzed by Northern blotting using a GlmZ-specific probe. For comparison, strains R1279 (wild-type; lanes 1, 2) and strain Z37 ( $\Delta rapZ$ ) were also employed. The blot obtained upon re-probing with a 5S specific probe is shown as loading control. The data show, that upon induction of its gene with arabinose Strep-RapZ triggers processing of GlmZ like wild-type RapZ (compare lanes 1, 3, 6, 8). **(B)** Glucosamine (GlcN) downshift induces the GlmY/GlmZ cascade in a  $\Delta glmS$  mutant. Strains Z479 and Z555 were grown in LB supplemented with arabinose and 0.2% glucosamine until an  $OD_{600} = 0.3$ . Cells were washed and split into two cultures and growth was continued in the presence and absence of GlcN, respectively (top). Total RNA was isolated from samples harvested at the indicated times and analyzed by Northern blotting using probes specific for GlmY and GlmZ. The 5S re-probed membranes are shown as loading controls, respectively. The data show that withdrawal of GlcN ceases growth of the  $\Delta glmS$  cells (top). Concomitantly GlmY and unprocessed GlmZ strongly accumulate. **(C)** Glucosamine (GlcN) downshift induces expression of a plasmid-encoded *glmS'-lacZ* reporter fusion in the  $\Delta glmS$  mutant. Strains Z479 and Z555 carrying the *glmS'-lacZ* reporter plasmid pBGG16 were grown in LB supplemented with arabinose and 0.2% glucosamine and the appropriate antibiotics until an  $OD_{600} = 0.3$ . Cells were washed and split into two cultures and growth was continued in the presence and absence of GlcN, respectively. After 3 hours cells were harvested and the  $\beta$ -galactosidase activities were determined. **(D)** Depletion of intracellular GlcN6P does not affect the cellular amount of RapZ. Strains Z592 and Z593 ( $\Delta glmS$ ) carry a C-terminally FLAG-tagged RapZ allele encoded at the authentic locus of *rapZ*. The cells were grown under the same conditions as described in (B), but arabinose was omitted. Samples harvested at the indicated times were analyzed by Western blotting using anti-FLAG antiserum. To demonstrate specificity of the antiserum, strain R1279 containing no FLAG peptide served as negative control (lane 1).



**Fig. S4.22. Effect of GlmY and GlmZ over-expression on the *glmS* transcript in the strain carrying the recombinant *strep-rapZ* gene.** Northern blotting of total RNA prepared from the strains used in Fig. 4.6 D. A probe directed against *glmS* was used. The ethidium bromide stained agarose gel is shown as a loading control below the blot.



**Fig. S4.23. Analysis of the elution fractions of the copurification experiment shown in Fig. 4.6 E by Western blotting using anti-RapZ antiserum.** Strains Z479 ( $\Delta rapZ$ ,  $attB::strep-rapZ$ ) and Z555 ( $\Delta rapZ$ ,  $\Delta glmS$ ,  $attB::strep-rapZ$ ) were grown in LB containing GlcN and arabinose until  $OD_{600} = 0.3$ . Subsequently, the cultures were split and growth was continued in the absence or presence of GlcN as indicated. After 2 h cultures were harvested and subjected to the copurification protocol using StrepTactin affinity chromatography. Presence of Strep-RapZ in the elution fractions was tested by Western blotting using anti-RapZ antiserum. The copurifying RNAs analyzed in Fig. 4.6 E were isolated from elution fractions 2, respectively.

**SUPPLEMENTAL EXPERIMENTAL PROCEDURES*****Growth conditions and construction of strains and plasmids***

LB was used as standard medium for cultivation of *E. coli*. Cells were grown routinely under agitation (200 r.p.m.) at 37°C. When necessary, antibiotics were added to the medium (ampicillin 100 µg/ml, kanamycin 30 µg/ml, chloramphenicol 15 µg/ml, spectinomycin 75 µg/ml, tetracycline 12.5 µg/ml). 0.2% L-arabinose was added for induction of genes under control of the  $P_{Ara}$  promoter. The *glmS* mutant strain was grown in LB supplemented with 0.2% (w/v) glucosamine if not otherwise indicated.

The used *E. coli* strains are listed in Table 1, including a description of their relevant genotypes. Established alleles tagged with an antibiotic resistance marker were moved to other strains by general transduction using phage T4GT7 (Wilson *et al.*, 1979). The  $P_{Ara}::strep-rapZ$  allele was first established on plasmid pYG13 and subsequently integrated into the phage  $\lambda attB$ -site on the chromosome of strain Z37 by site-specific recombination yielding strain Z479. Recombination was achieved using helper plasmid pLDR8 as previously described (Diederich *et al.*, 1992). Briefly, an origin-less DNA-fragment encompassing the  $P_{Ara}::strep-rapZ$  allele, the *bla*  $\beta$ -lactamase resistance gene and the  $\lambda attP$ -site was isolated by NotI digest. The DNA-fragment was self-ligated and subsequently introduced into strain Z37 carrying the temperature-sensitive  $\lambda$ -integrase expression plasmid pLDR8. Recombinants were obtained upon selection on ampicillin-plates at 42°C. FLAG-tagging of gene *rapZ* was performed by epitope tagging using a PCR-fragment obtained with primers BG853/BG854 and plasmid pSUB11 as template as described before (Uzzau *et al.*, 2001). Antibiotic resistance genes encompassed by FLP recognition sites were removed by making use of the temperature-sensitive FLP recombinase expression plasmid pCP20 as described previously (Datsenko and Wanner, 2000). All strain constructions were verified by diagnostic PCRs.

DNA-cloning was carried out in *E. coli* strain DH5 $\alpha$  following standard procedures. The plasmids and oligonucleotides used in this study are listed in Tables 2 and 3, respectively. For the construction of pBGG61 carrying *rapZ* under  $P_{Ara}$  promoter control, *rapZ* was PCR-amplified using primers BG209/BG212 and plasmid pFDX4296 (Kalamorz *et al.*, 2007) as template. The PCR fragment was digested with SacI and XbaI and inserted between the SacI/XbaI-sites on plasmid pBAD33. Plasmid pYG30 is isogenic with pBGG61, but carries the quadruple mutation K270A-K281A-R282A-K283A in *rapZ*. It was constructed by PCR-amplification of *rapZ* using primers BG209/BG920 and ligation of the fragment between the SacI/XbaI-sites of plasmid pBAD33. Plasmid pYG29 carries the K270A-K281A-R282A-K283A exchanges within *strep-rapZ*. For its construction, *rapZ* was PCR-amplified using primers BG399/BG920. The PCR-fragment was digested with BamHI/XbaI and the resulting smaller DNA-fragment encompassing the 3' end of *rapZ* was used to replace the BamHI/XbaI fragment in plasmid pBGG164. For construction of plasmid pYG13-M, a DNA-fragment encompassing the *araC* gene and the  $P_{Ara}$ -promoter was isolated by digestion of plasmid pBAD18-cm with ClaI/SacI and inserted between the ClaI/SacI-sites in the multiple cloning site of plasmid pLDR10 (Diederich *et al.*, 1992). Subsequently, a PCR-fragment encompassing *strep-rapZ* fragment was amplified by PCR using primers BG765/BG766 and pBGG164 as template and cloned into the SacI site of pYG13-M yielding plasmid pYG13. Mutations in *glmZ* were introduced on plasmids using the combined chain reaction (CCR; (Bi and Stambrook, 1998)). To this end, 5'-phosphorylated oligonucleotides BG745, BG1029 and BG1030 carrying the desired mutations, respectively, were used together with primers BG235 and BG746 in CCRs. CCRs are modified PCR reactions, in which the phosphorylated mutagenesis primers are incorporated during amplification of a DNA fragment by the thermo-stable DNA-

Ampligase (Epicentre). The *SacI*/*XbaI*-digested PCR fragments were inserted between the *SacI*/*XbaI* sites of plasmid pBAD30 resulting in plasmids pBGG403, pYG51 and pYG52, respectively. For construction of the two-hybrid constructs pBGG348 and pBGG349, which carry fusions of *rapZ* to the 3' end of *cyaA-T25* and *cyaA-T18*, *rapZ* was PCR-amplified using primers BG637/BG639 and cloned between the *XbaI*/*KpnI* sites on plasmids pKT25 and pUT18C, respectively. Plasmids pYG42 and pYG94 are isogenic, but encode fusions of the RapZ quadruple mutant to the C-terminus of CyaA-T18 and CyaA-T25, respectively. They were constructed by replacement of the *BamHI*/*KpnI* fragments of plasmid pBGG349 and pBGG348 with a PCR-fragment that was obtained using primers BG399/BG973 and plasmid pYG29 as template and digested with *BamHI*/*KpnI*. Plasmid pYG45, which encodes a fusion of enolase to the C-terminus of CyaA-T18, was generated by cloning the PCR fragment obtained with primers BG1025/BG1026 between the *XbaI*/*KpnI*-sites of pUT18C. For construction of plasmid pYG95, which encodes the T25-enolase fusion, the *XbaI*-*KpnI* fragment of plasmid pYG45 was ligated to the *XbaI*/*KpnI*-digested plasmid pKT25. Plasmids pYG97, pYG98 and pYG99, which code for various fragments of RNase E fused to the C-terminus of CyaA-T18, were constructed by using the primer pairs BG1110/BG1114 for amplification of the 5' half of *rne*, BG1113/BG1111 for amplification of the 3' half of *rne* and BG1110/BG1111 for amplification of the full-length *rne* gene. The PCR fragments were cloned between the *XbaI*/*KpnI* sites on plasmid pUT18C. Plasmid pYG46 was generated by amplification of gene *ptsN* using primers BG1017/BG1018 and subsequent ligation of the fragment with plasmid pUT18C, which was opened by *XbaI*/*KpnI* digestion. Plasmid pYG23 was constructed by ligation of the *BamHI*-*NdeI* fragment of pBGG149 encompassing the  $\lambda P_L::glmY$  cassette to the vector backbone of pBAD18-cm that was isolated by digestion with *BamHI* and *NdeI*. Plasmid pYG24 was constructed by insertion of the *SacI*-*XbaI* fragment of plasmid pBGG84 between the *SacI*/*XbaI*-sites on plasmid pBAD18-cm. Plasmid constructions were verified by DNA sequencing.

### **Bacterial two-hybrid (BACTH) assays**

The BTH system used in this study is based on reconstitution of adenylate cyclase CyaA activity in *E. coli* strain BTH101, which lacks an endogenous functional *cyaA* gene (Karimova *et al.*, 1998; Karimova and Ladant, 2005). The proteins to be tested are fused to the T18- and T25-fragments of the CyaA protein from *Bordetella pertussis*, respectively. Without any fusion partner, the T18- and T25-domains are incapable to interact. However, when fused to proteins that interact with each other, the T18- and T25-domains assemble into a functional CyaA protein, which synthesizes cAMP. cAMP synthesis is monitored by expression of cAMP-CRP dependent genes such as *lacZ* encoding  $\beta$ -galactosidase. Therefore,  $\beta$ -galactosidase activities reflect the strength of interaction between the proteins fused to the CyaA domains. Fusions of RNase E or of C- or N-terminally truncated RNase E variants to the T25-CyaA fragment were expressed from plasmids pT25-rne, pT25-RE1 and pT25-RE3, respectively. These plasmids (Singh *et al.*, 2009) are derivatives of plasmid pT25, which carries the T25-*cyaA* fragment under control of the *lacUV5* promoter on a p15A-type plasmid (Karimova *et al.*, 1998). Fusions of the same RNase E variants to the T18-CyaA fragment were expressed from plasmids pYG99, pYG97 and pYG98, respectively. The latter plasmids are derivatives of the ColEI-type plasmid pUT18C, which expresses the T18-*cyaA* fragment from the *lacUV5* promoter. Plasmid pUT18C and Plasmid pKT25 were used for the construction of fusions of T18-CyaA and T25-CyaA, respectively, to the N-terminus of potential interaction partners of RNase E (RapZ, RapZ<sub>quad</sub>, Enolase and PtsN). Plasmid pKT25 is similar to plasmid pT25, but carries the *neo* rather than the *cat* antibiotic resistance gene (Karimova *et al.*, 2001). Enolase, which is a previously characterized interaction partner of RNase E (Miczak *et al.*, 1996; Worrall *et al.*, 2008), served as positive control. Strain BTH101 was co-

transformed with the respective plasmids expressing the T18- and T25-fusion proteins to be tested. For monitoring interaction of the fusion proteins phenotypically, the co-transformants were spotted onto LB plates containing the appropriate antibiotics, 1 mM IPTG and 40 µg/ml 5-Brom-4-chlor-3-indoxyl-β-D-galactopyranosid (X-Gal) as indicator for β-galactosidase activity. The plates were incubated at 30°C. A blue coloring of the emerging colonies indicates interaction of the respective T18-CyaA and T25-CyaA fusion proteins. For quantitative β-galactosidase assays, the co-transformants were grown overnight in LB containing 0.5 mM IPTG and the respective antibiotics and were subsequently diluted into 10 ml fresh medium to an OD<sub>600</sub> = 0.1. The cultures were grown at 30°C to stationary growth phase and the β-galactosidase activities were determined.

### **Protein purification**

Strep-tagged recombinant proteins were purified essentially as described previously (Lüttmann *et al.*, 2012) with some deviations. Plasmids pBGG164, pYG29 and pBGG217 were used for overproduction of Strep-RapZ, Strep-RapZ<sub>quad</sub> and Strep-PtsN in the host strain Z106, respectively. The transformants were grown in 1 l LB medium to an OD<sub>600</sub> = 0.5 to 0.8. Synthesis of the recombinant proteins was induced by addition of 1 mM IPTG and growth was subsequently continued for 1 h. Cells were harvested and washed in HEPES buffer (50 mM HEPES, 150 mM NaCl, 1 mM EDTA, pH 8.0) and lysed by passage through a French pressure cell at 18.000 psi. The crude lysates were cleared by low speed centrifugation (8000 r.p.m., 10 min, 4°C) and ultracentrifugation (35000 r.p.m., 1 h, 4°C). The cleared lysates were loaded onto columns containing 1 ml Strep-Tactin matrix (IBA, Germany) that was pre-equilibrated with HEPES buffer. The columns were washed four times using HEPES buffer. Finally, the Strep-tagged proteins were eluted in three steps using 3×1 ml HEPES buffer with 2.5 mM desthiobiotin. The elution fractions containing the pure proteins were dialysed two times for 24 h against dialysis buffer (10 mM Tris-HCl, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 2 mM DTT, pH 7.0). In the second dialysis step, the buffer additionally contained 25% (v/v) glycerol. The purified proteins were stored at -20°C until further use. The N-terminally His<sub>6</sub>-tagged catalytic domain of RNase E (amino acids 1-529 of RNase E) was purified from strain BL21 (DE3) carrying plasmid pRne529-N by Ni<sup>2+</sup>-NTA affinity chromatography as described previously (Callaghan *et al.*, 2003). *E. coli* Hfq protein was purified as described previously (Worrall *et al.*, 2008).

### **In vitro transcription and labeling of RNA**

For internal labeling of RNAs with α-<sup>32</sup>P-UTP, the RNAs were transcribed *in vitro* from PCR fragments. First, DNA fragments were generated, which carried the sRNA genes under control of the T7 promoter, respectively. Primers combinations BG446/BG447 and BG446/BG448 were used for amplification of DNA fragments encoding the full-length and the processed form of GlmY, respectively, using plasmid pBGG149 as a template. Similarly, primers BG444/BG445 were used for amplification of the full-length *glmZ* gene and primers BG444/BG471 for amplification of the DNA fragment corresponding to processed GlmZ from plasmid pBGG84 as template. Mutant alleles of *glmZ* were amplified using primers BG444/BG445 and plasmids pBGG403 (*glmZ-G132T*), pYG51 (*glmZ-G137A*), and pYG52 (*glmZ-G139T*) as templates. The sRNA gene *sraC* was amplified with primers BG527/BG528 from chromosomal DNA as template. *In vitro* transcription reactions contained 100 to 400 ng of the respective PCR fragment, 40 units T7-RNA polymerase (Roche Diagnostics, Germany) 10 mM ATP, 10 mM GTP, 10 mM CTP, 0.1 mM UTP and 20 µCi α-<sup>32</sup>P-UTP (6000Ci/mMol) in transcription buffer in a volume of 20 µl. The reactions were incubated for at least 2 h at 37°C. Unincorporated nucleotides were removed using Illustra MicroSpin G-50 Columns (GE

healthcare). After addition of 2×RNA loading dye (95% Formamide; 18 mM EDTA; 0.025% SDS, 0.01% Xylene cyanol, 0.01% Bromophenol blue) the samples were separated on 7M urea/1×TBE/8% polyacrylamide gels (300 V for 2 h at RT). The wet gels were exposed to phospho-imaging for visualization of the labeled RNA. The RNA was cut out and extracted from the gel by incubation overnight in RNA elution buffer (0.1 M sodium acetate pH 5.2, 10 mM EDTA pH 8.0, 0.1% SDS) under shaking at 1400 rpm. Subsequently, the solution was extracted with phenol:chloroform:isoamyl-alcohol (25:24:1) and the RNA was precipitated with ethanol:sodium acetate pH 5.2 (30:1) for 2 h at -20°C and finally dissolved in 20 µl RNase-free water. RNA concentrations were determined spectrophotometrically using a NanoDrop and appropriate dilutions were prepared in RNase-free water.

5'-end labeling of GlmY and GlmZ was performed using polynucleotide kinase and  $\gamma$ -<sup>32</sup>P-ATP. First, DNA fragments carrying *glmY* and *glmZ* under T7 promoter control were amplified by PCR using oligonucleotides BG971/BG447 and BG972/BG445 and plasmids pBGG149 and pBGG84 as templates, respectively. These amplifications added an additional guanosine residue to the 5'ends of the sRNA genes. The PCR fragments were used as templates for *in vitro* transcription using the SP6/T7 transcription kit (Roche Diagnostics) according the manufacturer's instructions. The RNA was precipitated as described above. 40 pmol of heat denatured RNAs (95°C for 1 min and subsequently 5' on ice) were dephosphorylated by incubation with 10 units alkaline phosphatase (FastAP, Thermo) for 45 min at 37 °C in the supplied buffer. The solutions were extracted with phenol:chloroform:isoamyl-alcohol (25:24:1) and the RNAs were precipitated. The dephosphorylated RNAs were 5'-labeled by incubation with 20 units polynucleotide kinase and 20 µCi  $\gamma$ -<sup>32</sup>P-ATP for 15 min at 50°C in PNK buffer B (Fermentas). Unincorporated nucleotides were removed and the labeled RNAs were isolated from polyacrylamide gels as described above.

#### ***Prediction of RNA binding residues in RapZ***

The software BindN (<http://bioinfo.ggc.org/bindn/>; (Wang and Brown, 2006)) was used for the prediction of putative RNA binding residues in RapZ. A summary of the results of this analysis is given below:

**Input sequence length:** 284 amino acids  
**Predicted binding sites:** 25 residues  
**User-defined specificity:** 96.00%  
**Estimated sensitivity:** 22.72%

### Overview

```

Sequence:  MVLMI VSGRSGSGK SVALRALEDMGFYCVDNLPVVLLPDLARTLADREISA AVSIDVRNM
Prediction: -----++++-+------
Confidence: 999998139989199355089899986999897999977780158426816771538025

Sequence:  PESPEIFEQAMSNLPDAFSPQLLFLDADRNTLIRRYSDTRRLHPLSSKNLSLES AIDKES
Prediction: -----++-+-----
Confidence: 454689868983584375375997858702224100000883804010141742357364

Sequence:  DLLEPLRSRADLIVDTSEMSVHELAEMLRTRLLGKRERELTMVFESFGFKHGIPIDADYV
Prediction: -----+-----
Confidence: 876356010775698443639799978803032001483838998478823775869949

Sequence:  FDVRF L PNP H WDPKLRPMTGLDKPVAAFLDRHTEVHNF IYQTRSYLELWLPMLETNNRSY
Prediction: -----+-----
Confidence: 999388445523416145157616985478264594469343021869797784121821

Sequence:  LTVAIGCTGGK HRSVYIAEQLADYFRSRGKNVQSRHRTLEKRKP
Prediction: -----+--+-----+--+--+--+--+--+--+--+--+
Confidence: 51878638018182728877986348192990808188107872
    
```

\*\*\* Prediction: binding residues are labeled with '+' and in red;  
 non-binding residues labeled with '-' and in green.  
 \*\*\* Confidence: from level 0 (lowest) to level 9 (highest).

## SUPPLEMENTAL TABLES

Table S1. *E. coli* strains used in this study

| Name       | Genotype  | Reference or construction              |
|------------|---|--|
| BL21 (DE3) | F <sup>-</sup> <i>ompT lon gal dcm hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>)</i> λ(DE3 [ <i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i> ]) | (Studier and Moffatt, 1986)            |
| BTH101     | F, <i>cyoA-99, araD139, galE15, galK16, rpsL1 (Str<sup>R</sup>), hsdR2, mcrA1, mcrB1</i>  | (Karimova and Ladant, 2005)            |
| DH5α       | φ80d <i>lacZΔM15, recA1, endA1, gyrA96, thi-1, hsdR17 (r<sub>K</sub><sup>-</sup>, m<sub>K</sub><sup>+</sup>), supE44, relA1, deoR, Δ(lacZYA-argF) U169</i>          | Laboratory stock                       |
| IBPC633    | as N3433, but <i>rnc105, nadB51::Tn10 (tet)</i>   | (Regnier and Hajnsdorf, 1991)          |
| IBPC750    | <i>thi-1, argG6, argE3, his-4, xyl-5, rpsL, ΔlacX74, mlc, ΔglmS::tet</i>  | (Plumbridge and Vimr, 1999)            |
| IBPC935    | as N3433, but <i>rng::cat</i>   | (Bardey <i>et al.</i> , 2005)          |
| JW4130     | Δ( <i>araD-araB</i> )567, Δ <i>lacZ</i> 4787(:: <i>rrnB-3</i> ), <i>LAM-</i> , <i>rph-1, Δ(rhaD-rhaB)568, Δhfq-722::kan, hsdR514</i>                                | (Baba <i>et al.</i> , 2006)            |
| N3431      | as N3433, but <i>rne3071(Ts)</i>  | (Goldblum and Apririon, 1981)          |
| N3433      | <i>HfrH, lacZ43, λ, relA1, spoT1, thi1</i>  | (Goldblum and Apririon, 1981)          |
| R1279      | <i>CSH50 Δ(pho-bgl)201 Δ(lac-pro) ara thi</i>   | (Schnetz <i>et al.</i> , 1996)         |
| TM338      | <i>W3110 mlc rne-3xFLAG- cat</i>  | (Morita <i>et al.</i> , 2004)          |
| Z8         | as R1279, but <i>attB::[aadA, glmS-5':lacZ], strp<sup>R</sup>, F'(pro<sup>+</sup>)</i>  | (Kalamorz <i>et al.</i> , 2007)        |
| Z28        | as R1279, but Δ <i>rapZ, attB::[aadA, glmS-5':lacZ], strp<sup>R</sup>, F'(pro<sup>+</sup>)</i>  | (Kalamorz <i>et al.</i> , 2007)        |
| Z37        | as R1279, but Δ <i>rapZ</i>   | (Kalamorz <i>et al.</i> , 2007)        |
| Z38        | as R1279, but Δ <i>glmZ::cat, attB::[aadA, glmS-5':lacZ], strp<sup>R</sup>, F'(pro<sup>+</sup>)</i>   | (Kalamorz <i>et al.</i> , 2007)        |
| Z44        | as R1279, but Δ <i>glmZ::cat</i>  | (Kalamorz <i>et al.</i> , 2007)        |
| Z45        | as R1279, but Δ <i>glmZ</i>   | (Kalamorz <i>et al.</i> , 2007)        |
| Z64        | as R1279, but <i>rne-3xFLAG- cat</i>  | T4GT7 (TM338)→ R1279; this work        |
| Z95        | as R1279, but Δ <i>glmY::cat</i>  | (Reichenbach <i>et al.</i> , 2008)     |
| Z96        | as R1279, but Δ <i>glmY</i>   | (Reichenbach <i>et al.</i> , 2008)     |
| Z105       | as R1279, but Δ <i>glmY::cat, ΔglmZ</i>   | (Reichenbach <i>et al.</i> , 2008)     |
| Z106       | as R1279, but Δ <i>glmY, ΔglmZ</i>  | Z105 cured from <i>cat</i> ; this work |
| Z115       | as R1279, but Δ <i>rapZ, ΔglmY::cat</i>   | (Reichenbach <i>et al.</i> , 2008)     |
| Z116       | as R1279, but Δ <i>rapZ, ΔglmZ::cat</i>   | (Reichenbach <i>et al.</i> , 2008)     |
| Z479       | as R1279, but Δ <i>rapZ, attB::[araC, P<sub>Ara</sub>::strep-rapZ, bla]</i>   | pYG13/NotI→ Z37; this work             |
| Z555       | as R1279, but Δ <i>rapZ, ΔglmS::tet, attB::[araC, P<sub>Ara</sub>::strep-rapZ, bla]</i>   | T4GT7 (IBPC750)→ Z479; this work       |
| Z591       | as R1279, but <i>rapZ-3xFLAG::kan</i>   | PCR BG853 + 854→ R1279; this work      |
| Z592       | as R1279, but <i>rapZ-3xFLAG</i>  | Z591 cured from <i>kan</i> ; this work |
| Z593       | as R1279, but <i>rapZ-3xFLAG, ΔglmS::tet</i>  | T4GT7 (IBPC750)→ Z592; this work       |
| Z610       | as R1279, but <i>rapZ-3xFLAG, ΔglmY::cat</i>  | T4GT7 (Z95)→ Z592; this work           |
| Z612       | as R1279, but <i>rapZ-3xFLAG, ΔglmZ::cat</i>  | T4GT7 (Z44)→ Z592; this work           |
| Z663       | as R1279, but Δ <i>hfq::kan</i>   | T4GT7 (JW4130)→ R1279; this work       |
| Z664       | as R1279, but Δ <i>hfq</i>  | Z663 cured from <i>kan</i> ; this work |

*strep*, sequence encoding the Strep-tag epitope.

**Table S2.** Plasmids used in this study

| Name      | Relevant structure  | Reference or construction          |
|-----------|---|------------------------------------|
| pBAD18-cm | <i>P<sub>Ara</sub></i> , MCS 2, <i>cat</i> , ori ColEI  | (Guzman <i>et al.</i> , 1995)      |
| pBAD30    | <i>P<sub>Ara</sub></i> , MCS 2, <i>bla</i> , ori p15A   | (Guzman <i>et al.</i> , 1995)      |
| pBAD33    | <i>P<sub>Ara</sub></i> , MCS 2, <i>cat</i> , ori p15A   | (Guzman <i>et al.</i> , 1995)      |
| pBGG16    | <i>glmS'</i> - <i>lacZ</i> , <i>neo</i> , <i>attP</i> , <i>aada</i> , ori p15A                            | (Kalamorz <i>et al.</i> , 2007)    |
| pBGG61    | <i>rapZ</i> (-17 to +855) under <i>P<sub>Ara</sub></i> -control in pBAD33                                 | this work                          |
| pBGG84    | <i>glmZ</i> (-100 to +254) under <i>P<sub>Ara</sub></i> -control in pBAD30                                | (Kalamorz <i>et al.</i> , 2007)    |
| pBGG149   | $\lambda P_{L::glmY}$ (+1 to + 233), <i>bla</i> , ori ColEI   | (Reichenbach <i>et al.</i> , 2008) |
| pBGG164   | <i>strep-rapZ</i> under <i>P<sub>tac</sub></i> control in pBGG237   | (Lüttmann <i>et al.</i> , 2012)    |
| pBGG217   | <i>strep-ptsN</i> under <i>P<sub>tac</sub></i> control in pBGG237   | (Lüttmann <i>et al.</i> , 2012)    |
| pBGG237   | <i>lacI<sup>f</sup></i> , <i>P<sub>tac</sub>-streptag</i> -MCS, <i>rrnBT1/T2</i> , <i>bla</i> , ori ColEI | (Lüttmann <i>et al.</i> , 2012)    |
| pBGG261   | encodes T25-PtsN fusion in pKT25  | (Lüttmann <i>et al.</i> , 2009)    |
| pBGG348   | encodes T25-RapZ fusion in pKT25  | this work                          |
| pBGG349   | encodes T18-RapZ fusion in pUT18C   | this work                          |
| pBGG403   | as pBGG84, but <i>glmZ</i> with G132T mutation  | this work                          |
| pKT25     | <i>P<sub>lac</sub>::cyaA-T25</i> (aa 1-224), MCS, <i>neo</i> , ori p15A                                   | (Karimova <i>et al.</i> , 2001)    |
| pRne529-N | <i>His<sub>6</sub>-rne</i> (+1 to +1587) in pET16b, <i>bla</i> , ori ColEI                                | (Callaghan <i>et al.</i> , 2003)   |
| pT25      | <i>P<sub>lac</sub>::cyaA-T25</i> (aa 1-224), MCS, <i>cat</i> , ori p15A                                   | (Karimova <i>et al.</i> , 1998)    |
| pT25-Rne  | encodes T25-RNase E (full length) fusion in pT25  | (Singh <i>et al.</i> , 2009)       |
| pT25-RE1  | encodes T25-RNase E (aa 1-597) fusion in pT25   | (Singh <i>et al.</i> , 2009)       |
| pT25-RE3  | encodes T25-RNase E (aa 499-1061) fusion in pT25  | (Singh <i>et al.</i> , 2009)       |
| pUT18C    | <i>P<sub>lac</sub>::cyaA-T18</i> (aa 225-399), MCS, <i>bla</i> , ori ColEI                                |                                    |
| pYG13-M   | <i>araC</i> , <i>P<sub>Ara</sub></i> , <i>bla</i> , <i>lattP</i> , <i>cat</i> , ori ColEI                 | this work                          |
| pYG13     | <i>rapZ</i> (-27 to +855) with N-terminal Strep-tag under <i>P<sub>Ara</sub></i> control in pYG13-M       | this work                          |
| pYG23     | <i>glmY</i> (+1 to + 233) under $\lambda P_{L}$ control, <i>cat</i> , ori ColEI                           | this work                          |
| pYG24     | <i>glmZ</i> (-100 to +254) under <i>P<sub>Ara</sub></i> -control in pBAD18-cm                             | this work                          |
| pYG29     | as pBGG164, but <i>rapZ</i> with K270A-K281A-R282A-K283A mutations  | this work                          |
| pYG30     | as pBGG61, but <i>rapZ</i> with K270A-K281A-R282A-K283A mutations   | this work                          |
| pYG42     | encodes T18-RapZ <sub>quad</sub> fusion in pUT18C; RapZ with K270A-K281A-R282A-K283A mutations            | this work                          |
| pYG45     | encodes T18-enolase fusion in pUT18C  | this work                          |
| pYG46     | encodes T18-PtsN fusion in pUT18C   | this work                          |
| pYG51     | as pBGG84, but <i>glmZ</i> with G137A mutation  | this work                          |
| pYG52     | as pBGG84, but <i>glmZ</i> with G139T mutation  | this work                          |
| pYG94     | encodes T25- RapZ <sub>quad</sub> fusion in pKT25   | this work                          |
| pYG95     | encodes T25-enolase fusion in pKT25   | this work                          |
| pYG97     | encodes T18-RNase E (aa 1-597) fusion in pUT18C   | this work                          |
| pYG98     | encodes T18-RNase E (aa 499-1061) fusion in pUT18C  | this work                          |
| pYG99     | encodes T18-RNase E (full length) fusion in pUT18C  | this work                          |

aa, amino acid; Ori, origin of replication; *SD*, Shine-Dalgarno sequence; MCS, multiple cloning site; *strep*, sequence encoding the Strep-tag epitope.

