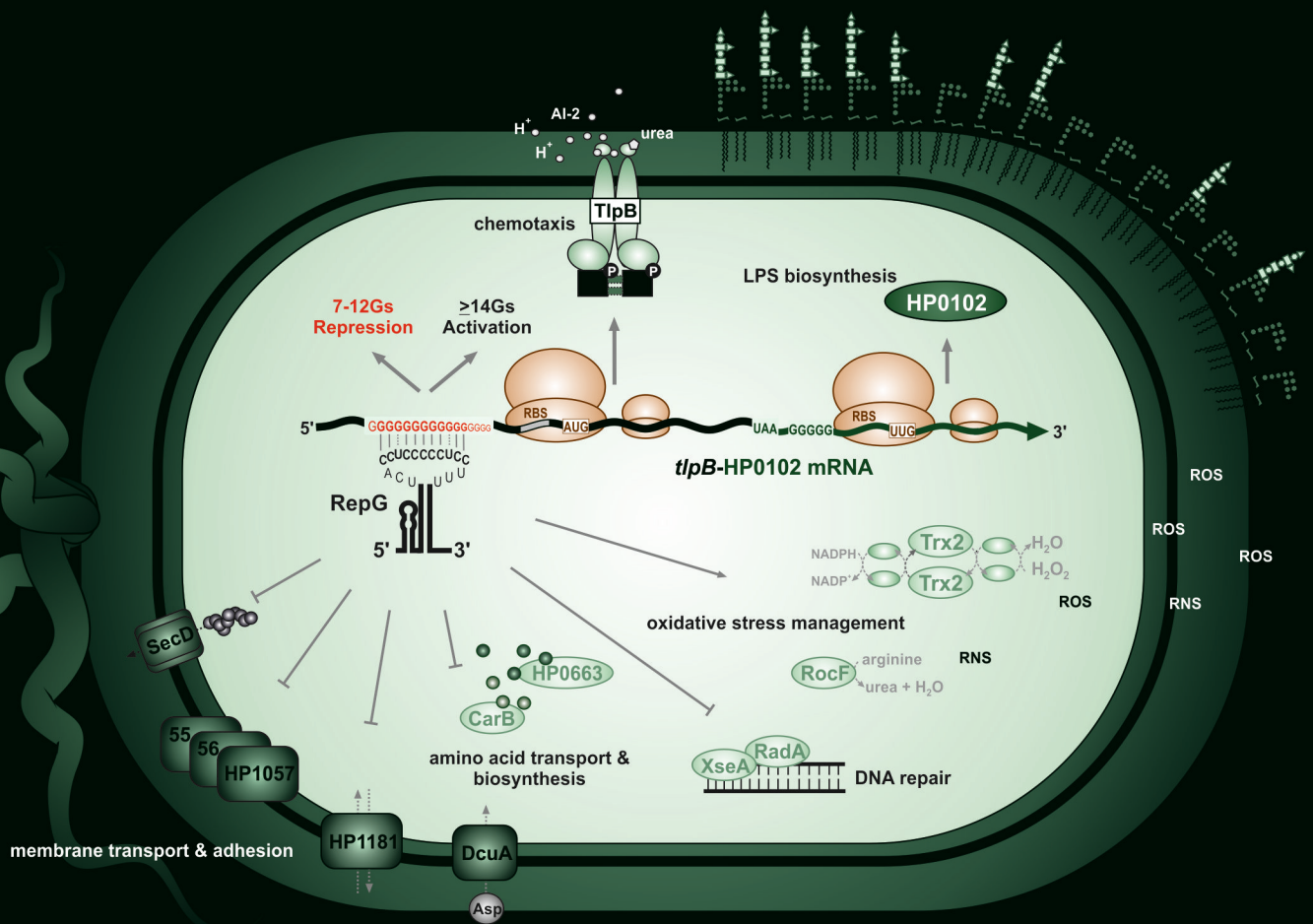# Functional characterization of the abundant and conserved small regulatory RNA RepG in *Helicobacter pylori*

Dissertation

Sandy Ramona Pernitzsch

Julius-Maximilians-Universität Würzburg, 2015

Functional characterization of the abundant and conserved small regulatory RNA RepG in *Helicobacter pylori*

Funktionelle Charakterisierung der abundanten und
konservierten kleinen regulatorischen RNA RepG
in *Helicobacter pylori*

Dissertation

zur Erlangung des naturwissenschaftlichen Doktorgrades (Dr. rer. nat.)
der *Graduate School of Life Sciences*,
Julius-Maximilians-Universität Würzburg

Sektion: Infektion und Immunität

vorgelegt von

Sandy Ramona Pernitzsch

aus Oranienburg

Würzburg, 2015

Vorgelegt von: Diplom Biologin Sandy Ramona Pernitzsch

Angefertigt im: Zentrum für Infektionsforschung (ZINF),
                        Julius-Maximilians-Universität Würzburg

Tag der Einreichung: 20. Mai 2015

Tag der Verteidigung:

Gutachter:

Dr. Cynthia M. Sharma   Zentrum für Infektionsforschung (ZINF),
                           Julius-Maximilians-Universität Würzburg

Prof. Dr. Dagmar Beier   Lehrstuhl für Mikrobiologie, Biozentrum,
                           Julius-Maximilians-Universität Würzburg

Mitglieder der Promotionskommission:

Dr. Cynthia M. Sharma   Zentrum für Infektionsforschung (ZINF),
                           Julius-Maximilians-Universität Würzburg

Prof. Dr. Dagmar Beier   Lehrstuhl für Mikrobiologie, Biozentrum,
                           Julius-Maximilians-Universität Würzburg

Dr. Tobias Ölschläger   Institut für Molekulare Infektionsbiologie (IMIB),
                           Julius-Maximilians-Universität Würzburg

Prof. Dr. Jörg Vogel   Institut für Molekulare Infektionsbiologie (IMIB),
                           Julius-Maximilians-Universität Würzburg

Prüfungsvorsitzender:

Prof. Dr. Thomas Dandekar Lehrstuhl für Bioinformatik, Biozentrum,
                           Julius-Maximilians-Universität Würzburg

Affidavit/Eidesstattliche Erklärung

I hereby confirm that my thesis entitled "Functional characterization of the abundant and conserved small regulatory RNA RepG in *Helicobacter pylori*" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or material applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, May 20th, 2015

_____

Sandy Ramona Pernitzsch

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation „Funktionelle Charakterisierung der abundanten und konservierten kleinen regulatorischen RNA RepG in *Helicobacter pylori*" eigenständig, d.h. insbesondere ohne die Hilfe oder Unterstützung von kommerziellen Promotionsberatern, angefertigt habe. Ergänzend bestätige ich, dass ich keine anderen als die von mir angegebenen Quellen oder Hilfsmittel verwendet habe.

Ich erkläre außerdem, dass diese Dissertation weder in gleicher noch in ähnlicher Form bereits in einem Prüfungsverfahren vorgelegen hat.

Würzburg, den 20. Mai, 2015

_____

Sandy Ramona Pernitzsch

Dedicated to my parents, Ramona and Holger Pernitzsch.

Summary

Bacterial small non-coding RNAs (sRNAs) play fundamental roles in controlling and fine-tuning gene expression in a wide variety of cellular processes, including stress responses, environmental signaling and virulence in pathogens. Despite the identification of hundreds of sRNA candidates in diverse bacteria by genomics approaches, the mechanisms and regulatory capabilities of these posttranscriptional regulators have most intensively been studied in Gram-negative Gammaproteobacteria such as *Escherichia coli* and *Salmonella*. So far, almost nothing is known about sRNA-mediated regulation (riboregulation) in Epsilonproteobacteria, including the major human pathogen *Helicobacter pylori*. *H. pylori* was even thought to be deficient for riboregulation as none of the sRNAs known from enterobacteria are conserved in *Helicobacter* and since it lacks the major RNA chaperone Hfq, which is crucial for sRNA function as well as stability in many bacteria. Nonetheless, more than 60 *cis*- and *trans*-acting sRNA candidates were recently identified in *H. pylori* by a global RNA sequencing approach, indicating that this pathogen, in principle, has the capability to use riboregulation for its gene expression control. However, the functions and underlying mechanisms of *H. pylori* sRNAs remained unclear.

This thesis focused on the first functional characterization and target gene identification of a *trans*-acting sRNA, RepG (Regulator of polymeric G-repeats), in *H. pylori*. Using *in-vitro* and *in-vivo* approaches, RepG was shown to directly base-pair with its C/U-rich terminator loop to a variable homopolymeric G-repeat in the 5' untranslated region (UTR) of the *tlpB* mRNA, thereby regulating expression of the chemotaxis receptor TlpB. While the RepG sRNA is highly conserved, the length of the G-repeat in the *tlpB* mRNA leader varies among different *H. pylori* isolates, resulting in a strain-specific *tlpB* regulation. The modification of the number of guanines within the G-stretch in *H. pylori* strain 26695 demonstrated that the length of the homopolymeric G-repeat determines the outcome of posttranscriptional control (repression or activation) of *tlpB* by RepG. This length-dependent targeting of a simple sequence repeat by a *trans*-acting sRNA represents a new twist in sRNA-mediated regulation and a novel mechanism of gene expression control, since it uniquely links phase variation by simple sequence repeats to posttranscriptional regulation.

In almost all sequenced *H. pylori* strains, *tlpB* is encoded in a two gene operon upstream of HP0102, a gene of previously unknown function. This study provided evidence that HP0102 encodes a glycosyltransferase involved in LPS O-chain and Lewis x antigen production. Accordingly, this glycosyltransferase was shown to be essential for mice colonization by *H. pylori*. The coordinated posttranscriptional regulation of the *tlpB*-HP0102 operon by antisense base-pairing of RepG to the phase-variable G-repeat in the 5' UTR of the

tlpB mRNA allows for a gradual, rather than ON/OFF, control of HP0102 expression, thereby affecting LPS biosynthesis in *H. pylori*. This fine-tuning of O-chain and Lewis x antigen expression modulates *H. pylori* antibiotics sensitivity and thus, might be advantageous for *Helicobacter* colonization and persistence.

Whole transcriptome analysis based on microarray and RNA sequencing was used to identify additional RepG target mRNAs and uncover the physiological role of this riboregulator in *H. pylori*. Altogether, *repG* deletion affected expression of more than 40 target gene candidates involved various cellular processes, including membrane transport and adhesion, LPS modification, amino acid metabolism, oxidative and nitrosative stress, and nucleic acid modification. The presence of homopolymeric G-repeats/G-rich sequences in almost all target mRNA candidates indicated that RepG hijacks a conserved motif to recognize and regulate multiple target mRNAs in *H. pylori*.

Overall, this study demonstrates that *H. pylori* employs riboregulation in stress response and virulence control. In addition, this thesis has successfully established *Helicobacter* as a new model organism for investigating general concepts of gene expression control by Hfq-independent sRNAs and sRNAs in bacterial pathogens.

Zusammenfassung

Bakterielle kleine, nicht-kodierende RNAs (sRNAs, engl. für small RNAs) spielen eine fundamentale Rolle in der Kontrolle und Feinabstimmung der Genexpression in Bakterien. Sie sind an einer Vielzahl von zellulären Prozessen, einschließlich der Adaption an unterschiedliche Stress- sowie Umweltbedingungen und der Virulenz von bakteriellen Pathogenen, beteiligt. Trotz der Identifizierung von Hunderten von sRNA-Kandidaten in diversen Bakterien durch genomweite Untersuchungsmethoden, wurden die regulatorischen Eigenschaften und Mechanismen dieser posttranskriptionellen Regulatoren bisher hauptsächlich in Gram-negativen Gammaproteobakterien wie *Escherichia coli* und *Salmonella* untersucht. Bislang ist nur wenig über sRNA-basierte Regulation (Riboregulation) in Epsilonproteobakterien, einschließlich dem weitverbreiteten Humanpathogen *Helicobacter pylori*, bekannt. Es wurde sogar angenommen, dass *H. pylori* über keine Art der Riboregulation verfügt, da keine der enterobakteriellen sRNAs in *Helicobacter* konserviert sind. Zudem konnte in diesem Erreger kein Homolog für das RNA-Chaperon Hfq, welches in vielen Bakterien essentiell für die Funktion und Stabilität von sRNAs ist, identifiziert werden. Nichtsdestotrotz wurden mit Hilfe einer globalen RNA-Sequenzierungsstudie, die auf der Sequenzierung primärer Transkripte in einem Hochdurchsatzverfahren basiert, kürzlich mehr als 60 *in cis*- und *in trans*-agierende sRNA-Kandidaten in *H. pylori* identifiziert. Diese Transkriptomanalyse deutet darauf hin, dass *H. pylori* prinzipiell die Fähigkeit hat Riboregulation zur Kontrolle seiner Genexression zu nutzen. Die Funktionen und Mechanismen von sRNAs in *H. pylori* sind jedoch immer noch unklar.

In der vorgelegten Arbeit wurde erstmals eine *in trans*-agierende sRNA, RepG (<u>Re</u>gulator of <u>p</u>olymeric <u>G</u>-repeats), in *Helicobacter* charakterisiert sowie dessen zelluläre Zielgene identifiziert. Mit Hilfe diverser *in-vitro* und *in-vivo* Analysen konnte gezeigt werden, dass der C/U-reiche Transkriptionsterminatorloop von RepG direkt an eine variable, repetitive G-Sequenz in der 5' untranslatierten Region (UTR) der *tlpB* mRNA bindet. Durch diese direkte sRNA-mRNA Interaktion wird die Expression des Chemotaxis Rezeptors TlpB reguliert. Im Gegensatz zu einer hohen Konservierung der Sequenz der RepG sRNA, variiert die Länge des G-Stretches im 5' UTR der *tlpB* mRNA zwischen unterschiedlichen *H. pylori* Isolaten. Diese Längenvariation resultiert in einer Stamm-spezifischen Regulation der TlpB Expression. Die Modifikation der Anzahl der Guanin-Basen im G-Stretch des *H. pylori* Stammes 26695 demonstrierte, dass die Länge der repetitiven G-Sequenz das Ergebnis der posttranskriptionellen Regulation (Repression oder Aktivierung) von *tlpB* durch RepG beeinflusst. Die hier beschriebene Längen-abhängige Interaktion zwischen einer *in trans*-agierenden sRNA und einer einfachen, repetitiven Sequenz repräsentiert nicht nur ein neues

Konzept für die Genregulation durch sRNAs, sondern stellt auch einen neuen Mechanismus der Genexpressionskontrolle dar. Darüber hinaus, veranschaulicht die hier beschriebene sRNA-mRNA Interaktion eine bislang einzigartige Verknüpfung von Phasenvariation durch hochvariabel, repetitive Sequenzen mit Genregulation durch sRNAs.

In nahezu allen sequenzierten *H. pylori* Stämmen ist das *tlpB* Gen in einem Operon zusammen mit einem Gen mit bisher unbekannter Funktion, HP0102, kodiert. In dieser Arbeit konnte gezeigt werden, dass HP0102 für eine Glykosyltransferase kodiert, die an der Synthese der O-Seitenketten des LPS und des Lewis x Antigens in *H. pylori* beteiligt ist. Darüber hinaus konnte demonstriert werden, dass diese Glykosyltransferase für die Kolonisierung des murinen Magens durch *H. pylori* essentiell ist. Die koordinierte, posttranskriptionelle Regulation des *tlpB*-HP0102 Operons, welche durch antisense Basenpaarung zwischen RepG und der phasen-variablen, repetitiven G-Sequenz im 5' UTR der *tlpB* mRNA vermittelt wird, ermöglicht eine graduelle Kontrolle der Genexpression von HP0102, und somit Einflussnahme auf die LPS Biosynthese in *H. pylori*. Diese Feinabstimmung der LPS O-Seitenketten und Lewis x Antigen Expression beeinflusst die Resistenz von *H. pylori* gegen diverse Antibiotika und könnte somit sowohl für die Kolonisierung als auch für die persistente Infektion des Wirts durch *H. pylori* vorteilhaft sein.

Um Einblicke in die physiologische Funktion von RepG zu gewinnen, wurden in einer genom-weiten Transkriptomanalyse mittels Microarray und RNA-Sequenzierung weitere Zielgene von RepG bestimmt. Insgesamt beeinflusste die Deletion von *repG* die Expression von mehr als 40 potentiellen Zielgenen, welche an diversen zellulären Prozessen beteiligt sind, wie z.B. Membrantransport und Adhäsion, Aminosäure- und Nukleinsäure-Metabolismus, oxidative und nitrosative Stressantwort sowie LPS Modifizierung. Die Identifizierung von homopolymeren G-Stretchen bzw. G-reichen Sequenzen in allen Ziel-mRNAs deutet darauf hin, dass RepG ein konserviertes Motiv bindet, um mehrere Zielgene in *H. pylori* zu erkennen und zu regulieren.

Zusammenfassend zeigt diese Arbeit, dass *H. pylori* Riboregulation basierend auf sRNAs nutzt, um seine Genexpression in unterschiedlichen Stress- und Virulenzbedingungen zu regulieren. Darüber hinaus hat diese Studie *Helicobacter* als neuen Modelorganismus für die Untersuchung genereller Wirkungsweisen Hfq-unabhängiger sRNAs und sRNAs in bakteriellen Pathogenen etabliert.

Table of contents

Abbreviation index

A	adenine
aa	amino acid
Amp	ampicillin
APS	ammonium persulfate
asRNA	antisense RNA
ATP	adenosine triphosphate
BSA	bovine serum albumin
C	cytosine
cDNA	complementary DNA
CDS	coding sequence
CFU	colony forming units
Cm	chloramphenicol
CO_2	carbon dioxide
CTP	cytidine triphosphate
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide
DPP	2'2 dipyridyl
ds	double-stranded
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetate
Erm	erythromycin
EtOH	ethanol
FBS/FCS	fetal bovine serum/fetal calf serum
$Fe(II)SO_4$	iron (II)-sulfate
G	guanine
Gen	gentamicin
gDNA	genomic DNA
GFP	green fluorescent protein
GO	gene onthology
GT	glycosyltranferase
GTP	guanosine triphosphate
H_2O	water (distilled)
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
IGB	Integrated Genome Browser
IGR	intergenic region
Kan	kanamycin
LPS	lipopolysaccharide
mRNA	messenger RNA
miRNA	microRNA
N_2	nitrogen
NaCl	sodium chloride

NaOH	sodium hydroxide
NiCl$_2$	nickel chloride
OD$_{600nm}$	optical density at a wavelength of 600 nm
OE	overexpression
ORF	open reading frame
P	promoter
PAA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
P:C:I	phenol:chloroform:isoamyl alcohol
PCR	polymerase chain reaction
Phu	Phusion polymerase
RBP	RNA-binding protein
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
RNA-seq	RNA sequencing
RNS	reactive nitrogen species
ROS	reactive oxygen species
RT	room temperature
rRNA	ribosomal RNA
rNTP	ribonucleotide
sRNA	small regulatory/non-coding RNA
siRNA	short interfering RNA
SD	Shine-Dalgarno
SDS	sodiumdodecylsulfate
spp.	species
ss	single-standed
SSR	simple sequence repeat
Str	streptomycin
T	thymine
TAE	Tris/Acetate/EDTA
TBE	Tris/Borate/EDTA
TEMED	tetramethylethylendiamin
T$_m$	melting temperature
tmRNA	transfer-messanger RNA
Tris	tris-(hydroxymethyl)-aminomethan
tRNA	transfer RNA
TSS	transcriptional start site
U	uracil
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
vol	volume
v/v	volume/volume
w/v	weight/volume in g/ml
WT	wildtype

Units

%	percent
° C	degree Celsius
A	ampere
bp	base pair(s)
Ci	Curie
Da	Dalton
g	gravity
g	gram
h/hrs	hour(s)
l	liter
M	molar
min	minute(s)
molar	gram molecule
N	normality (volumetic)
nt	nucleotide(s)
pH	minus the decimal logarithm of the hydrogen concentration
rpm	rounds per minute
s/sec	second(s)
u	unit
V	Volt
W	Watt

Multiples

M	mega (10^6)
k	kilo (10^3)
m	milli (10^{-3})
μ	micro (10^{-6})
n	nano (10^{-9})
p	pico (10^{-12})

Aim and organization of this thesis

In bacteria, small regulatory RNAs are a versatile class of posttranscriptional regulators that are involved in gene expression control in response to various stress conditions or during virulence control (Caldelari *et al.*, 2013, Michaux *et al.*, 2014). Although sRNAs have been intensively investigated during the last years, the current knowledge of their regulatory potential and mechanisms are mainly based on work in the model organisms *Escherichia coli, Salmonella* and other Gammaproteobacteria. Using biocomputational predictions and experimental methods, a wealth of potential sRNAs have been identified on a global scale in diverse bacterial species (Sharma & Vogel, 2009, Croucher & Thomson, 2010, Sorek & Cossart, 2010). For example, a global transcriptome study using RNA sequencing revealed a surprising number of more than 60 sRNA candidates, including potential regulators of *cis-* and *trans*-encoded target mRNAs, in the major human pathogen *Helicobacter pylori* (Sharma *et al.*, 2010). Albeit this study demonstrated that *H. pylori* has the capacity for extensive riboregulation, it remained unclear whether sRNAs are functional in this pathogen. Thus, the current knowledge about their mechanisms of gene expression control, target genes and own transcriptional regulators is still very limited in *H. pylori.* In line with this, it is largely unknown how and/or to what extent sRNA-mediated gene expression control impacts on the physiology and virulence of Epsilonproteobacteria.

This study aims at the functional characterization and target gene identification of the highly abundant and conserved sRNA, RepG (<u>Re</u>gulator of <u>p</u>olymeric <u>G</u>-repeats), in *H. pylori*. This 87-nt long sRNA was previously identified as one of the most abundant transcripts in *H. pylori* strain 26695 (Sharma *et al.*, 2010). Bioinformatics-based predictions for potential target mRNAs indicated that RepG might bind with its C/U-rich terminator loop to a homopolymeric G-repeat in the 5' UTR of the *tlpB*-HP0102 mRNA, encoding the acid-sensing chemotaxis receptor TlpB and the hypothetical protein HP0102, respectively. Transcriptome and proteome analyses upon deletion of *repG* revealed sRNA-mediated repression of *tlpB* at the transcript and protein level (Sharma *et al.*, 2010), suggesting that this chemotaxis receptor represents the first *trans*-encoded target gene of RepG. The presented work (I) investigates the underlying molecular mechanisms of target gene regulation by RepG and (II) identifies/defines the regulon of this first example of a *trans*-acting sRNA in *Helicobacter*. Furthermore, the role of RepG-regulated genes (*tlpB* and HP0102) in virulence- and survival-associated phenotypes is analyzed to gain insights into the physiological function of RepG and its target genes.

Organization of this thesis

First, the biological background of gene expression control through sRNAs in bacteria is reviewed in Chapter 1. Afterwards, bacterial adaptation strategies with the emphasis on phase variation through simple sequence repeats and the here studied model organism, *Helicobacter pylori,* are introduced.

Chapter 2 focuses on the external conditions under which RepG expression is regulated and investigates the underlying molecular mechanism of RepG-mediated *tlpB* regulation. One central issue is how and/or to which extent the length of homopolymeric G-repeat in the *tlpB* mRNA leader influences RepG-mediated posttranscriptional control of this chemotaxis receptor in *Helicobacter*. A novel mechanism linking phase-variable gene expression by simple sequence repeats to posttranscriptional regulation by sRNAs is proposed. Most of the work presented and discussed in Chapter 2 has already been published in the following paper: Pernitzsch SR, Tirier SM, Beier D, Sharma CM (2014). *A variable homopolymeric G-repeat defines sRNA-mediated posttranscriptional regulation of a chemotaxis receptor in Helicobacter pylori*. PNAS, 111(4):E501-10.

In *H. pylori*, phase-variable simple sequence repeats have been shown to control expression of genes involved in the direct interaction with host cells, *e.g.* bacterial surface structures. This raises the question, why a homopolymeric G-repeat is located in the mRNA leader of the TlpB chemotaxis receptor. The *tlpB* gene is encoded in a two gene operon upstream of HP0102. Work presented in Chapter 3 aims at the functional characterization of this so far uncharacterized protein and shows that HP0102 encodes a glycosyltransferase, which is involved in *H. pylori* lipopolysaccharide O-chain and Lewis x antigen production. Accordingly, HP0102 represents as a novel colonization factor of *H. pylori*. In addition, Chapter 3 focuses on the molecular mechanism of RepG-mediated co-regulation of the *tlpB*-HP0102 operon and its impact on modulating *H. pylori* LPS production.

In Chapter 4, multi-target regulation by RepG is investigated by the analysis of global mRNA changes using microarray and RNA sequencing upon deletion of *repG* in two *H. pylori* strains. Bioinformatics-based predictions including a motif search are used to define a RepG-binding site (homopolymeric G-repeats/G-rich sequences), which is present in almost all target mRNA candidates. Work in this Chapter shows that RepG is a versatile riboregulator that controls, in addition to the *tlpB*-HP0102 operon, expression of genes involved in *H. pylori* membrane transport, LPS modification and amino acid as well as nucleic acid metabolism through hijacking a conserved G-rich motif.

The individual experimental results are discussed at the end of each Chapter, followed by an overall conclusion about the key observations and future perspectives in Chapter 5. Experimental and biocomputational methods that were used in this thesis are described in Chapter 6.

Contributions by others

The work described in this PhD thesis was conducted under supervision of Dr. Cynthia M. Sharma in the Young Investigator Group "Deep Sequencing Approaches to Pathogenesis" at the Research Center for Infectious Diseases (ZINF) of the Julius-Maximilians-University of Würzburg, Germany. Parts of the work described in this thesis, which have been contributed by others and/or have been conducted in collaborations with others, are indicated below.

- Electrophoretic gel-mobility shift assays with diverse *repG* promoter fragments and purified HPG27_385 (HP1043) were performed by Susanne Bauer and Prof. Dr. Dagmar Beier, Biocenter, University of Würzburg, Germany.
- Parts of the *H. pylori* cloning, northern blot and western blot experiments were done with technical assistance of Anika Lins and Belinda Aul (AG Sharma), and by Stephan M. Tirier, Nina Kapica and Elisabeth Schönwetter during their student internships under supervision of Sandy R. Pernitzsch. Experiments to investigate RepG expression, *i.e.* northern blot analysis of *H. pylori* 26695 grown in absence or presence of metal ions, and *H. pylori* 26695 WT and Δ*arsS* mutant grown under acidic stress conditions, were supported by the work of Stephanie Stahl and Giacomo Golfieri (former lab members of the AG Sharma).
- Mouse infection experiments with *H. pylori* strain X47-2AL were performed by Marie Robbe-Saule and PhD Hilde De Reuse, Institut Pasteur, Paris, France.
- Microarray experiments and analysis were done together with the Microarray Core Facility headed by Dr. Hans Mollenkopf (Max Planck Institute for Infection Biology, Berlin, Germany) and Dr. Kai Papenfort (former lab member AG Vogel, IMIB, University of Würzburg, Germany).
- cDNA libraries were constructed by *vertis* Biotechnology AG in Freising-Weihenstephan, Germany.
- High-throughput sequencing was performed by the team of Dr. Richard Reinhardt at the Max Planck Institute for Plant Breeding Research, Cologne, Germany.
- Processing of RNA sequencing data and bioinformatics analyses were mainly performed by Dr. Konrad U. Förstner (AG Sharma/AG Vogel, and SysMed Core Unit of the Medical Faculty, University of Würzburg, Germany). Peak detection of enriched cDNA reads in diverse RNA sequencing libraries was done by Thorsten Bischler (AG Sharma/AG Vogel, Bioinformatics of the ZINF/IMIB, University of Würzburg, Germany).

1. Biological background

1.1. Small regulatory RNAs in bacteria

Bacteria can rapidly adapt to changing conditions by altering their gene expression. Generally, regulation of gene expression occurs at multiple layers, including diverse control mechanisms at the transcriptional, posttranscriptional (translational) and posttranslational level. Messenger RNA (mRNA) abundances are mainly adjusted to changing environmental conditions by altered transcription through *e.g.* modifications of the DNA bending (structure and epigenetic modifications), transcription factors that either activate or repress promoter activities, or specificity factors such as sigma factors, which modify the promoter specificity of the RNA polymerase. In addition, posttranscriptional regulation, including the processing, localization, stability or turnover, and translation of mRNAs, adds substantial complexity to the control of gene expression. Following translation, the activity and/or stability of proteins can be modified at the posttranslational level by *e.g.* proteolytic cleavage, protein modification and interactions with other macromolecules such as chaperones and RNAs.

Conventionally, RNA is thought to contain the information required for the translation machinery, thus connecting the genome to the proteome. However, a large fraction of the bacterial genome is transcribed into functional RNA molecules that are not translated into proteins. These so-called non-coding RNAs include the highly abundant ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), which are essential for mRNA translation, as well as other house-keeping RNAs (Wassarman *et al.*, 1999). In addition, a heterologous class of non-coding RNAs act as regulatory molecules and influence transcription, translation, mRNA stability and DNA maintenance or silencing by various mechanisms (Waters & Storz, 2009). Among others, this class includes regulatory RNA elements in the 5' untranslated region (UTR) of mRNAs such as riboswitches and RNA thermometers, which affect expression of downstream encoded open reading frames in *cis* (reviewed in Wachter, 2014). Another unique class of recently discovered regulatory RNAs are CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-derived RNAs, which together with the Cas proteins constitute an adaptive immune system in bacteria (reviewed in Barrangou & Marraffini, 2014). In addition, small regulatory RNAs (sRNAs) are an emerging class of riboregulators that act on all levels of gene expression control, ranging from transcription initiation to translation and protein activity (Waters & Storz, 2009).

Bacterial sRNAs are encoded sense/antisense to open reading frames (ORFs) or within intergenic regions (IGR), and vary in both size (50 – 500 nt) and secondary structure. In general, they do not contain open reading frames, although some of them can also encode for functional small proteins (dual-function sRNAs, reviewed in Vanderpool *et al.*, 2011).

Classically, sRNA genes are transcribed from an independent promoter and contain a conserved, rho-independent transcription terminator; however, recent studies reported some of them to be derived from untranslated regions of coding genes (Miyakoshi *et al.*, 2015). Even though sRNAs are often conserved in related bacterial genomes (*e.g.* enterobacteria), different bacterial species and phyla might have evolved their own repertoire of small riboregulators (Dugar *et al.*, 2013).

The majority of the functionally characterized sRNAs exert their regulatory function at the posttranscriptional level through direct antisense base-pairing interactions with target mRNAs (antisense RNAs; will be discussed below). However, sRNAs can also bind to proteins and regulate their activity. For example, sRNAs of the CsrB/C family antagonize the activity of the translational regulator CsrA (c̲arbon s̲torage r̲egulator A) by mimicking the structure of its default RNA substrate, *i.e.* GGA-containing stem loops, and thereby sequester/titrate this RNA-binding protein away from its target mRNAs like a sponge (reviewed in Babitzke & Romeo, 2007). In contrast to CsrB/C, the abundant 6S RNA does not exert its activity posttranscriptionally, but rather influences the transcriptional process by its interaction with the housekeeping form of the RNA polymerase (reviewed in Cavanagh & Wassarman, 2014). The secondary structure of the 6S RNA mimics an open promoter complex and in turn, titrates/sequesters the RNA polymerase bound to the housekeeping sigma factor RpoD (δ^{70} in *E. coli*), in order to facilitate transcription by alternative sigma factors during nutrient starvation.

1.2.   Base-pairing small RNAs

Bacterial sRNAs often act as antisense RNAs (asRNAs) and can be divided into *cis*- and *trans*-encoded sRNAs according to their location. Whereas *cis*-encoded sRNAs originate from the opposite DNA strand and share full complementarity with their target transcripts, *trans*-encoded sRNAs are encoded at distinct genomic locations and regulate target mRNAs by short and imperfect base-pairing interactions.

Cis-encoded antisense RNAs on plasmids, transposons, and phages that control stability or maintenance of such mobile genetic elements, were among the first characterized regulatory RNAs in prokaryotes (Brantl, 2007). Such *cis*-encoded antisense RNAs, as well as several chromosomally-encoded examples, have been shown to affect gene expression by inhibition of maturation of primers required for plasmid replication, transcription attenuation, inhibition of translation, or promotion of RNA degradation and/or cleavage (Brantl, 2007; Brantl, 2012). Using tiling arrays or RNA sequencing analyses, recent genome-wide transcriptome studies of several Gram-negative and -positive bacteria suggest that naturally occurring antisense transcription from the chromosome is more widespread

than anticipated in bacteria, and revealed a wealth of *cis*-encoded antisense RNAs (small and/or long) in prokaryotes (Georg & Hess, 2011). Short and long *cis*-encoded antisense RNAs regulate a variety of functions such as virulence, toxins, motility, and biofilm formation (reviewed in Thomason & Storz, 2010 and Georg & Hess, 2011). However, it is still under debate whether all of these *cis*-encoded antisense RNAs are functional or are simply the result of spurious transcription (Wade & Grainger, 2014).

Most of the functionally characterized sRNAs are *trans*-encoded antisense RNAs that originate from intergenic regions. Exceptions notwithstanding, many *trans*-encoded sRNAs are assumed to share general properties and appear to have conceptually similar domains, at least in enterobacteria (Storz *et al.*, 2011). First, a structured 3' end followed by poly-U facilitates rho-independent transcription termination and protects sRNAs from degradation by 3' exonucleases. A second domain represents the binding site for the RNA chaperone Hfq (host factor of the bacteriophage Qβ), which is required for effective target base-pairing and stabilization of most of the *trans*-encoded sRNAs. In enterobacteria, Hfq binds preferentially single-stranded and A/U-rich regions within sRNAs (Vogel & Luisi, 2011, Tree *et al.*, 2014). However, this domain might be absent in *trans*-encoded sRNAs of *e.g.* Gram-positive bacteria, in which Hfq seems to be dispensable for sRNA-mediated gene expression control (Hammerle *et al.*, 2014). A third region is utilized for base-pairing to the target RNAs. Although loop-loop contacts between sRNAs and target mRNAs have been described (Argaman & Altuvia, 2000, Boisset *et al.*, 2007), an increasing number of Hfq-associated sRNAs use a single-stranded and conserved domain for target mRNA binding (Peer & Margalit, 2011). This is reminiscent to the "seed-pairing" mechanism known from eukaryotic microRNAs (reviewed in Bartel, 2009). Although the optimal length and composition of the bacterial "seed region" have not yet been defined, various studies have shown that several sRNAs employ a single "seed region" for multi-target recognition and regulation. For example, the GcvB sRNA in *Salmonella* was shown to directly base-pair to more than 20 mRNAs by a G/U-rich element (Sharma *et al.*, 2011). The majority of these targets are involved in amino acid uptake as well as biosynthesis, implicating that GcvB represents a regulatory node in amino acid metabolism. Likewise, many other sRNAs are key actors in the regulation of different metabolic processes, under diverse growth conditions, in adaption to stress or environmental changes, and in microbial pathogenesis (reviewed in Beisel & Storz, 2010, Michaux *et al.*, 2014 and Caldelari *et al.*, 2013). For example, the conserved enterobacterial sRNAs RyhB, OxyS, Spot42, FnrS and RybB are integral components of regulatory networks involved in gene expression control in response to iron starvation, oxidative stress, different carbon sources, anaerobicity and membrane perturbations, respectively. Thereby, sRNAs seems to generally fine-tune or time target gene expression rather than mediating a strict ON/OFF switches.

Although the majority of the functionally characterized sRNAs repress translation and/or induce degradation of *cis*- or *trans*-encoded target mRNAs (Waters & Storz, 2009), some sRNAs can activate gene expression (Frohlich & Vogel, 2009). Remarkably, sRNAs that control expression of multiple target mRNAs can mediate both positive and negative regulation. For example, the RyhB sRNA in *E. coli* was shown to repress non-essential iron-using proteins such as the superoxide dismutase (*sodB*) and enzymes of the tricarboxylic acid cycle (*acnA* and *fumA*), whereas it promotes siderophore production (reviewed in Oglesby-Sherrouse & Murphy, 2013). Likewise, RNAIII not only activates translation of the α-hemolysin (*hla*) mRNA, but also inhibits translation of several targets, including *e.g.* the pleiotropic transcriptional factor Rot in *Staphylococcus aureus* (Novick *et al.*, 1993, Boisset *et al.*, 2007).

1.2.1. Repression of gene expression by sRNAs

The bacterial sRNAs characterized to date mainly bind directly within or adjacent to the region required for translation initiation, *i.e.* the Shine-Dalgarno sequence and translational start codon (Figure 1.1 A). Sequestration and masking of the ribosome binding site (RBS) and/or start codon through sRNAs prevents binding of the 30S ribosomal subunit and thus, formation of the translation initiation complex. Consequently, this results in reduced protein synthesis. Likewise, sRNA binding to an upstream ORF within the 5' UTR of an mRNA has been shown to not only mediate repression of the upstream ORF/short peptide, but also of the translationally coupled, downstream encoded ORF(s) (Vecerek *et al.*, 2007). A large number of *trans*-encoded sRNAs require the RNA chaperone Hfq for target gene regulation (Vogel & Luisi, 2011).

Small RNA-mRNA duplexes are frequently subjected to degradation by diverse ribonucleases, including the dominant catalyst of general mRNA turnover in Gammaproteobacteria, endoribonuclease E (RNase E, reviewed in Saramago *et al.*, 2014). Although translational repression through sRNA-mediated sequestration of the RBS can be sufficient for gene silencing (Morita *et al.*, 2006), the subsequent degradation of the sRNA-mRNA complex provides an efficient mechanism to make gene silencing irreversible. Two pathways are proposed by which sRNA base-pairing to target mRNAs trigger RNase E-dependent transcript degradation in enterobacteria (Figure 1.1 B, reviewed in Lalaouna *et al.*, 2013). First, target mRNAs might become more sensitive to RNase E attacks after binding to sRNAs as a result of the loss of protection conferred by translating ribosomes (Figure 1.1 B, *upper panel*). Secondly, following translation inhibition, an active recruitment of RNase E on the target mRNA triggers formation of a sRNA/Hfq/RNase E complex
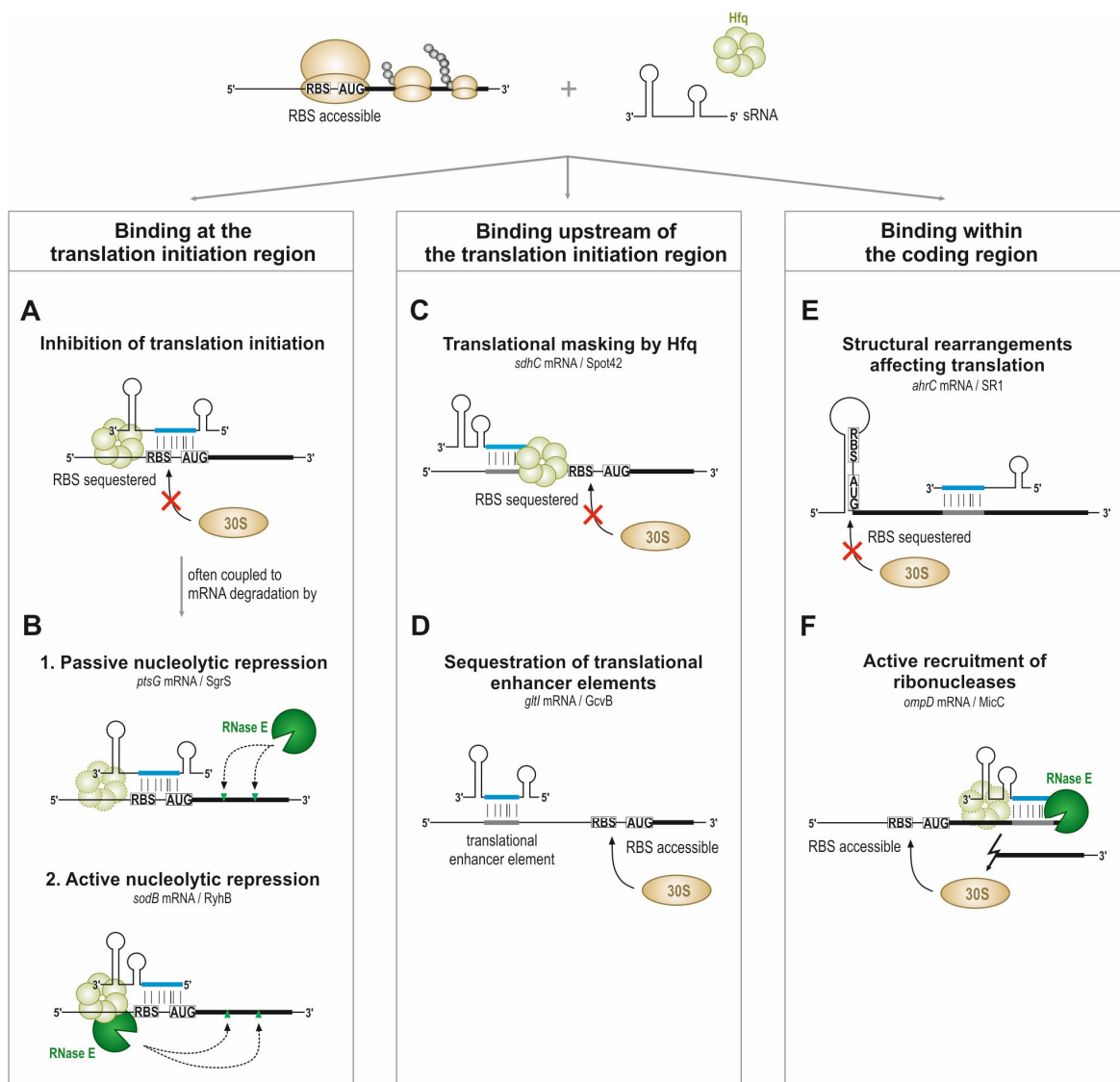
Figure 1.1: Mechanisms of sRNA-mediated repression of gene expression. (A) Most of the characterized sRNAs bind within or adjacent to the region relevant for translation initiation, *i.e.* ribosome binding site (RBS) and translational start codon (AUG), and block entry of the 30S ribosomal subunit on target mRNAs. Various *trans*-encoded sRNAs require the major RNA chaperone Hfq for stability and target mRNA binding. Inhibition of translation is often coupled to target mRNA degradation. **(B)** (*Upper panel*) SgrS interferes with translation initiation of the *ptsG* mRNA, which in turn, becomes more sensitive to degradation by RNase E due to the loss of protection from translating ribosomes. (*Lower panel*) Binding of RyhB to the RBS of the *sodB* mRNA interferes with translation initiation and actively recruits RNase E for target mRNA degradation. **(C)** Spot42 binds too far upstream of the *sdhC* RBS to interfere with translation initiation. However, concomitant recruitment of Hfq and redirection of the RNA chaperone close to the RBS results in inhibition of ribosome binding. **(D)** GcvB inhibits translation of the *gltI* mRNA by binding and sequestration of translation enhancer elements within the 5' UTR (upstream of RBS). **(E)** SR1 prevents *ahrC* mRNA translation initiation by the induction of structural rearrangements downstream of the RBS (binding in coding region). **(F)** MicC binds to the *ompD* mRNA deep in the coding region and recruits RNase E and thus, induces nucleolytic cleavage in the target mRNA.

that favors ribonucleolytic attacks and leads to distal cleavages in the coding region of the target mRNAs (Figure 1.1 B, *lower panel*).

Binding of sRNAs far upstream or downstream of the translation initiation region can also repress translation of the target mRNA. For example, base-pairing of the *E. coli* Spot42 sRNA to the *sdhCDAB* mRNA far upstream of the RBS results in recruitment of the RNA chaperone Hfq, which in turn binds at an A/U-rich region in close vicinity of the *sdhC* RBS and thereby inhibits translation initiation (Figure 1.1 C, Desnoyers & Masse, 2012). In contrast to the canonical model in which Hfq recruits an sRNA that blocks translation initiation (Figure 1.1 A), in this case Hfq itself is directly involved in the translational repression of the target mRNA, and the sRNA acts only as a recruitment factor for the RNA chaperone (Figure 1.1 C). In addition, the GcvB and IstR-1 sRNAs have been reported to repress translation of their target mRNAs by binding upstream of the RBS either through sequestration of translational enhancer elements (Figure 1.1 D, Sharma *et al.*, 2007) or through masking of a ribosome stand-by site (Darfeuille *et al.*, 2007).

To date, the only known example of an sRNA that inhibits translation initiation through binding within the coding region is the SR1 sRNA from *Bacillus subtilis*. Base-pairing of SR1 about 100-nt downstream of the AUG in the *ahrC* mRNA induces structural alterations not only in the complementary regions, but also in regions adjacent to the *ahrC* RBS, resulting in inhibition of translation initiation (Figure 1.1 E, Heidrich *et al.*, 2007). There is growing evidence that multiple bacterial sRNAs can also directly induce target mRNA degradation without interfering with translation initiation or elongation. The Hfq-dependent MicC sRNA in *Salmonella,* for instance, binds deep within the coding region of the *ompD* mRNA and promotes rapid mRNA decay by recruiting RNase E, which cleaves the mRNA immediately downstream of the MicC binding site and thus, repress expression of the porin OmpD (Figure 1.1 F, Pfeiffer *et al.*, 2009).

Apart from sRNA-mediated target mRNA regulation at the posttranscriptional level, sRNAs and the RNA chaperone Hfq have also been shown to regulate gene expression co-transcriptionally, *e.g.* by induction or interference with rho-dependent transcription termination (Bossi *et al.*, 2012, Rabhi *et al.*, 2011).

1.2.2. sRNA-mediated activation of gene expression

Small RNAs have also been shown to activate genes by a variety of direct or indirect mechanisms (reviewed in Frohlich & Vogel, 2009). For example, several sRNAs act as direct translational activators by the so-called "anti-antisense" mechanism (Figure 1.2 A). In these cases, translation of the target mRNA is intrinsically inhibited by an RNA secondary structure (*e.g.* hairpin, stem loop) in the 5' UTR of the mRNA that occludes the RBS and/or

start codon. Activating sRNAs liberate the occluded RBS, usually by pairing with the anti-SD or anti-AUG sequences of the inhibitory structure. As translation and ribosome occupancy increase, ribonucleases might stand a lesser chance for cleavage, whereby the stability and steady-state levels of the activated mRNA might also increase. First proposed for *S. aureus* RNAIII (Novick *et al.*, 1993), this regulatory mechanism was only brought to a comprehensive understanding through studies on the *rpoS* mRNA, which is translationally activated by three sRNAs (ArcZ, DsrA and RprA) in *E. coli* and *Salmonella* (*e.g.* Mandin & Gottesman, 2010, Papenfort *et al.*, 2009). Similar to sRNA-mediated repression of gene expression, several activating sRNAs require Hfq for stability and function as well (Soper *et al.*, 2010).

In addition to translational activation, sRNA-dependent mRNA stability control is a crucial element of activation by *trans*-encoded sRNAs. For example, binding of the *Clostridium perfringens* VR sRNA to the 5' UTR of the *colA* mRNA induces an endonucleolytic cleavage event upstream of the translation initiation region that results in (I) liberation of the formerly sequestered *colA* RBS, hence permitting ribosome access similar to the "anti-antisense" mechanism, and (II) formation of an alternative, shorter hairpin structure that promotes *colA* mRNA stabilization (Figure 1.2 B, Obana *et al.*, 2010). Similarly, FasX sRNA from *Streptococcus pyrogenes* was shown to stabilize *ska* mRNA by binding upstream of the *ska* RBS (Figure 1.2 C, Ramirez-Pena *et al.*, 2010). In the absence of FasX, *ska* mRNA is rapidly degraded. FasX pairing makes the formerly single-stranded 5' UTR of the *ska* mRNA double-stranded, *e.g.* as if by generating a 5' hairpin, and thus, prevents attacks by RNases.

Cis-encoded antisense RNAs have also been shown to promote target mRNA stability by inducing processing events (Figure 1.2 D). For example, the GadY sRNA in *E. coli* is encoded on the opposite strand in the IGR between *gadX* and *gadW*, which are transcribed as a rather unstable dicistronic mRNA. Base-pairing of GadY to the *gadXW* mRNA initiates processing mediated in part by RNase III (and other RNases) and thereby promotes stability of both the *gadW* mRNA and *gadX* transcript (Opdyke *et al.*, 2004, Opdyke *et al.*, 2011).

Apart from activation through direct base-pairing, sRNAs can also indirectly promote gene expression. For example, induction of a "trap-mRNA" (mRNA with sRNA binding site) could lead to selective sequestration and degradation of a regulatory sRNA, thereby abolishing the sRNA-based silencing of its cognate target mRNA (*e.g.* MicM and *ybfMN/chbBCARFG* mRNA, Overgaard *et al.*, 2009). In addition, GlmY sRNA in *E. coli* was shown to indirectly activate *glmS* mRNA (glucosamine-6-phosphate synthase) by protecting GlmZ sRNA from degradation (reviewed in Gopel *et al.*, 2014). GlmZ promotes *glmS* mRNA translation by "anti-antisense" mechanism. Its levels are controlled at the level of RNA decay
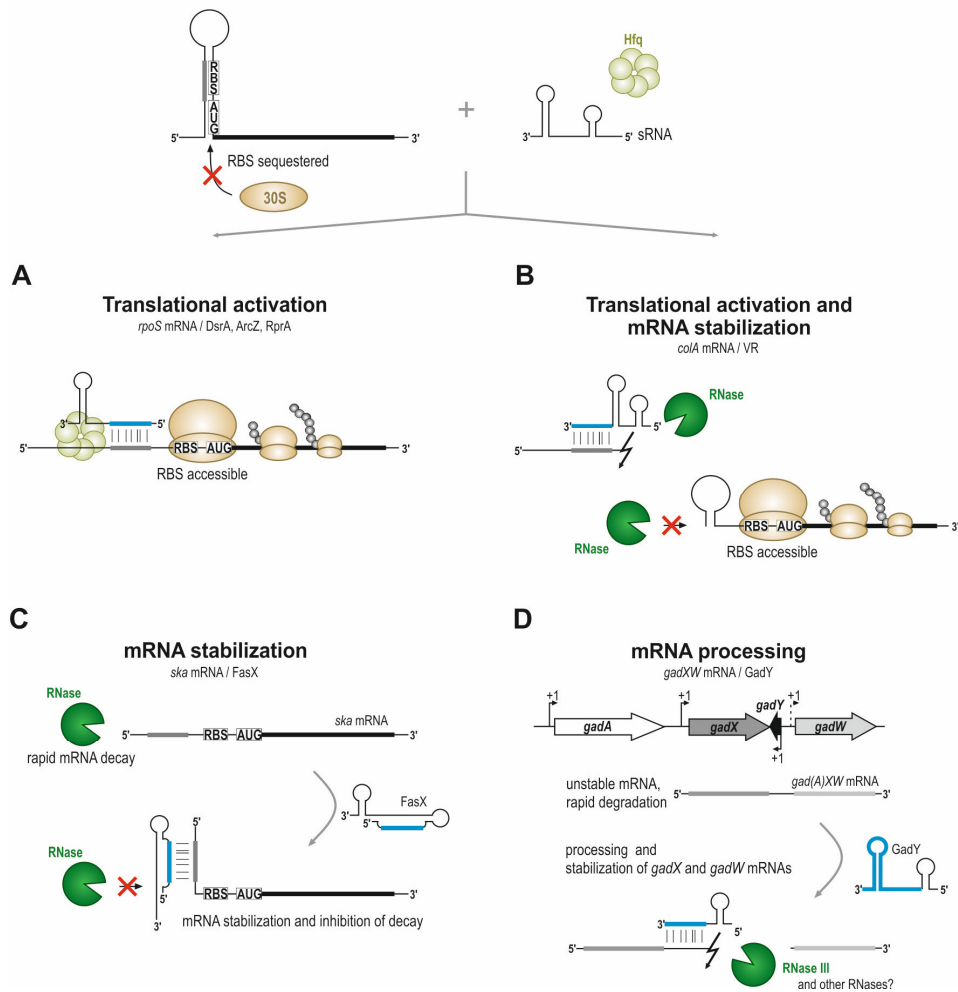
Figure 1.2: Mechanisms of sRNA-mediated activation of gene expression. (A) Many sRNAs directly promote translation of target mRNAs by an "anti-antisense" mechanism. An intrinsic stem-loop structure occludes the RBS and thereby inhibits translation initiation of the *rpoS* mRNA. Small RNAs (*e.g.* DsrA, ArcZ, RprA) function as structural competitors that bind to the anti-SD/anti-AUG sequence in the inhibitory stem and thereby liberate the formerly sequestered RBS, allowing ribosome association. **(B)** VR base-pairing to the 5' UTR of the *colA* mRNA induces processing upstream of the *colA* RBS. This leads to a new mRNA species with an accessible RBS and stabilizing short hairpin structure. **(C)** When FasX sRNA is absent, *ska* mRNA is rapidly degraded. FasX binding to the extreme 5' end of the *ska* mRNA generates a 5' proximal stabilizing structure that protects *ska* mRNA from nucleolytic cleavage. **(D)** GadY sRNA, which is encoded on the opposite strand in the *gadX–gadW* IGR, interacts with the 3' UTR of *gadX* to stabilize both transcripts.

by the RNase adaptor protein RapZ and RNase E. Under specific conditions, highly similar GlmY accumulates and sequesters RapZ by an RNA mimicry mechanism, increasing GlmZ levels and thus, *glmS* expression.

1.2.3. Posttranscriptional regulation of operons by sRNAs

Bacterial genes are organized into operons, or clusters of co-regulated genes. These often functionally coupled/related genes are regulated in a manner that they are all turned ON or OFF together. Analogous to transcription factors, several sRNAs have been described to regulate genes encoded in operons (reviewed in Balasubramanian & Vanderpool, 2013). Small RNA activity has been shown either to result in simultaneous activation or repression of all genes in the operon (coordinate regulation), or to uncouple gene expression of single genes within polycistrons, resulting in discoordinate regulation of the operon. Coordinate posttranscriptional regulation of polycistronic mRNAs can be achieved by either a single sRNA-mRNA base-pairing interaction that regulates translation of an upstream cistron, resulting in polar effects on downstream genes, or sRNA base-pairing to multiple translation initiation sites. For example, the enterobacterial SgrS sRNA was shown to coordinately regulate *manXYZ* mRNA, the products of which are involved in uptake of several sugars, by binding to two distinct regions on the polycistronic mRNA, preventing translation of either *manX* or *manYZ*, and promoting cooperative mRNA degradation (Rice & Vanderpool, 2011, Rice *et al.*, 2012). In contrast, discoordinate regulation can occur when sRNA-mRNA base-pairing affects translation of one cistron in the operon, but has no effect on the others. The iron-responsive RyhB sRNA, for instance, has been shown to mediate discoordinate regulation of the *iscRSUA* operon encoding for a transcription factor (*iscR*) and genes that are involved in the synthesis and assembly of the Fe-S clusters present in many *E. coli* proteins. While antisense base-pairing of RyhB to the *iscR-iscS* intergenic region inhibits translation and triggers selective mRNA degradation of *icsSUA, iscR* mRNA is protected from nucleolytic cleavage by a secondary stem-loop structure at its 3' end (Desnoyers *et al.*, 2009). This type of regulation is useful when only selected products of an operon need to be synthesized under particular conditions, while other products are needed constitutively.

1.2.4. sRNAs involved in bacterial pathogenesis

Genome-wide expression analyses under various infection-related conditions have begun to identify potential virulence associated sRNAs in diverse pathogens (Yan *et al.*, 2013, Kroger *et al.*, 2013). *In-vitro* studies with *Salmonella*, pathogenic *E. coli* as well as *S. aureus* suggested that various sRNAs contribute to bacterial pathogenicity as they are either encoded on pathogenicity islands itself and/or regulate expression of virulence factors or virulence-associated genes (reviewed in Caldelari *et al.*, 2013, Papenfort & Vogel, 2014). That an sRNA-mRNA interaction can be decisive for the regulation of virulence factors and/or adaptation to the host has been first demonstrated for the RNAIII, which controls the switch between expression of surface proteins and secreted toxins in *S. aureus* (Novick *et al.*,

1993). Other examples include the *Salmonella* pathogenicity island 1 (SP-1)-encoded InvR (Pfeiffer *et al.*, 2007) and DapZ (Chao *et al.*, 2012) sRNAs, the lipopolysaccharide (LPS)-controlling MgrR sRNA (Moon *et al.*, 2013) and the virulence effector-targeting SgrS sRNA (Papenfort *et al.*, 2013). However, inhibition of many of these well-characterized riboregulators often do not lead to a pronounced growth defect in *in-vivo* models of infection (Barquist *et al.*, 2013, Chaudhuri *et al.*, 2013).

1.3. Auxiliary factors required for sRNA regulation

Non-coding RNAs act in concert with a variety of RNA-binding proteins (RBPs) and ribonucleases (RNases) to elicit posttranscriptional gene expression control. The Sm-like RNA chaperone Hfq is considered to be a key player in sRNA-mediated regulation of target mRNAs in many bacteria (reviewed in Vogel & Luisi, 2011). This small hexameric RNA-binding protein is widely conserved in bacteria, *i.e.* Hfq is encoded in about 50 % of all eubacterial species (Chao & Vogel, 2010). Hfq not only protects bound sRNAs from endonucleolytic cleavage prior to target recognition, but also facilitates sRNA-mRNA interactions and/or recruits specific RNases for target mRNA degradation (Wagner, 2013). In addition, Hfq can regulate gene expression in the absence of sRNAs, *e.g.* by influencing polyadenylation or translation of mRNAs (Vogel & Luisi, 2011). According to its important role as a global posttranscriptional regulator, deletion of *hfq* causes pleiotropic phenotypes including impaired stress regulation as well as reduced virulence in several bacterial pathogens (Chao & Vogel, 2010). In contrast to many Gram-negative bacteria, Hfq seems to be dispensable for sRNA-mediated regulation in Gram-positive bacteria. Although many of them contain an Hfq homolog, deletion of *hfq* in *S. aureus* and *B. subtilis*, for instance, has no obvious phenotype and is apparently not required for sRNA-mRNA interactions and/or sRNA stability (Boisset *et al.*, 2007, Geissmann *et al.*, 2009, Hammerle *et al.*, 2014).

Several other RBPs, such as the previously mentioned translational regulator CsrA and the RNase adaptor protein RapZ, are involved in posttranscriptional gene expression control by sRNAs. Furthermore, the recent findings that additional proteins with so far unknown functions or those unrelated to RNA metabolism are involved in posttranscriptional regulation indicate that we are far away from knowing all RBPs and their roles in bacteria (*e.g.* YbeY and RodZ; Pandey *et al.*, 2011, Mitobe *et al.*, 2011).

In addition to RBPs, RNases are involved in sRNA-mediated regulation in bacteria and have been shown to influence the processing and/or turnover of these molecules (Viegas *et al.*, 2007). In Gammaproteobacteria, RNase E is the dominant catalyst of general mRNA turnover, but is also involved in the processing of ribosomal and transfer RNAs (Carpousis *et al.*, 2009). This single-strand specific endoribonuclease has a preference for

5' monophosphate termini and A/U-rich sequences of RNAs, and initiates the decay of the majority of mRNAs. RNase E is not only a crucial component in sRNA-induced decay of target mRNAs (reviewed in Lalaouna *et al.*, 2013), but also contributes to the turnover of sRNAs (Schmidt & Delihas, 1995). Generally, RNase E is part of a multiprotein complex called RNA degradosome, which is composed of a 3'-exoribonuclease (PNPase), a DEAD-box RNA helicase (RhlB), a glycolytic enzyme (enolase), and other proteins (*e.g.* Hfq), depending on physiological conditions (Bandyra *et al.*, 2013). RNase E is not ubiquitous in bacteria, but two functional RNase E orthologs, RNase J1 and RNase J2, have been identified in the RNA degradosome of Gram-positive bacteria (reviewed in Lehnik-Habrink *et al.*, 2012).