**Table S3.** Oligonucleotides used in this study

| Primer | Sequence <sup>a</sup>   | Res. Sites | Position <sup>b</sup>                             |
|--------|---|------------|---|
| BG209  | TCC <u>CCCGGGAGCT</u> CGTGAGGAGAAACAGTACATG   | SmaI, SacI | <i>rapZ</i> -17 to +3                             |
| BG212  | GGCTCTAGAGGTACCTCATGGTTTACGTTTTCCAG   | XbaI, KpnI | <i>rapZ</i> + 855 to + 834                        |
| BG235  | GGC <u>GAGCTCT</u> CAGGAAGTTATTACTCAGGAAGC  | SacI       | <i>glmZ</i> -100 to -77                           |
| BG399  | CTCGTACTCATATGTGGAGCCACCCGAGTTCGAAAAAGCTAGCATGGTACTGATGATCGTCAG<br>CGG              | NdeI, NheI | <i>rapZ</i> +1 to +23                             |
| BG444  | CTAATACGACTCACTATAGGGAGAGTAGATGCTCATTCCATCTCTTATG                                   |            | <i>glmZ</i> +1 to +25                             |
| BG445  | AAAAAACGCCTGCTCTTATTACGGAGC   |            | <i>glmZ</i> +207 to + 179                         |
| BG446  | CTAATACGACTCACTATAGGGAGAAAGTGGCTCATTACCGACTTATGTC                                   |            | <i>glmY</i> +1 to +25                             |
| BG447  | AACAAAGCCGGGAATTACCCGGC   |            | <i>glmY</i> +184 to +161                          |
| BG448  | AAGGCGGTGCCTAACTCGACG   |            | <i>glmY</i> +148 to +127                          |
| BG471  | CAACAAGTGGGTGCTTCACTC   |            | <i>glmZ</i> +154 to +135                          |
| BG527  | CTAATACGACTCACTATAGGGAGAAAGTCAAGCGAAGGAAATG   |            | <i>sraC</i> +1 to +19                             |
| BG528  | ATCACCAGAACGGGCGG   |            | <i>sraC</i> +249 to +234                          |
| BG637  | GCGTCTAGAGATGGTACTGATGATCGTCAGCG  | XbaI       | <i>rapZ</i> +1 to +22                             |
| BG639  | CGCGGTACCTCATGGTTTACGTTTTCCAGCG   | KpnI       | <i>rapZ</i> +855 to +833                          |
| BG709  | ATAGAGCGAGGACGCTAACAG   |            | <i>csrC</i> +1 to +21                             |
| BG710  | CTAATACGACTCACTATAGGGAGAAAGAAAAAGGCGACAGATTACTC                                     |            |   |
| BG745  | [P]-CATGGACACAACCTTTGAGTGAAGCACCCAC   |            | <i>glmZ</i> +120 to +149                          |
| BG746  | CGTCTCTAGATTCCTTCTACCCGGAGGC  | XbaI       | <i>glmZ</i> +254 to +236                          |
| BG765  | GGC <u>GAGCTCTGCA</u> CTTTAGACGTTGTGAGGAGAAACAGTACATGTGGAGCCACCCGCG                 | SacI, Sall | <i>rapZ</i> -27 to -1, <i>strep-tag</i> +1 to +18 |
| BG766  | GGC <u>GAGCTCT</u> CTAGATCATGGTTTACGTTTTCCAGCG                                      | SacI, XbaI | <i>rapZ</i> + 855 to + 832                        |
| BG853  | GGTAAAAACGTCCAGTCACGCCATCGTACGCTGAAAAACGTAAACCAGACTACAAAGACCAT<br>GACGG             |            | <i>rapZ</i> +805 to +852                          |
| BG854  | ATGCATGCCAGCTTGTGGTATTCAACAGTTTGGTGGACGGTCATATGAATATCCTCCTTAG<br>TTCCTATTCC         |            | <i>ptsO</i> +48 to +3                             |
| BG920  | GGCTCTAGATTATCATGGTGCAGCTGCTTCCAGCGTACGATGGCGTGACTGGACGTTTGCACCG<br>CGCGAGCGGAAGTAG | XbaI       | <i>rapZ</i> +855 to +789                          |
| BG971  | CTAATACGACTCACTATAGGGAGAGAGTGGCTCATTACCGACTTATGTC                                   |            | <i>glmY</i> +1 to +25                             |
| BG972  | CTAATACGACTCACTATAGGGAGAGGTAGATGCTCATTCCATCTCTTATG                                  |            | <i>glmZ</i> +1 to +25                             |
| BG973  | CGTGGTACCTCATGGTGCAGCTGCTTCC  | KpnI       | <i>rapZ<sub>quad</sub></i> +855 to +837           |
| BG1017 | GCGTCTAGAGATGACAAATAATGATACAACCTC   | XbaI       | <i>ptsN</i> +1 to +22                             |
| BG1018 | CGCGGTACCTTACGCTTATCCGGAGTAC  | KpnI       | <i>ptsN</i> +492 to +473                          |
| BG1025 | GCGTCTAGAGATGTCCAAAATCGTAAAAATCAT   | XbaI       | <i>eno</i> +1 to +23                              |
| BG1026 | CGCGGTACCTTATGCCTGGCCTTTGATCT   | KpnI       | <i>eno</i> +1299 to +1279                         |
| BG1029 | [P]- GACACCATGGACACAACGTTGAATGAAGCACCCACTTGTG                                       |            | <i>glmZ</i> +115 to +155                          |
| BG1030 | [P]-CATGGACACAACGTTGAGTTAAGCACCCACTTGTGTC   |            | <i>glmZ</i> +120 to +157                          |
| BG1110 | GCGTCTAGAGATGAAAAGAATGTTAATCAACGC   | XbaI       | <i>rne</i> +1 to +23                              |
| BG1111 | CGCGGTACCTTACTCAACAGTTGCGGAC  | KpnI       | <i>rne</i> +3186 to 3167                          |
| BG1113 | GCGTCTAGAGAGTACATGCTGCCGAAGC  | XbaI       | <i>rne</i> +1495 to +1514                         |
| BG1114 | CGCGGTACCTTATTTCCGGTGTGTTGCTCGG   | KpnI       | <i>rne</i> +1791 to +1774                         |

<sup>a</sup>Restriction sites are underlined; Positions deviating from the wild-type sequence are in bold; [P] indicates 5'-phosphorylation of the oligonucleotide. The recognition site for T7 RNA polymerase is underlined by a dashed line; <sup>b</sup>Positions are relative to the first nucleotide of the respective gene.

## Molecular requirements for Hfq binding and processing by RNase E within the regulatory GlmY/GlmZ sRNA cascade

Göpel, Y., Khan, M.A., and Görke, B.

*The results presented in this Chapter are unpublished.*

*A manuscript dealing with the results presented here is currently in preparation.*

### *Author contributions:*

This study was designed by Y.G. and B.G.. Y.G. performed gel shift assays with Strep-RapZ and Hfq. Gel retardation analysis of hybrid GlmYZ-H2 in regard to RapZ binding was conducted by M.A.K. Protein purification was performed by Y.G. and M.A.K.. RNA isolation and northern blotting were conducted by M.A.K. and Y.G.. RNase E *in vitro* cleavage assays were done by Y.G.. Strains and plasmids were constructed by M.A.K. under the supervision of Y.G.. Y.G. wrote the manuscript, B.G. offered insightful discussion of the manuscript.

## ABSTRACT

Homologous small RNAs GlmY and GlmZ act in a hierarchical regulatory cascade to fine-tune expression of glucosamine-6-phosphate (GlcN6P) synthase GlmS thereby mediating GlcN6P homeostasis. Though vastly similar, GlmY and GlmZ employ different modes of action. Whereas GlmZ is Hfq-dependent and acts by base-pairing to activate the *glmS* transcript, GlmY does not depend on Hfq and acts indirectly by sequestering RapZ. Adaptor protein RapZ is strictly required for processing of GlmZ by RNase E and its subsequent degradation. In contrast, although highly similar, GlmY is not degraded by this machinery. However, the reason for this difference is unknown. Thus, the regulatory GlmYZ-cascade provides an excellent system to study molecular requirements for binding of sRNAs to RNA chaperon Hfq as well as for recognition processing of sRNAs by RNase E. By using a library of hybrid GlmY/GlmZ sRNAs, we demonstrate that the 3' end of GlmZ, but not of GlmY, provides a high affinity binding site for Hfq functional *in vitro* and *in vivo*. In contrast, the molecular determinants for recognition and efficient cleavage by RNase E reside within the 5' region of sRNA GlmZ. We demonstrate that swapping of the lateral bulge in the central stem loop of the two sRNAs is sufficient to enable processing of the non-substrate sRNA GlmY by RNase E *in vitro* independently of its 5' phosphorylation state.

## INTRODUCTION

Post-transcriptional regulation employing small regulatory RNAs (sRNAs) has emerged as a wide spread mechanism of gene regulation not only in eukaryotes but also in bacteria and archaea. As the number of known sRNAs increases, it becomes apparent that sRNAs are involved in virtually all physiological circuits controlling for example metabolite homeostasis or providing a quick response to various stresses (Storz *et al.*, 2011; Gottesman and Storz, 2011). Thus, due to the regulatory potential of sRNAs, their activities have to be tightly controlled. At the level of transcription, the expression of sRNAs is elaborately controlled by specific transcriptional regulators, two-component systems as well as alternative sigma factors (Göpel and Görke, 2012a; Mandin and Guillier, 2013). As opposed to transcriptional control, some insight into regulation at the level of decay has only recently been obtained (Göpel *et al.*, 2013).

At least in Gram negative bacteria, most *trans*-acting base-pairing sRNAs require the hexameric RNA chaperon Hfq either for annealing to their specific target transcripts and/or for stability (Urban and Vogel, 2007; Vogel and Luisi, 2011). The ring-shaped Hfq hexamer possesses three distinct binding surfaces that aid in facilitating annealing of cognate mRNA/sRNA pairs and stabilizing resulting complexes. The distal surface effectively binds A-rich sequences and ARN repeats (R denotes a purine and N denotes any nucleotide) whereas the proximal face has been shown to interact with protruding poly-(U) sequences, often succeeding Rho-independent terminators of sRNAs (Link *et al.*,

2009; Otaka *et al.*, 2011; Sauer and Weichenrieder, 2011). The recently discovered lateral interface, or rim, provides a third binding site that is likely to protect sRNAs from degradation and expedites duplex formation with target mRNAs (Sauer *et al.*, 2012). Although Hfq interacts with a multitude of sRNAs and the 3' poly-(U)-stretch has been shown to be crucial for recognition by Hfq, it is still not completely understood how Hfq discriminates between the various sRNAs that compete for binding.

Major regulators of sRNA abundance are the endoribonucleases RNase III, RNase E and polynucleotide phosphorylase PNPase that acts as a 3' to 5' exoribonuclease (Storz *et al.*, 2011; Vogel and Luisi, 2011). While RNase III cleaves double stranded RNA and thus plays a major role in turn-over of highly structured RNAs, RNase E seems to have little sequence specificity and preferably cleaves within single-stranded A/U-rich regions (Carpousis, 2007; Kim *et al.*, 2004). In *E. coli*, RNase E composes the scaffolding core of a multi-enzyme RNA degradosome that cooperatively promotes turn-over of bulk RNA (Górna *et al.*, 2012). The degradosome, and to a lesser extent RNase III, are involved in rapid and coupled degradation of target transcripts upon base-pairing with their cognate sRNA repressors (Massé *et al.*, 2003). However, how sRNAs are processed independently of their targets is less understood. Initial cleavage might either lead to subsequent degradation by exoribonucleases such as PNPase or might even enhance the regulatory potential of the sRNA by generating a mature variant (Andrade *et al.*, 2012; Papenfort *et al.*, 2009; Soper *et al.*, 2010).

RNase E employs two distinct modes of substrate recognition. First, catalytic activity of RNase E can be activated by interaction of the 5' mono-phosphorylated end of the substrate RNA with its sensory domain (Callaghan *et al.*, 2005). In case of sRNA MicC for example, the coupled decay with its target *ompD* was reported to involve *trans*-stimulation of RNase E activity by a 5' mono-phosphate of sRNA MicC (Bandyra *et al.*, 2012). Thus, sRNAs are not only capable to induce degradation passively by blocking access for ribosomes but might even act as allosteric activators of RNase E activity (Mackie, 2013b). However, this mode of action relies on the generation of 5' mono-phosphorylated RNA species to initiate decay, for example by pyrophosphohydrolase RppH (Deana *et al.*, 2008). Second, RNase E is capable to recognize some target transcripts independently of their phosphorylation state in a process referred to as 'direct entry' (Kime *et al.*, 2010; Mackie, 2013a). Yet, the molecular requirements that discriminate if an (s)RNA can be recognized by RNase E via direct entry are poorly understood. Whether there are structural requirements that have to be met or specific protein co-factors that are involved in direct entry is still to be determined for most substrates of RNase E. Recently, an example for such a specific co-factor was identified. RNase E adaptor protein RapZ was shown to promote cleavage of sRNA GlmZ and thus mediate glucosamine-6-phosphate (GlcN6P) homeostasis (Göpel *et al.*, 2013). Other examples may include the CsrD protein and RHON1. CsrD is involved in the Csr-system and is required for turn-over of sRNAs CsrB and CsrC by RNase E (Suzuki *et*

*al.*, 2006). However, if this process involves interaction between CsrD and RNase E remains to be clarified. Further, RNA-binding protein RHON1 of *Arabidopsis thaliana* was recently shown to target specific plastid transcripts for processing by RNase E (Stoppel *et al.*, 2012).

Unlike other homologous sRNAs, GlmY and GlmZ act in a regulatory cascade that feed-back controls synthesis of glucosamine-6-phosphate synthase GlmS (Görke and Vogel, 2008; Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban and Vogel, 2008). Exclusively GlmZ can base-pair with an inhibitory structure within the 5' UTR of its target mRNA *glmS*, thus facilitating translation. As GlmY does not possess the nucleotide sequence complementary to the *glmS* hairpin and is unable to bind RNA chaperon Hfq with high affinity it acts indirectly to activate *glmS* expression (Reichenbach *et al.*, 2008; Göpel *et al.*, 2013). Under conditions of ample GlcN6P GlmZ is processed by RNase E, which removes the base-pairing nucleotides and initiates degradation. This process absolutely relies on adaptor protein RapZ that binds to GlmZ and recruits the processing machinery through direct interaction with RNase E. Upon depletion of GlcN6P, sRNA GlmY accumulates and acts as an anti-adaptor sequestering protein RapZ, thus stabilizing sRNA GlmZ. GlmZ is now available for ternary complex formation with Hfq and *glmS* that leads to synthesis of GlmS. Therefore, both sRNAs although vastly similar act by different mechanisms, i.e. Hfq-dependent base-pairing vs. Hfq-independent protein sequestration. Similarly, while GlmY sequesters RapZ and thereby protects GlmZ from cleavage by RNase E, GlmY itself is no substrate of RNase E and is not processed in this pathway *in vivo* (Göpel *et al.*, 2013). Thus, the GlmY/Z cascade provides an ideal opportunity to study the molecular requirements for Hfq-binding and RNase E-dependent processing.

In this work, we created a library of GlmYZ hybrid sRNAs to investigate the molecular determinants for recognition by Hfq and RNase E. Here, we show that although the poly-(U)-tail succeeding the Rho-independent terminator of GlmZ is important for binding of Hfq, it is not sufficient. Rather, the single-stranded A/U-rich region adjacent to the terminator stem and the poly-(U)-sequence are both required for high affinity binding by Hfq. The ability to efficiently bind Hfq can be transferred to the Hfq-independent sRNA GlmY by swapping 3' ends with GlmZ. Further, we show that this substitution renders the hybrids dependent on Hfq for stability, indicating that this module is sufficient to create an sRNA hybrid that is able to compete for Hfq-binding *in vivo*. In contrast, the 3' ends of the sRNAs did not contribute the recognition and processing by RNase E. Rather, we identified a specific region within the second stem loop of GlmZ that is of utmost importance for cleavage by RNase E and transferring this region to GlmY is sufficient to induce cleavage of this non-substrate sRNA.

## MATERIALS AND METHODS

### ***Growth conditions and construction of strains and plasmids***

LB served as standard medium for cultivation of *E. coli*. Cells were grown under agitation (200 r.p.m.) at 37°C and for selection antibiotics were added to the medium in the following concentrations: ampicillin 100 µg/ml, kanamycin 30 µg/ml, chloramphenicol 15 µg/ml. For induction of expression of genes placed under the control of the  $P_{tac}$  promoter 1 mM IPTG was added.

The *E. coli* strains used in this study and their relevant genotypes are given in Table 5.1. For construction of the  $\Delta glmY \Delta glmZ \Delta hfq$  triple mutant strain Z865, the previously established *hfq::kan* allele was transduced into the pre-existing  $\Delta glmY \Delta glmZ$  double mutant Z106 using phage T4GT7 (Wilson *et al.*, 1979). In a second step, the kanamycin resistance cassette flanked by FLP recognition sites was removed from strain Z863 using the temperature-sensitive FLP recombinase expression plasmid pCP20 as described before (Datsenko and Wanner, 2000). All mutations were verified by diagnostic PCRs.

Cloning of DNA was performed routinely in *E. coli* strain DH5 $\alpha$  according to standard procedures (Sambrook and Russell, 2001). The plasmids and oligonucleotides employed in this study are listed in Tables 5.2 and 5.3, respectively. For construction of plasmids pYG83 and pYG84 carrying *glmY* or *glmZ* under control of the hybrid  $\lambda P_L/P_{tac}$  promoter, sRNA genes were amplified by PCR using primers BG1100/BG1101 for amplification of *glmY* and BG1102/BG1103 for *glmZ*, respectively. Following digestion with *AatII/EcoRI*, the DNA-fragments were ligated between the *AatII/EcoRI* sites of plasmid pBR-pLac. This plasmid allows expression of the sRNAs from their authentic start site. Plasmids pYG85 and pYG86 encompassing hybrids GlmYZ-H2 and GlmYZ-H1, were constructed by using PCR products obtained with primers BG1102/BG1101 and pBGG331 as template for GlmYZ-H2 as well as primers BG1100/BG1103 and template pBGG332 for GlmYZ-H1. PCR products were subjected to restriction with *AatII/EcoRI* and subsequently inserted into pBR-pLac previously opened with *AatII/EcoRI*. For construction of plasmids pYG87 and pYG88, DNA fragments encompassing the sequence for hybrids GlmYZ-H3 and GlmYZ-H4 were obtained by PCR using primers BG1100/BG1104 and pBGG149 as template for GlmYZ-H3 as well as BG1100/BG1105 together with pBGG332 as template for GlmYZ-H4. After digestion with *AatII/EcoRI*, the obtained DNA fragments were ligated between the *AatII/EcoRI* sites of plasmid pBR-pLac.

Construction of plasmids pBGG331 and pBGG332 required two separate PCR reactions, the products of which were later fused using a third, overlap-extension PCR. For pBGG331, the fragment encompassing *glmZ* (-100 to +146) was obtained with primers BG235/BG603 using pBGG84 as template. The second fragment encompassing *glmY* (+144 to +202) was acquired by PCR using primers BG604/BG602 together with pBGG296 as template. Both fragments were fused by overlap-extension PCR: in a first step the two separate fragments were annealed making use of the 20 bp complementary overhang introduced by primer BG603. In a second step, the fusion PCR product was obtained with primers BG235/BG602. The resulting fused DNA-fragment was isolated from an agarose-gel, digested with *SacI/ApaI* and integrated in pBAD30 opened with the same restriction enzymes, thus, making use of the internal *ApaI* site. Plasmid pBGG332 was constructed essentially as described for plasmid pBGG331, but using different primer combinations. For the first fragment encompassing *glmY* (+1 to +143) primers BG289/BG605 were used together with pBGG296 as template and the second DNA fragment corresponding to *glmZ* (+146 to +254) was obtained by PCR using primers BG606/BG602 and pBGG84 as template. Fusion of both fragments was achieved by amplification with primers BG289/BG602 and the resulting fusion insert was integrated between the *SacI/ApaI* of pBAD30.

For construction of plasmid pBGG296, the *glmY* gene (+1 to + 202) was PCR-amplified using primers BG289/BG557 and digested using *SacI/XbaI* restriction enzymes. The resulting fragment was then integrated between the *SacI/XbaI* sites of plasmid pBAD18-cm.

Plasmids pYG103, pYG104, pYG105 and pYG106 were obtained by gene synthesis of the respective insert sequences (*glmYZ-H5*, *glmYZ-H6*, *glmYZ-H7* and *glmYZ-H8*, respectively) followed by sub-cloning. Gene synthesis, as well as sub-cloning of plasmids pYG103 and pYG106 using the *AatII/EcoRI* of pBR-pLac was provided by Eurofins MWG operon (Ebersberg, Germany). Sub-cloning of plasmids pYG104 and pYG105 was achieved by PCR amplification of the synthesized inserts using primers BG1100/BG1101 for *glmYZ-H6* and BG1102/BG1103 for *glmYZ-H7*. Subsequently, the PCR fragments were restricted with *AatII/EcoRI* and inserted into pBR-pLac previously opened with *AatII/EcoRI*. Plasmid constructions were verified by analytical digestion and DNA sequencing.

### **RNA isolation and Northern Blot analysis**

Isolation of total RNA and Northern blot analysis were performed as described previously (Reichenbach *et al.* 2008, Reichenbach *et al.* 2009). Digoxigenin-labeled RNA probes were obtained by *in vitro* transcription of PCR products.

### **Protein purification**

Strep-tagged recombinant RapZ was purified as described before (Lüttmann *et al.*, 2012). Briefly, plasmid pBGG164 encoding for Strep-RapZ was introduced into strain Z106 and transformants were grown in 1 l LB medium to an  $OD_{600} = 0.5$  to 0.8. Expression and over-production of Strep-RapZ was induced by addition of 1 mM IPTG and growth was continued for 1 hour. Cells were harvested and washed in buffer W (100 mM Tris-HCl at pH 8.0, 150 mM NaCl, 1 mM EDTA) and lysed using a French pressure cell at 18.000 psi. Lysates were cleared by low speed centrifugation (10 min, 4°C, 8000 r.p.m.) and subsequent ultracentrifugation (1 h, 6°C, 35000 r.p.m.). Cleared lysates were loaded onto columns containing 1 ml Strep-Tactin matrix (IBA, Germany) pre-equilibrated with buffer W. Following four washing step using 10 ml buffer W, each; Strep-RapZ was eluted in three steps using 3×1 ml buffer W containing 2.5 mM desthiobiotin. Elution fractions were pooled and dialyzed twice for 16 h against dialysis buffer (10 mM Tris-HCl, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 2 mM DTT, pH 7.0) and 25% (v/v) glycerol was added to the buffer prior to the second dialysis. Purified Strep-RapZ was stored at -70°C, for long term storage or -20°C for short term storage.

For Ni<sup>2+</sup>-NTA affinity chromatography of the N-terminally His<sub>6</sub>-tagged catalytic domain of RNase E (amino acids 1-529) plasmid pRne529-N was introduced into BL21 (DE3) and RNase E NTD was purified as described previously (Callaghan *et al.*, 2003).

Purified Hfq was a gift from J. Vogel (University of Würzburg, Germany).

### **Radioactive labeling of sRNAs**

Radioactive labeling of sRNAs was performed as described previously (Göpel *et al.*, 2013). In general, the RNAs were transcribed *in vitro* from PCR fragments and simultaneously internally labeled with  $\alpha$ -<sup>32</sup>P-UTP. To obtain PCR products encompassing the sRNA genes under control of the T7 promoter, the following primer combinations and templates were used. Primers BG446/BG447 together with template pBGG149 yielded fragments encoding *GlmY*. For amplification of fragments encoding hybrids *GlmYZ-H4*, *GlmYZ-H6* and *GlmYZ-H7*, the same primer combination was used together with pYG88, pYG104 or pYG105, respectively. Similarly, primers BG444/BG445 were used for amplification of the *glmZ* gene from plasmid

pYG84. For amplification of hybrid GlmYZ-H8 primers BG444/BG445 were used together with plasmid pYG106. Hybrid GlmYZ-H2 was amplified from pYG85 using primers BG444/BG447 and DNA-fragments encoding for hybrids GlmYZ-H1 and GlmYZ-H3 were obtained by PCR using primers BG446/BG445 and templates pYG86 or pYG87, respectively.

*In vitro* transcriptions were performed in 20 µl reactions containing transcription buffer, 200 ng of the respective PCR fragment, 40 units T7-RNA polymerase (Roche Diagnostics, Germany), 10 mM ATP, GTP, and CTP, as well as 0.1 mM UTP and 20 µCi  $\alpha$ -<sup>32</sup>P-UTP (6000Ci/mMol). After 4-5 h at 37°C, unincorporated nucleotides were removed using Illustra MicroSpin G-50 Columns (GE healthcare). Samples containing RNA loading dye (2×RNA-loading dye: 95% Formamide; 18 mM EDTA; 0.025% SDS, 0.01% Xylene cyanol, 0.01% Bromophenol blue) were separated on 7M urea/1×TBE/8% polyacrylamide gels at 300 V for 2 h at RT. Gels were exposed to phosphorImaging for visualization of the labeled RNA and RNA was excised. For RNA extraction, the gel fragments were incubated overnight at 4°C in RNA elution buffer (0.1 M sodium acetate pH 5.2, 10 mM EDTA pH 8.0, 0.1% SDS). Following phenol:chloroform:isoamyl-alcohol (25:24:1) extraction, the RNA was precipitated with ethanol:sodium acetate pH 5.2 (30:1) for 3 h at -20°C and finally dissolved in 20 µl RNase-free water. Appropriate dilutions were prepared in RNase-free water.

### **Electrophoretic mobility shift assays (EMSA)**

EMSA experiments were performed as described before (Göpel *et al.*, 2013; Sittka *et al.*, 2007). In short, approximately 4 nM heat-denatured, <sup>32</sup>P-labeled sRNA and 1 µg of yeast tRNA (final concentration: 4.3 µM) were incubated for 30 minutes at 30°C with increasing concentrations of purified Strep-RapZ or Hfq. All EMSA reactions were carried out a volume of 10 µl containing 1x structure buffer (10 mM Tris at pH 7, 100 mM KCl, 10 mM MgCl<sub>2</sub>).

Protein dilutions were calculated for the respective monomers and prepared in 1x structure buffer. After addition of 2 µl of native loading buffer (50% glycerol, 0.5× TBE, 0.2% bromphenolblue), samples were separated on native gels (8% polyacrylamide/1×TBE) in 0.5×TBE at 300 V, 4°C for ~3 hours. Gels were dried and analyzed by PhosphorImaging using a Storm 860 Imager and ImageQuant software (Amersham Biosciences).

### **RNase E *in vitro* cleavage assay**

RNase E *in vitro* cleavage assays were performed essentially as described before (Corcoran *et al.*, 2012; Göpel *et al.*, 2013). Briefly, ~40 nM internally <sup>32</sup>P-labeled sRNAs and GlmYZ-hybrids were denatured at 70°C for 2 minutes followed by incubation on ice for 5 minutes and re-naturing at 30°C for 5 minutes in 8 µl 1x reaction buffer (25 mM Tris-HCl at pH 7.5, 50 mM NaCl, 50 mM KCl, 10 mM MgCl<sub>2</sub>, 1 mM DDT) with 0.1 mg/ml yeast tRNA (Ambion). Next, 1 µl 1x reaction buffer containing Strep-RapZ was added as indicated, in mock samples, Strep-RapZ was omitted. Incubation was continued for 10 min and RNase E NTD (1-529 aa) was added in the indicated final concentrations. After 20 min, cleavage reactions were stopped by addition of 4 u Proteinase K (Fermentas) in PK-buffer (100 mM Tris/HCl pH 7.5, 12.5 mM EDTA, 150 mM NaCl, 1% SDS). The reaction was incubated for 10 minutes at 50°C in order to allow for proteolysis of RNase E. Subsequently, 2x RNA loading dye (95% Formamide; 18 mM EDTA; 0.025% SDS, Xylene Cyanol, and Bromophenol Blue) was added and samples were separated on denaturing 7M urea/TBE/8% polyacrylamide gels. Radioactively labeled RNA was visualized on dried gels by phosphorImaging.

**Table 5.1.** *E. coli* strains used in this study

| Name         | Genotype   | Reference or construction  |
|--------------|--|--|
| JW4130       | $\Delta(\text{araD-araB})567, \Delta\text{lacZ4787}::(\text{rrnB-3}), \text{LAM-}, \text{rph-1}, \Delta(\text{rhaD-rhaB})568, \Delta\text{hfq-722}::\text{kan}, \text{hsdR514}$  | (Baba <i>et al.</i> , 2006)  |
| DH5 $\alpha$ | $\phi 80d \text{ lacZ}\Delta\text{M15}, \text{recA1}, \text{endA1}, \text{gyrA96}, \text{thi-1}, \text{hsdR17} (\text{r}_{\text{K}}, \text{m}_{\text{K}}^+), \text{supE44}, \text{relA1}, \text{deoR}, \Delta(\text{lacZYA-argF}) \text{U169}$ | Laboratory collection  |
| R1279        | CSH50 $\Delta(\text{pho-bgl})201 \Delta(\text{lac-pro}) \text{ara thi}$  | (Schnetz <i>et al.</i> , 1996)   |
| Z106         | As R1279, but $\Delta\text{glmZ}, \Delta\text{glmY}$   | (Göpel <i>et al.</i> , 2013)   |
| Z664         | As R1279, but $\Delta\text{hfq}$   | (Göpel <i>et al.</i> , 2013)   |
| Z863         | As R1279, but $\Delta\text{glmZ}, \Delta\text{glmY}, \Delta\text{hfq}::\text{kan}$   | This work, $\Delta\text{hfq}::\text{kan}$ transduced from JW4130 to Z106 |
| Z865         | As R1279, but $\Delta\text{glmZ}, \Delta\text{glmY}, \Delta\text{hfq}$   | This work, Z863 cured from <i>kan</i>                                    |

**Table 5.2.** Plasmids used in this study

| Name      | Relevant structure   | Reference or construction          |
|-----------|--|------------------------------------|
| pBAD30    | $P_{\text{ara}}$ -promoter, pACYC-ori; MCS 2, <i>bla</i>   | (Guzman <i>et al.</i> , 1995)      |
| pBAD18-cm | $P_{\text{ara}}$ -promoter, pBR322-ori; MCS 2, <i>cat</i>  | (Guzman <i>et al.</i> , 1995)      |
| pBR-pLac  | $\lambda\text{P}_{\text{L}}/\text{P}_{\text{lacO-1}}, \lambda\text{P}_{\text{R}}, \text{ColEI-ori}, \text{tet}, \text{bla}$  | (Guillier and Gottesman, 2006)     |
| pBGG84    | <i>glmZ</i> (-100 to +254) under $P_{\text{ara}}$ control in pBAD30  | (Kalamorz <i>et al.</i> , 2007)    |
| pBGG149   | ori pMB1, MCS, <i>bla</i> , $\lambda\text{P}_{\text{L}}::\text{glmY}$  | (Reichenbach <i>et al.</i> , 2008) |
| pBGG164   | <i>strep-rapZ</i> under $P_{\text{lac}}$ control in pBGG237  | (Lüttmann <i>et al.</i> , 2012)    |
| pBGG237   | <i>lacI<sub>q</sub></i> , $P_{\text{lac-streptag}}$ -MCS, <i>rrnBT1/T2</i> , <i>bla</i> , ori ColEI  | (Lüttmann <i>et al.</i> , 2012)    |
| pBGG296   | <i>glmY</i> (+1 to +202) under $P_{\text{ara}}$ control in pBAD18-cm   | This work, Laboratory collection   |
| pBGG331   | hybrid <i>glmYZ-H2</i> : <i>glmZ</i> (-100 to +146) fused to <i>glmY</i> (+144 to +202) in pBAD30  | This work, Laboratory collection   |
| pBGG332   | hybrid <i>glmYZ-H1</i> : <i>glmY</i> (+1 to +143) fused to <i>glmZ</i> (+147 to +254) in pBAD30  | This work, Laboratory collection   |
| pYG83     | <i>glmY</i> (+1 to +202) in pBR-plac   | This work                          |
| pYG84     | <i>glmZ</i> (+1 to +226) in pBR-plac   | This work                          |
| pYG85     | hybrid <i>glmYZ-H2</i> : <i>glmZ</i> (+1 to +146) fused to <i>glmY</i> (+144 to +202) in pBR-plac  | This work                          |
| pYG86     | hybrid <i>glmYZ-H1</i> : <i>glmY</i> (+1 to +143) fused to <i>glmZ</i> (+147 to +226) in pBR-plac  | This work                          |
| pYG87     | hybrid <i>glmYZ-H3</i> : <i>glmY</i> (+1 to +155) fused to <i>glmZ</i> (+173 to +207), <i>glmY</i> (+185 to +203) in pBR-plac  | This work                          |
| pYG88     | hybrid <i>glmYZ-H4</i> : <i>glmY</i> (+1 to +143) fused to <i>glmZ</i> (+147 to +173) and <i>glmY</i> (+156 to +203) in pBR-plac   | This work                          |
| pYG103    | hybrid <i>glmYZ-H5</i> : <i>glmZ</i> (+1 to +65) fused to <i>glmY</i> (+67 to +203) in pBR-plac  | This work                          |
| pYG104    | hybrid <i>glmYZ-H6</i> : <i>glmY</i> (+1 to +71) fused to <i>glmZ</i> (+73 to +146) and <i>glmY</i> (+144 to +203) in pBR-plac   | This work                          |
| pYG105    | hybrid <i>glmYZ-H7</i> : <i>glmY</i> (+1 to +71) fused to <i>glmZ</i> (+73 to +85), <i>glmY</i> (+85 to +120), <i>glmZ</i> (+124 to +146) and <i>glmY</i> (+144 to +203) in pBR-plac | This work                          |
| pYG106    | hybrid <i>glmYZ-H8</i> : <i>glmZ</i> (+1 to +72) fused to <i>glmY</i> (+72 to +84), <i>glmZ</i> (+86 to +123), <i>glmY</i> (+121 to +143) and <i>glmZ</i> (+147 to +226) in pBR-plac | This work                          |

Ori, origin of replication, MCS, multiple cloning site, P, promoter <sup>a</sup>Positions given are relative to the first nucleotide of the respective gene. Gene names according to <http://ecocyc.org/>

**Table 5.3.** Oligonucleotides used in this study

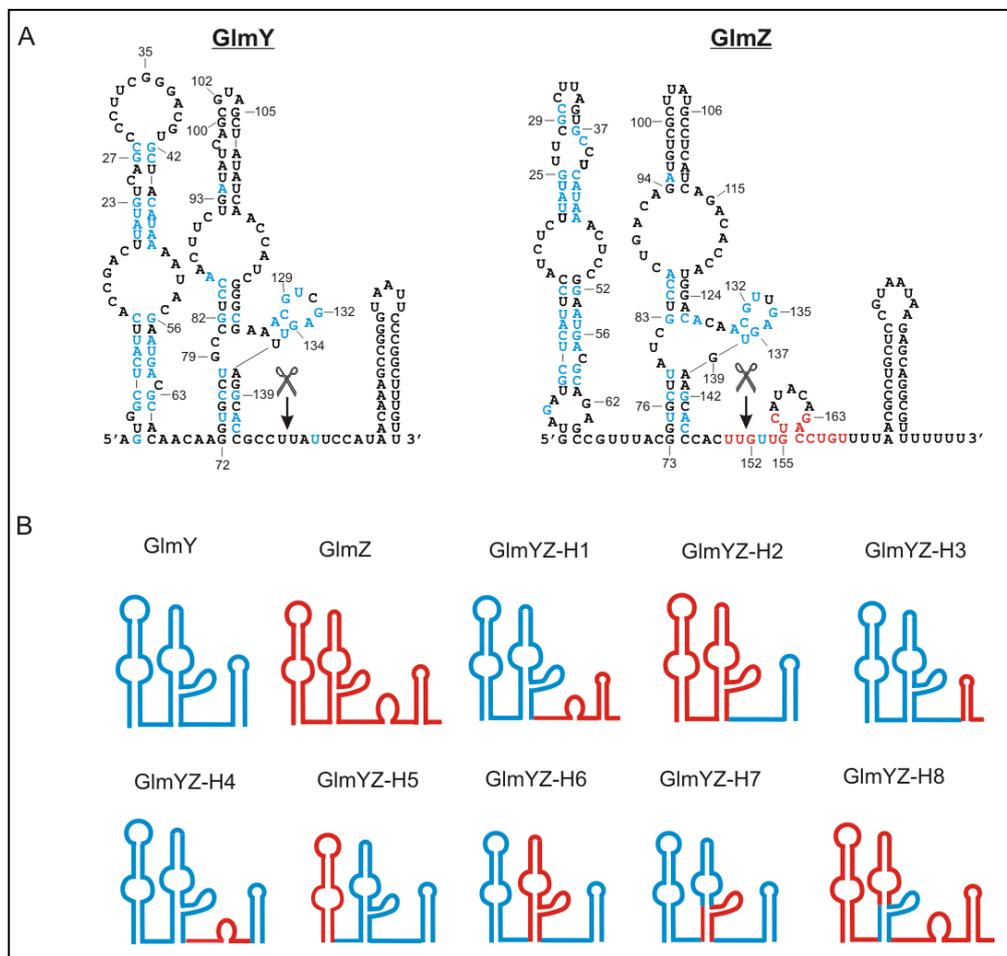
| Primer | Sequence <sup>a</sup>  | Res.Sites    | Position <sup>b</sup>                       |
|--------|--|--------------|---|
| BG235  | <u>GGCGAGCTCT</u> CAGGAAGTTATTACTCAGGAAGC  | <u>SacI</u>  | <i>glmZ</i> (-100 to -76)                   |
| BG289  | <u>GGCGAGCTC</u> AGTGGCTACATCCCGAC   | <u>SacI</u>  | <i>glmY</i> (+1 to +18)                     |
| BG444  | <u>CTAATACGACTCACT</u> ATAGGGAGAGTAGATGCTCATTCCATCTCTTATG  |              | <i>glmZ</i> (+1 to +25)                     |
| BG445  | AAAAAAACGCCTGCTCTTATTACGGAGC   |              | <i>glmZ</i> (+207 to +179)                  |
| BG446  | <u>CTAATACGACTCACT</u> ATAGGGAGAGTAGATGCTCATTCCCGACTTATGTC   |              | <i>glmY</i> (+1 to +25)                     |
| BG447  | AACAAGCCGGGAATTACCCGGC   |              | <i>glmY</i> (+184 to +161)                  |
| BG557  | <u>GGCTCTAGAGAGGGGAAGTT</u> CAGATACAAC   | <u>XbaI</u>  | <i>glmY</i> (+202 to +182)                  |
| BG602  | GAGATCCAGTTCGATGTAACCC   |              | pBAD30 (+2027 to +2006)                     |
| BG603  | ACCCGGCTTTGTTATGGAATAAGGCGG <b>GTGCTTCACTCAACGTTGTG</b>  |              | <i>glmZ</i> (+146 to +126)                  |
| BG604  | GCCTTATCCATAACAAAGCCG  |              | <i>glmY</i> (+144 to +165)                  |
| BG605  | AAAACAGGTCTGTATGACAACAAGTGG <b>GTGCCTAACTCGACGTTTC</b>   |              | <i>glmY</i> (+143 to +123)                  |
| BG606  | CACTTGTGTCATACAGACCTG  |              | <i>glmZ</i> (+146 to +168)                  |
| BG1100 | GCGGGACGTCAGTGGCTCATTACCGACTTATG   | <u>AatII</u> | <i>glmY</i> (+1 to +23)                     |
| BG1101 | GGCGAATTCGAGGGGAAGTTCAGATACAACAAAGC  | <u>EcoRI</u> | <i>glmY</i> (+202 to +177)                  |
| BG1102 | GCGGGACGTCGTAGATGCTCATTCCATCTCTTATG  | <u>AatII</u> | <i>glmZ</i> (+1 to +25)                     |
| BG1103 | GGCGAATTCGGGGCCCTTCTGATACATAAAAA   | <u>EcoRI</u> | <i>glmZ</i> (+226 to +203)                  |
| BG1104 | GGCGAATTCGAGGGGAAGTTCAGATACA <b>AAAAAACGCCTGCTCTTAT</b><br><b>TACGGAGCAGGCGTTATGGAATAAGGCGGTGC</b>           | <u>EcoRI</u> | <i>glmY</i> (+202 to +177 and +155 to +139) |
| BG1105 | GGCGAATTCGAGGGGAAGTTCAGATACAACAAAGCCGGGAATTACCC<br>GGCTTTGTT <b>AAAAACAGGTCTGTATGACAACAAGTGGGTGCCTAACTCG</b> | <u>EcoRI</u> | <i>glmY</i> (+202 to +156 and +143 to +131) |

<sup>a</sup>Restriction sites are underlined, dotted lines correspond to the T7 recognition sequence, sequences in bold deviate from respective gene given on the right; <sup>b</sup>Positions are relative to the first nucleotide of the respective gene.