Besides RNase E, the double-strand specific RNase III was shown to degrade, mainly in concert with the sRNA RNAIII, several mRNAs encoding virulence factors in *S. aureus* (*e.g.* Boisset *et al.*, 2007). Likewise, RNase III is also important for sRNA-mediated target regulation in diverse Gram-negative bacteria. For example, the interaction between the sRNA IstR-1 and its target *tisAB* entails RNase III-dependent cleavage for inactivation of mRNA translation in *E. coli* (Vogel *et al.*, 2004). Moreover, the decay of the sRNA RyhB was shown to mainly depend on RNase III, in contrast to the RNase E-dependent degradation of the *sodB* target mRNA (Afonyushkin *et al.*, 2005).

1.4.   Methods to identify and study sRNA targets

RNA molecules that act as regulators have been first identified in *E. coli* more than three decades ago, long before the first microRNAs (miRNAs) and short interfering RNAs (siRNAs) were discovered in eukaryotes (Stougaard *et al.*, 1981, Tomizawa *et al.*, 1981). To date, more than 150 sRNAs are identified and validated in *E. coli* (Raghavan *et al.*, 2011), a number that is even exceeded by the nearly 300 sRNA candidates known in *Salmonella* (Kroger *et al.*, 2013). Especially, the development of high-throughput DNA sequencing technology and its application to massively parallel sequencing of cDNA molecules (derived from RNA) has provided a powerful method for both mapping and quantifying the transcriptional outputs of sequenced genomes (Wang *et al.*, 2009). This method, termed RNA sequencing (RNA-seq), has drastically improved the view of the extent and complexity of prokaryotic transcriptomes, including the discovery of a wealth of novel sRNAs in various bacterial species (Croucher & Thomson, 2010).

A full understanding of the physiological function of a given sRNA requires the identification of its cellular interaction partners. For this purpose, a variety of experimental and biocomputational approaches have been developed in the last years (reviewed in Sharma & Vogel, 2009). Generally, sRNA-dependent changes in gene expression are monitored by genetic screens or, most commonly, comparative transcriptome and proteome

analyses of sRNA overexpression and/or deletion mutants and the wildtype. Target regulation at the mRNA level can be detected by a variety of methods such as northern blot hybridization, primer extension, and quantitative real-time PCR (qRT-PCR) at the single gene level. As is now appreciated that many sRNAs act on multiple targets, global methods involving *e.g.* microarrays and RNA-seq are recommended for target mRNA identification. Investigations of deletion and/or constitutive overexpression of sRNAs are prone to report indirect regulation because of potential off-targets and pleiotropic effects on overall gene expression, especially if a transcriptional regulator is a direct target of an sRNA. Therefore, approaches using transient overexpression or pulse-expression of sRNAs from inducible promoters in combination with whole transcriptome analysis have been developed to identify direct targets and avoid downstream effects. In these studies it is assumed that direct targets are affected before indirect targets (in terms of time after induction; Sharma *et al.*, 2011, Masse *et al.*, 2005)*.* Nowadays, RNA-seq is the premier method for comprehensive as well as quantitative expression profiling on bacterial transcriptomes under various stress conditions or between mutant and wild-type strains (reviewed in Creecy & Conway, 2015). Although microarrays have revolutionized the study of transcriptomics and proved useful in determining (*e.g.* sRNA-mediated) gene expression profiles, RNA sequencing by comparison is more sensitive, provides absolute quantity levels, is not affected by on-chip sequence biases, and gives additional information on variations of gene expression levels or processing. Accordingly, RNA-seq and ribosome profiling, which is a recently developed strategy that enables monitoring of protein translation directly by deep sequencing of ribosome protected RNA fragments, so-called ribosome footprints (Ingolia *et al.*, 2009), have been successfully applied to examine sRNA-mediated changes in mRNA abundances and protein synthesis, respectively (Guo *et al.*, 2014).

In addition to ribosome profiling, most of the quantitative methods which have been successfully used to define the protein content of a given cell/sample, including SDS-PAGE (1D and 2D) and/or SILAC approaches (stable isotope labeling by amino acids in cell culture), can be used to identify sRNA-mediated changes in protein abundances on a global scale. Contradictory to ribosome profiling, these methods detect changes in steady state protein levels rather than providing information about differences in the actively translated mRNA fractions. At single gene level, western blot analysis is mainly used to detect sRNA-mediated protein alterations. Chromosomal epitope-tagging helps to detect regulated proteins when no specific antibody is available. Also, chromosomal or plasmid-borne reporter fusions, including mRNA fusions to *lacZ* (encoding for β-galactosidase), *gfp* (green fluorescent protein) or *luc* (luciferase), are commonly used to determine target regulation.

A convenient way to examine posttranscriptional target regulation is to uncouple expression of the sRNA and the mRNA from transcriptional control. Chromosomal promoter

exchanges as well as a *gfp*-based two plasmid system, in which both the sRNA and target mRNA fusions are under control of constitutive or inducible promoters (optionally), have been successfully applied to investigate posttranscriptional sRNA-mRNA interactions in *E. coli* (Urban & Vogel, 2007). However, for base-pairing RNAs, compensatory base-pair mutations are the monetary standard for the validation of direct sRNA-mRNA interactions. Mutations in either the interaction site of the sRNA or the target site should abolish regulation, whereas compensatory mutations in both sites should restore regulation. In addition, *in-vitro* experiments are frequently used to validate the interactions between two RNAs and provide insights into the mechanism of actions of sRNA-mediated regulation. These biochemical methods include gel-shift and structure probing assays with *in-vitro* transcribed RNAs, toeprinting assays that allow for investigating the formation of the translation initiation complex in absence or presence of an sRNA, and *in-vitro* translation assays with reconstituted ribosomes (reviewed in Sharma & Vogel, 2009).

Although the majority of methods focus or are broadly applicable for base-pairing antisense RNAs, several strategies have been developed for the global investigation of RNA-protein complexes. These include co-immunoprecipitation of epitope-tagged RBPs and identification of RNA binding partners by RNA sequencing (RIP-seq). For example, RIP-seq analysis of FLAG-tagged Hfq revealed various sRNAs and mRNAs that are specifically bound and/or stabilized by this RNA chaperone in *Salmonella* (Sittka *et al.*, 2008). More recently developed approaches, which use *in-vitro* and *in-vivo* ultra-violet (UV)-crosslinking of RNA-protein interactions and co-immunoprecipitation (CLIP) strategies, allow for the investigation of RNA-protein complexes at a very high resolution and specificity (Konig *et al.*, 2011). For an unbiased, genome-wide identification of novel RBPs, an orthogonal approach using aptamer-tagged RNAs and affinity chromatography can be used, in which recovered RNA-protein complexes are analyzed by mass spectrometry (Corcoran *et al.*, 2012b).

In addition to experimental approaches, several bioinformatics-based algorithms are available to predict putative sRNA targets (reviewed in Backofen & Hess, 2010, Wright *et al.*, 2014). In principle, most of the algorithms search for regions in both the mRNA and the sRNA that are complementary to each other. RNA secondary structures and thus, the accessibility of RNA interactions sites, potential co-folding of RNA duplexes, sequence conservation and position within the RNA molecule, *e.g.* close to the RBS in mRNA, are considered to predict potential target mRNAs. However, *in-silico* identification of sRNA targets appears to be challenging as the partial and incomplete sequence complementarity of sRNA-mRNA interactions is difficult to evaluate. In addition, the high heterogeneity of sRNAs in length, sequence and structure as well as their mechanisms of target recognition and regulation complicate sRNA target prediction.

1.5. The major human pathogen *Helicobacter pylori* as a model organism

While most of the human body is colonized by highly complex microbial communities, the stomach represents an exceptional niche. Due to its harsh and acidic milieu, it is a challenging barrier to orally ingested microorganisms and thus, was long thought to be sterile. However, the discovery of the Gram-negative and microaerophilic Epsilonproteobacterium *Helicobacter pylori* caused a change in the understanding of the human stomach as an ecological niche for bacteria (Marshall & Warren, 1984). Uniquely, *H. pylori* thrives in the hostile and acidic environment of the human stomach where it can lead to peptic ulcer disease and chronic superficial gastritis, which are characterized by a distinct breach or inflammation of the gastric mucosa, respectively (Cover & Blaser, 2009). It is considered to be a major human pathogen as more than half of the human's population are infected by *H. pylori*. Usually acquired during childhood, *H. pylori* gastric infections generally persist life-long in the absence of antibiotic treatment. Decades of colonization by this pathogen have been associated with the development of gastric adenocarcinoma and gastric mucosa-associated lymphoid tissue (MALT) lymphoma. Therefore, *H. pylori* was the first bacterium classified as a class I carcinogen by the World's Health Organization (WHO) in 1994 (International Agency for Research on Cancers, 1994). While about 80 % of the infections remain asymptomatic, up to 20 % of *H. pylori* carriers develop peptic ulcers and only 1-2 % develop precancerous lesions and ultimately, gastric cancer. The different clinical outcomes have been attributed to the high genetic diversity of *H. pylori*, but also to host predisposition such as age, nutrition and other life style factors (Haley & Gaddy, 2015).

Most clinical isolates and *in-vitro* cultured *H. pylori* cells are spiral-shaped and possess a unipolar bundle of two to six flagella. However, a conversion from the spiral to coccoid morphology can be observed under diverse stress conditions including nutrient starvation, oxidative or acidic stress, and antibiotic treatment (Cellini, 2014). This might represent an adaptation strategy for bacterial survival and immune evasion.

The clinical importance of *H. pylori* and resulting studies of its pathogenicity have contributed much to the understanding of its virulence mechanisms and turned *H. pylori* into a model organism for bacterial pathogenesis (Salama *et al.*, 2013). *H. pylori* is of major importance as a bacterial inducer of gastric cancer, but it also represents a system to study characteristics of chronic and persistent infections, and the associated immune modulatory mechanisms (Monack, 2013). An arsenal of *H. pylori* virulence factors has been identified that are important to avoid/neutralize the bactericidal acid in the gastric lumen and survive near to, attach to and subvert the human gastric epithelium as well as the immune system (Figure 1.3, Fischer *et al.*, 2009). In addition, the great level of genetic heterogeneity among diverse *H. pylori* strains and clinical isolates was shown to contribute to persistent infection and immune evasion (Suerbaum & Josenhans, 2007).

1.5.1. Colonization of the gastric mucosa – *H. pylori* adaptation to the human host

Because *H. pylori* is a neutrophilic bacterium, its ability to survive the acidic conditions of the gastric lumen is mainly dependent on the production of a very potent urease enzyme, which converts urea into ammonia and CO_2. The generation of ammonia provides both acid-neutralizing and acid-buffering and thus, enables *H. pylori* to raise the pH in its microenvironment and keep its periplasmic pH near neutrality (Stingl & De Reuse, 2005). Initial colonization of the gastric niche does not only require *H. pylori* to be acid-resistant, but also to rapidly orient itself towards the mucus layer of the gastric epithelium (primary site of infection), in which the pH is close to neutral. Therefore, flagellar-based motility and chemotactic behavior of *Helicobacter* are crucial for host colonization and virulence (Spohn & Scarlato, 2001). The major chemotactic signal in the mucus layer is the pH gradient (Schreiber *et al.*, 2004). *H. pylori* senses chemoattractants or -repellents by utilizing four methyl-accepting chemotaxis proteins, *i.e.* membrane-spanning TlpA, TlpB, TlpC and cytoplasmic TlpD, that control the direction of the flagellar motor rotation, and thus, the swimming behavior of *H. pylori* (reviewed in Lertsethtakarn *et al.*, 2011).

Adherence to the gastric mucosa and surface of epithelial cells is another crucial step required for both *H. pylori* colonization and pathogenesis (Oleastro & Menard, 2013), but it also plays an important role for protection against the acidic pH and exfoliation (Smolka & Backert, 2012). The interaction with the mucus layer and/or host cells is modulated through several paralogous outer membrane proteins (OMPs), which are often expressed phase-variably and may lead to differential outcome in infected patients. Individually, none of the known adhesins is essential for colonization; however, their expression enhances the ability of *H. pylori* to adhere and thus, colonize the stomach (Yamaoka *et al.*, 2002).

Besides OMPs, *H. pylori* adhesion to the gastric mucus and extracellular matrix proteins is also mediated by lipopolysaccharides (Valkonen *et al.*, 1997, Reeves *et al.*, 2008). This family of phosphorylated glycolipids is found in the outer leaflet of the outer membrane of Gram-negative bacteria. LPS is essential for the physical integrity and function of the bacterial membrane (Holst *et al.*, 1996) and – as the main surface antigen of Gram-negative bacteria – also play an important role in the interaction with the environment and/or host (Raetz & Whitfield, 2002). In host-pathogen interactions, LPS structures are potent immunomodulating and immunostimulating compounds, which possess a broad spectrum of endotoxic properties and thereby contribute to the pathogenicity of bacterial pathogens. The biological activity of bacterial LPS is mostly marked by its capability to activate specific host cell receptors, *e.g.* Toll-like receptor TLR4, which in turn triggers the host immune response (reviewed in Janssens & Beyaert, 2003). Although *H. pylori* LPS also has properties similar to those of other Gram-negative bacteria, it possesses significantly lower
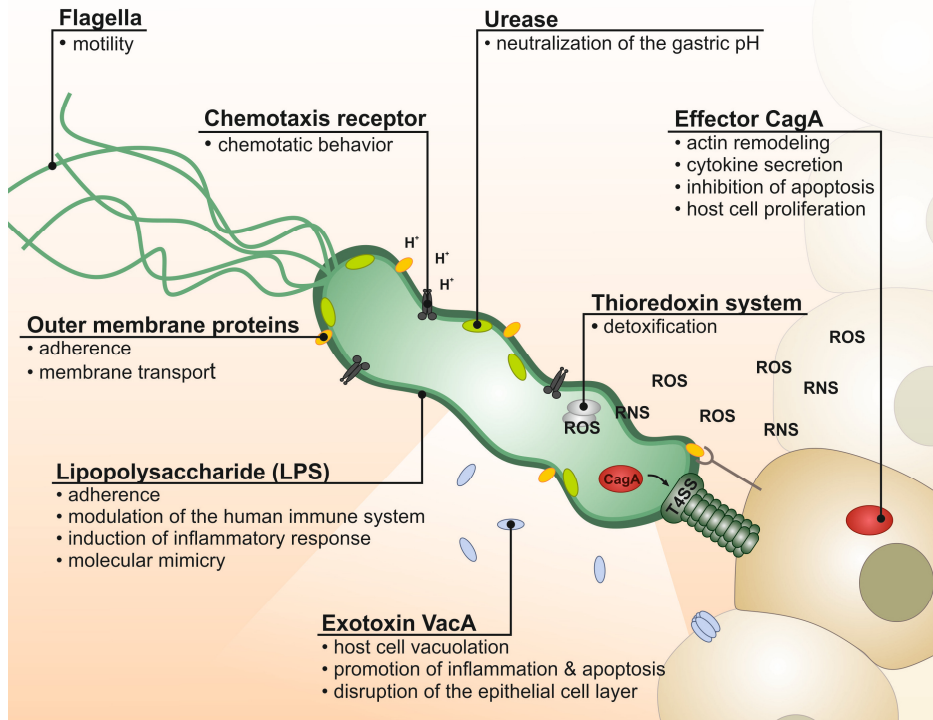
Figure 1.3: *H. pylori* **colonization and persistence factors.** During initial colonization of the gastric lumen, urease-dependent ammonia production locally raises the pH and promotes bacterial survival. Chemotactic behavior and flagellar motility are crucial for the orientation of *H. pylori* towards the mucus layer of the gastric epithelial cells. Outer membrane proteins and LPS structures mediate adhesion to the mucus layer and/or host cells. Cell-associated bacteria alter gastric host cell behavior through the vacuolating cytotoxin VacA and the cytotoxin-associated gene CagA. VacA is secreted by a type V autotransporter. The major effector protein CagA is injected into host cells by a type IV secretion system (T4SS). Multiple proteins, including the components of the thioredoxin oxidant system, are essential for *H. pylori* to combat reactive oxygen (ROS) and nitrogen species (RNS).

endotoxic/immunological activity (up to 1,000 fold lower) when compared to the enterobacterial LPS. Structural variations and modifications of different LPS components have been implicated to reduce the immunological response (*e.g.* through escape of recognition by TLR4), hence aiding persistence of *H. pylori* and the development of chronic infections (Moran, 2001). Another important contribution of LPS to *H. pylori* pathogenesis concerns the molecular mimicry of carbohydrate structures present on human epithelial and blood cells, so-called Lewis antigens. Phase-variable expression of Lewis antigens gives rise to antigenic diversity and thus, provides an effective strategy to escape from the humoral immune response of the human host (Appelmelk & Vandenbroucke-Grauls, 2000).

Besides factors that contribute to immune evasion, *H. pylori* actively manipulates host cells and promotes its own persistence through the activity of two secreted toxins, the vacuolating cytotoxin A (VacA) and the cytotoxin-associated gene A (CagA) (Palframan *et al.*, 2012, Backert *et al.*, 2010). VacA is a multi-functional pore-forming toxin, which is secreted

by a type V autotransporter mechanism and leads to host cell vacuolation, disruption of cell junctions, promotion of apoptosis, inflammation, and inhibition of proliferation. The immunodominant antigen CagA is so far the only known effector protein of *H. pylori*. It is secreted by a type IV secretion system (T4SS), which is encoded together with *cagA* on a about 38-kb multi-operon locus (*cag* pathogenicity island) in the chromosome. Upon translocation into host cells, CagA induces a variety of changes in host signaling cascades, which in turn lead to alterations in *e.g.* cell polarity, cell structure as well as motility or proliferation.

During persistent infection, *H. pylori* induces a chronic inflammatory host response. Multiple detoxifying systems including the catalase KatA, superoxide dismutase SOD and arginase RocF protect *Helicobacter* against reactive oxygen (ROS) and nitrogen species (RNS) imposed by the host immune system (Wang *et al.*, 2006a). In addition, *H. pylori* depends on a thioredoxin (Trx) oxidant system, which mediates the reduction of ROS-detoxifying enzymes and hence, is important for colonization and pathogenicity (Lu & Holmgren, 2014).

1.5.2. Genetic diversity and phase variation contribute to *Helicobacter* pathogenicity

Persistent colonization of the human host by *H. pylori* is facilitated through its extensive genetic diversity due to an elevated mutation rate, impaired DNA repair system, horizontal gene transfer, frequent recombination events, and phase variation (Suerbaum & Josenhans, 2007, Salama *et al.*, 2013). Phase variation represents a high-frequent and stochastic mechanism of reversible genotype switching, usually "all-or-none" or "ON/OFF", that contributes to phenotypic heterogeneity within bacterial populations (van der Woude, 2011). Besides a variety of mechanisms including conservative or site-specific recombination, insertion and extension events or epigenetic changes through DNA methylation, phase variation can occur due to highly mutable DNA sequences. These so-called contingency loci are often associated with genes involved in LPS biosynthesis, surface exposed proteins, pili as well as flagella, and DNA restriction or modification (van der Woude & Baumler, 2004). In addition to deletions, gene conversions and point mutations, variable simple sequence repeats (SSRs) have been shown to be the major source of phase variation in these loci (Moxon *et al.*, 2006, Zhou *et al.*, 2014).

Phase variation of SSRs is in most cases independent of recombination and occurs during DNA replication or DNA repair through slipped-strand mispairing and polymerase slippage, leading to repeat length variation (van der Woude & Baumler, 2004). Depending on their location, length variation of SSRs can either affect translation through the introduction of frame-shift mutations within coding regions leading to premature

translation termination or altered C-termini of proteins (intragenic SSRs) or influence transcription by changing the spacing of promoter elements or transcription factor-binding sites (intergenic SSRs, Figure 1.4). Whereas the mechanisms and roles of SSRs on gene regulation at the DNA level are established, effects on mRNA stability or posttranscriptional control are less understood.

The impact of phase variation on the success of a bacterial population depends on two factors: (I) on the role/function of the gene that phase-varies, and (II) on the growth environment, which provides a certain selective pressure (Bayliss, 2009). In host-pathogen interactions, phase variation is implicated in mediating immune evasion and modulation, host colonization and biofilm formation (van der Woude & Baumler, 2004). Accordingly, SSRs have also been described to affect *H. pylori* virulence and long-term adaption to the gastric niche (*e.g.* Appelmelk *et al.*, 1999, Solnick *et al.*, 2004).
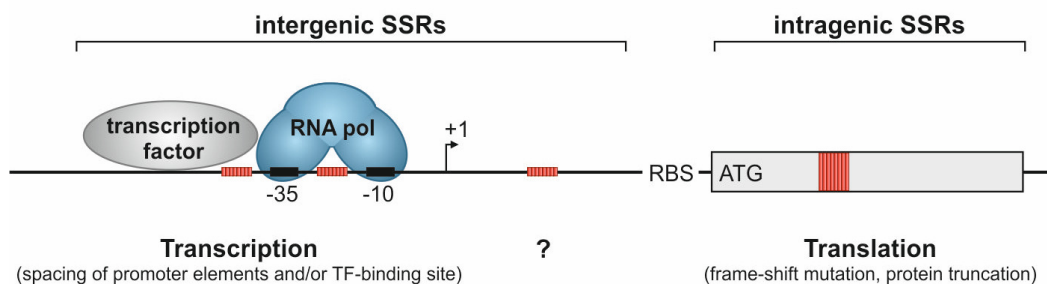
Figure 1.4: Phase variation occurs due to slipped-strand mispairing at simple sequence repeats. Schematic representation of the location, relative to a gene, at which SSRs (red) such as homopolymeric or tandem repeats can cause phase variation. Length variation of intragenic SSRs, which are located within the ORFs (gray box), can affect translation due to frame-shift mutation or protein truncation. Intergenic SSRs are located either in the 5' UTRs (+1 marks transcriptional start site (TSS)), between promoter elements (-10 and -35 box) and/or promoter elements and transcription factor (TF)-binding sites. Length variation of intergenic SSRs influences binding of the RNA polymerase (RNA pol) and/or TF by changing the spacing between the promoter elements or TF-binding sites. Intergenic SSRs in the 5' UTR might cause phase variation by yet unknown mechanisms. RBS – ribosome binding site; ATG – translational start site.

Regarding the presence of simple sequence or tandem repeats within open reading frames and/or promoter elements, more than 50 phase-variable gene candidates have been identified in *H. pylori* (Tomb *et al.*, 1997, Salaun *et al.*, 2004). For example, different lengths of poly-C tracts within 5' coding regions of genes encoding for diverse fucosyltransferases, which are important for the generation of Lewis epitopes, allow *H. pylori* to produce diverse LPS antigens (Appelmelk & Vandenbroucke-Grauls, 2000). In addition, expression of several *H. pylori* outer membrane proteins, including the major adhesins BabA and SabA, can be

switched ON and OFF by phase-variable dinucleotide repeats (CT-tracts) within their coding regions (Solnick *et al.*, 2004, Yamaoka, 2008). Likewise, length variation of a homopolymeric C-repeat in the coding region of the flagellar basal body gene *fliP* was shown to regulate motility in *H. pylori* strain 26695 (Josenhans *et al.*, 2000). Contrary to the ON/OFF switch through intragenic SSRs, intergenic SSRs have been shown to gradually/stepwise affect transcription of various genes in *Helicobacter*. For example, length variation of a promoter-proximal homopolymeric T-tract has been shown to result in multiphasic expression of SabA, thereby altering *H. pylori* adherence to the human gastric mucosa (Aberg *et al.*, 2014).

Phase variation can not only affect expression of single genes, but also multiples genes through SSR-mediated ON/OFF switching of DNA methyltransferases of restriction and modification (R-M) systems, so referred as the "phase-varion" (Srikhanta *et al.*, 2010). In *Helicobacter*, phase-variable DNA methyltransferases have been shown to be important for host colonization by randomizing virulence factor expression through global changes in methylation status of multiple promoter elements (Gauntlett *et al.*, 2014).

1.5.3. Regulation of gene expression in *H. pylori*

The gastric niche can be regarded as a hostile and yet dynamic environment, in which *H. pylori* encounters a variety of diverse environmental stressors. These include pH fluctuations within the gastric lumen (1-2 to 6.5, Sachs *et al.*, 2003), osmotic stress, nutrient limitations, rapidly changing availabilities of metal ions, and oxidative stress caused by the immune system of the host. Thus, despite being adapted to its host, *Helicobacter* requires regulatory networks to respond and adapt to different conditions during colonization by changing its gene expression.

H. pylori strain 26695 was among the first bacteria whose genomes were sequenced (Tomb *et al.*, 1997), which in turn, contributed much to the understanding of its genetic repertoire of virulence and regulatory genes. The small *H. pylori* genome has a size of about 1.7 megabases, which is only one-third of the *E. coli* genome (Blattner *et al.*, 1997), and possibly reflects the high degree of specialization of this bacterium. Only a small number of transcriptional regulators were identified in *H. pylori*, including three sigma factors (RpoD, FliA and RpoN), four complete two-component systems (TCSs) as well as two orphan response regulators, and a few other genes with regulatory functions (de Vries *et al.*, 2001). Interestingly, *H. pylori* lacks homologs of stress-related sigma factors (RpoH, RpoS) and additional key regulators of other Gram-negative bacteria such as OxyR, Crp, and FNR, which are involved in gene regulation in response to oxidative stress, carbon starvation and anaerobic conditions, respectively (Tomb *et al.*, 1997). The small genome size and the absence of important regulatory systems known from other bacteria suggest two

possibilities. Either, regulation is only rudimentarily present or yet undefined genes or components, such as non-coding RNAs, play a role in the adaptation of *H. pylori* to the environment or its human host.

In the past decade, enormous efforts have been undertaken to elucidate and understand the regulatory networks of *H. pylori*. Genome-wide changes in *H. pylori* gene expression induced by environmental stimuli or mutation of putative transcriptional regulators have been successfully analyzed in various transcriptome studies, mostly using microarrays in combination with qRT-PCR, primer extension, *in-vitro* promoter binding, and gel-shift assays (reviewed in Josenhans *et al.*, 2007). These studies have provided fundamental insights into gene expression control at the transcriptional level and revealed that acidic milieu of the human stomach, for instance, represents one of the most important triggers for gene expression in *H. pylori* (Merrell *et al.*, 2003a, Wen *et al.*, 2003). The majority of genes involved in acid acclimatization are controlled by the acid-sensing ArsRS TCS, which is composed of a sensory histidine kinase, ArsS, and its cognate response regulator, ArsR (reviewed in Pflock *et al.*, 2006a). In response to acid, ArsS promotes phosphorylation of ArsR, which regulates expression of target genes such as urease.

Besides two component systems, *H. pylori* possesses two orphan response regulators with atypical receiver domains, namely HP1043 (HsrA) and HP1021. These transcription factors lack a cognate sensor kinase and do not require phosphorylation for their activation (Schar *et al.*, 2005). Both regulators serve fundamental functions in *H. pylori* as HP1043 is essential and deletion of HP1021 causes severe growth defects (Beier & Frank, 2000). Although the target genes of both HP1043 and HP1021 are largely unknown, recent studies indicated that they might be involved in acetone metabolism, chromosome replication, growth stage regulation and oxidative stress management (Olekhnovich *et al.*, 2013, Olekhnovich *et al.*, 2014, Pflock *et al.*, 2007, Donczew *et al.*, 2015).

Transcription factors such as ferric-uptake regulator Fur and the nickel-uptake regulator NikR, which use metal ions such as nickel (Ni^{2+}) and iron (Fe^{2+}) as co-factors, have been shown to control gene expression in response to different environmental stimuli including acid, heat shock and availability of essential nutrients or ions (Danielli & Scarlato, 2010). Overall, an efficient use of the few regulatory systems in *H. pylori* is accomplished by sensing and responding to multiple signals, and through cross-talk of the transcriptional gene regulators.

1.5.4. The non-coding RNA repertoire of *H. pylori*

Although transcriptome studies provided a global view on *H. pylori* transcriptional networks, almost nothing is known about posttranscriptional gene expression control,

sRNAs and their functions in this emerging human pathogen. None of the enterobacterial sRNAs, except for the housekeeping RNAs such as the transfer-messenger RNA (tmRNA), RNase P RNA (M1 RNA), and signal recognition particle RNA (SRP/4.5S RNA), are conserved in *Helicobacter*. As it also lacks the endonucleolytic RNase E and the major RNA chaperone Hfq, *H. pylori* was even regarded as an organism without riboregulation that only possesses basic regulatory circuits for gene expression control (Mitarai *et al.*, 2007). In addition, bioinformatics-based predictions and a small scale cDNA cloning approach identified only few natural antisense transcripts and sRNA candidates in *Helicobacter* (Xiao *et al.*, 2009a, Xiao *et al.*, 2009b). However, a novel differential RNA sequencing approach (dRNA-seq) selective for the 5' end of primary transcripts that allows for the definition of a global map of transcription start sites (TSS) and operons in *H. pylori* strain 26695, revealed a very complex and compact transcriptional output from the small *Helicobacter* genome (Sharma *et al.*, 2010). Moreover, it led to the discovery of massive antisense transcription as well as an unexpected high number of more than 60 sRNA candidates including potential regulators of *cis*- and *trans*-encoded mRNA targets. This indicates that *H. pylori* uses riboregulation for its gene expression control.

Several of the newly identified *H. pylori* sRNAs seem as abundant as housekeeping RNAs, and most of them are conserved between different *Helicobacter* species, but are absent from bacterial species outside of the phylum of Epsilonproteobacteria (Sharma *et al.*, 2010). Moreover, *H. pylori* sRNAs typically possess specific structural motifs that are divergent from those found in enterobacteria, indicating that *H. pylori* and other Epsilonproteobacteria, such as the food-borne pathogen *Campylobacter jejuni*, might have evolved a unique sRNA repertoire (Dugar *et al.*, 2013). In contrast to their limited transcription factor inventories, Epsilonproteobacteria, in principle, harbor a similar capacity for riboregulation compared to *E. coli,* respective to their smaller genome sizes.

One of the most abundant transcripts among the novel sRNA candidates was identified as a homolog of the 6S RNA, which was previously thought to be absent in the Epsilonproteobacteria based on bioinformatics predictions (Barrick *et al.*, 2005, Sharma *et al.*, 2010). Despite only little sequence conservation to the *E. coli* 6S RNA, the 180-nt long RNA from *H. pylori* can fold into the characteristic long hairpin structure mimicking an open promoter complex (Trotochaud & Wassarman, 2005). Since *H. pylori* lacks the stationary sigma-factor RpoS, the function of 6S RNA in *H. pylori* remains unknown until now.

More than 900 *cis*-encoded antisense RNAs have been identified in *H. pylori* strain 26695 (Sharma *et al.*, 2010). These include *bona fide* sRNAs as well as overlapping 5' and 3' regions of mRNAs from contingous genes that are trancribed in divergent/opposite directions. With at least one antisense TSS associated with about half of all ORFs, the fraction of genes associated with asRNAs in *H. pylori* is among the highest compared to other

bacteria (Georg & Hess, 2011). Whether all of the identified asRNAs are functional or rather represent spurious transcription still needs to be clarified.

In principle, base-pairing RNAs can have the capability to regulate gene expression in *H. pylori*. For example, expression of an artificial asRNA has been successfully used to repress the essential *ahpC* gene, encoding alkyl hydroperoxide reductase (Croxen *et al.*, 2007). In addition, the recent characterization of a naturally-occurring 292-nt long *cis*-encoded asRNA, 5' *ureB*-sRNA, from the opposite strand of the urease operon (*ureAB*), further demonstrated functionality of asRNAs in *Helicobacter* and their potential role in acid adaptation (Wen *et al.*, 2011).

Several of the novel identified sRNAs are candidates for putative *trans*-encoded riboregulators. For example, the highly abundant and conserved sRNA, RepG (<u>Re</u>gulator of <u>p</u>olymeric <u>G</u>-repeats, previously known as HPnc5490) has been suggested to affect expression of a chemotaxis receptor in *H. pylori* strain 26695 by direct base-pairing interactions (Sharma *et al.*, 2010). However, a detailed functional characterization for the majority of *trans*-acting sRNA candidates is still pending and almost nothing is known about their transcriptional regulators, mechanisms of action and physiological functions in *H. pylori*.

1.5.4.1.   Protein factors involved in *H. pylori* riboregulation

Only a few RNA-protein interactions have been studied in *H. pylori*. The essential tmRNA has been shown to interact with its protein co-factor SmpB, and both are required for *trans*-translation and translational control (Thibonnier *et al.*, 2008, Thibonnier *et al.*, 2010). Although *H. pylori* encodes a homolog of CsrA (Barnard *et al.*, 2004), neither the full repertoire of CsrA targets nor homologs of the CsrA-antagonizing sRNAs, CsrB/C, have been identified in Epsilonproteobacteria so far. Because Epsilonproteobacteria lack an Hfq homolog, *Helicobacter* proteins of unknown function or those associated with other cellular processes such as bacterial membrane binding, translation, cell cycle, or virulence could play a role in posttranscriptional regulation. In line with this, the zinc-ribbon domain containing protein HP0958 (FlgZ) has been described as a potential posttranscriptional regulator of *H. pylori* motility genes (Douillard *et al.*, 2008). Furthermore, a direct interaction of the aconitase (AcnB), a major enzyme of the TCA cycle, with the 3' UTR of the cell-wall modifying peptidoglycan deacetylase (*pgdA*) mRNA was shown to increase the stability and expression of *pgdA* (Austin & Maier, 2013).

Using either affinity purification of aptamer-tagged RNAs or RIP-seq, RNA-protein interactions between the ribosomal proteins S1 and various mRNAs as well as sRNAs were identified in *H. pylori* strain 26695 (Rieder *et al.*, 2012). Ribosomal protein S1 might be a

candidate protein that does not only facilitate translation initiation, but might also act as a general RNA chaperone (Hajnsdorf & Boni, 2012).

Because all Epsilonproteobacteria appear to lack RNase E, it is unclear whether or which endoribonuclease participates in sRNA-mediated mRNA decay in these bacteria. In *H. pylori,* most of the mRNA degradation seems to be carried out by a minimal RNA degradosome consisting of a homolog of RNase J and the only DExD-box RNA helicase of *H. pylori*, RhpA (Redko *et al.*, 2013). In line with the strong phenotypes observed upon RNase J deletion in Gram-positive bacteria, *H. pylori* RNase J is essential for growth under standard growth conditions. The biochemical characterization of the *H. pylori* RNase J enzyme showed that it contains both 5'-3' exonucleolytic and endonucleolytic activity similar to its *B. subtilis* ortholog (Dorleans *et al.*, 2011). In general, *H. pylor*i seems to possess the majority of the key enzymes (RNase J, RNase Y, PNPase, RppH and RNase III) that are known to be involved in mRNA degradation in the Gram-positive *B. subtilis* (Condon, 2010, Lundin *et al.*, 2003). However, a large number of processes that occur during sRNA-mediated decay and rRNA maturation await characterization and assignment to specific enzymes or encoding genes in *Helicobacter*.
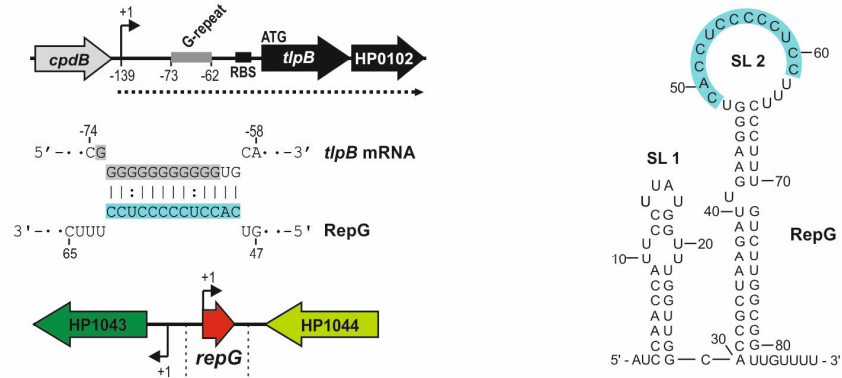
2. A variable homopolymeric G-repeat determines RepG-mediated posttranscriptional regulation of the chemotaxis receptor TlpB

The conserved 87-nt long RepG sRNA was originally identified as one of the most abundant transcripts in the dRNA-seq study of *H. pylori* strain 26695 (Sharma *et al.*, 2010) and has been predicted to regulate expression of the acid-sensing chemotaxis receptor TlpB. The RepG sRNA is transcribed from the intergenic region between HP1043, an orphan response regulator, and HP1044, a protein of unknown function, and is predicted to fold into a two stem-loop structure (Figure 2.1 A). *In-silico* predictions for potential target mRNAs using the TargetRNA program (Tjaden *et al.*, 2006) indicated that the RepG sRNA might base-pair with its C/U-rich terminator loop to a homopolymeric G-repeat in the 5' UTR of the dicistronic mRNA encoding the chemotaxis receptor TlpB (HP0103) and the hypothetical protein HP0102. Studying transcriptome as well as proteome changes upon deletion of *repG* revealed sRNA-mediated down-regulation of *tlpB* on mRNA and protein level in *H. pylori* strain 26695, indicating that this chemotaxis receptor might represent the first *trans*-encoded target of RepG (Sharma *et al.*, 2010). Work presented in this Chapter is focused on the molecular mechanism of RepG-mediated *tlpB* regulation as well as its functional implications, and provides insights into the external conditions under which RepG expression is regulated. As RepG represents the first bacterial sRNA that directly interacts with a homopolymeric G-repeat in its target mRNA, this Chapter investigates whether the length of this SSR affects RepG-mediated posttranscriptional gene expression control of *tlpB*. A novel connection between sRNA-mediated gene expression control and phenotypic variation through variable repeats is proposed.

2.1. The highly abundant sRNA RepG is broadly conserved in *Helicobacter*

Biocomputational searches for RepG homologs in 31 different *Helicobacter pylori* strains, *Helicobacter acinonychis, Helicobacter cetorum* (MIT_00_7128) and *Helicobacter mustelae* revealed that the RepG sRNA, its genomic context, and its predicted secondary structure are highly conserved (Figures 2.1 B and C). In particular, the C/U-rich terminator loop of RepG, corresponding to the *tlpB* interaction site, is remarkably conserved, even in more distant species such as the ferret-colonizing species *H. mustelae* (Figure 2.1 B). Using northern blot analysis, expression of RepG homologs was confirmed in various *Helicobacter* strains in exponential growth phase; however, strain-specific variations in abundance and band patterns of RepG were observed (Figure 2.2 A). For example, in spite of a highly similar *repG* sequence between the two strains, multiple RepG-derived bands were detected in *H. pylori*

A

+1 G-repeat ATG RBS

cpdB tlpB HP0102

-139 -73 -62

-74 ... -58

5' -- CG ... CA -- 3' *tlpB* mRNA

GGGGGGGGGGGUG

||:|||||||||

CCUCCCCCUCCAC RepG

3' -- CUUU ... UG -- 5'

65 47

+1

HP1043 *repG* HP1044

+1

SL 2

SL 1 RepG

B

-10 +1 20 40 60 80

Helicobacter pylori

26695, SJM180, P12, Shi470, Sat464, Lithuania75, Cuz20, PeCan4, v225d, Puno120, Puno135, 35A, F16, F30, F32, F57, 51, 52, B38, 83, India7, SNT49, B8, G27, J99, HPAG1, 2017, 2018, 908, Gambia94/24, SouthAfrica7, Hac, Hce, Hmu

secondary structure SL 1 SL 2 *tlpB* binding site

C

Helicobacter pylori

26695, P12, Lithuania75, Cuz20, India7, B8, G27

SJM180, PeCan4, Shi470, Sat464, J99, Hac, Hce, Hmu

flhA HP1042 *repG* HP1043 HP1044 acoE rimP rbfA infB

HPSJM_02140 HPSJM_02135

HPPC_02035

Hac_1148 Hac_1149 Hac_1150 Hac_1151 Hac_1152

HCW_03545 HCW_03535 HCW_03525 03530

truD alr HMU_12360 recR dnaJ rbn dnaB HMU_12300 HMU_12290
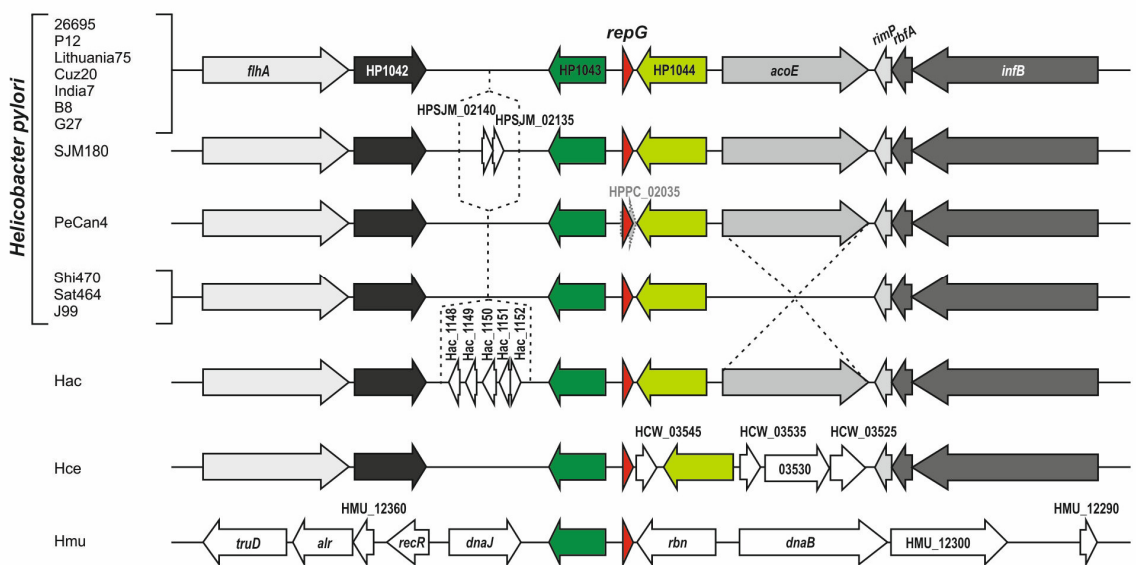
Figure legend on next page

Figure 2.1: Genomic context and sequence alignment of *repG* homologs in different *H. pylori* strains, *H. acinonychis* (Hac), *H. cetorum* (Hce) and *H. mustelae* (Hmu). (A) The 87-nt long RepG sRNA is encoded between an orphan response regulator, HP1043, and a hypothetical protein, HP1044, and is predicted to fold into a two stem-loop structure. A C/U-rich single-stranded region (marked in blue) in the RepG terminator loop (SL 2) was predicted to base-pair with a G-repeat (marked in gray) in the 5′ UTR of the dicistronic *tlpB*-HP0102 mRNA (dotted line) encoding a chemotaxis receptor and a hypothetical protein, respectively. Transcriptional start sites (TSS, +1; Sharma *et al.*, 2010) are indicated by arrows, and the *tlpB* G-repeat and RBS by gray and black bars, respectively. Numbers indicate the distance to the *tlpB* start codon (ATG). **(B)** Arrows below the *repG* alignment indicate the predicted stem-loop structures SL 1 and SL 2 of RepG. The *tlpB* binding site in SL 2 is indicated by a blue bar. The alignment is based on *Helicobacter* genome sequences published until January 2014. **(C)** Homologs in different *Helicobacter* strains are illustrated by the same colors, whereas unrelated genes are indicated in white. For simplicity, only one strain of *H. cetorum* is shown (MIT_00_7128). Gene insertions or deletions are marked by dotted lines. A potential ORF (gray arrow) is annotated next to *repG* in *H. pylori* strain PeCan4.

strain 26695 (Figure 2.2 A, lane 1), while only a single band was observed in G27 (Figure 2.2 A, lane 11). Primer extension assays revealed that the different bands in strain 26695 correspond to slightly different RepG versions that vary at their 5' end (87, 85, 84-nt long; Figure 2.2 B). The dominant, 87-nt long RepG species in 26695 corresponds to a transcript starting with an "A" at the same position that was determined as the transcriptional start site (TSS) in the dRNA-seq study from Sharma *et al.*, 2010. In strain G27, the TSS of RepG is shifted by one nucleotide downstream as compared to the main TSS in *H. pylori* strain 26695. Despite these differences, a similar RepG expression profile over growth was observed in both strains (Figure 2.14 A, see section 2.10.).

2.2.   RepG – the first example of a *trans*-acting small RNA in *H. pylori*

Bioinformatics-based target predictions and preliminary investigations of a *repG* deletion mutant indicated that RepG represses the expression of *tlpB* at both the mRNA and protein level in *H. pylori* strain 26695 (Sharma *et al.*, 2010). The chemotaxis receptor TlpB is required for chemo-repulsion from acid and the quorum-sensing molecule autoinducer-2 (AI-2) (Croxen *et al.*, 2006, Rader *et al.*, 2011), and has been implicated in colonization and inflammation during mice and gerbil infections (Croxen *et al.*, 2006, McGee *et al.*, 2005, Williams *et al.*, 2007). To study the putative regulation of *tlpB* by RepG in further detail*,* the Δ*repG* mutant of *H. pylori* strain 26695 was complemented with wild-type (C$_{RepG}$) or various mutant RepG sRNAs expressed from the native P$_{repG}$ promoter at the unrelated *rdxA* locus. The following sRNA mutants were tested: SL 2, which expresses only the second stem loop of RepG; ΔCU, in which the C/U-rich region of the RepG terminator loop was replaced by an extra-stable tetra loop; 3xG and 1xG*, in which three or one* C residue(s) in the predicted
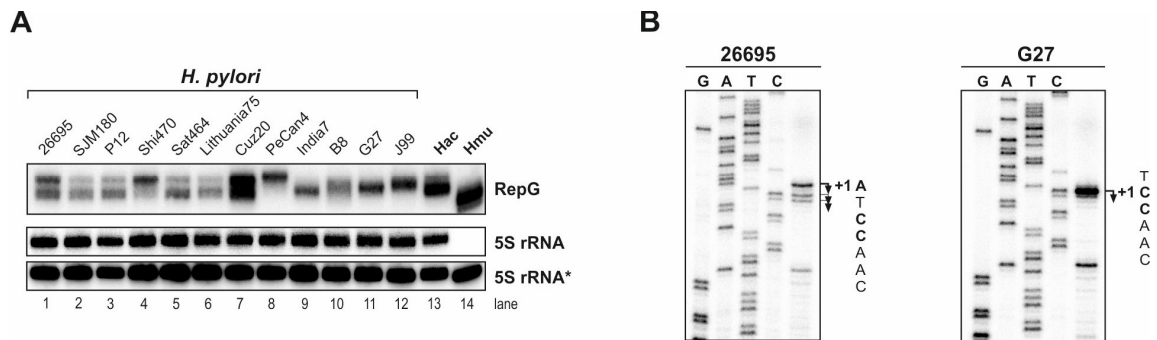
Figure 2.2: Strain-specific abundances and band patterns of RepG homologs in diverse *Helicobacter* strains. (A) Northern blot analysis of RepG using ^{32}P-labeled oligonucleotide CSO-0003 in diverse *H. pylori* strains, *H. acinonychis* (Hac), and *H. mustelae* (Hmu) at exponential growth phase. 5S rRNA served as loading control and was probed with two oligonucleotides: JVO-0485 (5S rRNA) for *H. pylori* and *H. acinonychis,* and CSO-0053 (5S rRNA*) for *H. mustelae*, respectively. **(B)** Determination of the RepG transcriptional start site using primer extension in *H. pylori* strains 26695 and G27. Total RNA was isolated from *H. pylori* strains 26695 (*left panel*) and G27 (*right panel*) grown to exponential growth phase (OD$_{600nm}$ of ∼ 0.8). After DNase I treatment, 10 μg of total RNA was used in primer extension assays with ^{32}P end-labeled oligo JVO-5126. A sequencing ladder corresponding to the *repG* upstream region served as reference (lanes G, A, T, C). The identified TSS of RepG (+1) are indicated by arrows and bold nucleotides.

interaction site was/were exchanged to G(s) (Figure 2.3 A, Table 6.12). *H. pylori* strain 26695 wildtype (WT), Δ*repG*, and sRNA complementation (C$_{RepG}$, SL 2, ΔCU, 3xG and 1xG*) mutants were grown to exponential growth phase and total RNA as well as protein samples were analyzed by northern blot, and SDS-PAGE or western blot, respectively. Analysis of whole cell protein fractions of by one-dimensional SDS-PAGE showed that complementation of Δ*repG* with the wild-type sRNA and the SL 2 mutant, both of which harbor the predicted C/U-rich *tlpB* interaction site, restore *tlpB* repression to wild-type levels (Figure 2.3 B). Despite a three-fold lower transcript level, the SL 2 mutant regulates *tlpB* to the same extent as the wild-type sRNA, indicating that RepG levels are not limiting for *tlpB* regulation under the conditions examined. In contrast, despite similar sRNA levels compared to the wildtype, deletion of the C/U-rich binding site (ΔCU) and introduction of triple or single* point mutations (3xG and 1xG*) abolished *tlpB* regulation, confirming that the sRNA terminator loop is essential for repression of this chemotaxis receptor. Western blot analysis using an antiserum against all four chemotaxis receptors of *H. pylori* confirmed that the TlpB protein was about five-fold repressed in the presence of RepG, whereas the protein levels of the other chemotaxis receptors TlpA, TlpC, and TlpD remained unaltered upon deletion of the sRNA. Comparison of Δ*tlpB* and Δ*tlpB*/Δ*repG* double deletion mutants confirmed that the up-regulated band was indeed derived from TlpB. Overall, these data demonstrate that the conserved C/U-rich sequence of RepG is the *tlpB* interaction site. Moreover, the simple stem-

28
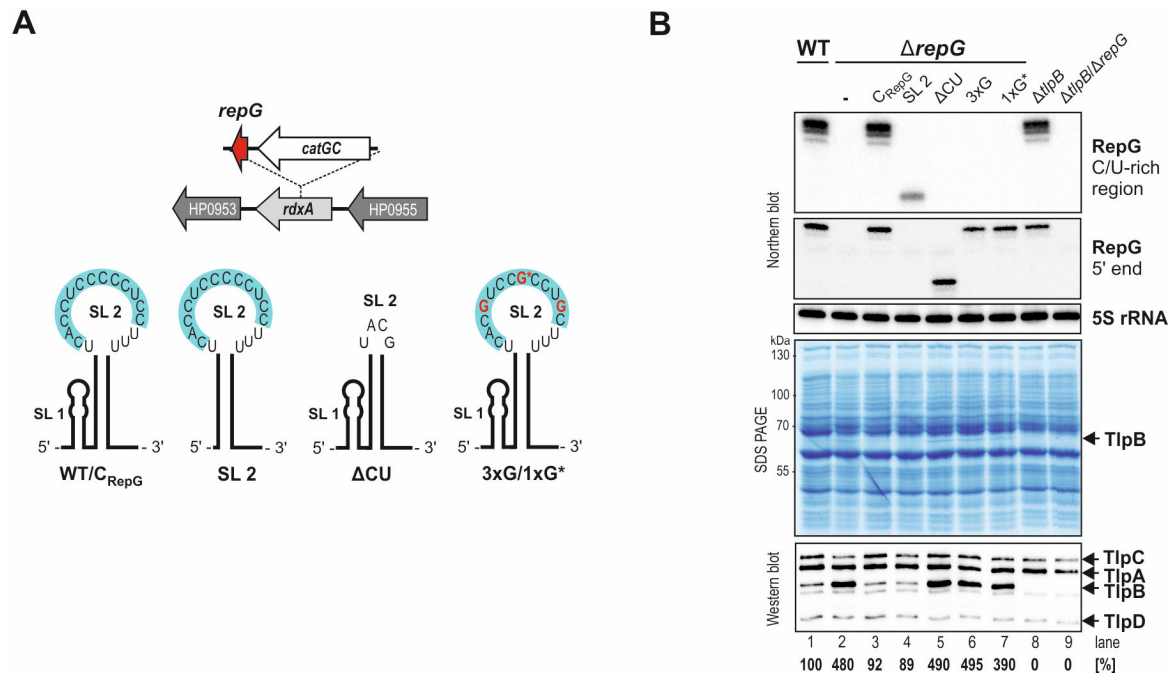
Figure 2.3: The RepG terminator harbors the C/U-rich *tlpB* interaction site and is sufficient to repress *tlpB* expression. (A) The Δ*repG* mutant was complemented with RepG wildtype (C_RepG) or several mutant sRNAs in the *rdxA* locus. The sRNA mutant SL 2 consists of only the second stem loop (nucleotides 30-87) of RepG. In ΔCU the *tlpB* binding site (marked in blue) was replaced by an extra-stable tetra loop (UACG). Triple or single* C to G point mutations at position 52, 56 and 60 (3xG) or at position 56 (1xG*) are indicated in red. **(B)** *H. pylori* 26695 wildtype (WT) and indicated mutant strains were grown to exponential growth phase, and RNA and protein samples were analyzed by northern blot, and SDS-PAGE or western blot, respectively. RepG was detected with CSO-0003 (binds to the C/U-rich loop) and JVO-2134 (binds to RepG 5' end). 5S rRNA was used as loading control (JVO-0485). Whole cell protein fractions (OD_{600nm} of 0.1 for Coomassie gel or 0.01 for western blot) were directly stained with Coomassie or chemotaxis receptors TlpA, B, C and D were detected by a polyclonal rabbit anti-TlpA22 antiserum.

loop structure of 58 nt (SL 2), corresponding to the RepG terminator, is sufficient to act as a regulatory RNA and thus, could be used for the optimized design of synthetic RNA regulators, which commonly consist of multiple domains (Na *et al.*, 2013).

2.3. RepG represses *tlpB* expression at the posttranscriptional level

To uncouple RepG function from transcriptional control of *tlpB* and investigate whether RepG regulates expression of this chemotaxis receptor at the posttranscriptional level, the *tlpB* promoter was exchanged with the unrelated *cagA* promoter of the major effector protein CagA. Therefore, the endogenous *tlpB* promoter of *H. pylori* strain 26695 was replaced by insertion of the *cagA* promoter in the chromosome upstream of the *tlpB*
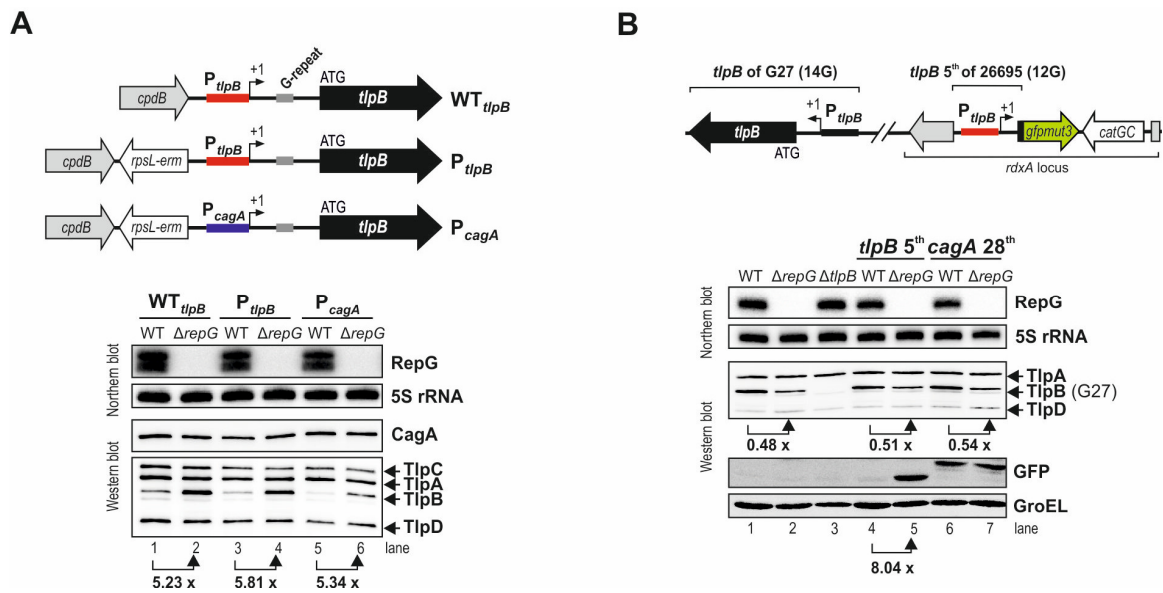
Figure 2.4: Posttranscriptional RepG-mediated *tlpB* regulation. (A) *(Upper panel)* Schematic illustration of the *tlpB* locus including its promoter region (WT*tlpB*). The *cagA* promoter region (blue bar) together with an *rpsL-erm* resistance cassette (P*cagA*) or *rpsL-erm* alone (P*tlpB*) were inserted upstream of the TSS (+1) and promoter (red bar) of *tlpB*. The *tlpB* G-repeat is marked by a gray box. (*Lower panel*) *H. pylori* 26695 strains with either the wild-type *tlpB* locus (WT*tlpB*) or mutants (P*tlpB*, P*cagA*) in the wild-type (WT) or *repG* deletion (Δ*repG*) background were grown to exponential growth phase, and RNA and protein samples were analyzed by northern and western blot, respectively. 5S rRNA and CagA protein served as controls. **(B)** (*Upper panel*) The *H. pylori* 26695 *tlpB* promoter region (red bar), its 5' UTR and the first five amino acids of the *tlpB* coding region were fused to *gfpmut3* and inserted into the *rdxA* locus of *H. pylori* G27. (*Lower panel*) *H. pylori* G27 WT, Δ*repG* and Δ*tlpB* mutant strains, as well as WT and Δ*repG* strains, which carry either the *tlpB* 5th::*gfpmut3* or *cagA* 28th::*gfpmut3* fusions were grown to exponential phase, and RNA and protein samples were analyzed by northern and western blot, respectively. *H. pylori* G27 expresses only three chemotaxis receptors, TlpA, B, and D.

transcriptional start site (Sharma *et al.*, 2010, Delany *et al.*, 2002) together with an *rpsL-erm* resistance cassette (Figure 2.4 A, *upper panel*). In contrast to *tlpB*, *cagA* is constantly transcribed over growth (Delany *et al.*, 2002) and western blot analysis for endogenous CagA confirmed that *cagA* expression itself is not affected by RepG (Figure 2.4 A). As a control, the resistance cassette alone was inserted upstream of the *tlpB* promoter (P*tlpB*) and western blot analysis confirmed that this did not interfere with neither TlpB expression nor its regulation by RepG (Figure 2.4 A, lanes 1-2 and 3-4). When *tlpB* was transcribed from the *cagA* promoter (P*cagA*), a two- to three-fold reduction in TlpB protein level was observed compared to the wildtype (WT*tlpB*) and control (P*tlpB*) strain. Nevertheless, deletion of *repG* resulted in a five-fold increase of TlpB protein level (Figure 2.4 A, lanes 5-6), indicating that RepG-mediated regulation of *tlpB* occurs at the posttranscriptional level.