## RESULTS

**Construction of a library of *GlmY-GlmZ* hybrid sRNAs to investigate molecular determinants for Hfq binding and RNase E processing**

Conserved small RNAs *GlmY* and *GlmZ* are highly homologous in sequence and structure, as they share 63% sequence identity and fold into similar secondary structures (Fig. 5.1 A; Reichenbach *et al.*, 2008). However, whereas *GlmZ* is a base-pairing small RNA that requires Hfq, *GlmY* is Hfq-independent and acts through binding of protein RapZ. Further, with the aid of RapZ, sRNA *GlmZ* is processed by RNase E. *GlmY*, however, is not degraded in this manner (Göpel *et al.*, 2013). To narrow down which regions within *GlmZ* are involved in Hfq-binding and recognition by RNase E, we constructed a library of sRNA hybrids by swapping distinct sequences between both sRNAs (Fig. 5.1 B).



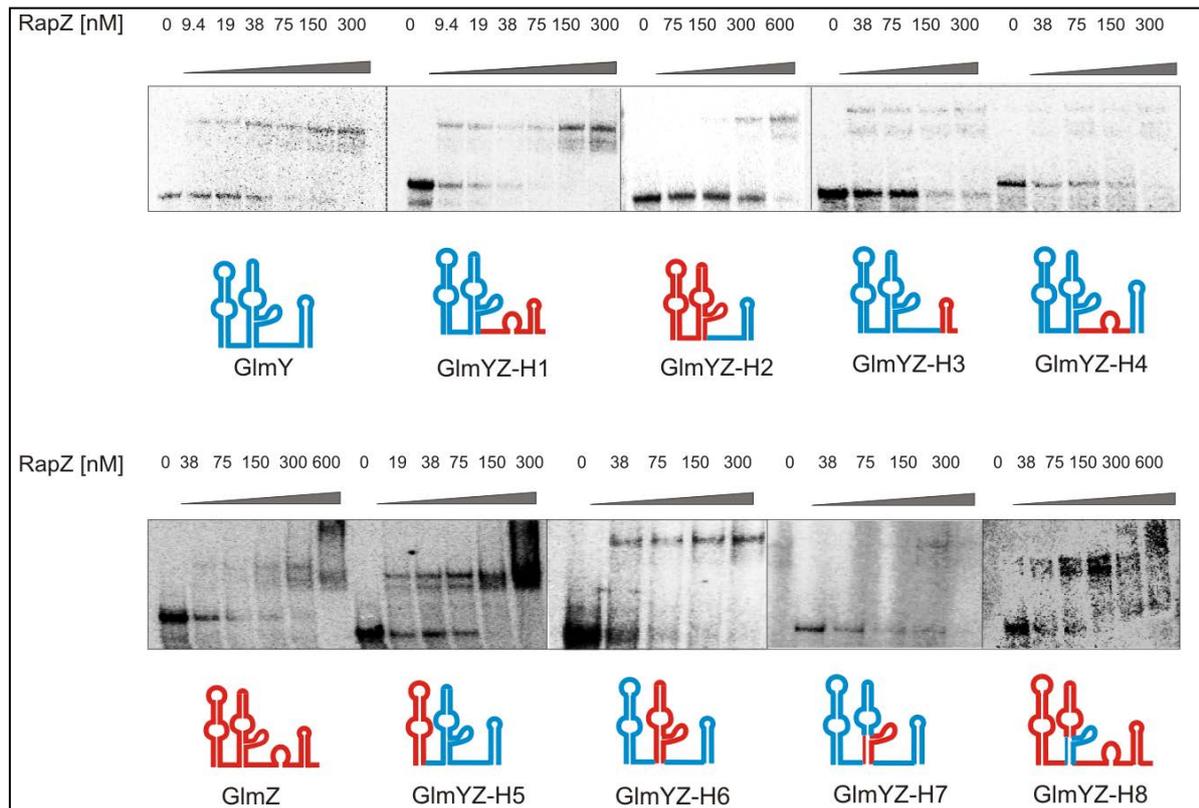
**Figure 5.1: Schematic view of the library of *GlmY-GlmZ* sRNA hybrids constructed to investigate molecular requirements for Hfq-binding and RNase E dependent processing.** **A.** secondary structures of *GlmY* and *GlmZ* sRNAs. Residues that are fully conserved are highlighted in blue (Reichenbach *et al.*, 2008; Göpel *et al.*, 2013), residues involved in base-pairing with the *gls* 5'UTR are marked in red. Scissors indicate the processing sites in *GlmY* and *GlmZ*. **B.** schematic view of the *GlmY-GlmZ* hybrid library encoded in plasmids pYG83 (*GlmY*), pYG84 (*GlmZ*), pYG86 (*GlmYZ-H1*), pYG85 (*GlmYZ-H2*), pYG87 (*GlmYZ-H3*), pYG88 (*GlmYZ-H4*), pYG103 (*GlmYZ-H5*), pYG104 (*GlmYZ-H6*), pYG105 (*GlmYZ-H7*) and pYG106 (*GlmYZ-H8*). Sequences and features corresponding to *GlmY* are shown in blue, features corresponding to sequences in *GlmZ* are drawn in red.

As GlmY and GlmZ share similar 5' stem loop structures, we started by swapping the entire 3' ends since they diverge greatly considering the overall similarity of both sRNAs. In case of GlmZ, the 3' end encompasses a Rho independent terminator with an accessible poly-U-tail as well the *glmS* base-pairing site. The base-pairing nucleotides themselves as well as their surrounding residues are enriched in AU-residues, which is a typical feature of Hfq binding and RNase E cleavage sites (Kim *et al.*, 2004; Lin-Chao *et al.*, 1994; Vogel and Luisi, 2011).

By a PCR-based approach, we fused the 3' end of GlmZ to the 5' end of GlmY exactly after the second stem loop and vice versa thus obtaining hybrids GlmYZ-H1 and GlmYZ-H2. Following initial analysis of these two hybrids we performed other substitutions transplanting smaller regions of the GlmZ 3' end. Replacement of the Rho-independent terminator of GlmY with the terminator within GlmZ yielded hybrid GlmYZ-H3. Hybrid GlmYZ-H4 was obtained by substituting the sequence between the second stem loop and the terminator of GlmY with the corresponding sequence of GlmZ. This hybrid, as well as hybrid GlmYZ-H1, encompasses the base-pairing site present in GlmZ for activation of *glmS*. In hybrid GlmYZ-H5 the first stem loop within GlmY is replaced by the first stem loop of GlmZ. Hybrid GlmYZ-H6 carries the second stem loop of GlmZ instead of that of GlmY. Last, we replaced only the lateral bulges of the second stem loop and the lower part of the stem in GlmY with the corresponding sequence in GlmZ (GlmYZ-H7) or vice versa (GlmYZ-H8, compare Fig. 5.1 B for hybrid structures).

### ***RapZ specifically binds the GlmYZ hybrids in vitro***

In a previous study, we could show that purified RapZ *in vitro* binds to radio labeled transcribed GlmY and GlmZ sRNAs with high affinity but does not specifically bind non-cognate sRNAs (Göpel *et al.*, 2013). To exclude, that swapping of GlmY and GlmZ specific modules with each other should result in drastic structural changes within the hybrids, we tested high affinity binding to RapZ. We therefore incubated the *in vitro* transcribed and internally labeled hybrids with increasing concentrations of purified RapZ protein and performed electrophoretic mobility shift assays. We then analyzed binding efficiencies of RapZ for the various hybrid sRNAs in comparison to GlmY and GlmZ sRNAs (Fig. 5.2).



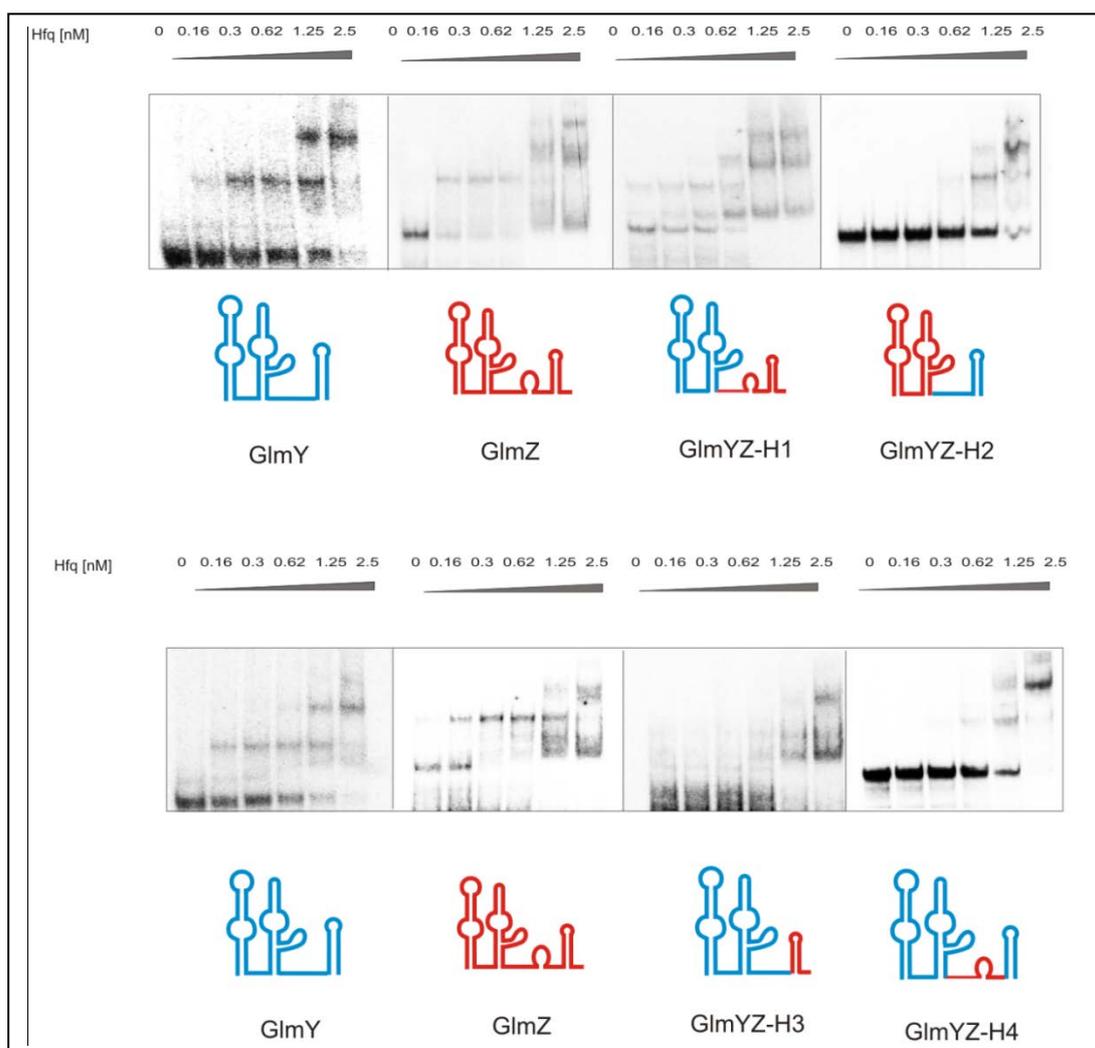
**Figure 5.2: The sRNA hybrids can be bound by RapZ with similar affinities as GlmY and GlmZ.** EMSA using  $\alpha$ - $^{32}$ P-UTP labeled sRNA hybrids with increasing concentrations of purified Strep-RapZ. The final protein concentrations are as indicated; the schematic view represents the respective sRNA or hybrid used in the assay. Solid lines indicate the combination of independent assays on separate gels; dotted lines indicate EMSAs from the same gel with changed exposure settings.

With increasing concentrations of Strep-RapZ all tested hybrids were efficiently bound. As previously shown, RapZ seems to possess a slightly higher affinity for GlmY as for GlmZ (Göpel *et al.*, 2013). Therefore, we repeated the gel shift assays with GlmY and GlmZ sRNAs for direct comparison. The apparent  $K_d$  for binding of GlmY was determined in the range of  $\sim 38$  nM and the  $K_d$  for binding of GlmZ was approximately 75 nM. RapZ exhibited similarly high binding affinities towards all tested hybrids, with the exception of the GlmYZ-H2 hybrid. Note that binding affinities varied between independent experiments and the GlmZ sRNA used together with GlmYZ-H2 also showed a delay in binding with an apparent  $K_d$  of  $\sim 300$  nM (Khan, Göpel and Görke, unpublished). This difference can most likely be attributed to variations in the activity of different RapZ preparations. Therefore, we judged that swapping of domains within the sRNAs did not significantly alter the overall structure of the hybrids and proceeded to analyze the influence of the various sRNA domains in respect to binding by Hfq and recognition by RNase E.

### ***The 3' end of GlmZ, but not of GlmY, provides a high affinity binding site for Hfq***

As RNA chaperon Hfq binds to a multitude of sRNAs and mRNAs within the cell, the question of specificity determinants is of great importance to understand how Hfq is able to discriminate

between substrates and non-substrates. To investigate, which modules of a small RNA are absolutely required for binding by Hfq, we used the library of GlmYZ hybrids. Modules of sRNA GlmZ, which depends on Hfq for functionality and stability, were swapped with the corresponding sequences of Hfq-independent sRNA GlmY thus creating GlmYZ hybrid sRNAs. We then analyzed the binding affinity of purified Hfq towards the aforementioned hybrids *in vitro* by gel retardation experiments (Fig. 5.3).



**Figure 5.3: Dissection of the molecular requirements for binding by Hfq points to the 3' end of GlmZ as determinant for Hfq binding.** EMSA using  $\alpha$ - $^{32}$ P-UTP labeled sRNA hybrids with increasing concentrations of purified Hfq. The final protein concentrations are as indicated above; the schematic view represents the respective sRNA or hybrid used in the assay. Solid lines denote the combination of independent assays on separate gels.

Small RNA GlmZ was specifically bound by Hfq at very low protein concentrations with an apparent  $K_d$  ranging between  $\sim 0.16$  and  $0.3$  nM. In contrast, sRNA GlmY exhibited an approximately ten-fold higher apparent  $K_d$  of  $\sim 1.23$  to  $2.5$  nM. This difference in binding affinities is in agreement with previously published results indicating a ten to 15-fold lower affinity of Hfq for GlmY than for GlmZ (Göpel *et al.*, 2013). However, it is worth noting that the Hfq concentration required for complex formation in the experiments conducted in this study are generally much lower than in the

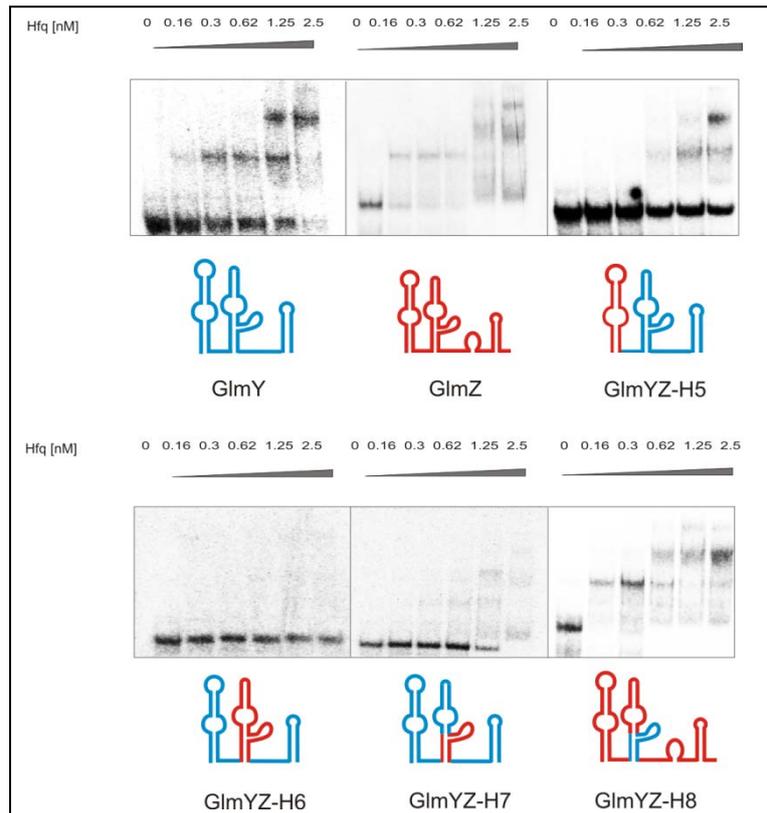
previously published study. These differences might be attributed to variations between the individual Hfq preparations used in the two studies.

Interestingly, hybrid GlmYZ-H1 encompassing the 5' end of GlmY fused to the 3' end of GlmZ was efficiently bound by Hfq with an apparent  $K_d$  of  $\sim 0.3$  nM and thus exhibited highly increased binding properties as compared to GlmY (Fig. 5.3). Hence, the ability to bind to Hfq resides in the 3' end of sRNA GlmZ. Corroborating the importance of this finding, swapping of the 3' end of GlmZ with that of GlmY abrogated Hfq binding of the resulting hybrid GlmYZ-H2. This sRNA chimera was only bound by Hfq at the highest protein concentration, thus exhibiting an apparent  $K_d$  of 2.5 nM similar to GlmY (Fig. 5.3).

To narrow down the molecular determinants responsible for Hfq binding within the 3' end of GlmZ we further dissected this region by analyzing the effects of transplanting the Rho-independent terminator or the base-pairing region alone (hybrids GlmYZ-H3 and GlmYZ-H4, respectively). Hfq exhibited a slightly higher binding affinity towards these two hybrids as compared to GlmY with app.  $K_d$  values between 0.62 and 1.25 nM as evaluated from several independent assays. However, neither of the two modules alone is sufficient to confer a high binding affinity, as apparent by comparison with hybrid GlmYZ-H1 that exhibited a two to four-fold higher binding capacity (Fig. 5.3).

#### ***The 5' stem loop structures of GlmY and GlmZ are not required for efficient binding by Hfq***

To investigate whether exchanges of domains in the 5' end of the sRNAs also affect binding by Hfq, we again performed gel retardation experiments. This time we used the GlmYZ hybrids with substitutions in the 5' stem loop structures: hybrids GlmYZ-H5 and H6 that encompass the GlmY sequence with either stem loop one or stem loop two derived from GlmZ, as well as hybrids GlmYZ-H7 and H8 carrying swapped lateral bulges (Fig. 5.4). As Hfq is thought to bind to single stranded, A/U-rich regions and extended poly-U-sequences of Rho-independent terminators (Sauer and Weichenrieder, 2011; Vogel and Luisi, 2011), we hypothesized that the elaborate 5' stem loop structures of either GlmY or GlmZ might not considerably affect the Hfq binding capabilities of the sRNA hybrids.



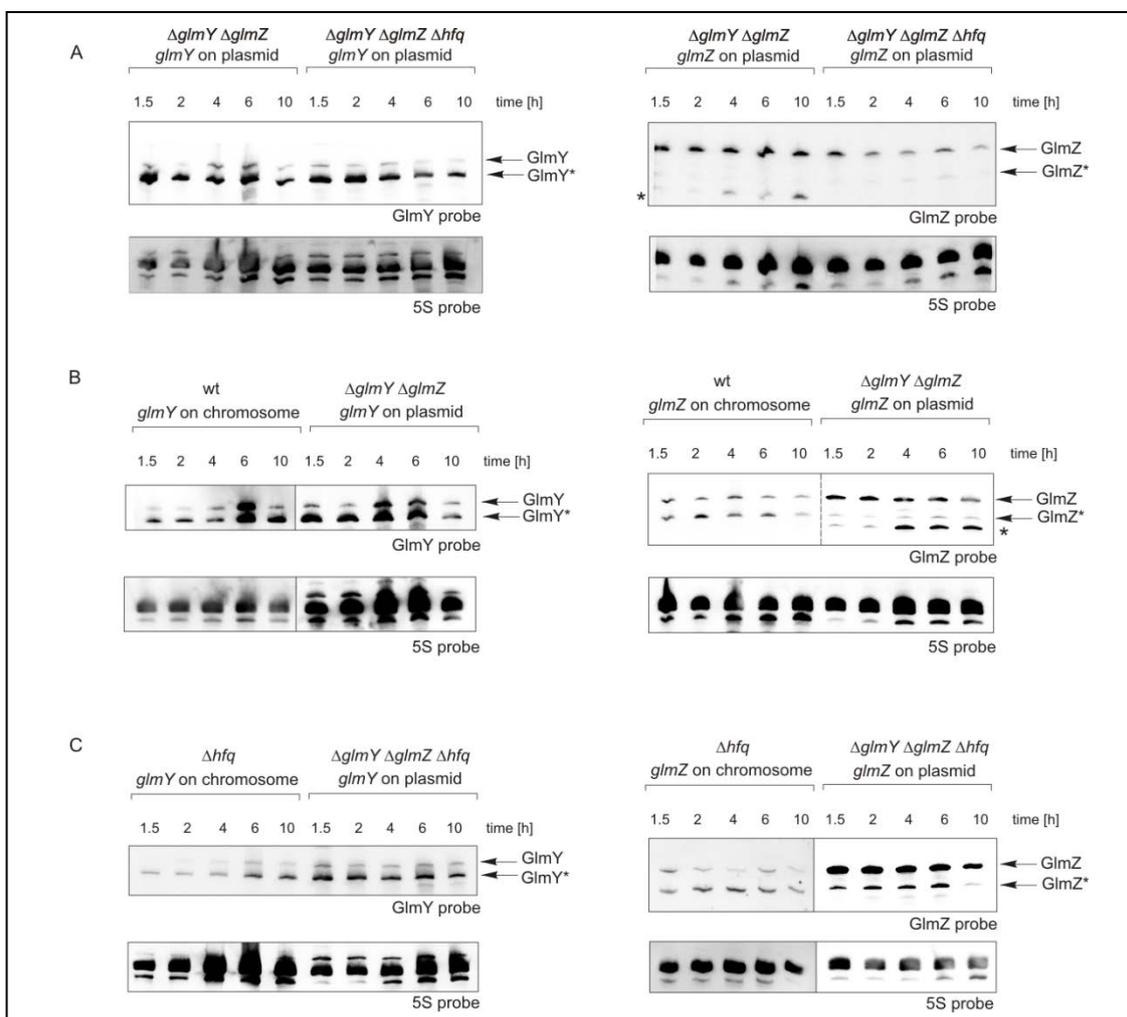
**Figure 5.4: The 5' stem loop structures of GmY and GmZ do not contribute to the capability to bind Hfq.** EMSA using  $\alpha$ - $^{32}$ P-UTP labeled sRNA hybrids with increasing concentrations of purified Hfq. The final protein concentrations are as indicated above; the schematic view represents the respective sRNA or hybrid used in the assay. Solid lines denote the combination of independent assays on separate gels.

The substitution of one of the first or second stem loop in GmY with the corresponding stem loop of GmZ did not improve the ability of the resulting hybrids GmYZ-H5 and GmYZ-H6 to bind Hfq *in vitro* (Fig. 5.4). Both hybrids exhibited similarly low app.  $K_d$ s of  $\geq 2.5$  nM as observed for GmY. Further, the hybrids with swapped lateral bulges possessed nearly identical Hfq binding capabilities as the sRNA of which they are derived, i.e. GmYZ-H7 displayed an app.  $K_d$  of  $\sim 1.25$  nM and GmYZ-H8 exhibited an app.  $K_d$  of  $\sim 0.16$  nM. Thus, the 5' stem loop structures are not required for recognition and binding by Hfq.

In conclusion, the Hfq binding affinities of the various sRNAs as determined by EMSA can be arranged in the following order from highest to lowest: GmZ  $\geq$  GmYZ-H8  $>$  GmYZ-H1  $>$  GmYZ-H3  $>$  GmYZ-H4  $>$  GmYZ-H2  $\geq$  GmYZ-H7  $\geq$  GmYZ-H5  $\geq$  GmYZ-H6  $\geq$  GmY. Taken together, our data suggest that two different features, the extended poly-(U)-tail and the single-stranded base-pairing region within GmZ confer the ability to specifically bind Hfq, whereas the 5' stem loops do not seem to contribute to Hfq binding.

**Stability of plasmid-encoded GlmZ, but not of GlmY, requires Hfq in vivo**

A previous study demonstrated that endogenous GlmZ is destabilized in a  $\Delta hfq$  deletion mutant and thus required Hfq for stability *in vivo*. GlmY on the other hand, was not affected by the *hfq* deletion and was therefore thought to be independent of Hfq *in vivo* (Göpel *et al.*, 2013). In order to investigate Hfq-dependencies of the GlmYZ-hybrids *in vivo*, we first tested the behavior of plasmid encoded GlmY and GlmZ upon deletion of *hfq* in comparison to the endogenous sRNAs. Therefore, we performed northern blot experiments with endogenous and plasmid encoded GlmY and GlmZ in absence and presence of Hfq (Fig. 5.5). The plasmids encoding for GlmY and GlmZ were introduced in wild type and *hfq* deletion strains lacking the endogenous sRNA alleles.



**Figure 5.5: GlmZ is destabilized in cells lacking *hfq*.** **A.** Northern blot experiment addressing the stabilities of plasmid encoded GlmY (pYG83) and GlmZ (pYG84) in absence and presence of Hfq in strains lacking endogenous *glmY* and *glmZ* (Z106 denoted  $\Delta glmY \Delta glmZ$  and Z865 denoted  $\Delta glmY \Delta glmZ \Delta hfq$ , respectively). **B.** Northern blot comparing chromosomal and plasmid encoded GlmY and GlmZ in presence of Hfq. RNA samples were obtained from strains R1279 (wild type) and Z106/pYG83 ( $\Delta glmY \Delta glmZ$ , *glmY* on plasmid) for GlmY or Z106/pYG84 ( $\Delta glmY \Delta glmZ$ , *glmZ* on plasmid) for GlmZ, respectively. **C.** Northern experiment comparing chromosomal and plasmid encoded GlmY and GlmZ in absence of Hfq. RNA samples were obtained from strains Z664 ( $\Delta hfq$ ) and Z865/pYG83 ( $\Delta glmY \Delta glmZ \Delta hfq$ , *glmY* on plasmid) for GlmY or Z865/pYG84 ( $\Delta glmY \Delta glmZ \Delta hfq$ , *glmY* on plasmid) for GlmZ, respectively. Cells were grown in LB over the course of 10 hours, total RNA was isolated from samples harvested at the indicated times and 2.5  $\mu$ g RNA was analyzed by northern blot using the indicated RNA probes. Re-probing using a probe directed against 5S rRNA served as loading control (lower panels). Asterisks denote unspecific processing products, solid lines indicate the combination of independent assays; dotted lines indicate blots with changed exposure settings.

Whereas the amount of plasmid encoded GlmY was not affected by a deletion of the *hfq* gene, the amounts of full-length GlmZ were drastically reduced in the  $\Delta hfq$  mutant (Fig. 5.5 A). As the amounts of GlmY remained constant under all tested conditions, we excluded that Hfq influences transcription from the artificial  $\lambda P_L/P_{lac}$  hybrid promoter of the plasmids encoding for GlmY, GlmZ and the various sRNA hybrids. Thus, in agreement with previous results, GlmZ requires Hfq for its stability *in vivo*, even when over-expressed from a plasmid (Fig. 5.5. A; Göpel *et al.*, 2013). In the  $\Delta glmY \Delta glmZ (hfq^+)$  strain, expression of plasmid encoded GlmZ led to the appearance of an uncharacterized, smaller variant of GlmZ (marked with an \*). This variant possessed a lower molecular weight as the described processed form of GlmZ (Fig. 5.5 B; right) and resembled a previously observed variant that appeared upon inactivation of RNase E (Göpel *et al.*, 2013).

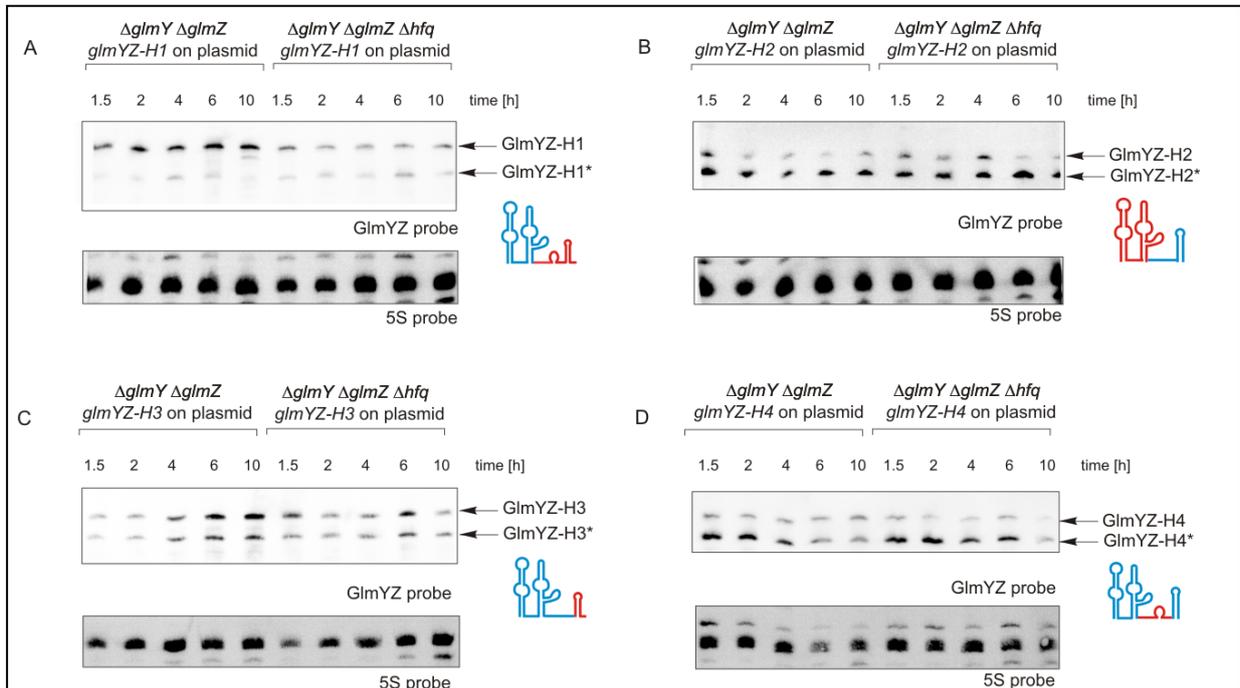
In order to compare the behavior of plasmid encoded GlmY and GlmZ to the endogenous sRNAs, northern experiments were repeated and both species were analyzed in parallel (Fig. 5.5. B and C). Endogenous GlmY accumulated upon transition to stationary phase due to activation of the  $\sigma^{54}$ -dependent *glmY* promoter by the GlrK/GlrR two-component system (Fig. 5.5 B, left; Reichenbach *et al.*, 2009). In contrast, the amount of plasmid encoded GlmY decreased upon entry of stationary phase due to cessation of  $\sigma^{70}$ -dependent transcription that might decrease the activity of the  $\lambda P_L/P_{lac}$  promoter of the plasmid. These effects were also observed in the respective *hfq* deletion strains, again indicating that GlmY does not require Hfq for stability (Fig. 5.5 C, left). In presence of Hfq, endogenous and plasmid encoded GlmZ amounts were decreasing only slightly in stationary phase (Fig. 5.5 B, right). This observation can be attributed to the overall decrease of  $\sigma^{70}$ -dependent transcription during stationary growth since GlmZ was shown to be expressed from a  $\sigma^{70}$ -dependent promoter (Göpel *et al.*, 2011). Again, the smaller GlmZ species was visible in wild type cells expressing GlmZ from a plasmid and its amount increased in stationary phase (Fig. 5.5 B, right). Deletion of the *hfq* gene led to destabilization of full-length GlmZ, while amounts of the processed variant increased (Fig. 5.5 C, right). Interestingly, the smaller GlmZ variant could not be detected in the *hfq*-deficient strain that expressed GlmZ from a plasmid.