Translational reporter fusions based on *gfp* or *lacZ* have been successfully used to study sRNA-mediated gene expression control and to define mRNA and sRNA interaction

sites *in vivo* (Urban & Vogel, 2007, Sharma & Vogel, 2009, Mandin & Gottesman, 2009). Therefore, a GFP variant, *gfpmut3*, which has been previously used in transcriptional fusions on reporter plasmids in *Helicobacter* (Carpenter *et al.*, 2007), was applied to construct a translational reporter fusion in the chromosome of *H. pylori*. To confirm that the leader of the *tlpB* mRNA is sufficient for RepG-mediated repression, the first five amino acids of the coding region, the promoter region and the 5' UTR of *tlpB* from *H. pylori* strain 26695 were fused to *gfpmut3*. Because transformation of *gfpmut3* seems to be toxic for *H. pylori* strain 26695, but not for G27, the translational *tlpB* 5th::*gfpmut3* reporter fusion (*tlpB* mRNA leader sequence of strain 26695) was introduced in the *rdxA* locus of *H. pylori* strain G27 (Figure 2.4 B, *upper panel*; for sequence details see Table 6.9). Western blot analysis of the TlpB::GFP fusion protein in *H. pylori* G27 showed an approximately eight-fold up-regulation upon *repG* deletion (Figure 2.4 B, lanes 4-5). As a control, expression of a CagA::GFP fusion protein (*cagA* 28th::*gfpmut3*) was not affected by *repG* deletion (Figure 2.4 B, lanes 6-7), which is in agreement with what was observed for endogenous CagA (Figure 2.4 A). Together, our *in-vivo* results show that RepG represses *tlpB* at the posttranscriptional level by interacting with its 5' UTR.

In contrast to the repression of the translation of *tlpB* 5th::*gfpmut3* fusion from strain 26695, the endogenous TlpB protein level of *H. pylori* strain G27 was about two-fold decreased upon *repG* deletion (Figure 2.4 B, lanes 4-5). Since RepG is highly conserved and the *tlpB* 5' UTRs are overall very similar in both strains but carry different G-repeat lengths, namely 12 and 14Gs, the G-repeat length might determine strain-specific RepG-mediated *tlpB* regulation, which was investigated later on (see below).

2.4. RepG and *tlpB* mRNA base-pair directly

In-vitro experiments are widely used to identify and validate sRNA-mRNA interactions, including the determination of binding affinities by gel-shift assays with *in-vitro* transcribed RNAs. To further test for a direct interaction between the C/U-rich terminator loop of RepG und the homopolymeric G-repeat in the 5' UTR of *tlpB*, gel-shift assays were performed with the 5' end-labeled *tlpB* mRNA leader from *H. pylori* strain 26695 and increasing amounts of RepG and *vice versa* (Figure 2.5; for sequence details see Table 6.12). Affinity electrophoresis experiments revealed an RNA-RNA complex formation at a molecular ratio of 1 : 15.6 (4 nM : 62.5 nM), whereas modifications in the *tlpB* binding site of RepG (ΔCU, 3xG, 1xG*) abolished RepG-*tlpB* mRNA interaction (Figure 2.5 B). In line with this, RepG did not shift with increasing amounts of *tlpB* mRNA leader variants, which either lack the G-stretch (ΔG) or contained triple and single* nucleotide exchanges in the homopolymeric
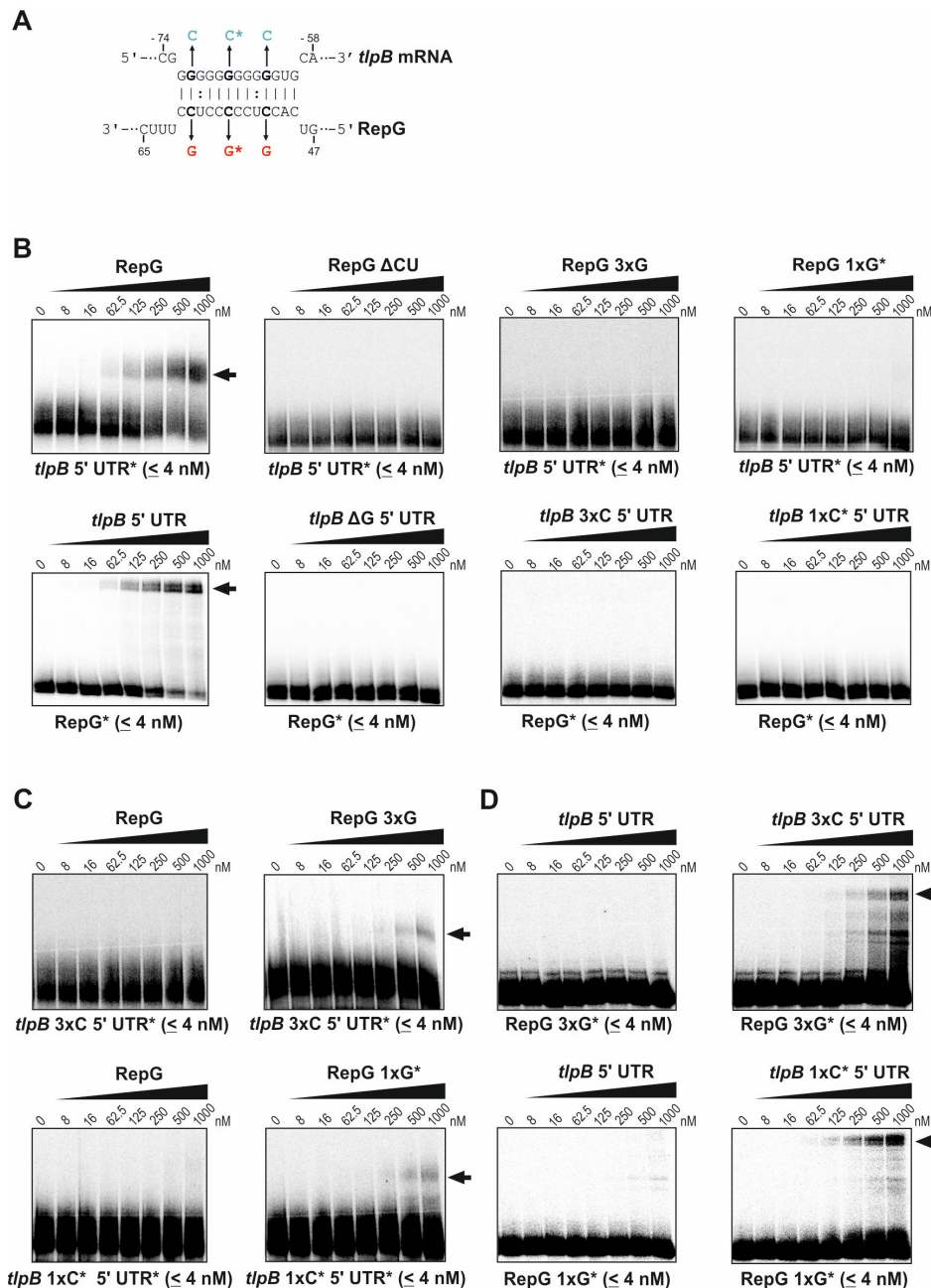
Figure 2.5: Validation of a direct interaction between RepG and the *tlpB* mRNA leader using gel-mobility shift assays. **(A)** The C/U-rich terminator loop of RepG targets a homopolymeric G-repeat in the *tlpB* mRNA leader. Triple and single* nucleotide exchanges in the *tlpB* binding site of RepG and compensatory base-pair exchanges in the G-repeat of the *tlpB* 5' UTR are indicated in red and blue, respectively. **(B)** (*Upper panel*) About 0.04 pmol (4 nM final concentration) *in-vitro* transcribed, 5' end ^{32}P-labeled *tlpB* mRNA leader was incubated without or with increasing concentrations (final concentrations are indicated) of unlabeled RepG or RepG variants (ΔCU, 3xG and 1xG*) for 15 minutes at 37 °C. RNA-RNA complex formation was investigated by direct loading of the samples to a native 6 % PAA gel. (*Lower panel)* In a reverse experiment, about 0.04 pmol labeled RepG was incubated with the wild-type *tlpB* leader or *tlpB* leader variants (ΔG, 3xC, 1xC*). **(C)** For *in-vitro* analyses of the compensatory base-pair exchanges, labeled *tlpB* mRNA leader variants (3xC, 1xC*) were incubated with wild-type RepG or RepG variants (3xG, 1xG*). **(D)** *Vice versa* gel-mobility shift assays, in which RNA-RNA complex formation between the RepG variants 3xG or 1xG* and the wild-type or *tlpB* mRNA leader variants with compensatory mutations (3xC, 1xC*) was examined.

G-repeat (3xC, 1xC*). Similarly, 5' end-labeled *tlpB* 5' UTR variants (3xC, 1xC*) and RepG sRNA mutants (3xG, 1xG*) did not shift with the respective wild-type RNAs (Figures 2.5 C and D). Only gel shifts with RNA pairs containing compensatory base-pair exchanges, *i.e.* RepG 3xG with *tlpB* 3xC or RepG 1xG* with *tlpB* 1xC*, restored the interaction between RepG and the *tlpB* mRNA leader, albeit with lower affinities compared to the respective wild-type versions (Figures 2.5 C and D). Overall, gel shifts confirmed a direct interaction between the homopolymeric G-repeat in the *tlpB* 5' UTR and the C/U-rich terminator loop of RepG *in vitro* and that mutations in the interaction sites abolish RNA-duplex formation.

2.5.   The C/U-rich RepG terminator loop interacts directly with the homopolymeric G-repeat in the *tlpB* mRNA leader

To map the sRNA and mRNA interaction site *in vitro*, structure probing assays of *in-vitro* transcribed, 5' end-labeled RepG in the absence or presence of the unlabeled *tlpB* mRNA leader were performed using in-line probing (Regulski & Breaker, 2008) as well as enzymatic and chemical cleavages (Figure 2.6). The in-line probing technique uses the natural instability of RNA under basic conditions to elucidate its secondary structure characteristics. Because single-stranded or unstructured RNA regions undergo spontaneous cleavage of phosphoester linkages faster than structured regions, a reduction in spontaneous RNA degradation can not only be used for secondary structure probing, but also to monitor ligand binding or RNA-RNA interactions. Cleavage patterns in the in-line probing assays agreed with single- and double-stranded regions according to the two biocomputationally predicted stem loops of RepG (Figures 2.6 A and C). Similarly, structure mapping of labeled RepG using RNase T1 (cleaves single-stranded RNA after guanine residues), lead (II)-acetate (cleaves single-stranded nucleotides), and RNase III (cleaves double-stranded RNA) confirmed the predicted RepG secondary structure (Figures 2.6 B and C). However, slight protection of several lead (II)-cleavage sites as well as some RNase III sites were observed in the predicted single-stranded, 17-nt long terminator loop harboring the *tlpB* interaction site. A similar cleavage pattern was observed in structure probing assays with the RepG terminator loop (SL 2) alone, arguing against the formation of an intramolecular structure between the first (SL 1) and second stem loop (SL 2; data not shown). These data indicate that nucleotides within the C/U-rich loop of RepG might be involved in a tertiary structure. Nonetheless, a clear footprint was observed in the RepG terminator loop region in the in-line and lead (II)-acetate probing upon addition of the unlabeled *tlpB* mRNA leader, suggesting that the C/U-rich terminator loop is indeed involved in the sRNA-mRNA interaction (Figures 2.6 A, lanes 5-6 and 2.6 B, lanes 8-9).
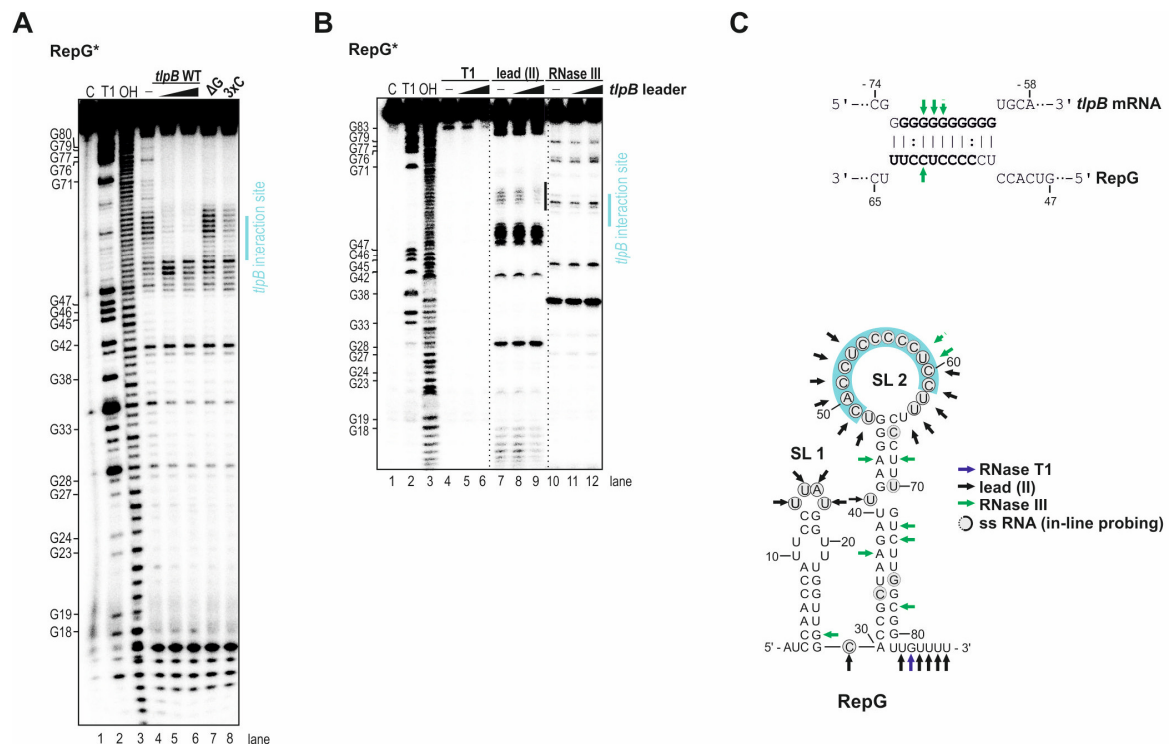
Figure 2.6: *In-vitro* structure probing of RepG in absence or presence of the *tlpB* mRNA leader.
(A) In-line probing of about 0.2 pmol ^{32}P-labeled RepG in the absence (lane 4) or presence of either 20 nM (lane 5) or 200 nM *tlpB* mRNA leader (lane 6), or 200 nM of *tlpB* leader mutants ΔG (lane 7) and 3xC (lane 8). Spontaneous cleavage of single-stranded regions was analyzed on 10 % PAA gel under denaturing conditions. **(B)** About 0.1 pmol (10 nM final concentration) *in-vitro* transcribed and ^{32}P-labeled RepG* was treated in the absence (lanes 4, 7, 10) or presence of 100 nM (lanes 5, 8, 11) or 1000 nM (lanes 6, 9, 12) unlabeled *tlpB* mRNA leader with RNase T1, lead (II)-acetate or RNase III. Protection from lead (II)-acetate in the predicted *tlpB* interaction site (blue bar) of RepG is indicated in black. Untreated RNA (lane C), partially alkali- (lane OH) or RNase T1- (lane T1) digested RepG served as ladders in structure mapping experiments. **(C)** (*Upper panel*) Predicted 11-bp long duplex between the C/U-rich loop of RepG and the G-repeat in the *tlpB* 5' UTR based on structure probing assays. Identified RNase III cleavage sites in the RepG-*tlpB* mRNA duplex are indicated by arrows and nucleotides that are protected from cleavage in the in-line probing assays are marked in bold. (*Lower panel*) Putative secondary structure of RepG based on structure mapping assays. Arrows indicate RNase T1, lead (II)-acetate, and RNase III cleavage sites according to the cleavage pattern from the structure probing assays (A-B). Locations of spontaneous cleavage of nucleotides in single-stranded RNA regions from in-line probing experiments are circled in gray. The terminator loop of RepG (SL 2) harbors the C/U-rich *tlpB* binding site (marked in blue).

Consistent with the predicted RepG-*tlpB* mRNA interaction, this protection from spontaneous cleavages within the C/U-rich region was not observed with a *tlpB* mutant RNA that lacks the G-repeat (ΔG) and was only slightly visible upon addition of the 3xC *tlpB* mutant RNA (Figure 2.6 A, lanes 7-8).

In a reciprocal experiment, the structure of the 5' end-labeled *tlpB* mRNA leader (-139 to + 78 nt relative to the annotated start codon) was mapped in absence or presence of unlabeled RepG (Figure 2.7). In combination with RNA secondary structure predictions

using RNAstructure (Mathews, 2006), structure mapping results indicated a stem loop in front of the ribosome binding site (RBS) and start codon (AUG) of the 139-nt long *tlpB* mRNA leader. Interestingly, despite being predicted as single-stranded, only minor cleavage events were obtained within the G-repeat in the in-line probing assay (Figure 2.7 A). Furthermore, a protection against RNase T1 and lead (II)-cleavages as well as two RNase III cleavage sites (Figure 2.7 B) indicate that the G-repeat might fold into a potential intra- or intermolecular structure, which could not be resolved with the applied methods. Apart from this potentially structured region, the cleavage patterns in the in-line and lead (II)-probing assays indicate a rather flexible or multiple conformations of the *tlpB* leader (Figure 2.7 C). Nonetheless, addition of increasing concentrations of RepG resulted in a footprint in the in-line probing as well as additional RNase III cleavages in the homopolymeric G-repeat (Figures 2.7 A and B), indicating RepG-*tlpB* complex formation. Furthermore, several structural rearrangements, especially in the stem-loop structure upstream of the *tlpB* RBS, were observed upon RepG-*tlpB* mRNA interaction. In agreement with the gel-mobility shift assays (Figure 2.5 B), the footprint as well as RNase III cleavages, and structural rearrangements in the *tlpB* mRNA leader were not observed upon addition of ΔCU or 3xG RepG mutant RNAs. In summary, the structure probing results support an interaction between the C/U-rich terminator loop of RepG and the G-repeat in the *tlpB* 5' UTR (Figure 2.6 C). Note that the interacting nucleotides based on the structure probing results are slightly shifted compared to the predicted RepG-*tlpB* mRNA interaction (Figure 2.1 A).

2.6.   Posttranscriptional RepG-mediated *tlpB* regulation varies in different *H. pylori* strains

Conservation analysis revealed that RepG and especially its C/U-rich terminator loop are highly conserved among different *H. pylori* strains (Figure 2.1 B). In contrast, despite high conservation of the promoter (-10), RBS and translational start codon, the homopolymeric G-repeat in the *tlpB* mRNA leader varies in length among diverse *H. pylori* strains ranging from 6 to 16 guanines (Figure 2.8 A). In *H. pylori* strain 26695, in which *tlpB* is repressed by the RepG sRNA, the G-repeat comprises 12 guanines. In contrast, the G-stretch is completely absent in *H. pylori* strains SJM180, Puno120 and Puno 135 or contains a duplication of two G-repeats of 17Gs separated by TGGTTTT in strain 908. Length variation of SSRs has been shown to either affect translation by introducing frame-shift mutations within coding regions (intragenic SSR) or transcription by changing the spacing of promoter elements(intergenic SSR) (Moxon *et al.*, 2006). Previously, the homopolymeric G-stretch in *tlpB* mRNA leader has been described as an intergenic, promoter-associated SSR (Saunders *et al.*, 1998). However, the global transcriptome map of *H. pylori* strain 26695 revealed that

Figure 2.7: *In-vitro* **mapping of the *tlpB* mRNA leader and RepG-*tlpB* mRNA interaction site. (A)** In-line probing experiment with about 0.2 pmol ^{32}P-labeled *tlpB* mRNA leader in the absence (lane 4) or presence of either 20 nM (lane 5) or 200 nM RepG (lane 6), or 200 nM of the RepG mutants ΔCU (lane 7) and 3xG (lane 8). **(B)** *In-vitro* structure mapping of the RepG-*tlpB* mRNA interaction using RNase T1, lead (II)-acetate, and RNase III. About 0.1 pmol ^{32}P-labeled *tlpB* mRNA leader was treated with RNase T1, lead (II)-acetate and RNase III in the absence (-) or presence of either 100 nM or 1000 nM RepG or 1000 nM RepG ΔCU or 3xG mutant RNAs. RepG-mediated protection from RNase T1 cleavage is indicated by blue bars. RNase III cleavage sites in the G-repeat and structural rearrangements in the *tlpB* 5' UTR are indicated by green stars and black bars, respectively. **(C)** Potential secondary structure of the *tlpB* mRNA leader based bioinformatics-based predictions and structure mapping assays (A-B). The homopolymeric G-repeat (red) of the *tlpB* mRNA leader was protected from cleavage, indicating the formation of an intramolecular structure (box).

this G-stretch is located in the 5' UTR (Sharma *et al.*, 2010), and thus, could not lead to frame-shift mutations and presumably does not affect transcription of *tlpB*. Based on the observed sRNA-mediated regulation of *tlpB* expression, variations in the G-repeat length

might influence posttranscriptional regulation of *tlpB* by RepG. Therefore, *repG* was deleted in *H. pylori* strains B8, Cuz20, Shi470, Lithuania75, India7, J99 and G27, which harbor a G-repeat length ranging from 8 to 14Gs. Western blot analysis of whole protein extracted from *repG* deletion mutants and their respective *H. pylori* wild-type backgrounds showed that the basal protein levels of the four chemotaxis receptors vary slightly among the different *H. pylori* strains; however, only the TlpB protein level was affected by *repG* deletion (Figure 2.8 B). Two- to seven-fold increased TlpB protein levels were detected in *repG* deficient mutants compared to their parental wild-type strains in *H. pylori* B8, Cuz20, Shi470, Lithuania75 and India7, which have homopolymeric G-repeats of 8 to 12Gs. However, deletion of *repG* had only a slight effect on TlpB protein level in strain J99, which has a 13G-long repeat.

In contrast to the RepG-mediated *tlpB* repression for *H. pylori* strains with a G-repeat composed of 8 to 12Gs, TlpB protein level was about two-fold down-regulated upon deletion of *repG* in strain G27, which harbors a 14G-long repeat. This is in agreement with the observed decrease of endogenous TlpB level in G27 during the GFP reporter fusion experiments (Figure 2.4 B), and suggests that longer G-repeat lengths could lead to activation of TlpB expression by RepG. Despite slight variations in their 5' ends, RepG homologs in strains 26695 and G27 are highly conserved, especially in the C/U-rich *tlpB* interaction site, and show a similar expression pattern over growth (Figures 2.1 B and 2.14 A). Complementation of the sRNA deficient mutant in *H. pylori* strain G27 with *repG* originating from strain 26695 restored wild-type protein level of TlpB in strain G27 (Figure 2.8 C). This indicates that the difference in RepG-mediated *tlpB* regulation is based on differences in the G-repeat length, rather than on sRNA variations. In line with this, RepG from G27 is able to efficiently repress expression of a *tlpB* 5th::*gfpmut3* reporter fusion from strain 26695 that contains a 12G-long repeat (Figure 2.4 B). Overall, these data indicate that variations in the G-repeat length could influence sRNA-mediated regulation of *tlpB*.

2.7. The length of the homopolymeric G-repeat determines posttranscriptional regulation of *tlpB* and defines an optimal window for RepG-mediated repression

The above mentioned results strongly suggest that the length of the G-repeat affects RepG-mediated regulation, as little differences were obtained in RepG primary sequence and expression across *Helicobacter* strains. However, strain-specific features other than the homopolymeric G-repeat may contribute to the differences in *tlpB* regulation observed above. To test whether the high level of genetic diversity among the *H. pylori* strains or the different G-repeat length can lead to the observed strain-specific *tlpB* regulation, the

37
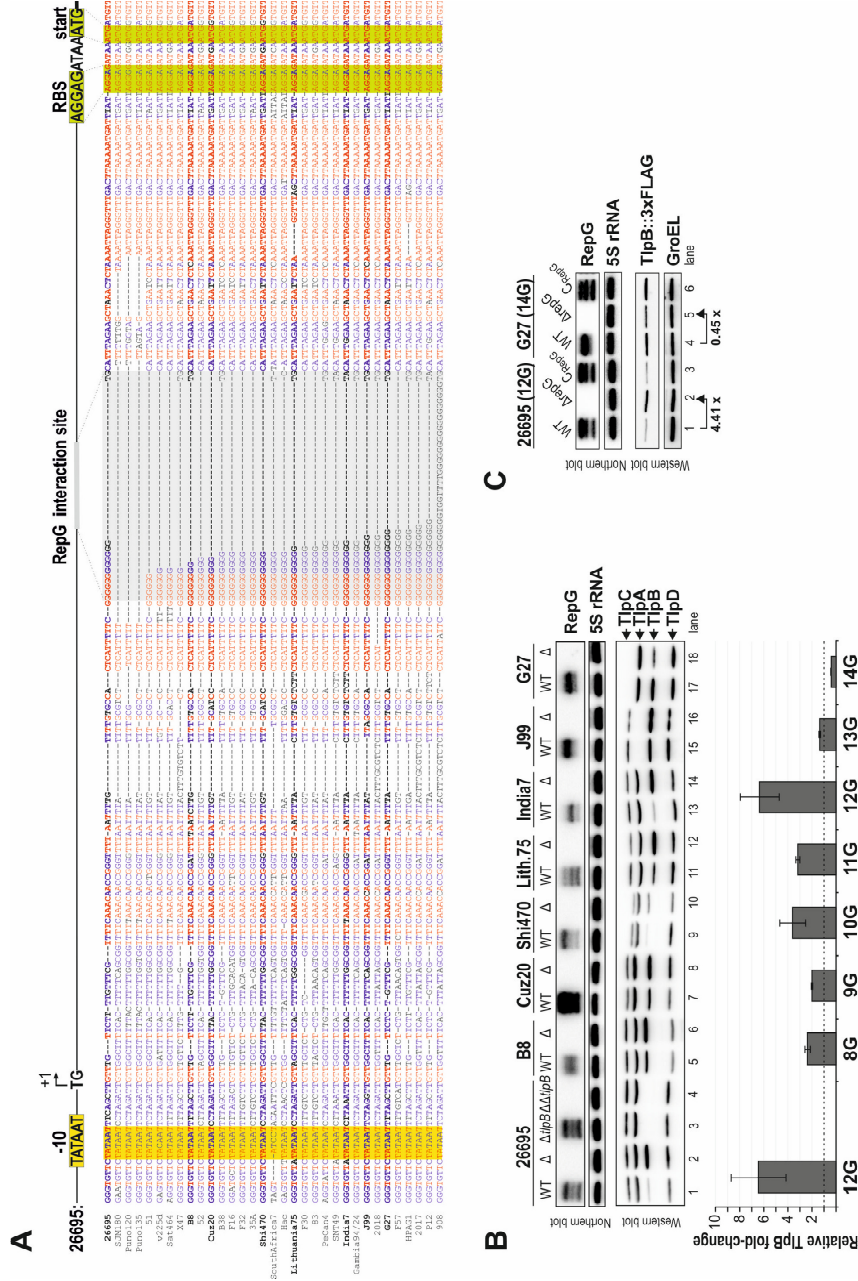
Figure 2.8: *H. pylori* strains with variable G-repeat lengths in the *tlpB* leader show differences in RepG-mediated regulation. (A) Sequence alignment of *tlpB* homologs, including promoter regions (-10 box), TSSs (+1), 5' UTRs with RBSs and annotated start codons (ATG), and the first nucleotides of the ORFs from diverse *H. pylori* strains and *H. acinonychis* (Hac). The RepG interaction site according to strain 26695 is boxed in gray. **(B)** (*Upper panel*) Northern and western blot analyses for RepG-mediated *tlpB* regulation in diverse *H. pylori* WT and Δ*repG* mutants (bold in A) at exponential growth phase. (*Lower panel*) Relative fold-changes of TlpB protein levels upon *repG* deletion in indicated *H. pylori* strains (*upper panel*) are represented in the bar diagram (based on two/three biological replicates). **(C)** The *tlpB* mRNA leaders of strains 26695 and G27 comprise a 12 and 14-nt long G-repeat, respectively. Protein and RNA samples from *H. pylori* 26695 and G27 WT, Δ*repG*, and complementation with RepG from strain 26695 (C$_{RepG}$), which carried a *tlpB*::3xFLAG gene, were harvested at exponential growth phase and analyzed by northern and western blot, respectively. TlpB::3xFLAG was detected with anti-FLAG antibody and GroEL served as loading control.

G-stretch length was modified in the *tlpB* mRNA leader of *H. pylori* strain 26695 at its native locus, which normally contains 12Gs. Therefore, the G-repeat in the 5' UTR of the *tlpB* mRNA was either deleted completely (ΔG) or its length was modulated from 6 to 16 guanines (6-16G) in a *H. pylori* 26695 *tlpB*::3xFLAG strain background (Figure 2.9 A). In order to avoid polar effects on expression of the downstream encoded gene, HP0102, a markerless cloning strategy was used for FLAG-tagging of TlpB (see Chapter 6, Material and methods). Western blot analysis of TlpB::3xFLAG levels in these G-stretch variants in both the wild-type and the Δ*repG* deletion background revealed that RepG-mediated *tlpB* regulation is dependent on the length of the G-repeat (Figures 2.9 B and C). Whereas a lack of the G-stretch had only a minor influence on the TlpB protein level compared to *H. pylori* 26695 wildtype (12G), TlpB expression was increased in a variant carrying 6Gs. A gradual decrease in the TlpB protein level was observed with an increasing number of guanines in the *tlpB* leader variants in the wildtype, reaching a minimum for 9 to 11Gs (Figure 2.9 C, *upper panel*). Further extension of the G-stretch length ranging from 12 to 16 guanines (12-16G) again resulted in a transient increase in TlpB protein level. While deletion of *repG* did not affect TlpB expression in *tlpB* leader variants lacking the G-repeat (ΔG) or comprising 6Gs and 13Gs, increased TlpB protein levels were detected for a G-stretch length of 7 to 12Gs upon sRNA deletion, indicating an optimal window for RepG-mediated *tlpB* repression (Figure 2.9 C, *lower panel*). In line with the observation in *H. pylori* strain G27, a homopolymeric repeat of 14 to 16 guanines (14-16G) in the *tlpB* mRNA leader resulted in a slight down-regulated TlpB protein levels when *repG* was deleted. Since RepG is expressed at similar levels in all *tlpB* mRNA leader mutants (Figure 2.9 B), differences in TlpB expression are likely a result of the variation in the G-repeat length rather than altered sRNA levels. Overall, RepG mediates both, repression and activation of the chemotaxis receptor TlpB dependent on the length of a G-repeat in the mRNA leader.

## 2.8. The length of the homopolymeric G-repeat influences RepG-*tlpB* mRNA interaction

To investigate whether the different G-repeat lengths influence the RepG-*tlpB* mRNA interaction, gel-shift assays and in-line probing experiments were performed with RepG and different *tlpB* mRNA leader variants. Gel-shift assays with 5' end-labeled RepG in the absence or presence of unlabeled *tlpB* leader variants either lacking the homopolymeric G-repeat (ΔG) or comprising different G-stretch lengths (6-16G), showed that RepG efficiently base-pairs with *tlpB* mRNA leaders, which contain a repeat of 9 to 14Gs, with the strongest affinity for variants with 10 to 13Gs (Figure 2.10 A). In contrast, shorter
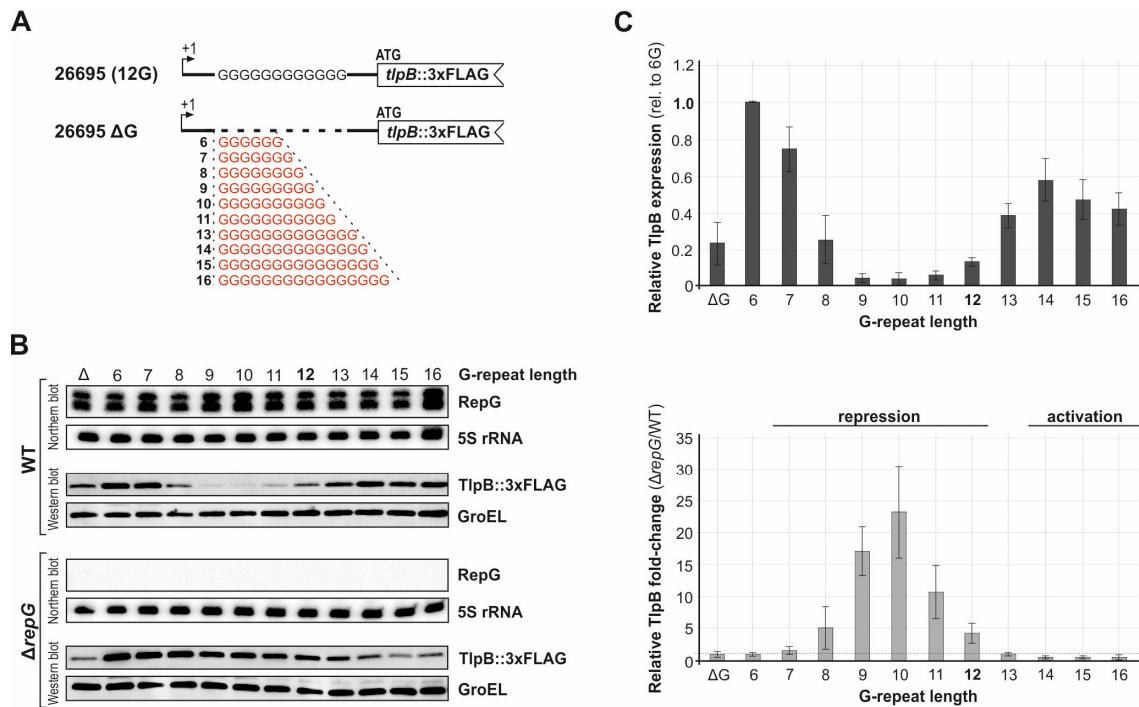
Figure 2.9: Variation of the homopolymeric G-repeat in *H. pylori* strain 26695 determines *tlpB* regulation by RepG. (A) Scheme of the *tlpB* mRNA leader mutants, which either lack the homopolymeric G-repeat (ΔG) or comprise different G-repeat lengths ranging from 6 to 16 guanines (6-16G). All mutants were constructed in a *H. pylori* 26695 *tlpB*::3xFLAG strain. **(B)** Western and northern blot analyses of *tlpB* leader mutants (ΔG, 6-16G) in WT or Δ*repG* backgrounds at exponential growth phase. TlpB::3xFLAG protein was detected using an anti-FLAG antibody; GroEL served as loading control. **(C)** (*Upper panel*) Quantification of the relative TlpB::3xFLAG protein levels in the different *tlpB* leader mutants determined by western blot (B) in *H. pylori* 26695 wildtype. The TlpB protein level in the *tlpB* 6G leader was used as reference and set to 1. (*Lower panel*) Relative fold-changes of TlpB::3xFLAG levels upon *repG* deletion in the *tlpB* leader mutants determined by western blot (B) when compared to the respective WT backgrounds (based on two/three biological replicates). Regulation of the 26695 wild-type *tlpB* mRNA leader (12G) is shown in bold.

(ΔG, 6-8G) or longer G-repeats (14-16G) abolished or reduced the interaction between RepG and the *tlpB* mRNA leader. Differences in the footprint strength observed in the terminator loop of RepG upon addition of different G-repeat variants in in-line probing experiments confirmed that the strength of the interaction between both RNAs is influenced by the *tlpB* G-repeat length (Figure 2.10 B). Reciprocal gel-shift experiments with selected 5' end-labeled *tlpB* variants (ΔG, 10G and 12-14G) and increasing concentrations of RepG confirmed that the chosen *tlpB* leaders bind RepG with different affinities (Figures 2.10 C and D). Overall, the pattern of dissimilar binding affinities for different G-repeat variants closely correlates with the observed pattern of RepG-mediated *tlpB* regulation *in vivo*, indicating that the G-repeat length influences the interaction with RepG and thus, posttranscriptional *tlpB* regulation.
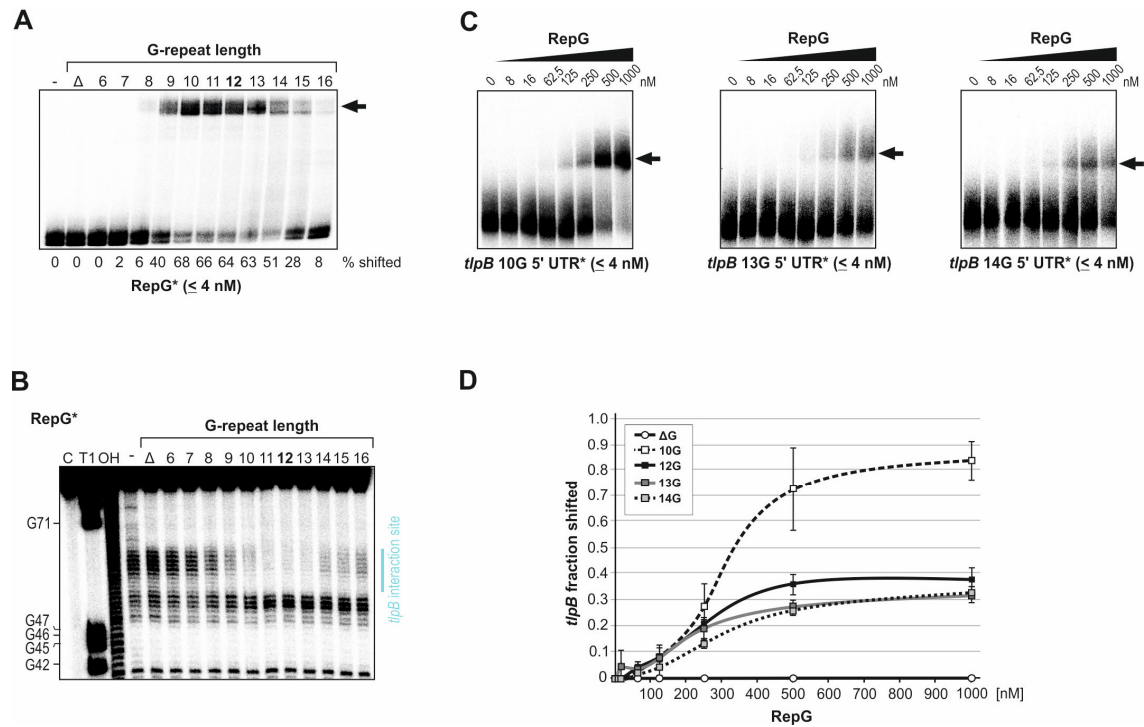
Figure 2.10: Analysis of the interaction between *tlpB* leader variants with different G-repeat lengths and RepG using gel-mobility shifts and in-line probing assays. (A) Gel-shift assays with about 0.04 pmol ^{32}P-labeled RepG in the absence or presence of 1000 nM unlabeled *tlpB* leader variants (ΔG, 6-16G). The arrow indicates RNA-RNA duplex formation, and the amount of shifted RepG* for each variant is given in percent. The results of one representative experiment (out of three) are shown. **(B)** In-line probing of about 0.2 pmol ^{32}P-labeled RepG in the absence or presence of 20 nM *tlpB* mRNA leader variants with indicated G-repeat lengths. The footprint in the RepG terminator loop, which is observed upon addition of several *tlpB* variants, is marked by a blue bar and corresponds to the *tlpB* interaction site. **(C)** Gel-shift assays with 0.04 pmol 5' end-labeled *tlpB* mRNA leaders (10G, 13G and 14G) without or with increasing concentrations of unlabeled RepG (final concentrations are indicated) **(D)** Quantification of the *tlpB* mRNA leader fraction that was shifted when incubated with increasing concentrations of RepG based on the gel-mobility shift assays (C and Figure 2.5 B). Averages and standard deviations were calculated from at least two independent experiments.

## 2.9. RepG regulates *tlpB* translation dependent on the G-repeat length

To study the influence of the G-repeat length on RepG-mediated posttranscriptional control of *tlpB*, the underlying molecular mechanism was further examined. In line with the observation of an increased TlpB protein level upon *repG* deletion in *H. pylori* strain 26695 (Figure 2.3 B), previous quantitative RT-PCR data indicated that the *tlpB* mRNA is also up-regulated in the Δ*repG* mutant compared to the wildtype (Sharma *et al.*, 2010). To examine a potential effect of RepG on *tlpB* mRNA stability, the half-life ($t_{1/2}$) of the *tlpB* mRNA was determined in *H. pylori* 26695 wildtype, *repG* deletion (Δ*repG*), and complementation (C_{RepG}) mutants (Figure 2.11 A). Rifampicin stability assays showed that the *tlpB* mRNA was less

A

H. pylori 26695

B

H. pylori G27
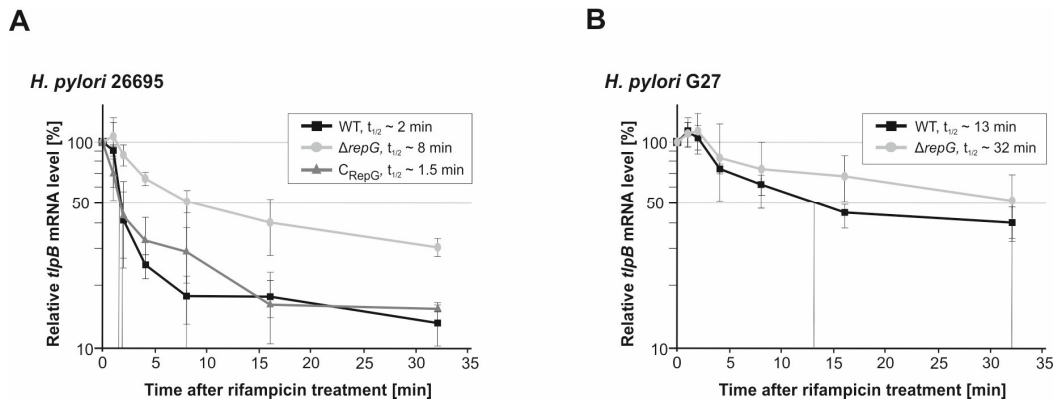
Figure 2.11: RepG reduces *tlpB* mRNA stability. (A) The *tlpB* mRNA half-life at exponential growth phase was determined in *H. pylori* 26695 WT, Δ*repG*, and C$_{RepG}$ using rifampicin assays and quantitative RT-PCR. The *tlpB* mRNA abundance at 0 min was set to 100 % and percentage of *tlpB* mRNA remaining at indicated time points after rifampicin treatment was plotted. The time points at which 50 % of *tlpB* mRNA remained (dotted lines) were used to determine the half-lives (t$_{1/2}$) of the *tlpB* mRNA in the three strains. Averages and standard deviations are based on three biological replicates. **(B)** Determination of the *tlpB* mRNA half-life in *H. pylori* G27 WT and Δ*repG*.

stable in the wildtype (t$_{1/2}$ WT ~ 2 minutes) and complementation strain (t$_{1/2}$ C$_{RepG}$ ~ 1.5 minutes) than in the *repG* deletion mutant (t$_{1/2}$ Δ*repG* ~ 8 minutes), indicating that RepG reduces *tlpB* mRNA stability in *H. pylori* strain 26695. Because RepG-mediated *tlpB* control was strain-specific with opposite outcomes in different *H. pylori* strains (Figure 2.8), the half-life of the *tlpB* mRNA was also determined in strain G27 (Figure 2.11 B). Different decay rates were observed in both wild-type strain backgrounds with *tlpB* mRNA half-lives of 2 and 13 minutes in 26695 and G27, respectively. Despite the fact that in G27 RepG positively affects *tlpB* expression on the protein level (Figures 2.8 B and C), rifampicin assays revealed that the *tlpB* mRNA was slightly less stable in the presence (t$_{1/2}$ WT ~ 13 minutes) than in the absence of *repG* (t$_{1/2}$ Δ*repG* ~ 32 minutes). This suggests that activation of TlpB in *H. pylori* strain G27 is not caused by RepG-mediated *tlpB* transcript stabilization, but may rather occurs at the level of translation.

Changes in mRNA stability might be due to sRNA-mediated interference with translation, as the latter is often coupled to mRNA degradation *in vivo*. To investigate whether RepG interferes with translation initiation, toeprinting assays with *in-vitro* transcribed *tlpB* mRNA leader of *H. pylori* strain 26695 (12G, -139 to +78 according to the annotated start codon) and 30S subunits of *E. coli* ribosomes were performed (Figure 2.12 A). In line with ribosome binding blocking reverse transcription, binding of the 30S ribosomal subunit to the *tlpB* mRNA leader resulted in a tRNA[f-Met]-dependent termination site (shortened cDNA fragment = toeprint) at the +14/+15 position relative to the second potential start codon (AUG) of the *tlpB* mRNA (C++; Figure 2.12 A, lane 3). This indicates that
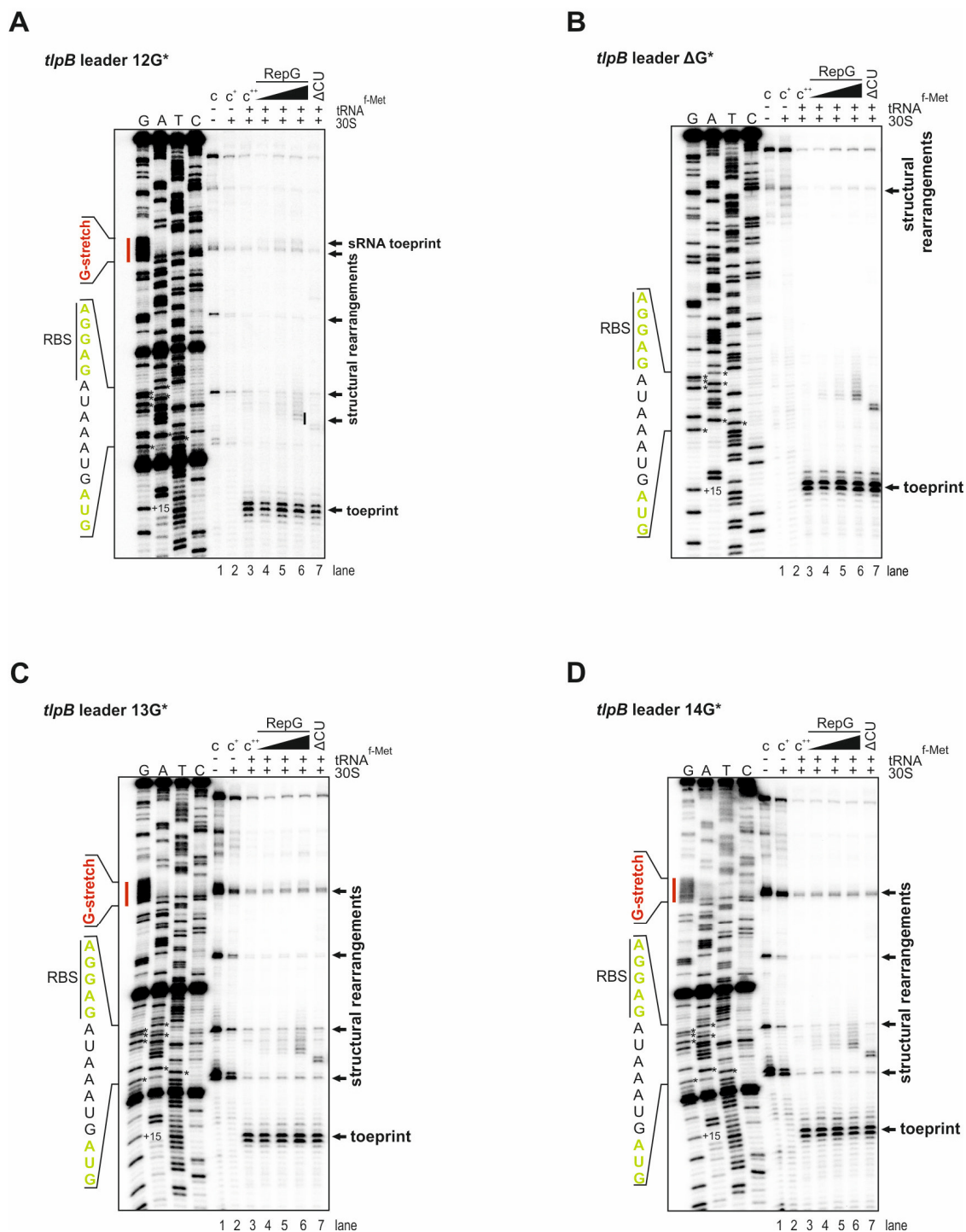
Figure 2.12: RepG does not interfere with translation initiation of the *tlpB* mRNA *in vitro*. Toeprinting assays of *in-vitro* synthesized *tlpB* mRNA leader variants of *H. pylori* strain 26695 (ΔG, 12G, 13G and 14G; 0.2 pmol) in absence (lane 3) or presence of 100, 200 and 1000 nM RepG (lanes 4-6) or 1000 nM RepG ΔCU (lane 7). "-/+" indicates addition of 30S subunit and initiator tRNA[f-Met]. The homopolymeric G-stretch, RBS, and the AUG start codon (2nd codon) are indicated by a red line or green letters and stars, respectively. The toeprint at position +15 respective to the start codon (+3 to annotated AUG), structural rearrangements as well as a potential RepG toeprint are also highlighted by arrows.

the second AUG rather than the annotated start codon is recognized by the *E. coli* ribosome subunits. Moreover, diverse fragments that likely reflect sterical hindrances impeding reverse transcription progression, such as strong RNA secondary structures, were detected at numerous sites in the 5' UTR in case the reaction was performed on *tlpB* alone (C^{--}; Figure 2.12 A, lane 1). However, these signals disappeared upon the addition of 30S and tRNA$^{f\text{-Met}}$ (C^{-+} and C^{++}; Figure 2.12 A, lanes 2-3). Independent of these structural rearrangements, the "toeprint" signal intensity was not affected by increasing concentrations of RepG or RepG ΔCU (Figure 2.12 A, lanes 4-7). Although a putative sRNA toeprint at the homopolymeric G-repeat confirmed that RepG binds to the *tlpB* mRNA under the conditions used in the toeprint assay, even a 50-fold excess of the sRNA did not interfere with translation initiation *in vitro* (Figure 2.12 A, lane 6). Similarly, RepG had no effect on 30S subunit binding in toeprint assays with *tlpB* leader variants that either lack the homopolymeric G-repeat (ΔG) or contain 13 to 14Gs (Figures 2.12 B-D). Noteworthy, multiple cDNA fragments between the RBS and AUG of the *tlpB* mRNA leader variants were detected upon increasing concentration of RepG (*e.g.* highlighted by a black bar in Figure 2.12 A, lane 6), potentially reflecting previously observed structural rearrangements in the 5' UTR of the *tlpB* mRNA (Figure 2.7).

In order to examine whether RepG binding represses translation of *tlpB* at post-initiation steps, *in-vitro* transcribed mRNAs of *tlpB*::3xFLAG and the reporter fusion *tlpB* 5th::*gfpmut3* were translated in absence or presence of RepG using reconstituted *E. coli* ribosomes, and protein synthesis was monitored on western blots (Figure 2.13 A). Reduced TlpB::3xFLAG or TlpB::GFP protein levels were detected upon addition of increasing concentrations of wild-type RepG whereas mutant RepG RNAs had no effect on protein synthesis. Translation of a control mRNA (*cagA* 28th::*gfpmut3* fusion) was not affected by RepG, confirming a specific effect on *tlpB* translation by RepG. Overall, the *in-vitro* translation assays closely recapitulated the observed regulation *in vivo* (Figures 2.3 B and 2.4 B) and indicate that RepG mainly regulates *tlpB* expression at the translational level.

Next, *in-vitro* translation reactions were performed with different *tlpB* leader variants (ΔG, 10-14G) in the absence or presence of RepG (Figure 2.13 B). In line with the *in-vivo* results (Figure 2.9), RepG reduced translation for *tlpB* variants with G-repeat lengths of 10 to 12Gs, had no effect on *tlpB* mRNAs that either lack the G-stretch or contain 13Gs, and slightly increased translation of the 14G-long *tlpB* mRNA. Since the RepG binding site is still present in the *tlpB* leader and binding still occurs (although to lesser extent, Figure 2.10), longer G-repeats (13-14G) might fold into a structure that affects translation of *tlpB* and thereby lead to the reversal of RepG-mediated regulation. Overall, these data indicate that the G-repeat length determines the outcome of RepG-mediated posttranscriptional regulation of *tlpB* and that activation or repression occurs mainly at the translational level.
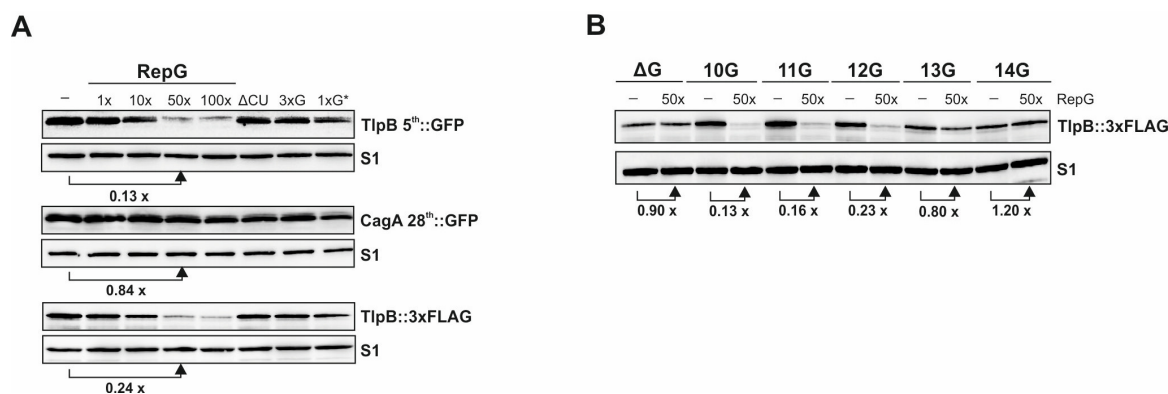
Figure 2.13: RepG mainly represses *tlpB* expression at the translational level. (A) Western blot of TlpB::GFP, TlpB::3xFLAG or CagA::GFP proteins synthesized during *in-vitro* translation assays with 0.1 µM *in-vitro* transcribed *tlpB* 5th::*gfpmut3*, *tlpB*::3xFLAG or *cagA* 28th::*gfpmut3* mRNAs in the absence or presence of 0.1, 1, 5, and 10 µM RepG (1- to 100-fold excess). As control, the effect of 10 µM of RepG mutants ΔCU, 3xG or 1xG* on *tlpB* translation was examined. TlpB::GFP as well as CagA::GFP, and TlpB::3xFLAG were detected by anti-GFP and anti-FLAG antibody, respectively. The ribosomal protein S1 served as loading control. **(B)** *In-vitro* translation assays with *in-vitro* synthesized mRNAs of FLAG-tagged *tlpB* mRNA leader variants ΔG, 10-14G in the absence (-) or presence of 50-fold excess (50x) of RepG. For (A) and (B) one representative western blot (out of two/three experiments) is shown.

2.10. Transcriptional regulation of *repG*

Typically, the expression of sRNAs is activated or repressed in response to environmental stimuli to control gene expression through posttranscriptional mechanisms. To examine the regulation of RepG itself, its expression was monitored under various growth conditions and in available transcriptional regulator mutants of *H. pylori* strains 26695 and/or G27.

2.10.1. RepG accumulates in late exponential growth phase

Expression of RepG was monitored under standard conditions, *i.e.* over growth in nutrient-rich BHI medium under microaerobic conditions at 37°C and 140 rpm. In BHI medium, RepG expression peaked in late-exponential phase, was low in stationary phase, and accumulated in the coccoid form of *H. pylori* strains 26695 and G27 (Figure 2.14 A). Intracellular steady-state transcript levels are influenced by *de-novo* synthesis, processing, transport, interaction with target mRNAs or protein factors, and transcript decay rates. To examine whether variations in RepG levels are due to growth stage-dependent alterations in its stability, the half-life of RepG was determined in *H. pylori* strain 26695 at mid- (OD_{600nm} of 1) and late-exponential phase (OD_{600nm} of 2) (Figure 2.14 B). While the half-life of RepG in mid-exponential growth phase was about four minutes, RepG levels remained nearby constant
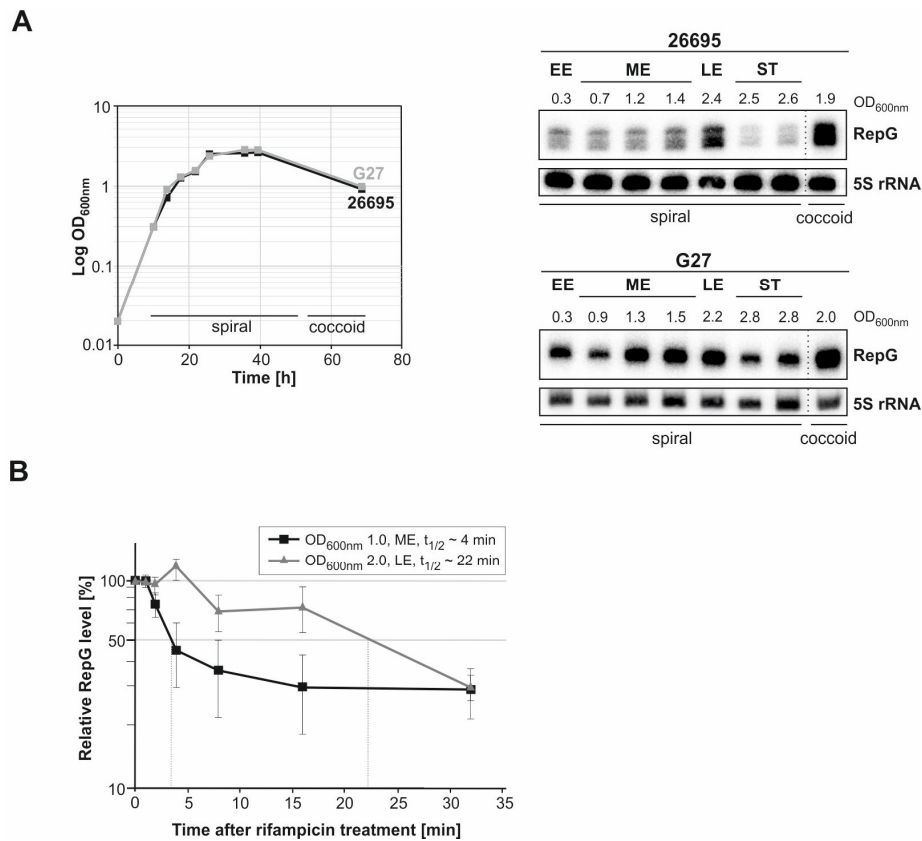
Figure 2.14: Expression and stability of the RepG sRNA over growth. (A) Expression of *repG* was analyzed in *H. pylori* strains 26695 and G27 by northern blot analysis. RNA samples were taken over growth at different OD_{600nm}, which are indicated in the growth curve obtained for 26695 (black) and G27 (gray) in BHI medium (*left panel*). Both strains change their morphology from spiral to coccoid shape after about 60 hours of growth; (EE – early-exponential, ME – mid-exponential, LE – late-exponential, and ST – stationary growth phase). **(B)** Quantification of *in-vivo* stability of RepG in mid- or late-exponential growth phase. The half-life of RepG in different growth phases was determined in *H. pylori* strain 26695 using rifampicin assays and northern blot analysis. RepG sRNA levels prior to addition of rifampicin (0 min) was set to 100 %, and the percentages of sRNA remaining at indicated time points after rifampicin treatment were plotted versus the time. The time points at which 50 % of RepG remained were used to calculate the half-life ($t_{1/2}$; based on at least two biological replicates).

after rifampicin treatment in late-exponential phase ($t_{1/2} \geq 20$ minutes). This indicates that at least one source of the observed RepG accumulation in late-exponential growth phase is its increased transcript stability.

### 2.10.2. HP1043 – a putative transcriptional regulator of *repG*

Since *repG* is encoded adjacent to the orphan response regulator HP1043 (HsrA) in all analyzed *Helicobacter* strains (Figure 2.1 C), a potential role of this transcriptional regulator in the control of *repG* expression was investigated. The dRNA-seq data of *H. pylori* strain
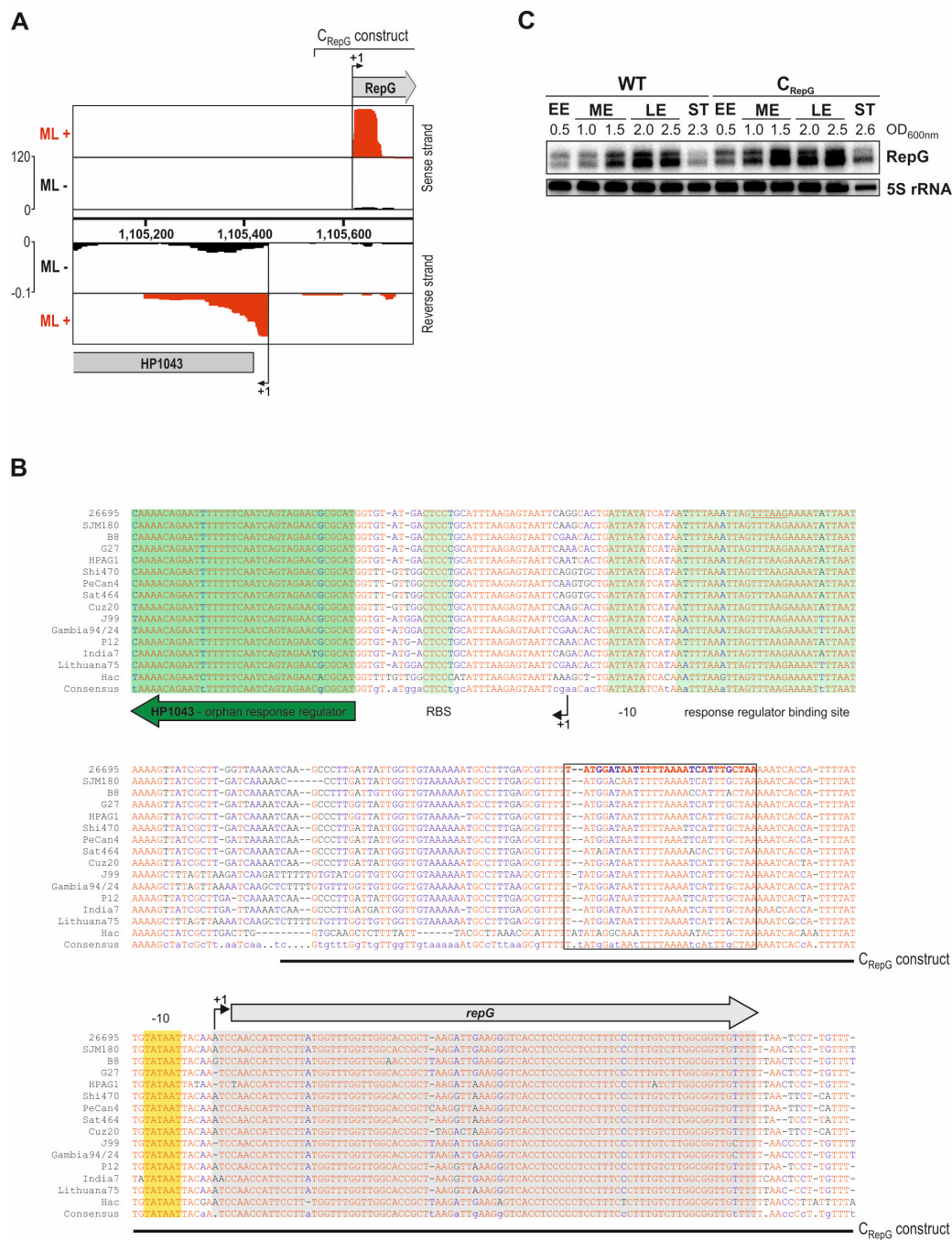
Figure 2.15: The orphan response regulator HP1043 and the RepG sRNA are unlikely to influence each other's expression by transcriptional interference. (A) Screen shot of the IGB genome browser showing the intergenic region of HP1043 and RepG with the dRNA-Seq cDNA coverage plots of the –TEX and +TEX libraries of *H. pylori* strain 26695 from mid-log (ML) growth (Sharma *et al.*, 2010). Gray arrows represent the annotated ORF of HP1043 and the *repG* gene, while black arrows indicate TSS. **(B)** Sequence alignment of HP1043 and *repG* promoter regions in diverse *H. pylori* strains and *H. acinonychis* (Hac). HP1043, the response regulator (RR) binding site, and its -35 (underlined) and -10 promoter box are highlighted in green. TSS are indicated by +1. The promoter of *repG* is marked in yellow. A putative HP1043 binding site in the *repG* promoter region is boxed and shown in bold letters **(C)** Expression of *repG* was analyzed over growth in *H. pylori* 26695 WT and RepG complementation strain (C$_{RepG}$) using northern blot analysis. LE – early-exponential, ME – mid-exponential, LE – late-exponential, ST – stationary growth phase.

26695 revealed that the RepG sRNA and the mRNA of HP1043 are divergently transcribed and that both transcripts do not overlap (Figure 2.15 A). A conserved transcriptional start site and promoter region have been identified for both HP1043 and *repG* (Figure 2.15 B, Sharma *et al.*, 2010). In addition, HP1043 has been shown to bind an inverted sequence repeat (A<u>TT</u>AxTx<u>T</u>xTxx<u>TTAA</u>xCxx<u>A</u>xxTx<u>AA</u>Ax), also referred to as the response regulator binding site (Delany *et al.*, 2002), which overlaps with the -35 box of the HP1043 promoter. Based on the spatial separation of the promoter regions of HP1043 and *repG*, it appears unlikely that the two genes would influence each other in *cis, e.g.* via transcriptional interference. Indeed, comparative expression profiling between the *H. pylori* 26695 wildtype and the *trans*-complemented *repG* strain (C$_{RepG}$) argue against a direct coupling of *repG* and HP1043 expression. That is, wild-type expression of RepG was observed over growth for the complementation strain (Figure 2.15 C), in which the sRNA is expressed under its native promoter from the unrelated *rdxA* locus, but far distant from to the HP1043 promoter (underlined sequence in Figure 2.15 B).

Rather than a direct transcriptional coupling of the two genes, HP1043 could control RepG expression in *trans*, namely via a binding of dimeric HP1043 to the *repG* promoter region. In collaboration with Prof. Dr. Dagmar Beier (Biocenter, University of Würzburg, Germany), electrophoretic gel-mobility shift assays with recombinant HPG27_385 protein (HP1043 homolog of *H. pylori* strain G27) and biotinylated PCR products covering the promoter region of the RepG sRNA (P$_{repG*}$, 118 nt upstream and 22 nt downstream of *repG* TSS of *H. pylori* strain G27) were performed (Figure 2.16 A). A shift of the P$_{repG*}$ PCR fragment upon addition of HPG27_385 protein demonstrated that the response regulator (RR) binds to the *repG* promoter region *in vitro* (Figure 2.16 B). After initial RR/P$_{repG*}$ complex formation, increasing concentrations of unlabeled PCR products were added and their ability to compete with RR/P$_{repG*}$ interaction was determined. Whereas competition with unlabeled P$_{repG}$ PCR products resulted in abrogation of the shift signal (Figure 2.16 B, lanes 3-4), addition of an unlabeled PCR product that covers the unrelated promoter region of the DNA-directed RNA polymerase (*rpoA*) hardly affected complex formation (Figure 2.16 B, lanes 9-10). This indicates that the *repG* promoter region is specifically recognized and bound by the response regulator. Furthermore, competition experiments with non-labeled PCR products that comprising defined fragments of the *repG* promoter (USC = unspecific competitor from -118 nt to -60 nt upstream of the TSS, SC = specific competitor from -59 nt to +22 nt relative to the TSS) showed that the response regulator binds within a 60-nt window upstream of the *repG* transcriptional start site (Figure 2.16 B, lanes 5-8). In line with the previously described HP1043 binding motif (Delany *et al.*, 2002), a putative response regulator binding site from -57 to -29 relative to the TSS of *repG* was predicted (Figure 2.16 A), which is conserved among several *H. pylori* strains (Figure 2.15 B).
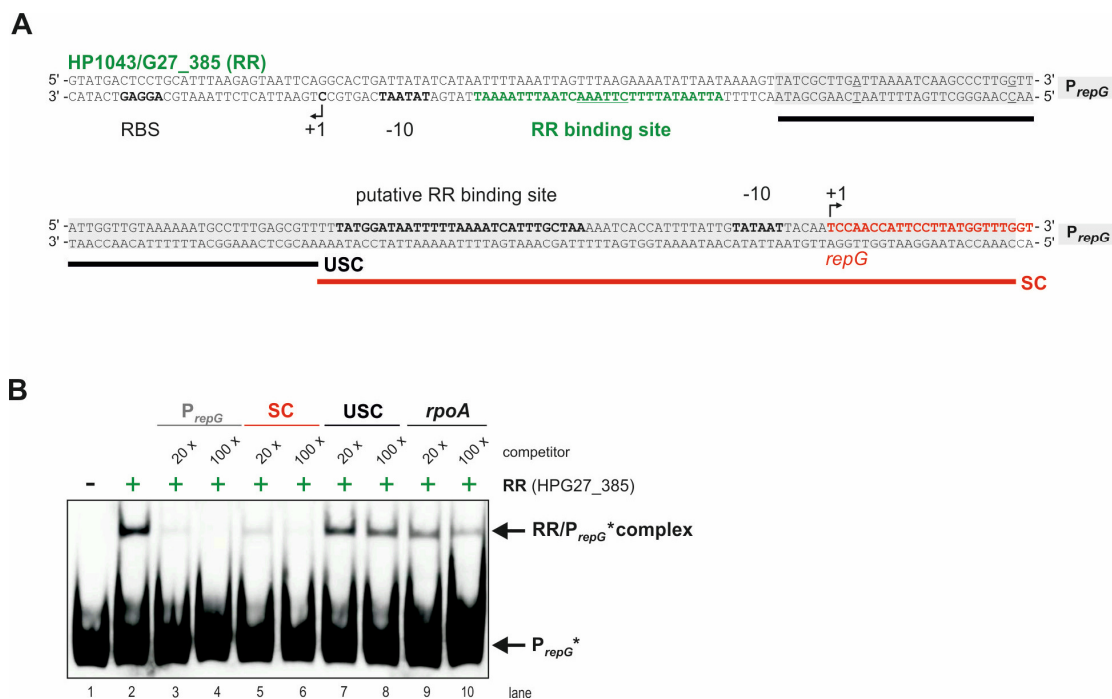
Figure 2.16: Electrophoretic gel-mobility shifts of HPG27_385 (HP1043 homolog) and *repG* promoter region. (A) IGR between the orphan response regulator (RR) and *repG* (red letters) including both promoter regions in *H. pylori* G27; +1 marks TSS. The -10 boxes and HP1043 RBS are shown in bold. The RR binds its own promoter region at the RR-binding site (green), which overlaps with the -35 box (underlined). The sequences of the biotinylated PCR products used in (B) are shown in different colors. The full *repG* promoter fragment (P*repG**) is boxed in gray. USC (black) = unspecific competitor from -118 nt to -60 nt upstream of the *repG* TSS, SC (red) = specific competitor from -59 nt to +22 nt relative to the *repG* TSS. A potential RR binding site in the *repG* promoter is indicated by bold letters. **(B)** Band-shift assays of 0.02 pmol biotinylated *repG* promoter fragment (P*repG**) in absence (lane 1) or presence of 2 µg of HPG27_385 (lane 2). For competition experiments, 20- to 100-fold excess of unlabeled PCR fragments comprising the promoter region of the DNA-directed RNA-polymerase subunit (*rpoA*, lanes 9-10), the full *repG* promoter region (P*repG*, lanes 3-4), 59 nt upstream of *repG* TSS (SC, lanes 5-6) or 61 nt upstream of putative RR binding site (USC, lanes 7-8) were added after initial complex formation between P*repG** and HPG27_385. Arrows indicate either the complex of HPG27_385 and the *repG* promoter or the free, labeled sRNA promoter fragment (P*repG**).

To investigate whether HP1043 or its homolog, HPG27_385, affect *repG* expression *in vivo*, HP1043 overexpression mutants (HP1043-OE/HPG27_385-OE) were constructed, which harbor an additional copy of the response regulator transcribed from the *cagA* promoter in the unrelated *rdxA* locus of *H. pylori* strains 26695 and G27 (kindly provided by Prof. Dr. Dagmar Beier). Introduction of an additional gene copy resulted in a two- to three-fold increase in HP1043/HPG27_385 protein levels. However, despite the capability of HPG27_385 to bind to the *repG* promoter *in vitro* (Figure 2.16 B), wild-type RepG levels were detected in the HP1043/HPG27_385-OE strains (Figure 2.17 A). This indicates that the response regulator does not control RepG expression *in vivo*, at least under the examined conditions.
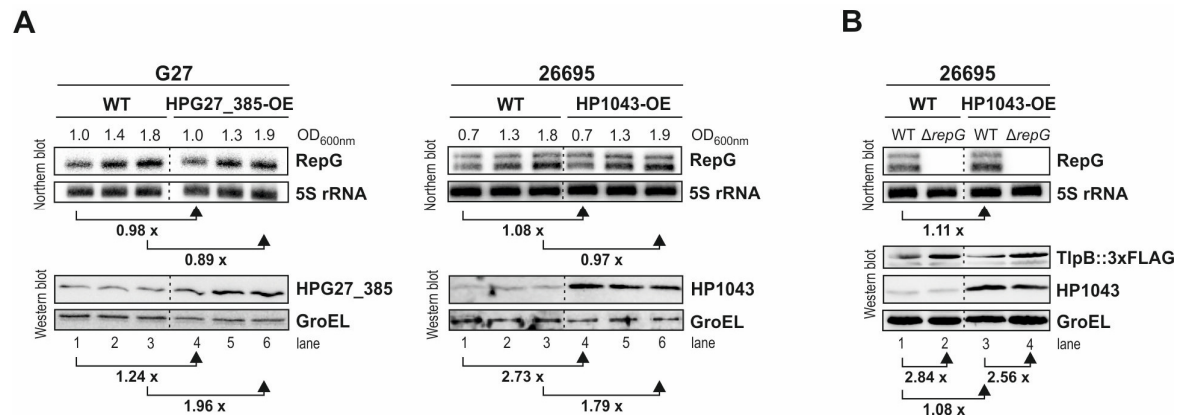
Figure 2.17: Overexpression of the response regulator HP1043/HPG27_385 does not affect RepG expression in *H. pylori* strains 26695 and G27. (A) Western blot and northern blot analysis of protein and RNA samples taken over growth (at indicated OD_{600nm}) from *H. pylori* G27 and 26695 wildtype (WT) and HPG27_385/HP1043 overexpression mutants (OE). For western blot analysis, protein samples corresponding to an OD_{600nm} of 0.006 (G27) or 0.01 (26695) were separated on 12 % SDS-PAGE gel, blotted onto PVDF-membrane, and detected using polyclonal anti-HP1043 antiserum. The chaperone GroEL was used as loading control. **(B)** RepG-mediated repression of the epitope-tagged TlpB in *H. pylori* 26695 WT or HP1043-OE was analyzed by western blot. TlpB::3xFLAG was detected by an anti-FLAG antibody.

Although the target genes of HP1043 are largely unknown, DNA-footprinting experiments with recombinant HP1043 suggested its binding to the promoter of *tlpB* (Delany *et al.*, 2002). Nevertheless, northern and western blot analysis of whole protein and RNA samples taken in exponential growth phase from *H. pylori* 26695 wildtype, Δ*repG*, HP1043-OE and HP1043-OE/Δ*repG* mutants, which carry a *tlpB*::3xFLAG at the native locus in the chromosome, revealed that neither RepG transcript nor TlpB::3xFLAG protein levels are affected by the overexpression of the response regulator (Figure 2.17 B, lanes 1 and 3). Furthermore, TlpB::3xFLAG protein levels were similarly increased (about two- to three-fold) upon *repG* deletion in the 26695 wildtype and HP1043-OE mutant (Figure 2.17 B), indicating that the response regulator is dispensable for RepG-mediated *tlpB* repression.

Noteworthy, the increased expression of the response regulator was only transient in HP1043-OE/G27_385-OE mutants as it dropped to wild-type levels after several passages on plate and/or extended growth in liquid culture (data not shown). This is in line with the previously described tight expression control of HP1043 at the transcriptional, posttranscriptional and posttranslational level (Muller *et al.*, 2007) and suggests that *H. pylori* counteracts changes in the amount of this protein. Due to the essentiality of the response regulator and the problems in manipulating its expression, the here described results of the regulation of *repG* expression by HP1043 *in vivo* are still preliminary. Further studies including the construction of a conditional mutant with weakly-expressed HP1043 homologs from other species such as HPMG439 and Cj0355 from *Helicobacter pullorum* and

C. jejuni, respectively (Bauer *et al.*, 2013, Muller *et al.*, 2007), are required to assess a potential role of HP1043 in the regulation of *repG*.
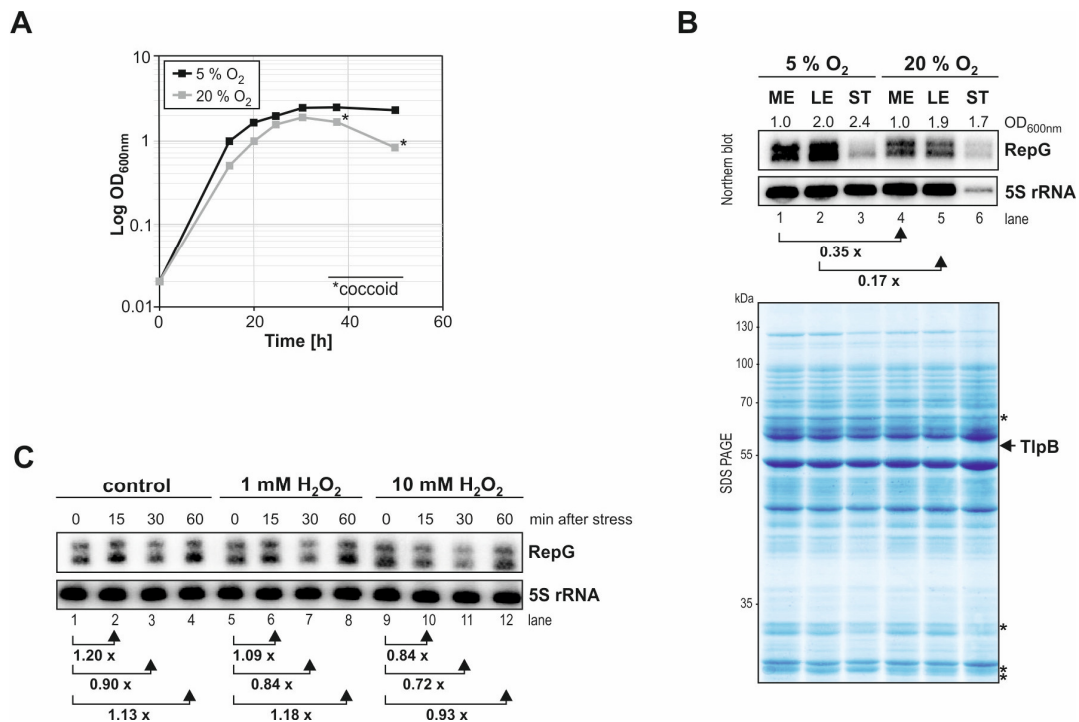

2.10.3. Oxygen stress represses RepG expression

During host colonization, *H. pylori* induces a chronic inflammatory host response, which often results in the generation of reactive oxygen (ROS) and nitrogen species (RNS) by gastric epithelial cells (Nardone *et al.*, 2004). Compared to other Gram-negative bacteria, *H. pylori* lacks homologs of the oxidative stress response regulatory genes *rpoH*, *rpoS*, *soxRS*, and *oxyR* (Tomb *et al.*, 1997). However, the orphan response regulator HP1043 has been assumed to be involved in oxidative stress management in *H. pylori* (Olekhnovich *et al.*, 2014). To examine whether RepG is involved in the combat against oxidative stress, RepG expression was monitored in *H. pylori* strain 26695 grown (I) under microaerobic and atmospheric conditions or (II) in the absence or presence of the powerful oxidant hydrogen peroxide (H_2O_2).

The optimal *in-vitro* growth condition for *H. pylori* is between 5 to 8 % oxygen (O_2) partial pressure (Bury-Mone *et al.*, 2006). However, recent studies suggested that *H. pylori* is a capnophilic aerobe as its growth can be stimulated by atmospheric oxygen levels in the presence of 10 % carbon dioxide (CO_2) (Park & Lee, 2013). *H. pylori* strain 26695 was grown under microaerobic (5 % O_2, 10 % CO_2) and atmospheric oxygen (20 % O_2, 10 % CO_2) conditions in order to simulate long-term oxidative stress. *H. pylori* showed a modest growth defect under atmospheric O_2 levels (Figure 2.18 A). In addition, an increase in the O_2 partial pressure inhibited long-term survival of *H. pylori* as deduced from a rapid drop in the OD_{600nm} and a partial transformation from spiral to coccoid morphology ($\sim$ 30 % of bacterial cells) when bacteria entered stationary phase. Northern blot analysis revealed about three to five-fold reduced RepG levels in exponential phase at higher O_2 tensions (Figure 2.18 B). Since oxygen stress not only affected *H. pylori* morphology and survival, but also impaired with the total RNA content of the cells (as indicated by a decrease in 5S rRNA, Figure 2.18 B), no statements could be made on how different O_2 tensions would affect RepG expression in stationary phase. Besides, only minor changes in the overall protein content were observed for *H. pylori* grown under microaerobic and atmospheric oxygen conditions (marked by asterisks in Figure 2.18 B). Despite RepG levels being slightly reduced, TlpB protein levels were not affected by different O_2 partial pressures.

The decreased RepG levels observed under long-term oxidative stress might be due to slower growth and thus, changes in the overall transcription or sRNA turnover rates. Therefore, RepG expression was also determined upon H_2O_2-induced oxidative stress

Figure 2.18: Atmospheric oxygen tensions repress RepG expression. (A) *H. pylori* strain 26695 was grown under microaerobic (5 % O₂, 10 % CO₂, black) and atmospheric oxygen (20 % O₂, 10 % CO₂, gray) conditions. *H. pylori* changed from spiral to coccoid morphology when exposed to atmospheric O₂ levels for more than 30 hours. Whole RNA and protein samples were taken over growth and used for expression analysis of RepG by northern blot (B – *upper panel*). (ME- mid-exponential, LE – late exponential, ST – stationary). **(B)** (*Lower panel*) Protein samples corresponding to an OD$_{600nm}$ of 0.1 were separated on 12 % SDS-PAGE and stained with Coomassie. Oxygen-responsive changes in protein abundances are indicated by asterisks. **(C)** *H. pylori* strain 26695 was grown to exponential growth phase (OD$_{600nm}$ of ~ 0.8) and split into three independent subcultures. One subculture was left untreated (control), whereas the other two were shifted to mild (1 mM H₂O₂) or strong (10 mM H₂O₂) oxidative stress. RNA samples were taken prior to or at indicated time points after H₂O₂ treatment.

(Figure 2.18 C). In line with the decrease in RepG expression under high O₂ tensions (Figure 2.18 B), northern blot analysis revealed slightly reduced RepG levels (~ 1.4-fold) when *H. pylori* was exposed to 10 mM H₂O₂ for 30 minutes (Figure 2.18 C, lanes 1,3 and 9,11).

Overall, these data indicate that RepG expression might be oxygen-responsive; however, the exact trigger of RepG expression and its transcriptional regulator(s), *e.g.* HP1043, remain elusive. RepG expression analyses under oxidative and nitrosative stress conditions induced by chemical or physical triggers, such as the superoxide generator paraquat, UV-radiation, S-nitrosoglutathione, and redox-active antibiotics (*e.g.* metronidazole), in the wildtype and/or conditional HP1043 mutants will be required to investigate a potential role of RepG and/or HP1043 in the oxidative stress management of *Helicobacter*. In addition, viability assays with the wildtype and *repG* mutant could be used to examine whether RepG affects the sensitivity of *H. pylori* to ROS/RNS.

2.10.4. Expression of RepG is induced under acidic stress conditions

One of the most important triggers for gene expression in *H. pylori* is the acidic milieu of the human stomach (Merrell *et al.*, 2003a, Wen *et al.*, 2003). To investigate whether RepG is expressed in response to different pH, *H. pylori* strain 26695 was exposed to acidic (pH 5.0) or alkaline (pH 8.5) stress conditions (Figure 2.19 A). Northern blot analysis of RNA samples taken at indicated time points after pH adjustment/shift of the growth medium revealed that RepG expression is pH-dependent. Whereas RepG is rapidly induced (about two-fold) under acidic stress, sRNA expression is decreased (about five-fold) when *H. pylori* encounters alkaline environments.
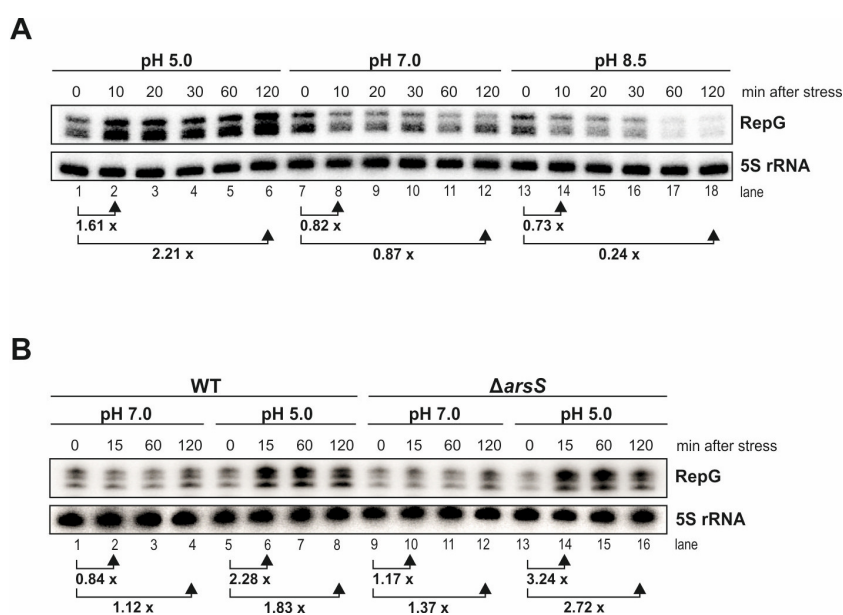
Figure 2.19: Induction of RepG under acidic stress is independent of the acid-sensing ArsRS two-component system. (A) *H. pylori* strain 26695 was grown in BHI medium (pH 7.0) to exponential growth phase (OD_{600nm} of ~ 0.8) and split into three independent subcultures. One subculture was left untreated, whereas the media of the other two were adjusted to pH 5.0 (*left*) or pH 8.5 (*right*) with 37 % HCl or 1 M NaOH, respectively. Samples for RNA extraction were withdrawn prior to or at indicated time points after pH adjustment. Expression of RepG was determined by northern blot analysis. **(B)** RNA samples were taken from *H. pylori* 26695 wildtype (WT) and *arsS* deletion mutant (Δ*arsS*) grown under neutral pH (pH 7.0) and acid stress (pH 5.0) conditions. Experimental set-up as described in A.

The majority of genes involved in acid adaptation in *H. pylori* are controlled by the acid-sensing ArsRS two-component system (TCS) (reviewed in Pflock *et al.*, 2006a). While the response regulator ArsR is essential, the histidine kinase gene *arsS* can be deleted without affecting *Helicobacter's* growth (Beier & Frank, 2000, McDaniel *et al.*, 2001). To experimentally assess whether or not the ArsRS TCS is responsible for pH-dependent

induction of RepG, sRNA expression levels in the presence of acidic stress were compared between *H. pylori* 26695 wildtype and an *arsS* deletion mutant (Figure 2.19 B). Upon the shift to pH 5.0, increased sRNA levels were detected in both the wildtype and *arsS* deletion mutant, indicating that acid-mediated RepG induction is independent of the ArsRS TCS.

2.10.5. Iron-dependent activation of the RepG sRNA

H. pylori infectivity is largely dependent on the expression and maturation of metallo-enzymes, which are involved in acid adaptation, respiration and detoxification. The activity of these proteins depend on the presence of specific transition metal ions such as nickel (Ni^{2+}) and iron (Fe^{2+}). Consequently, these metal ions are important for the persistence of *H. pylori* in the gastric niche (Bury-Mone *et al.*, 2004, van Vliet *et al.*, 2004). To investigate whether RepG expression is affected by the availability of nickel or iron, *H. pylori* 26695 wildtype was grown in the absence or presence of nickel chloride ($NiCl_2$), iron (II)-sulfate ($Fe(II)SO_4$) and high-affinity iron chelator 2,2' dipyridyl (DPP) (Figure 2.20 A). Almost no change in RepG abundance ($\sim$ 1.2-fold) was observed when *H. pylori* was challenged with $NiCl_2$ (Figure 2.20 A, lane 2). This suggests that in *H. pylori* strain 26695, transcription from the *repG* promoter and/or RepG stability are nickel-independent. In contrast, expression of RepG was about two-fold induced under iron-replete conditions (Figure 2.20 A, lane 3), indicating that the *repG* gene might be iron-responsive. Accordingly, slightly reduced RepG levels ($\sim$ 1.4-fold) were observed when *H. pylori* encountered iron limitation (Figure 2.20 A, lane 4).

In *H. pylori*, the ferric-uptake regulator Fur is a key player in iron homeostasis and has been shown to repress and/or activate gene expression in response to iron availability, low pH, and oxidative and salt stress (reviewed in Danielli & Scarlato, 2010 and Pich & Merrell, 2013). To examine whether Fur is involved in iron-dependent induction of RepG, sRNA expression levels under standard growth conditions were compared between the *H. pylori* 26695 wildtype and a *fur* deletion mutant (Δfur) (Figure 2.20 B). Upon deletion of *fur*, slightly reduced RepG sRNA levels ($\sim$ 1.6-fold) were detected, suggesting that the iron homeostasis regulator might activate *repG* transcription. However, no conserved binding motif for either the apo- (iron-free) or the holo- (iron-bound) form of Fur (Fur-box motifs: TAATAATnATTATTA, TGATAATnATTATCA or $TCATTn_{10}TTAAATGA$, Pich *et al.*, 2012, Agriesti *et al.*, 2014 and Carpenter *et al.*, 2013) could be identified within or upstream of the *repG* core promoter region. At this point of the study, it cannot be excluded that the deletion of *fur* indirectly affects RepG expression at the transcriptional and/or posttranscriptional level, *e.g.* through Fur-mediated transcriptional control of transcriptional regulators
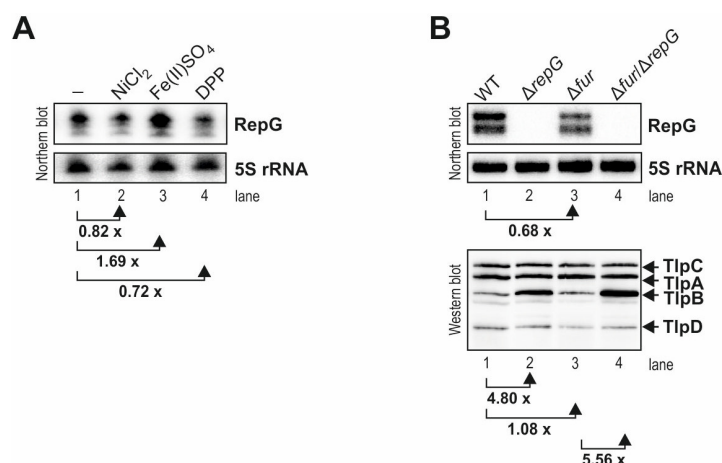
Figure 2.20: RepG expression is induced under iron-replete conditions and might be activated by the ferric-uptake regulator Fur. (A) *H. pylori* strain 26695 was grown in BHI medium to exponential growth phase (OD_{600nm} of ~ 0.6) in absence or presence of 20 μM nickel chloride ($NiCl_2$), 100 μM iron (II)-sulfate ($Fe(II)SO_4$) and 60 μM of the high-affinity iron chelator 2,2' dipyridyl (DPP). Expression of RepG was determined using northern blot analysis with 5' end ^{32}P-labeled CSO-0003. **(B)** RNA samples were taken from *H. pylori* 26695 WT, Δ*repG*, Δ*fur* and Δ*fur*/Δ*repG* grown to exponential growth phase and RepG levels were analyzed as described in A. Protein samples corresponding to an OD_{600nm} of 0.01 were loaded on 12 % SDS-PAGE, blotted to PVDF membrane and the chemotaxis receptors were detected with a polyclonal rabbit anti-TlpA22 antiserum.

such as NikR, ArsRS, RpoN, FlgRS and CheA (Danielli *et al.*, 2006), or an RNA-binding protein that might influence RepG stability. Expression profiling of RepG in the wildtype and the Δ*fur* mutant under iron-restricted or iron-replete conditions, low pH and oxidative stress will help to understand if/to what extent Fur controls transcription of *repG*. In addition, electrophoretic gel-mobility shift assays with purified Fur and the *repG* promoter sequence as well as genetic approaches (*e.g.* the exchange of the *repG* promoter) could be used to investigate whether the *repG* gene is a direct target of the ferric-uptake regulator.

Although expression of the chemotaxis receptor *tlpB* is regulated in response to iron availabilies (Merrell *et al.*, 2003b), its transcriptional control seems to be independent of Fur (Ernst *et al.*, 2005). Accordingly, deletion of *fur* did not affect TlpB protein levels in *H. pylori* strain 26695, albeit slightly reduced RepG sRNA levels were detected in the Δ*fur* mutant (Figure 2.20 B, lanes 1 and 3). This suggests that RepG expression levels are not limiting for *tlpB* regulation under the examined conditions. In line with this, about five-fold increased TlpB protein levels were detected in both the wildtype and the Δ*fur* mutant upon *repG* deletion (Figure 2.20 B, lanes 2 and 4), indicating that Fur is dispensable for RepG-mediated TlpB repression in *H. pylori* strain 26695, at least under standard growth conditions.

2.11. Protein factors potentially involved in the RepG-*tlpB* mRNA interaction

In-vitro translation assays suggested that RepG is able to control *tlpB* expression independent of additional protein factors (Figure 2.13). However, this might not necessarily recapitulate the *in-vivo* situation. A preliminary affinity purification approach using *in-vitro* transcribed and 5' aptamer-tagged *tlpB* mRNA leaders derived from either *H. pylori* strains 26695 (12G) or G27 (14G) and whole cell lysate of *Helicobacter* suggested that amongst others, RNase J potentially interacts with the *tlpB* mRNA (Figure 2.21 A). In *H. pylori,* mRNA degradation is thought to be predominantly mediated by a minimal RNA degradosome consisting of RNase J and the DExD-box RNA helicase RhpA (Redko *et al.*, 2013). However, it is yet unclear whether RNase J is also involved in sRNA-mediated posttranscriptional regulation. Initial expression profiling of *tlpB* in a conditional *H. pylori* B128 RNase J mutant, which lacks the functional N-terminal domain of RNase J (P$_i$ΔN-*rnj*; kindly provided by PhD Hilde De Reuse, Institute Pasteur, Paris, France), suggested that this ribonuclease might be involved in *tlpB* mRNA degradation (Figure 2.21 B). Surprisingly, despite *tlpB* mRNA levels being increased, no changes were detected at the TlpB protein level upon RNase J inactivation (Figure 2.21 C) In *H. pylori* strain B128, the *tlpB* mRNA leader contains an 8-nt long homopolymeric G-repeat (McClain *et al.*, 2009), which may be suggestive of a repression of the chemotaxis receptor by RepG in this strain background. In line with RNase J being dispensable for RepG stability (Figure 2.21 C), RepG-mediated posttranscriptional control of *tlpB* might be responsible for unaltered TlpB protein levels in the P$_i$ΔN-*rnj* mutant when compared to the wildtype.

Ribosomal protein S1 was previously shown to interact with various mRNAs and sRNAs in *H. pylori* strain 26695 (Rieder *et al.*, 2012). Accordingly, the 5' UTR of the *tlpB* mRNA and the RepG sRNA were found to potentially bind to S1 (Figures 2.21 A and 2.22 A). In addition, preliminary expression analysis of RepG in an *rpsA* deletion mutant (S1, Rieder *et al.*, 2012) revealed that S1 might affect RepG expression and/or stability (Figure 2.22 B). As the Δ*rpsA* mutant displayed reduced growth rates and acquired compensatory mutations after prolonged *in-vitro* passaging of *H. pylori* on solid medium (data not shown), further experiments will be required to address whether decreased RepG levels in the Δ*rpsA* mutant are due to slower growth and thus, changes in the overall transcription or sRNA turnover rates, or if S1 specifically stabilizes RepG. Interestingly, deletion of *rpsA* and thus, reduced RepG levels did not significantly affect TlpB protein levels (Figure 2.22 B).
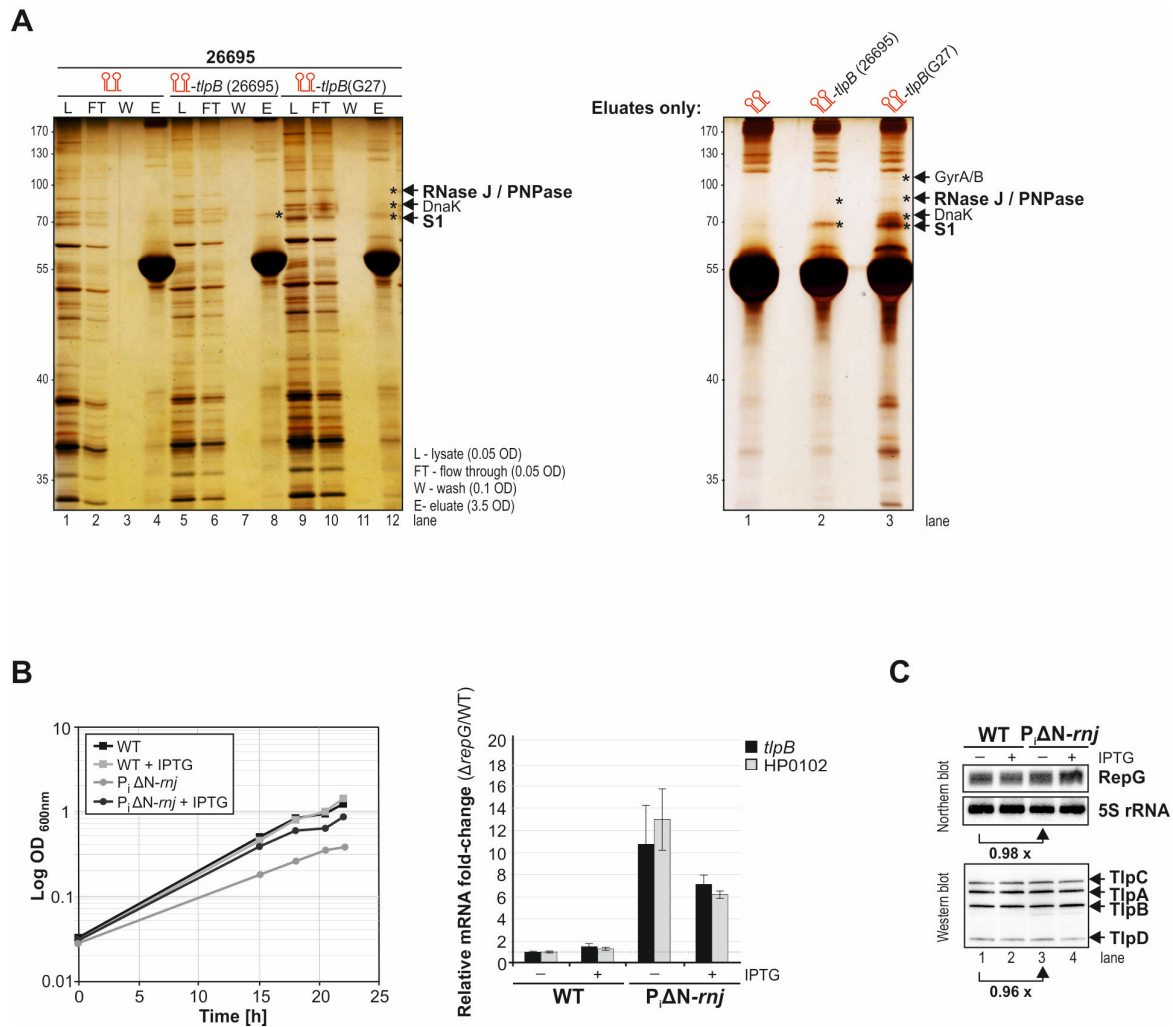
Figure 2.21: RNase J and ribosomal protein S1 potentially interact with the *tlpB* mRNA. (A) Affinity chromatography of aptamer-tagged, *in-vitro* transcribed *tlpB* mRNA leaders (originating from *H. pylori* strains 26695 and G27). Aptamer-tagged RNAs were incubated with whole cell lysate of *H. pylori* strain 26695, followed by affinity purification according to Rieder *et al.*, 2012. *In-vitro* transcribed MS2-tag alone was used as control. (*Left panel*) Protein fractions of the different steps of the affinity purification were separated on 12 % SDS-PAGE and stained with silver. (*Right panel*) Prominent protein bands in the elution fraction were excised and analyzed by mass spectrometry. **(B)** Growth characteristics of *H. pylori* B128 WT (wildtype containing pILL2157) and RNase J-deficient mutant $P_i\Delta N$-*rnj* (Δrnj containing pHP135; pILL2157 with N-terminal truncation of *rnj* gene under control of an IPTG-inducible promoter - P_i) in BHI medium. Growth in absence (-IPTG) or presence of inducer (+IPTG) was used to control expression of ΔN-*rnj*. Total RNA and protein samples were taken at early exponential growth phase (OD_{600nm} of ~ 0.4). **(C)** Quantitative RT-PCR analysis of *tlpB* and HP0102 expression in *H. pylori* B128 WT and $P_i\Delta N$-*rnj*. After DNase I treatment, RNA samples were used in one-step qRT-PCR reactions with specific primer pairs for *tlpB* and HP0102. Wild-type mRNA levels were set to 1, and relative fold-changes in the mutant strain are shown as bars. Values are shown as mean ± standard derivations from at least two experiments. **(D)** Using northern blot analysis, RepG expression was monitored in the wildtype and $P_i\Delta N$-*rnj* mutant. In addition, protein samples corresponding to an OD_{600nm} of 0.01 were loaded on 10 % SDS-PAGE, blotted to PVDF membrane and the chemotaxis receptors were detected by a polyclonal rabbit anti-TlpA22 antiserum.
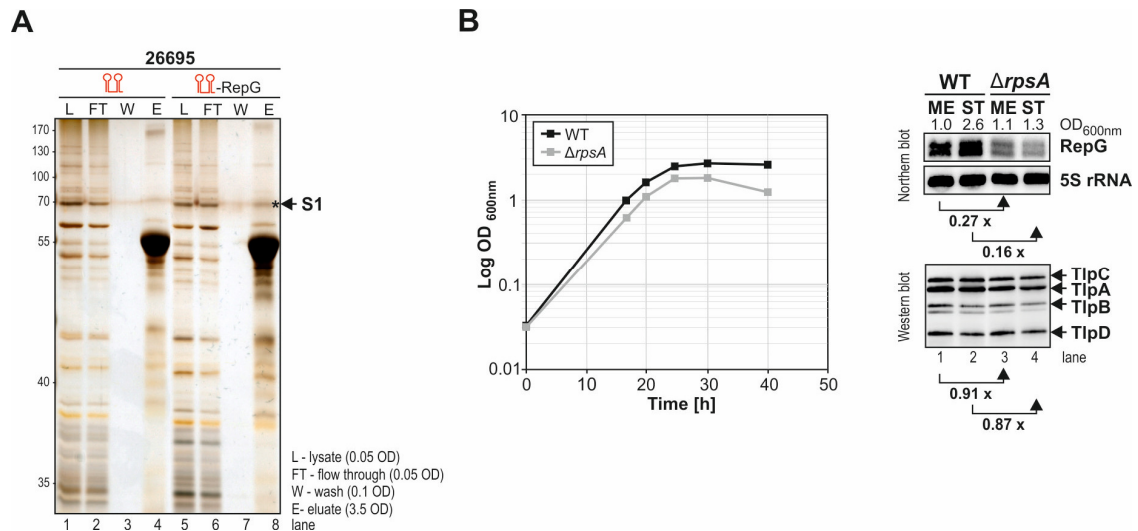
Figure 2.22: Ribosomal protein S1 affects RepG stability. (A) *In-vitro* transcribed, aptamer-tagged RepG was incubated with whole cell lysate of *H. pylori* strain 26695, followed by affinity chromatography according to Rieder *et al.*, 2012. The co-eluted protein band marked by an asterisk was identified as ribosomal protein S1 by mass spectrometry. **(B)** *H. pylori* strain 26695 wildtype (WT) and S1-deficient mutant (Δ*rpsA*) were grown in nutrient-rich BHI medium, and total RNA and protein samples were taken at mid-exponential (ME) and stationary growth phase (ST). Using northern blot and western blot analysis, expression of RepG and TlpB was monitored by P^{32}-labeled oligonucleotide CSO-0003 and a polyclonal rabbit anti-TlpA22 antiserum, respectively.

2.12. Discussion: A simple sequence repeat determines sRNA-mediated gene regulation in *Helicobacter pylori*

Phase variation of hypermutable SSRs is a widespread and stochastic mechanism to generate phenotypic variation within a population and thereby, contributes to host adaptation of bacterial pathogens. While several examples of SSRs have been reported to impact on transcriptional activity or affect the coding potential, work presented in this Chapter of the thesis uncovered an SSR that determines posttranscriptional regulation by a bacterial sRNA. In particular, dependent on its length, the G-repeat in the 5' UTR of the *tlpB* mRNA mediates both repression and activation of this chemotaxis receptor through RepG. This modulation of *tlpB* expression by length variation of an SSR represents a new twist in sRNA-mediated regulation and connects it with gene expression control and phenotypic variation through variable repeats. This exemplifies how mutations within the bacterial genome can impact on posttranscriptional control mechanisms and represents a novel paradigm for the interdependence of the fundamental layers of gene expression. The posttranscriptional control mediated by sRNA binding to an SSR within a 5'UTR allows for a gradual modulation of gene expression and is in contrast to the digital ON/OFF switches associated with intragenic SSRs that typically cause frame-shift mutations within ORFs.

Likewise, intergenic SSRs mainly result in strong, moderate or low gene expression, due to changing the spacing of promoter elements or transcription factor binding sites. Since phase variation by SSRs has been shown to facilitate host adaptation of numerous bacterial pathogens (Bayliss, 2009, Moxon *et al.*, 2006), SSR-dependent regulation of gene expression at the posttranscriptional level might turn out to be an important means to adjust and fine-tune virulence gene expression during infection not only in *Helicobacter,* but also in other pathogens.

RepG binds to a G-repeat in the *tlpB* 5' UTR and affects translation of *tlpB* mRNA. This thesis represents the first functional characterization of a *trans*-acting sRNA in *Helicobacter*. *In-vitro* translation assays indicated that RepG mainly influences *tlpB* expression at the level of translation (Figure 2.13). However, since the *tlpB* G-repeat, which represents the RepG interaction site, is located far upstream of the RBS (Figure 2.8), repression of *tlpB* expression might be based on structural changes or binding to an upstream regulatory element rather than a direct masking of the RBS. Indeed, sRNAs have been shown to bind far upstream of the RBS and affect gene expression through sequestering of translational enhancer elements or ribosome stand-by sites (Sharma *et al.*, 2007, Darfeuille *et al.*, 2007).

The thermodynamic stability of mRNA secondary structures near the RBS and start codon has been shown to affect mRNA translation, with translation efficiency being inversely correlated to secondary structure stability (Gu *et al.*, 2010). In line with this, sRNA-mediated structural changes within the target mRNA can lead to inhibition of translation, transcript destabilization, transcription attenuation or termination (Waters & Storz, 2009, Lalaouna *et al.*, 2013). This might also apply for RepG-mediated regulation of *tlpB* expression, as structural rearrangements within the *tlpB* mRNA leader, especially in the hairpin structure upstream of the RBS (position -23 to -52 relative to the annotated AUG), were observed upon the RepG-*tlpB* mRNA interaction (Figure 2.7 B). Although *in-vitro* toeprinting assays indicated that RepG is rather unlikely to interfere with translation initiation (Figure 2.12), RepG-mediated structural rearrangements might affect efficient translation elongation of the *tlpB* mRNA. Interestingly, in comparison to the wild-type *tlpB* 5' UTR of *H. pylori* 26695 (12G), RepG interaction with *tlpB* leader variants that contain a G-repeat of 13 or 14Gs resulted in less pronounced structural changes within the *tlpB* mRNA (compare Figure 2.7 B and Appendix, Figure 13.1). This suggests that structural changes within the *tlpB* mRNA might be linked to the outcome of RepG-mediated *tlpB* regulation.

The homopolymeric G-repeat of *tlpB* folds into an inter- or intramolecular structure. RepG represents the first example of a *trans*-acting sRNA that interacts with a homopolymeric G-repeat in its target mRNA. *In-vitro* structure probing indicated that the *tlpB* G-repeat might form an inter- or intramolecular structure (Figure 2.7). In addition, *in-*

vivo analyses of triple (3xC) and single (1xC*) G to C exchanges in the G-stretch, which were used successfully in the compensatory base-pair exchanges for the gel-shift assays (Figure 2.5), revealed that these mutations affect *tlpB* expression *per se* (Figure 2.23 A). Furthermore, structure probing assays with *tlpB* mRNA leader variants 3xC and 1xC* showed that the introduced cytosine residue(s) is/are accessible to lead (II)-acetate cleavage (Figure 2.23 B), indicating that mutations within the G-repeat might interfere with the formation of an inter- or intramolecular structure that is required for efficient *tlpB* translation (*i.e.* the G-repeat can function as translational enhancer element).

Repetitive guanine-rich DNA and RNA sequences have the ability to form non-canonical, stable structures referred to as G-quadruplexes, which are composed of planar assembled guanine residues that pair via Hoogsteen hydrogen-bondings (Millevoi *et al.*, 2012, Bochman *et al.*, 2012). In eukaryotes, G-quadruplexes have been implicated in many biological processes, including DNA maintenance, telomere homeostasis, epigenetic regulation in recombination, mRNA splicing as well as gene expression regulation. For example, RNA G-quadruplexes in 5' UTRs of eukaryotic mRNAs have been demonstrated to mediate both repression and activation of gene expression by *e.g.* compromising the assembly of the translation machinery and through functioning as structural determinants of internal ribosome entry sites, respectively (Bugaut & Balasubramanian, 2012). Likewise, introduction of an artificial G-quadruplex forming sequence close to the RBS in bacterial mRNAs has been shown to affect gene expression in *E. coli* (Wieland & Hartig, 2009). In addition, a guanine-rich DNA sequence in the promoter of the pilin locus in *Neisseria gonorrhoeae* has been shown to form a G-quadruplex and to be required for antigenic variation (Cahoon & Seifert, 2009). Interestingly, transcription of a *cis*-encoded antisense RNA that originated within the G-rich sequence has been implicated to be crucial for the formation of this DNA G-quadruplex and antigenic variation (Cahoon & Seifert, 2013). More recently, eukaryotic RNA G-quadruplexes have been proposed to be involved in posttranscriptional regulation by non-coding RNAs. In particular, an *in-vitro* study of two G-rich sequences within the 3' UTR of the PSD-95 mRNA showed that G-quadruplexes are regulators of microRNA binding sites in humans (Stefanovic *et al.*, 2015). In line with this, targeting of RNA G-quadruplexes by antisense RNAs has been shown to inhibit or promote G-quadruplex folding and thus, enhance or repress translation of specific mRNAs *in vivo* (Rouleau *et al.*, 2015). Therefore, binding of RepG to the homopolymeric G-repeat might influence formation of the intra- or intermolecular structure and thus, activate or repress *tlpB* mRNA translation. Future studies will be required to determine the exact structure of the G-repeat in the *tlpB* 5' UTR, its interaction with RepG, and to resolve RepG-mediated structural rearrangements, which could be the underlying mechanism of the translational regulation of *tlpB*.

A

B
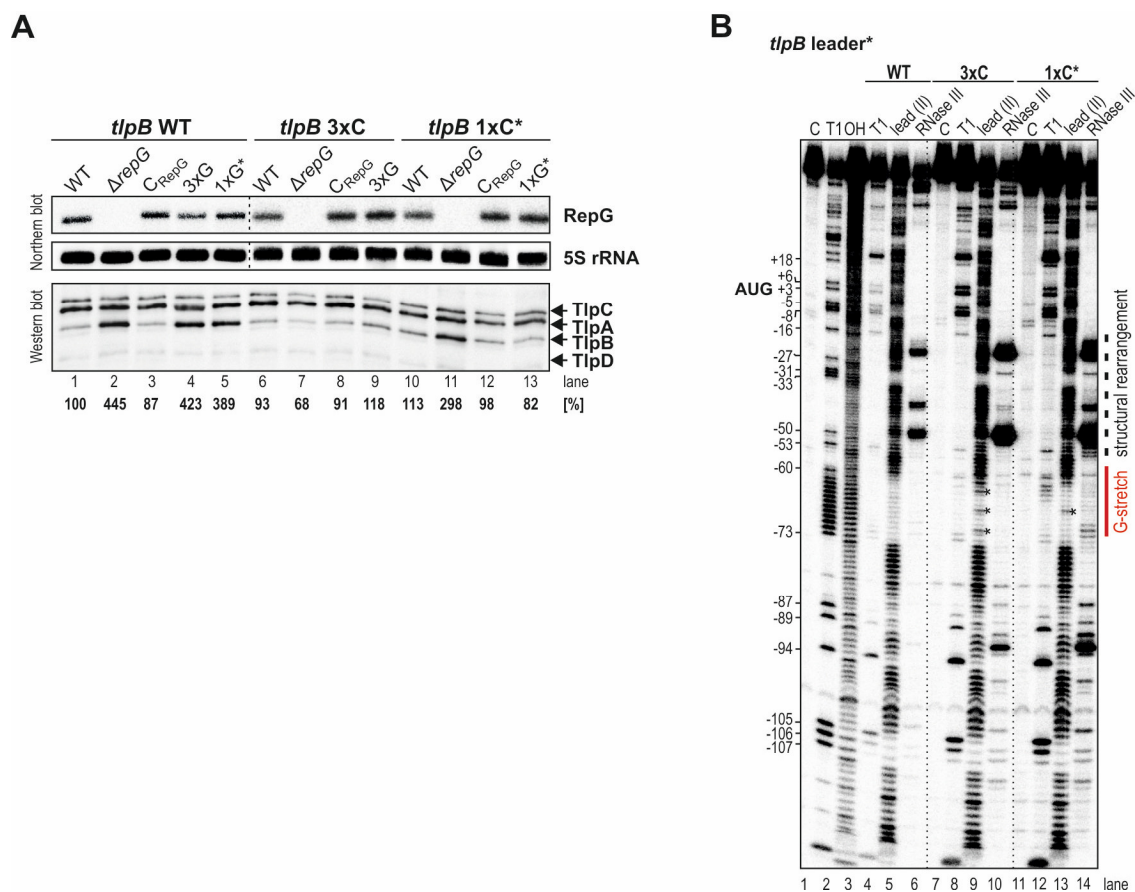
Figure 2.23: Mutations within the homopolymeric G-repeat affect *tlpB* mRNA translation *per se*. (A) *H. pylori* 26695 WT or Δ*repG* mutant containing the wild-type *tlpB* leader, *tlpB* 3xC or *tlpB* 1xC*, and the wild-type RepG (WT, C_{RepG}) or mutant sRNAs (3xG, 1xG*) were grown to exponential growth phase (OD_{600nm} of ∼ 0.7), and expression levels of RepG and TlpB were determined by northern and western blot, respectively. **(B)** *In-vitro* structure probing of wild-type *tlpB* mRNA leader, *tlpB* 3xC and *tlpB* 1xC*. About 0.1 pmol ^{32}P-labeled *tlpB* mRNA leader or its variants were treated with RNase T1, lead (II)-acetate and RNase III. While the guanines of the homopolymeric G-repeat are protected, triple or single G to C substitution(s) are/is accessible for T1 and lead (II)-acetate cleavages (indicated by asterisk). Untreated RNA (lane C), partially alkali- (lane OH) or RNase T1- (lane T1) digested wild-type *tlpB* mRNA served as ladders.

RepG-mediated *tlpB* mRNA decay. Following translational inhibition, interacting RNAs often become substrates for endoribonucleases such as RNase E and RNase III (Lalaouna *et al.*, 2013). Likewise, RepG-mediated repression of *tlpB* translation might be coupled to increased transcript degradation, leading to the observed reduction in *tlpB* mRNA stability in *H. pylori* strain 26695 in the presence of RepG (Figure 2.11 A). While RepG positively affects *tlpB* expression at the protein level in *H. pylori* strain G27 (Figures 2.8 B and C), it destabilizes the *tlpB* mRNA (Figure 2.11 B). This suggests that, in addition to the regulation at the translational level, antisense base-pairing of RepG might trigger *tlpB* mRNA degradation by active recruitment of RNases. Following this hypothesis, RepG-induced *tlpB*

mRNA decay might be independent from its translational control as indicated by the inverse outcome of RepG-mediated *tlpB* regulation at transcript and protein level in *H. pylori* G27. Target gene repression by active nucleolytic repression and/or induction of RNase-dependent cleavages have been described for several enterobacterial sRNAs, which either block translation initiation or bind their target mRNAs within the coding region (Prevost *et al.*, 2011, Pfeiffer *et al.*, 2009).

Preliminary data suggested that RNase J and ribosomal protein S1 potentially interact with the *tlpB* mRNA and/or RepG, thereby affecting mRNA and/or sRNA stability (Figures 2.21 and 2.22). Future studies will be required to determine the role and mechanism of RNase J and ribosomal protein S1 in posttranscriptional control of gene expression in *Helicobacter*, with particular focus on RepG-mediated *tlpB* regulation.

The RepG sRNA levels are altered in response to acid, iron and oxidative stress. To better understand the physiological role of RepG in *H. pylori*, it is important to identify the environmental signals and mechanisms that control its transcription and/or turnover. RepG expression and stability are regulated in response to nutrient availabilities and/or in a cell density-dependent manner (Figure 2.14). Since TlpB is assumed to be involved in sensing the signal molecule AI-2 (Rader *et al.*, 2011), there might be a potential link between the (growth phase-dependent) posttranscriptional regulation of *tlpB* by RepG and quorum-sensing in *H. pylori*. However, this awaits further examinations. Future studies will be required to unveil whether growth phase-dependent variations in RepG sRNA stability (Figure 2.14 B) are due to coupled sRNA-target mRNA degradation or changes in expression of protein factors such as the ribosomal protein S1.