In conclusion, when expressed from a plasmid the sRNAs behave similar to chromosomally encoded GlmY and GlmZ. However, as their amounts are increased due to over-expression, the processed form of GlmZ was more prominent in the *hfq*-deficient strain. The requirement for Hfq for stability of GlmZ was, thus, manifested in a drastic reduction of the amount of unprocessed GlmZ.

#### ***The need for protection by Hfq in vivo can be transferred to GlmY by swapping 3' ends with GlmZ***

Next, we addressed the question whether the GlmYZ hybrids require Hfq for stability *in vivo*. To investigate the stability of the GlmYZ hybrids in absence and presence of Hfq we again performed

northern blot analysis (Fig. 5.6). To this end, we introduced plasmids expressing the hybrid sRNAs in strains lacking the endogenous *glmY* and *glmZ* genes in presence and absence of Hfq.



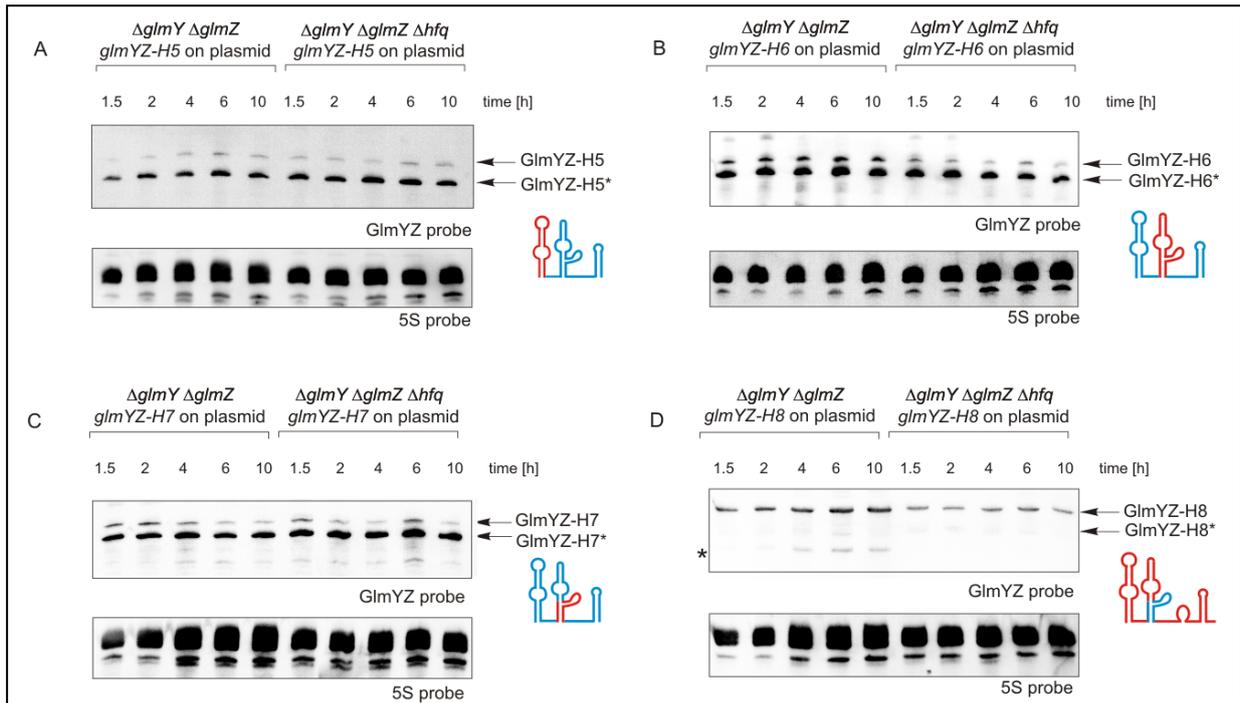
**Figure 5.6: The requirement for Hfq for stability *in vivo* resides within the 3' end of GlmZ.** **A.** Northern blot experiment addressing the stability of hybrid GlmYZ-H1 in absence and presence of Hfq analyzing transformants of Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ) with pYG86. **B.** Northern blot analyzing the stability of hybrid GlmYZ-H2 (pYG85) in Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ). **C.** Hybrid GlmYZ-H3 (pYG87) was examined by northern analysis in respect to its stability in presence (Z106,  $\Delta glmY \Delta glmZ$ ) and absence (Z865,  $\Delta glmY \Delta glmZ \Delta hfq$ ) of Hfq. **D.** Northern experiment investigating the requirement for Hfq for stability of hybrid GlmYZ-H4 (pYG88) in Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ). Cells were grown in LB over the course of 10 hours, total RNA was isolated from samples harvested at the indicated times and 2.5  $\mu$ g RNA was analyzed by northern blot using a mixture of GlmY and GlmZ RNA probes. Re-probing using a probe directed against 5S rRNA served as loading control (lower panels).

Similar to GlmZ, amounts of the full-length form of hybrid GlmYZ-H1 were reduced in *hfq*-deficient cells (Fig. 5.6 A, compare 5.5 A, right). Thus, GlmYZ-H1 which encompasses the 5' stem loops of GlmY fused to the 3' end of GlmZ was destabilized in an *hfq* deletion mutant as compared to the respective wild type strain. In contrast, the reverse hybrid GlmYZ-H2 did not rely on Hfq for stability *in vivo*, since the amounts of this hybrid remained constant in absence and presence of Hfq (Fig. 5.6 B). Hence, the 3' end of GlmZ seems to be important for stabilization of this hybrid by Hfq *in vivo*. We next addressed the question whether parts of the 3' end, i.e. the Rho-independent terminator with the extended poly-(U)-tail and/or the single-stranded base-pairing region, are sufficient to bestow the need for Hfq to the hybrids. To this end, we tested the GlmYZ-hybrids H3 and H4 that correspond to GlmY fused to the terminator present in GlmZ or GlmY carrying the single-stranded base-pairing region of GlmZ, respectively (Fig. 5.6 C and D). These two hybrids showed only slight but significant reduction of the RNA amounts after 10 hours in the  $\Delta hfq$  strain as compared to the wild type. Taken together, these observations corroborate the *in vitro* results that individual parts of the 3' end of

GlmZ are not sufficient to confer the ability to bind Hfq. Rather, the entire 3' sequence of GlmZ is required for high affinity binding of Hfq *in vitro* and protection by Hfq *in vivo*.

**The 5' stem loop structures of GlmY and GlmZ are not required for protection by Hfq *in vivo***

To further validate our findings that binding of Hfq *in vitro* is not dependent on the 5' stem loops of GlmY and GlmZ we further analyzed hybrids GlmYZ-H5 to H8 that carry substitutions in one of the stem loops (Fig. 5.7).



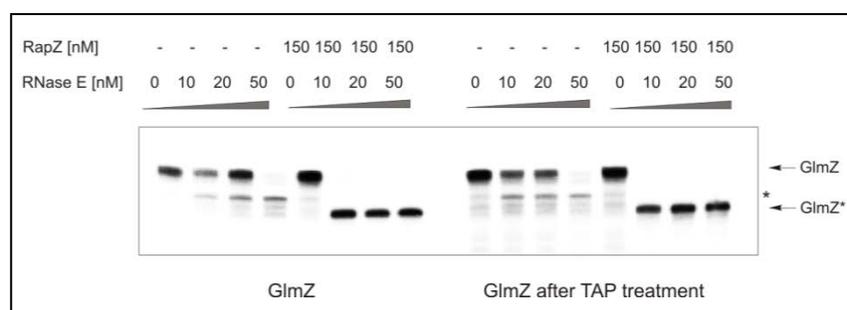
**Figure 5.7: The 5' stem loop structures are not required for Hfq-mediated protection of the sRNA hybrids *in vivo*.** A. Northern blot experiment investigating the stability of hybrid GlmYZ-H5 in absence and presence of Hfq analyzing transformants of Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ) with pYG103. B. Northern blot analyzing the stability of hybrid GlmYZ-H6 (pYG104) in Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ). C. Hybrid GlmYZ-H7 (pYG105) was assessed by northern analysis regarding its stability in presence (Z106,  $\Delta glmY \Delta glmZ$ ) and absence (Z865,  $\Delta glmY \Delta glmZ \Delta hfq$ ) of Hfq. D. Northern experiment assessing the requirement for Hfq for stability of hybrid GlmYZ-H8 (pYG106) in Z106 ( $\Delta glmY \Delta glmZ$ ) and Z865 ( $\Delta glmY \Delta glmZ \Delta hfq$ ). Cells were grown in LB over the course of 10 hours, total RNA was isolated from samples harvested after the indicated time intervals and 2.5  $\mu$ g RNA was analyzed by northern blot using a mixture of GlmY and GlmZ RNA probes. Re-probing using a probe directed against 5S rRNA served as loading control (lower panels).

Analogous to GlmY, the amounts of hybrids GlmYZ-H5, GlmYZ-H6 and GlmYZ-H7 remained constant under all conditions tested and overall amounts were unaltered in an *hfq* deletion mutant (Fig. 5.7 A-C, compare Fig. 5.5 A, left). These hybrids encompass the 3' end of GlmY but carry exchanges within the two 5' stem loop structures: GlmYZ-H5 carries the first stem loop of GlmZ, in hybrid GlmYZ-H6 the second stem loop is derived from GlmZ and hybrid GlmYZ-H7 contains the lateral bulge and lower part of the second stem loop of GlmZ. Last, we investigated the influence of the lateral bulge within the second stem loop. This region was suggested to play a major role in binding of RapZ and RNase E dependent processing (Göpel *et al.*, 2013). However, no major difference was observed, when this region in GlmZ was swapped with the corresponding sequence in GlmY (GlmYZ-H8). Like GlmZ, this

hybrid exhibited reduced stability in *hfq* deficient cells (Fig. 5.7 D, compare Fig. 5.5 A, right). In sum, these experiments are in agreement with the *in vitro* studies suggesting that the 5' structures of the sRNAs are not important for Hfq binding and protection from degradation *in vivo*.

### **Processing of GlmZ by RNase E is independent of a 5' mono-phosphate**

Previous work demonstrated that processing of GlmZ by RNase E strictly requires adaptor protein RapZ and that RapZ is essential for specific and efficient cleavage of sRNA GlmZ *in vitro* (Göpel *et al.*, 2013). However, the mechanism by which RapZ acts to induce cleavage of GlmZ by RNase E is yet to be clarified. As RNase E was shown to display enhanced cleavage efficiencies towards some 5' mono-phosphorylated substrates (Deana *et al.*, 2008; Mackie, 2013a, b), one possibility is that RapZ acts as a pyrophosphatase specific for sRNA GlmZ. Since  $^{32}\text{P}$ -labeled GlmZ is derived from internal labeling during *in vitro* transcription, all GlmZ molecules tested in the *in vitro* cleavage approach carry a 5' tri-phosphate. To assess whether the presence of a 5' mono-phosphate would enhance cleavage efficiency or even render RapZ dispensable, we performed *in vitro* cleavage assays comparing cleavage efficiencies for 5' tri-phosphorylated and for 5' mono-phosphorylated GlmZ in absence and presence of RapZ (Fig. 5.8).



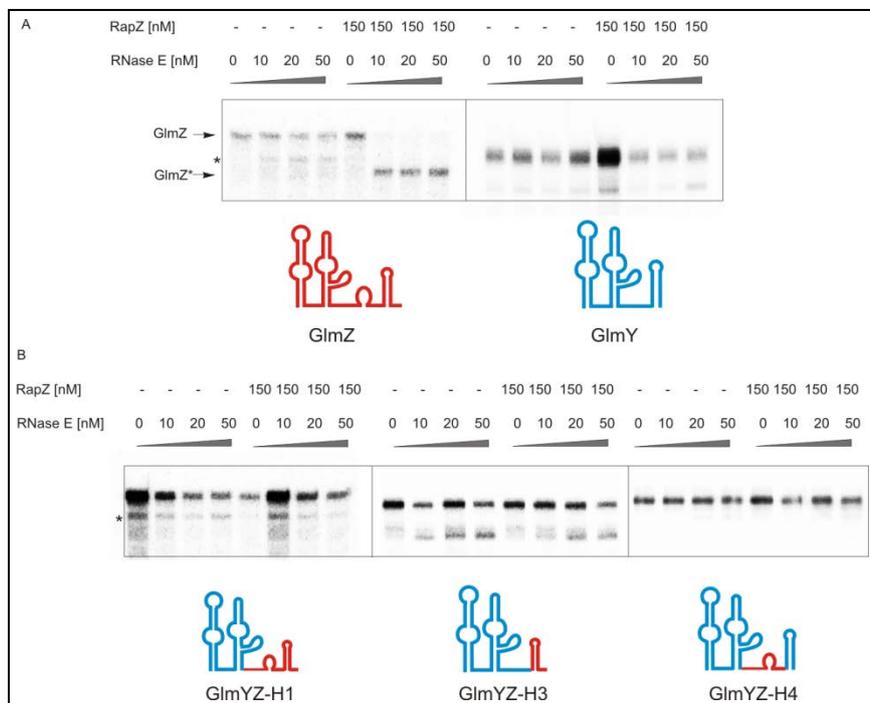
**Figure 5.8: Generation of a 5' mono-phosphate prior to cleavage by RNase E does not enhance cleavage efficiency.** RNase E *in vitro* cleavage assay assessing the stabilities of  $\alpha$ - $^{32}\text{P}$ -UTP-labeled GlmZ with a 5' tri-phosphate and GlmZ with a 5' mono-phosphate upon incubation with increasing amounts of RNase E NTD (amino acids 1-529) in absence and presence of 150 nM RapZ. To generate the 5' monophosphate,  $\alpha$ - $^{32}\text{P}$ -UTP-labeled GlmZ was incubated for 3 hours with 5 u TAP at 37°C according to manufacturer's instructions. Following TAP treatment, the RNA was isolated by phenol:chloroform:isoamyl-alcohol (25:24:1) extraction and subsequently precipitated with EtOH:NaAc (30:1). The RNA was dissolved in RNase free water and RNase E cleavage assays were performed as described previously (Göpel *et al.*, 2013). Asterisks denote unspecific cleavage products.

To generate 5' mono-phosphorylated GlmZ, radio-labeled *in vitro* transcribed GlmZ was incubated with TAP (Efstratiadis *et al.*, 1977) prior to the RNase E *in vitro* cleavage reaction. Similar to 5' tri-phosphorylated GlmZ (Fig. 5.8, left), mono-phosphorylated GlmZ was only cleaved with high efficiency when RapZ was present in the reaction (Fig. 5.8, right). Note that incubation of GlmZ with RNase E already resulted in unspecific cleavage and that the amounts of full length GlmZ were strongly reduced upon incubation with the highest amounts of RNase E. This might be attributed to the use of different RNase E preparations that vary in activity. Nevertheless, the specific cleavage product was only obtained in reactions that also contained RapZ. As the processing efficiency was

unaffected by the 5' phosphorylation state of GlmZ and RapZ was still required for specific processing, it may be excluded that RapZ acts as a pyrophosphate on GlmZ.

***In contrast to Hfq-binding, processing of GlmZ by RNase E is not dependent on features within its 3' end***

Whereas GlmZ was shown to be subject to processing by RapZ/RNase E *in vivo*, GlmY is not degraded by this pathway (Göpel *et al.*, 2013). In order to investigate the molecular discriminates that are essential for recognition and processing by RNase E, we performed *in vitro* cleavage assays with *in vitro* transcribed, radio-labeled sRNAs and the purified N-terminal catalytical domain of RNase E in absence and presence of RapZ (Fig. 5.9).



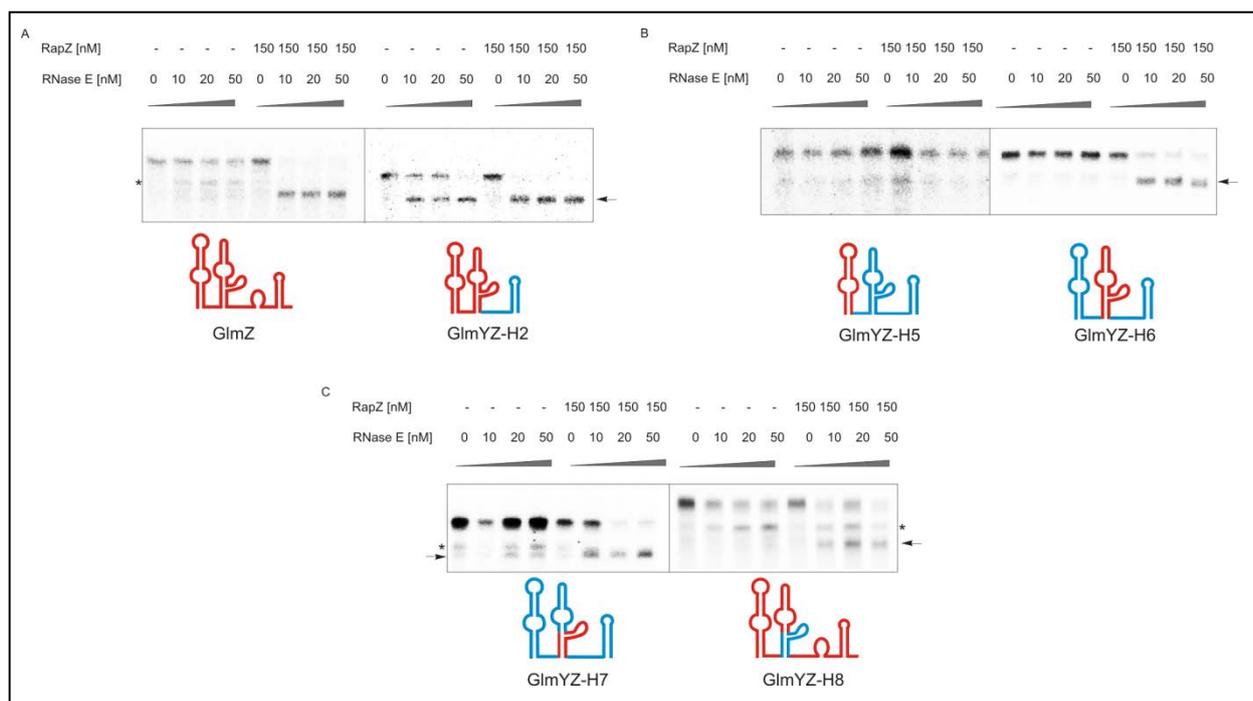
**Figure 5.9: Transplantation of the 3' end of GlmZ does not confer efficient and specific processing of GlmY by RNase E and RapZ.** **A.** RNase E *in vitro* cleavage assay assessing the stabilities of GlmZ and GlmY upon incubation with increasing amounts of RNase E NTD (amino acids 1-529) in absence and presence of 150 nM RapZ. **B.** RNase E *in vitro* cleavage experiment addressing the ability of RNase E NTD to cleave hybrids GlmYZ-H1, GlmYZ-H3 and GlmYZ-H4 in absence and presence of 150 nM RapZ. Small RNAs were *in vitro* transcribed and simultaneously labeled with  $\alpha$ - $^{32}$ P-UTP using DNA-templates carrying a 5' T7-recognition sequence. DNA fragments were derived from PCR using the following template plasmids: pBGG149 for GlmY, pBGG84 for GlmZ, pYG86 for GlmYZ-H1, pYG87 for GlmYZ-H3 and pYG88 for GlmYZ-H4. Asterisks denote unspecific processing products; solid lines indicate the combination of independent assays.

As previously demonstrated, upon addition of RapZ, GlmZ is efficiently processed by RNase E *in vitro* (Göpel *et al.*, 2013; Figs. 5.8, 5.9 A). In contrast, GlmY remained stable even at highest RNase E concentrations and in presence of RapZ (Fig. 5.9 A). Thus, GlmY is not a substrate for RNase E as has been indicated from *in vivo* data showing that GlmY is not significantly influenced by an inactivation of RNase E (Göpel *et al.*, 2013). Next, we tested whether GlmYZ hybrids encompassing the GlmZ 3' end or parts of this domain, i.e. the terminator and the base-pairing region, could be processed by RNase E and RapZ *in vitro*. Surprisingly, even in presence of RapZ, RNase E was unable to cleave

hybrids GlmYZ-H1, GlmYZ-H3 and GlmYZ-H4 (Fig. 5.9 B). Interestingly, hybrids GlmYZ-H1 and GlmYZ-H4 are completely stable even though they possess the RNase E cleavage site at position 151 (Göpel *et al.*, 2013). Only hybrid GlmYZ-H3 showed a smaller variant, the amounts of which increased with increasing RNase E concentrations (Fig. 5.9 B, middle). Nevertheless, this hybrid was not efficiently cleaved as compared to GlmZ (Fig. 5.9 A) and cleavage occurred independent of RapZ. It is possible that the addition of the GlmZ specific terminator with its extended poly-(U)-tail to this hybrid might provide excess for RNase E even in absence of RapZ. Hence, in contrast to Hfq-binding, the ability for cleavage by RNase E is not conferred by the 3' end of GlmZ.

***The second stem loop of GlmZ, but not of GlmY, provides a recognition site for efficient processing by RNase E***

To test, whether features within the 5' end of the sRNAs influence recognition and cleavage by RNase E *in vitro*, we next examined processing efficiencies of GlmYZ hybrids that contain substitutions within the 5' stem loop structures (Fig. 5.10).



**Figure 5.10: The lateral bulge within the second stem loop of GlmZ is strictly required for efficient and specific processing by RNase E and RapZ.** **A.** RNase E *in vitro* cleavage assay assessing the stabilities of GlmZ and GlmYZ-H2 upon addition of increasing amounts of RNase E NTD (amino acids 1-529) in absence and presence of 150 nM RapZ. **B.** RNase E *in vitro* cleavage experiment addressing the ability of RNase E NTD to cleave hybrids GlmYZ-H5 and GlmYZ-H6 in absence and presence of 150 nM RapZ. **C.** *In vitro* processing assay of GlmYZ-H7 and GlmYZ-H8 using increasing amounts of RNase E NTD in absence and presence of 150 nM RapZ. Small RNAs were *in vitro* transcribed and simultaneously labeled with  $\alpha$ - $^{32}$ P-UTP using DNA-templates carrying a 5' T7-recognition sequence. DNA fragments were derived from PCR using the following template plasmids: pBGG84 for GlmZ, pYG85 for GlmYZ-H2, pYG103 for GlmYZ-H5, pYG104 for GlmYZ-H6, pYG105 for GlmYZ-H7 and pYG106 for GlmYZ-H8. Arrows mark specific processing products, asterisks denote unspecific processing products; solid lines indicate the combination of independent assays.

With increasing concentration of RNase E NTD, hybrid GlmYZ-H2 was efficiently cleaved and even in the absence of RapZ, complete cleavage was obtained at the highest RNase E concentration (Fig. 5.10

A). This was surprising since GlmYZ-H2 encompasses the 5' stem loops of GlmZ fused to the 3' end of GlmY and therefore lacks the GlmZ-specific RNase E cleavage site. However, GlmY also comprises an A/U-rich single stranded region that could potentially be cleaved by RNase E (Fig. 5.1 A). Further, addition of RapZ increased the cleavage efficiency. Similar to wild type GlmZ, complete processing was obtained at the lowest RNase E concentration of 10 nM (Fig. 5.10 A). Thus, features within the 5' stem loop structures of GlmZ seem to confer the ability for cleavage by RNase E.

To narrow down, which structural features within the 5' end of GlmZ are required for recognition and processing by RNase E, we assessed cleavage efficiencies for hybrids GlmYZ-H5 and GlmYZ-H6 (Fig. 5.10 B). Hybrid GlmYZ-H5 corresponds to GlmY with the first stem loop substituted for the corresponding stem loop of GlmZ. Analogous to GlmY, this hybrid was completely stable even in the presence of RapZ and RNase E. In contrast, hybrid GlmYZ-H6, in which the second stem loop of GlmY is replaced with that of GlmZ, is effectively cleaved by RNase E and RapZ (Fig. 5.10 B). The cleavage efficiency for this hybrid strongly resembles the efficiency with which GlmZ is cleaved (Fig. 5.10 A). Hence, the ability to pose as substrate for RNase E resides within the second stem loop of GlmZ.

One striking difference between the second stem loop in GlmY and that of GlmZ is the structure of the lateral bulge (Fig. 5.1). Previously, data suggested that the lateral bulge within the second stem loop of GlmZ plays a critical role for processing by RNase E. Single nucleotide exchanges in this region led to a drastic decrease in RNase E cleavage efficiency of the resulting sRNA mutants (Göpel *et al.*, 2013). Therefore, we transplanted the lateral bulge of GlmZ onto GlmY and tested the resulting hybrid GlmYZ-H7 for *in vitro* cleavage (Fig. 5.10 C). Interestingly, this substitution was sufficient to confer effective cleavage to this hybrid. Approximately 50% cleavage efficiency was obtained at an RNase E concentration of 10 nM in presence of RapZ and complete cleavage was observed at 20 nM RNase E (Fig. 5.10 C). Vice versa, exchange of the lateral bulge of GlmZ for the sequence present in GlmY (hybrid GlmYZ-H8) impaired processing by RNase E but did not abolish cleavage. Further, an unspecific cleavage product appeared even when RapZ was added (Fig. 5.10 C). This shorter variant of GlmZ was previously observed, when GlmZ was incubated with increasing concentrations of highly active RNase E but disappeared upon addition of RapZ. Interestingly, mutants of GlmZ carrying single nucleotide exchanges of G-residues within the lateral bulge were also shown to be strongly impaired in their ability to pose as substrate for RNase E and often exhibited a similar unspecific cleavage product (Göpel *et al.*, 2013).

In sum, the GlmZ-specific lateral bulge is adequate to bestow effective RNase E processing, but other features within the second stem loop are likely to contribute to recognition by RNase E.

**DISCUSSION**

Though highly similar in sequence and structure, homologous sRNAs GlmY and GlmZ differ greatly in their mode of action. Whereas GlmZ acts by base-pairing to activate its target transcript *glmS*, thus allowing synthesis of the enzyme, GlmY acts indirectly by sequestration of RNase E adaptor protein RapZ. Since RapZ is indispensable for degradation of GlmZ, binding of GlmY to RapZ, in turn, stabilizes GlmZ. However, GlmY is not degraded by RapZ and RNase E. It is also worth noting, that whereas RNase E-dependent processing inactivates GlmZ due to removal of the base-pairing nucleotides, the processed form of GlmY is the active form *in vivo* and is sufficient to counteract processing of GlmZ *in vitro* (Göpel *et al.*, 2013). Thus, we employed the GlmYZ sRNAs in order to study discriminants for recognition and cleavage by RNase E *in vitro*. Further, we aimed to expand the knowledge about molecular requirements for efficient Hfq-binding by exploiting the fact that GlmZ strictly depends on RNA-chaperon Hfq for its function and stability, whereas GlmY is an Hfq-independent sRNA (Urban and Vogel, 2007; Göpel *et al.*, 2013).

To assess the questions raised above, we constructed a library of GlmYZ hybrids that comprises various sRNA chimeras with substitutions within the 5' and 3' modules of both sRNAs (Fig. 5.1). To ensure that the swapping of domains between both sRNA species did not alter the overall structure of the hybrids, we first assessed binding to RNase adapter protein RapZ. Previously, RapZ was shown to bind both, GlmY and GlmZ, with high affinity *in vivo* and *in vitro*, but did not specifically interact with unrelated sRNAs (Göpel *et al.*, 2013). Similarly to GlmY and GlmZ, the sRNA hybrids were bound by RapZ with high affinity, thus indicating that the overall structure was not drastically altered by substitution of domains in GlmY with the homologous regions in GlmZ and vice versa (Fig. 5.2). Hence, the library of sRNA hybrids is a suitable tool to investigate determinants of Hfq-binding and identify requirements for cleavage by RNase E *in vitro*.

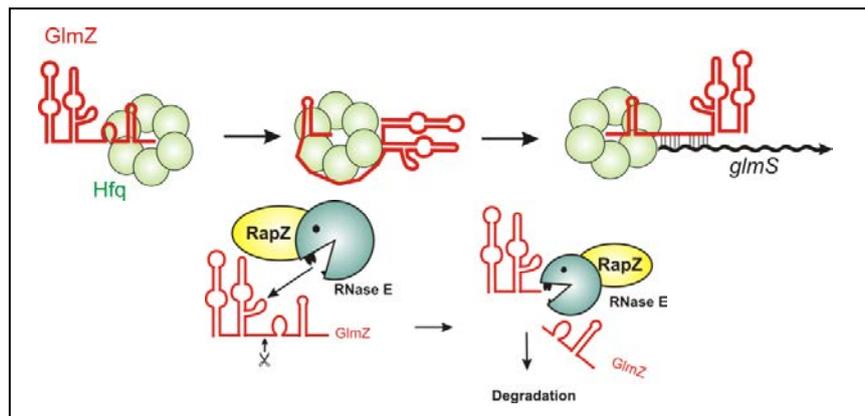
By analyzing the Hfq-binding affinities of the various chimeras of sRNAs GlmY and GlmZ, we could demonstrate that the ability of GlmZ to bind Hfq with high affinity resides within its 3' end. In contrast, the 5' ends of the sRNAs did not contribute to Hfq-binding (Fig. 5.4). These conclusions are supported by the observation that hybrid GlmYZ-H1 that comprises of the 5' stem loops of GlmY fused to the 3' end of GlmZ was bound by Hfq with an affinity that was almost comparable to the affinity with which GlmZ is bound (Fig. 5.3, app.  $K_d$  (GlmYZ-H1)  $\sim 0.3$  nM vs.  $K_d$  (GlmZ)  $\sim 0.16$  nM). However, in case of the inverse hybrid, GlmYZ-H2, Hfq-binding was abrogated (Fig. 5.3). One striking difference between the 3' modules of both sRNAs is the existence of a protruding poly-(U)-tail succeeding the Rho-independent terminator of GlmZ. In contrast, GlmY does not possess a protruding poly-(U)-tail as the U-rich sequence at the 3' end is completely sequestered within the

terminator stem (Fig. 5.1 A). It was previously discovered, that length and accessibility of the poly-(U) stretch within an sRNA is crucial for binding of Hfq (Otaka *et al.*, 2011; Sauer and Weichenrieder, 2011). This initial recognition of the 3' poly-(U)-sequence is thought to allow selectivity for Hfq-dependent sRNAs over other RNA species despite their high degree of structural variation (Sauer and Weichenrieder, 2011). Further, this difference in accessibility of a U-rich sequence as entry site for Hfq might account for the fact, that GlmY is not efficiently bound by Hfq (Fig. 5.3; Göpel *et al.*, 2013). However, substitution of the Rho-independent terminator of GlmY with the terminator of GlmZ was not sufficient to completely confer the ability to bind Hfq with high affinity to the resultant hybrid GlmYZ-H3. Nonetheless, binding affinities were enhanced for hybrid GlmYZ-H3 as well as for hybrid GlmYZ-H4 that encompasses GlmY with the GlmZ-specific single-stranded base-pairing region (Fig. 5.3). Hence, our data indicate that these two domains provide independent binding sites for Hfq, but both of these modules are required to achieve high affinity binding to Hfq.

This result is in agreement with previous data indicating that an internal U-rich sequence preceding the Rho-independent terminator and a protruding poly-(U)-tail are both required to form an Hfq-binding module within an sRNA. Further, this module was proposed to be structurally distinct of the base-pairing region, but both modules may overlap (Ishikawa *et al.*, 2012). Interestingly, GlmZ contains a UGUUUU sequence adjacent to the terminator stem and thus fulfills the requirements for an independent 3' Hfq-binding module. In contrast, the corresponding sequence of GlmY (UCCAUC) encompasses two cytosine-residues, which have been suggested to strongly reduce binding by Hfq (Ishikawa *et al.*, 2012, Fig. 5.1). Thus, the UGUUUU sequence in GlmZ is likely to contribute to Hfq-binding and therefore might explain the requirement for the complete 3' end or at least the sequence preceding the Rho-independent terminator for efficient binding by Hfq *in vitro* (Fig. 5.3).

In concordance, northern blot analysis of the stabilities of the hybrid sRNAs in presence and absence of Hfq revealed that the exchange of the 3' end of GlmY with the 3' end of GlmZ conferred the need for Hfq for stability to the resultant hybrid GlmYZ-H1. In contrast, the inverse hybrid GlmYZ-H2 did not rely on Hfq for stability (Fig. 5.6 A and B). Again, hybrids GlmYZ-H3 and H4 carrying only parts of the 3' module were only slightly affected by the absence of Hfq (Fig. 5.6 C and D), whereas substitutions within the 5' end of the sRNAs did not influence the stability of the resultant chimeras (Fig. 5.7 A-D). Hence, in agreement with *in vitro* binding studies, the substitution of the 3' end of GlmY with the 3' end of GlmZ renders the hybrid dependent on Hfq. Our data are in agreement with a model, in which initial binding of Hfq occurs by its proximal face at the protruding 3' poly-(U)-sequence of GlmZ, possibly involving the UGUUUU-sequence adjacent to the Rho-independent terminator (Ishikawa *et al.*, 2012; Sauer and Weichenrieder, 2011). In accordance with a study by Sauer *et al.*, 2012, protection against RNases *in vivo*, or in this case, against RapZ and RNase E, might

occur by extension of the interaction with Hfq along its lateral surface, possibly involving the A/U-rich single-stranded region (Fig. 5.11). So far, it is unclear whether this extended interaction along the rim leads to changes within the structures of the 5' stem loops or if the interaction with the 26-nt long single-stranded region is sufficient to stabilize GlmZ *in vivo*. However, it is tempting to speculate that rearrangements within the second stem loop upon Hfq-binding might protect GlmZ from binding of RapZ, which was shown to be dependent on structural features within this region of GlmZ and GlmY (Göpel *et al.*, 2013). Upon binding of the *glmS* mRNA to the distal face of Hfq, GlmZ might be released and thus duplex formation is facilitated (Fig. 5.11).