Expression of RepG was shown to be altered in response to oxidative stress, acid and different iron-availabilities and (Figures 2.18, 2.19 and 2.20). In addition, preliminary data suggested that *repG* expression might be activated by the iron-homeostasis regulator Fur (Figure 2.20 B). Fur is a global regulator in *H. pylori* that has been shown to affect transcript levels of more than 200 gene loci, which are involved in iron homeostasis and the combat against acidic and oxidative stress (Danielli *et al.*, 2006). Therefore, alterations in RepG levels under each of the examined growth/stress conditions could be connected to its potential transcriptional regulation by Fur. A prominent example of a Fur-regulated sRNA in enterobacteria is RyhB, which either represses multiple mRNAs that encode nonessential iron-utilizing proteins or activates genes that are involved in siderophore production (Salvail & Masse, 2012). Interestingly, RyhB regulation is not limited to iron utilization and homeostasis in *E. coli*, but has also been shown to contribute to the response to oxidative stresses in *Salmonella* (Calderon *et al.*, 2014), and control motility, chemotaxis and biofilm formation in *V. cholera* (Mey *et al.*, 2005). Although the *ryhB* gene is conserved in several

pathogenic bacteria, no homolog could be identified in *Helicobacter* so far. It remains to be elucidated whether or not RepG might represent a functional homolog of RyhB in *H. pylori*.

Given that sRNAs are frequently encoded adjacent to their transcriptional regulators (*e.g.* Vanderpool & Gottesman, 2004, Urbanowski *et al.*, 2000), the orphan response regulator HP1043 might control *repG* expression over growth and/or in response to oxidative stress. An activating effect of this response regulator on *tlpB* expression has been suggested based on an *in-vitro* promoter binding study and correlated expression profiles of HP1043 and *tlpB* over growth and under oxidative stress (Delany *et al.*, 2002, Olekhnovich *et al.*, 2014). This might suggest a potential feed-forward loop of *tlpB* regulation involving HP1043 and RepG.

In *H. pylori* strain 26695, RepG represses *tlpB* at both the transcript and protein level. Based on this, an anti-correlation between the sRNA and the target mRNA or protein levels would be expected. Surprisingly, however, RepG and *tlpB* mRNA levels correlate over growth (Figure 2.14, Delany *et al.*, 2002). Moreover, expression of both transcripts is co-regulated with an induction under iron-sufficiency in exponential phase (Figure 2.20, Ernst *et al.*, 2005) and a repression under high oxygen tensions (Figure 2.18 B, Park & Lee, 2013). Taking into account that RepG is highly abundant and its levels were typically not limiting for *tlpB* regulation under examined conditions, RepG might represent a housekeeping sRNA that prevents target mRNA overexpression or counteracts transcriptional responses at the posttranscriptional level. In addition, the correlation of the RepG and *tlpB* expression profiles might suggest a common transcriptional regulator or physiological function for both transcripts under similar growth/stress conditions. Further studies will be required to understand if/to what extent acidic, iron and oxidative stresses influence RepG-mediated TlpB regulation.

A possible role for phase variation of the *Helicobacter* chemotaxis receptor TlpB in virulence. Motility and chemotaxis are essential for virulence and efficient colonization of the host by *Helicobacter* (Spohn & Scarlato, 2001). *H. pylori* strain 26695 carries four methyl-accepting chemotaxis receptors, TlpA, TlpB, TlpC, and TlpD. The here presented work showed that RepG specifically regulates *tlpB* expression, whereas the other chemotaxis receptors are not affected (Figure 2.3 B). TlpB has been shown to sense quorum sensing molecules and pH, whereby both the signal molecule AI-2 and acid act as chemo-repellants (Croxen *et al.*, 2006, Rader *et al.*, 2011). Moreover, TlpB has been implicated in colonization and inflammation during mice and gerbil infections (Croxen *et al.*, 2006, McGee *et al.*, 2005, Williams *et al.*, 2007). The length of the G-repeat in the *tlpB* 5' UTR determines the outcome of sRNA-mediated *tlpB* regulation in different *H. pylori* strains (Figure 2.8). Analysis of *tlpB* sequences of sequential *H. pylori* isolates from human patients (Devi *et al.*, 2010, Avasthi *et al.*, 2011) and from strains re-isolated from animal colonization

experiments (Behrens *et al.*, 2013, Salaun *et al.*, 2005) indicated that the G-repeat not only varies between strains from different patients but also between isolates from the same host, suggesting that *tlpB* could undergo phase variation during infection (Appendix, Table 13.1). Phase-variable gene expression in *H. pylori* has been shown to be associated with genes involved in surface structures and thus, in host recognition, adhesion (Appelmelk *et al.*, 1999, Solnick *et al.*, 2004, Yamaoka *et al.*, 2002, Kennemann *et al.*, 2012), motility (Josenhans *et al.*, 2000), or in DNA restriction and modification (de Vries *et al.*, 2002, Srikhanta *et al.*, 2010). How differential *tlpB* expression is connected to host adaptation remains to be shown.

As the pH gradient is the major chemotactic signal in the mucosa of the stomach (Bucker *et al.*, 2012), RepG-mediated *tlpB* regulation might be important for the ability of *H. pylori* to sense the surrounding environment and thus, for a successful colonization of suitable niches/microenvironments within the stomach. Preferential colonization of distinct areas of the stomach has been reported for different *H. pylori* strains in animal models (Akada *et al.*, 2003). Also, chemotaxis and motility are assumed to promote *H. pylori* colonization of injury sites and thereby, biases injured tissues towards sustained gastric damage (Aihara *et al.*, 2014). Acid not only acts as a chemo-repellent, but also leads to profound changes in *H. pylori* motility, as acid-exposed bacteria display significantly higher speed of active movements (Merrell *et al.*, 2003a). The role of the chemotaxis receptor TlpB for motility of *H. pylori* is controversially discussed in the literature (Croxen *et al.*, 2006, McGee *et al.*, 2005). In our hands, deletion of *tlpB* in motile *H. pylori* strains India7 and G27 resulted in slightly reduced motility (Figure 2.24 A). According to the G-repeat length in the *tlpB* mRNA leader, RepG represses and activates *tlpB* expression in India7 (12G) and G27 (14G), respectively (Figure 2.8). Preliminary data suggested that neither RepG nor RepG-mediated control of *tlpB* affects *H. pylori* motility in these strain backgrounds (Figure 2.24 A). Nevertheless, further studies will be required to investigate the effect of RepG-mediated *tlpB* expression control on the chemotactic behavior of *H. pylori*, *e.g.* in response to low pH or AI-2.

H. pylori mutants deficient for *tlpB* are impaired in colonization in the mouse model (Croxen et al., 2006), suggesting that this acid-sensing chemotaxis receptor might be involved in *H. pylori* acid acclimatization. Preliminary data showed that deletion of *tlpB* did not affect acid-sensitivity of *H. pylori* strain 26695 (Figure 2.24 B). In contrast, mutants lacking *repG* displayed impaired long-term survival rates under low pH conditions, suggesting that RepG, independent from its regulation of *tlpB,* is required for the adaptation to acidic stress. This might be linked to RepG-mediated expression control of additional, so far unknown, target genes.
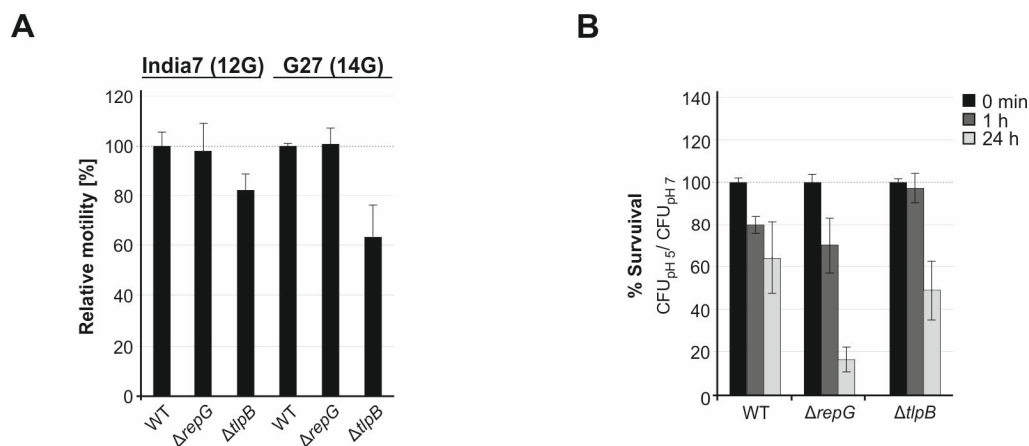
Figure 2.24: Deletion of *repG* does not affect *H. pylori* motility, but impairs long-term survival under acidic stress. (A) Motility of WT, Δ*repG* and Δ*tlpB* mutants of *H. pylori* strains India7 and G27 was assayed on 0.4 % semi-solid Brucella broth agar after 5 days of incubation. The swimming/swarming diameter of the wildtype was used as reference and set to 100 %. **(B)** *H. pylori* strain 26695 WT, Δ*repG* and Δ*tlpB* were grown in to early exponential growth phase (OD_{600nm} of ~ 0.5) and split into two subcultures. While one subculture was left untreated (pH ~ 7.0), the media of the other one was adjusted to pH 5.0 with 37 % HCl. Colony forming units (CFU) per ml for indicated time points were determined by serial dilutions. The survival (%) of each strain was assayed by comparison of CFU/ml obtained under pH 5 to those of the untreated culture (for each time point). Error bars indicate standard deviations of two biological replicates.

RepG is one of the most conserved sRNAs in *Helicobacter*, particularly its C/U-rich terminator loop, indicating that RepG might use this loop region to interact with other mRNAs. As will be discussed in Chapter 4, whole transcriptome analyses indeed indicate that multiple genes are affected upon *repG* deletion. Consequently, a mutation in the highly conserved C/U-rich region would abolish the global regulatory function of RepG. In contrast, variation of the *tlpB* G-repeat length and its influence on RepG regulation facilitates uncoupling of a single target from a sRNA regulon through modification of its targeting site.

Mechanisms and importance of intergenic SSRs in 5' UTRs. This study showed that the length of a G-repeat in an mRNA leader affects expression of a chemotaxis receptor through posttranscriptional regulation by a sRNA. The only other example of a 5' UTR-associated G-repeat so far has been described for the UspA1 adhesin in *Moraxella catarrhali* and has been shown to influence *uspA1* mRNA levels (Lafontaine *et al.*, 2001). In addition, the length of a heteropolymeric tetranucleotide repeat in the leader of *uspA2* mRNA was shown to affect mRNA stability and protein level of this adhesin and thereby, contributed to serum resistance in *M. catarrhali* (Attia & Hansen, 2006). However, in both cases the underlying mechanism remained unclear, but it is possible that also these SSRs might be targeted by sRNAs. Length variation of simple A- or T-repeats, and dinucleotide (AT)-repeats embedded in the spacer region of the RBS and AUG start codon has been successfully used to fine-tune

gene expression in *E. coli* within artificial regulatory networks (Egbert & Klavins, 2012), highlighting the regulatory potential of SSRs in untranslated regions of mRNAs. Moreover, a current study demonstrated that phase-variable SSRs within leader peptides can also control expression of downstream encoded genes (Danne *et al.*, 2014). In particular, a variable GCAGA-repeat in a leader peptide of the *pil1* mRNA, encoding for pili structures that are important for the host colonization by *Streptococcus gallolyticus*, was shown to regulate transcription of *pil1* through ribosome-induced destabilization of a premature transcription terminator. Apart from SSRs, a phase-variable invertible element in the *cwpV* mRNA leader of *Clostridium difficile* has been shown to determine transcription elongation through formation of an intrinsic transcription terminator depending on the orientation of the DNA element (Emerson *et al.*, 2009).

In addition to length variation of SSRs by slipped-strand mispairing during DNA replication, transcription slippage can also cause the incorporation or deletion of nucleotides in SSRs of mRNAs (Wagner *et al.*, 1990, Baranov *et al.*, 2005, Gueguen *et al.*, 2014). Albeit mainly reported for long poly-A and poly-T stretches, it might be possible that the for- and back-slippage of the RNA-polymerase at the homopolymeric G-repeat in the *tlpB* mRNA leader results in a heterogeneous population of mRNAs without mutating the *H. pylori* chromosome. Further studies will be required to investigate whether or not transcriptional slippage might also contribute to RepG-mediated *tlpB* regulation in different *H. pylori* strains.

Besides in *Helicobacter,* length variations of poly-G tracts have also been observed under selective environmental conditions and passage through animals in other Epsilonproteobacteria such as *C. jejuni* (Jerome *et al.*, 2011, Bayliss *et al.*, 2012). These G-repeats could also be potential target sites of sRNAs, which have recently been identified in this pathogen (Dugar *et al.*, 2013). Comparisons of homopolymeric SSR locations with the global transcriptional start site maps of *H. pylori* and *C. jejuni* (Sharma *et al.*, 2010, Dugar *et al.*, 2013) showed that the majority of SSRs are found in promoter or coding regions, but revealed about 11 genes that carry a SSR in their 5' UTR and thus, might act by influencing posttranscriptional regulation (Appendix, Table 13.2). Besides base-pairing with translation initiation regions, bacterial sRNAs can also regulate gene expression by targeting coding sequences (Pfeiffer *et al.*, 2009). Therefore, several of the intragenic SSRs could also be targeted by *trans*-encoded sRNAs. Moreover, the previous transcriptome study also identified several *cis*-encoded antisense RNAs to SSRs in *H. pylori* strain 26695 (Sharma *et al.*, 2010). Overall, this new mode of gene regulation through homopolymeric repeats is likely to be more widespread and SSRs not only in 5' UTRs, but also within the coding sequence could be targeting sites of sRNAs

3. Small RNA-mediated gradual gene expression control of a novel LPS biosynthesis and colonization factor in *Helicobacter pylori*

The conserved small RNA RepG represents the first *trans*-acting sRNA described in bacteria that targets an SSR, namely a homopolymeric G-repeat within the leader of *tlpB* chemotaxis receptor mRNA. In *H. pylori*, phase-variable SSRs are genrally associated with genes encoding products that influence the interaction with the host, *e.g.* bacterial surface and/or motility-associated structures. This raises the question, why *tlpB* expression is coupled to the phase-variable G-repeat.

In almost all sequenced *H. pylori* strains, *tlpB* is encoded in a two-gene operon upstream of HP0102, encoding a gene of unknown function. Previous data indicated that RepG co-regulates *tlpB* and HP0102 on the transcript level in *H. pylori* strain 26695 (Sharma *et al.*, 2010). Therefore, the G-repeat in the *tlpB* 5' UTR might be related to the function of HP0102. Work presented in this Chapter is focused on the functional characterization of the hypothetical protein HP0102 and provides insights into the underlying mechanism of RepG-mediated co-regulation of both *tlpB* and HP0102. In addition, the role of RepG-mediated target gene (*tlpB* and HP0102) regulation in *H. pylori* virulence and survival-associated phenotypes is examined to gain further insights into the physiological function of RepG and its target genes.

3.1. Expression of the conserved *tlpB*-HP0102 operon is regulated by RepG

Using BLAST (Basic Local Alignment Search Tool, NCBI) and SyntTax webserver (Oberto, 2013), biocomputational searches for *tlpB* and HP0102 homologs were conducted in various *Helicobacter* species (spp.) and other Epsilonproteobacteria, including *Campylobacter* spp., *Wolinella* spp., *Sulfuricurvum* spp., *Arcobacter* spp., *Sulfurospirillum* spp., and *Nautilia profundicola* (approximately 120 genomes). In almost all sequenced *Helicobacter* spp.*,* the chemotaxis receptor and HP0102 are encoded in an operon (Figure 3.1 A). Exceptions are *H. pylori* strain Puno120 – in whose genome two genes belonging to a type-II DNA restriction and modification system are inserted downstream of *tlpB* –, and non-*H. pylori* strains such as *H. felis*, *H. bizzozeronii*, and *H. mustelae*. In *H. mustelae*, for instance, two pseudogenes with limited sequence homology to HP0102 are encoded more than 150 kbp apart from the *tlpB* homolog. Despite being highly conserved in the genus of *Helicobacter*, no homologs of *tlpB* and HP0102 (homology $\geq$ 30 %) could be identified in the other Epsilonproteobacteria. Also, *tlpB* and HP0102 are missing in various non-*Helicobacter pylori* strains such as *H. pullorum*, *H. bilis*, *H. cinaedi* and *H. hepaticus*. RepG is highly conserved in different *H. pylori* strains (Figure 2.1 B). Apart from *H. acinonychis*, *H. cetorum* and
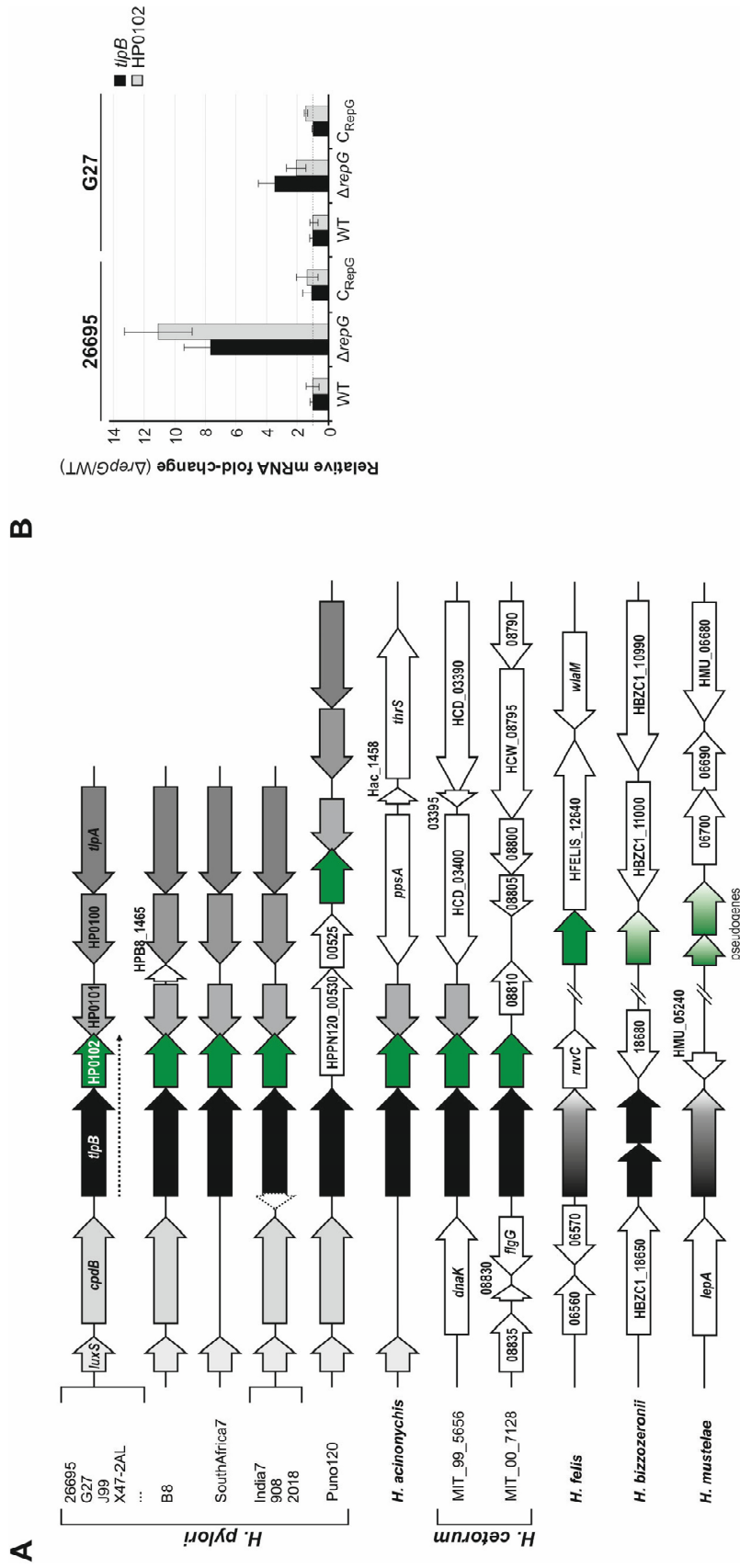
Figure 3.1: Genomic context and RepG-mediated regulation of the dicistronic *tlpB*-HP0102 mRNA. (A) Gene synteny analysis of the *tlpB*-HP0102 operon (dotted line) encoding a chemotaxis receptor and a protein of unknown function. Homologs in different *Helicobacter* spp. are illustrated by the same colors, whereas unrelated genes are indicated in white. A potential open reading frame (dotted, white arrow) is anotated in the *tlpB* 5′ UTR in *H. pylori* strains India7, 908 and 2018. Gradually colored arrows indicate a limited sequence homology, whereas fully colored arrows represent homologs with more than 50 % sequence conservation. Different *H. pylori* strains are grouped according to their genomic context. (B) Quantitative RT-PCR analysis of RepG-dependent expression of *tlpB* and HP0102 mRNAs in *H. pylori* strains 26695 and G27. After DNase I treatment, RNA samples isolated from wildtype (WT), *repG* deletion (Δ*repG*) and complementation (C$_{RepG}$) mutants were used in one-step qRT-PCR reactions with specific primer pairs for *tlpB* and HP0102. Wild-type mRNA levels of *tlpB* and HP0102 were set to 1, and relative fold-changes in mutant strains are shown as bars. Values are shown as mean ± standard derivations from at least three independent experiments.

H. mustelae, no RepG sRNA homologs (homology $\geq$ 30 %) could be identified in other non-*H. pylori* spp. or closely-related Espilonproteobacteria by a BLAST search.

To confirm that expression of the hypothetical protein HP0102 is co-regulated with *tlpB* by RepG, transcript levels of the dicistronic *tlpB*-HP0102 mRNA were determined in *H. pylori* 26695 wildtype, *repG* deletion ($\Delta repG$) and complementation (C_{RepG}) strains at exponential growth phase (Figure 3.1 B). Quantitative RT-PCR revealed about eight- to ten-fold increased transcript levels for both *tlpB* and HP0102 upon *repG* deletion. The similar deregulation of *tlpB* and HP0102 in $\Delta repG$ is consistent with both genes being encoded in an operon. Furthermore, complementation of $\Delta repG$ restored wild-type levels of *tlpB* and HP0102, indicating that the entire *tlpB*-HP0102 operon is repressed by RepG in *H. pylori* strain 26695. The effect of *repG* deletion on *tlpB*-HP0102 mRNA levels was also tested in *H. pylori* strain G27. While RepG positively affects *tlpB* expression at the protein level (Figures 2.8 B and C), slightly (about two- to four-fold) increased *tlpB*-HP0102 transcript levels were detected upon *repG* deletion in this strain background (Figure 3.1 B). This fully agrees with the observed RepG-mediated *tlpB* mRNA destabilization in *H. pylori* strains 26695 and G27 (Figure 2.11), and suggests that base-pairing of RepG to the G-repeat in the *tlpB* mRNA leader negatively affects the stability of the entire *tlpB*-HP0102 mRNA. Overall, these data show that RepG co-regulates the expression of the chemotaxis receptor TlpB and hypothetical protein HP0102 in different *H. pylori* strains at the transcript level.

A putative ribosome binding site was identified in the 30-nt long intergenic region (IGR) between *tlpB* and HP0102 (Figure 3.2 A), suggesting that both genes are co-transcribed, but not necessarily co-translated in *H. pylori*. In addition, a G-rich sequence (G^{HP0102}, A$\underline{GGGGG}$AG in *H. pylori* 26695) that might serve as a putative RepG interaction site was identified upstream of the potential HP0102 RBS in the *tlpB*-HP0102 IGR (Figure 3.2 B, prediction of RepG-G^{HP0102} interaction is based on *H. pylori* 26695). In contrast to the high conservation of the *tlpB* stop codon, putative RBS and annotated start codon of HP0102, the remainder of the IGR varies among different *H. pylori* strains (Figure 3.2 A). In general, the size of the G^{HP0102}-repeat varies from three to six guanines. However, taking various *H. pylori* strains into account, no striking correlation between the lengths of the G-repeat in the *tlpB* mRNA leader and the G^{HP0102}-repeat in the *tlpB*-HP0102 IGR was found. Whether or not the G-repeat in the *tlpB*-HP0102 IGR affects RepG-mediated regulation of the chemotaxis receptor and/or HP0102 has been investigated in section 3.5.

3.2. HP0102 is required for *H. pylori* colonization of the murine stomach

To identify a potential function of HP0102, database searches for functional protein association networks of HP0102 were performed using STRING (Search Tool for the
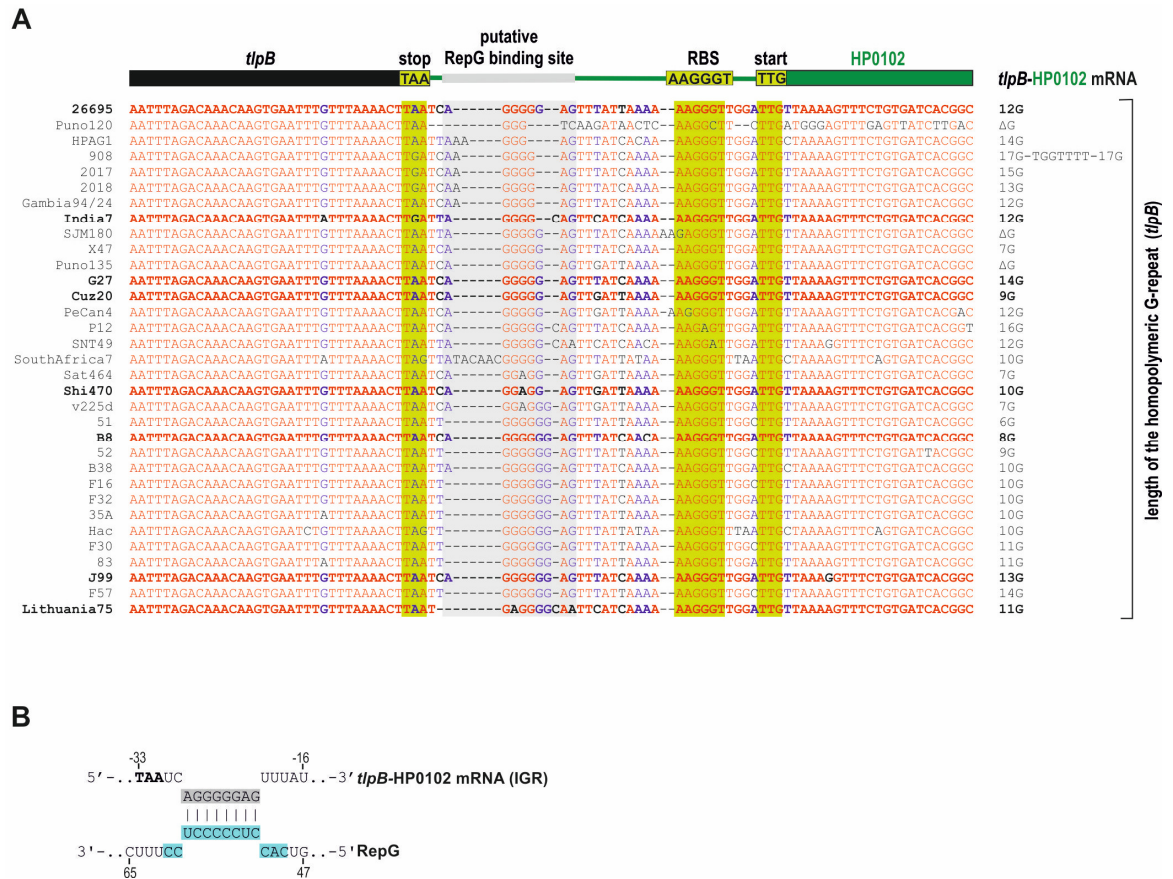
Figure 3.2: Sequence alignment of the intergenic region between *tlpB* and HP0102 in diverse *H. pylori* strains and *H. acinonychis* (Hac). The *tlpB* stop codons (TAA, TGA or TAG) as well as putative RBSs and annotated start codons (TTG) of HP0102 are highlighted in light green. The putative RepG interaction site (G^{HP0102}) is boxed in gray. The length of the homopolymeric G-repeat in the *tlpB* 5' UTR of each strain is indicated on the right. RepG-mediated *tlpB* regulation was previously investigated in strains shown in bold (Figure 2.8). **(B)** In *H. pylori* strain 26695, RepG was predicted to base-pair to G^{HP0102} (gray) in the *tlpB*-HP0102 IGR. The *tlpB* stop codon is shown in bold letters. Numbers indicate positions relative to the annotated HP0102 start codon. The *tlpB* interaction site in the terminator loop of RepG is shown in blue.

Retrieval of Interacting Genes/Proteins, Franceschini et al., 2013), MINT (Molecular INTeraction database, Chatr-aryamontri *et al.*, 2007), DIP (Database of Interacting Proteins, Xenarios *et al.*, 2002), and APID (Agile Protein Interaction DataAnalizer, Prieto & De Las Rivas, 2006) databases. A search for conserved interactions between pairs of proteins by interolog mapping (Yu *et al.*, 2004) indicated that HP0102 might interact with proteins that are involved in lipopolysaccharide and/or cell wall biosynthesis. In addition, using the Pfam database (Protein families database, Finn *et al.*, 2014), a glycosyltransferase family 2 (GT-2)-domain was predicted for HP0102. Glycosyltransferases catalyze glycosidic bound formation using sugar donors containing either a nucleoside or lipid phosphate group. They have been shown to act on a wide range of substrates, including cell wall components, LPS

structures or intermediates, proteins and even nucleic acids. In particular, the GT-2 domain is found in glycosyltransferases that transfer sugar units from UDP-glucose, UDP-N-acetylglucosamine or GDP-mannose to a range of substrates including cellulose, dolichol phosphate and teichoic acid. Accordingly, Logan and colleagues proposed that HP0102 might encode for one glycosyltransferase of the LPS assembly pathway of *H. pylori* strain 26695 (Logan *et al.*, 2005).

Lipopolysaccharides are important pathogenicity and virulence factors of *H. pylori* and contribute to the severity and chronicity of the infection (Moran, 2008). Likewise, motility and chemotaxis have been shown to facilitate effective colonization of the stomach by *H. pylori* (Spohn & Scarlato, 2001). In agreement with this, a large-scale, transposon mutagenesis-based screen for novel virulence-associated factors identified *tlpB* and HP0102 as candidate genes that are important for the colonization and/or persistence of *H. pylori* in mouse infection experiments (Baldwin *et al.*, 2007). Thus, sRNA-mediated co-regulation of the dicistronic *tlpB*-HP0102 mRNA (Figure 3.1 B) suggests a potential involvement of RepG in *H. pylori* pathogenicity.

To assess whether RepG and/or its targets, the chemotaxis receptor *tlpB* and the putative glycosyltransferase HP0102, contribute to *H. pylori* virulence or survival in a mouse model, *in-vivo* infection experiments with *H. pylori* strain X47-2AL were performed in collaboration with PhD Hilde De Reuse (Institut Pasteur, Paris, France). In this mouse-adapted *H. pylori* strain, the syntenic organization of the *tlpB*-HP0102 operon and its surrounding regions are reminiscent to that of strains 26695, J99, and G27 (Figure 3.1 A). The RepG sRNA, *tlpB*, HP0102 or the *tlpB*-HP0102 operon were deleted in *H. pylori* strain X47-2AL (Figure 3.3 A). To avoid potential polar effects of Δ*tlpB* on the expression of HP0102, the *tlpB* coding region was replaced by a non-polar gentamicin resistance cassette (*aac*(3)-IV), leaving the *tlpB* promoter and 5' UTR intact. An analogous cloning strategy was used for the construction of the *tlpB*-HP0102 double deletion mutant (Δ*tlpB*-HP0102) and deletion of HP0102 alone (ΔHP0102). Please note that in these mutants, the *aac*(3)-IV cassette was inserted 77 nt upstream of the HP0102 stop codon to avoid interference with HP0101 expression (HP0102 and HP0101 overlap in the chromosome, Figure 3.3 A). RepG was deleted in the Δ*tlpB*, ΔHP0102 and Δ*tlpB*-HP0102 mutants and complemented by introduction of the *H. pylori* X47-2AL *repG* gene under control of its native promoter at the unrelated *rdxA* locus. Likewise, Δ*tlpB*-HP0102 and ΔHP0102 mutants were complemented with either the entire *tlpB*-HP0102 operon (*tlpB*-HP0102) or the glycosyltransferase gene HP0102 alone in the *rdxA* locus (Figure 3.3 B). In case of the latter, the HP0102 coding sequence and 30-nt long *tlpB*-HP0102 IGR were fused to the *tlpB* promoter.

The G-repeat in the *tlpB* mRNA leader is composed of seven guanines in *H. pylori* strain X47-2AL (Figure 2.8 A). In line with the observation that G-stretch lengths of 7 to
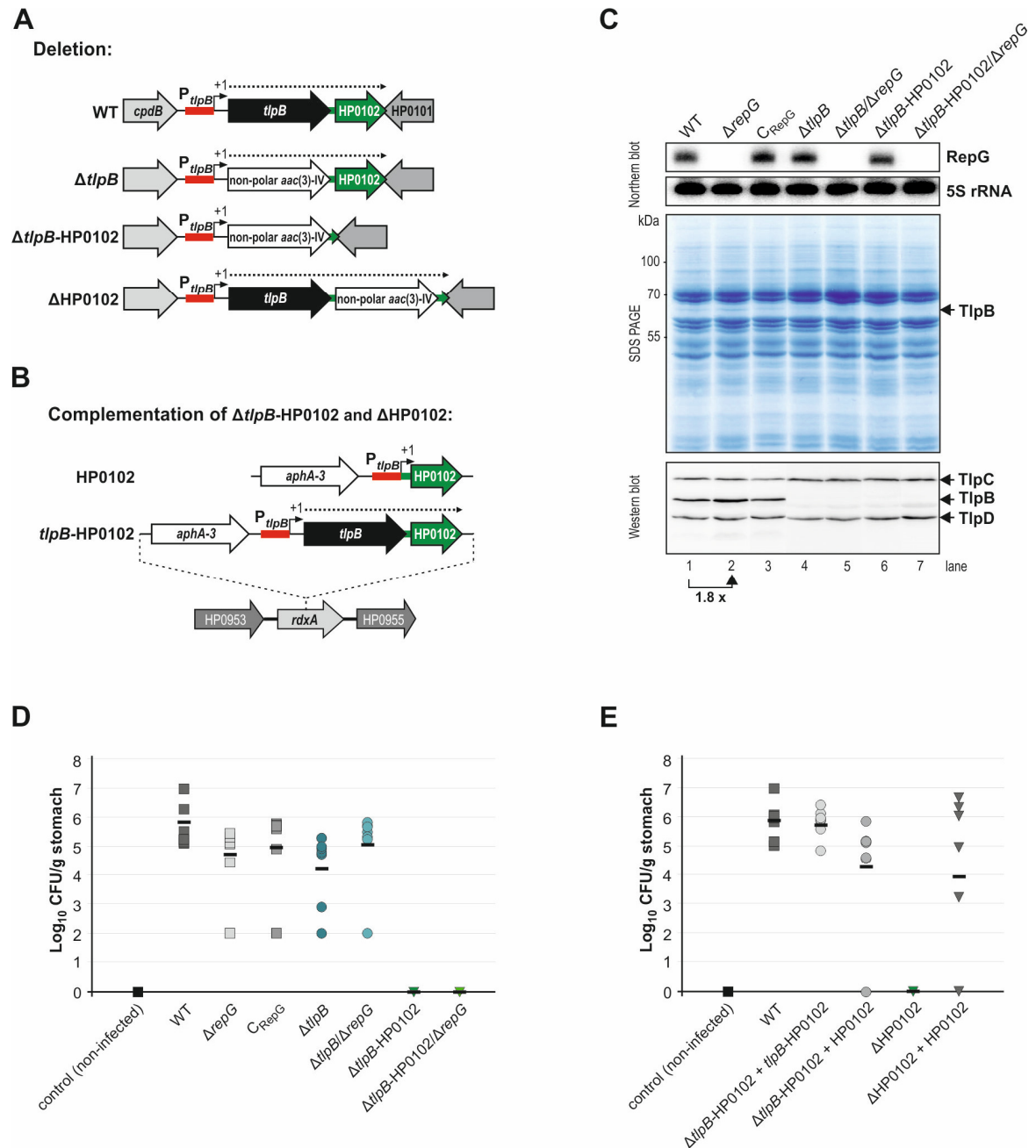
Figure 3.3: HP0102 is required for colonization of the stomach in mouse infection studies with _H. pylori_ strain X47-2AL. (A-B) Schematic representation of the cloning strategy used for the construction of _H. pylori_ X47-2AL mutants, which have been used in mouse infection studies. The non-polar gentamicin cassette (_aac_(3)-IV) was used to avoid polar effects of mutant construction. The TSS of the _tlpB_-HP0102 operon is denoted as +1, the _tlpB_ promoter (P_{tlpB}) is shown in red. **(C)** _H. pylori_ X47-2AL wildtype and mutants were grown to exponential growth phase, and RNA as well as protein samples were analyzed by northern blot, and SDS-PAGE or western blot, respectively. RepG was detected with CSO-0003 and the chemotaxis receptors were detected by a polyclonal rabbit anti-TlpA22 antiserum. **(D-E)** About 10^7 bacteria of _H. pylori_ X47-2AL wildtype or mutants were orogastrically administrated to NMRI Swiss mice. As a control, mice were infected with peptone broth only. Four weeks post infection, mice were sacrificed and colony forming units (CFU) per gram of stomach weight were calculated by serial dilutions and plating assays. Each symbol indicates the _H. pylori_ X47-2AL colonization titer in the stomach of a single mouse. The horizontal bars represent the geometric mean for each group of data.

12Gs mediate RepG-dependent *tlpB* repression (Figure 2.9), *repG* deletion in *H. pylori* strain X47-2AL resulted in about two-fold increased TlpB levels (Figure 3.3 C, lanes 1-2), whereas the levels of the other chemotaxis receptors remained unaltered. In addition, wild-type levels of the TlpB protein were detected in the RepG complementation strain (Figure 3.3 C, lane 3). Analysis of Δ*tlpB* and Δ*tlpB*-HP0102 as well as Δ*tlpB*/Δ*repG* and Δ*tlpB*-HP0102/Δ*repG* double deletion mutants confirmed the specificity of the antiserum against TlpB in *H. pylori* X47-2AL (Figure 3.3 C, lanes 4-7). All constructed *H. pylori* X47-2AL mutants displayed wild-type growth characteristics (data not shown).

In-vivo infection studies with the above-mentioned *H. pylori* X47-2AL mutant strains revealed that the *repG* and *tlpB* deletion mutants were slightly attenuated (1- to 1.5-log lower CFU/g) in their ability to colonize murine stomachs when compared to the wildtype Figure 3.3 D). This suggests that both *repG* and *tlpB* are dispensable for *H. pylori* colonization in mice. Accordingly, the Δ*tlpB*/Δ*repG* double deletion mutant displayed similar colonization loads as the two single deletion mutants (Δ*tlpB* or Δ*repG*). Compared to the wildtype, both Δ*repG* and C_{RepG} mutants showed slightly reduced colonization loads, suggesting that a fold-change of 1- to 1.5-log in CFU/g might represent intrinsic variations in mouse infection studies. In contrast to Δ*repG* and Δ*tlpB*, deletion of the *tlpB*-HP0102 operon completely abolished murine stomach colonization as no bacteria could be recovered from the stomach of mice that had been infected with the Δ*tlpB*-HP0102 and Δ*tlpB*-HP0102/Δ*repG* double deletion mutants (Figure 3.3 D). This indicates that HP0102, but not *tlpB*, is required for *H. pylori* strain X47-2AL colonization of mice. Mouse infection experiments with ΔHP0102 and its complementation (ΔHP0102 + HP0102) confirmed that HP0102 is responsible for the observed colonization defect (Figure 3.3 E). Complementation of the HP0102-deficient strains (Δ*tlpB*-HP0102 and ΔHP0102) with HP0102 alone only partially rescued the colonization defect observed for the respective deletion mutants (Figures 3.3 D and E). This might be due to altered HP0102 protein levels in the complementation mutants when compared to the wildtype, *e.g.* caused by the fusion of the IGR between *tlpB* and HP0102 to the *tlpB* promoter (Figure 3.3 B). Overall, these *in-vivo* infection studies identified the putative glycosyltransferase HP0102 as an important virulence factor for *H. pylori* colonization in mice.

3.3. The glycosyltransferase encoded by HP0102 is involved in biosynthesis of O-specific polysaccharide chains

The cell envelope of *H. pylori,* like that of other Gram-negative bacteria, contains lipopolysaccharides and is composed of three principal components: the glycolipid moiety
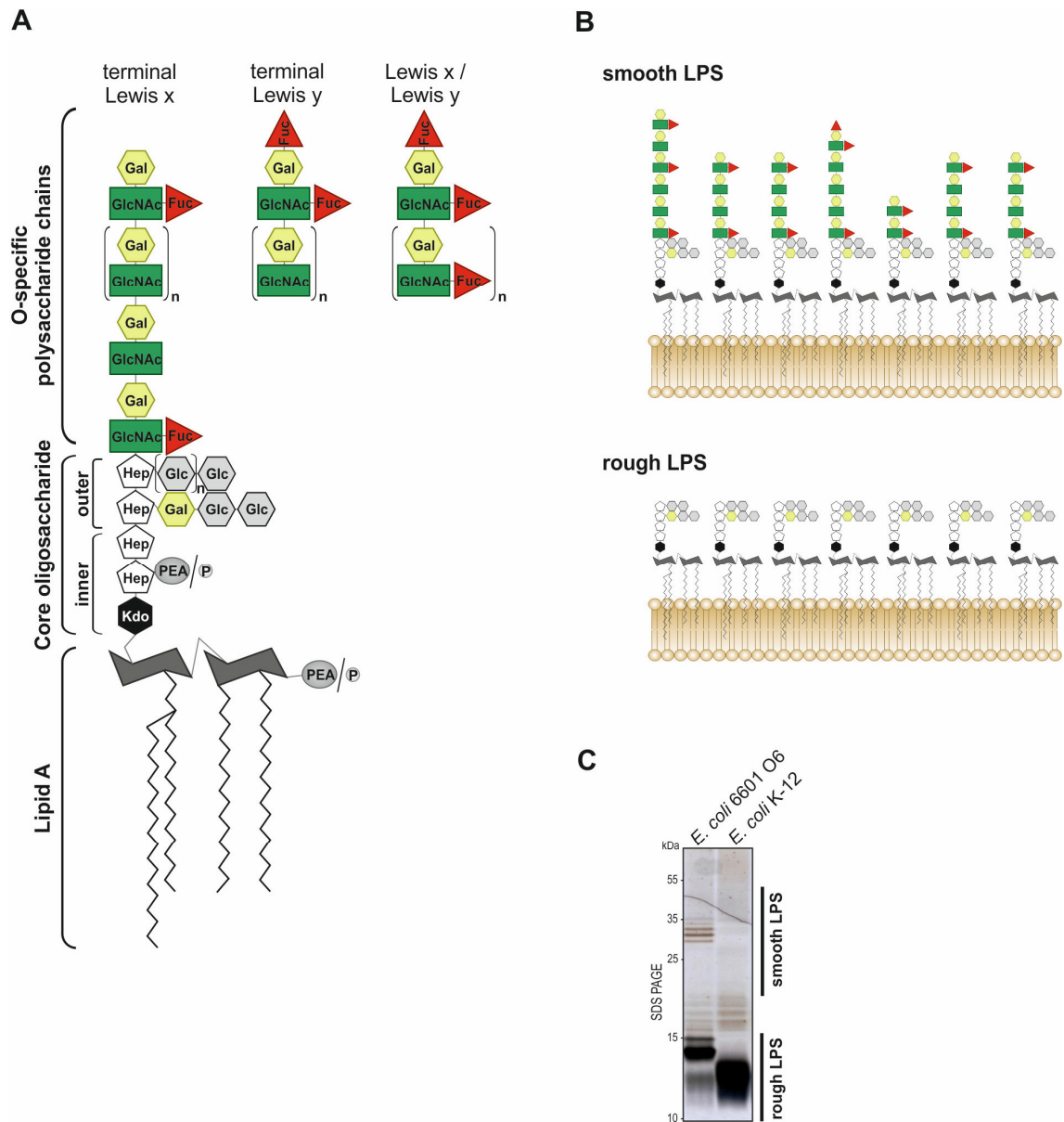
Figure 3.4: Schematic representation of the *H. pylori* lipopolysaccharide. (A) *H. pylori* LPS is composed of lipid A, the conserved core oligosaccharide, and variable O-specific polysaccharide chains (O-antigens). The number (n) of O-chain polysaccharide units may vary slightly within and between *H. pylori* cells. A typical O-antigen chain is glycosylated with multiple internal Lewis x units, and possesses either Lewis x or Lewis y at the terminal position. The Lipid A-core can be modified by phosphorylation (P) or addition of phosphoethanolamine (PEA). Kdo – 3-deoxy-D-manno-2-octulosonic acid, Hep – heptose, Gal – galactose, Glc – glucose, GlcNAc – N-acetylglucosamine, Fuc – fucose units. **(B)** *H. pylori* strains that possess smooth, high molecular weight LPS express all three LPS compartments, whereas rough, low molecular weight LPS bacteria do not produce O-antigen chains. **(C)** Silver-staining of LPS isolated from semi-rough (6601 O6) and rough-LPS producing (K-12) *E. coli* strains.

lipid A, the core oligosaccharide, and the O-specific polysaccharide chain (Figure 3.4 A). Each of these domains has different structural and functional properties (reviewed for

H. pylori in Moran, 2007). The innermost lipid A is a glucosamine-based glycolipid that anchors the LPS molecule in the outer membrane and endows LPS with a range of immunological and endotoxic activities. The core oligosaccharide is ketosidically linked to lipid A by an eight-carbon sugar, 3-deoxy-D-manno-2-octulosonic acid (Kdo), and was shown to be essential for the penetration/permeability properties of the outer membrane (Holst *et al.*, 1996). While the inner core in *H. pylori* is mainly composed of heptoses (including Kdo), a heteropolymer of neutral sugars, *e.g.* glucose and galactose, provides the structural basis for the outer core. The O-specific polysaccharide (O-chain) represents the outermost component of the LPS and contains an elongated, partially fucosylated polysaccharide backbone composed of alternating N-actetylglucosamine and galactose units in *H. pylori*. The O-chains contribute to antigenicity and serospecificity of native LPS and, - as the main surface antigens (O-antigens) - play an important role in *H. pylori* virulence (Moran, 2008). While the lipid A and core oligosaccharide are conserved, the composition of the O-chain varies among *H. pylori* strains. In most *H. pylori* strains, Lewis x and/or Lewis y antigens, which mimics the carbohydrate structures present on human epithelial and blood cells, are expressed in certain regions of LPS O-chains (reviewed in Appelmelk & Vandenbroucke-Grauls, 2001). Fresh clinical isolates of *H. pylori* produce high-molecular weight, smooth-form LPS, whereas *H. pylori* strains that have been extensively passaged on solid media produce low molecular-weight, rough-form LPS lacking O-polysaccharide chains (Figures 3.4 B and C, Moran *et al.*, 1992).

Since HP0102 is predicted to encode a glycosyltranferase, it might be involved in either the assembly of monosaccharide sugar units in the O-chain or addition of sugar modifications to the lipid A-core in *H. pylori*. To address whether HP0102 is a component of the LPS biosynthesis pathway in *H. pylori* strain X47-2AL, the LPS patterns of the wildtype and mutant strains, which had been used in the mouse infection experiments (Figure 3.3), were investigated by silver staining and western blot analysis with a Lewis x antigen-specific antibody. *H. pylori* X47-2AL wildtype displayed high-molecular ($\geq$ 15kDa), smooth-form LPS composed of the lipid A-core moiety and O-specific polysaccharide chains containing Lewis x antigens (Figure 3.5, lane 1). The $\Delta repG$, C_{RepG}, $\Delta tlpB$ and $\Delta tlpB/\Delta repG$ mutants showed similar LPS profiles to that of the wildtype with comparable expression levels of Lewis x antigens (Figure 3.5, lanes 1-5). In contrast, only low-molecular weight, rough-form LPS without Lewis x antigens was produced by the mutant strains devoid of HP0102 expression ($\Delta tlpB$-HP0102 and ΔHP0102) (Figure 3.5, lanes 6-7 and 10-11). Complementation of the $\Delta tlpB$-HP0102 and ΔHP0102 mutants with either the *tlpB*-HP0102 operon or HP0102 alone both restored synthesis of smooth LPS and Lewis x antigen profile similar to that of the wildtype (Figure 3.5, lanes 8-9 and 12). According to the LPS patterns of rough-LPS producing *E. coli* strains (Figure 3.4 C), these results demonstrate that the
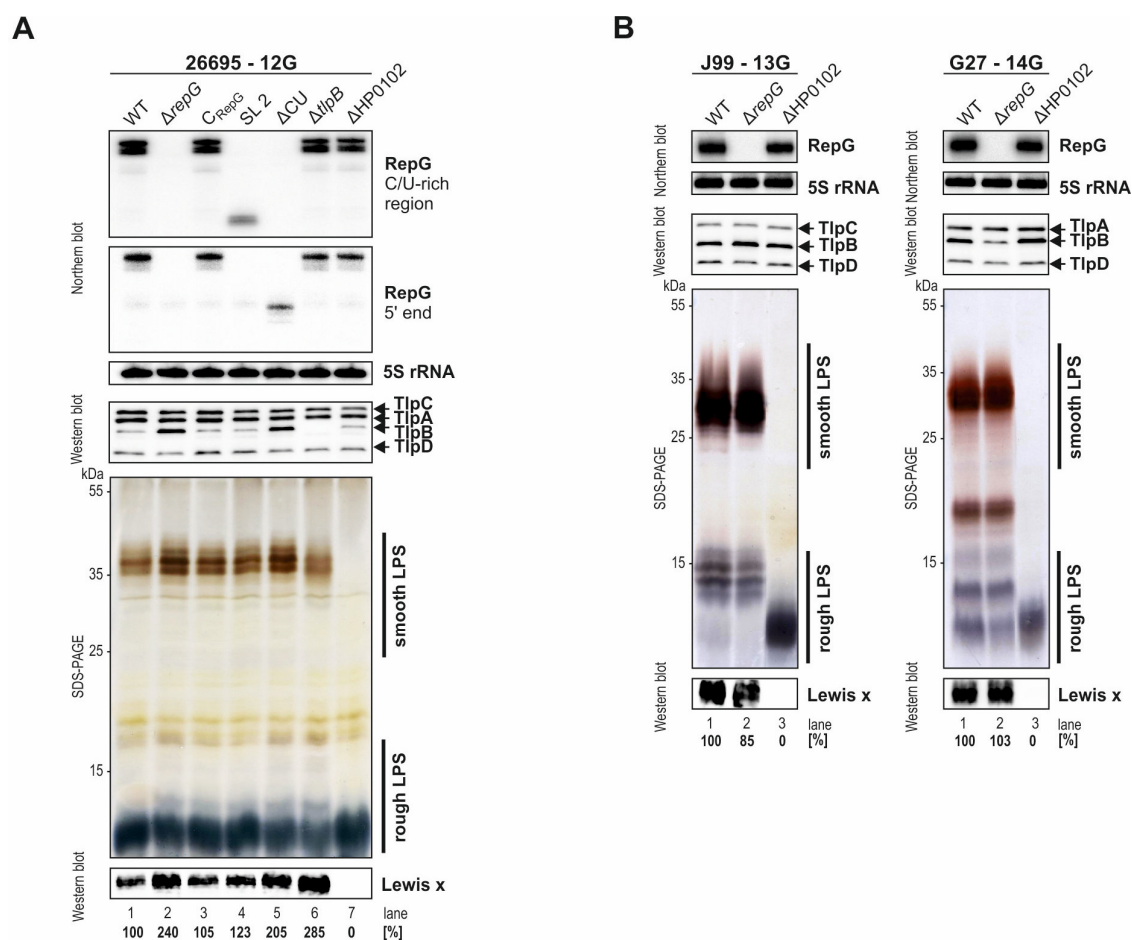
glycosyltransferase HP0102, but not *tlpB*, is an essential component in the LPS biosynthesis pathway in *H. pylori* strain X47-2AL and is responsible for the formation of the O-antigens.
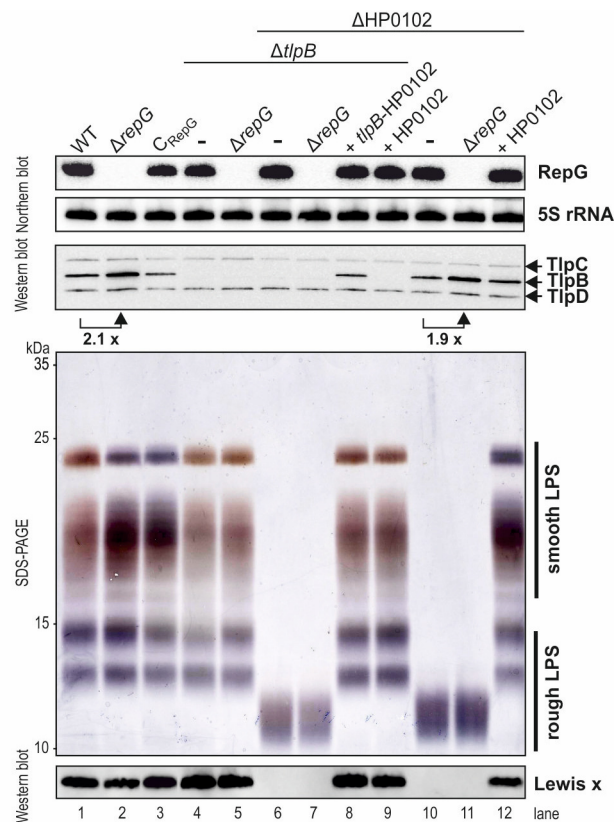
Figure 3.5: LPS isolated from *H. pylori* X47-2AL Δ*tlpB*-HP0102 and ΔHP0102 mutants contain no O-antigens. *H. pylori* strain X47-2AL wildtype and indicated mutant strains were grown to exponential growth phase. RNA and protein samples were analyzed by northern and western blot, respectively. RepG sRNA and chemotaxis receptors were detected as described in Figure 3.3. In parallel, Proteinase K-digested protein samples (OD$_{600nm}$ of 0.1) were separated on 15 % SDS-PAGE gels. Lipopolysaccharides were directly detected by silver staining, or electro-blotted to PVDF membrane and probed with anti-Lewis x antibody.

3.4. The glycosyltransferase HP0102 is essential for smooth LPS biosynthesis in various *Helicobacter pylori* strains

Unlike enterobacterial LPS, *H. pylori* produces smooth LPS with relatively constant O-chain length (Moran, 1995). However, variations in LPS band patterns, mobility and in-gel straining properties have been observed between different *H. pylori* strains, indicating strain-specific LPS sugar-chain modifications (Moran *et al.*, 1992). The LPS of the wildtype, Δ*repG* and ΔHP0102 deletion mutants of *H. pylori* strains 26695, J99, and G27 were analyzed in order to examine whether (I) the function of the glycosyltransferase HP0102 is

Figure 3.6: The glycosyltransferase HP0102 is involved in smooth LPS production in various *H. pylori* **strains. (A)** *H. pylori* 26695 wildtype, Δ*repG* and sRNA complementation strains (C~RepG~, SL 2, ΔCU), as well as Δ*tlpB* and ΔHP0102 mutants were grown to exponential growth phase. Total RNA and protein samples were analyzed by northern blot and western blot, respectively. LPS patterns and Lewis x antigen expression were investigated by silver staining and western blot using Lewis x antigen-specific antibody. **(B)** The effect on the LPS biosynthesis upon *repG* and HP0102 deletion was investigated in *H. pylori* strains J99 and G27.

conserved and (II) if RepG-mediated *tlpB* regulation affects HP0102 expression and thus, LPS biosynthesis. Irrespective of the parental strain background, mutant strains lacking HP0102 express only rough LPS without O-chains and Lewis x antigens (Figure 3.6).

The length of the homopolymeric G-repeat in the *tlpB* mRNA leader determines the outcome of RepG-mediated expression control of the chemotaxis receptor (Figures 2.8 and 2.9). While RepG did not influence TlpB protein levels in *H. pylori* strain J99 (13G), it represses and activates *tlpB* expression in 26695 (12G) and G27 (14G), respectively (Figure 2.8 B). Deletion of *repG* did not affect LPS structures and Lewis x antigen expression in strains J99 and G27 (Figure 3.6 B). However, increased band intensities in smooth LPS and elevated Lewis x antigen levels were observed upon *repG* deletion in *H. pylori* strain 26695 (12G) (Figure 3.6 A, lanes 1-2). Complementation of Δ*repG* with wild-type (C~RepG~) or mutant

RepG (SL2 and ΔCU) showed that the terminator loop of RepG is sufficient to repress both *tlpB* and HP0102, resulting in decreased O-chains and Lewis x antigen expression in this strain background (Figure 3.6 A, lanes 1-5). Although *H. pylori* 26695 wildtype and Δ*tlpB* mutant possess similar LPS patterns, increased Lewis x antigens levels were observed (Figure 3.6 A, lanes 1 and 6). This might be due to a polar effect on HP0102 expression in the Δ*tlpB* mutant (insertion of *erm-rpsL* resistance cassette including its own promoter in the *tlpB* locus).

3.5. RepG represses expression of HP0102 protein in *H. pylori* strain 26695

RepG-mediated *tlpB* expression control is significantly stronger in *H. pylori* strain 26695 (12G, five-fold) than in X47-2AL (7G, two-fold) (Figures 2.3 B and 3.3 C), which might explain why deletion of *repG* did only affect LPS O-chain patterns in *H. pylori* 26695 (Figures 3.5 and 3.6 A). As RepG was shown to repress expression of both *tlpB* and HP0102 at transcript level in *H. pylori* strain 26695 (Figure 3.1 B), this strain was used for investigations into the molecular mechanism of RepG-mediated co-regulation of the chemotaxis receptor and glycosyltransferase.

To assess whether HP0102 regulation by RepG is also reflected at the protein level, the first ten amino acids of the HP0102 coding region from *H. pylori* strain 26695 were fused to *gfpmut3.* To preserve the *tlpB*-HP0102 operon structure, the translational HP0102::GFP reporter fusion was introduced together with the upstream-encoded gene *tlpB*, including the *tlpB* promoter and 5' UTR (26695, 12G), into the *rdxA* locus of *H. pylori* strain G27 (*tlpB*-HP0102, Figure 3.7 A). While HP0102::GFP fusion protein levels were about two-fold up-regulated upon *repG* deletion, comparable TlpB protein levels were observed in *H. pylori* G27 wildtype and Δ*repG* (Figure 3.7 B, lanes 3-4 and 3.7 C). Because RepG can mediate both activation of the endogenous G27 *tlpB* (14G) and repression of the ectopic *tlpB* of the 26695 (12G) reporter construct, western blot analysis profiles represent the sum of these regulatory events. To examine whether translation of full-length *tlpB* is required for RepG-mediated HP0102 regulation, the 20th amino acid of the *tlpB* coding region was fused to its stop codon, resulting in a non-functional *tlpB*mini mini-gene within the *tlpB*mini-HP0102 10th::*gfpmut3* fusion (*tlpB*mini-HP0102, Figure 3.7 A)*.* Using western blot analysis, about three-fold increased HP0102::GFP fusion protein levels were detected in the *H. pylori* G27 Δ*repG* mutant carrying *tlpB*mini-HP0102 10th::*gfpmut3* reporter fusion when compared to the respective wild-type background (Figure 3.7, lanes 5-6 and 3.7 C). Conversely, endogenous TlpB levels (G27) were decreased upon RepG deletion. Thus, RepG-mediated HP0102 regulation is independent of the function of the chemotaxis receptor.
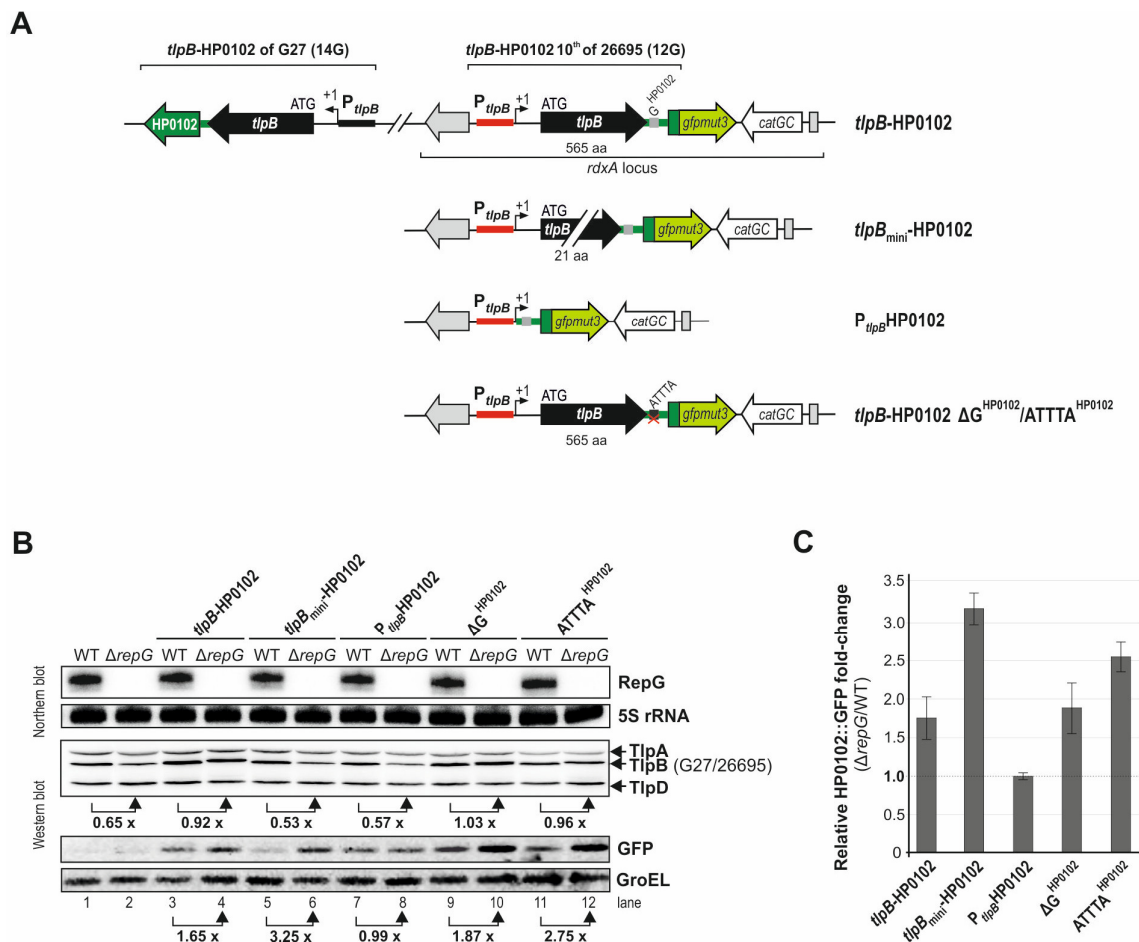
Figure 3.7: RepG represses HP0102 expression at the protein level. (A) Schematic representation of translational HP0102::GFP fusion constructs. The 10th codon of HP0102 of *H. pylori* strain 26695 was fused to *gfpmut3* and inserted together with *tlpB* (full length (565 aa) or mini gene (21 aa)), its 5' UTR and promoter (P_{tlpB}) into the *rdxA* locus of *H. pylori* G27 (*tlpB*-HP0102 or *tlpB*mini-HP0102). The *tlpB*-HP0102 IGR (containing a putative additional RepG binding site, G^{HP0102}) was fused together with the HP0102::*gfpmut3* coding region to P_{tlpB}. Also, the G^{HP0102}-repeat in the *tlpB*-HP0102 IGR was either deleted (ΔG^{HP0102}) or exchanged (ATTTA^{HP0102}). **(B)** *H. pylori* G27 wildtype, $\Delta repG$ as well as WT and $\Delta repG$ strains that carry the indicated reporter fusions (A) were grown to exponential phase, and RNA and protein samples were analyzed by northern (RepG, CSO-0003) and western blot (HP0102::GFP with anti-GFP antibody and TlpB with anti-TlpA22 antiserum), respectively. Note that TlpB levels in the strains that carry the *tlpB*-HP0102 (lanes 3-4), ΔG^{HP0102} (lanes 9-10) or ATTTA^{HP0102} (lanes 11-12) fusions are the sum of both, the endogenous G27 TlpB and ectopically expressed 26695 TlpB levels. **(C)** Quantification of relative HP0102::GFP protein levels observed in B (based on two biological replicates).

The IGR between *tlpB* and HP0102 contains a short homopolymeric G-repeat (G^{HP0102}) that might present an additional RepG interaction site in *H. pylori* strain 26695 (Figure 3.2). To investigate whether the G^{HP0102}-repeat is involved in RepG-mediated HP0102 regulation, the HP0102 10th::*gfpmut3* coding region and the 30 nt-long IGR between *tlpB* and HP0102 were fused to the *tlpB* promoter (P_{tlpB}HP0102, Figure 3.7 A). Comparable HP0102::GFP protein levels were observed in *H. pylori* G27 wildtype and $\Delta repG$ mutant

which carry the P$_{tlpB}$HP0102 10th::*gfpmut3* reporter fusion (Figure 3.7 B, lanes 7-8 and 3.7 C). This suggests two things: (I) the putative RBS in the *tlpB*-HP0102 IGR is sufficient to mediate HP0102 translation, and (II) RepG does not affect HP0102 expression by interaction with the G^{HP0102}-repeat. Accordingly, the G^{HP0102}-repeat of the *tlpB*-HP0102 10th::*gfpmut3* reporter fusion was either deleted (ΔG^{HP0102}) or exchanged to an ATTTA-stretch (ATTTA^{HP0102}). Neither the deletion (ΔG^{HP0102}) nor the exchange of the G^{HP0102}-repeat (ATTTA^{HP0102}) in the *tlpB*-HP0102 10th::*gfpmut3* reporter fusion did affect repression of HP0102::GFP through RepG (Figures 3.7 B, lanes 9-12 and 3.7 C). Variations in the HP0102::GFP protein abundance were observed in different GFP-reporter fusions, *e.g. tlpB*-HP0102 compared to ΔG^{HP0102} (Figure 3.7 B, lanes 3-4 and 9-10), suggesting that the length/composition of IGR between *tlpB* and HP0102 might influence HP0102 translation. However, this observation awaits further investigation.

Taken together, these results demonstrate that RepG co-regulates *tlpB* and HP0102 expression not only at the transcript level, but also at the protein level. In addition, data obtained from mutational analysis of the short homopolymeric G-repeat in the *tlpB*-HP0102 IGR suggest that the G^{HP0102}-repeat is dispensable for RepG-mediated HP0102 regulation.

3.6. The G-repeat length in the *tlpB* mRNA leader determines RepG-mediated HP0102 regulation and smooth LPS production

In order to investigate whether the length of the G-repeat affects HP0102 regulation by RepG, the transcript level of the dicistronic *tlpB*-HP0102 mRNA was determined in the previously-used *tlpB* mRNA leader variants of *H. pylori* strain 26695 wildtype and Δ*repG* (ΔG, 6-16G; Figure 2.9). Quantitative RT-PCR analysis revealed that RepG-mediated regulation of the *tlpB*-HP0102 mRNA is dependent on the G-repeat length (Figure 3.8). While a lack of the G-repeat (ΔG) in the *tlpB* mRNA leader had only a minor influence on the *tlpB* and HP0102 mRNA levels compared to *H. pylori* 26695 wildtype (12G), expression of both genes was increased (about five-fold) in the 6G-variant (Figure 3.8 A). A gradual decrease in the *tlpB* and HP0102 transcript levels was observed with an increasing number of guanines in the *tlpB* mRNA leader in the wild-type background, reaching a minimum for 8 to 12Gs. Further extension of the G-stretch from 13 to 16Gs resulted again in a gradual increase in *tlpB*-HP0102 mRNA abundance. Except for 16Gs, the levels of *tlpB* and HP0102 transcripts detected correlate well within different *tlpB* leader variants. While deletion of *repG* did not significantly affect *tlpB* and HP0102 expression in *tlpB* leader variants lacking the G-repeat (ΔG) or comprising a 6G- or 14G-long repeat, increased *tlpB* and HP0102 mRNA levels were observed in the *tlpB* leader variants 7-13G and 15-16G upon *repG*
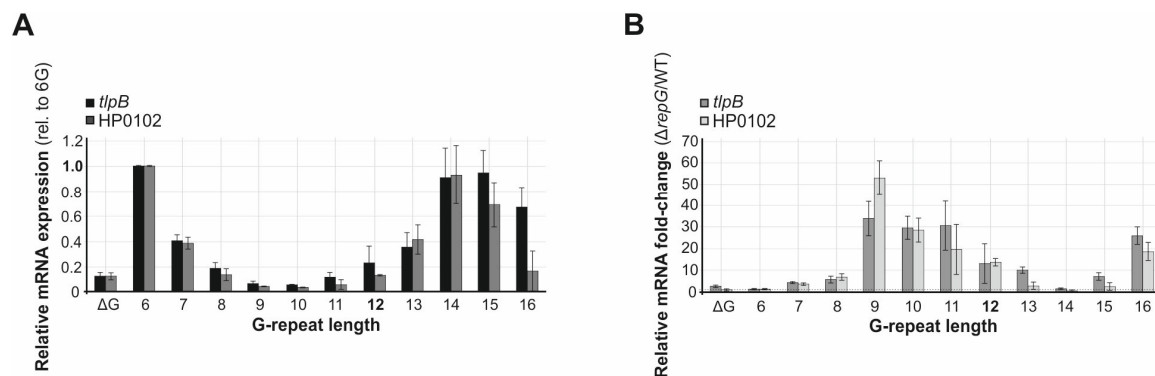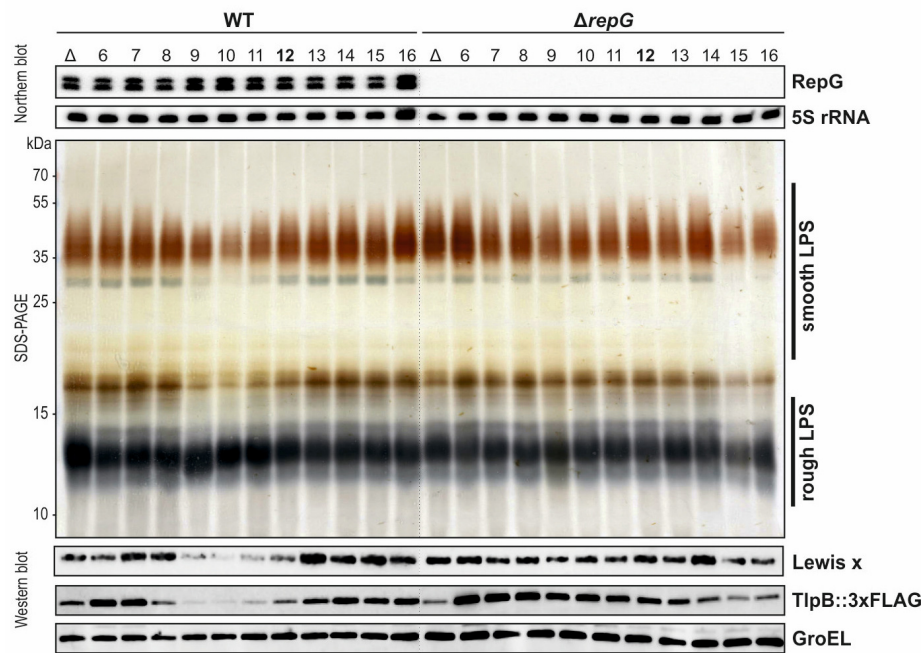
Figure 3.8: Variations in the homopolymeric G-repeat length determines RepG-mediated regulation of the dicistronic *tlpB*-HP0102 mRNA in *H. pylori* strain 26695. (A) Quantitative RT-PCR of relative *tlpB* and HP0102 mRNA levels in different *tlpB* leader mutants (ΔG, 6-16G) in the *H. pylori* 26695 wild-type background. *tlpB* and HP0102 mRNA levels in the *tlpB* 6G-leader variant were used as reference and set to 1. **(B)** Relative *tlpB* and HP0102 mRNA fold-changes upon *repG* deletion in *tlpB* leader mutants compared to the respective wild-type backgrounds. Error bars indicate standard deviation between two biological replicates.

deletion (Figure 3.8 B). Although this greatly correlates with the optimal window for RepG-mediated TlpB repression (7 to 12Gs, Figure 2.9 C), different or even opposing *tlpB* transcript abundances and protein levels were observed for longer G-repeat variants (13 to 16G) upon *repG* deletion (Figure 3.8 B). For example, although RepG was shown to promote *tlpB* expression on the protein level (Figure 2.9 C), its represses *tlpB*-HP0102 mRNA levels in the G-repeat leader variants 14G, 15G and 16G (Figure 3.8 B). The underlying mechanism for this divergent RepG-mediated regulation on transcript and protein level still needs to be clarified.

Dependent on the length of the homopolymeric G-repeat, RepG controls expression of the HP0102 mRNA. This suggests that variations in the G-repeat length might also affect LPS biosynthesis. Analysis of LPS patterns of the *tlpB* mRNA leader mutants (ΔG, 6-16G) in *H. pylori* 26695 wild-type and Δ*repG* background revealed that all investigated strains express smooth LPS. However, different band intensities of the O-chains and Lewis x antigen levels were detected (Figure 3.9). For example, while *tlpB* leader variants ΔG, 6-8G and 13-16G showed increased band intensities, lower amounts O-chains and Lewis x antigens were detected for variants with 9 to 11-nt long G-repeats when compared to the *H. pylori* 26695 wildtype (12G). In the *tlpB* leader variants 9-12G, deletion of *repG* resulted in significantly increased O-chain and Lewis x antigen expression. In contrast, LPS patterns remained unaltered in the *tlpB* leader variants ΔG, 6-8G, 13G and 14G and showed reduced LPS biosynthesis rates in *tlpB* leader variants with longer G-repeats (15 to 16Gs) upon *repG* deletion. Slight variations notwithstanding, LPS as well as Lewis x antigen expression and

Figure 3.9: The length of the homopolymeric G-repeat influences RepG-mediated *tlpB*-HP0102 co-regulation and thus, smooth LPS production in *H. pylori* 26695. The LPS patterns and Lewis x antigen levels in *tlpB* leader mutants (ΔG, 6-16G) of *H. pylori* 26695 wild-type and Δ*repG* mutant background grown to exponential phase were analyzed by silver staining and western blot analysis with anti-Lewis x antibody, respectively. Expression of TlpB::3xFLAG levels was analyzed by western blot with anti-FLAG antibody. GroEL served as a loading control.

TlpB protein levels closely correlates in the different G-repeat variants. In particular, RepG was shown to significantly repress TlpB protein levels and smooth LPS biosynthesis/O-antigen levels in *tlpB* leader variants that comprise a 9 to 12G-long repeat (optimal window for RepG-mediated repression). Overall, these data demonstrate that the length of the homopolymeric G-repeat influences RepG-mediated co-regulation of TlpB and HP0102 and, in turn, smooth LPS in *H. pylori* strain 26695.

3.7. Inactivation of HP0102 leads to strong autoagglutination, reduced motility, and lower survival under high-salt stress

LPS is essential for the integrity and functionality of the outer membrane of Gram-negative bacteria. To test whether changes in LPS structures through the loss-of-function of HP0102 and/or RepG-mediated HP0102 expression control are associated with phenotypic characteristics related to alterations in membrane integrity, autoagglutination activity, motility and sensitivity to sodium chloride-induced osmotic stress were compared between the wildtype and mutants in different *H. pylori* strains.
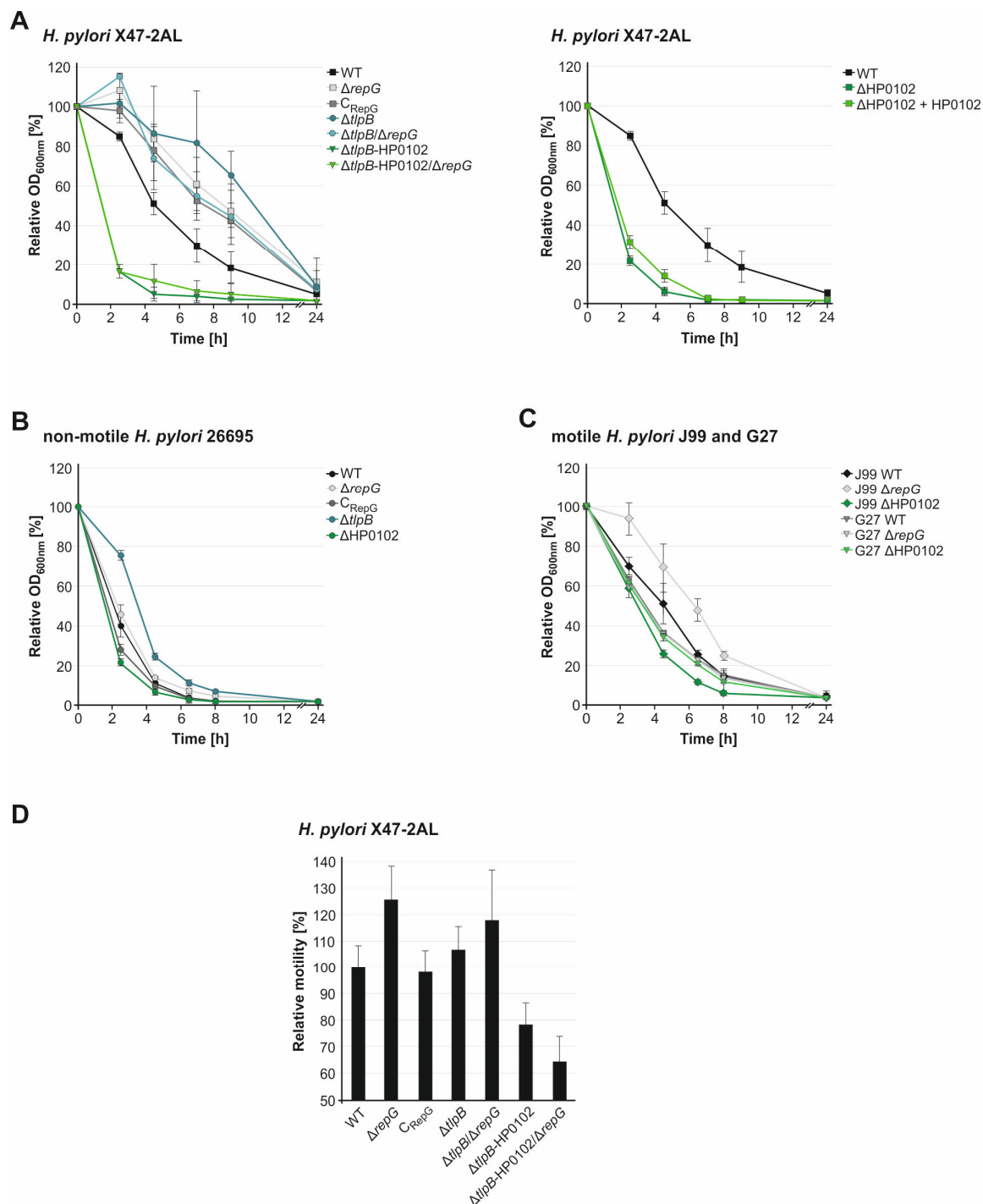
Figure 3.10: Deletion of HP0102 results in strong autoagglutination and reduced motility.
(A-C) Autoagglutination kinetics of different *H. pylori* wild-type and mutant strains at exponential phase. Cells were adjusted to an OD_{600nm} of 1.0 in 1 x phosphate-buffered saline (PBS) and incubated under microaerobic conditions for 24 h at 37 °C without shaking. After indicated time points the OD_{600nm} of the top 100 μl was measured. The OD_{600nm} at 0 min was set to 100 % and percentage of OD_{600nm} remaining at indicated time points was plotted. **(D)** The motility of the *H. pylori* X47-2AL wildtype and mutant strains was assayed by stab-inoculation on 0.4 % agar semi-solid Brucella broth plates. The diameter of the halo around the site of inoculation was measured after 5 days of incubation at 37 °C under microaerobic conditions. The colony diameter of the wildtype was used as reference and set to 100 %. Bars represent the mean determinations of triplicate measurements of relative colony diameter ± standard deviation.

Autoagglutination is caused by the interaction of aggregate-forming bacteria and is considered as a marker for the bacterial interaction with host cells and virulence in *Campylobacter* and *Helicobacter*, although this property varies considerably among strains and isolates (Misawa & Blaser, 2000 and Moran, 1995). Autoagglutination assays of *H. pylori* X47-2AL wildtype and mutant strains revealed slower autoagglutination kinetics for the Δ*repG*, Δ*tlpB* and Δ*tlpB*/Δ*repG* mutants when compared to the wildtype (Figure 3.10 A). In contrast, HP0102-deficient strains (Δ*tlpB*-HP0102, Δ*tlpB*-HP0102/Δ*repG* and ΔHP0102) exhibit a stronger bacterial aggregation phenotype, *i.e.* faster autoagglutination kinetics, than the wildtype. These data indicate that variations in the surface properties (LPS patterns) of investigated *H. pylori* mutants might be associated with variations in autoagglutination activities, *e.g.* faster autoagglutination kinetics might be linked to rough LPS expression in *H. pylori* strains lacking the glycosyltransferase HP0102. However, strain-specific autoagglutination characteristics (Δ*repG* or ΔHP0102 vs. wildtype in X47-2AL, 26695, J99 and G27) as well as potential secondary effects that might be caused by mutant construction impeded a direct link between HP0102 and/or RepG-mediated changes in LPS patterns and autoaggluntiation kinetics at this point of the thesis. For example, while deletion of HP0102 in *H. pylori* strains X47-2AL, 26695 and J99 resulted in faster autoagglutination kinetics compared to the parental control, wild-type autoagglutination characteristics were observed in G27 upon HP0102 deletion (Figures 3.10 A-C). In addition, although complementation of ΔHP0102 in *H. pylori* strain X47 restores smooth LPS production (Figure 3.5), it could not return autoagglutination kinetics to wildtype (Figure 3.10 A, *right panel*). Likewise, complementation of Δ*repG* did not result in *H. pylori* X47-2AL wild-type autoagglutination characteristics (Figure 3.10 A, *left panel*). This might suggest that among others, the destruction of the *rdxA* locus could potentially affect autoagglutination kinetics in *H. pylori.* Further studies, including *e.g.* the construction of a Δ*rdxA* mutant or complementation at a different genomic locus, will be required to investigate whether autoagglutination kinetics are directly linked to HP0102 and thus, to smooth LPS biosynthesis in different *H. pylori* strains. Please note that the motility behavior might also affect autoagglutination activties, *i.e.* non-motile *H. pylori* strain 26695 autoagglutinated faster than motile strains (Figures 3.10 A-C, see below).

Motility and chemotactic behavior of the *H. pylori* X47-2AL wildtype and mutant strains was assayed by stab inoculation assays on soft-agar plates. As reported for *H. pylori* strain SS1 (McGee *et al.*, 2005), *H. pylori* X47-2AL mutants lacking *tlpB* retained wild-type motility and chemotactic abilities (Figure 3.10 D). In contrast, while Δ*repG* and Δ*tlpB*/Δ*repG* double deletion mutants slightly increased outward migration, motility of the Δ*tlpB*-HP0102 and Δ*tlpB*-HP0102/Δ*repG* strains was significantly reduced compared to the wildtype. It still
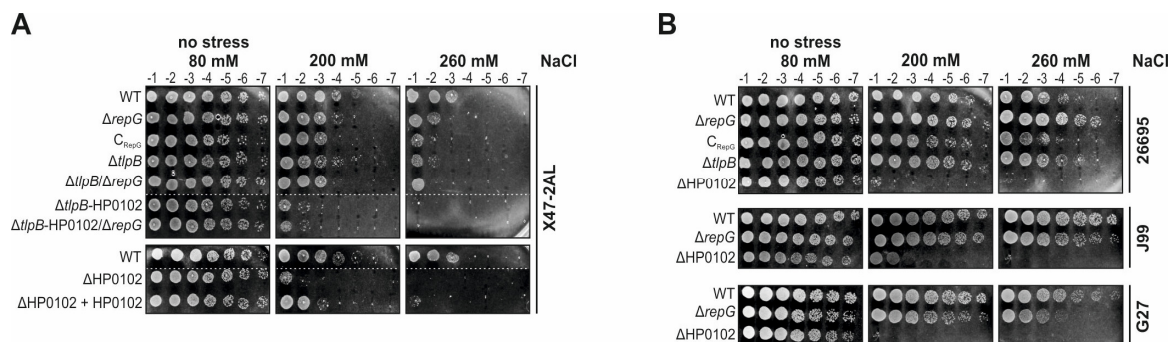
Figure 3.11: The glycosyltransferase HP0102 contributes to *H. pylori* survival under sodium chloride-induced membrane stress. (A) Survival of *H. pylori* strain X47-2AL wildtype and mutant strains on GC-agar plates with or without increased NaCl concentrations. Cells were grown to exponential growth phase (OD$_{600nm}$ of 1) and ten-fold dilutions of the indicated strains were spotted on GC-agar plates containing 80 mM (no stress), 200 mM (mild stress) or 260 mM (harsh stress) sodium chloride. Plates were incubated for 3 to 5 days at 37 °C under microaerobic conditions. The results shown are representative of at least two biological replicates. **(B)** Survival assays of the wildtype and mutants of *H. pylori* strains 26695, J99 and G27 under high-salt conditions.

needs to be verified whether deletion of HP0102 alone reduces *H. pylori* motility, and whether this phenotype could be rescued upon complementation.

Within the host or environment, bacterial pathogens encounter many threats to the integrity of their cell envelope, including changes in temperature, pH, and osmolarity. To assess whether mutations in *repG*, *tlpB*, the *tlpB*-HP0102 operon, or HP0102 alone affect bacterial survival under elevated salt concentrations in *H. pylori* strains X47-2AL, 26695, J99 and G27, viable counts/growth for each mutant strain and wild-type control were enumerated after plating the bacteria to agar plates with defined salt concentrations (Figure 3.11). A drastic defect in growth/survival on high salt media was observed for *H. pylori* mutants lacking either the *tlpB*-HP0102 operon or HP0102 alone (Figures 3.11 A and B), indicating that that HP0102 expression and thus, smooth LPS production might be associated with *H. pylori* survival under osmotic stress. However, complementation of the ΔHP0102 mutant only partially restored wild-type survival in *H. pylori* strain X47-2AL (Figure 3.11 A). While deletion of *tlpB* only slightly affected survival in all investigated *H. pylori* strain backgrounds, deletion of *repG* resulted in strain-specific sensitivities to sodium chloride-induced osmotic stress. In particular, deletion of *repG* facilitated survival on high-salt (260 mM NaCl) agar plates in *H. pylori* strain 26695, whereas an increased sensitivity was observed in G27 (Figure 3.11 B). As the smooth LPS biosynthesis and Lewis x antigens are unaffected by RepG in *H. pylori* strain G27 (Figure 3.6 B), this phenotype might be linked to regulation of other – so far undefined – target genes of RepG. Only slight effects were observed in the Δ*repG* mutants of *H. pylori* strains X47-2AL and J99.

Overall, the phenotypic characterization of either double or individual mutants demonstrates that disruption of the *tlpB*-HP0102 operon and HP0102, but not that of *tlpB* alone, strongly affects cell surface-related phenotypes in diverse *H. pylori* strains. Together, all data suggest that LPS O-chain production is dependent on the glycosyltransferase HP0102, and that smooth LPS of *H. pylori* is required for maintaining the integrity of the bacterial cell envelope.

3.8. RepG-mediated repression of HP0102 results in increased antibiotic sensitivity

The cell wall and outer membrane of Gram-negative bacteria provides an innate permeability barrier that limits the penetration of some antibiotics. Accordingly, modifications in the membrane permeability of bacteria by *e.g.* modifications in the LPS structure have been shown to prevent/affect the access of some antibiotics to their targets at required concentrations (Chen & Groisman, 2013). Therefore, deletion of HP0102 and thus, modifications in the LPS structure could affect *H. pylori* antibiotic sensitivity. To test this, the wildtype and mutants (*e.g.* Δ*repG*, C_{RepG}, Δ*tlpB*, ΔHP0102 and ΔHP0102 + HP0102) of *H. pylori* strains X47-2AL, 26695, J99, and G27 were characterized for rifampicin and polymyxin B resistance by disk diffusion assays. Rifampicin is a large hydrophobic antibiotic, which diffuses across the outer and cytoplasmic membrane, and inhibits the DNA-dependent RNA polymerase. In contrast, polymyxin B alters the permeability of bacterial cell walls by interaction with a negatively charged site on the lipid A of the LPS layer and interruption of the cytoplasmic membrane. Irrespective of the parental strain background, mutants lacking HP0102 displayed higher sensitivities to rifampicin and polymyxin B than the respective wild-type backgrounds (Figure 3.12). In general, polymyxin B had a larger inhibitory/killing effect on HP0102-deficient mutants than rifampicin (relative to the respective WT), which might be linked to the different mechanism of action and/or targets of these antibiotics. Complementation of the ΔHP0102 mutant in *H. pylori* strain X47-2AL restored wild-type levels of antibiotic susceptibility, suggesting that the loss of the glycosyltransferase HP0102 and thus, modification in the LPS structure (*i.e.* rough LPS) might be directly linked to antibiotic sensitivity of *H. pylori*. While no significant effect on antibiotic sensitivity was observed upon *repG* deletion in *H. pylori* strains X47-2AL, J99 and G27, the Δ*repG* mutant of strain 26695 was slightly less susceptible to polymyxin B compared to the wildtype. This suggests that increased HP0102 expression and thus, O-chain production upon *repG* deletion might contribute to the integrity and permeability of the bacterial membrane.
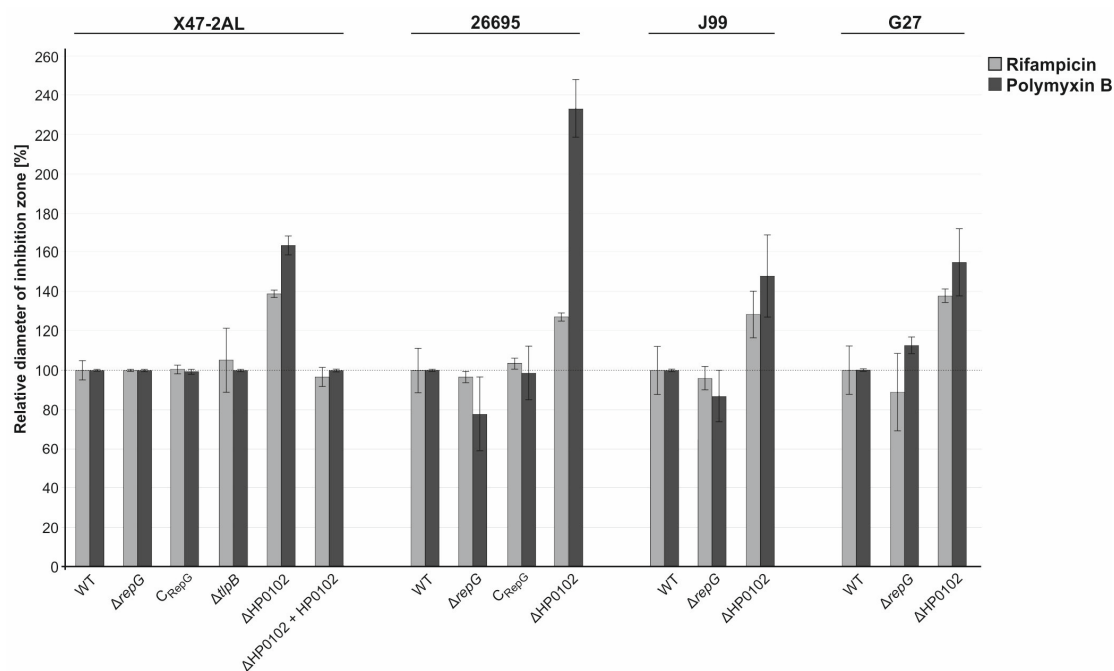
Figure 3.12: *H. pylori* **ΔHP0102 mutants are more sensitive to rifampicin and polymyxin B.**
Antibiotic sensitivity testing of *H. pylori* X47-2AL, 26695, J99 and G27 wildtype and indicated mutant
strains grown to exponential growth phase using disk diffusion assays. *H. pylori* cells corresponding
to an OD_{600nm} of 0.01 were plated on GC-agar plates and disks impregnated with 10 µg/ml rifampicin
or 300 units polymyxin B were placed on top. The inhibition zone was measured after 3 days of
incubation under microaerobic conditions at 37 °C. The inhibition zone diameter of the wildtype was
used as reference and set to 100 %. Bars represent the mean determinations of duplicate
measurements of relative inhibition zone diameter (relative to the respective WT of each *H. pylori*
strain) ± standard deviation.

To investigate whether posttranscriptional control of HP0102 through RepG also
contributes to antibiotic sensitivity, *tlpB* leader variants in the wild-type or Δ*repG* mutant
background of *H. pylori* strain 26695 were characterized for polymyxin B resistance.
Previously, TlpB protein levels and HP0102-dependent smooth LPS expression showed
strongest increase in the *tlpB* 6G-leader variant when compared to the *H. pylori* 26695
wildtype (12G, Figures 2.9 and 3.9). Considering that this might be correlated to the "lowest"
sensitivity of *H. pylori* 26695 to polymyxin B, the inhibition zone diameter of the 6G variant
was used as reference. Disk diffusion assays of the *tlpB* leader variants in the *H. pylori* 26695
wild-type background revealed similar inhibition zone diameters for variants ΔG, 6-8G or
13-16G (Figure 3.13, *left panel*). In contrast, *H. pylori* mutants containing 9 to 12G-long
repeats were about about 1.5- fold more susceptible to polymyxin B compared to the 6G-
leader variant. In addition, deletion of *repG* did not significantly affect the antibiotic
susceptibility in the *tlpB* leader variants ΔG, 6-8G and 13-16G; however, increased
resistance to polymyxin B was detected for *tlpB* leader mutants 9-12G (Figure 3.13, *right
panel*). This suggests that RepG-mediated repression of the glycosyltransferase HP0102 and

thus, reduction in LPS O-chain production correlates with an increase in antibiotic resistance.
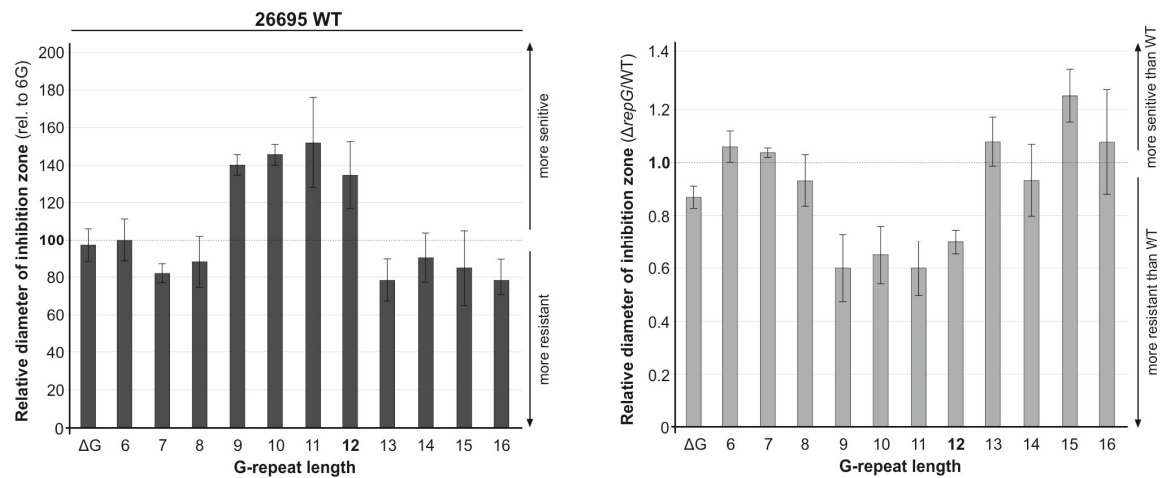
Figure 3.13: RepG-mediated HP0102 repression is correlated with an increase in sensitivity to polymyxin B. (*Left panel*) Antibiotic sensitivity testing of *H. pylori* 26695 *tlpB* mRNA leader variants (ΔG, 6G to 16G) in wild-type background using disk diffusion assays (see also Figure 3.12). The inhibition zone diameter of the *tlpB* 6G-leader variant was used as reference and set to 100 %. (*Right panel*) Relative fold-changes in antibiotic sensitivity (inhibition zone diameter) upon *repG* deletion in diverse *tlpB* leader variants compared to the respective wild-type background. Error bars indicate standard deviations among two/three biological replicates.

3.9. Discussion: The length of an SSR determines sRNA-mediated gradual expression control of both the chemotaxis receptor TlpB and the glycosyltransferase HP0102

In *H. pylori*, phase-variable SSRs have been shown to regulate gene expression of bacterial surface structures (LPS and outer membrane proteins) and in turn, host recognition and adhesion (*e.g.* Appelmelk *et al.*, 1999, Yamaoka *et al.*, 2002). This Chapter provides a possible answer to the question: Why is the expression of the chemotaxis receptor TlpB coupled to a homopolymeric G-repeat? Because it is encoded in an operon together with the glycosyltransferase HP0102, which represents a novel LPS biosynthesis factor in *H. pylori.* Using genetic and biochemical approaches, this glycosyltransferase was shown to be essential for LPS O-chain and Lewis x antigen production in diverse *H. pylori* strains. In addition, HP0102-dependent smooth LPS production was shown to be required for the integrity and permeability of the bacterial membrane, and more importantly, to be crucial for colonization of the murine stomach by *H. pylori*.

Besides the functional characterization of HP0102, work presented in this Chapter demonstrates that antisense base-pairing of RepG to the homopolymeric G-repeat in the *tlpB* mRNA leader controls gene expression of both the chemotaxis receptor and the glycosyltransferase. In line with previous results (Chapter 2), the length of the G-repeat determines the outcome of posttranscriptional regulation of *tlpB* and HP0102. In particular, RepG represses *tlpB*-HP0102 expression when the G-repeat is composed of 7 to 12Gs and activates it for G-stretches longer than 14Gs (Figure 3.14). In contrast to the ON/OFF switch of LPS-modifying enzymes through SSRs in coding regions, such posttranscriptional control of HP0102 allows for a gradual control of LPS biosynthesis. This fine-tuning might be required for adaptation of *H. pylori* to a specific niche in the human host. Furthermore, it provides the opportunity to interconnect environmental signals, such as acidic stress, to LPS O-chains assembly and Lewis x antigens synthesis.
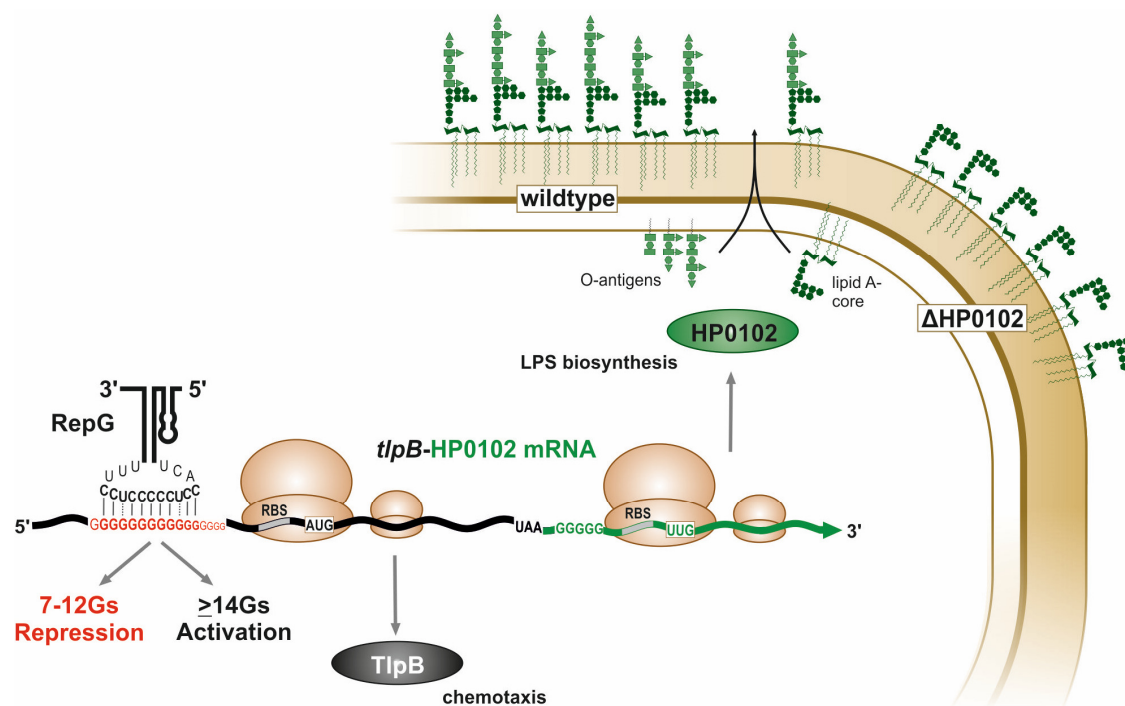
Figure 3.14: The G-repeat length in the *tlpB* mRNA leader determines the outcome of RepG-mediated co-regulation of the chemotaxis receptor TlpB and the glycosyltransferase HP0102. *tlpB* and HP0102 are encoded in an operon. The chemotaxis receptor TlpB is assumed to play a role in pH-taxis and quorum sensing. The glycosyltransferase HP0102 is involved in smooth LPS biosynthesis (O-chains and Lewis x antigens). Antisense base-pairing of the C/U-rich terminator loop of RepG to the homopolymeric G-repeat in the 5' UTR of the *tlpB* mRNA results in posttranscriptional co-regulation of *tlpB* and HP0102 at the transcript and protein level. Depending on the G-repeat length, RepG mediates both repression (7-12Gs) and activation ($\geq$ 14Gs) of *tlpB* and HP0102.

Coordinated regulation of the *tlpB*-HP0102 operon through RepG. In bacteria, operonic gene organization is a mechanism that ensures coordinated expression of genes encoding for related or coupled functions. Besides co-regulation at the transcriptional level, activity of sRNAs has been shown to result in simultaneous activation or repression of all genes in an operon (coordinate regulation), or to uncouple gene expression of single genes within polycistrons, resulting in discoordinate operon regulation (reviewed in Balasubramanian & Vanderpool, 2013). RepG was shown to co-regulate expression of both *tlpB* and HP0102 at the transcript and protein level (Figures 3.1 B and 3.7). Although a putative additional RepG interaction site (G^{HP0102}) was predicted upstream of the HP0102 RBS, RepG base-pairing to the homopolymeric G-repeat in the *tlpB* mRNA leader is sufficient for coordinated/coupled regulation of the entire *tlpB*-HP0102 operon (Figure 3.7). While being about eight- to ten-fold up-regulated at the transcript level (Figures 3.1 B), TlpB and HP0102 protein levels were only two- to five-fold altered in *H. pylori* strain 26695 upon *repG* deletion (Figures 2.3 B, 2.4 B, 2.9 and 3.7). Various sources of this discrepancy can be proposed, including variations in the mRNA or protein stability of epitope-tagged TlpB and/or GFP-reporter fusion constructs and, as discussed in Chapter 2, an active recruitment of RNases upon RepG and *tlpB*-HP0102 mRNA interaction. Further studies will be required to unveil the underlying mechanism of RepG-mediated *tlpB*-HP0102 co-regulation. For example, mutagenesis of the *tlpB* start codon could provide further insight whether or not *tlpB* and HP0102 are translationally coupled, and whether RepG-mediated alterations in the efficiency of *tlpB* translation accompanies HP0102 expression control.

The glycosyltransferase HP0102 is involved in O-chain biosynthesis or modification of the core oligosaccharide. Although LPS molecules are located on the bacterial outer membrane, their synthesis is initiated in the cytoplasm. In *H. pylori*, like in many other Gram-negative bacteria, the O-specific polysaccharide chain and lipid A-core moiety are assembled separately at the inner leaflet of the inner membrane. After translocation through the inner membrane, the O-chain is transferred to the lipid A-core structure, followed by translocation through the outer membrane by a lipoprotein transport system. While little is known about the periplasmatic components that are required for LPS translocation, *H. pylori* is assumed to use a novel LPS assembly pathway that is evolutionarily connected to protein N-glycosylation (Figure 3.15, Hug *et al.*, 2010). The O-specific polysaccharide chain is assembled onto an undecaprenyl-phosphate (UndP) lipid carrier by various specific glycosyltransferases. Among them, diverse galactosyl- (GalTs) and N-acetylglucosaminyltransferases (GlcNAcTs) mediate the sequentially addition of the monosaccharide units (galactose, N-acetylglucosamine), forming the linear O-chain backbone (*e.g.* Logan *et al.*, 2005). These O-chain backbones can be decorated at selected
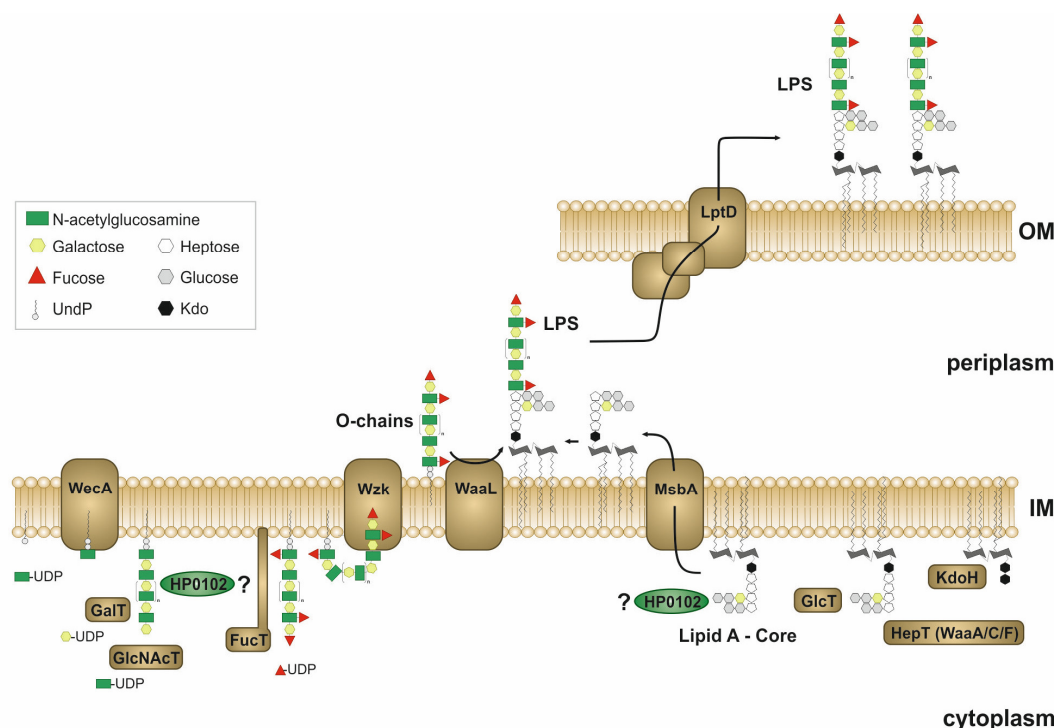
Figure 3.15: LPS biosynthesis pathway in *H. pylori*. The lipid A-core and O-chains are assembled independently from each other in the cytoplasm. *Lipid A-core:* Heptose and glucose units are assembled to the lipid A by multiple heptosyl- (HepTs) and glucosyltransferases (GlcTs). *O-chains:* The O-chains are assembled onto a polyisoprenoid membrane anchor (UndP). The initiating glycosyltransferase WecA transfers an N-acetylglucosamine unit from a nucleotide-activated donor (uridine diphosphate (UDP)-GlcNAc) to the lipid carrier, providing the platform for the O-chain synthesis. The linear O-chain backbone is assembled by processive galactosyl- (GalTs) and N-acetylglucosaminyltransferases (GlcNAcTs). Fucosyltransferases (FucTs) attach fucose units to the O-chain backbone, generating Lewis antigens. O-chains and the lipid A-core moiety are translocated through the inner membrane (IM) by the flippase Wzk and transmembrane protein MsbA, respectively, and are assembled to each other by O-chain ligase WaaL. The LPS molecule is then transported to the outer leaflet of the outer membrane (OM). The glycosyltransferase HP0102 might be involved in synthesis of the O-chain backbone or modification of the core oligosaccharide.

locations though the activity of various fucosyltranferases (FucTs), producing the surface Lewis antigens (Appelmelk & Vandenbroucke-Grauls, 2000). Independent from the O-chain biosynthesis, multiple heptosyl- (HepTs) and glucosyltransferases (GlcTs) mediate the assembly of the inner and outer core oligosaccharide through the addition of heptose and glucose units to the lipid A moiety, respectively (*e.g.* Langdon *et al.*, 2005). Modifications of the lipid A are mediated by Kdo-hydrolases (KdoH, Stead *et al.*, 2010).

Unlike most bacteria, the genes involved in LPS biosynthesis in *H. pylori* are not arranged in a single cluster, but rather found in various locations distributed throughout the chromosome. Overall, the *H. pylori* genome encodes approximately 30 genes that have been shown or predicted to be involved in LPS biosynthesis (Tomb *et al.*, 1997). Besides proteins responsible for LPS translocation, more than 20 putative glycosyltransferases have been

identified in *H. pylori* strain 26695 (Logan *et al.*, 2005). Among them, HP0102 encodes a glycosyltransferase of the GT-2 family. Considering that members of the GT-2 family typically possess β-glucosyl- and N-acetylglucosaminyltransferase activities and *H. pylori* ΔHP0102 mutants displayed only rough LPS without O-chains and Lewis x antigens (Figures 3.5 and 3.6), the glycosyltransferase HP0102 might be involved in the assembly of N-actetylglucosamine to the O-chain backbone (Figure 3.15). In line with this, HP0102 was predicted to be homologous to β-glycosyltransferases identified in *E. coli* (β-GlcT, WaaV), *Yersinia pestis* (β-GlcT, WbyL), *Streptococcus pneumoniae* (β3-GlcNAcT, Cps14I and β4-GalT, Cps14J), *Sinorhizobium meliloti* (β6-GlcT, ExoO), and *Rhizobium etli* (ORF4) (Vinuesa *et al.*, 1999, Skurnik *et al.*, 2000). All of these enzymes have been shown or predicted to be involved in LPS O-chain biosynthesis. For example, mutagenesis of the glycosyltransferase locus ORF2-ORF4 resulted in rough LPS-producing *R. etli* (Vinuesa *et al.*, 1999). Similarly, *H. pylori* mutants that are deficient for either O-chain-producing galactosyl- or N-acetylglucosaminyltransferases have been shown to lack O-chains and Lewis antigen expression (Logan *et al.*, 2000, Logan *et al.*, 2005).

Alternatively, HP0102 might contribute to the biosynthesis of the LPS core. HP0102 shares homology with the *E. coli* β-glycosyltransferases WaaV, which catalyzes the addition of glucose to the core oligosaccharide (Heinrichs *et al.*, 1998). This sugar-linkage generates core oligosaccharide specificity and is crucial for the organization of *E. coli* LPS. Likewise, an α-1,6-polymeric glucan chain was identified in the core oligosaccharide of many *H. pylori* strains (Altman *et al.*, 2003, Logan *et al.*, 2005). Mutational analyses of glycosyltransferases involved in the biosynthesis of the LPS core, especially of the α-1,6-glucan polymer, have also been shown to result in a truncated LPS molecule devoid of Lewis antigens (Langdon *et al.*, 2005, Logan *et al.*, 2005). Therefore, HP0102 might be (also) involved in the modification of the *H. pylori* LPS core, which is required for linkage between the O-chains and the lipid A-core moiety.

Also, deletions of other LPS biosynthesis components, such as the O-chain ligase WaaL, have been shown to result in rough LPS-producing *H. pylori* strains (Hug *et al.*, 2010). In principle, HP0102 could therefore also encode an enzyme that catalyzes similar reactions, *e.g.* the transfer of the O-chain to the lipid A core. However, based on *in-silico* predictions (SignalP, HMMTOP), HP0102 lacks any kind of transmembrane domain and/or signal peptide sequence, indicating that this glycosyltransferase is localized in the cytoplasm and participates in either O-chain or LPS core biosynthesis. Further studies including LPS mass spectrometry, chemical analysis, and nuclear magnetic resonance (NMR) analysis are required to determine the exact LPS structure of ΔHP0102 mutants and the linkage specificity of the glycosyltransferase HP0102.

HP0102 – a novel colonization factor in *H. pylori*. This Chapter provides evidence that the glycosyltransferase HP0102 represents a novel colonization factor in *H. pylori*. Mouse infection studies revealed a strong colonization defect for HP0102-deficient mutants, which produce only rough LPS without O-chains and Lewis x antigens (Figures 3.3 D and E). In line with previous studies (Moran *et al.*, 2000), this demonstrates that LPS O-specific polysaccharide chains and Lewis antigens are crucial for host colonization by *H. pylori*. LPS and especially the O-antigens, promote *H. pylori* adherence to epithelial cells (Valkonen *et al.*, 1997, Edwards *et al.*, 2000, Fowler *et al.*, 2006). Therefore, the loss of LPS O-chains and Lewis x antigen expression in ΔHP0102 mutants might have resulted in an impaired ability of these mutants to bind to the gastric epithelium and thus, hindered initial colonization. Nevertheless, *H. pylori* core oligosaccharides of rough LPS are also able to bind protein factors within the gastric mucosa (Reeves *et al.*, 2008).

LPS is crucial for the physical integrity and function of membranes of Gram-negative bacteria. Accordingly, HP0102-dependent modification in the LPS structure of diverse *H. pylori* strains caused pleiotropic effects on bacterial cell physiology (Figures 3.10-3.12). Therefore, the inability of *H. pylori* strain X47-2AL to colonize the murine stomach might be also linked to the observed motility defect of HP0102-deficient mutants and/or altered autoagglutination kinetics. Autoagglutination is considered to be a marker for bacterial adherence to and/or invasion of host cells, and its importance for bacterial virulence has been described in various pathogens, including *Yersinia enterocolitica*, *E. coli*, and *C. jejuni* (Roggenkamp *et al.*, 1995, Knutton *et al.*, 1999, Guerry *et al.*, 2006). Interestingly, bacterial cell densities and quorum sensing, *i.e.* the production of the signal molecule AI-2, have been shown to affect autoagglutination kinetics in *C. jejuni* (Jeon *et al.*, 2003). In line with this, deletion of AI-2-sensing chemotaxis receptor *tlpB* was shown to slow down autoagglutination kinetics in diverse *H. pylori* strains (Figures 3.10 A-C).

In contrast to HP0102, *tlpB* hardly influenced *H. pylori* X47-2AL colonization of the murine stomach (Figure 3.3 D). The role of this chemotaxis receptor in *H. pylori* colonization is controversially discussed in the literature. While Δ*tlpB* mutants were defective for colonization of highly permissive C57BL/6 interleukin-12 (IL-12) (p40-/-)-deficient mice (Croxen *et al.*, 2006), McGee and coworkers observed that *H. pylori* strains lacking *tlpB* colonize gerbil stomachs to wild-type levels, but exhibit differences in gastric inflammation (McGee *et al.*, 2005). In the latter case, the authors verified HP0102 expression in the Δ*tlpB* mutant. Therefore, it is possible that Δ*tlpB* mutant construction in the study of Croxen *et al.*, resulted in a polar effect on the HP0102 gene expression and thus, could explain the colonization defect of the Δ*tlpB* mutant.

In many bacterial pathogens, a wealth of sRNAs have been found to play profound roles not only in physiology, but also pathogenicity (Caldelari *et al.*, 2013, Papenfort & Vogel,

2014). Here, RepG was shown to regulate expression of the glycosyltransferase HP0102, a novel virulence-associated factor that is essential for *H. pylori* colonization of the murine stomach. This indicates that sRNAs and, in particular, RepG could be involved in *H. pylori* virulence control. However, no significant colonization phenotype has been observed for the *H. pylori* X47-2AL Δ*repG* mutant in *in-vivo* infection studies (Figure 3.3 D). Similarly, inhibition of many well-characterized enterobacterial sRNAs, which have been shown to regulate genes that are important for bacterial pathogenicity, often do not lead to a pronounced growth defect in *in-vivo* models of infection (Barquist *et al.*, 2013, Chaudhuri *et al.*, 2013). Depending on the length of the homopolymeric G-repeat in the *tlpB* mRNA leader, RepG co-regulates *tlpB* and HP0102 in a rheostat-like fashion, *i.e.* mild repression for 7 to 8Gs and stronger repression for 9 to 12Gs (Figures 3.8 and 3.9). Since the G-repeat of *H. pylori* strain X47-2AL is only composed of 7Gs, colonization defects upon deletion of *repG* might be less pronounced than in *H. pylori* strains harboring a 10 to 12-nt long G-repeat. Future studies will be required to investigate whether a "stronger" RepG-mediated *tlpB*-HP0102 repression, *i.e.* through modification of the G-repeat length from 7 to 10Gs in the *tlpB* mRNA leader of *H. pylori* X47-2AL wild-type and Δ*repG* mutant background, will influence colonization of the murine stomach by *Helicobacter.* In line with this, it would be interesting to investigate the switching frequency of the G-repeat and thus, outcome of RepG-mediated *tlpB*-HP0102 regulation *in vivo* (*e.g.* during mouse infections studies).