**Figure 5.11: The 3' end of GlmZ comprises a high affinity binding site for Hfq, whereas determinants for cleavage by RNase E reside within the 5' end of GlmZ.** Model for Hfq-binding and recognition of GlmZ by RNase E and RapZ. Hfq binding stabilizes sRNA GlmZ and facilitates duplex formation with *glmS*. Initial binding by the proximal face of Hfq presumably occurs at the 3' poly-(U)-tail and a U-rich sequence preceding the Rho-independent terminator of GlmZ. Binding might then be extended along the rim (lateral face), thus protecting GlmZ from degradation. Upon binding of the *glmS* mRNA to the distal face, GlmZ might be released from its interactions with residues along the rim. Hence, base-pairing with *glmS* is facilitated. Under conditions of ample glucosamine-6-phosphate, GlmZ is subject to degradation. RNase E presumably recognizes the structure of the lateral bulge within the second stem loop of GlmZ. Together with adaptor protein RapZ, RNase E presumably recognizes the structure of the lateral bulge within the second stem loop of GlmZ and subsequently cleaves a downstream A/U-rich sequence. After initial cleavage, GlmZ is subject to degradation.

It is worth noting, that the sRNA species constitutively expressed from plasmids behaved somewhat differently as compared to the endogenous sRNAs (Fig. 5.5 B and C). Whereas the constitutive expression of plasmid encoded GlmY ensured a constant amount of the sRNA that only slightly decreased upon entry of stationary phase, chromosomally encoded GlmY strongly accumulates during transition to stationary phase due to up-regulation of its  $\sigma^{54}$ -dependent promoter by the GlrK/GlrR two-component system (Fig. 5.5 A-C; Reichenbach *et al.*, 2009). Similarly, as the overall amounts of GlmZ are increased when expressed from a plasmid, the dependency on Hfq was detectable by a decrease in the abundance of the full length form of GlmZ (Fig. 5.5 A). Notably, over-expression of GlmZ in wild type cells led to the appearance of an uncharacterized shorter variant of GlmZ (Fig. 5.5 A, B). This species is reminiscent of a shorter variant previously observed upon inactivation of RNase E (Göpel *et al.*, 2013). However, the origin of this species is so far unknown.

In a second approach, we used the library of sRNA hybrids to determine molecular requirements for recognition and processing by RNase E. In *E. coli*, RNase E is an essential enzyme organizing the RNA

degradosome protein complex and promoting turn-over of various RNA species (Górna *et al.*, 2012). However, the mechanism by which substrate specificity of RNase E is achieved is still unclear. Analogous to de-capping of eukaryotic mRNA, RNase E requires the generation of 5' monophosphates to initiate degradation of some of its substrate RNAs (Deana *et al.*, 2008). Another subset of target RNAs, however, is recognized independently of its 5' phosphorylation state presumably through structural features (Kime *et al.*, 2010; Mackie, 2013b). Recently, protein RapZ was described to act as an adaptor protein enabling RNase E to degrade sRNA GlmZ independently of a 5' monophosphate (Fig.5.8; Göpel *et al.*, 2013). Thus, protein co-factors might contribute to substrate recognition by direct entry. In this study, we exploited the fact that sRNA GlmZ is efficiently cleaved *in vitro* by RNase E in presence of RapZ, whereas the homologous sRNA GlmY is completely stable as GlmY is not a substrate for RNase E (Fig. 5.9; Göpel *et al.*, 2013).

*In vitro* cleavage assays of the various GlmYZ hybrids revealed that, in contrast to Hfq-binding, the 3' ends of GlmY and GlmZ were not involved in recognition and processing by RNase E in presence of RapZ (Fig. 5.9). However, swapping of the 5' end of GlmY with that of GlmZ resulted in a hybrid (GlmYZ-H2) that was efficiently processed. Interestingly, complete processing was observed at highest RNase E concentrations even in absence of RapZ but was significantly enhanced upon addition of RapZ (Fig. 5.10). It is possible that the structure surrounding the cleavage site might be subtly altered in this hybrid since the single stranded region is considerably shorter than in GlmZ and may not be able to interact with residues of the stem loops to form intricate structures. Following studies of the influence of the separate 5' stem loop structures, revealed that the molecular discriminants for RNase E-dependent processing reside within the second stem loop of sRNA GlmZ, more precisely, in the lateral bulge contained within the second stem loop (Fig. 5.10, Fig. 5.11). This region is sufficient to confer the ability to pose as RNase E substrate to sRNA GlmY. The resultant hybrid was effectively processed, although complete processing required twice as much RNase E as processing of GlmZ or a hybrid encompassing GlmY with the complete second stem loop of GlmZ (GlmYZ-H6; Fig. 5.10). Further, an inverse hybrid encompassing GlmZ with the lateral bulge of GlmY was less efficiently processed as compared to the efficiency for processing of GlmZ (GlmYZ-H8; Fig. 5.10). This is in agreement with previous data indicating that the G-residues within the lateral bulge of GlmZ are required for efficient processing (Göpel *et al.*, 2013).

Interestingly, GlmY possesses G-residues in conserved positions, thus the swap of lateral bulges left the G-residues intact, but changed the surrounding nucleotides and possibly the structure of the bulge (Fig. 5.1). A major difference is a protruding C-residue in the lateral bulge of GlmY (C131). In GlmZ, this position is reserved for a U-residue (U134). However, if these residues are important for recognition by RNase E remains to be tested. Taken together, our findings indicate that the lateral

bulge of sRNA GlmZ is of utmost importance for RNase E-dependent processing. However, other features within the second stem loop of GlmZ may contribute to recognition by RNase E, as the exchange of the lateral bulge in GlmZ for the corresponding sequence in GlmY impaired cleavage but did not completely abolish processing (Fig. 5.10 C). Additionally, the second stem loop of GlmY might have a slightly different fold than that of GlmZ. This might possibly also explain the observation that RapZ, which binds to the central stem loops of both sRNAs, has a slightly higher affinity for sRNA GlmY (Göpel *et al.*, 2013).

Thus, recognition of an sRNA as a substrate relies on highly specific structural features, as RNase E discriminates between the highly homologous sRNAs GlmY and GlmZ (Fig. 5.9). In contrast, the fact that RNase E, in concert with RapZ, was capable to cleave the GlmYZ hybrids containing the second stem loop or lateral bulge of GlmZ again confirms the observation that RNase E has little sequence specificity in regard to cleavage sites and that the presence of a single-stranded A/U-rich region is sufficient for cleavage (Carpousis, 2007; Kim *et al.*, 2004). A recent study reported that an increase of the number of unpaired nucleotides at the site of cleavage might have stimulatory effects on recognition by direct entry (Mackie, 2013b). However, in case of GlmY and GlmZ, 12 unpaired nucleotides as present in GlmY seem sufficient for cleavage of hybrids carrying 3' ends derived from GlmY. Thus, the direct entry mechanism may rely on different structural features in various substrate RNAs. The fact that the sRNAs in the *in vitro* assay all carried 5' tri-phosphates and possess a stable stem loop at the 5' end indicated that the 5' phosphorylation state has no influence on processing of GlmZ. Supporting this notion, we found that treatment with tobacco acid pyrophosphatase did not enhance the cleavage efficiency of RNase E for sRNA GlmZ and that efficient processing of 5' monophosphorylated GlmZ still relied on the presence of RapZ (Fig. 5.8). Hence, we excluded that RapZ might act as a specific pyrophosphorylase converting the 5' tri-phosphate of GlmZ to a monophosphate. However, the exact mode of action for RapZ is subject to current investigations in our laboratory.

## ACKNOWLEDGEMENTS

We are thankful to Jörg Vogel for the generous gift of Hfq and to Marcus Resch for help with purification of RNase E. Birte Reichenbach is acknowledged for construction of plasmid pBGG296 and Sabine Zeides is thanked for construction of plasmids pBGG331 and pBGG332. Jörg Stülke is thanked for laboratory space. Y.G. received a fellowship of the Dorothea Schlözer Program of the Göttingen University. This work was supported by grants of the DFG priority program SPP1258 "Sensory and regulatory RNAs in prokaryotes" to B.G..

## 6. DISCUSSION

Post-transcriptional control of gene expression by small regulatory RNAs has emerged as a major principle of gene regulation in bacteria. Due to their vital role in practically all physiological circuits, including regulation of metabolic pathways and adaptation to various stress conditions, sRNAs must be elaborately controlled (Storz *et al.*, 2011; Gottesman and Storz, 2011). Expression of sRNAs can be intricately controlled by specific transcriptional regulators, two-component systems and global regulators, such as alternative sigma factors (Gogol *et al.*, 2011; Göpel and Görke, 2012a). As opposed to transcriptional control, regulation at the level of decay is less understood. This work focuses on the regulation of homologous sRNAs GlmY and GlmZ at the level of biogenesis and decay. Both sRNAs act in a hierarchical cascade to attune expression of glucosamine-6-phosphate (GlcN6P) synthase GlmS to the concentration of its enzymatic product. Thus, GlmY and GlmZ are of utmost importance to mediate GlcN6P homeostasis. Interestingly, in *E. coli*, the amount of GlmY is extensively controlled at the level of synthesis, while sRNA GlmZ abundance is controlled at the level of decay.

### 6.1 Transcriptional control of small RNAs GlmY and GlmZ

#### **Transcriptional control of small RNAs GlmY and GlmZ in *Enterobacteriaceae***

In *E. coli*, expression of sRNA GlmY is controlled by two perfectly overlapping promoters, a seemingly unregulated  $\sigma^{70}$ -dependent promoter that is active mainly during exponential growth, and a  $\sigma^{54}$ -dependent promoter (Reichenbach *et al.*, 2009). During transition to stationary phase, the  $\sigma^{54}$ -promoter of *glmY* is strongly activated by its cognate two-component system GlrK/GlrR. Signal perception leads to auto-phosphorylation of sensor kinase GlrK. Subsequently, the phosphoryl-group is transferred to aspartyl-residue 56 in response regulator GlrR. Upon phosphorylation, GlrR binds to three conserved sequence motifs far upstream of the *glmY* promoter and activates transcription in concert with integration host factor IHF (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011). Interestingly, *in silico* studies predicted three distinct types of promoter architectures existent in *Enterobacteriaceae* (Reichenbach *et al.*, 2009). By extending the investigation of *glmY* and *glmZ* expression to *Yersinia pseudotuberculosis* and *Salmonella enterica* serovar Thyphimurium as representatives of these groups, we were able to verify the following scenarios: (I) In *Y. pseudotuberculosis* expression of both sRNAs, GlmY and GlmZ, is driven solely from  $\sigma^{54}$ -promoters; (II) perfectly overlapping  $\sigma^{70}$ - and  $\sigma^{54}$ -dependent promoters control expression of *glmY* in *E. coli* and of both sRNA genes in *S. typhimurium*; (III) in contrast, *glmZ* of *E. coli* is constitutively expressed from a  $\sigma^{70}$ -promoter. As the  $\sigma^{54}$ -dependent promoters are regulated by GlrK/GlrR, *glmY* and *glmZ* presumably compose a regulon in *Y. pseudotuberculosis* and the majority of enterobacteria.

However, this regulon seems to be in transition to a  $\sigma^{70}$ -dependent system in a subset of species, such as *E. coli* and *S. typhimurium* (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011). This peculiar promoter arrangement might have the advantage to diversely regulate transcription of the sRNAs due to different stimuli. For example, the  $\sigma^{70}$ -dependent promoter might be sufficient to ensure basal expression of the sRNAs for regulation of glucosamine-6-phosphate homeostasis during exponential growth. Further, the  $\sigma^{54}$ -dependent promoter might enhance sRNA transcription in response to activation of GlrK/GlrR and thus supply GlmY (and GlmZ) for an alternative purpose, the nature of which is yet to be determined. Moreover, it is possible that this promoter organization is a result of the ongoing transition of  $\sigma^{54}$ -dependent promoters toward a  $\sigma^{70}$ -dependent system. As many  $\sigma^{70}$ -promoters have been shown to possess extensive functional overlap with  $\sigma^{24}$ - and  $\sigma^{32}$ -dependent genes in *E. coli*, transcription from two overlapping promoters could serve a more general function, namely to increase transcription of  $\sigma^{70}$ -dependent genes (Wade *et al.*, 2006). A previous study reported that 14% of all  $\sigma^{54}$ -dependent genes in *E. coli* are not restricted to  $\sigma^{54}$ , but can be transcribed by  $\sigma^{70}$ -RNA polymerase, at least *in vitro* (Zhao *et al.*, 2010). Whether these genes possess overlapping or consecutive promoters is as of yet unknown.

#### **What is the purpose of stationary phase induction of GlmY (and GlmZ)?**

In stationary phase, expression from  $\sigma^{70}$ -dependent promoters in general and, from the *E. coli* *glmZ*- and *glmUS*-promoter in particular, is strongly reduced (Reichenbach, 2009). This is achieved by 6S RNA that sequesters and stores the  $\sigma^{70}$ -RNA polymerase during stationary phase by mimicking an open promoter complex (Wassarman, 2007). Furthermore, expression of anti-sigma factor Rsd is activated in stationary phase (Jishage and Ishihama, 1998). Rsd interacts with sigma factor  $\sigma^{70}$  and releases the RNA polymerase core enzyme, thus replenishing the pool of RNA polymerase for alternative sigma factors, especially  $\sigma^5$  directing transcription of genes important in stationary phase. As cellular growth ceases, GlmZ and GlmS are no longer needed to provide precursors for the cell wall and lipopolysaccharides. In contrast, sRNA GlmY strongly accumulates in stationary phase (Reichenbach *et al.*, 2009). Thus, it is conceivable that GlmY and in the majority of enterobacteria also GlmZ, serve additional functions under these conditions. Since *glmY* and *glmZ* are the only known genes controlled by the GlrK/GlrR two-component system, it is likely that the aforementioned additional functions of the sRNAs are directly related to stimuli sensed by GlrK. Supporting this notion, an *in silico* analysis of possible GlrR binding sites within the *E. coli* chromosome did not reveal any other genes that might be regulated by GlrR (Göpel and Görke, unpublished).

However, the signal perceived by GlrK has yet to be identified, but the GlrK/GlrR two-component system has been implicated in virulence in some pathogenic bacteria. Virulence of these pathogens

has been shown to rely on cell contact (Flamez *et al.*, 2008; Reading *et al.*, 2007; Reading *et al.*, 2009). However, genes for *glrK* and *glrR* are conserved in various enterobacterial species; most of them are non-pathogenic (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011; Fig. S2.3, S2.4). Furthermore, drastic stationary phase induction of *glmY* expression can be observed in *E. coli*, even when grown in LB medium (Reichenbach *et al.*, 2009; Fig. 3.2 B). Thus, the stimulus sensed by GlrK is likely to be rather general as opposed to a species-specific virulence signal. Supporting this conclusion, a recent study suggested that cells lacking *glmY* are significantly more susceptible to cell envelope stress than wild type cells (Hobbs *et al.*, 2009). Coherently, the expression of *glmY* seems to be somewhat increased upon administration of cell envelope stress using sub-lethal amounts of SDS and EDTA (Künzl, Göpel and Görke, unpublished). Hence, cells might respond to envelope stress signals with modestly increased amounts of GlmY. This may also offer an explanation for the drastic reduction of virulence upon deletion of *glrK/glrR* in pathogens that rely on cell contact with their host cells and indicate the involvement of sRNA(s) GlmY (and GlmZ) in these processes. Thus, in a next step, the stimulus sensed by GlrK should be identified in order to draw conclusions as to additional functions of GlmY and GlmZ. Further, the implication of this system in regard to cell envelope stress response and virulence in correlation with host cell contact remains to be demonstrated.

#### **Outer membrane protein YfhG stimulates GlrK/GlrR-dependent expression of *glmY***

The genes encoding sensor kinase GlrK and response regulator GlrR are highly conserved in *Enterobacteriaceae* and located directly downstream of the *glmY* gene (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011; Fig. S2.2). Interestingly, a third gene, *yfhG*, is localized in between the *glrK* and *glrR* genes. As genes that encode proteins with related functions often co-localize, outer membrane lipoprotein YfhG may be involved in GlrK/GlrR two-component signaling. A functional connection between YfhG and GlrK/GlrR has been previously suggested for enterohemorrhagic *E. coli* (EHEC), however the nature of the connection remained elusive (Reading *et al.*, 2007; Reading *et al.*, 2009).

Here, we could demonstrate that YfhG, similar to GlrR, is strictly required for activity of the  $\sigma^{54}$ -dependent *glmY* promoter as a deletion of the *yfhG* gene almost completely abolished transcription of *glmY* (Fig. 3.1 A). Coherently, *yfhG* on a plasmid complemented a *yfhG* deletion to a level of *glmY* transcription rate that exceeded wild type levels by two-fold (Fig. 3.1 B). As mentioned above, it is tempting to speculate that the GlrK/GlrR two-component system may sense membrane integrity or composition. However, as the signal sensed by GlrK has yet to be identified, it might be possible that signal perception occurs indirectly via outer membrane protein YfhG. In this scenario, YfhG or interaction between YfhG and GlrK might be an indicator for the composition and flexibility of the

outer membrane. Thus, the GlrK/GlrR system might be a three-component system requiring YfhG for functionality.

### **Expanding the complexity of transcriptional control of sRNA GlmY: protein RapZ modulates *glmY* expression in *E. coli***

The *rapZ* gene is an “evolutionary old” gene that is conserved in a number of bacterial phyla (Pompeo *et al.*, 2011). In contrast, small RNAs GlmY and GlmZ are only highly conserved in enterobacteria (Reichenbach *et al.*, 2009; Göpel *et al.*, 2011; Fig. S2.3, S2.4). Hence, it is plausible that RapZ and its homologs serve diverse functions in different bacteria, but may even share a wide spread common function. Previously, RapZ was reported to be involved in the control of  $\sigma^{54}$ -dependent transcription of *glmY* (Reichenbach, 2009). In this work we presented more detailed data on the role of RapZ as modulator of  $\sigma^{54}$ -dependent *glmY* expression.

Activity of the *glmY* promoter is strongly reduced upon deletion of the *rapZ* gene and fails to induce *glmY* expression at the onset of stationary phase (Fig. 3.2 B). Consistently, over-expression of *rapZ* enhanced *glmY* expression ~two to three-fold (Figs. 3.2 C,D; 3.5 B; 3.6 A). In contrast, transcription from other  $\sigma^{54}$ -dependent promoters, the *zraP* promoter and the  $P_{glnA2}$  promoter of *glnA*, was not influenced by a deletion of the *rapZ* gene. However, at least in the case of the *zraP* promoter, over-expression of *rapZ* modestly enhanced transcription (Fig. 3.4). Thus, RapZ might be capable to affect transcription from more than one  $\sigma^{54}$ -dependent promoter, though not all promoters might strictly require RapZ for activity.

Since RapZ is known to bind sRNAs GlmY and GlmZ, its effect on the  $\sigma^{54}$ -dependent *glmY* promoter might be indirectly mediated through a so far unknown sRNA. Previous work reported that regulation of *glmY*-transcription by RapZ still occurred in mutants lacking the *glmY* and *glmZ* genes (Reichenbach, 2009). Thus, it is unlikely that the RapZ-mediated control of *glmY* expression serves as an auto-regulatory feedback-loop within the regulatory GlmYZ-cascade. Further, the function of RapZ in *glmY* promoter control seems to be completely independent of its recently described function as an RNase E adaptor protein promoting turn-over of sRNA GlmZ (Göpel *et al.*, 2013). Corroborating this conclusion, a RapZ quadruple mutant, that is unable to bind GlmY and GlmZ, as well as the *Bacillus* homologue YvcJ are able to complement a *rapZ* deletion in regard to *glmY* transcription rates, but not in respect to regulation of *glmS* expression (Fig. 3.5 B and C; Göpel *et al.*, 2013). Moreover, expression of *glmY* is independent of RNA chaperon Hfq (Fig. 3.5 A). As most base-pairing sRNAs rely on Hfq for functionality and stability, at least in Gram-negative bacteria (Urban and Vogel, 2007; Vogel and Luisi, 2011), the involvement of a so far unidentified base-pairing sRNA in *glmY* promoter control seems unlikely.

RapZ absolutely depends on the *glmY*-specific  $\sigma^{54}$ -activator protein GlrR to modulate *glmY* expression. Thus, it is possible that RapZ acts indirectly on GlrR rather than independent of the two-component system. In principle, two general mechanisms could be envisioned for the RapZ-mediated regulation of *glmY* transcription from its  $\sigma^{54}$ -dependent promoter: First, RapZ may act by altering the activity of a sigma factor, in this case  $\sigma^{54}$ , which also requires GlrR for open complex formation at the *glmY*-promoter (Reichenbach *et al.*, 2009). In this case, more than one  $\sigma^{54}$ -dependent promoter might be influenced by RapZ. However, other features of the promoter, e.g. number of IHF or activator binding sites might influence the degree of modulation by RapZ. Or second, RapZ might act as a modulator of the activity of a transcriptional regulator, e.g. response regulator GlrR. Thus the positive influence of RapZ on other  $\sigma^{54}$ -dependent promoters might possibly be attributed to regulatory cross-talk between RapZ and other response regulators. Note that RapZ was constitutively expressed from a low copy number plasmid, a pSC101-derivative (Kalamorz *et al.*, 2007), and thus RapZ-levels were elevated in complementation experiments as compared to cells expressing endogenous *rapZ* from chromosome. All  $\sigma^{54}$ -dependent promoters require specific transcriptional activator proteins, many of these are response regulators of two-component systems (Reitzer and Schneider, 2001). As response regulators share similar structural features (Jung *et al.*, 2012), RapZ might be able to interact with related domains, for example the  $\sigma^{54}$  interaction domain, within response regulators to enhance their activity. Yet, only a subset of these, in this scenario GlrR, might strictly depend on RapZ for activity. However, the molecular mechanism underlying the regulation of *glmY* transcription by RapZ still remains to be clarified (see below).

#### **Acetylation is required for activity of the *glmY*-promoter**

Recently, modulation of protein activity by acetylation has emerged as a new regulatory principle in bacteria (Wolfe, 2005). Protein acetylation was reported to regulate enzyme activities in central metabolism, influence chemotaxis and to add to the layer of control of the RNA polymerase holoenzyme aiding in stress resistance (Lima *et al.*, 2011; Lima *et al.*, 2012; Starai *et al.*, 2002; Yan *et al.*, 2008). Whereas protein modification by acetylation and its effects on gene expression and central metabolism are well studied in eukaryotes, the extend of protein acetylation and its role in bacteria and archaea is poorly understood (Albaugh *et al.*, 2011; Hu *et al.*, 2010).

A genetic screen identified sirtuin deacetylase CobB as a modulator of  $\sigma^{54}$ -dependent expression of *glmY* (Künzl, Göpel and Görke, unpublished). Upon over expression of *cobB*, the activity of the *glmY*-promoter was drastically reduced in minimal as well as complex medium (Fig. 3.7; Künzl, Hoffmann, Göpel and Görke, unpublished). It first remained unclear how acetylation fitted into the aforementioned complex regulation of *glmY* expression.

Gcn5-related acetyl-transferase YfiQ (Pat) and its homologs were suggested to be the major acetyl transferases in bacteria, as they are widely conserved (Hu *et al.*, 2010). These enzymes catalyze N<sup>ε</sup>-acetylation of lysine-residues that can be reverted by deacetylases such as CobB. In fact, YfiQ (Pat) and CobB have been shown to inversely regulate the activity of metabolic enzymes, response regulators and even of RNA polymerase (Hu *et al.*, 2010; Hu *et al.*, 2013; Lima *et al.*, 2011; Starai *et al.*, 2002). Thus, it is conceivable that there is a larger regulatory overlap between CobB and YfiQ, i.e. CobB and YfiQ may regulate the activity of many other proteins by reversible acetylation. However, here we showed that in contrast of CobB, YfiQ (Pat) is not involved in modulation of the activity of the *glmY* promoter (Fig. 3.7 F). In contrast, mutants lacking the genes encoding acetate kinase AckA and phospho-transacetylase Pta, showed a stark reduction in the activity of the *glmY* promoter (Fig.3.8 B). Enzymes AckA and Pta generate the high energy intermediate acetyl-phosphate from acetate or acetyl-CoA, respectively (Fig. 3.8 A, Wolfe, 2005).

### **RapZ is acetylated *in vivo***

Interestingly, a recently conducted global proteomics study by Zhang and colleagues suggested that RapZ might be acetylated *in vivo*, presumably at lysine residue K251 in close proximity of the predicted RNA-binding domain in RapZ homologs of *Enterobacteriaceae* (Fig. 3.5, Göpel *et al.*, 2013; Zhang *et al.*, 2009). In this work, we demonstrated by western blot that RapZ is indeed acetylated *in vivo* (Fig. 3.8 C).

As mentioned above, acetylation of RapZ was observed to be dependent on acetyl-P generated by the AckA-Pta pathway (Fig. 3.8 A). Upon disruption of the AckA-Pta pathway, which leads to drastically reduced levels of acetyl-P (Mizrahi *et al.*, 2006), the activity of the *glmY* promoter was strongly reduced as well. This inhibitory effect may at least in part be attributed to reduced auto-phosphorylation of GlrR with acetyl-P, as response regulator GlrR showed enhanced DNA-binding activity after treatment with acetyl-P and was, thus, suggested to auto-phosphorylate with acetyl-P *in vitro* (Göpel *et al.*, 2011; Fig. 2.7). Regardless, over-expression of *rapZ* restored high *glmY'-lacZ* expression levels even in absence of AckA and Pta (Fig. 3.8 B). Further, we found that the acetylation level of RapZ increased over time when incubated with acetyl-phosphate *in vitro* (Fig. 3.8 D). Interestingly, a recently published story reported that most acetylation relies on the availability of acetyl-P *in vivo* and that a majority of proteins are chemically acetylated by acetyl-P *in vitro* (Weinert *et al.*, 2013). Thus, RapZ most likely belongs to the class of acetylated proteins that require acetyl-P and are chemically acetylated independently of Gcn5-related acetyl-transferase YfiQ. Further, acetyl-P dependent acetylation was shown to increase in stationary growth phase; however the relevance of acetylation by acetyl-P remains elusive (Weinert *et al.*, 2013).

So far, there is no evidence that GlrK or GlrR may be acetylated (Hoffmann, Göpel and Görke, unpublished). In sum, our findings suggest that acetyl-P might not only be the source for auto-phosphorylation of GlrR, but may also be important for non-enzymatic acetylation of RapZ. Both of these processes are likely to contribute to the regulation of *glmY* expression. It is tempting to speculate that acetylation of RapZ may provide a functional switch as the proposed acetylation site is located close to the predicted RNA-binding site (Fig. 3.5, Göpel *et al.*, 2013; Zhang *et al.*, 2009). Further, control of the  $\sigma^{54}$ -dependent *glmY* promoter and feedback regulation of *glmS* expression are two distinct and independent functions of RapZ (see Chapter 3). Hence, acetylated RapZ might be active in promoter control, whereas non-acetylated RapZ may function within the GlmY/GlmZ sRNA cascade.

To test these hypotheses, first the predicted acetylation site should be verified. To this end, the lysine residue at position 251 could be replaced with an alanine residue that can no longer be acetylated. This RapZ variant should then be tested in its ability to activate *glmY* expression and its acetylation state should be determined. If this lysine residue is indeed the (sole) acetylation site in RapZ, a drastic difference should be observed when comparing the acetylation state of wild type RapZ and the aforementioned variant. Further, this mutant should be tested in regard to its function within the GlmY/GlmZ sRNA cascade to again confirm that both functions are independent of each other. In case, that RapZ should be acetylated at more than one residue or the predicted acetylation site should not be acetylated *in vivo*, the acetylated residue(s) could be detected by mass spectrometry.

Second, the function of acetylation and the molecular mechanism by which RapZ stimulates the *glmY* promoter remain to be clarified. Are the Walker A and B motifs in RapZ required for stimulation of the *glmY* promoter? And does RapZ require acetylation to bind and/or hydrolyze nucleotides? To answer these questions, mutants of RapZ that carry nucleotide exchanges within the Walker A motif should be assessed in respect to their ability to stimulate the *glmY* promoter. Further, the rate of ATP and GTP hydrolysis could be measured using purified acetylated RapZ and the acetylation deficient variant. Further, it should be determined by different methods whether RapZ is capable to interact with RpoN or GlrR to determine the mechanism by which RapZ activates the *glmY* promoter. Co-purification approaches, bacterial two hybrid assays and *in vitro* gel retardation experiments could be used to investigate a possible protein-protein interaction between these three regulators. Though RapZ failed to bind a DNA fragment encompassing the *glmY*-promoter in previous EMSAs (Reichenbach, 2009), it could not be excluded, that RapZ interacts with GlrR. Further, a GlrR/RapZ complex might bind to the *glmY* promoter region and thus effectively stimulate transcription. Preliminary phosphoryl-transfer assays indicated that RapZ does not stimulate auto-phosphorylation

of GlrK or enhance phosphoryl-transfer to GlrR (data not shown). However, RapZ might modulate GlrR activity by other means. Moreover, RapZ may also act by altering the activity of RpoN or conferring stronger selectivity for certain types of  $\sigma^{54}$ -dependent promoters. Thus, further experiments are needed to shed light on the mechanism underlying transcriptional regulation of *glmY*.

## **6.2 Targeted decay of small RNA GlmZ by RNase adaptor RapZ and counteraction by sRNA GlmY**

### **RapZ is a novel RNA-binding protein**

Protein RapZ was shown to be involved in regulation of *glmS* expression by the hierarchical cascade composed of small RNAs GlmY and GlmZ (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008; Urban and Vogel, 2008). In mutants lacking RapZ, GlmZ accumulates in its full-length form and is no longer subject to processing. Genetic analyses revealed that RapZ is involved in the step of signal transmission from GlmY to GlmZ (Fig. S4.1 B; Reichenbach *et al.*, 2008). Initial data suggested that RapZ most likely acts by binding GlmY and GlmZ *in vivo* and *in vitro* (Göpel, Waldmann and Görke, unpublished).

In this work we demonstrated that RapZ is indeed a novel kind of sRNA binding protein that specifically binds GlmY and GlmZ sRNAs, presumably by contacting shared structural features within the central stem loop of both sRNAs (Fig. 4.4; Göpel *et al.*, 2013). Moreover, data indicate that the RNA-binding domain is located at the C-terminus of RapZ. Interestingly, a putative RNA-binding motif can be predicted exclusively for RapZ homologues from bacteria that also contain GlmY and GlmZ, thus indicating that the molecular mechanism underlying the GlmYZ regulatory cascade may be conserved in these bacteria (Figs., 4.3 A, S4.8; Göpel *et al.*, 2013). A combination of four amino acid exchanges in the putative C-terminal RNA-binding domain substituting lysine residues 270, 281 and 283 as well as arginine residue 282 with an alanine, abrogated RNA binding activity of the resultant RapZ quadruple mutant (RapZ<sub>quad</sub>, Figs. 4.3 C, S4.12; Göpel *et al.*, 2013). These observations support the notion that the RNA-binding domain resides within the C-terminus of RapZ. Furthermore, GlmY and GlmZ were shown to displace each other from RapZ when one or the other sRNA was over-expressed thus indicating a titration mechanism for signal transduction within the GlmYZ-cascade (Fig. 4.6 D; Göpel *et al.*, 2013). Interestingly, the preference for binding of GlmY and GlmZ changed with the intracellular glucosamine-6-phosphate (GlcN6P) concentration: under conditions of ample GlcN6P more GlmZ was bound to RapZ, whereas more GlmY was bound to RapZ when GlcN6P became limiting (Fig. 4.6 E; Göpel *et al.*, 2013). These observations corroborate previous findings that GlmY accumulates upon GlcN6P starvation and subsequently stabilizes GlmZ (Reichenbach *et al.*, 2008). Hence, our collective data suggest that depending on the GlcN6P levels within the cell, RNA-

binding protein RapZ mediates signal transduction from GlmY to GlmZ by differentially binding to both sRNAs (Göpel *et al.*, 2013).

### **RNase E adaptor protein RapZ targets sRNA GlmZ for processing to RNase E and is counteracted by sRNA mimicry GlmY**

As mentioned above, expression of glucosamine-6-phosphate synthase GlmS is controlled by the GlmY/Z sRNA cascade (Reichenbach *et al.*, 2008; Urban and Vogel, 2008). Since GlmY accumulates upon depletion of GlcN6P by a post-transcriptional mechanism (Reichenbach *et al.*, 2009), this sRNA cascade provides a feedback mechanism to attune *glmS* expression to the intracellular GlcN6P level. Interestingly, GlmZ is subject to processing by RNase E, which removes the base-pairing nucleotides and initiates degradation of GlmZ. Our data indicate that processing of GlmZ is dependent on the intracellular GlcN6P concentration. Intriguingly, sRNA GlmZ *per se* is not a substrate for RNase E and processing of GlmZ strictly requires protein RapZ (Göpel *et al.*, 2013; Fig. 4.6 A and B; Fig. S4.1). Under GlcN6P limiting conditions, GlmY sequesters RapZ. In turn, GlmZ is stabilized and activates the *glmS* transcript with the aid of RNA chaperon Hfq (Urban and Vogel, 2008; Reichenbach *et al.*, 2008).

Here, we demonstrated that the molecular mechanism for regulation of *glmS* expression within the GlmYZ cascade depends on RNase adaptor protein RapZ that recruits GlmZ to its processing machinery. Further, we proposed an anti-adaptor function for sRNA GlmY. This function is exerted by GlmY upon GlcN6P starvation and leads to activation of *glmS* by sequestration of the adaptor RapZ (Göpel *et al.*, 2013; Fig. 4.6 D and E). Thus, RapZ might be the first of many proteins programming a general ribonuclease for processing of a specific target (s)RNA.

The molecular mechanism of targeted processing of sRNA GlmZ is strongly reminiscent of the principle of controlled proteolysis in bacteria (Bougdour *et al.*, 2008). In eukaryotes, proteolysis is initiated by ubiquitylation that marks regulatory proteins for degradation by the proteasome (Finley *et al.*, 2012). However, a comparable mechanism does not exist in bacteria. Rather, degradation of proteins relies on specific sequence motifs within substrates or on dedicated adaptor proteins (Bougdour *et al.*, 2008; Hengge, 2009).

Controlled proteolysis employs adaptor proteins that interact with regulatory proteins and recruit the degrading protease complex, thus allowing for degradation of the target protein. Specific anti-adaptor proteins can counteract degradation of the regulator by sequestering the adaptor and thus stabilizing the regulatory protein. A well-studied example is sigma factor  $\sigma^S$ , encoded by the *rpoS* gene. RpoS is the master regulator of stationary phase gene expression and general stress response (Battesti *et al.*, 2011). RpoS is targeted to degradation by ClpXP protease through interaction with

the phosphorylated form of adaptor protein RssB. Several anti-adaptor proteins responding to various stress-related stimuli have been shown to stabilize RpoS in response to different stress conditions (Bougdour *et al.*, 2008). However, since the anti-adaptor proteins do not share any similarity it is currently not known how many anti-adaptors exist and which structural features are required for interaction with the adaptor RssB (Battesti *et al.*, 2011).

As described above, the mechanism employed by the GlnYZ cascade is very similar, with the exception that an sRNA is targeted for degradation by a general ribonuclease and an anti-adaptor sRNA mimicry sequesters the adaptor protein (Göpel *et al.*, 2013). Such an elaborate regulatory circuit provides multiple junctions for integration of additional signals, as exemplified by the recruitment of the  $\sigma^S$ -regulon to various stress conditions by employing a multitude of anti-adaptor proteins, each associated with specific environmental cues (Bougdour *et al.*, 2008). It is tempting to speculate, that similar scenarios might apply for adaptor proteins promoting turn-over of specific (s)RNAs, thus allowing for stabilization of the RNA under certain conditions. However, it is of utmost importance to increase the knowledge about adaptor proteins and the mechanisms they employ, to shed light on the regulatory potential of such a system and whether there is additional input by other regulatory networks or environmental and intracellular cues.

#### **Decay of plastid RNAs in *Arabidopsis* and the protein binding CsrB/CsrC sRNAs in *E. coli* is dependent on further putative RNase E adaptor proteins**

RNase E in *Arabidopsis thaliana* chloroplasts, for example, was recently shown to interact with the RNA-binding protein RHON1 (Stoppel *et al.*, 2012). Interestingly, RHON1 is absolutely required for turn-over of plastid RNAs, as a *rhon1* deficient plant shows the same phenotype as an *rne* deficient plant. Thus, RHON1 was suggested to facilitate efficient cleavage by RNE presumably through determining sequence specificity and delivering target transcripts to a degradosome like complex (Stoppel *et al.*, 2012).

Another example is the *E. coli* protein CsrD that is involved in regulation of the Csr-System. The Csr (Rsm) system coordinates adaptation among major physiological phases or modes of growth, e.g. non-virulent vs. virulent, or sessile vs. motile (Babitzke and Romeo, 2007; Heroven *et al.*, 2012; Romeo *et al.*, 2012). Therefore, this system fine-tunes gene expression to modulate relative fluxes of competing metabolic pathways. This is achieved by the translational regulator CsrA, an RNA-binding protein that modulates translation of its mRNA targets by directly binding to specific GGA-motives often in the vicinity of the translational start sites (Dubey *et al.*, 2005; Romeo *et al.*, 1993; Romeo, 1998). Small RNAs CsrB and CsrC possess a multitude of stem loop structures exhibiting GGA-motifs and are thus able to sequester CsrA in response to certain stimuli that enhance the expression of

these sRNAs (Dubey *et al.*, 2005). Further, this system is subject to a tight control: an auto-regulatory loop between CsrA and the BarA/UvrY two-component system positively regulates CsrB expression, while CsrB and CsrC negatively regulate CsrA (Suzuki *et al.*, 2002; Weilbacher *et al.*, 2003). Protein CsrD was proposed to control degradation of the CsrB and CsrC sRNAs in an RNase E dependent manner (Suzuki *et al.*, 2006). The authors predict that binding of CsrD to CsrB and CsrC alters the structure of these sRNAs, thus converting them to substrates that can be recognized by RNase E. Interestingly, decay of CsrB is dependent on CsrD in absence of CsrA, i.e. like GlmZ, free CsrB sRNA is degraded with the help of the adaptor protein. In contrast, whereas CsrA/CsrC complexes rely on CsrD for degradation, unbound CsrC seems to be degraded via a different pathway independent of CsrD (Suzuki *et al.*, 2006). Although there are striking similarities between the proposed mode of action for the assumed adaptor protein CsrD and the recently identified mechanism that is employed by RapZ, many questions regarding the action of CsrD are still unsolved. Hence, it remains to be determined in future studies whether CsrD is a specific or a general adaptor protein and whether decay of CsrB/CsrC requires direct interaction of CsrD with RNase E.

#### **Adaptor proteins as a means to confer substrate specificity to general Ribonucleases?**

The work presented in this part of the thesis, has two main implications for (s)RNA turn-over: first, specific RNA-binding proteins may be more common in bacteria as previously assumed and second, specific adaptor proteins may confer substrate specificity to general RNases and provide a means for targeted processing of certain RNAs.

The fundamental function of ribonucleases in RNA maturation, processing and degradation ensures cellular survival and prosperity (Górna *et al.*, 2012; Li *et al.*, 2002). Thus, ribonucleases are subject to tightly regulated control at DNA, RNA and protein level (Bardwell *et al.*, 1989; Diwa *et al.*, 2000; Mackie, 2013a). It is known for quite some time that the function of ribonucleases is affected by physiological and environmental conditions and is closely linked to the central metabolism (Bernstein *et al.*, 2004; Del Favero *et al.*, 2008; Górna *et al.*, 2012; Newman *et al.*, 2012). However, the mechanisms by which activity and substrate specificity of ribonucleases are regulated in response to intrinsic and extrinsic cues are still poorly understood.

RNase E of *E. coli*, for example, can recognize its substrates by two distinct mechanisms. The first mechanism relies on interaction of the 5' mono-phosphorylated terminus of the target RNA with a sensory region within the N-terminal catalytic domain of RNase E (Mackie, 2013a). This interaction allosterically activates RNase E for cleavage of the substrate (Callaghan *et al.*, 2005). Interestingly, recent findings discovered that base-pairing sRNAs are also able to stimulate the catalytic activity of RNase E and thus induce the coupled degradation of sRNA/mRNA duplexes. Hence, as transcripts

that are not associated with translating ribosomes are prone to degradation in *E. coli*, sRNAs can either induce decay passively by blocking access for ribosomes but may also act as direct activators of RNase E activity (Baker and Mackie, 2003; Bandyra *et al.*, 2012; Mackie, 2013b). An alternative mechanism, referred to as 'direct entry', is believed to rely on structural features and the fold of the target RNA thus bypassing the need for a 5' mono-phosphate (Kime *et al.*, 2010). However, the exact nature of these structural requirements is so far unknown for most RNA substrates. Recently, protein RapZ was identified as an adaptor protein enabling RNase E to cleave sRNA GlmZ independently of its 5' phosphorylation state (Göpel *et al.*, 2013; Fig. 5.8). Thus, it is conceivable that protein co-factors may contribute to substrate recognition by direct entry.

General RNA binding proteins, such as Hfq, are known to bind and protect certain RNAs from cleavage by their degrading machineries (Sauer *et al.*, 2012; Vogel and Luisi, 2011). Rather than protection achieved through binding of global RNA-binding proteins, recently discovered RNase adaptor proteins may confer dynamic regulation of ribonuclease activity itself (Kim *et al.*, 2008; Stoppel *et al.*, 2012). These regulators can either inhibit RNase activity by altering complex formation with ancillary proteins, or changing the ability of an RNase to multimerize (Gao *et al.*, 2006; Zhao *et al.*, 2006). Furthermore, recent discoveries added RNA specific adaptors, such as RapZ, to the repertoire of ribonuclease regulators (Göpel *et al.*, 2013; Stoppel *et al.*, 2012; Suzuki *et al.*, 2006). These proteins may bind their target RNAs and facilitate efficient cleavage by the associated RNase presumably through protein-protein interaction and either allosteric activation of ribonucleolytic activity or by altering the structure of the targeted RNA (Göpel *et al.*, 2013; Stoppel *et al.*, 2012; Suzuki *et al.*, 2006). Thus, dedicated RNase adaptor proteins possess the potential to determine substrate specificity and integrate intra- and extracellular cues to modulate gene expression at the level of RNA decay.

#### **Other Ribonuclease binding regulators globally modulate RNA turn-over**

Recently, RraA and RraB were isolated in gene bank screening approaches aiming to identify negative regulators for RNase E activity (Gao *et al.*, 2006; Lee *et al.*, 2003). These repressors of RNase E activity were shown to differentially influence various subsets of RNase E target transcripts. RraA was shown to act as a global inhibitor of RNase E activity (Lee *et al.*, 2003). Whereas 336 transcripts are equally affected by RraA or RraB, RraA uniquely affects the steady-state levels of 371 transcripts and 85 distinct transcripts are influenced by RraB. Interestingly, RraA directly interacts with RNase E by binding to both C-terminal RNA-binding sites: the RNA binding domain (RBD) and the arginine-rich region 2 (AR2) (Gao *et al.*, 2006; Górna *et al.*, 2012; Lee *et al.*, 2003). This process was shown to be dependent on DEAD-box RNA helicase RhlB. Proteins containing DEAD-boxes often function to

unwind RNA or DNA duplexes and possess the capacity to remodel protein–nucleic acid interactions employing the free energy obtained from NTP binding and hydrolysis (Pyle, 2008). Thus, it was proposed that RraA stimulates the ATPase activity of RhlB, subsequently leading to remodeling of the degradosome–RNA complex. The now available RNA binding sites RBD and AR2 may then be masked by RraA to impede binding of substrate RNA molecules (Górna *et al.*, 2010). The second inhibitor RraB also interacts with the C-terminal domain, presumably with a region within the coiled-coil domain of RNase E (Gao *et al.*, 2006). As both regulators have been implicated in modulation of the composition of the degradosome and *rraA* expression was shown to be dependent on the general stress sigma factor RpoS ( $\sigma^S$ ) and RNase E, RraA and RraB may globally alter RNA metabolism in response to intra- and extracellular cues (Gao *et al.*, 2006; Lee *et al.*, 2003).

A further example for a global regulator of RNase activity is the *E. coli* YmdB ribonuclease-binding protein that was reported to modulate activity of RNase III (Kim *et al.*, 2008). It was suggested that YmdB and RNase III form heterodimers, which seems to affect RNase III cleavage specificity. YmdB is specific for members of the RNase III family and also inhibits RNase III of *Streptomyces coelicolor*, but does not influence *E. coli* RNase E or RNase G. Vice versa, effectors of RNase E, such as RraA, are without effect on RNase III (Kim *et al.*, 2008). Interestingly, expression of *ymdB* is also dependent on RpoS and YmdB was shown to accumulate in stationary phase or upon cold shock (Kim *et al.*, 2008). Thus, YmdB may be yet another ribonuclease regulator directing specificity of RNase III in stationary phase or upon cold shock.

In sum, ancillary proteins guiding ribonuclease activity and substrate specificity might be more common than previously thought. The identification of novel RNase adaptor proteins may provide explanations on how substrate specificity can be conferred to global RNA processing enzymes. To globally identify new adaptor proteins that may interact with an *E. coli* ribonuclease, the RNases may be expressed as fusion proteins containing one subdomain of the adenylate cyclase of *Bordetella pertussis*. Supplying a gene bank within the bacterial two-hybrid system, a screen could be conducted to globally identify interaction partners of the *E. coli* RNases.

### **On the impact of alternative degradosome compositions on RNA turn-over**

In *E. coli*, the C-terminal domain of RNase E provides a scaffold for the organization of a multi-enzyme RNA degrading complex, the degradosome (Górna *et al.*, 2012). The canonical degradosome consists of endoribonuclease RNase E, polynucleotide phosphorylase PNPase, the ATP-dependent RNA helicase RhlB and a glycolytic enzyme, enolase (Carpousis *et al.*, 1994; Miczak *et al.*, 1996; Vanzo *et al.*, 1998). However, various other proteins have been found to transiently associate with the degradosome either to globally alter RNA-turn-over as suggested for RraA and RraB or to recruit the

degradosome for degradation of specific sRNA/mRNA complexes (Gao *et al.*, 2006; Ikeda *et al.*, 2011; Lee *et al.*, 2003). In the latter case, Hfq has been shown to associate with the degradosome at a region overlapping the RhlB interaction site, thus forming an alternative degradosome that can be targeted to co-degrade sRNA/mRNA duplexes (Ikeda *et al.*, 2011; Massé *et al.*, 2003). Other known interaction partners of the degradosome that might lead to the formation of an alternative RNA degrading complex are the cold-shock associated helicase CsdA (Prud'homme-Genereux *et al.*, 2004), or the ribosomal L4 protein (Singh *et al.*, 2009). These two 'minor' components have been suggested to adapt RNA turn-over to cold shock conditions or, in the case of L4, to nutrient-related stress conditions, such as the stringent response (Kaberdin and Lin-Chao, 2009).

Thus, it is tempting to speculate that RNase adaptor proteins such as RapZ, or possibly CsrD might also be alternative components transiently interacting with the degradosome. At least turn-over of sRNA CsrB assisted by CsrD seemed to require the assembly of the degradosome *in vivo* (Suzuki *et al.*, 2006). Whether RapZ and/or CsrD might be regulators affecting turn-over of other transcripts by altering the composition of the degradosome is so far unclear. To answer the questions raised above, turn-over of GlmZ could be addressed in a degradosome assembly mutant. These experiments should reveal whether the degradosome is required for degradation of sRNA GlmZ, or whether the N-terminal catalytic domain might suffice *in vivo* as observed *in vitro*.

Interestingly, previous micro array analysis suggested that the availability of RapZ might influence the expression of heat shock, as well as cold shock proteins, some chaperons and some biosynthetic genes (Kalamorz, 2008). Taking in account, that RapZ may also influence some of these genes by altering their transcription rates, it might be plausible that RapZ affects at least a subset of the genes with altered expression levels by affecting the stability of their mRNAs. This mode of action was previously proposed for the L4 ribosomal protein that was suggested to stabilize stress-responsive RNAs upon interaction with RNase E (Kaberdin and Lin-Chao, 2009; Singh *et al.*, 2009). Thus, it is tempting to speculate, that RapZ might possess a more global role of RNA turn-over reflecting location of its gene in the pleiotropic *rpoN*-operon. Nevertheless, more experiments are needed to test these notions, for example northern and western blot experiments could be conducted to verify the proposed genes with altered expression levels. Further, mRNA stabilities of these transcripts should be assessed in strains lacking the ability to assemble the degradosome and/or deficient in *rapZ*. Further, processing and translation efficiencies for selected transcripts could be assessed *in vitro* in absence and presence of RNase E or the degradosome, and RapZ.

### **6.3 The GlmY/Z-cascade as a unique model system**

#### **The GlmY/Z-cascade as model system to study molecular requirements for Hfq-binding to sRNAs and determinants for processing by RNase E**

Despite the striking similarities in sequence and structure of the homologous sRNAs GlmY and GlmZ, both regulatory RNAs vastly differ in their mode of action. GlmZ activates expression of its target mRNA *glmS* by base-pairing with an inhibitory structure masking the ribosome binding site of *glmS* (Kalamorz *et al.*, 2007; Urban and Vogel, 2008; Görke and Vogel, 2008). This activation of *glmS* expression is dependent on RNA chaperon Hfq (Kalamorz *et al.*, 2007; Urban and Vogel, 2008). Furthermore, Hfq is required for stability of sRNA GlmZ *in vivo* (Fig. 4.1; Göpel *et al.*, 2013). In contrast, sRNA GlmY acts indirectly on *glmS*, since GlmY lacks the nucleotides required for base-pairing within the *glmS* 5' UTR (Reichenbach *et al.*, 2008; Urban and Vogel, 2008). Concomitantly, GlmY does neither bind Hfq with high affinity *in vitro* nor require Hfq for its stability *in vivo* (Fig. 4.1; Göpel *et al.*, 2013). Rather, GlmY acts by stabilizing the full-length, and thus active, form of sRNA GlmZ (Reichenbach *et al.*, 2008; Urban and Vogel, 2008; Göpel *et al.*, 2013). Under conditions of ample GlcN6P, GlmZ is bound by adaptor protein RapZ and recruited to its processing machinery. Subsequently, RNase E initiates decay of GlmZ. When GlcN6P becomes limiting, sRNA GlmY accumulates by a post-transcriptional mechanism and sequesters adaptor protein RapZ (Reichenbach *et al.*, 2009). Thus, GlmZ is stabilized and activates *glmS* expression in an Hfq-dependent manner (Urban and Vogel, 2008; Göpel *et al.*, 2013). Although highly similar, GlmY is not a substrate for RNase E, and is completely stable in an *in vitro* RNase E cleavage assay, even in presence of RapZ (Fig. 5.9; Göpel *et al.*, 2013).

Thus, GlmZ possesses unique features rendering it an sRNA dependent on Hfq for stability and functionality, as well as an sRNA target for processing by RNase E. In contrast, these features are lacking in homologous sRNA GlmY, which is remarkable due to fact that both sRNA share a sequence identity of 63% (Figs. 4.4 and 5.1 A). Hence, the GlmYZ cascade provides an ideal system to study molecular determinants for Hfq binding and discriminants required for recognition and cleavage by RNase E and RapZ.

In this work, we swapped sequences between sRNAs GlmY and GlmZ, thus creating a library of GlmYZ hybrid sRNAs that were then used as a tool to study the molecular requirements for these two processes (Fig. 5.1 B). To determine the integrity of the overall structure of the hybrid sRNAs we assessed binding affinities of RapZ toward these chimeras. We recently demonstrated that RapZ binds GlmY and GlmZ at a central stem loop structure with high affinity *in vivo* and *in vitro*, but RapZ does not specifically interact with non-cognate sRNAs (Göpel *et al.*, 2013). As all hybrids were bound

with similar affinities as the wild type GlmY and GlmZ sRNAs (Fig. 5.2), we reasoned that the overall structure should be fairly similar as well and that the hybrids could provide a suitable tool for the investigation of the questions raised above.

### **On the discriminants for Hfq-binding**

Our analysis of Hfq-binding affinities of the various hybrids of sRNAs GlmY and GlmZ revealed an essential role for the 3' end of GlmZ in binding of Hfq (Fig. 5.3). As opposed to the vital role of the 3' end, the 5' region of GlmZ was found to be dispensable for high affinity binding by Hfq (Fig. 5.4). These conclusions are based on the observation that hybrid GlmYZ-H1 carrying the 5' stem loops of GlmY fused to the 3' end of GlmZ was efficiently bound by Hfq with an affinity almost comparable to wild type GlmZ (Fig. 5.3). In contrast, in the inverse hybrid, GlmYZ-H2, binding of Hfq was strongly impaired (Fig. 5.3). Upon further dissection of the 3' end, we could conclude that neither the base-pairing sequence nor the Rho-independent terminator of GlmZ alone were sufficient to confer the ability to bind Hfq to GlmY, but rather both of these regions are required for efficient binding by Hfq (Fig. 5.3).

Taking into account, that the 5' region of GlmZ is of minor importance, we find that our observations fit recently established models to describe binding of Hfq to sRNAs. Previous studies revealed the significance of the accessibility and length of the poly-(U) stretch of Rho-independent terminators for initial binding of Hfq to an sRNA thus allowing selectivity despite the grand structural diversity of Hfq-dependent sRNAs (Otaka *et al.*, 2011; Sauer and Weichenrieder, 2011). A discrepancy in accessibility of the protruding U-rich sequence is the most striking difference within the 3' end of sRNAs GlmY and GlmZ (Figs. 4.4 C,D and 5.1 A). Whereas sRNA GlmZ possesses an accessible poly-(U)-tail, the U-rich sequence succeeding the Rho-independent terminator of GlmY is fully sequestered within the stable terminator stem (Fig. 5.1 A). Another remarkable disparity between GlmY and GlmZ is the absence of the base-pairing sequence in GlmY. As a result, the single-stranded A/U-rich region in GlmY encompasses only 12 nucleotides, while GlmZ possesses a single-stranded A/U-rich sequence of 26 nucleotides in length (Fig. 5.1 A).

Interestingly, neither of these two regions, the terminator and the single-stranded base-pairing region, alone was sufficient to confer the ability to bind Hfq with high affinity to the resultant hybrid sRNAs (Fig. 5.3). This is in perfect agreement with a previous study by Ishikawa and colleagues that suggested that a functional Hfq-binding module within an sRNA consists of both, an internal U-rich region preceding the Rho-independent terminator and the terminator with its accessible poly-(U)-tail (Ishikawa *et al.*, 2012). While GlmZ contains a UGUUUU sequence preceding its Rho-independent

terminator, the corresponding sequence of GlmY (UUCCA) does not meet the requirements for an Hfq-binding module (Fig. 5.1; Ishikawa *et al.*, 2012). In contrast, cytosine-residues contained within the sequence preceding the terminator stem were shown to strongly impair binding by Hfq (Ishikawa *et al.*, 2012). In conclusion, the single-stranded A/U-rich base-pairing region, or at least the UGUUUU motif as well as the terminator structure of GlmZ are both required to create a high affinity Hfq-binding site.

### **Transplantation of the 3' end of GlmZ renders a formerly Hfq-independent sRNA dependent on Hfq for stability**

Corroborating our results from the *in vitro* Hfq-binding studies, northern experiments assessing the stabilities of the hybrid sRNAs in presence and absence of RNA chaperon Hfq *in vivo* revealed that the dependence on Hfq for stability of GlmZ resides within the entire 3' end (Fig. 5.6). Substitution of the 3' end in GlmY with that of GlmZ resulted in a hybrid sRNA, GlmYZ-H1, that was bound by Hfq with high affinity (Fig. 5.3) and relied on Hfq for stability *in vivo* (Fig. 5.6). In contrast, the inverse hybrid, GlmYZ-H2, was neither bound by Hfq nor required Hfq for stability (Figs. 5.3, 5.6). As observed in the *in vitro* gel retardation assays, hybrids carrying individual parts of the 3' end, i.e. the single-stranded A/U-rich region (GlmYZ-H4) or the terminator stem loop (GlmYZ-H3), were slightly but significantly less stable in absence of Hfq (Fig. 5.6). This again confirms that the entire 3' module of GlmZ is required as a platform for efficient Hfq-binding *in vivo* and is, thus in complete agreement with data obtained by *in vitro* gel shift experiments (Figs. 5.3, 5.6). Again, domain swapping between GlmY and GlmZ did not alter the stability of the resultant hybrid sRNAs in absence of Hfq as compared to the respective wild type (Fig. 5.7).

Applying our data to a recently proposed model for the mode of sRNA binding by Hfq (Ishikawa *et al.*, 2012; Sauer and Weichenrieder, 2011), GlmZ is most likely initially bound to the proximal face of Hfq via its protruding poly-(U)-tail. This initial recognition is liable to require at least part of the single-stranded A/U-rich region, possibly the UGUUUU motif preceding the Rho-independent terminator (Ishikawa *et al.*, 2012; Fig. 5.11). Since the entire 3' end of GlmZ is required for stabilization of the sRNA by Hfq *in vivo*, it is plausible that protection of GlmZ from RNase E and RapZ relies on the extended interaction of GlmZ with the lateral binding surface of Hfq, as proposed by Sauer and colleagues (Sauer *et al.*, 2012; Fig. 5.11).

It is tempting to speculate that this augmented interaction results in alterations of the secondary structure of GlmZ, thus preventing recognition by RapZ. This might be feasible since RapZ seems to contact structural features within the central stem loop of GlmY and GlmZ (Fig. 4.4; Göpel *et al.*, 2013). Preliminary data indicate that binding of GlmZ by RapZ leads to a highly stable complex.

Further, complex formation between GlmZ and RapZ or GlmZ and Hfq is mutually exclusive as GlmZ can neither be replaced from a preformed RapZ/GlmZ complex by Hfq nor does addition of Hfq lead to the formation of a ternary RapZ/GlmZ/Hfq complex (data not shown). Hence, binding of RapZ seals the fate of GlmZ towards decay, while binding of Hfq protects GlmZ and allows for activation of *glmS* expression. Binding of the *glmS* mRNA to the distal face of Hfq may lead to release of GlmZ from the rim and thus facilitate *glmS*/GlmZ duplex formation (Fig. 5.11).

### **On the determinants for recognition and processing by RNase E**

As mentioned above, apart from substrate recognition by its 5' mono-phosphorylated terminus, the understanding of the requirements for recognition by direct entry is rather limited. As protein RapZ was recently described to act as an adaptor protein guiding sRNA GlmZ for degradation by RNase E (Göpel *et al.*, 2013), it is plausible that protein co-factors might contribute to substrate recognition by direct entry. In this work, we exploited the observation that sRNA GlmZ is efficiently cleaved *in vitro* by RNase E and RapZ, while sRNA GlmY, though vastly similar in sequence and structure, is completely stable (Fig. 5.9).

As opposed to the molecular requirements for Hfq-binding, the 3' end of GlmZ is not required for recognition by RNase E as judged from *in vitro* cleavage experiments assessing the stabilities of the various GlmYZ hybrid sRNAs in presence of RNase E and RapZ (Fig. 5.9). In contrast, the exchange of the 5' end of GlmY with that of GlmZ led to efficient cleavage of the resultant hybrid GlmYZ-H2 (Fig. 5.10 A). Further dissection of the 5' end revealed that the molecular determinants for recognition of GlmZ by RNase E and RapZ reside within the central stem loop (Fig. 5.10). Intriguingly, substitution of the lateral bulge in the second stem loop of GlmY with that of GlmZ resulted in a hybrid that was specifically cleaved by RNase E in presence of RapZ (Fig. 5.10 C, GlmYZ-H7). This is remarkable since only the lateral bulge and lower stem corresponded to GlmZ, whereas the rest of the sequence of this hybrid corresponded to GlmY (Fig. 5.1). However, it is worth noting that complete processing of this hybrid required twice the amount of RNase E necessary for complete processing of wild type GlmZ (Fig. 5.10 A,C). For the inverse hybrid, GlmYZ-H8, cleavage was not abolished but significantly impaired as compared to wild type GlmZ (Fig. 5.10 A,C).

Previous data suggested that the G-residues within the lateral bulge of GlmZ are important for recognition by RNase E, as mutants carrying single nucleotide exchanges in these G-residues showed impaired processing efficiencies while binding to RapZ was not effected (Figs. 4.6 C, S4.18 C; Göpel *et al.*, 2013). As GlmY also contains G-residues at these conserved positions (Fig. 5.1), it is feasible that the context of the G-residues or the fold of the lateral bulge might be important for recognition by RNase E. One striking discrepancy is a protruding C-residue in the lateral bulge of GlmY (C131) at a

position that is reserved for a U-residue (U134) in GlmZ (Fig. 5.1). Whether these residues indeed contribute to recognition by RNase E remains to be tested in future studies.

In sum, our findings suggest that the lateral bulge of GlmZ is essential for recognition by RNase E. Yet, other structural features within the central stem loop of GlmZ are likely to contribute to recognition by RNase E. Supporting this notion, the substitution of the lateral bulge in GlmZ for that GlmY only impaired cleavage but did not completely abolish processing (Fig. 5.10 C). Our data are thus in agreement with previous observations that specific loop structures are important for recognition of a subset of target RNAs by RNase E, however, cleavage might occur elsewhere (Kime *et al.*, 2010; Schuck *et al.*, 2009).

### How does RapZ act on a molecular level?

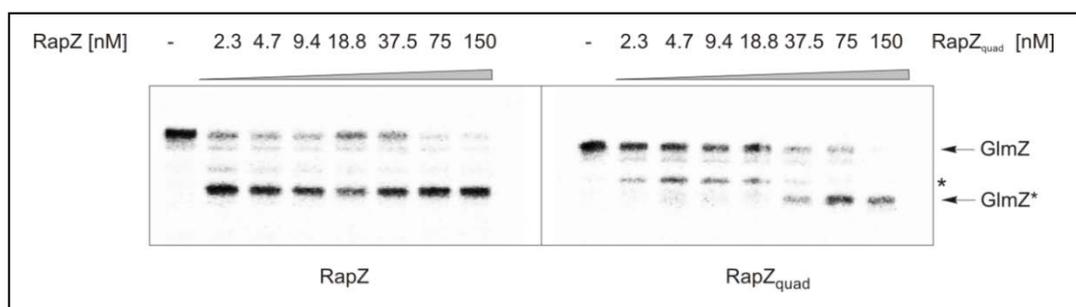
Previously, RapZ has been identified as an adaptor protein that directly interacts with the N-terminal domain of RNase E and promotes specific and efficient cleavage of sRNA GlmZ *in vivo* and *in vitro* (Göpel *et al.*, 2013). Full activation of RNase E *in vitro* required a ~three-fold excess of RapZ over RNase E, which is in agreement with a proposed homotrimeric RapZ complex (Resch *et al.*, 2013). However, how RapZ acts on a molecular level to induce cleavage of GlmZ by RNase E remains elusive. RNase E was shown to act by two distinct modes of substrate recognition: either cleavage is directly stimulated by a 5' mono-phosphate within the RNA or target transcripts are recognized by direct entry independently of the 5' phosphorylation state (Callaghan *et al.*, 2005; Deana *et al.*, 2008; Kime *et al.*, 2010; Mackie, 2013a). GlmZ tested in RNase E *in vitro* cleavage assays carries a 5' tri-phosphate as <sup>32</sup>P-labeled GlmZ is derived from internal labeling during *in vitro* transcription. Thus, as one possibility, RapZ might act as a GlmZ-specific pyrophosphatase. To address this question, we performed *in vitro* processing assays comparing cleavage efficiencies of 5' PPP-GlmZ and a 5' P-GlmZ variant that was treated with tobacco acid pyrophosphatase prior to the RNase E cleavage reaction (Fig. 5.8). However, RNase E displayed very similar cleavage efficiencies for 5' mono-phosphorylated and 5' tri-phosphorylated GlmZ. Moreover, cleavage still required RapZ indicating that RapZ does not simply act by converting the 5' tri-phosphate of GlmZ to a mono-phosphate (Fig. 5.8). Further, GlmZ contains stable stem loop structures at its 5' end that sequester its 5' terminus thus limiting access for pyrophosphatases or RNase E. In fact, in absence of RapZ, GlmZ is completely stable *in vivo* (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008). Thus, it is rather unlikely that degradation of GlmZ is at all dependent on its 5' phosphorylation state.

Hence, GlmZ most likely belongs to those substrates recognized by direct entry (Kime *et al.*, 2010; Mackie, 2013a). Previous studies suggested that the fold of the target transcript or specific stem loop

structures of the respective target RNAs are important for recognition by RNase E, however, cleavage might occur elsewhere (Kime *et al.*, 2010; Schuck *et al.*, 2009). For GlmZ, we could determine a region within the central stem loop that is sufficient to induce processing of the non-substrate sRNA GlmY by RNase E *in vitro* (see chapter 5; Fig. 5.10). Nevertheless, RapZ is still required for RNase E either to gain access to that specific region, to bind to this structure, or to stimulate cleavage. In principle, two possible mechanisms could be anticipated for the mode of action of adaptor protein RapZ. First, binding of RapZ to GlmZ induces structural changes that allow for recognition by RNase E and thus convert GlmZ to a direct substrate for RNase E. Second, RapZ might act as an allosteric activator for RNase E and binding of GlmZ only serves to deliver the sRNA to the degradosome located at the membrane.

In case of the first scenario, binding of RapZ to GlmZ would be expected to result in conceivable structural changes within GlmZ. However, initial structure probing experiments using RNase T1 only showed slightly altered accessibility for nucleotides within the central stem loop and adjacent single stranded region upon binding of RapZ (Fig. 4.4; Göpel *et al.*, 2013). Further, single nucleotide exchanges within the lateral bulge of the second stem loop in GlmZ did not alter binding of RapZ, but strongly impaired processing by RNase E (Figs. S4.18; 4.6 C; Göpel *et al.*, 2013). Thus, it seems unlikely that RapZ acts by simply re-structuring sRNA GlmZ upon binding.

To corroborate this conclusion, we performed *in vitro* cleavage assays comparing cleavage efficiencies for GlmZ when limiting amounts of RNase E (10 nM) and increasing amounts of either RapZ or RapZ<sub>quad</sub> were added to the reaction (Fig. 6.1). We reasoned that if binding and structural re-organization of GlmZ were important for cleavage by RNase E, the RapZ quadruple mutant, that is unable to bind GlmY and GlmZ (Figs. 4.3, S4.12; Göpel *et al.*, 2013), should not be able to induce cleavage. In contrast, when contemplating the second scenario that RapZ may act as an allosteric activator, the inability to bind GlmZ should not drastically impair the ability to induce processing *in vitro*.



**Figure 6.1: The RNA binding function of RapZ is not strictly required to induce cleavage of GlmZ by RNase E *in vitro*.** RNase E *in vitro* cleavage assay assessing the stability of  $\alpha$ - $^{32}$ P-UTP-labeled GlmZ upon incubation with 10 nM RNase E NTD (amino acids 1-529) and increasing amounts of RapZ or RapZ<sub>quad</sub>. Final protein concentrations are indicated above. RNase E cleavage assays were performed as described previously (Göpel *et al.*, 2013). Asterisks denote unspecific cleavage products, solid lines indicate the combination of two independent experiments.

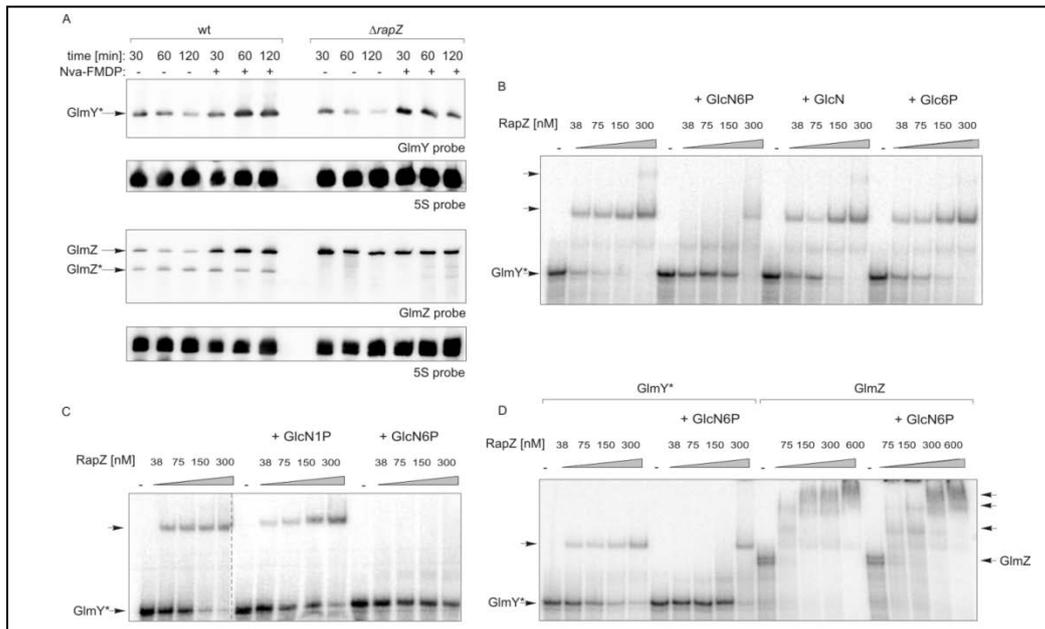
Interestingly, RapZ<sub>quad</sub> was able to promote cleavage of GlmZ, even though higher protein concentrations were needed for complete cleavage as compared to wild type RapZ (Fig. 6.1). Further, the first specific processing products were obtained at a concentration of 2.3 nM RapZ, whereas an ~16-fold higher protein concentration, 37.5 nM, was required when RapZ<sub>quad</sub> was used. At the highest concentration of 150 nM RapZ or RapZ<sub>quad</sub>, respectively, GlmZ was completely processed (Fig. 6.1). Note that standard RNase E cleavage assays are performed using increasing amounts of RNase E and 150 nM RapZ (Göpel *et al.*, 2013). Under these conditions RapZ and RapZ<sub>quad</sub> induced cleavage of GlmZ to the same extent (data not shown). Hence, at least *in vitro* the RNA-binding function of RapZ is not required for processing of GlmZ by RNase E. However, this activity may well be required *in vivo* to deliver GlmZ to RNase E, which is located at the membrane (Liou *et al.*, 2001; Miczak *et al.*, 1991). This is supported by the fact that RapZ<sub>quad</sub> fails to complement a *rapZ* deletion *in vivo* (Fig. 4.3; Göpel *et al.*, 2013).