Small RNA-mediated LPS modification results in increased antibiotic susceptibility. Besides effects on motility and autoagglutination, HP0102-dependent modifications in LPS have been shown to be associated with an increased sensitivity of *H. pylori* to membrane stress and antibiotics treatment. In addition, RepG-mediated repression of HP0102 (for *tlpB* leader variants 9-12G) resulted in a decrease of in O-chains and Lewis x antigen production, thereby causing *H. pylori* 26695 to be more susceptible to the antimicrobial peptide polymyxin B. Such sRNA-mediated posttranscriptional control of genes involved in LPS modification and/or transcriptional regulators thereof, has also been shown to affect antibiotic resistance in enterobacteria. For example, the MgrR and ArcZ sRNAs have been shown to directly repress *eptB*, a gene encoding an LPS (lipid A-core)-modifying enzyme that contributes to resistance to antimicrobial peptides in *E. coli* (Moon & Gottesman, 2009, Moon *et al.*, 2013). In addition, MicA sRNA was shown to indirectly affect LPS modification and thus, resistance to antimicrobial peptides in *Salmonella* and *E. coli* through regulation of the PhoPQ two-component system (Coornaert *et al.*, 2010).

Modifications of LPS, especially of the lipid A moiety, have been shown to affect the sensitivity of *H. pylori* to polymyxin B and other cationic antimicrobial peptides such as human β-defensin 2 (Stead *et al.*, 2010, Cullen *et al.*, 2011). Human β-defensin 2 and 3 are important components of the first-line innate mucosal defense system and possess high

killing activity against *H. pylori* (George *et al.*, 2003, Bauer *et al.*, 2012). Therefore, it would be interesting to investigate whether RepG-mediated HP0102 repression might alter the response of *H. pylori* to these antimicrobial peptides. This will provide new insights into the importance of sRNAs in posttranscriptional gene regulation of bacterial surface structures and its physiological impacts.

Posttranscriptional, gradual control of LPS biosynthesis. In *H. pylori*, LPS and Lewis antigen profiles vary extensively among different strains and isolates (*e.g.* Monteiro *et al.*, 2000). Moreover, O-antigen structures of *H. pylori* strains are assumed to "be adapted" to the individual human host, enabling the establishment of a chronic infection or colonization of a specific niche (Nilsson *et al.*, 2006). Depending on the expression of diverse fucosyltransferases, *H. pylori* produces LPS O-chains with different Lewis epitopes (Moran, 2008, Appelmelk *et al.*, 2000). While usually possessing Lewis x and/or Lewis y antigens (80-90 % of the strains), *H. pylori* strains/isolates have also been shown to express Lewis a, b, and c, sialyl-Lewis x and H-1 as well as blood group A and B antigens. This variable expression of O-chains and Lewis antigens has been implicated in bacterial adhesion to host cells and gastric colonization, immune evasion by molecular mimicry, modulation of the host response through interaction with immune cells, and induction of gastric autoimmunity (Moran, 2008, Chmiela *et al.*, 2014).

Several enzymes of the LPS biosynthesis pathway, but especially the fucosyltransferases and glycosyltransferases, are phase-variably expressed in *H. pylori* (Bergman *et al.*, 2006). Among them, all three FucTs (HP0379, HP0651, HP0093/94) are subjected to translational frame-shifting in homopolymeric C-repeats or TAA-tracts within their coding regions, thereby causing enzymes to switch between an active and silent stage (Appendix, Table 13.2; Appelmelk *et al.*, 1999, Wang *et al.*, 1999). In addition, GlcTs that are involved in the assembly of the O-chain backbone (HP0619) and modification of the core oligosaccharide (HP0208), respectively, or the O-chain flippase Wzk (HP1206) also contain SSRs within their coding region, indicating a phase-variable ON/OFF switch of gene expression (Oleastro *et al.*, 2006, Oleastro *et al.*, 2010 and Langdon *et al.*, 2005). Recently, the identification of a homopolymeric C-repeat in the ORF of Kdo-hydrolase HP0579/580 suggested that phase-variable modifications of the lipid A moiety might also contribute to host adaptation in *H. pylori* (Stead *et al.*, 2010). While the majority of LPS-modifying genes in *H. pylori* show rather an ON/OFF expression, posttranscriptional control of HP0102 by RepG allows for a gradual expression of this glycosyltransferase. In *H. pylori* strain 26695, two other examples of LPS-modifying enzymes, which are associated with an intergenic SSR have been identified (Appendix, Table 13.2). In particular, homopolymeric A-repeats are present in the promoter region of the FucT HP0651 and the 5' UTR of the core-modifying

GlcT HP0208. Whether these repeats affect promoter strength and/or are targeting sites for posttranscriptional gene regulation by sRNAs still needs to be clarified.

It has been suggested that the host Lewis phenotype may preferentially select for the expression of particular Lewis determinants by *H. pylori* (Pohl *et al.*, 2009). Thus, the host milieu promotes selection of bacterial strains with particular characteristics that facilitate adaptation and survival in the gastric mucosa of the individual host, and shape the bacterial community structure (Skoglund *et al.*, 2009). In addition, acid-induced variation in LPS O-antigen expression has been observed during *in-vitro* growth of *H. pylori* (Moran *et al.*, 2002), suggesting that environmental triggers such as acid can induce phase variation or at least, affect transcription of enzymes involved in their synthesis (McGowan *et al.*, 1998). While *tlpB*-HP0102 mRNA levels are not affected by acid stress, RepG expression is pH-dependent (Figure 2.19 A). Therefore, RepG-mediated regulation of *tlpB* and HP0102 might represent a mechanism to posttranscriptionally control the chemotactic behavior and LPS biosynthesis of *H. pylori* under *e.g.* acidic/alkaline conditions. The coupling and coordinate expression control of both the chemotaxis receptor and the glycosyltransferase might be important for the adaptation to the gastric niche. In particular, posttranscriptional control of *tlpB* might affect the negative chemotactic response of *H. pylori* to low pH (Croxen *et al.*, 2006), and therefore, its ability to remain in/orient itself towards an optimal pH zone. Likewise, HP0102-dependent LPS modifications might contribute to *H. pylori* survival under acidic stress conditions. *H. pylori* mutants lacking LPS O-chains (*e.g.* Δ*wbcJ*) have been shown to be more sensitive to acid stress than the wildtype (McGowan *et al.*, 1998), suggesting that structural changes in the LPS are important for *H. pylori* to respond or withstand acid shock. The control of *tlpB* and HP0102-dependent LPS modifications by RepG provides an additional layer of regulation necessary in the changing environments inside or outside of the host.

Also, seen from another perspective, RepG-mediated repression of *tlpB* and HP0102 might provide a selective pressure, *e.g.* under a given environment, thereby promoting genetic adaptation through phase variation (*i.e.* change in the G-repeat length). In line with this, natural competence of *H. pylori* (*e.g.* DNA uptake) was shown to be reduced in mutants producing only rough LPS (Hug *et al.*, 2010), suggesting a connection between the genetic variability and LPS production/modifications in *H. pylori*.

Overall, it seems that *H. pylori* LPS synthesis and modification is a balance act between maintaining membrane integrity, mediating the adhesion to the host cells and escape from the immune system. This Chapter provides evidence that posttranscriptional regulation through sRNAs represents an important layer of regulation in the complex LPS biosynthesis and modification of *H. pylori*.

4. The conserved and abundant riboregulator RepG controls gene expression by antisense base-pairing to G-rich sequences

While initial studies reported interactions of sRNAs with only a single or, at most, a few target mRNAs, it is now evident that sRNAs can regulate expression of multiple target mRNAs in distinct regulons, thus integrating posttranscriptional sRNA activities in extensive regulatory networks. Thereby, sRNAs have been shown to be key regulators of diverse metabolic processes and in microbial virulence.

The intensive study of the interaction between RepG and the *tlpB*-HP0102 mRNA has provided first insights into molecular mechanisms of sRNA-mediated gene expression control in Epsilonproteobacteria (Chapters 2 and 3). However, it remains elusive whether RepG controls additional target mRNAs in *H. pylori*. The identification of the cellular interaction partners of RepG, besides the *tlpB*-HP0102 operon, will provide further insights into the biological functions of this riboregulator and its implications in various physiological responses in *Helicobacter*. In this Chapter, microarray and RNA-seq is applied for global target gene identification of RepG in *H. pylori* strains 26695 and G27. Whole transcriptome analysis reveals that RepG acts as a multi-target riboregulator that controls expression of genes involved in membrane transport and adhesion, LPS biosynthesis, amino acid metabolism, and nucleic acid modification. Similar to the observation that RepG binds directly to a G-repeat within the 5' UTR of the *tlpB*-HP0102 mRNA, homopolymeric G-stretches/G-rich sequences are proposed as putative sRNA binding sites in the additional target mRNAs. This suggests that the C/U-rich terminator loop of RepG is used for multi-target regulation by binding to G-rich sequences.

4.1. Microarray-based analysis of RepG-mediated gene expression changes

To identify additional target mRNAs of RepG (besides the *tlpB*-HP0102 operon), RepG-mediated gene expression changes were monitored between *H. pylori* 26695 wildtype (WT), *repG* deletion (Δ*repG*) and sRNA complementation (C$_{RepG}$) strains using whole-genome microarrays (Microarray Core Facility, Dr. Hans Mollenkopf, Max Planck Institute for Infection Biology, Berlin, Germany). These custom-designed microarrays cover open reading frames, intergenic regions as well as regulatory RNA elements such as non-coding RNAs and 5' UTRs. Among the about 1,600 *Helicobacter* open reading frames represented on the microarrays, about 30 transcripts were altered more than two-fold in the Δ*repG* mutant compared to the wildtype. Gene regulation by RepG was considered significant if altered transcript levels were detected in the Δ*repG* mutant compared to both WT and
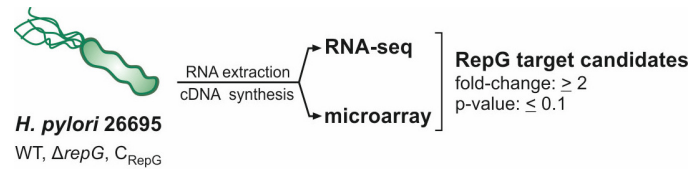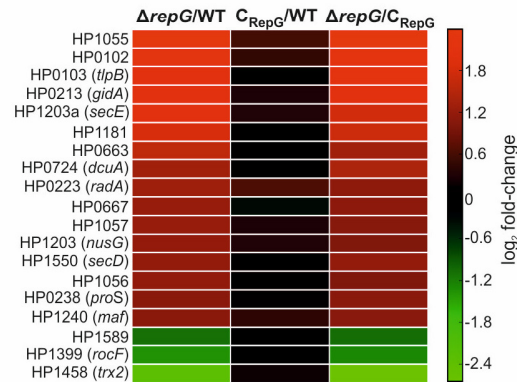
Figure 4.1: Identification of RepG targets using RNA-seq and microarray analysis. (*Upper panel*) Total RNA was isolated from *H. pylori* 26695 wildtype (WT), Δ*repG* and RepG complementation (C$_{RepG}$) strains grown to exponential growth phase (OD$_{600nm}$ of ~ 1.0; two biological replicates). Whole transcriptome analysis was performed by microarrays or RNA-seq. (*Lower panel*) Expression profiles of genes (horizontal bars) significantly regulated by RepG (fold-change ≥ 2, p-value ≤ 0.1) are shown in a heat-map. Green indicates down-regulation, black no regulation and red up-regulation upon *repG* deletion.

C$_{RepG}$ (≥ two-fold, p-value of ≤ 0.1), and if RNA abundances were similar between WT and sRNA complementation. Overall, RepG significantly affected the expression of 19 genes in *H. pylori* strain 26695 (Figure 4.1, Table 4.1). Consistent with previous data, the transcript abundance of the dicistronic *tlpB*-HP0102 mRNA was about four-fold (log$_2$ fold-change: 2) increased upon *repG* deletion. Please note that the fold-changes in *tlpB*-HP0102 mRNA abundance upon sRNA deletion were significantly stronger in qRT-PCR experiments (eight- to ten-fold, Figure 3.1 B) than in microarray analysis (four-fold, Figure 4.1 and Table 4.1). Such differences in fold-changes have been described in diverse transcriptome studies and might be caused by technical merits and drawbacks of each method (Git *et al.*, 2010).

Additional RepG target candidates encode for diverse membrane proteins including the amino acid transporter *dcuA* (HP0724), the multidrug-efflux pump HP1181, one component of a protein secretion system (protein translocase subunit *secD*) and still undefined outer membrane proteins (*e.g.* HP1057-HP1055). Furthermore, expression of genes involved in amino acid metabolism, such as the chorismate synthase (HP0663) and proline-tRNA ligase *proS*, as well as in DNA-repair (*radA*) and translation (*gidA*) are

controlled by RepG. Almost all of the 19 target candidates showed elevated transcript levels upon *repG* deletion, suggesting that RepG mostly acts as a repressor. The only transcripts positively affected by RepG were HP1589, *rocF* and *trx2*, which encode a hypothetical protein, the arginase RocF and the redox protein thioredoxin 2 (Trx2), respectively. While expression of HP1589 and *rocF* was only slightly (about two-fold) reduced, deletion of *repG* resulted in an about six-fold decrease in *trx2* mRNA abundance. RocF and Trx2 are components of anti-oxidant systems in *Helicobacter*, which are involved in detoxification of organic peroxides and other reactive oxygen species as well as nitrogen intermediates (reviewed in Wang *et al.*, 2006a). The reduction of such harmful radicals prevents DNA and protein damages. Accordingly, several components of the thioredoxin system have been shown to be important for colonization of the human host. Similar to *rocF* and *trx2*, other genes affected by RepG have also been associated with *Helicobacter* colonization defects in animal models or predicted to be potential drug candidates for antibiotic therapy (Baldwin *et al.*, 2007, Dairi, 2009). For example, the aspartate transporter *dcuA* is a component of a deaminase-transport system involved in amino acid uptake and has been shown to be crucial for the colonization in an animal model (Leduc *et al.*, 2010). Overall, this transcriptome study demonstrates that RepG regulates multiple target genes and might be involved in virulence control in *Helicobacter*.

4.2. RNA-seq reveals additional RepG target mRNAs

Nowadays, RNA-seq is considered to be the premier method for comprehensive as well as quantitative transcription profiling under various stress conditions or between mutant and wild-type strains. Compared to microarrays, RNA-seq is superior in detecting low abundant transcripts and in differentiating between biological isoforms as well as genetic variants (Creecy & Conway, 2015). In order to complement the number of RepG target mRNA candidates identified by microarrays, RNA-seq analysis was applied for whole transcriptome profiling of *H. pylori* 26695 wildtype, Δ*repG* and RepG complementation strains. To allow for an unbiased comparison between both methods (microarray vs. RNA-seq), the same RNA samples as mentioned previously (section 4.1.) were converted into cDNA libraries and subjected to high-throughput sequencing by Illumina using the Genome Analyzer IIx. Sequencing data were processed by the automated RNA-seq processing pipeline READemption (Forstner *et al.*, 2014). Different sequencing runs were found to be highly reproducible between biological replicates (Appendix, Figure 13.2 A). Genes that were differentially expressed and displayed significantly altered transcript abundances in the RNA-seq libraries (*e.g.* Δ*repG* vs. WT) were determined by DESeq (Appendix, Figure 13.3 A; Anders & Huber, 2010). Similar to the microarray analysis, a two-fold change

Table 4.1: RepG-mediated whole transcriptome changes in *H. pylori* strain 26695 analyzed by microarray and RNA-seq. Genes that showed more than two-fold change in transcript levels (p-value ≤ 0.1) in the Δ*repG* mutant (Δ) compared to the wildtype (WT) and RepG complementation strain (C_RepG/C) are listed in the table. Values represent fold-changes in gene expression levels in Δ*repG* vs. WT and Δ*repG* vs. C_RepG. Negative and positive values correspond to down- and up-regulation of the transcript, respectively. RepG target candidates identified by both RNA-seq and microarray are highlighted in gray.

ID	Name	Description	Microarray (Δ/WT)/(Δ/C)	RNA-seq (Δ/WT)/(Δ/C)	qRT-PCR (Δ/WT)
HP0199		uncharacterized protein		+6.16/+5.12	
HP1055		uncharacterized outer membrane protein	+5.12/5.23	+4.73/+4.21	+9.89
HP0663		chorismate synthase	+3.02/2.50	+4.72/+3.87	+7.79
HP0201	*plsX*	phosphate acyltransferase		+4.51/+4.20	
HP0223	*radA*	DNA repair protein	+2.44/+2.21	+4.50/+3.73	+5.57
HP1467		uncharacterized protein		+4.50/+4.11	
HP1181		multidrug-efflux transporter	+3.55/+3.33	+4.21/+3.49	+4.68
HP0308		uncharacterized protein		+4.11/+3.85	
HP0341		uncharacterized protein		+4.07/+2.40	
HP0580as		putative antisense transcript to HP0580		+3.74/+3.12	
HP1550	*secD*	protein translocase subunit SecD	+2.26/+2.27	+3.73/+3.37	+2.77
HP0238	*proS*	proline-tRNA ligase	+2.16/+2.18	+3.66/+3.11	
HP0303	*obgD*	GTPase		+3.62/+3.34	
HP0213	*gidA*	tRNA uridine 5-carboxymethyl-aminomethyl modification enzyme	+4.12/+3.94	+3.55/+3.26	
HP1398	*Ald*	alanine dehydrogenase		+3.54/+3.82	
HP1380	*tyrA*	prephenate dehydrogenase (TyrA)		+3.36/+2.99	
HP1023		uncharacterized protein		+3.34/+2.99	
HP0677		uncharacterized membrane protein	+2.34/+2.17	+3.22/+2.75	
HP0102		glycosyltransferase	+4.83/+4.47	+3.08/+2.52	+11.12
HP1056		uncharacterized outer membrane protein	+2.22/+2.01	+3.04/+2.48	+4.27
HP0777	*pyrH*	uridine monophosphate kinase		+2.87/+2.52	
HP0103	*tlpB*	methyl-accepting chemotaxis protein	+4.14/+4.50	+2.83/+2.52	+7.68
HP0259	*xseA*	exodeoxyribonuclease VII large subunit		+2.77/+2.06	+3.39
HP0724	*dcuA*	anaerobic C4-dicarboxylate transporter	+2.50/+2.66	+2.68/+2.48	+4.69
HP1288		putative uncharacterized protein		+2.65/+2.57	
HP0330	*ilvC*	ketol-acid reductoisomerase		+2.64/+2.61	
HP1577	*yaeE*	ABC-type transport system, permease		+2.62/+2.55	
HP0036		uncharacterized protein		+2.49/+2.11	
HP1131	*atpC*	ATP synthase epsilon chain		+2.49/+2.53	
HP1057		uncharacterized outer membrane protein	+2.33/+2.01	+2.47/+2.18	+3.07
HP0955	*lgt*	prolipoprotein diacylglyceryl transferase		+2.44/+2.06	
HP1468	*ilvE*	branched-chain-amino-acid aminotransferase		+2.41/+2.39	
HP0718		uncharacterized membrane protein		+2.34/+2.63	
HP0961	*gpsA*	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase		+2.28/+2.03	
HP0499		phospholipase A1		+2.26/+2.16	
HP0830	*gatA*	glutamyl-tRNA amidotransferase subunit A		+2.25/+2.62	
HP0580		sialidase, Kdo-hydrolase		+2.18/+2.01	+2.51
HP1323	*rnhB*	ribonuclease HII		+2.09/+2.20	+1.73
HP0919	*carB*	carbamoyl-phosphate synthetase ammonia chain		+2.03/+2.37	+3.96

ID	Name	Description	Microarray (Δ/WT)/(Δ/C)	RNA-seq (Δ/WT)/(Δ/C)	qRT-PCR (Δ/WT)
HP0616	*cheV2*	chemotaxis coupling protein		+2.01/+2.16	+2.30
HP1240	*maf*	Maf-like septum formation protein	+2.14/+2.10	+1.85/+2.10	
HP1203	*nusG*	transcription antitermination protein	+2.28/+2.03	+1.03/-1.03	
HP1203a	*secE*	protein translocase subunit	+4.10/+3.42	-1.01/-1.04	
HP1589		hypothetical protein	-2.12/-2.07	-1.23/+1.01	
HP1399	*rocF*	arginase	-2.56/-2.43	-2.00/-2.00	
HP0200as		putative antisense transcript to ribosomal protein L32		-2.12/-2.13	
HP0307		putative uncharacterized protein		-2.27/-2.13	
HPt29		tRNA		-2.56/-2.22	
HP1458	*trx2*	thioredoxin 2	-6.56/-6.25	-3.70/-2.78	-2.94
HPnc5490	*repG*	conserved small RNA RepG	-100/-100	-110/-100	

in transcript level in Δ*repG* vs. WT as well as Δ*repG* vs. C$_{RepG}$, and a p-value of ≤ 0.1 were used to identify RepG target candidates. Using this analysis, 45 genes were differentially expressed in the absence of the sRNA, with 40 being repressed and 5 being activated by RepG (Figures 4.2 A and B, Table 4.1). Almost all RepG target candidates determined by microarray analysis were also identified in the RNA-seq study (15 out of 19, Figures 4.2 A and B) and quantitative fold-changes greatly correlated between both methods (Figure 4.2 C). The remaining four RepG targets from the microarray approach (HP1203, HP1203a, HP1240 and HP1589) were either not affected by RepG deletion or did not pass the settings in the RNA-seq analysis (Table 4.1). In addition to the 15 genes identified by both methods, the RNA-seq approach led to the discovery of 30 additional RepG target candidates (Figures 4.2 A and B, Table 4.1).

To gain further insights into the biological function of RepG, a gene ontology analysis of the regulated genes was carried out by Dr. Konrad U. Förstner (University of Würzburg, Germany) using the PyloriGene database for *H. pylori* strain 26695 (http://genolist.pasteur.fr/PyloriGene/). Genes involved in amino acid, fatty acid and phospholipid metabolism belong to the most enriched functional categories of RepG target candidates (45 target mRNAs from RNA-seq as input, Figure 4.2 D). Furthermore, multiple identified RepG target mRNAs have been implicated in nucleotide synthesis and were associated with the cell envelope.

In agreement with the observation that RepG target genes such as HP0102 and *dcuA* are required for the colonization of the mouse model (Chapter 3, Leduc *et al.*, 2010), additional target mRNAs from the RNA-seq study have also been shown to be important for *H. pylori* virulence and adaption to the human host. For example, deletion of the chemotaxis receptor coupling protein *cheV2* (HP0616) has been associated with smooth swimming behavior and modest chemotaxis defects in *H. pylori* strain SS1 (Lowenthal *et al.*, 2009).
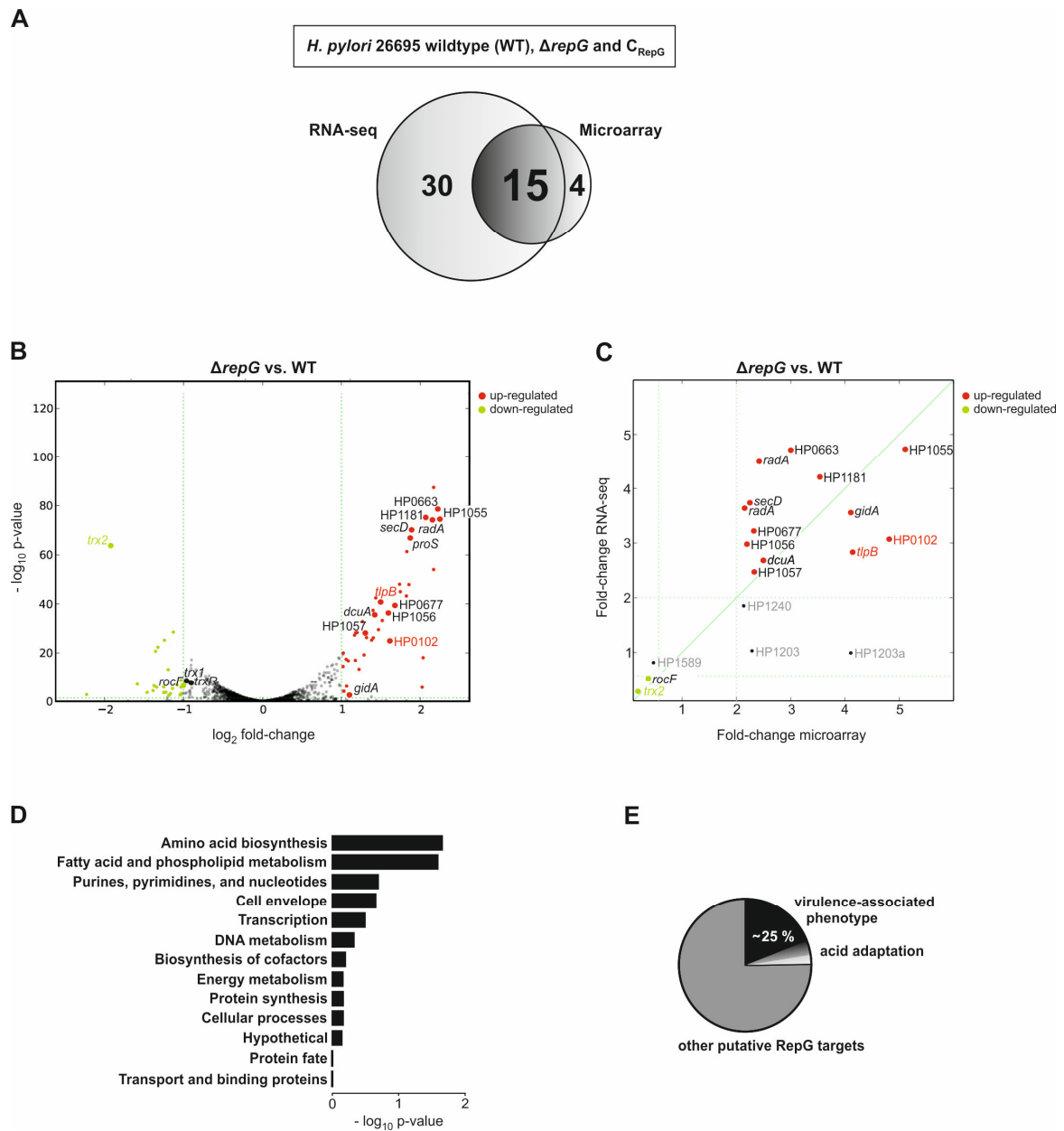
Figure 4.2: Additional RepG target transcripts identified by RNA-seq. (A) Results of the RNA-seq and microarray analyses of *H. pylori* 26695 WT, Δ*repG* and C$_{RepG}$. About 45 and 19 potential RepG target mRNAs were identified by RNA-seq and microarray analysis, respectively. **(B)** A volcano plot representation of the differentially expressed genes in *H. pylori* 26695 Δ*repG* vs. WT (RNA-seq). The y-axis represents negative log$_{10}$ of p-values (a higher value indicates greater significance) and the x-axis shows the difference in expression between Δ*repG* and WT (log$_2$ scale). The significance cut-off was set to a p-value of ≤ 0.1 and the biological cut-off was set to a fold-change of ± two-fold (-1 $\leq$ log fold-change $\geq$1). Regulated genes are labeled on the plot (RNA-seq). Big dots and gene ID indicate genes also identified by microarrays (Figure 4.1). **(C)** Correlation scatter plot of genes with changed transcript abundances identified by microarray and RNA-seq (reference: 19 genes from microarray analysis). RepG target candidates, which have been found to be only regulated in the microarray analysis, are shown in gray. **(D)** Gene ontology enrichment analysis of differentially expressed genes identified by RNA-seq (45 RepG targets as input). **(E)** Nearby one-fourth of the 45 putative RepG target mRNAs are associated with a virulence phenotype in the animal model or are important for *H. pylori* acid adaptation (Baldwin *et al.*, 2007, Dairi, 2009).

Moreover, the sialidase HP0580 together with HP0579 have been shown to display Kdo-hydrolase activity, which is important for lipid A modification and O-antigen expression (Stead *et al.*, 2010). Overall, about one-fourth of the 45 RepG targets have been shown to be regulated under acidic stress and/or were associated with *H. pylori* pathogenicity (Figure 4.2 E).

In summary, microarray and RNA sequencing analyses have been successfully applied for quantitative gene expression profiling and the identification of RepG-mediated changes at the transcript level in *H. pylori* strain 26695. These studies suggest that the RepG regulon might comprise more than 40 target mRNAs that are involved in membrane transport, adhesion, LPS biosynthesis, amino acid metabolism and nucleic acid modification.

## 4.3.   RepG preferentially binds to G-rich sequences within its target mRNAs

For manual and in-depth analysis of RNA-seq data, sequencing reads of each cDNA library (WT, Δ*repG* and C$_{RepG}$) were visualized as graphs representing the number of normalized mapped cDNA reads per nucleotide of the *H. pylori* 26695 genome in the Integrated Genome Browser (IGB, http://genoviz.sourceforge.net/; Figure 4.3). RepG represses expression of the *tlpB*-HP0102 operon in *H. pylori* strain 26695. Accordingly, a higher cDNA read coverage of the *tlpB*-HP0102 operon region was observed in the Δ*repG* mutant when compared to the wildtype or RepG complementation mutant (Figure 4.3 A, *lower panel on the left*). An opposite cDNA coverage was observed for the *trx2* gene, which is activated by RepG (Figure 4.3 A, *lower panel on the right*). Manual inspection of cDNA read coverage plots of the Δ*repG* mutant compared to the wildtype/C$_{RepG}$ indicated a specific enrichment pattern of cDNA reads covering homopolymeric G-repeats or G-rich sequences within RepG target mRNA candidates (Figures 4.3 A, 4.4 and 4.5). In general, such enrichment patterns of cDNA reads in high-throughput sequencing experiments are used for the identification of RNA-protein interaction sites using RIP-seq or CLIP-based techniques (Zhang *et al.*, 2015) and/or are observed upon enzymatic/chemical treatment of RNA samples prior to cDNA library construction (*e.g.* TEX or DMS treatment; Sharma *et al.*, 2010, Ding *et al.*, 2014). However, sequencing reads obtained from cDNA libraries of untreated RNA samples are usually distributed equally throughout genes, resulting in a more or less uniform coverage of the RNA. Albeit stable RNA structures can result in higher read coverage, *e.g.* at transcription terminator sites or within stable secondary structures, the enrichment of sequencing reads at G-rich sequences in the Δ*repG* library is unique and has not been observed in other RNA-seq based transcriptome studies of sRNA deletion mutants in *H. pylori* (*e.g.* ArsZ sRNA, P. Tan & Dr. C. M. Sharma, unpublished). Considering that the C/U-rich terminator loop of

Figure 4.3: RepG-mediated whole transcriptome changes analyzed by RNA-seq. (A) (*Upper panel*) cDNA sequencing reads of the wildtype (WT), *repG* deletion (Δ*repG*) and complementation (C$_{RepG}$) libraries were aligned to the genome of *H. pylori* strain 26695 and are visualized as coverage plots representing the number of reads (left scale) per nucleotide in the Integrated Genome Browser (IGB, Affymetrix). A representative screen shot of the IGB shows cDNA reads mapped along the whole *Helicobacter* genome on sense (leading) and reverse (lagging) strand. Genomic coordinates are indicated in the middle. RepG target candidates (15, identified by both microarray and RNA-seq) are boxed and labeled with gene ID. (*Lower panel*) Representative screen shots of the *tlpB*-HP0102 and *trx2* regions on the lagging strand of *H. pylori* 26695. Gray arrows represent the annotated ORFs and black ones denote published TSS (+1). **(B)** RepG was predicted to interact with G-rich sequences in the *tlpB* coding sequence. The *tlpB* interaction site in the terminator loop of RepG is highlighted in blue. Numbers indicate position in the *tlpB* coding region with respect to the annotated start codon.

RepG is highly conserved in *Helicobacter*, the G-rich sequences within RepG target mRNA candidates may represent RepG interaction sites. Bioinformatics-based predictions for RepG and target mRNA interaction sites (CopraRNA, Wright *et al.*, 2013 and RNAhybrid,

Rehmsmeier *et al.*, 2004) greatly overlapped with the enriched cDNA reads in the Δ*repG* mutant library (will be discussed for selected RepG targets below). For example, cDNA reads were accumulated at the homopolymeric G-repeat within the leader of the *tlpB* mRNA (Figure 4.3 A, *lower panel on the left*), which was previously shown to be the RepG target site. In addition, two other regions within the coding region of *tlpB* were found to be enriched for cDNA reads in the Δ*repG* sequencing library. These regions contain G-rich sequences and were predicted to interact with the C/U-rich terminator loop of RepG (Figure 4.3 B). However, whether these additional sRNA-binding sites might be involved in RepG-mediated control of *tlpB* and HP0102 await further investigation.

Similar to *tlpB*, a specific enrichment of cDNA reads was also observed for other RepG target mRNA candidates upon *repG* deletion (Figures 4.4 and 4.5). For instance, cDNA reads were enriched at a repetitive G-rich sequence in the 5' coding region of *dcuA*, which was also predicted to interact with the C/U-rich terminator loop of RepG (Figure 4.4 A). Comparable observations were also made for RepG target candidates that are encoded within putative operons such as HP1181-HP1182 and HP1057-HP1055. Albeit an internal TSS has been described in the HP1181 coding region (Sharma *et al.*, 2010), the multidrug-efflux pump HP1181 might be encoded in an operon together with the hypothetical protein HP1182. While expression of HP1181 is repressed by RepG, no significant change in the HP1182 transcript level was observed upon *repG* deletion. In line with this, enriched cDNA read counts and two putative RepG interaction sites were only identified in the coding region of HP1181, but not within HP1182 (Figure 4.4 B). In contrast to this putative discoordinate regulation of neighboring genes, all genes of the HP1057-HP1055 operon were co-regulated by RepG (Figure 4.4 C). The detection of multiple peaks within the HP1057-HP1055 operon indicates that RepG might repress expression of these outer membrane proteins by targeting each gene of the polycistronic mRNA at a repetitive G-rich sequence.

Specific enrichment patterns of cDNA reads were also observed for RepG target candidates that were only identified by RNA-seq, including the chemotaxis receptor coupling protein *cheV2* (HP0616) and carbamoyl-phosphate synthetase *carB* (HP0919) (Figure 4.5). In the latter case, RepG was predicted to bind to a homopolymeric G-repeat (9G) located within the *carB* coding region (Figure 4.5 A). Although an experimental validation of a direct RepG-*carB* interaction is still pending, this might represent an example of an intragenic SSR that is targeted by RepG. In line with the previously described optimal window of RepG-mediated repression (7 to 12Gs, Chapter 2), an increased *carB* mRNA level was detected upon *repG* deletion. The *carB* gene is encoded in an operon with two hypothetical proteins, HP0920 and HP0918. Besides targeting the *carB* coding region,
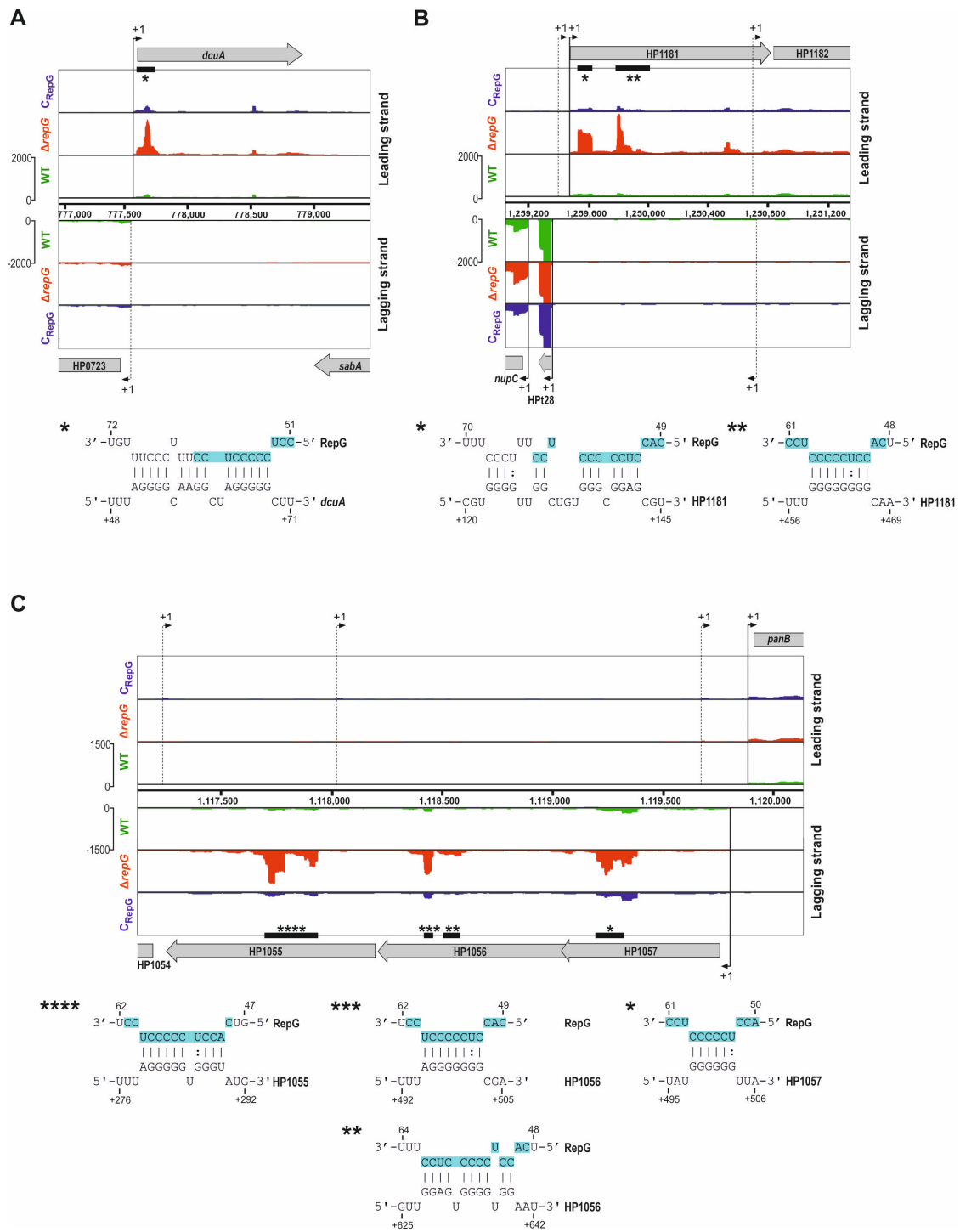
Figure 4.4: Peaks of cDNA reads in the Δ*repG* mutant greatly overlap with predicted RepG-target mRNA interaction sites. (A-C) cDNA reads from the RNA-seq analysis of the WT, Δ*repG* and RepG complementation (C$_{RepG}$) strains were mapped to the *dcuA* (HP0724, A), HP1181-HP1182 (B) and HP1057-HP1055 (C) regions in the *H. pylori* 26695 genome. Gray and black arrows represent the annotated ORFs and published TSS (+1; filled line – primary TSS, dotted line – secondary/antisense/internal TSS), respectively. Specific enrichment patterns for cDNA reads (peaks, thick black lines) in Δ*repG* were detected by a peak calling algorithm and greatly overlap with putative RepG-target mRNA interaction sites predicted by CopraRNA or RNAhybrid (denoted by */**/***/****, below IGB screenshot). Numbers indicate the position within the coding regions with respect to the annotated start codons.
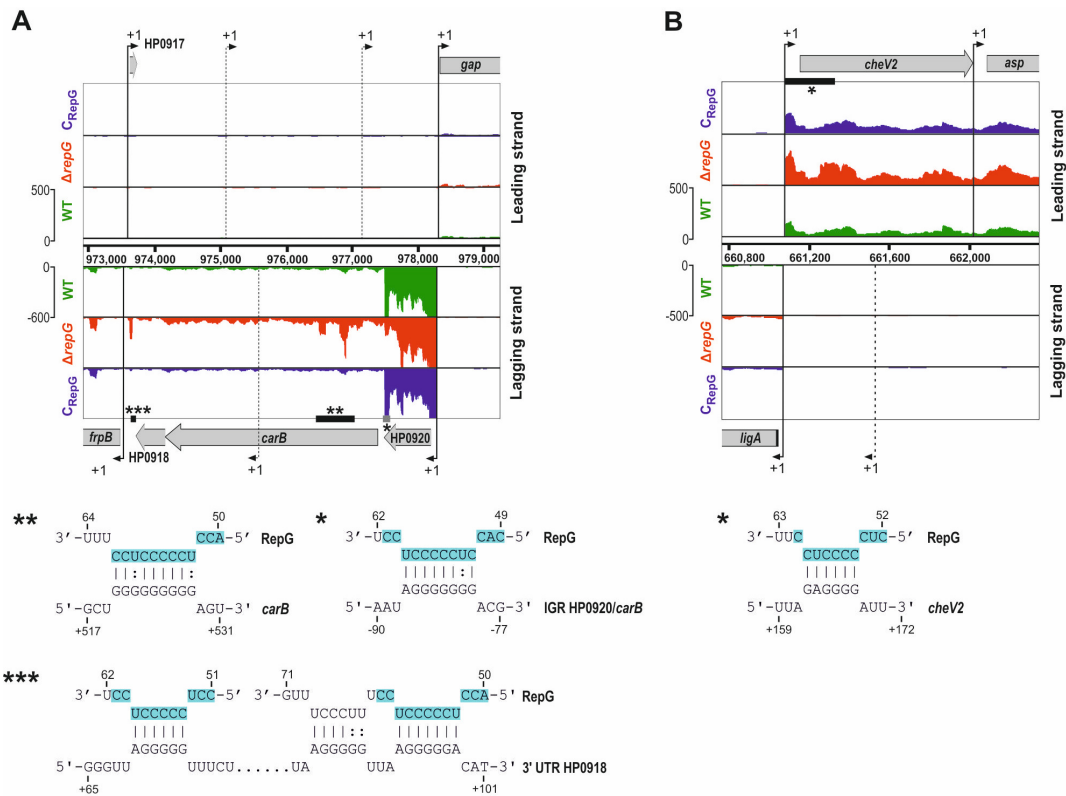
Figure 4.5: Accumulation of cDNA reads in the Δ*repG* mutant and predicted RepG-target mRNA interaction sites in the *carB* and *cheV2* mRNAs. (A-B) cDNA reads from the RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$ were mapped to the *carB* (HP0919, A) and *cheV2* (HP0616, B) regions in the *H. pylori* 26695 genome. Gray and black arrows represent the annotated ORFs and published TSS (+1; filled line – primary TSS, dotted line – secondary or antisense TSS), respectively. Specific enrichment patterns for cDNA reads in Δ*repG* (thick black lines) greatly overlap with putative RepG-target mRNA interaction sites (denoted by */**/***, below IGB screenshot). **(A)** Please note that cDNA reads in the IGR between HP0920 and *carB* accumulated in WT and C$_{RepG}$ sequencing libraries, but not in Δ*repG* (peak, thick gray line). A putative interaction between this region and RepG was also predicted by RNAhybrid (*). Numbers indicate position upstream or within the *carB* coding region with respect to the annotated start codon.

RepG was also predicted to interact with a G-rich sequence (a<u>GGGGGGG</u>) in the IGR between HP0920 and *carB*. In contrast to the cDNA peak within the *carB* coding region (9G-repeat), cDNA reads accumulated in the wildtype and C$_{RepG}$, but not in Δ*repG* mutant. Since HP0920 expression was not altered upon *repG* deletion, it is possible that this G-rich sequence may be also important for RepG-mediated regulation of carbamoyl-phosphate synthetase. Further studies will be required to investigate whether only one or both G-stretches are required for RepG-mediated *carB* expression control. In addition, yet another cDNA peak, and thus, potential RepG interaction site was detected in the putative 3' UTR of HP0918. The putative 3' UTR of HP0918 overlaps with the antisense encoded HP0917, encoding for a small peptide (23 aa). No significant changes in HP0918 or HP0917

mRNA abundances were detected in Δ*repG* compared to the WT or C$_{RepG}$ (p-values exceeded the cut-off, data not shown). Therefore, it is unclear whether RepG-dependent alterations in the abundance of the 3' UTR of HP0918 and/or a putative 3' UTR-derived sRNA are involved in gene expression control in any way. Nonetheless, this may represent an interesting example for further investigations as it suggests that RepG might control gene expression of *cis*-encoded RNAs (asRNAs or overlapping transcript regions). In line with this, expression control of the Kdo-hydrolase HP0580 and putative, *cis*-encoded HP0580as transcript through RepG will be discussed later on (see section 4.8).

The great overlap between the predicted RepG-target mRNA interaction sites and accumulations of cDNA reads in the Δ*repG* mutant suggested that this local enrichment of read counts could be used to define a RepG binding motif. Using peak calling, a transcriptome-wide screen for peaks of cDNA reads and thus, putative RepG interaction sites was performed (T. Bischler & Dr. C. M. Sharma, unpublished). Peak calling is a computational method which is used to identify areas in the genome that have been enriched for aligned sequencing reads and was originally developed to map RNA-protein interaction sites in RIP- and CLIP-seq studies (Konig *et al.*, 2011). In the here applied peak detection algorithm the genome of *H. pylori* 26695 was divided in user-defined genomic intervals, so-called windows, and the average expression values (calculated based on the number of mapped cDNA reads per window) of the Δ*repG* RNA-seq library were compared to the wildtype. Windows with a significant enrichment in cDNA reads in the Δ*repG* vs. WT were merged into peak regions reflecting potential RepG binding sites (for further details see Chapter 6, Material and methods). Predicted peak regions are available as annotations for manual inspection in the IGB (black boxes in Figures 4.3-4.5) or as sequence information table (Appendix, Table 13.5), and greatly overlap as well as define boundaries of cDNA enrichment patterns. Please note that multiple cDNA peaks and thus, putative RepG interaction sites were identified within one target mRNA, *e.g.* HP1181 and HP1056 (Figures 4.4 B and C). Overall, about 207 peaks were detected as significantly enriched in the RNA-seq data of *H. pylori* 26695 Δ*repG* mutant compared to the wildtype (Figure 4.6 A). At least one peak was detected within all RepG target candidates, which showed increased mRNA levels upon sRNA deletion. In contrast to that, RepG-mediated activation of gene expression did not correlate with peak detection of cDNA reads in the Δ*repG* mutant (*e.g. trx2*). Nearby one-third of the detected peaks (70 out of 207) were associated with RepG target mRNAs, whereas the other peaks (137) were assigned to genes that did not pass the chosen settings of the RNA-seq study. This indicates that additional, so far unknown RepG targets that show only slight changes in mRNA expression levels might have been missed in the RNA-seq study due to a setting of a two-fold change threshold.

A

B
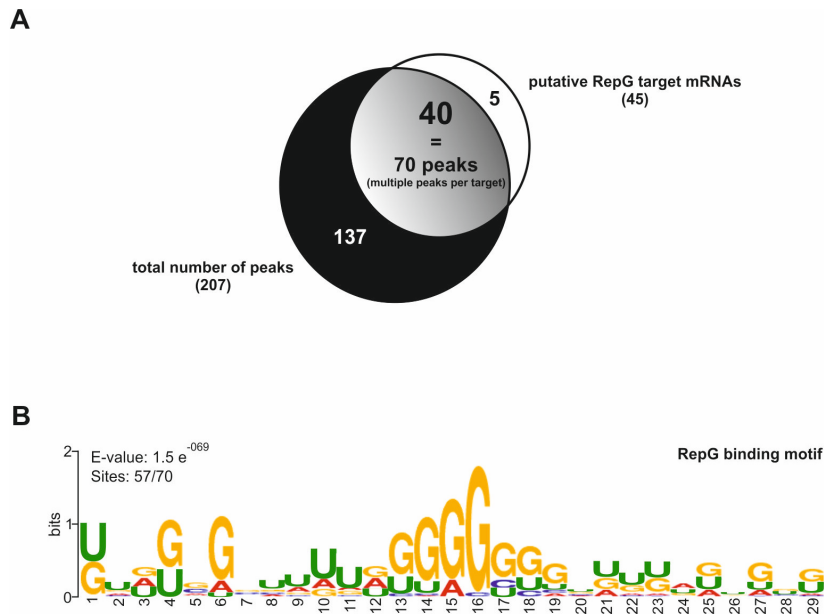
Figure 4.6: RepG preferentially interacts with G-rich sequences in its target mRNAs. (A) About one-third of the 207 predicted peaks of cDNA reads in the Δ*repG* mutant are located within RepG target candidates (70 peaks in 40 target mRNAs). **(B)** A consensus motif for the G-rich RepG target sites was determined by MEME (Multiple Em for Motif Elicitation, http://meme.nbcr.net/meme/cgi-bin/meme.cgi; Bailey *et al.*, 2009).

In order to predict a RepG binding motif, only peak sequences, which had been detected within RepG target mRNAs (70; Appendix, Table 13.5), were used as input for MEME motif identification (Bailey *et al.*, 2009) using the following parameters: single motif distribution in sequence: zero or one, number of different motifs: 10, minimum motif width: 6, and maximum motif width: 50. The upper size restriction of 50 nt was selected to cover the size range of the RepG terminator loop including the C/U-rich *tlpB* interaction site. This allows the identification of up to ten motifs with a length between 6 and 50 nucleotides within the input sequence. The best RepG binding motif identified by MEME is shown in Figure 4.6 B. It is 29 nt in length and present in 81 % of the input sequences (57 out of 70). The majority of the RepG target candidates share a common motif composed of a homopolymeric G-repeat and/or G-rich sequences (flanking), which could be targeted by the C/U-rich terminator loop of RepG and thereby mediate regulation of these targets.

Using FIMO motif search tool (Find Individual Motif Occurrence, MEME suite; Grant *et al.*, 2011), the occurrence of the best RepG binding motif (Figure 4.6 B) was determined in the *H. pylori* 26695 genome. About 2767 significant motif occurrences (p-value of less than 10^{-5}) were detected, indicating a similar G-rich motif to be present in or close to about 920 open reading frames in *H. pylori* strain 26695. This exceeds by far the number of cDNA peaks enriched in the Δ*repG* mutant and raises the question how specificity of RepG base-pairing to G-rich sequences is facilitated in *H. pylori*.

4.4. Common and strain-specific RepG target mRNAs

Although RepG is highly conserved among diverse *H. pylori* strains, a strain-specific *tlpB* regulation was observed (Figure 2.8). While RepG mediates repression of the chemotaxis receptor in *H. pylori* strain 26695, it activates TlpB protein expression in strain G27. This suggests that RepG might control distinct sets of target genes in different *H. pylori* strain backgrounds. To investigate common and strain-specific transcriptome changes caused by *repG* deletion, an RNA-seq study was performed with *H. pylori* G27 wildtype (WT), *repG* deletion mutant (Δ*repG*) and its complementation (C$_{RepG}$) grown to exponential growth phase (biological replicates; Appendix, Figure 13.4). In general, cDNA library construction and sequencing as well as RNA-seq data processing and analysis were performed as described in section 4.2. (see also Chapter 6, Material and methods). Overall, 54 transcripts showed significantly changed transcript levels upon *repG* deletion ($\geq$ two-fold change and p-value of $\leq$ 0.1), with 52 being repressed and only two being activated by RepG (Figure 4.7 A, Table 4.2). In latter cases, RNA-seq revealed an about two- and four-fold increase in the mRNA levels of G27_1151 (ribosomal protein L33, *rpmD*) and HPG27_82 (5-methyl-
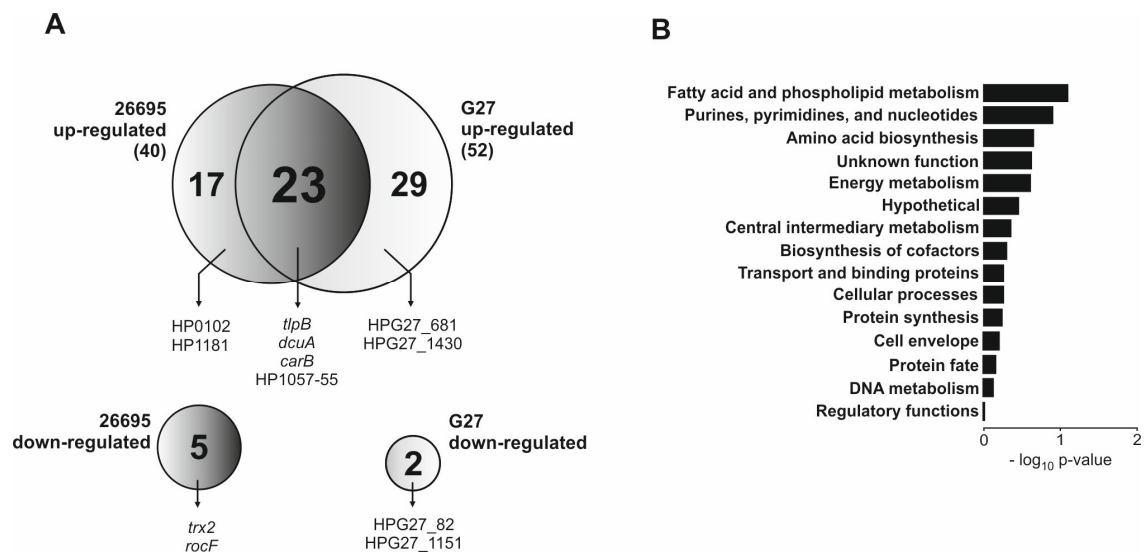
Figure 4.7: RNA-seq revealed putative candidates for common and strain-specific RepG target mRNAs in *H. pylori* strains 26695 and G27. (A) Circle diagram showing the overlap between RepG-mediated gene expression changes in *H. pylori* strains 26995 and G27. More than half of the genes that were up-regulated upon *repG* deletion in 26695 and G27 are controlled in both strains, *e.g. tlpB*, *dcuA,* HP1057-HP1055 and *carB*. RepG represses expression of 17 and 29 genes in in each of the two strains, respectively. Only 5 and 2 genes are activated by RepG in a strain-specific manner. None of the RepG target mRNAs are invers regulated between both strains. **(B)** Gene ontology enrichment analysis of differentially expressed genes in *H. pylori* G27 Δ*repG* (vs. WT) identified by RNA-seq (54 targets as input).

thioadenosine/S-adenosylhomocysteine nucelosidase), respectively. Similar to *H. pylori* strain 26695, genes involved in amino acid, fatty acid and phospholipid metabolism belong to the most enriched functional categories of RepG target candidates in G27 (Figure 4.7 B). In addition, comparable enrichment patterns of cDNA reads covering G-rich sequences were observed for RepG target mRNAs in G27 and 26695 (see below; Figures 4.8, 4.10 and Appendix, Figures 13.6-13.8).

A Best-Reciprocal-Hit clustering algorithm based on all-against-all nucleotide sequence comparison was employed to define orthologous genes of *H. pylori* strains 26695 and G27 (Dr. Konrad U. Förstner; University of Würzburg, Germany). Based on this, significantly up- and down-regulated RepG target candidates were compared between both strains (Figure 4.7 A), and RepG-mediated changes in gene expression were validated for chosen targets by quantitative RT-PCR (see below, Figure 4.9; Tables 4.1 and 4.2). No contradictory regulation of RepG target candidates, *i.e.* activation in one and repression in the other strain, was observed when comparing the sRNA regulons of *H. pylori* strains 26695 and G27 (Figure 4.7 A). Although orthologs have been annotated in both strains, a strain-specific activation of gene expression through RepG was observed, *e.g.* for *trx2* (will be discussed below, see section 4.7. and Appendix, Figure 13.5). Albeit strain-specific up-regulation of 17 and 29 target mRNAs in the Δ*repG* mutants of 26695 and G27, respectively, RepG repressed expression of more than 20 target genes independent of the strain background. Among them, expression of the amino acid transporter *dcuA* was about three- to four-fold enhanced in the absence of *repG* (Figure 4.8 A, Tables 4.1 and 4.2). In line with this, the putative G-rich RepG binding site within the 5' coding sequence of *dcuA* is greatly conserved among 26695 and G27 (Figure 4.8 B). Also, *carB* mRNA levels were up-regulated in the Δ*repG* mutants of both strains (Appendix, Figures 13.6 A-B; Tables 4.1 and 4.2). Although the enrichment patterns of cDNA reads in Δ*repG* greatly correlate in *H. pylori* strain 26695 and G27 (one peak at the 9G-repeat in *carB* coding region and another covering the G-rich stretch in HP0920-*carB* IGR), predicted RepG-binding sites are not completely conserved. In particular, the G-rich sequence in the HP0920-*carB* IGR contains an AG insertion (AGGGGGGG (26695) → AGGG<u>AG</u>GGGG (G27)) and a single G to A conversion is found within the homopolymeric G-repeat in the *carB* coding region. This might suggest that these mutations are dispensable for RepG targeting and/or that flanking G-rich sequences might contribute RepG-*carB* mRNA interaction. However, this awaits further investigations. Similar to *dcuA* and *carB*, elevated transcript levels were also detected for the outer membrane proteins HP1057-HP1055 independent of the strain background, albeit different fold-changes were observed for the last gene of the operon, HP1055 (Appendix, Figure 13.7; Tables 4.1 and 4.2). The putative RepG interaction sites are largely conserved in the HP1057-1055 operon.

Table 4.2: RepG-mediated whole transcriptome changes in *H. pylori* strain G27 analyzed by RNA-seq. Genes that showed more than two-fold change in transcript levels (p-value ≤ 0.1) in the Δ*repG* mutant (Δ) compared to the wildtype (WT) and RepG complementation strain (C$_{RepG}$/C) are listed in the table. Values represent fold-changes in gene expression levels in Δ*repG* vs. WT and Δ*repG* vs. C$_{RepG}$. Negative and positive values correspond to down- and up-regulation of the transcript. Genes that were significantly regulated upon *repG* deletion in *H. pylori* strains 26695 and G27 are highlighted in gray.