In sum, our data indicate that RapZ neither acts by converting the 5' tri-phosphate of GlmZ to a mono-phosphate (Fig. 5.8), nor that the RNA binding function is required for induction of cleavage *in vitro* (Fig. 6.1). Hence, it seems plausible that RapZ might rather act as an allosteric activator directly stimulating the enzymatic activity of RNase E. However, further experiments are needed to clarify the molecular mechanism by which RapZ promotes cleavage by RNase E.

#### **Which factor serves as cellular GlcN6P sensor?**

So far, it is not understood how the GlcN6P concentration is sensed by the GlmYZ system. Taking into account that GlmY is required for sensing of GlcN6P limitation and accumulates in response (Reichenbach *et al.*, 2008), the GlcN6P sensor might either be a factor governing GlmY abundance or it may be GlmY itself. A previous study could exclude an effect on the transcription rate of GlmY due to GlcN6P starvation, thus, accumulation of GlmY occurs by a post-transcriptional mechanism (Reichenbach *et al.*, 2009). Of all factors currently known to be involved in feedback-regulation of GlmS, only RapZ and GlmY itself remain as putative candidates for the GlcN6P sensor. It is possible that GlcN6P could interact with RapZ, which might alter the activity or sRNA binding preferences of the protein. Hence, the binding affinity towards GlmY may be increased upon GlcN6P limitation and GlmY might subsequently be protected by RapZ resulting in the observed accumulation of GlmY under these conditions (Reichenbach *et al.*, 2008; Fig. 6.1 A). In the latter case, binding of GlcN6P to GlmY might result in conformational changes of the sRNA and/or destabilization of GlmY. To understand how this cascade is initiated, it is crucial to identify the GlcN6P sensor. Preliminary results indicate that GlcN6P sensing still occurs in cells lacking *rapZ* (Fig. 6.2 A) and that GlcN6P impairs

binding of GlmY to RapZ *in vitro* (Fig.6.2 B-D). Taken together, our data indicate, that GlmY itself might bind to GlcN6P.



**Figure 6.2: GlmY itself might act as GlcN6P sensor in the cell.** **A.** Northern blot addressing the amounts of GlmY and GlmZ in absence and presence of Nva-FMDP in wild type (R1279) and  $\Delta rapZ$  (Z37) cells. Cells were grown in LB at 37°C, 200 r.p.m. to an  $OD_{600} \sim 0.3$  and split. Subsequently 100  $\mu\text{g}/\text{ml}$  Nva-FMDP or the equal amount of  $\text{ddH}_2\text{O}$  was added to the culture and growth was continued. Samples were collected after the indicated time, total RNA was isolated and analyzed by northern blot using DIG-labeled RNA probes as indicated. Re-probing against 5S rRNA served as loading control. **B.** EMSA using  $\alpha\text{-}^{32}\text{P}$ -UTP labeled processed GlmY\* with increasing concentrations of purified Strep-RapZ. The final protein concentrations are as indicated above. After heat denaturation and prior to incubation with Strep-RapZ, 15 mM of the indicated metabolite at pH 7.0 was added. **C.** EMSA using  $\alpha\text{-}^{32}\text{P}$ -UTP labeled processed GlmY\* with increasing concentrations of purified Strep-RapZ, final concentrations are given above. Prior to incubation with Strep-RapZ, 10 mM GlcN1P or GlcN6P at pH 7.0 was added. **D.** EMSA using  $\alpha\text{-}^{32}\text{P}$ -UTP labeled processed GlmY\* and full-length GlmZ with increasing concentrations of purified Strep-RapZ. The final protein concentrations are as indicated above. Prior to incubation with Strep-RapZ, 15 mM GlcN6P was added to samples containing GlmY\* and GlmZ. In untreated samples, 1x structure buffer was used. Asterisks denote the processed form of the respective sRNA; free unbound sRNAs are indicated and arrows denote protein/sRNA complexes.

Even though GlmZ is completely stable in its active form in cells deficient of *rapZ*, GlmY still accumulates upon induction of GlcN6P starvation using Nva-FMDP (Fig. 6.2 A). This compound acts as a glutamine analogue by covalently binding to the active site of the GlmS enzyme, thus inhibiting its activity. As previously shown, administration of Nva-FMDP resulted in GlcN6P depletion and subsequent up-regulation of the GlmYZ-cascade (Kalamorz *et al.*, 2007; Reichenbach *et al.*, 2008). Here we show, that this activation still occurs in absence of RapZ (Fig. 6.2 A). Thus, it is unlikely that RapZ is the GlcN6P sensor.

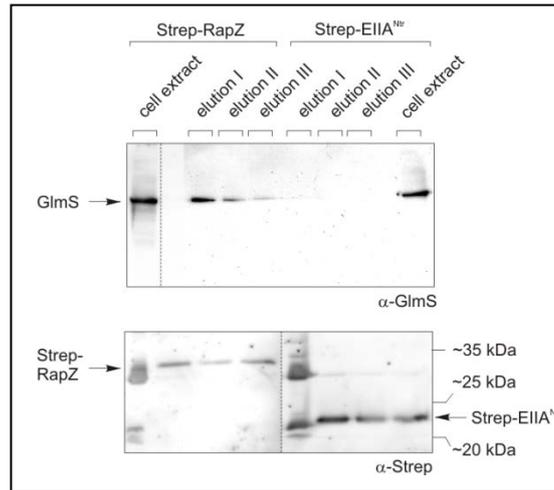
Further, addition of GlcN6P inhibited binding of GlmY to RapZ, whereas GlmZ was still bound by RapZ under the same conditions (Fig. 6.2 D). However, rather high concentrations of 10 mM GlcN6P were required to observe this effect *in vitro*. These concentrations exceed the GlcN6P concentration *in vivo* by  $\sim 10$ -fold (Bennett *et al.*, 2009). Interestingly, chemically similar compounds like glucosamine, glucose-6-phosphat and even glucosamine-1-phosphat applied in the same concentrations did not

significantly impair binding of GlmY to RapZ (Fig. 6.2 B, D). These preliminary results suggest that while GlmY is able to counteract GlmZ processing by sequestration of RapZ, GlmY itself might be subject to regulation by GlcN6P. Upon binding of GlcN6P to GlmY, the sRNA may become binding- incompetent for RapZ. This proposed mechanism is, thus, reminiscent of the regulation of *glmS* expression by the GlmS ribozyme/riboswitch in *Bacilli*. Here, GlcN6P directly binds to the ribozyme and acts as a co-factor in self-induced cleavage of the *glmS*-transcript (Collins *et al.*, 2007; Winkler *et al.*, 2004). It is tempting to speculate, that GlcN6P might influence GlmY in a similar fashion. Although GlmY might not be a ribozyme, its stability may be altered by GlcN6P, thus resulting in lower levels of GlmY under conditions of ample GlcN6P. Whether interaction with GlcN6P indeed induces structural changes and/or reduces GlmY stability *in vivo* is yet to be determined. First, filter-binding assays employing *in vitro* transcribed GlmY and radio-labeled GlcN6P could be conducted to confirm direct binding of GlcN6P to GlmY. Further, extended structure probing experiments with GlmY pre-incubated with increasing concentrations of GlcN6P might hint towards structural changes within GlmY upon interaction with GlcN6P. On the other hand, analysis of the half-life of GlmY under conditions of ample GlcN6P supply and GlcN6P starvation might indicate whether the interaction with GlcN6P destabilizes GlmY *in vivo*. In this case, the observed accumulation of GlmY may be indirect and might be attributed to stabilization of GlmY through interaction with RapZ.

### **On yet another role of RapZ**

In this work, we demonstrated that RapZ fulfills multiple roles within the bacterial cell: next to acting as an RNase E adaptor protein mediating GlcN6P homeostasis (Göpel *et al.*, 2013), RapZ acts as a modulator stimulating *glmY* transcription from its  $\sigma^{54}$ -dependent promoter (see chapter 3). Whether this stimulation of promoter activity requires activation of the specific response regulator GlrR or RpoN ( $\sigma^{54}$ ) itself or a so far unknown factor remains elusive. Interestingly, in a bacterial two-hybrid approach we found hints that RapZ may also interact with enzyme GlmS (Lüttmann, 2011). Further we performed co-purification assays assessing a possible interaction of Strep-RapZ with endogenous GlmS. To ensure synthesis of GlmS, we expressed Strep-RapZ from the chromosomal attachment site of phage  $\lambda$ , the *attB* site. This is required as over-production of RapZ from a plasmid strongly increases degradation of GlmZ and thus diminishes synthesis of the GlmS protein (Kalamorz *et al.*, 2007; Göpel *et al.*, 2013; Fig. 3.5 C). This would then impair the co-purification assay and lead to inconclusive results.

The co-purification experiments with chromosomally encoded Strep-RapZ revealed that GlmS specifically co-purifies with Strep-RapZ, but not with chromosomally encoded Strep-EIIA<sup>Ntr</sup> that was used as a control (Fig. 6.3).



**Figure 6.3: GlmS co-purifies with chromosomally encoded Strep-RapZ, but not with Strep-EIIA<sup>Ntr</sup>.** Western blot addressing the possible interaction of Strep-RapZ and Strep-EIIA<sup>Ntr</sup> expressed from the ectopic  $\lambda attB$  site of the *E. coli* chromosome with endogenous GlmS. Strains Z479 ( $\Delta rapZ$ ,  $P_{ara}::strep-rapZ-bla$  on chromosome) and Z703 ( $\Delta ptsN$ ,  $P_{ara}::strep-ptsN-bla$  on chromosome) were grown in 500 ml LB-Amp<sup>50</sup> and expression of strep-proteins was induced by addition of 0.2% L-arabinose. Growth was continued for 1-2 hours, cells were harvested by centrifugation and pellets stored by  $-20^{\circ}\text{C}$  until further use. Subsequently, cells were disrupted using a French pressure cell and purification of strep-proteins was performed as described before with few exceptions (Lüttmann *et al.*, 2012; Göpel *et al.*, 2013). A matrix volume of 250  $\mu\text{l}$  was used to bind strep-proteins and the proteins were later eluted from the Strep-Tactin columns with 200  $\mu\text{l}$  buffer E. After purification 10% (v/v) glycerin was added to the elution fractions, samples were supplied with laemmli loading buffer and analyzed by western blot. The western blots were detected using polyclonal  $\alpha$ -GlmS antibodies for investigation of a possible interaction and detection with  $\alpha$ -Strep antibodies was used as loading and purification control.

Chromosomally expressed Strep-RapZ and Strep-EIIA<sup>Ntr</sup> were successfully purified by Strep-Tactin affinity purification as detected by western blot using polyclonal antibodies directed against the Strep-epitope (Fig. 6.3 lower panel). Furthermore, when the samples were detected using an antibody directed against the GlmS protein, signals corresponding to GlmS were visible in elution fractions I to III obtained from Strep-RapZ purification (Fig. 6.3 upper panel). In contrast, though GlmS was present in the cell extract of strains carrying *strep-rapZ* as well as *strep-ptsN* on chromosome, no signal corresponding to GlmS was detectable in samples derived from purification of Strep-EIIA<sup>Ntr</sup> (Strep-PtsN; Fig. 6.3 upper panel).

Thus, co-purification experiments using physiological amounts of Strep-RapZ support the notion that RapZ and GlmS might interact within the cell. What is the purpose of this interaction? While RapZ is not essential for sensing of the GlcN6P concentration, it might be possible that an interaction between RapZ and GlmS contributes to the control of GlcN6P homeostasis. For instance, interaction with RapZ might alter the enzymatic activity of GlmS, either enhancing synthesis of GlcN6P or shutting off GlmS activity. Thus, this may be a layer of regulation that targets the pre-existing GlmS protein. In a different scenario, interaction between GlmS and RapZ might alter the sRNA binding affinity of RapZ, hence, providing a quick response to GlcN6P starvation. Whether the proposed interaction with GlmS requires binding of either GlmY or GlmZ is yet to be determined. Further experiments assessing the enzymatic activity of GlmS may provide an answer whether an interaction

with RapZ might indeed alter GlnS activity. If this should actually be the case, enzyme assays using the various RapZ mutants may provide answers to the involvement of GlnY and GlnZ, the impact of acetylation and the possible requirement of the ATPase/GTPase activity of RapZ for this process.

#### **6.4 Conclusion and perspectives**

In this work, we unraveled the molecular mechanism of signal transduction between homologous sRNAs GlnY and GlnZ as a titration mechanism involving the unique RNA-binding protein RapZ. Furthermore, we identified RapZ as an RNase E adaptor protein required to promote efficient and specific degradation of sRNA GlnZ and proposed an anti-adaptor function for GlnY. We further created GlnYZ hybrid sRNAs that provide a suitable tool for the molecular dissection of specificity determinants for binding by RNA chaperon Hfq and recognition by RNase E and RapZ. In addition, the work presented here, demonstrated the existence of three distinct promoter architectures of  $\sigma^{70}$ - and  $\sigma^{54}$ -dependent promoters driving the expression of *glnY* and *glnZ* in *Enterobacteriaceae*. Yet, various open questions still remain.

For instance, we could show that RapZ switches its RNA binding partners depending on the cellular GlcN6P concentration. However, the mechanism by which GlcN6P is sensed within this system remains elusive. Data accumulates indicating that GlnY itself might bind GlcN6P, hence acting as the sensor. *In vitro* experiments suggest that interaction with GlcN6P renders GlnY incapable for binding of RapZ. Thus, reminiscent of the *glnS* ribozyme in *Bacilli*, an RNA molecule might also be responsible for sensing and control of the GlcN6P concentration in Gram negative bacteria. Further experiments are needed to evaluate whether binding of GlcN6P alters the structure of GlnY and/or destabilizes the sRNA *in vivo*. Thus far, a non-coding RNA directly involved in sensing of a metabolite is unprecedented and would greatly expand the regulatory repertoire of small RNAs.

Further, the exact molecular mechanism by which RapZ activates RNase E for cleavage of GlnZ is so far unclear. Preliminary data point to the possibility that RapZ may act as an allosteric activator of RNase E catalytic activity. To gain insight into the mechanisms at work, a mutational analysis could be conducted to isolate mutants of RapZ that are impaired in their ability to promote processing of GlnZ, while RNA-binding might still occur. Furthermore, an attempt has been made to solve the crystal structure of RapZ in collaboration with the group of Ralf Ficner in Göttingen. So far, RapZ was crystallized and shown to form homo-trimeric complexes; however, the structure of RapZ still remains to be solved. As RapZ does not share significant similarities with other proteins whose structures are available, this would be the first structure solved for a p-loop ATPase protein of this class. The structure of a co-crystal of the N-terminal catalytic domain and RapZ could provide

valuable insight into the domains and specific residues involved in activation of the catalytic activity of RNase E.

In sum, ancillary proteins guiding ribonuclease activity and substrate specificity might be more common than previously thought. The identification of novel adaptor proteins and investigation of their mode of action could help to gain insight into the puzzling paradigm of how substrate specificity can be achieved for general RNA processing enzymes that need to recognize a multitude of structurally diverse substrate RNAs.

---

## 7. LIST OF REFERENCES

- Ades, S.E., Grigorova, I.L., and Gross, C.A.** (2003) Regulation of the alternative sigma factor sigma(E) during initiation, adaptation, and shutoff of the extracytoplasmic heat shock response in *Escherichia coli*. *J Bacteriol* **185**: 2512-2519.
- Aiba, H.** (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr. Opin. Microbiol.* **10**: 134-139.
- Ait-Bara, S., and Carpousis, A.J.** (2010) Characterization of the RNA degradosome of *Pseudoalteromonas haloplanktis*: conservation of the RNase E-RhlB interaction in the gammaproteobacteria. *J Bacteriol* **192**: 5413-5423.
- Albaugh, B.N., Arnold, K.M., and Denu, J.M.** (2011) KAT(ching) metabolism by the tail: insight into the links between lysine acetyltransferases and metabolism. *ChemBiochem* **12**: 290-298.
- Andrade, J.M., Pobre, V., Matos, A.M., and Arraiano, C.M.** (2012) The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA* **18**: 844-855.
- Angerer, A., Enz, S., Ochs, M., and Braun, V.** (1995) Transcriptional regulation of ferric citrate transport in *Escherichia coli* K-12. Fecl belongs to a new subfamily of sigma 70-type factors that respond to extracytoplasmic stimuli. *Mol Microbiol* **18**: 163-174.
- Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H., and Altuvia, S.** (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.* **11**: 941-950.
- Arraiano, C.M., Andrade, J.M., Domingues, S., Guinote, I.B., Malecki, M., Matos, R.G., Moreira, R.N., Pobre, V., Reis, F.P., Saramago, M., Silva, I.J., and Viegas, S.C.** (2010) The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol. Rev.* **34**: 883-923.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H.** (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**: 1-11.
- Babitzke, P., and Romeo, T.** (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr. Opin. Microbiol.* **10**: 156-163.
- Baker, K.E., and Mackie, G.A.** (2003) Ectopic RNase E sites promote bypass of 5'-end-dependent mRNA decay in *Escherichia coli*. *Mol Microbiol* **47**: 75-88.
- Bandyra, K.J., Said, N., Pfeiffer, V., Gorna, M.W., Vogel, J., and Luisi, B.F.** (2012) The Seed Region of a Small RNA Drives the Controlled Destruction of the Target mRNA by the Endoribonuclease RNase E. *Mol. Cell* **47**: 943-953.
- Bandyra, K.J., Bouvier, M., Carpousis, A.J., and Luisi, B.F.** (2013) The social fabric of the RNA degradosome. *Biochim Biophys Acta* **1829**: 514-522.

- Bardey, V., Vallet, C., Robas, N., Charpentier, B., Thouvenot, B., Mougín, A., Hajnsdorf, E., Regnier, P., Springer, M., and Branlant, C.** (2005) Characterization of the molecular mechanisms involved in the differential production of erythrose-4-phosphate dehydrogenase, 3-phosphoglycerate kinase and class II fructose-1,6-bisphosphate aldolase in *Escherichia coli*. *Mol. Microbiol.* **57**: 1265-1287.
- Bardwell, J.C., Regnier, P., Chen, S.M., Nakamura, Y., Grunberg-Manago, M., and Court, D.L.** (1989) Autoregulation of RNase III operon by mRNA processing. *Embo J* **8**: 3401-3407.
- Battesti, A., Majdalani, N., and Gottesman, S.** (2011) The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* **65**: 189-213.
- Beck, L.L., Smith, T.G., and Hoover, T.R.** (2007) Look, no hands! Unconventional transcriptional activators in bacteria. *Trends Microbiol.* **15**: 530-537.
- Becker, G., Klauck, E., and Hengge-Aronis, R.** (1999) Regulation of RpoS proteolysis in *Escherichia coli*: the response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. *Proc Natl Acad Sci U S A* **96**: 6439-6444.
- Beckmann, B.M., Burenina, O.Y., Hoch, P.G., Kubareva, E.A., Sharma, C.M., and Hartmann, R.K.** (2011) *In vivo* and *in vitro* analysis of 6S RNA-templated short transcripts in *Bacillus subtilis*. *RNA Biol* **8**: 839-849.
- Beisel, C.L., and Storz, G.** (2010) Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol. Rev.* **34**: 866-882.
- Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., and Rabinowitz, J.D.** (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**: 593-599.
- Berezikov, E., Thuemmler, F., van Laake, L.W., Kondova, I., Bontrop, R., Cuppen, E., and Plasterk, R.H.** (2006) Diversity of microRNAs in human and chimpanzee brain. *Nat. Genet.* **38**: 1375-1377.
- Bernstein, J.A., Lin, P.H., Cohen, S.N., and Lin-Chao, S.** (2004) Global analysis of *Escherichia coli* RNA degradosome function using DNA microarrays. *Proc Natl Acad Sci U S A* **101**: 2758-2763.
- Bi, W., and Stambrook, P.J.** (1998) Site-directed mutagenesis by combined chain reaction. *Anal. Biochem.* **256**: 137-140.
- Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y.** (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1474.
- Bordi, C., Lamy, M.C., Ventre, I., Termine, E., Hachani, A., Fillet, S., Roche, B., Bleves, S., Mejean, V., Lazdunski, A., and Filloux, A.** (2010) Regulatory RNAs and the HptB/RetS signalling pathways fine-tune *Pseudomonas aeruginosa* pathogenesis. *Mol. Microbiol* **76**: 1427-1443.

- Bougdour, A., Cunning, C., Baptiste, P.J., Elliott, T., and Gottesman, S.** (2008) Multiple pathways for regulation of sigmaS (RpoS) stability in *Escherichia coli* via the action of multiple anti-adaptors. *Mol. Microbiol.* **68**: 298-313.
- Bouvier, M., and Carpousis, A.J.** (2011) A tale of two mRNA degradation pathways mediated by RNase E. *Mol. Microbiol.* **82**: 1305-1310.
- Brantl, S.** (2007) Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr Opin Microbiol* **10**: 102-109.
- Brantl, S.** (2009) Bacterial chromosome-encoded small regulatory RNAs. *Future Microbiol.* **4**: 85-103.
- Bury-Moné, S., Nomane, Y., Reymond, N., Barbet, R., Jacquet, E., Imbeaud, S., Jacq, A., and Bouloc, P.** (2009) Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet* **5**: e1000651.
- Callaghan, A.J., Grossmann, J.G., Redko, Y.U., Ilag, L.L., Moncrieffe, M.C., Symmons, M.F., Robinson, C.V., McDowall, K.J., and Luisi, B.F.** (2003) Quaternary structure and catalytic activity of the *Escherichia coli* ribonuclease E amino-terminal catalytic domain. *Biochemistry* **42**: 13848-13855.
- Callaghan, A.J., Marcaida, M.J., Stead, J.A., McDowall, K.J., Scott, W.G., and Luisi, B.F.** (2005) Structure of *Escherichia coli* RNase E catalytic domain and implications for RNA turnover. *Nature* **437**: 1187-1191.
- Caramel, A., and Schnetz, K.** (1998) Lac and lambda repressors relieve silencing of the *Escherichia coli* *bgl* promoter. Activation by alteration of a repressing nucleoprotein complex. *J. Mol. Biol.* **284**: 875-883.
- Caron, M.P., Lafontaine, D.A., and Masse, E.** (2010) Small RNA-mediated regulation at the level of transcript stability. *RNA Biol.* **7**: 140-144.
- Caron, M.P., Bastet, L., Lussier, A., Simoneau-Roy, M., Massé, E., and Lafontaine, D.A.** (2012) Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc Natl Acad Sci U S A* **109**: E3444-3453.
- Carpousis, A.J., Van Houwe, G., Ehretsmann, C., and Krisch, H.M.** (1994) Copurification of *E. coli* RNase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* **76**: 889-900.
- Carpousis, A.J.** (2007) The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* **61**: 71-87.
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J.** (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *Embo J.* **31**: 4005-4019.

- Collins, J.A., Irnov, I., Baker, S., and Winkler, W.C.** (2007) Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev.* **21**: 3356-3368.
- Condon, C.** (2007) Maturation and degradation of RNA in bacteria. *Curr. Opin. Microbiol.* **10**: 271-278.
- Corcoran, C.P., Podkaminski, D., Papenfort, K., Urban, J.H., Hinton, J.C., and Vogel, J.** (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol. Microbiol.* **84**: 428-445.
- Darfeuille, F., Unoson, C., Vogel, J., and Wagner, E.G.** (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol. Cell.* **26**: 381-392.
- Datsenko, K.A., and Wanner, B.L.** (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U S A* **97**: 6640-6645.
- Davis, B.M., and Waldor, M.K.** (2007) RNase E-dependent processing stabilizes MicX, a *Vibrio cholerae* sRNA. *Mol. Microbiol.* **65**: 373-385.
- De Lay, N., Schu, D.J., and Gottesman, S.** (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J Biol Chem* **288**: 7996-8003.
- Deana, A., Celesnik, H., and Belasco, J.G.** (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**: 355-358.
- Dehal, P.S., Joachimiak, M.P., Price, M.N., Bates, J.T., Baumohl, J.K., Chivian, D., Friedland, G.D., Huang, K.H., Keller, K., Novichkov, P.S., Dubchak, I.L., Alm, E.J., and Arkin, A.P.** (2009) MicrobesOnline: an integrated portal for comparative and functional genomics. *Nucleic Acids Res.* **38**: D396-400.
- Del Favero, M., Mazzantini, E., Briani, F., Zangrossi, S., Tortora, P., and Deho, G.** (2008) Regulation of *Escherichia coli* polynucleotide phosphorylase by ATP. *J Biol Chem* **283**: 27355-27359.
- Desnoyers, G., and Massé, E.** (2011) Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes Dev* **26**: 726-739.
- Diederich, L., Rasmussen, L.J., and Messer, W.** (1992) New cloning vectors for integration in the *lambda* attachment site *attB* of the *Escherichia coli* chromosome. *Plasmid* **28**: 14-24.
- Diwa, A., Bricker, A.L., Jain, C., and Belasco, J.G.** (2000) An evolutionarily conserved RNA stem-loop functions as a sensor that directs feedback regulation of RNase E gene expression. *Genes Dev.* **14**: 1249-1260.
- Dong, T., Yu, R., and Schellhorn, H.** (2011) Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. *Mol Microbiol* **79**: 375-386.
- Dubey, A.K., Baker, C.S., Romeo, T., and Babitzke, P.** (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *Rna* **11**: 1579-1587.

- Durand, P., Golinelli-Pimpaneau, B., Moulleron, S., Badet, B., and Badet-Denisot, M.A.** (2008) Highlights of glucosamine-6P synthase catalysis. *Arch. Biochem. Biophys.* **474**: 302-317.
- Efstratiadis, A., Vournakis, J.N., Donis-Keller, H., Chaconas, G., Dougall, D.K., and Kafatos, F.C.** (1977) End labeling of enzymatically decapped mRNA. *Nucleic Acids Res* **4**: 4165-4174.
- Figueroa-Bossi, N., Valentini, M., Malleret, L., and Bossi, L.** (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev.* **23**: 2004-2015.
- Finley, D., Ulrich, H.D., Sommer, T., and Kaiser, P.** (2012) The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. *Genetics* **192**: 319-360.
- Flamez, C., Ricard, I., Arafah, S., Simonet, M., and Marceau, M.** (2008) Phenotypic analysis of *Yersinia pseudotuberculosis* 32777 response regulator mutants: new insights into two-component system regulon plasticity in bacteria. *Int. J. Med. Microbiol.* **298**: 193-207.
- Fröhlich, K.S., and Vogel, J.** (2009) Activation of gene expression by small RNA. *Curr. Opin. Microbiol.* **12**: 674-682.
- Gaal, T., Ross, W., Estrem, S.T., Nguyen, L.H., Burgess, R.R., and Gourse, R.L.** (2001) Promoter recognition and discrimination by EsigmaS RNA polymerase. *Mol Microbiol* **42**: 939-954.
- Gao, J., Lee, K., Zhao, M., Qiu, J., Zhan, X., Saxena, A., Moore, C.J., Cohen, S.N., and Georgiou, G.** (2006) Differential modulation of *E. coli* mRNA abundance by inhibitory proteins that alter the composition of the degradosome. *Mol. Microbiol.* **61**: 394-406.
- Georg, J., and Hess, W.R.** (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. *Microbiol. Mol. Biol. Rev.* **75**: 286-300.
- Gogol, E.B., Rhodius, V.A., Papenfort, K., Vogel, J., and Gross, C.A.** (2011) Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc. Natl. Acad. Sci. U S A* **108**: 12875-12880.
- Goldblum, K., and Apririon, D.** (1981) Inactivation of the ribonucleic acid-processing enzyme ribonuclease E blocks cell division. *J. Bacteriol.* **146**: 128-132.
- Göpel, Y., Lüttmann, D., Heroven, A.K., Reichenbach, B., Dersch, P., and Görke, B.** (2011) Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in *Enterobacteriaceae*. *Nucleic Acids Res.* **39**: 1294-1309.
- Göpel, Y., and Görke, B.** (2012a) Rewiring two-component signal transduction with small RNAs. *Curr. Opin. Microbiol.* **15**: 132-139.
- Göpel, Y., and Görke, B.** (2012b) Zusammenspiel von Zweikomponentensystemen und kleinen RNAs. *BIOspektrum* **07.12**: 702-705.

- Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J., and Görke, B.** (2013) Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev.* **27**: 552-564.
- Görke, B., and Rak, B.** (1999) Catabolite control of *Escherichia coli* regulatory protein BglG activity by antagonistically acting phosphorylations. *Embo J.* **18**: 3370-3379.
- Görke, B., Foulquier, E., and Galinier, A.** (2005) YvcK of *Bacillus subtilis* is required for a normal cell shape and for growth on Krebs cycle intermediates and substrates of the pentose phosphate pathway. *Microbiology* **151**: 3777-3791.
- Görke, B., and Stülke, J.** (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**: 613-624.
- Görke, B., and Vogel, J.** (2008) Noncoding RNA control of the making and breaking of sugars. *Genes Dev.* **22**: 2914-2925.
- Górna, M.W., Pietras, Z., Tsai, Y.C., Callaghan, A.J., Hernandez, H., Robinson, C.V., and Luisi, B.F.** (2010) The regulatory protein RraA modulates RNA-binding and helicase activities of the *E. coli* RNA degradosome. *RNA* **16**: 553-562.
- Górna, M.W., Carpousis, A.J., and Luisi, B.F.** (2012) From conformational chaos to robust regulation: the structure and function of the multi-enzyme RNA degradosome. *Q. Rev. Biophys.* **45**: 105-145.
- Gottesman, S., and Storz, G.** (2011) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb. Perspect. Biol.* **3**: a003798.
- Gruber, T.M., and Gross, C.A.** (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**: 441-466.
- Guillier, M., and Gottesman, S.** (2006) Remodelling of the *Escherichia coli* outer membrane by two small regulatory RNAs. *Mol. Microbiol.* **59**: 231-247.
- Guillier, M., Gottesman, S., and Storz, G.** (2006) Modulating the outer membrane with small RNAs. *Genes Dev.* **20**: 2338-2348.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J.** (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose P<sub>BAD</sub> promoter. *J. Bacteriol.* **177**: 4121-4130.
- Hames, C., Halbedel, S., Schilling, O., and Stülke, J.** (2005) Multiple-mutation reaction: a method for simultaneous introduction of multiple mutations into the *glpK* gene of *Mycoplasma pneumoniae*. *Appl. Environ. Microbiol.* **71**: 4097-4100.
- Helmann, J.D., and Chamberlin, M.J.** (1988) Structure and function of bacterial sigma factors. *Annu Rev Biochem* **57**: 839-872.

- Henderson, C.A., Vincent, H.A., Casamento, A., Stone, C.M., Phillips, J.O., Cary, P.D., Sobott, F., Gowers, D.M., Taylor, J.E., and Callaghan, A.J.** (2013) Hfq binding changes the structure of *Escherichia coli* small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of *rpoS*. *RNA* **19**: 1089-1104.
- Hengge, R.** (2009) Proteolysis of sigmaS (RpoS) and the general stress response in *Escherichia coli*. *Res Microbiol* **160**: 667-676.
- Heroven, A.K., Böhme, K., Rohde, M., and Dersch, P.** (2008) A Csr-type regulatory system, including small non-coding RNAs, regulates the global virulence regulator RovA of *Yersinia pseudotuberculosis* through RovM. *Mol. Microbiol.* **68**: 1179-1195.
- Heroven, A.K., Böhme, K., and Dersch, P.** (2012) The Csr/Rsm system of *Yersinia* and related pathogens: a post-transcriptional strategy for managing virulence. *RNA Biol* **9**: 379-391.
- Hobbs, E.C., Astarita, J.L., and Storz, G.** (2009) Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in *Escherichia coli*: analysis of a bar-coded mutant collection. *J Bacteriol* **192**: 59-67.
- Holmqvist, E., Unoson, C., Reimegard, J., and Wagner, E.G.** (2012) A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol. Microbiol.* **84**: 414-427.
- Hoover, T.R., Santero, E., Porter, S., and Kustu, S.** (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell* **63**: 11-22.
- Hsieh, Y.J., and Wanner, B.L.** (2010) Global regulation by the seven-component P<sub>i</sub> signaling system. *Curr. Opin. Microbiol.* **13**: 198-203.
- Hu, L.I., Lima, B.P., and Wolfe, A.J.** (2010) Bacterial protein acetylation: the dawning of a new age. *Mol Microbiol* **77**: 15-21.
- Hu, L.I., Chi, B.K., Kuhn, M.L., Filippova, E.V., Walker-Peddakotla, A.J., Basell, K., Becher, D., Anderson, W.F., Antelmann, H., and Wolfe, A.J.** (2013) Acetylation of the Response Regulator RcsB Controls Transcription from a Small RNA Promoter. *J Bacteriol* doi:10.1128/JB.00383-13.
- Ikeda, Y., Yagi, M., Morita, T., and Aiba, H.** (2011) Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli*. *Mol. Microbiol.* **79**: 419-432.
- Ishikawa, H., Otaka, H., Maki, K., Morita, T., and Aiba, H.** (2012) The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. *RNA* **18**: 1062-1074.
- Jain, C.** (2002) Degradation of mRNA in *Escherichia coli*. *IUBMB Life* **54**: 315-321.

- Jishage, M., Iwata, A., Ueda, S., and Ishihama, A.** (1996) Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol* **178**: 5447-5451.
- Jishage, M., and Ishihama, A.** (1998) A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc Natl Acad Sci U S A* **95**: 4953-4958.
- Joanny, G., Derout, J., Brechemier-Baey, D., Labas, V., Vinh, J., Regnier, P., and Hajnsdorf, E.** (2007) Polyadenylation of a functional mRNA controls gene expression in *Escherichia coli*. *Nucleic Acids Res.* **35**: 2494-2502.
- Joseph, P., Fantino, J.R., Herbaud, M.L., and Denizot, F.** (2001) Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis*. *FEMS Microbiol Lett* **205**: 91-97.
- Jung, K., Fried, L., Behr, S., and Heermann, R.** (2012) Histidine kinases and response regulators in networks. *Curr. Opin. Microbiol.* **15**: 118-124.
- Kaberdin, V.R., and Lin-Chao, S.** (2009) Unraveling new roles for minor components of the *E. coli* RNA degradosome. *RNA Biol* **6**: 402-405.
- Kalamorz, F., Reichenbach, B., März, W., Rak, B., and Görke, B.** (2007) Feedback control of glucosamine-6-phosphate synthase GlmS expression depends on the small RNA GlmZ and involves the novel protein YhbJ in *Escherichia coli*. *Mol. Microbiol.* **65**: 1518-1533.
- Kalamorz, F.** (2008) Regulation der Synthese der Glukosamin-6-Phosphat Synthase GlmS in *Escherichia coli* durch das neuartige Protein YhbJ. In *ph.D. thesis* Göttingen: Georg-August-Universität Göttingen.
- Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y.** (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**: D480-484.
- Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D.** (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc. Natl. Acad. Sci. U S A* **95**: 5752-5756.
- Karimova, G., Ullmann, A., and Ladant, D.** (2001) Protein-protein interaction between *Bacillus stearothermophilus* tyrosyl-tRNA synthetase subdomains revealed by a bacterial two-hybrid system. *J. Mol. Microbiol. Biotechnol.* **3**: 73-82.
- Karimova, G., and Ladant, D.** (2005) A bacterial two-hybrid system based on a Cyclic AMP signaling cascade. In *Protein-Protein interactions, A Molecular Cloning Manual*. Golemis, E. (ed). Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp. 499-515.
- Kawamoto, H., Morita, T., Shimizu, A., Inada, T., and Aiba, H.** (2005) Implication of membrane localization of target mRNA in the action of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in *Escherichia coli*. *Genes Dev* **19**: 328-338.

- Kawamoto, H., Koide, Y., Morita, T., and Aiba, H.** (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Mol. Microbiol.* **61**: 1013-1022.
- Keyhani, N.O., Bacia, K., and Roseman, S.** (2000a) The transport/phosphorylation of N,N'-diacetylchitobiose in *Escherichia coli*. Characterization of phospho-IIB(Chb) and of a potential transition state analogue in the phosphotransfer reaction between the proteins IIA(Chb) AND IIB(Chb). *J. Biol. Chem.* **275**: 33102-33109.
- Keyhani, N.O., Wang, L.X., Lee, Y.C., and Roseman, S.** (2000b) The chitin disaccharide, N,N'-diacetylchitobiose, is catabolized by *Escherichia coli* and is transported/phosphorylated by the phosphoenolpyruvate:glycose phosphotransferase system. *J. Biol. Chem.* **275**: 33084-33090.
- Kim, K.S., Sim, S., Ko, J.H., Cho, B., and Lee, Y.** (2004) Kinetic analysis of precursor M1 RNA molecules for exploring substrate specificity of the N-terminal catalytic half of RNase E. *J Biochem* **136**: 693-699.
- Kim, K.S., Manasherob, R., and Cohen, S.N.** (2008) YmdB: a stress-responsive ribonuclease-binding regulator of *E. coli* RNase III activity. *Genes Dev* **22**: 3497-3508.
- Kime, L., Jourdan, S.S., Stead, J.A., Hidalgo-Sastre, A., and McDowall, K.J.** (2010) Rapid cleavage of RNA by RNase E in the absence of 5' monophosphate stimulation. *Mol. Microbiol.* **76**: 590-604.
- Ladurner, A.G.** (2009) Chromatin places metabolism center stage. *Cell* **138**: 18-20.
- Langbein, I., Bachem, S., and Stülke, J.** (1999) Specific interaction of the RNA-binding domain of the *Bacillus subtilis* transcriptional antiterminator GlcT with its RNA target, RAT. *J. Mol. Biol.* **293**: 795-805.
- Lease, R.A., and Belfort, M.** (2000) A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc Natl Acad Sci U S A* **97**: 9919-9924.
- Lee, E.J., and Groisman, E.A.** (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol. Microbiol.* **76**: 1020-1033.
- Lee, K., Zhan, X., Gao, J., Qiu, J., Feng, Y., Meganathan, R., Cohen, S.N., and Georgiou, G.** (2003) RraA, a protein inhibitor of RNase E activity that globally modulates RNA abundance in *E. coli*. *Cell* **114**: 623-634.
- Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S., and Bassler, B.L.** (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118**: 69-82.

- Leonhartsberger, S., Huber, A., Lottspeich, F., and Böck, A.** (2001) The *hydH/G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system. *J. Mol. Biol.* **307**: 93-105.
- Li, Z., Reimers, S., Pandit, S., and Deutscher, M.P.** (2002) RNA quality control: degradation of defective transfer RNA. *Embo J* **21**: 1132-1138.
- Lima, B.P., Antelmann, H., Gronau, K., Chi, B.K., Becher, D., Brinsmade, S.R., and Wolfe, A.J.** (2011) Involvement of protein acetylation in glucose-induced transcription of a stress-responsive promoter. *Mol Microbiol* **81**: 1190-1204.
- Lima, B.P., Thanh Huyen, T.T., Basell, K., Becher, D., Antelmann, H., and Wolfe, A.J.** (2012) Inhibition of acetyl phosphate-dependent transcription by an acetylable lysine on RNA polymerase. *J Biol Chem* **287**: 32147-32160.
- Lin-Chao, S., Wong, T.T., McDowall, K.J., and Cohen, S.N.** (1994) Effects of nucleotide sequence on the specificity of rne-dependent and RNase E-mediated cleavages of RNA I encoded by the pBR322 plasmid. *J Biol Chem* **269**: 10797-10803.
- Link, T.M., Valentin-Hansen, P., and Brennan, R.G.** (2009) Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A* **106**: 19292-19297.
- Liou, G.G., Jane, W.N., Cohen, S.N., Lin, N.S., and Lin-Chao, S.** (2001) RNA degradosomes exist *in vivo* in *Escherichia coli* as multicomponent complexes associated with the cytoplasmic membrane via the N-terminal region of ribonuclease E. *Proc Natl Acad Sci U S A* **98**: 63-68.
- Liu, M.Y., Gui, G., Wei, B., Preston, J.F., 3rd, Oakford, L., Yuksel, U., Giedroc, D.P., and Romeo, T.** (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J Biol Chem* **272**: 17502-17510.
- Loewen, P.C., Hu, B., Strutinsky, J., and Sparling, R.** (1998) Regulation in the *rpoS* regulon of *Escherichia coli*. *Can J Microbiol* **44**: 707-717.
- Luciano, J., Foulquier, E., Fantino, J.R., Galinier, A., and Pompeo, F.** (2009) Characterization of YvcJ, a conserved P-loop-containing protein, and its implication in competence in *Bacillus subtilis*. *J. Bacteriol.* **191**: 1556-1564.
- Lukat, G.S., McCleary, W.R., Stock, A.M., and Stock, J.B.** (1992) Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. *Proc. Natl. Acad. Sci. U S A* **89**: 718-722.
- Lüttmann, D., Heermann, R., Zimmer, B., Hillmann, A., Rampp, I.S., Jung, K., and Görke, B.** (2009) Stimulation of the potassium sensor KdpD kinase activity by interaction with the phosphotransferase protein IIA<sup>Ntr</sup> in *Escherichia coli*. *Mol. Microbiol.* **72**: 978-994.
- Lüttmann, D.** (2011) Die regulatorischen Funktionen des paralogen Phosphotransferase Systems (PTS<sup>Ntr</sup>) in *Escherichia coli*. Göttingen: Georg-August-University.

- Lüttmann, D., Göpel, Y., and Görke, B.** (2012) The phosphotransferase protein EIIA(Ntr) modulates the phosphate starvation response through interaction with histidine kinase PhoR in *Escherichia coli*. *Mol. Microbiol.* **86**: 96-110.
- Maciag, A., Peano, C., Pietrelli, A., Egli, T., De Bellis, G., and Landini, P.** (2011) *In vitro* transcription profiling of the sigmaS subunit of bacterial RNA polymerase: re-definition of the sigmaS regulon and identification of sigmaS-specific promoter sequence elements. *Nucleic Acids Res* **39**: 5338-5355.
- Mackie, G.A.** (1998) Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**: 720-723.
- Mackie, G.A.** (2013a) RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* **11**: 45-57.
- Mackie, G.A.** (2013b) Determinants in the rpsT mRNAs recognized by the 5'-sensor domain of RNase E. *Mol Microbiol* **89**: 388-402.
- Magasanik, B.** (1993) The regulation of nitrogen utilization in enteric bacteria. *J Cell Biochem* **51**: 34-40.
- Majdalani, N., Chen, S., Murrow, J., St John, K., and Gottesman, S.** (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol. Microbiol.* **39**: 1382-1394.
- Majdalani, N., Hernandez, D., and Gottesman, S.** (2002) Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol. Microbiol.* **46**: 813-826.
- Mandin, P., and Guillier, M.** (2013) Expanding control in bacteria: interplay between small RNAs and transcriptional regulators to control gene expression. *Curr Opin Microbiol* **16**: 125-132.
- Massé, E., Escorcia, F.E., and Gottesman, S.** (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev.* **17**: 2374-2383.
- McCown, P.J., Winkler, W.C., and Breaker, R.R.** (2012) Mechanism and distribution of *glmS* ribozymes. *Methods Mol Biol* **848**: 113-129.
- Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juarez, K., Contreras-Moreira, B., Huerta, A.M., Collado-Vides, J., and Morett, E.** (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One* **4**: e7526.
- Mengin-Lecreulx, D., and van Heijenoort, J.** (1993) Identification of the *glmU* gene encoding N-acetylglucosamine-1-phosphate uridylyltransferase in *Escherichia coli*. *J. Bacteriol.* **175**: 6150-6157.

- Mengin-Lecreulx, D., and van Heijenoort, J.** (1994) Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. *J. Bacteriol.* **176**: 5788-5795.
- Mengin-Lecreulx, D., and van Heijenoort, J.** (1996) Characterization of the essential gene *glmM* encoding phosphoglucosamine mutase in *Escherichia coli*. *J. Biol. Chem.* **271**: 32-39.
- Merrick, M.J.** (1993) In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N). *Mol. Microbiol.* **10**: 903-909.
- Miczak, A., Srivastava, R.A., and Apirion, D.** (1991) Location of the RNA-processing enzymes RNase III, RNase E and RNase P in the *Escherichia coli* cell. *Mol Microbiol* **5**: 1801-1810.
- Miczak, A., Kaberdin, V.R., Wei, C.L., and Lin-Chao, S.** (1996) Proteins associated with RNase E in a multicomponent ribonucleolytic complex. *Proc. Natl. Acad. Sci. U S A* **93**: 3865-3869.
- Mika, F., and Hengge, R.** (2005) A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in *E. coli*. *Genes Dev* **19**: 2770-2781.
- Milewski, S.** (2002) Glucosamine-6-phosphate synthase--the multi-facets enzyme. *Biochim. Biophys. Acta.* **1597**: 173-192.
- Miller, J.** (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Mitrophanov, A.Y., and Groisman, E.A.** (2008) Signal integration in bacterial two-component regulatory systems. *Genes Dev.* **22**: 2601-2611.
- Mizrahi, I., Biran, D., and Ron, E.Z.** (2006) Requirement for the acetyl phosphate pathway in *Escherichia coli* ATP-dependent proteolysis. *Mol Microbiol* **62**: 201-211.
- Mohanty, B.K., and Kushner, S.R.** (2006) The majority of *Escherichia coli* mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucleic Acids Res* **34**: 5695-5704.
- Møller, T., Franch, T., Udesen, C., Gerdes, K., and Valentin-Hansen, P.** (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev.* **16**: 1696-1706.
- Morita, T., Kawamoto, H., Mizota, T., Inada, T., and Aiba, H.** (2004) Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in *Escherichia coli*. *Mol. Microbiol.* **54**: 1063-1075.
- Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A.R., and Schnetz, K.** (2007) Repression by binding of H-NS within the transcription unit. *J. Biol. Chem.* **282**: 23622-23630.
- Neidhardt, F.C., Bloch, P.L., and Smith, D.F.** (1974) Culture medium for enterobacteria. *J. Bacteriol.* **119**: 736-747.

- Newman, J.A., Hewitt, L., Rodrigues, C., Solovyova, A.S., Harwood, C.R., and Lewis, R.J.** (2012) Dissection of the network of interactions that links RNA processing with glycolysis in the *Bacillus subtilis* degradosome. *J Mol Biol* **416**: 121-136.
- Nishino, K., Senda, Y., and Yamaguchi, A.** (2008) The AraC-family regulator GadX enhances multidrug resistance in *Escherichia coli* by activating expression of mdtEF multidrug efflux genes. *J Infect Chemother* **14**: 23-29.
- Olejniczak, M.** (2011) Despite similar binding to the Hfq protein regulatory RNAs widely differ in their competition performance. *Biochemistry* **50**: 4427-4440.
- Opdyke, J.A., Kang, J.G., and Storz, G.** (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.* **186**: 6698-6705.
- Opdyke, J.A., Fozo, E.M., Hemm, M.R., and Storz, G.** (2011) RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J Mol Biol* **406**: 29-43.
- Otaka, H., Ishikawa, H., Morita, T., and Aiba, H.** (2011) PolyU tail of *rho*-independent terminator of bacterial small RNAs is essential for Hfq action. *Proc. Natl. Acad. Sci. U S A* **108**: 13059-13064.
- Ow, M.C., Liu, Q., and Kushner, S.R.** (2000) Analysis of mRNA decay and rRNA processing in *Escherichia coli* in the absence of RNase E-based degradosome assembly. *Mol Microbiol* **38**: 854-866.
- Panja, S., and Woodson, S.A.** (2012) Hfq proximity and orientation controls RNA annealing. *Nucleic Acids Res.* **40**: 8690-8697.
- Papenfort, K., Said, N., Welsink, T., Lucchini, S., Hinton, J.C., and Vogel, J.** (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol. Microbiol.* **74**: 139-158.
- Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C.K., and Vogel, J.** (2013) Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* **153**: 426-437.
- Paradis, S., Boissinot, M., Paquette, N., Belanger, S.D., Martel, E.A., Boudreau, D.K., Picard, F.J., Ouellette, M., Roy, P.H., and Bergeron, M.G.** (2005) Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase beta-subunit. *Int. J. Syst. Evol. Microbiol.* **55**: 2013-2025.
- Pichon, C., and Felden, B.** (2007) Proteins that interact with bacterial small RNA regulators. *FEMS Microbiol Rev.* **31**: 614-625.
- Plumbridge, J.** (1995) Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the *glmUS* operon and requires two separated NagC binding sites. *Embo J.* **14**: 3958-3965.

- Plumbridge, J., and Vimr, E.** (1999) Convergent pathways for utilization of the amino sugars N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by *Escherichia coli*. *J. Bacteriol.* **181**: 47-54.
- Pompeo, F., Luciano, J., Brochier-Armanet, C., and Galinier, A.** (2011) The GTPase function of YvcJ and its subcellular relocalization are dependent on growth conditions in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* **20**: 156-167.
- Postma, P.W., Lengeler, J.W., and Jacobson, G.R.** (1993) Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**: 543-594.
- Prévost, K., Desnoyers, G., Jacques, J.F., Lavoie, F., and Massé, E.** (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* **25**: 385-396.
- Prud'homme-Genereux, A., Beran, R.K., Iost, I., Ramey, C.S., Mackie, G.A., and Simons, R.W.** (2004) Physical and functional interactions among RNase E, polynucleotide phosphorylase and the cold-shock protein, CsdA: evidence for a 'cold shock degradosome'. *Mol Microbiol* **54**: 1409-1421.
- Pyle, A.M.** (2008) Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu Rev Biophys* **37**: 317-336.
- Rabus, R., Reizer, J., Paulsen, I., and Saier, M.H., Jr.** (1999) Enzyme I<sup>Ntr</sup> from *Escherichia coli*. A novel enzyme of the phosphoenolpyruvate-dependent phosphotransferase system exhibiting strict specificity for its phosphoryl acceptor, NPr. *J. Biol. Chem.* **274**: 26185-26191.
- Reading, N.C., Torres, A.G., Kendall, M.M., Hughes, D.T., Yamamoto, K., and Sperandio, V.** (2007) A novel two-component signaling system that activates transcription of an enterohemorrhagic *Escherichia coli* effector involved in remodeling of host actin. *J. Bacteriol.* **189**: 2468-2476.
- Reading, N.C., Rasko, D.A., Torres, A.G., and Sperandio, V.** (2009) The two-component system QseEF and the membrane protein QseG link adrenergic and stress sensing to bacterial pathogenesis. *Proc. Natl. Acad. Sci. U S A* **106**: 5889-5894.
- Regnier, P., and Hajnsdorf, E.** (1991) Decay of mRNA encoding ribosomal protein S15 of *Escherichia coli* is initiated by an RNase E-dependent endonucleolytic cleavage that removes the 3' stabilizing stem and loop structure. *J. Mol. Biol.* **217**: 283-292.
- Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E., and Görke, B.** (2008) The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucleic Acids Res.* **36**: 2570-2580.
- Reichenbach, B.** (2009) Regulation of glucosamine-6-phosphate synthase synthesis by a hierarchical acting cascade composed of two small regulatory RNAs in *Escherichia coli*. Göttingen: Georg-August-University.

- Reichenbach, B., Göpel, Y., and Görke, B.** (2009) Dual control by perfectly overlapping sigma 54- and sigma 70- promoters adjusts small RNA GlmY expression to different environmental signals. *Mol. Microbiol.* **74**: 1054-1070.
- Reitzer, L., and Schneider, B.L.** (2001) Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* **65**: 422-444.
- Repoila, F., and Darfeuille, F.** (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol. Cell* **101**: 117-131.
- Resch, M., Göpel, Y., Görke, B., and Ficner, R.** (2013) Crystallization and preliminary X-ray diffraction data analysis of YhbJ from *Escherichia coli*, a key protein involved in the GlmYZ sRNA regulatory cascade. *Acta Crystallographica Section F*: doi:10.1107/S1744309112048622.
- Rice, J.B., Balasubramanian, D., and Vanderpool, C.K.** (2012) Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proc Natl Acad Sci U S A* **109**: E2691-2698.
- Richards, G.R., and Vanderpool, C.K.** (2011) Molecular call and response: the physiology of bacterial small RNAs. *Biochim. Biophys. Acta* **1809**: 525-531.
- Romeo, T., Gong, M., Liu, M.Y., and Brun-Zinkernagel, A.M.** (1993) Identification and molecular characterization of csrA, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J Bacteriol* **175**: 4744-4755.
- Romeo, T.** (1998) Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol* **29**: 1321-1330.
- Romeo, T., Vakulskas, C.A., and Babitzke, P.** (2012) Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems. *Environ. Microbiol.*: doi: 10.1111/j.1462-2920.2012.02794.x.
- Saitou, N., and Nei, M.** (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.
- Salvail, H., and Massé, E.** (2012) Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis. *Wiley Interdiscip Rev RNA* **3**: 26-36.
- Sambrook, J., and Russell, D.** (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory.
- Sauer, E., and Weichenrieder, O.** (2011) Structural basis for RNA 3'-end recognition by Hfq. *Proc. Natl. Acad. Sci. U S A* **108**: 13065-13070.
- Sauer, E., Schmidt, S., and Weichenrieder, O.** (2012) Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc. Natl. Acad. Sci. U S A* **109**: 9396-9401.

- Scharf, B.E.** (2010) Summary of useful methods for two-component system research. *Curr. Opin. Microbiol.* **13**: 246-252.
- Schnetz, K., Stülke, J., Gertz, S., Krüger, S., Krieg, M., Hecker, M., and Rak, B.** (1996) LicT, a *Bacillus subtilis* transcriptional antiterminator protein of the BglG family. *J. Bacteriol.* **178**: 1971-1979.
- Schuck, A., Diwa, A., and Belasco, J.G.** (2009) RNase E autoregulates its synthesis in *Escherichia coli* by binding directly to a stem-loop in the rne 5' untranslated region. *Mol Microbiol* **72**: 470-478.
- Schwer, B., and Verdin, E.** (2008) Conserved metabolic regulatory functions of sirtuins. *Cell Metab* **7**: 104-112.
- Serganov, A., and Nudler, E.** (2013) A decade of riboswitches. *Cell* **152**: 17-24.
- Sharma, C.M., Darfeuille, F., Plantinga, T.H., and Vogel, J.** (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* **21**: 2804-2817.
- Sharma, C.M., Papenfort, K., Pernitzsch, S.R., Mollenkopf, H.J., Hinton, J.C., and Vogel, J.** (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**: 1144-1165.
- Singh, D., Chang, S.J., Lin, P.H., Averina, O.V., Kaberdin, V.R., and Lin-Chao, S.** (2009) Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A* **106**: 864-869.
- Sittka, A., Pfeiffer, V., Tedin, K., and Vogel, J.** (2007) The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. *Mol. Microbiol.* **63**: 193-217.
- Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C., and Vogel, J.** (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet.* **4**: e1000163.
- Smith, M.M.** (1991) Histone structure and function. *Curr Opin Cell Biol* **3**: 429-437.
- Sonnleitner, E., Abdou, L., and Haas, D.** (2009) Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U S A* **106**: 21866-21871.
- Sonnleitner, E., and Haas, D.** (2011) Small RNAs as regulators of primary and secondary metabolism in *Pseudomonas* species. *Appl. Microbiol. Biotechnol.* **91**: 63-79.
- Soper, T., Mandin, P., Majdalani, N., Gottesman, S., and Woodson, S.A.** (2010) Positive regulation by small RNAs and the role of Hfq. *Proc. Natl. Acad. Sci. U S A* **107**: 9602-9607.
- Soufi, B., Soares, N.C., Ravikumar, V., and Macek, B.** (2012) Proteomics reveals evidence of cross-talk between protein modifications in bacteria: focus on acetylation and phosphorylation. *Curr Opin Microbiol* **15**: 357-363.

- Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D., and Escalante-Semerena, J.C.** (2002) Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. *Science* **298**: 2390-2392.
- Stock, A.M., Robinson, V.L., and Goudreau, P.N.** (2000) Two-component signal transduction. *Annu. Rev. Biochem.* **69**: 183-215.
- Stoppel, R., Manavski, N., Schein, A., Schuster, G., Teubner, M., Schmitz-Linneweber, C., and Meurer, J.** (2012) RHON1 is a novel ribonucleic acid-binding protein that supports RNase E function in the *Arabidopsis* chloroplast. *Nucleic Acids Res.* **40**: 8593-8606.
- Storz, G., Vogel, J., and Wassarman, K.M.** (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell.* **43**: 880-891.
- Stratmann, T., Madhusudan, S., and Schnetz, K.** (2008) Regulation of the *yjjQ-bglJ* operon, encoding LuxR-type transcription factors, and the divergent *yjjP* gene by H-NS and LeuO. *J. Bacteriol.* **190**: 926-935.
- Studier, F.W., and Moffatt, B.A.** (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**: 113-130.
- Stülke, J., Martin-Verstraete, I., Zagorec, M., Rose, M., Klier, A., and Rapoport, G.** (1997) Induction of the *Bacillus subtilis ptsGHI* operon by glucose is controlled by a novel antiterminator, GlcT. *Mol. Microbiol.* **25**: 65-78.
- Suzuki, H.I., and Miyazono, K.** (2011) Emerging complexity of microRNA generation cascades. *J. Biochem.* **149**: 15-25.
- Suzuki, K., Wang, X., Weilbacher, T., Pernestig, A.K., Melefors, O., Georgellis, D., Babitzke, P., and Romeo, T.** (2002) Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of *Escherichia coli*. *J. Bacteriol.* **184**: 5130-5140.
- Suzuki, K., Babitzke, P., Kushner, S.R., and Romeo, T.** (2006) Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. *Genes Dev.* **20**: 2605-2617.
- Svenningsen, S.L., Tu, K.C., and Bassler, B.L.** (2009) Gene dosage compensation calibrates four regulatory RNAs to control *Vibrio cholerae* quorum sensing. *Embo J.* **28**: 429-439.
- Swinger, K.K., and Rice, P.A.** (2004) IHF and HU: flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* **14**: 28-35.
- Szurmant, H., White, R.A., and Hoch, J.A.** (2007) Sensor complexes regulating two-component signal transduction. *Curr Opin Struct Biol* **17**: 706-715.
- Tam, C., and Missiakas, D.** (2005) Changes in lipopolysaccharide structure induce the sigma(E)-dependent response of *Escherichia coli*. *Mol Microbiol* **55**: 1403-1412.
- Thao, S., Chen, C.S., Zhu, H., and Escalante-Semerena, J.C.** (2010) N<sup>epsilon</sup>-lysine acetylation of a bacterial transcription factor inhibits its DNA-binding activity. *PLoS One* **5**: e15123.

- Tintut, Y., Wang, J.T., and Gralla, J.D.** (1995) A novel bacterial transcription cycle involving sigma 54. *Genes Dev* **9**: 2305-2313.
- Tu, K.C., and Bassler, B.L.** (2007) Multiple small RNAs act additively to integrate sensory information and control quorum sensing in *Vibrio harveyi*. *Genes Dev.* **21**: 221-233.
- Ueta, M., Yoshida, H., Wada, C., Baba, T., Mori, H., and Wada, A.** (2005) Ribosome binding proteins YhbH and YfiA have opposite functions during 100S formation in the stationary phase of *Escherichia coli*. *Genes Cells* **10**: 1103-1112.
- Ueta, M., Ohniwa, R.L., Yoshida, H., Maki, Y., Wada, C., and Wada, A.** (2008) Role of HPF (hibernation promoting factor) in translational activity in *Escherichia coli*. *J Biochem* **143**: 425-433.
- Urban, J.H., Papenfort, K., Thomsen, J., Schmitz, R.A., and Vogel, J.** (2007) A conserved small RNA promotes discoordinate expression of the *glmUS* operon mRNA to activate GlmS synthesis. *J. Mol. Biol.* **373**: 521-528.
- Urban, J.H., and Vogel, J.** (2007) Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res.* **35**: 1018-1037.
- Urban, J.H., and Vogel, J.** (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol.* **6**: e64.
- Uzzau, S., Figueroa-Bossi, N., Rubino, S., and Bossi, L.** (2001) Epitope tagging of chromosomal genes in *Salmonella*. *Proc. Natl. Acad. Sci. U S A* **98**: 15264-15269.
- Valverde, C., Heeb, S., Keel, C., and Haas, D.** (2003) RsmY, a small regulatory RNA, is required in concert with RsmZ for GacA-dependent expression of biocontrol traits in *Pseudomonas fluorescens* CHA0. *Mol. Microbiol.* **50**: 1361-1379.
- Valverde, C., and Haas, D.** (2008) Small RNAs controlled by two-component systems. *Adv. Exp. Med. Biol.* **631**: 54-79.
- Vanderpool, C.K.** (2007) Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress. *Curr. Opin. Microbiol.* **10**: 146-151.
- Vanderpool, C.K., Balasubramanian, D., and Lloyd, C.R.** (2011) Dual-function RNA regulators in bacteria. *Biochimie* **93**: 1943-1949.
- Vanzo, N.F., Li, Y.S., Py, B., Blum, E., Higgins, C.F., Raynal, L.C., Krisch, H.M., and Carpousis, A.J.** (1998) Ribonuclease E organizes the protein interactions in the *Escherichia coli* RNA degradosome. *Genes Dev* **12**: 2770-2781.
- Venkatesh, G.R., Kembou Koungni, F.C., Paukner, A., Stratmann, T., Blissenbach, B., and Schnetz, K.** (2010) BglJ-RcsB heterodimers relieve repression of the *Escherichia coli* *bgl* operon by H-NS. *J. Bacteriol.* **192**: 6456-6464.

- Vogel, J., and Papenfort, K. (2006) Small non-coding RNAs and the bacterial outer membrane. *Curr. Opin. Microbiol.* **9**: 605-611.
- Vogel, J. (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Mol. Microbiol.* **71**: 1-11.
- Vogel, J., and Luisi, B.F. (2011) Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* **9**: 578-589.
- Wade, J.T., Roa, D.C., Grainger, D.C., Hurd, D., Busby, S.J., Struhl, K., and Nudler, E. (2006) Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat. Struct. Mol. Biol.* **13**: 806-814.
- Wang, L., and Brown, S.J. (2006) BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. *Nucleic Acids Res.* **34**: W243-248.
- Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.B., Zeng, R., Xiong, Y., Guan, K.L., Zhao, S., and Zhao, G.P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* **327**: 1004-1007.
- Wassarman, K.M. (2007) 6S RNA: a small RNA regulator of transcription. *Curr Opin Microbiol* **10**: 164-168.
- Waters, L.S., and Storz, G. (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615-628.
- Weilbacher, T., Suzuki, K., Dubey, A.K., Wang, X., Gudapaty, S., Morozov, I., Baker, C.S., Georgellis, D., Babitzke, P., and Romeo, T. (2003) A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.* **48**: 657-670.
- Weinert, B.T., Iesmantavicius, V., Wagner, S.A., Scholz, C., Gummesson, B., Beli, P., Nystrom, T., and Choudhary, C. (2013) Acetyl-Phosphate Is a Critical Determinant of Lysine Acetylation in *E. coli*. *Mol Cell* **51**: 265-272.
- Wen, Y., Feng, J., Scott, D.R., Marcus, E.A., and Sachs, G. (2011) A *cis*-encoded antisense small RNA regulated by the HP0165-HP0166 two-component system controls expression of *ureB* in *Helicobacter pylori*. *J. Bacteriol.* **193**: 40-51.
- Wigneshweraraj, S., Bose, D., Burrows, P.C., Joly, N., Schumacher, J., Rappas, M., Pape, T., Zhang, X., Stockley, P., Severinov, K., and Buck, M. (2008) Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Mol. Microbiol.* **68**: 538-546.
- Willkomm, D.K., and Hartmann, R.K. (2005) 6S RNA - an ancient regulator of bacterial RNA polymerase rediscovered. *Biol Chem* **386**: 1273-1277.
- Wilson, G.G., Young, K.Y., Edlin, G.J., and Konigsberg, W. (1979) High-frequency generalised transduction by bacteriophage T4. *Nature* **280**: 80-82.
- Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A., and Breaker, R.R. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**: 281-286.
- Wolfe, A.J. (2005) The acetate switch. *Microbiol Mol Biol Rev* **69**: 12-50.

- Wolfe, A.J., Parikh, N., Lima, B.P., and Zemaitaitis, B.** (2008) Signal integration by the two-component signal transduction response regulator CpxR. *J Bacteriol* **190**: 2314-2322.
- Worrall, J.A., Gorna, M., Crump, N.T., Phillips, L.G., Tuck, A.C., Price, A.J., Bavro, V.N., and Luisi, B.F.** (2008) Reconstitution and analysis of the multienzyme *Escherichia coli* RNA degradosome. *J. Mol. Biol.* **382**: 870-883.
- Xhemalce, B., Robson, S.C., and Kouzarides, T.** (2012) Human RNA Methyltransferase BCDIN3D Regulates MicroRNA Processing. *Cell* **151**: 278-288.
- Yakhnin, A.V., Baker, C.S., Vakulskas, C.A., Yakhnin, H., Berezin, I., Romeo, T., and Babitzke, P.** (2013) CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage. *Mol Microbiol* **87**: 851-866.
- Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., and Ishihama, A.** (2005) Functional characterization *in vitro* of all two-component signal transduction systems from *Escherichia coli*. *J. Biol. Chem.* **280**: 1448-1456.
- Yan, J., Barak, R., Liarzi, O., Shainskaya, A., and Eisenbach, M.** (2008) *In vivo* acetylation of CheY, a response regulator in chemotaxis of *Escherichia coli*. *J Mol Biol* **376**: 1260-1271.
- Yu, B.J., Kim, J.A., Moon, J.H., Ryu, S.E., and Pan, J.G.** (2008) The diversity of lysine-acetylated proteins in *Escherichia coli*. *J Microbiol Biotechnol* **18**: 1529-1536.
- Zakikhany, K., Harrington, C.R., Nimtz, M., Hinton, J.C., and Römling, U.** (2010) Unphosphorylated CsgD controls biofilm formation in *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.* **77**: 771-786.
- Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G.** (1998) The OxyS regulatory RNA represses *rpoS* translation and binds the Hfq (HF-I) protein. *Embo J* **17**: 6061-6068.
- Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C.F., Grishin, N.V., and Zhao, Y.** (2009) Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics* **8**: 215-225.
- Zhao, K., Liu, M., and Burgess, R.R.** (2005) The global transcriptional response of *Escherichia coli* to induced sigma 32 protein involves sigma 32 regulon activation followed by inactivation and degradation of sigma 32 *in vivo*. *J Biol Chem* **280**: 17758-17768.
- Zhao, K., Liu, M., and Burgess, R.R.** (2010) Promoter and regulon analysis of nitrogen assimilation factor, sigma54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. *Nucleic Acids Res.* **38**: 1273-1283.
- Zhao, M., Zhou, L., Kawarasaki, Y., and Georgiou, G.** (2006) Regulation of RraA, a protein inhibitor of RNase E-mediated RNA decay. *J Bacteriol* **188**: 3257-3263.

- Zhao, X., Koestler, B.J., Waters, C.M., and Hammer, B.K.** (2013) Post-transcriptional activation of a diguanylate cyclase by quorum sensing small RNAs promotes biofilm formation in *Vibrio cholerae*. *Mol Microbiol*: doi:10.1111/mmi.12325.
- Zimmer, B., Hillmann, A., and Görke, B.** (2008) Requirements for the phosphorylation of the *Escherichia coli* EIIA<sup>Ntr</sup> protein *in vivo*. *FEMS Microbiol. Lett.* **286**: 96-102.
- Zisoulis, D.G., Kai, Z.S., Chang, R.K., and Pasquinelli, A.E.** (2012) Autoregulation of microRNA biogenesis by let-7 and Argonaute. *Nature* **486**: 541-544.

## CURRICULUM VITAE

**Name:** Yvonne Göpel  
**Date of birth:** 18.04.1985  
**Place of birth:** Eisenach, Deutschland  
**Nationality:** German

**Professional address:** Institute for Microbiology and Genetics, Georg-August-University Göttingen,  
Grisebachstraße 8, D-37077 Göttingen

### School Education:

1989 to 1993 Grundschole Maxim Gorki, Förtha  
1995 to 2003 Ernst-Abbe-Gymnasium, Eisenach;  
2003 Abitur

### Stay abroad:

10/2003 – 5/2005 Rye, New York, USA  
2004-2005 part-time student at Manhattanville College, Purchase, NY

### Scientific Education:

10/2005 – 04/2010 Study of Biology (Diploma) at the Georg-August-University, Göttingen  
04/2010 Diploma in Biology  
2010 – 2013 PhD project:  
“GlmY and GlmZ: a hierarchically acting regulatory cascade composed of two small RNAs” in the group of PD Dr. Boris Görke, Department of General Microbiology, Institute for Microbiology and Genetics,  
Georg-August-Universität Göttingen