ID	Ortholog (26695)	Name	Description	RNA-seq (Δ/WT)/(Δ/C)	qRT-PCR (Δ/WT)
HPG27_636	HP0677		uncharacterized protein	+8.14/+8.80	
HPG27_184	HP0201	*plsX*	putative glycerol-3-phosphate acyltransferase	+6.48/+7.66	
HPG27_1038	HP0357		short chain alcohol dehydrogenase	+6.18/+7.31	
HPG27_474	HP0515		ATP-dependent protease peptidase subunit	+5.09/+6.15	
HPG27_195	HP0213	*gidA*	tRNA uridine 5-carboxymethylaminomethyl modification protein	+4.57/+4.75	
HPG27_282	HP0303	*obgE*	GTPase	+4.43/+5.13	
HPG27_238	HP0259	*xseA*	exodeoxyribonuclease VII large subunit	+4.31/+4.45	+2.93
HPG27_312	HP0330	*ilvC*	ketol-acid reductoisomerase	+4.13/+4.59	
HPG27_1106	HP1162		hypothetical protein	+4.05/+4.76	
HPG27_625	HP0663		chorismate synthase	+3.88/+4.23	+4.00
HPG27_1390	HP1467		outer membrane protein	+3.81/+3.81	
HPG27_204	HP0223	*radA*	DNA repair protein	+3.74/+4.08	+7.06
HPG27_185	HP0202	*fabH*	3-oxoacyl-(acyl carrier protein) synthase III	+3.66/+4.35	
HPG27_1076	HP1131	*atpC*	F$_0$F$_1$ ATP synthase subunit epsilon	+3.52/+4.15	
HPG27_1341	HP1418	*murB*	UDP-N-acetylenolpyruvoylglucosamine reductase	+3.46/+4.04	
HPG27_678	HP0724	*dcuA*	anaerobic C4-dicarboxylate transporter	+3.30/+3.76	+4.01
HPG27_734	HP0777	*pyrH*	uridine monophosphate kinase	+3.09/+3.78	
HPG27_95	HP0103	*tlpB*	methyl-accepting chemotaxis protein	+3.00/+3.15	+3.52
HPG27_257	HP0278		guanosine pentaphosphate phosphohydrolase	+2.91/+3.01	
HPG27_557	HP0597		penicillin-binding protein 1A	+2.77/+2.95	
HPG27_264	HP0285		hypothetical protein	+2.77/+2.94	
HPG27_183	HP0199		uncharacterized protein	+2.73/+3.92	
HPG27_1088	HP1143		hypothetical protein	+2.71/+3.65	
HPG27_788	HP0829		inosine 5-monophosphate dehydrogenase	+2.70/+3.31	
HPG27_372	HP1056		uncharacterized outer membrane protein	+2.67/+2.92	+4.05
HPG27_903	HP0955	*igt*	prolipoprotein diacylglyceryl transferase	+2.66/+3.04	
HPG27_218	HP0238	*proS*	prolyl-tRNA synthetase	+2.65/+4.45	
HPG27_868	HP0919	*carB*	carbamoyl-phosphate synthase large subunit	+2.65/+2.95	+2.59
HPG27_1002	HP0395		hypothetical protein	+2.65/+2.64	
HPG27_1488	HP1550	*secD*	preprotein translocase subunit SecD	+2.58/+3.00	+1.50
HPG27_1476	HP1538	*fbcH*	ubiquinol cytochrome c oxidoreductase	+2.55/+2.78	
HPG27_528	HP0569	*engD*	GTP-dependent nucleic acid-binding protein	+2.46/+2.81	
HPG27_681	HP0726		outer membrane protein	+2.41/+2.66	
HPG27_1061	HP1117		cysteine-rich protein X	+2.34/+2.72	
HPG27_1461	HP1398	*ald*	alanine dehydrogenase	+2.34/+2.30	
HPG27_33	HP0036		uncharacterized protein	+2.31/+2.40	
HPG27_287	HP0308		uncharacterized protein	+2.30/+2.71	
HPG27_1267	HP1318	*rplD*	50S ribosomal protein L4	+2.30/+2.63	
HPG27_1283	HP1335	*mnmA*	tRNA-specific 2-thiouridylase	+2.24/+2.31	
HPG27_1272	HP1323	*rnhB*	ribonuclease HII	+2.22/+2.78	+1.66
HPG27_317			cysteine-rich protein C	+2.22/+2.41	
HPG27_132	HP0145	*fixO*	cbb3-type cytochrome c oxidase subunit II	+2.17/+2.61	
HPG27_1185	HP1241		alanyl-tRNA synthetase	+2.15/+2.54	

ID	Ortholog (26695)	Name	Description	RNA-seq (Δ/WT)/(Δ/C)	qRT-PCR (Δ/WT)
HPG27_1430	HP1506		sodium/glutamate symport carrier protein	+2.13/+2.29	
HPG27_52	HP0057		uncharacterized protein	+2.09/+2.81	
HPG27_467	HP0509	glcD	glycolate oxidase subunit	+2.07/+2.34	
HPG27_1027	HP0370		biotin carboxylase	+2.06/+2.07	
HPG27_53			uncharacterized protein	+2.037+2.18	
HPG27_1391	HP1468	ilvE	branched-chain-amino-acid aminotransferase	+2.03/+2.29	
HPG27_217	HP0237		porphobilinogen deaminase	+2.00/+2.23	
HPG27_1524			uncharacterized protein	+2.00/+2.11	
HPG27_465	HP0507		uncharacterized protein	+2.00/+2.00	
HPG27_1151	HP1204	rpmD	50S ribosomal protein L33	-2.00/-2.00	
HPG27_82	HP0089		5-methylthioadenosine/S-adenosylhomocysteine nucleosidase	-4.26/-3.84	
HPnc5490	HPnc5490	repG	conserved small RNA RepG	-100	

Contradictory to a reduced TlpB protein level (Figure 2.8), the *tlpB* mRNA level was about three-fold increased upon RepG deletion in *H. pylori* strain G27 (Table 4.2). This is consistent with previous observations (Figures 2.11 and 3.1 B). In comparison to the relative fold-changes of *tlpB*-HP102 expression in 26695, the mRNA levels of the chemotaxis receptor and glycosyltransferase were only mildly affected by *repG* deletion in G27 (Table 4.1 and 4.2). While *tlpB* is considered a common RepG target gene in both strains (RNA-seq, fold-change ≥ 2), HP0102 mRNA levels were only significantly affected in 26695 (Figure 4.7 A). RNA-seq analysis and peak detection revealed a similar pattern of cDNA reads at conserved G-rich sequences in the *tlpB* coding region in the Δ*repG* mutant of both strains (compared to the wild-type background; Appendix, Figure 13.8). However, RepG-dependent enrichment of cDNA reads at the variable homopolymeric G-repeat in the 5' UTR of the *tlpB* mRNA was only observed in *H. pylori* strain 26695, but not in G27.
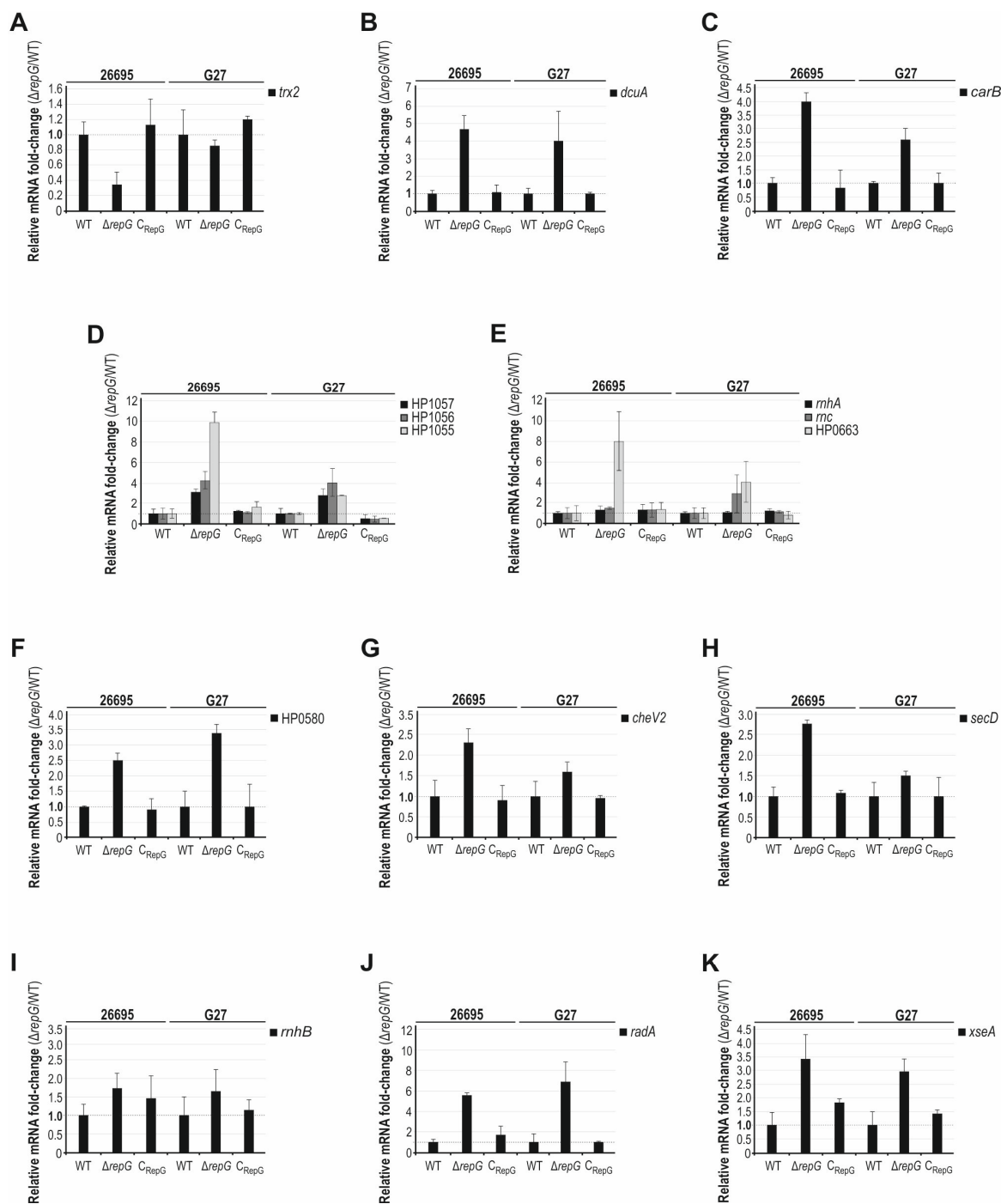
4.5. Validation of selected RepG target mRNAs

Using quantitative RT-PCR, selected RepG target mRNA candidates identified by microarray and RNA-seq were validated/confirmed in *H. pylori* strains 26695 and G27 (Figure 4.9). Despite slight variations in the relative mRNA fold-changes (Δ*repG* vs. WT), microarray, RNA-seq and qRT-PCR data greatly correlates with each other (Figure 4.9; Tables 4.1 and 4.2). Although the trend of regulation was almost always similar between both strains, qRT-PCR analysis revealed strain-specific mRNA fold-changes upon *repG* deletion. For example, while mRNA levels of HP1055 and HP0663 (each gene the last encoded in an operon) were about eight to ten-fold enriched in *H. pylori* 26695, deletion of *repG* resulted only in three- to four-fold increased transcript levels in strain G27 (Figures 4.9 D and E). This is a bit contradictory to the RNA-seq data, in which three- to four-fold increased HP1055 and
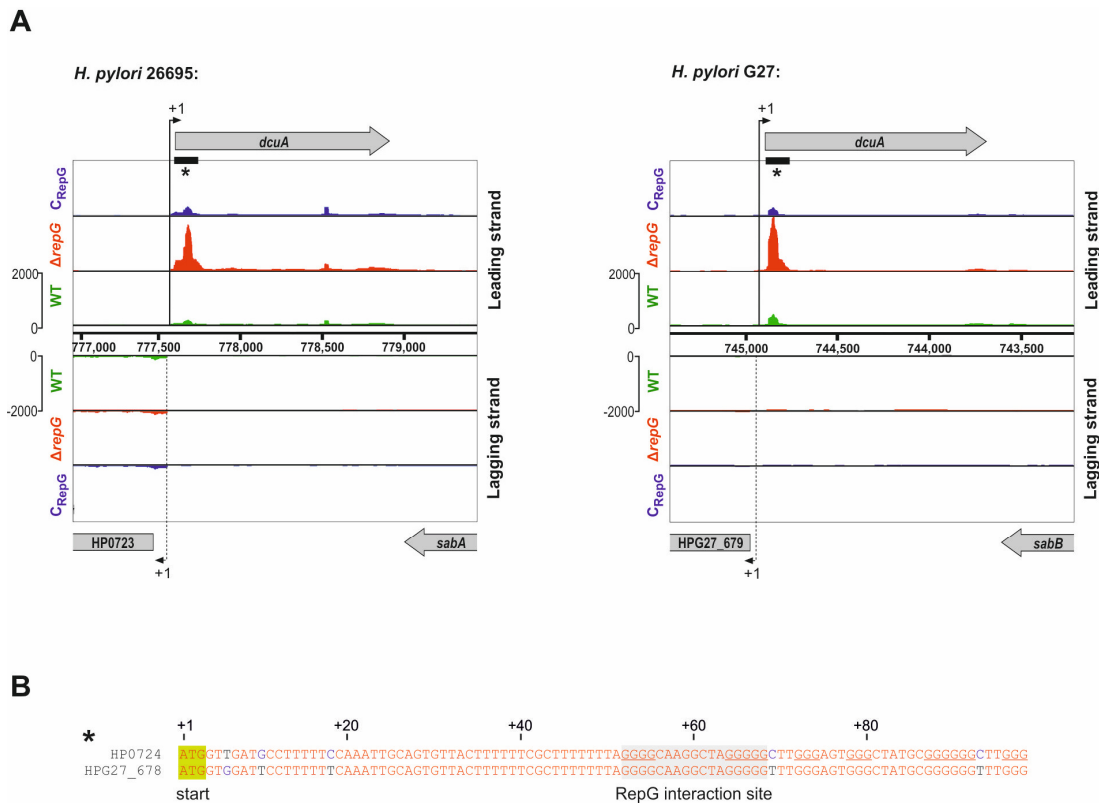
Figure 4.8: RepG represses expression of *dcuA* in *H. pylori* strains 26695 and G27. (A) IGB screenshot representing cDNA reads from the RNA-seq analysis of WT, Δ*repG* and complementation (C_RepG) mutants mapped to the *dcuA* region in the chromosome of *H. pylori* strains 26695 (*left panel*) and G27 (*right panel*). Gray and black arrows represent the annotated ORFs and published TSS (+1; filled line – primary TSS, dotted line – putative TSS), respectively. Specific enrichment patterns for cDNA reads (peaks, thick black lines) in Δ*repG* were detected greatly overlap with a putative RepG-target mRNA interaction sites predicted (denoted by *). **(B)** Sequence alignment of the 5' coding region of *dcuA* covering the predicted RepG binding site (gray) of *H. pylori* strains 26695 and G27. Numbers indicate positions in the coding region with respect to the annotated start codon of 26695 (ATG, green). Repetitive G-rich sequences (in RepG interaction site and flanking region) are underlined.

HP0663 mRNA abundances were observed upon *repG* deletion in both strain backgrounds (Tables 4.1 and 4.2). According to the RNA-seq data, qRT-PCR revealed that RepG significantly affects *cheV2* and *trx2* mRNA levels in 26695, but not in *H. pylori* strain G27 (1.5-fold, failed to pass the cut-off of two-fold change in mRNA abundance; Figures 4.9 A and G). Further studies, including whole proteome analysis, will be required to investigate whether the here identified strain-specific target mRNA candidates (*e.g. trx2*, *cheV2* and HP1181, see also sections 4.6 and 4.7) are indeed subjected to strain-specific posttranscriptional gene expression control by RepG.

Figure 4.9: Validation of selected RepG target mRNA candidates using qRT-PCR. DNase I-treated RNA samples isolated from WT, Δ*repG* and complementation (C$_{RepG}$) mutants of each strain background (26695, G27) were used in one-step qRT-PCR reactions. Wild-type mRNA levels were set to 1, and relative mRNA fold-changes in mutant strains compared to the WT are shown as bars. Values are shown as mean ± standard derivations from two/three independent experiments. *trx2* – thioredoxin 2, *dcuA* – anaerobic C4-dicarboxylate transporter, *carB* – carbamoyl-phosphate synthetase, HP1057-HP1055 – outer membrane proteins, *rnhA* – ribonuclease HI, *rnc* – ribonuclease III, HP0663 – chorismate synthase, HP0580 – Kdo-hydrolase, *cheV2* – chemotaxis coupling protein, *secD* – protein translocase subunit D, *rnhB* – ribonuclease HII, *radA* – DNA repair protein, *xseA* – exodeoxyribonuclease VII subunit.

In summary, comparative RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$ in *H. pylori* strains 26695 and G27 revealed putative candidates for common and strain-specific RepG targets. Based on RNA-seq analysis, about half of the RepG target mRNAs identified in 26695 and G27 are similarly regulated independent of the strain background. In both strains, RepG acts mainly as a repressor of genes involved in amino as well as fatty acid metabolism, and nucleotide synthesis. Furthermore, mRNAs encoding for components of the cell envelope (transport) are commonly affected by RepG, suggesting that this sRNA might be important for the definition of *H. pylori* surface properties.

4.6. RepG represses HP1181 expression at the transcript, but not significantly at the protein level

In order to investigate whether RepG (also) mediates repression of common or strain-specific target genes at the protein level, multiple candidates such as the membrane-associated proteins *dcuA*, HP1057-HP1055 and HP1181 were chosen for chromosomal FLAG-tagging. Although sequencing confirmed correct insertion of chromosomal FLAG-tags to respective genes of interest, various FLAG-tagged proteins were found to be not or not stably expressed under the examined conditions (data not shown). In particular, no significant signals could be detected by western blot analysis for DcuA::3xFLAG, HP1055::3xFLAG, HP1056::3xFLAG and HP1057::3xFLAG in the wildtype and *repG* deletion mutants of *H. pylori* strains 26695 and G27. This suggests that either these genes are not translated under standard growth conditions or that FLAG-tagging resulted in delocalization of these membrane-associated proteins, which is often then coupled to protein degradation. From the tested membrane proteins, only HP1181::3xFLAG was stably expressed and thus, used for further investigations.

The multidrug-efflux pump HP1181 is considered a strain-specific RepG target candidate, as mRNA levels were affected by *repG* deletion in 26695, but not G27 (RNA-seq and qRT-PCR, Figures 4.7 A and 4.10 A). Although the downstream encoded gene HP1182 was not significantly regulated in the RNA-seq study, qRT-PCR analysis showed slightly increased HP1182 mRNA level upon RepG deletion in *H. pylori* strain 26695 (Figure 4.10 A). In contrast, mRNA levels of neither HP1181 nor HP1182 were altered in *H. pylori* G27 Δ*repG* compared to the wildtype (lower than two-fold). In *H. pylori* strain 26695, two potential RepG interaction sites were predicted in the coding region of HP1181 (Figure 4.4 B). However, only one of them was found to be conserved strain G27 (Figures 4.10 B and C). Assuming that both G-rich sequences are potentially required for sRNA-binding and RepG-mediated repression, the disruption of one of these sites in G27 might explain strain-specific fold-changes in HP1181 mRNA abundance upon *repG* deletion.
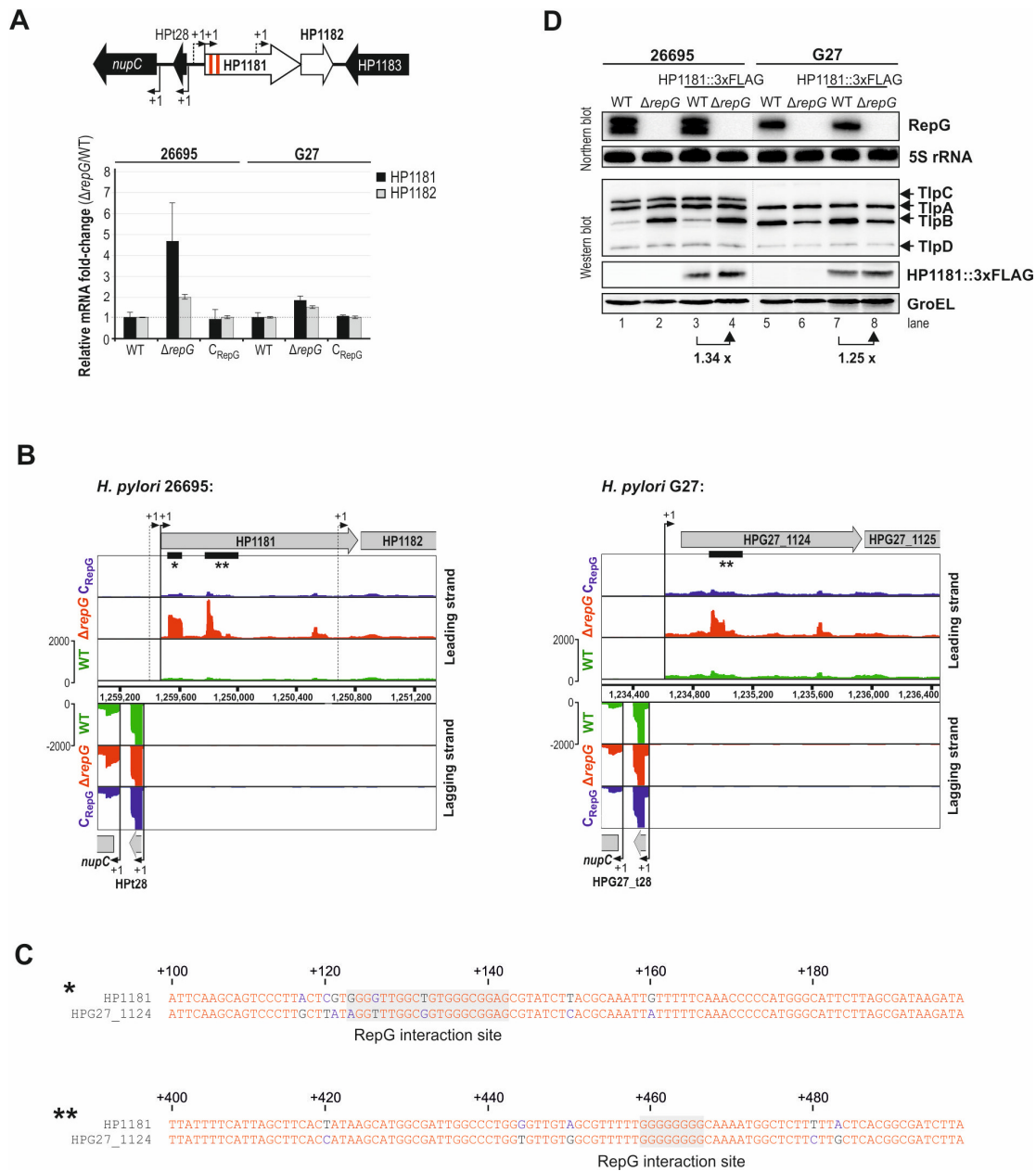
Figure 4.10: RepG represses expression of HP1181 at the transcript, but not significantly at the protein level. (A) (*Upper panel*) Genomic location of the HP1181-HP1182 operon. TSS (+1) are indicated by black arrows, and predicted RepG interaction sites are shown in red. (*Lower panel*) qRT-PCR analysis of RepG-dependent expression of the HP1181-HP1182 operon in WT, Δ*repG* and C$_{RepG}$ of *H. pylori* strains 26695 and G27. Wild-type mRNA levels were set to 1, and relative fold-changes (mutant vs. WT) are shown as bars (based on two biological replicates). **(B)** IGB screenshot of cDNA reads from the RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$ mapped to the HP1181-HP1182 region in strains 26695 (*left panel*) and G27 (*right panel*). **(C)** Sequence alignment of the two predicted RepG binding sites (gray) in HP1181 in strains 26695 and G27. Numbers indicate positions in the coding region with respect to the annotated start codon. **(D)** RNA and protein samples from *H. pylori* strains 26695 and G27 WT and Δ*repG*, as well as WT and Δ*repG* strains carrying HP1181::3xFLAG gene, were harvested at exponential growth phase (OD$_{600nm}$ of ~ 0.8) and analyzed by western and northern blot, respectively. HP1181::3xFLAG was detected with anti-FLAG antibody; GroEL served as loading control. RepG was probed with 5' end-labeled CSO-0003 and 5S rRNA with JVO-0485, respectively. The chemotaxis receptors were detected by TlpA-22 antiserum.

To test whether RepG also affects HP1181 protein levels, expression of HP1181::3xFLAG was examined in the wildtype and the Δ*repG* mutant of *H. pylori* strains 26695 and G27 at exponential growth phase. In both strain backgrounds, HP1181::3xFLAG protein levels were only mildly (1.3-fold) affected upon *repG* deletion (Figure 4.10 D). As transcript levels of the HP1181::3xFLAG mRNA have not yet been investigated in the wildtype and Δ*repG* mutants, a possible polar effect of the FLAG-tag on RepG-mediated regulation cannot be excluded. However, it might also indicate that not all RepG targets are similarly regulated at the transcript and protein level.

4.7. The conserved C/U-rich terminator loop of RepG activates expression of thioredoxin 2 (Trx2)

H. pylori possesses two thioredoxin proteins, Trx1 and Trx2, and a thioredoxin reductase, TrxR. The redox proteins Trx1/2 and the associated enzyme TrxR constitute a thiol-dependent reduction-oxidation system that catalyses the reduction of specific proteins, *e.g.* various ROS detoxification enzymes, by NADPH (reviewed in Wang *et al.*, 2006a). Reducing equivalents enter from NADPH to TrxR, which reduces Trx1 and Trx2. The reduced dithiol form of Trx1/2 then serves as an intermediate source of reducing power for numerous enzymatic reactions in the cell, including the reduction of peroxiredoxins such as the alkyl hydroperoxide reductase (AhpC), thiol peroxidase (Tpx) and bacterioferritin co-migratory protein (Bcp). Trx1 and TrxR are encoded in an operon and located far upstream of the *trx2* locus in nearby all sequenced *H. pylori* genomes. Expression of the dicistronic *trx1-trxR* mRNA was not affected by deletion of *repG* in *H. pylori* strains 26695 and G27 (data not shown). In contrast, qRT-PCR and northern blot revealed about three- to four-fold decreased *trx2* mRNA levels concomitant with *repG* deletion in strain 26695, but not G27 (Figures 4.9 A). In agreement with the RNA-seq results (Appendix, Figure 13.5 A-B), these data suggest that RepG might promote *trx2* expression in a strain-specific manner (at least at the transcript level).

Bioinformatics-based predictions for RNA-RNA interactions indicated that RepG might base-pair with its C/U-rich terminator loop to three G̲-rich s̲equences (GRS I-III) within the *trx2* mRNA of *H. pylori* strain 26695 (Figure 4.11 A). According to the previously identified RepG binding motif (Figure 4.6 B), all putative RepG-*trx2* interaction sites comprise a 5-nt long homopolymeric G-repeat (gt G̲G̲G̲G̲G̲a). One predicted RepG interaction site (GRS I) is located about 10 nt upstream of the RBS in the 5' UTR of the *trx2* mRNA, whereas the other two (GRS II and III) are found within the coding region. While the G̲G̲G̲G̲G̲a-motif in the *trx2* mRNA leader is 100 % conserved among *H. pylori* strains 26695
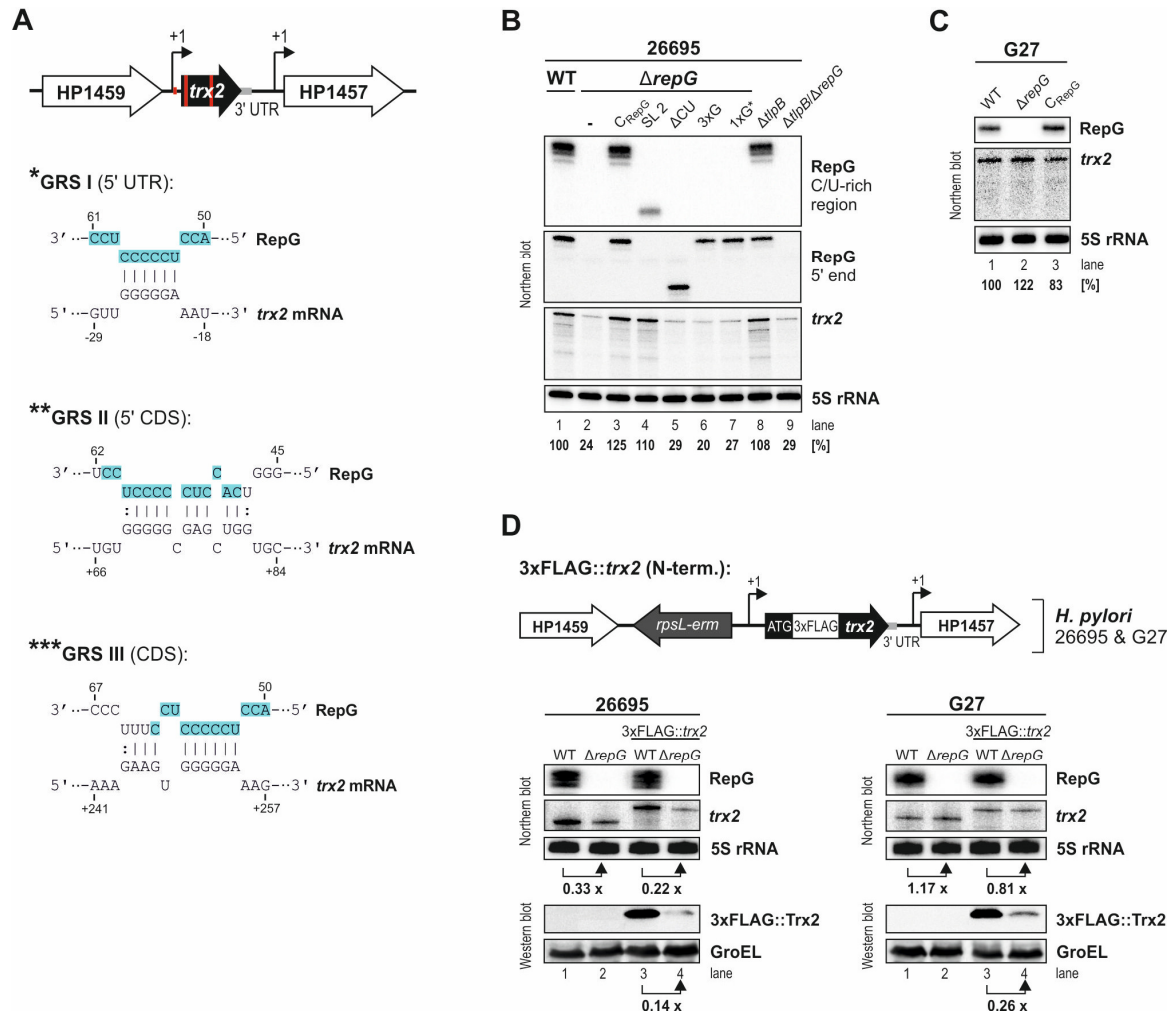
Figure 4.11: RepG activates *trx2* expression. (A) (*Upper panel*) Genomic location of *trx2*. TSS (+1) are indicated by black arrows. Positions of the three putative RepG interactions sites in the *trx2* mRNA are highlighted in red. (*Lower panel*) RepG was predicted to target three G-rich sequences (GRS) in the *trx2* mRNA, one in the 5' UTR and two in the coding region. Numbers indicate positions relative to the annotated *trx2* start codon. The *tlpB* interaction in the terminator loop of RepG is shown in blue. **(B)** *H. pylori* 26695 wildtype (WT), Δ*repG* and sRNA complementation strains (C$_{RepG}$, SL 2, ΔCU, 3xG and 1xG*), as well as Δ*tlpB,* and Δ*tlpB*/Δ*repG* double deletion mutants were grown to exponential growth phase and RNA samples were analyzed by northern blot. RepG was detected with CSO-0003 (binds to the C/U-rich loop) and JVO-2134 (binds to RepG 5' end). CSO-0896 was used for visualization of the *trx2* mRNA. 5S rRNA served as loading control (JVO-0485). **(C)** Northern blot analysis for *trx2* expression in *H. pylori* G27 at exponential growth phase. **(D)** (*Upper panel*) The *rpsL-erm* cassette along with a 3xFLAG epitope-tag was inserted after the start codon (ATG) into the *trx2* coding region of *H. pylori* strains 26695 and G27, resulting in N-terminal-tagged 3xFLAG-tagged Trx2 proteins. (*Lower panel*) *H. pylori* strains 26695 and G27 WT and Δ*repG,* as well as WT and Δ*repG* strains which carry the 3xFLAG::*trx2* fusions were grown to exponential phase, and RNA as well as protein samples were analyzed by northern and western blot, respectively. Size differences in the *trx2* mRNAs are due to introduction of the FLAG-tag sequence (~ 70 nt). Epitope-tagged Trx2 was detected with anti-FLAG antibody. GroEL served as loading control.

and G27, single point mutations were found within the other two RepG interaction sites (Appendix, Figure 13.5 C). Nucleotide exchanges in the RepG interaction sites are synonymous substitutions, meaning the amino acid sequence of Trx2 protein is not changed. However, these mutations might explain for the strain-specific activation of the *trx2* mRNA by RepG (Figures 4.11 B and C).

Direct base-pairing between the C/U-rich terminator loop of RepG and the homopolymeric G-repeat in the leader of the *tlpB* mRNA was shown to be sufficient to mediate regulation of the chemotaxis receptor (Chapter 2). Similarly, complementation of the Δ*repG* mutant with either the wildtype (C$_{RepG}$) or mutant sRNA SL 2, which consists only of the terminator loop, restored *trx2* activation by RepG in *H. pylori* strain 26695 (Figure 4.11 B, lanes 3-4). In contrast, deletion of the single-stranded C/U-rich binding site (ΔCU) and introduction of triple (3xG) or single C to G substitutions (1xG*) within the second stem loop abolished *trx2* regulation at the transcript level (Figure 4.11 B, lanes 5-7). This demonstrates that the conserved C/U-rich terminator loop of RepG is sufficient for the regulation of two target mRNAs, *tlpB* and *trx2*. However, while RepG represses the chemotaxis receptor in *H. pylori* strain 26695 at the transcript and protein level, it activates expression of the *trx2* mRNA. Northern blot analysis for *trx2* expression in Δ*tlpB* and Δ*tlpB*/Δ*repG* double deletion mutants revealed that the chemotaxis receptor is dispensable for RepG-mediated *trx2* activation (Figures 4.11 B, lanes 8-9), suggesting that both targets are regulated independently from each other.

To verify the influence of RepG on *trx2* expression at the protein level, a 3xFLAG epitope was fused to the N-terminal end of the *trx2* gene in *H. pylori* strains 26695 and G27. Cloning of an N-terminal 3xFLAG::*trx2* was preferentially used in this study because introduction of the epitope-tag together with the *rpsL-erm* resistance cassette to the *trx2* stop codon (C-terminal FLAG-tagging) would have resulted in disruption of the putative *trx2* 3' UTR (3' UTR indicated as gray bar in Figure 4.11 D, *upper panel*). Using western blot analysis, 3xFLAG::Trx2 protein levels were about five-fold decreased at exponential growth in *H. pylori* 26695 Δ*repG* mutant when compared to the wildtype (Figure 4.11 D, *lower panel*). Correspondingly reduced 3xFLAG::*trx*2 mRNA levels were detected in the *repG* deletion mutant, demonstrating that RepG activates *trx2* expression in *H. pylori* 26695 at the transcript and protein level. In contrast, 3xFLAG::*trx*2 mRNA levels were not significantly altered by RepG in *H. pylori* strain G27. Nevertheless, deletion of *repG* in *H. pylori* G27 carrying the N-terminal FLAG-tag also resulted in about four-fold decreased Trx2::3xFLAG protein levels. This suggests that RepG activates *trx2* expression at protein level in both *H. pylori* strains. However, steady-state *trx2* transcript levels are only affected significantly upon *repG* deletion in 26695.

4.8. Discussion: Targeting of homopolymeric G-repeats and G-rich sequences – a novel facet of sRNA-mediated gene expression control

Small RNAs control gene expression of multiple target mRNAs within distinct regulons and could have regulatory roles as broad as some transcription factors in bacteria (Beisel & Storz, 2010). This Chapter provides evidence that the highly abundant and conserved RepG sRNA acts as a multi-target riboregulator in *Helicobacter*. Deletion of *repG* was shown to significantly affect transcript levels of at least 45 genes in *H. pylori* strain 26695 (RNA-seq, Figure 4.2). Manual curation to compensate for co-regulated polycistronic genes, as in the case of the *tlpB*-HP0102 and HP1057-HP1055 operons, reduced the list to approximately 40 high-confidence RepG target mRNAs in this strain background. RepG generally acts as a repressor of gene expression control and regulates genes involved in membrane transport and adhesion, amino and fatty acid metabolism, LPS modification and nucleic acid restriction. Computational target prediction in combination with the unique enrichment patterns of cDNA sequencing reads in the *H. pylori* 26695 Δ*repG* mutant (*e.g.* Figure 4.4) suggested that RepG preferentially binds with its conserved C/U-rich terminator to homopolymeric G-repeats or G-rich sequences within its target mRNAs. A putative RepG binding site could be predicted for most of the deregulated mRNAs (Figure 4.6), suggesting that the majority of the identified target mRNAs might directly interact with RepG.

Nevertheless, constitutive overexpression or deletion of sRNAs fail to discriminate between direct and indirect target genes because of pleiotropic effects on overall expression, *e.g.* through sRNA-mediated control of a transcriptional regulator (Sharma & Vogel, 2009). Short-term pulse expression of sRNAs, *e.g.* from tightly controllable and inducible promoters, has been successfully applied to overcome these obstacles. An inducible system based on the TetR-repressor has been used to control mRNA expression in *H. pylori* (McClain *et al.*, 2013). However, a similar system for the transient overexpression of sRNAs is not yet available. Still, both the direct and indirect effects observed upon *repG* deletion are likely to contribute to the characterization of the biological role of this sRNA.

Overall, RepG was shown to impact more than 2 % of the *H. pylori* genome directly or indirectly under the examined growth conditions (*e.g.* 45 target candidates out of about 1,563 ORFs in 26695). This emphasizes the importance of sRNA-mediated, or in particular RepG-mediated, gene expression control in *H. pylori*. Likewise, enterobacterial sRNAs have been implicated in regulating up to 1 % of the genome in *E. coli* and *Salmonella* (Masse *et al.*, 2005, Sharma *et al.*, 2011).

The RepG regulon is largely conserved among different *H. pylori* strains. Comparison of RepG-mediated gene expression changes in *H. pylori* strains 26695 and G27 revealed common and putative strain-specific target candidates. About half of the target mRNAs

identified in each strain background are commonly regulated by RepG, whereas the other half appears to be controlled in a strain-specific manner (at least according to the filter criteria used in the RNA-seq analyses, Figure 4.7). Interestingly, in-depth analysis of the RNA-seq raw data and qRT-PCR analysis showed that the majority of "strain-specific" RepG targets are indeed mildly affected in one or the other strain (*e.g.* 1.5-fold change in mRNA levels in Δ*repG* vs. WT). In addition, although the transcript levels were only affected in *H. pylori* strain 26695 upon *repG* deletion, *trx2* expression was shown to be activated by RepG at the protein level in both strains, 26695 and G27 (Figure 4.11). This suggests that the RepG regulon is largely conserved among different *H. pylori* strains, or at least between 26695 and G27, albeit strain-specific differences in the strength of sRNA-mediated regulation have been observed for the same target mRNAs. RepG might serve analogous physiological functions in different *Helicobacter* strains as it (I) seems to control genes involved in similar cellular processes (*e.g.* fatty acid and amino acid metabolism, cell envelope biosynthesis) and (II) preferentially binds to homopolymeric G-repeats or G-rich sequences in its target mRNAs independent from the strain background.

RepG employs its conserved C/U-rich terminator loop for multi-target regulation. Except for the highly conserved C/U-rich terminator loop, the RepG sRNA is strongly structured and lacks extended single stranded regions. Thus, direct sRNA-mRNA interactions are energetically confined to its unstructured terminator loop region. Indeed, consistent with the putative RepG binding motif (Figure 4.6 B), the C/U-rich terminator of RepG seems to be the major determinant in sRNA-mediated gene expression control and can be considered as the "seed region" for the base-pairing of RepG to multiple mRNA targets. The importance of such highly conserved sRNA domains for multi-target recognition and regulation has been demonstrated for several sRNAs, *e.g.* GcvB in *Salmonella* (Sharma *et al.*, 2011). In contrast to RepG, the seed regions of many enterobacterial sRNAs are located at a single-stranded 5' end (Storz *et al.*, 2011, Papenfort *et al.*, 2010). Nevertheless, also sRNA loops have been shown to facilitate posttranscriptional gene expression control in *E. coli*, *e.g.* by formation of a kissing-loop complex (Argaman & Altuvia, 2000).

RepG exhibits striking structural features, *i.e.* the C/U-rich loop region, which also have been described for sRNAs in unrelated bacterial pathogens such as *S. aureus, Listeria monocytogenes* and *Xanthomonas campestris* pv. *vesicatory.* For example, albeit the overall structure is more complex, the RNAIII sRNA of *S. aureus* contains multiple C-rich stem loops, which can bind to the RBS of target mRNAs by loop-loop interactions and thereby, inhibit translation initiation and/or promote target mRNA degradation (Boisset *et al.*, 2007). In line with this, a conserved "UCCC"-motif was proposed in the loop regions of various other *S. aureus* non-coding RNAs (Geissmann *et al.*, 2009). Also, the LhrC4 sRNA of *L. monocytogenes* employs three C/U-rich sites, two of which are present in hairpin

structures, to repress translation of the LapB adhesion (Sievers *et al.*, 2014). In *Xanthomonas,* the sX13 sRNA represses expression of various virulence-associated genes by targeting short G-rich motifs in close proximity to the translational start codon (Schmidtke *et al.*, 2013). Interestingly, one of the C-rich loops of sX13 (UCCCCCU) is identical to the central region of the C/U-rich terminator loop of RepG. Recently, a potential functional RepG homolog was discovered in *C. jejuni* (PhD S. Svensson & Dr. C. M. Sharma, unpublished). Similar to RepG, this sRNA possesses a C/U-rich loop region and was predicted to interact with homopolymeric G-repeats and/or G-rich sequences in its target mRNA candidates. Remarkably, RepG is able to restore target mRNA regulation in the respective sRNA deletion mutant in *C. jejuni*, suggesting a common mechanism of sRNA-mediated gene expression control in Epsilonproteobacteria. Moreover, taking the other examples into account, structure-driven target gene regulation by C-rich sRNA loops seems to represent a general mechanism of posttranscriptional gene regulation in diverse bacteria.

Homopolymeric G-repeats and/or G-rich sequences are preferential RepG target sites.
The C/U-rich terminator of RepG has a high risk for promiscuous base-pair interactions with its target mRNAs as the uridines within or adjacent to the "seed region" (CCUCCCCCUCC) can permit two different base-pair interactions (U:A, U:G). Accordingly, RepG was predicted to bind G-rich sequences rather than to a defined nucleotide sequence in its target mRNAs (Figure 4.6 B). The core sequence of the RepG binding motif is composed of a 7 to 8G-long repeat and is in full agreement with the optimal G-repeat length (7 to 12Gs) required for RepG-mediated *tlpB* repression. However, the redundancy within the C/U-rich terminator loop of RepG might also facilitate thousands of "unspecific" RepG-mRNA contacts, especially at G-rich sequences (see FIMO analysis, section 4.3). Therefore, it is possible that sequences adjacent to the sRNA interaction site and/or structural determinants might confer specificity of RepG base-pairing. Often, short repetitive G-rich sequences were found in close proximity, adjacent to or within the RepG interaction sites in target mRNAs (Figure 4.8 and Appendix, Figures 13.6-13.8).

The homopolymeric G-repeat in the 5' UTR of the *tlpB* mRNA leader was proposed to function as a putative translational enhancer element (see discussion in Chapter 2). Likewise, G-rich sequences (RepG binding sites and/or adjacent ones) might promote translation initiation and/or elongation and thus, mRNA stability/abundance. Accordingly, RepG binding might interfere with this enhancer activity. Similar findings have been obtained for the *Salmonella* GcvB sRNA and the sX13 sRNA in *Xanthomonas*. While GcvB targets C/A-rich enhancer elements (Sharma *et al.*, 2011), sX13 was shown to repress target mRNAs by base-pairing to G-rich sequences (Schmidtke *et al.*, 2013). In contrast to GcvB and sX13, which bind to enhancer elements that are located upstream of the RBS or in close proximity to the translational start site, almost all of RepG interaction sites are located

within mRNA coding regions (Figure 4.12, except *e.g. tlpB*-HP0102). While initiation of translation is mainly controlled by structures within the 5' UTRs (Gu *et al.*, 2010), stable structures in open reading frames have been shown to contribute to ribosomal frame-shifts (Giedroc & Cornish, 2009, Endoh *et al.*, 2013), mRNA surveillance by no-go mRNA decay (Doma & Parker, 2006) and co-translational folding of nascent proteins (Komar, 2009). Albeit the majority of reported structures within coding regions suppress translation elongation, G-rich RepG targeting sites might in fact promote translation and/or mRNA stability in absence of RepG.
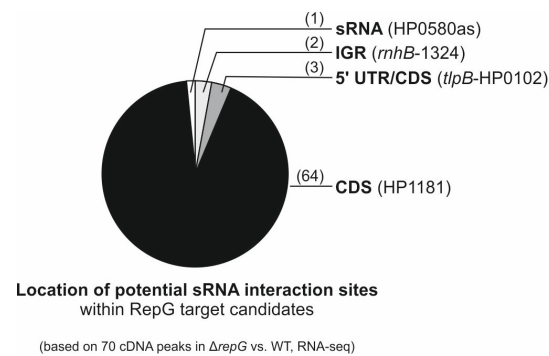
Figure 4.12: Location of potential RepG interaction sites (= cDNA peaks) within target mRNAs. The specific enrichment pattern of cDNA reads in Δ*repG* mutant compared to the wildtype (RNA-seq) greatly overlapped with predicted sRNA interaction sites within target mRNAs that are repressed by RepG (40 candidates, Figure 4.6 A). Putative RepG binding sites are found to be present within the mRNA leader (5' UTR/CDS), coding sequence (CDS) or in the intergenic region (IGR) between two genes, and in sRNAs. Numbers indicate RepG interaction sites in each category.

Small RNA base-pairing to the mRNA coding region has been shown to either result in rapid target mRNA decay by recruitment of RNases (Pfeiffer *et al.*, 2009, Frohlich *et al.*, 2012) or structural rearrangements that interfere with translation initiation (Heidrich *et al.*, 2007). Therefore, it is also possible that targeting of G-rich sequences by RepG might result in direct target mRNA degradation and/or structural rearrangements that block ribosome binding. Future investigations of selected RepG-target mRNA interactions will help to understand the exact mechanism of action and underlying molecular details. Whether or not structural determinants, such as RNA G-quadruplexes, might contribute to RepG-mediated target regulation still needs to be clarified.

Several target mRNAs possess more than one putative RepG interaction site, suggesting that RepG has the ability to control its target mRNAs by cooperative or coordinated binding to multiple G-rich sites, *e.g.* HP1181 and *trx2*. Mutations within one of the sRNA interaction sites in *H. pylori* strain G27 might lead to the observed strain-specific regulation of HP1181 and/or *trx2* through RepG at the transcript level. Small RNA binding to

multiple target sequences in bacterial mRNAs has been reported, but usually involves different regions of the respective sRNA. Examples are given by MicF that binds to the *lpxR* mRNA both at the RBS and in the coding region (Corcoran *et al.*, 2012a), and the polycistronic *manXYZ* mRNA, which is targeted at the RBS and an intergenic region by the SgrS sRNA (Desnoyers *et al.*, 2013). Like the MicF-*lpxR* example, RepG was predicted to bind not only to the homopolymeric G-repeat in the *tlpB* mRNA leader, but also to two additional G-repeats in its coding region. However, GFP-reporter assays indicated that the G-repeat in the 5' UTR is sufficient for RepG-mediated repression (Figure 2.4 B). Future studies should be designed to investigate the relevance of RepG binding to multiple sites within the same target transcript.

Except for the homopolymeric G-repeat in the 5' UTR of the *tlpB* mRNA, none of potential RepG interaction sites within the newly identified target mRNAs represents a recognizable, phase-variable SSR. Preliminary sequence alignments of chosen RepG targets revealed that the majority of the sRNA interaction sites, even if they possess a homopolymeric G-repeat, are conserved among different *H. pylori* strains (data not shown). If present, variations in the RepG interaction sites are based on single nucleotide polymorphisms rather than variations in the G-repeat lengths. Since almost all of the predicted RepG interactions sites are located in the coding region of target mRNAs (Figure 4.12), variations in the G-repeat length would lead to frame-shift mutations and consequently, ON/OFF switch of gene expression. This probably leads to a selection against length variation in the coding region-associated homopolmeric G-repeats/G-rich sequences.

RepG's putative physiological role in *H. pylori*. RepG seems to be a versatile riboregulator that is involved in diverse biological processes (Figure 4.13). About one-fourth of the putative RepG target genes have been implicated in *H. pylori* virulence or acid adaptation (Figure 4.2 E). This reflects the general high percentage of the *H. pylori* genome that has been devoted to the ability to colonize and persist within the murine stomach (Baldwin *et al.*, 2007). Nevertheless, without being exclusively restricted to the regulation of virulence factors, RepG might contribute to *H. pylori* pathogenicity and colonization of the host.

The *Helicobacter* chemotaxis core signal transduction machinery consists of the chemotaxis receptors TlpA, TlpB, TlpC and TlpD, the CheA/CheY two-component system and accessory proteins such as CheW and CheV, which physically couple CheA (sensor kinase) to the chemotaxis receptors (Lertsethtakarn *et al.*, 2011). *H. pylori* encodes three CheV proteins, CheV1-V3, which catalyze the removal of phosphate from the kinase CheA, thereby preventing phosphorylation of CheY (Jimenez-Pearson *et al.*, 2005). RepG specifically regulates expression of both the acid-sensing chemotaxis receptor TlpB and coupling protein CheV2 (*e.g.* repression in *H. pylori* strain 26695, Figures 4.3 and 4.5 B).
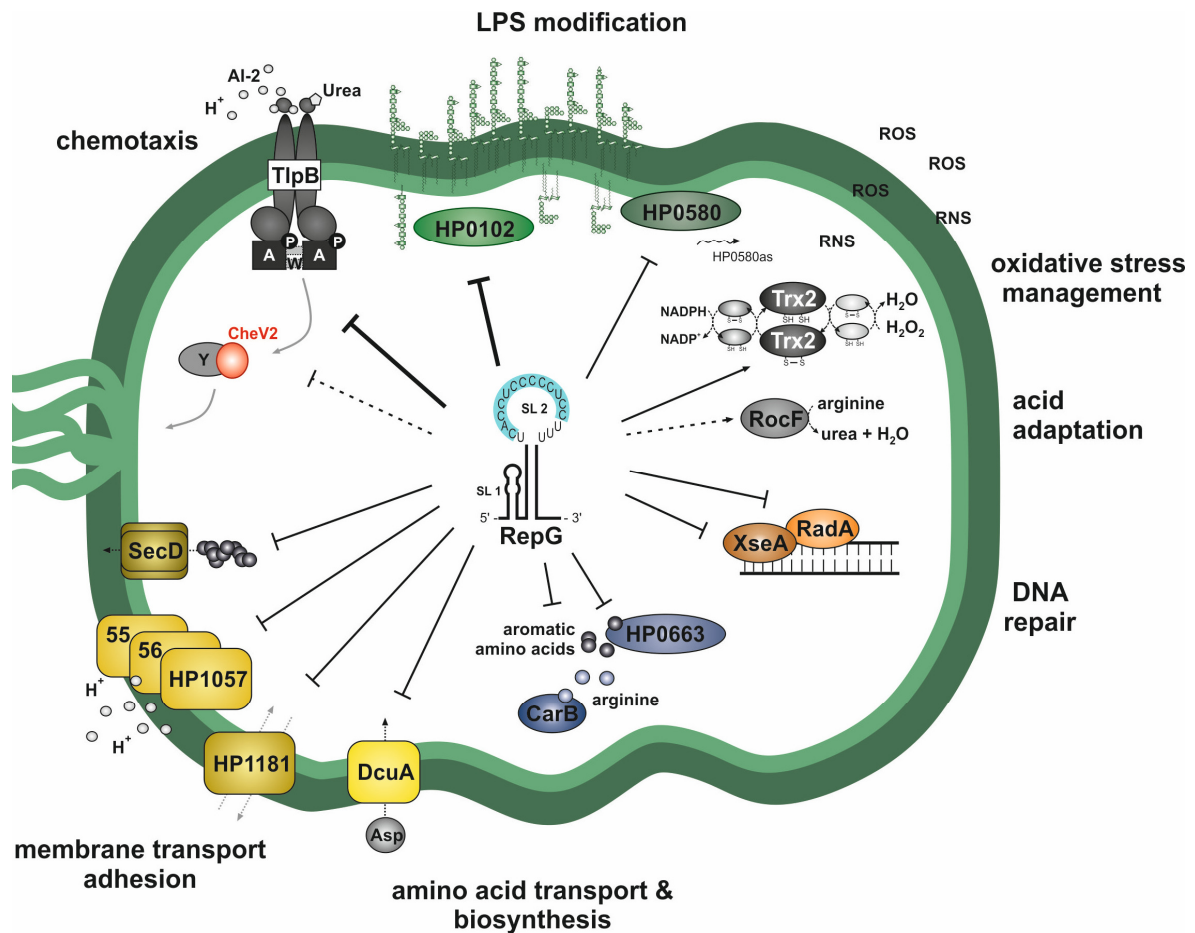
Figure 4.13: Overview about the RepG regulon. Using microarray and RNA-seq, more than 40 putative RepG target genes were identified in *H. pylori* strain 26695, including the *tlpB*-HP0102 operon. Newly identified RepG target gene candidates are involved in membrane transport and adhesion, amino acid metabolism, nucleic acid modification and LPS biosynthesis, oxidative stress management and/or acid adaptation. RepG represses the Kdo-hydrolase HP0580 and expression of the putative, *cis*-encoded sRNA (HP0580as), which is encoded antisense to the 5' coding region of HP0580. Although RepG mainly act as a repressor of gene expression control, the redox protein Trx2 and arginase RocF are positively affected by this sRNA. This scheme shows only representative RepG target gene candidates belonging to the indicated functional categories/cellular processes. Thick arrows indicate RepG-target gene interactions, which have been investigated in detail in this thesis.

According to *cheV2* mutant strains displaying an altered motility behavior (Lowenthal *et al.*, 2009), RepG-mediated repression of *tlpB* and *cheV2* might contribute to a specific chemotactic and swimming behavior of *H. pylori*. In line with this, deletion of *repG* was shown to affect the motility behavior of *H. pylori* strain X47-2AL (Figure 3.10 D).

The outer membrane of Gram-negative bacteria functions as a selective barrier that prevents the entry of many toxic molecules into the cell, plays a vital role in bacterial survival in diverse environments and is crucial for the adherence to host cells. Many enterobacterial sRNAs have been shown to impact the composition of the outer membrane at the posttranscriptional level, *e.g.* by fine-tuning the expression of outer membrane

proteins (OMPs) (Papenfort & Vogel, 2010). Likewise, RepG was shown to repress expression of various membrane proteins, which might be involved in bacterial adhesion and transport through the inner and outer membrane. Among them, the HP1057-HP1055 operon encodes three hypothetical outer membrane proteins that belong to a small sub-OMP family confined to *Helicobacter* (pfam01856). Several proteins of this family, including HP1057, were identified as candidate loci contributing to stomach colonization of mice (Baldwin *et al.*, 2007). In addition, RepG-mediated control of the multidrug-efflux pump HP1181 might affect *H. pylori* (26695) surface properties, *e.g.* membrane trafficking and transport of molecules. Although the exact function of HP1181 in *Helicobacter* is unknown, it shares sequence homology to multidrug-efflux pumps that are involved in fluoroquinolone resistance and tetracycline and/or carbohydrate ion transport in Gram-positive bacteria (Morrison *et al.*, 2003, Kaatz *et al.*, 1993). Besides direct targeting of (outer) membrane proteins, RepG might also indirectly control the membrane protein content of *H. pylori*. For example, a component of the Sec protein translocation machinery, *secD*, is repressed by RepG in *H. pylori* strain 26695. The Sec translocon transports secretory proteins across the inner membrane and inserts membrane proteins into the inner membrane of Gram-negative bacteria (Kudva *et al.*, 2013).

RepG might not only shape the outer membrane of *H. pylori* by modification of the protein content, but also by regulating genes involved in LPS biosynthesis, including HP0102 and HP0580. While the glycosyltransferase HP0102 is essential for LPS O-chain synthesis and Lewis x antigen production, Kdo-hydrolase HP0580 modifies the lipid A moiety. *H. pylori* mutants lacking either HP0102 or HP0580 have been shown to display strongly altered LPS O-chain patterns and Lewis x antigen expression (Figures 3.5 and 3.6, Stead *et al.*, 2010). Accordingly, RepG-mediated control of both HP0102 and HP0580 might synergistically contribute to alterations in the antibiotic susceptibilies of *H. pylori* strain 26695 upon *repG* deletion (Figures 3.12 and 3.13). The Kdo-hydrolase activity is dependent on a two-protein complex composed of HP0580 and HP0579. HP0580 is encoded in an operon together with the flagellar switch protein FliN/Y and three hypothetical proteins (HP0583-HP0581, Figure 4.14 A). A potential TSS in the HP0580 coding region implicates that transcription of the downstream-encoded HP0579 is uncoupled from that of the *fliN/Y*-HP0580 operon (Sharma *et al.*, 2010). RepG-mediated repression of HP0580 mRNA might be achieved by two different mechanisms. First, a putative G-rich targeting site was identified in the HP0580 coding region, suggesting a direct RepG-HP0580 mRNA interaction. Secondly, RepG might repress expression of a putative, *cis*-encoded antisense RNA (HP0580as). The 140-nt long HP0580as overlaps with two homopolymeric C-repeats within the 5' end of the HP0580 coding region (Figure 4.14 A). These poly-C tracts have been assumed to contribute to phase-variable ON/OFF expression of the Kdo-hydrolase

(Appendix, Table 13.2). Antisense RNAs can affect the expression of the complementary sense transcript at the level of transcription, mRNA stability and translation (Sesto *et al.*, 2013). Among others, asRNAs have been shown to stabilize the RNA expressed from the opposite strand, *e.g.* by inducing specific cleavage events or protecting target mRNAs from degradation by RNases (Opdyke *et al.*, 2004, Obana *et al.*, 2010, Ramirez-Pena *et al.*, 2010). So far, no obvious promoter or transcriptional start site have been identified for the putative HP0580as (not present in the transcriptome study of Sharma *et al.*, 2010), raising the question of whether this transcript is indeed expressed. Further studies will be required to unveil the nature of HP0580as and how it might be connected to the expression of the Kdo-hydrolase.

Similar to the enterobacterial GcvB sRNA (Sharma *et al.*, 2011), RepG represses multiple *H. pylori* genes important in amino acid metabolism. This includes proteins involved in amino acid biosynthesis including *carB* and HP0663, as well as amino acid transporters such as *dcuA*. Since RepG is highly expressed in fast growing cells in nutrient-rich BHI medium (Figure 2.14 A), it is likely that one function of RepG might be to limit energy-consuming amino acid production and uptake when nutrients are plentiful.

Oxidative DNA damage is considered as a major source of cell death and mutagenesis. *H. pylori* encodes for various proteins involved in DNA replication, recombination and repair (Wang *et al.*, 2006a). Among them, the DNA repair protein RadA and exodeoxyribonuclease subunit XseA were shown to be repressed by RepG at the transcript level (Figures 4.9 J-K). Both are assumed to play an important role in DNA repair after exposure to mutagenic agents, *e.g.* UV-radiation and ROS (Song & Sargentini, 1996, Repar *et al.*, 2013). Therefore, RepG-mediated control of DNA repair/recombination enzymes might be associated to ROS-induced DNA damage.

In *H. pylori* strain 26695, RepG represses expression of almost all identified target mRNAs. Exceptions are given by the *rocF* and *trx2* mRNA, encoding for an arginase and antioxidant protein, respectively. For both, at least one putative RepG interaction site was predicted in the mRNA (5' UTR and/or coding region; Figures 4.14 B and 4.11 A), suggesting that *rocF* and *trx2* mRNA levels might be positively affected by direct base-pair interactions with RepG. A recent whole proteome analysis of RepG-mediated changes in gene expression using SILAC (<u>s</u>table <u>i</u>sotope <u>l</u>abeling with <u>a</u>mino acids in <u>c</u>ell culture) confirmed activation of *rocF* and *trx2* through RepG at the protein level (independent from the here presented FLAG-tagging approach; Müller, Pernitzsch *et al.*, 2015, accepted). Therefore, both are interesting candidates to study RepG-mediated target gene activation and its underlying molecular mechanism(s).
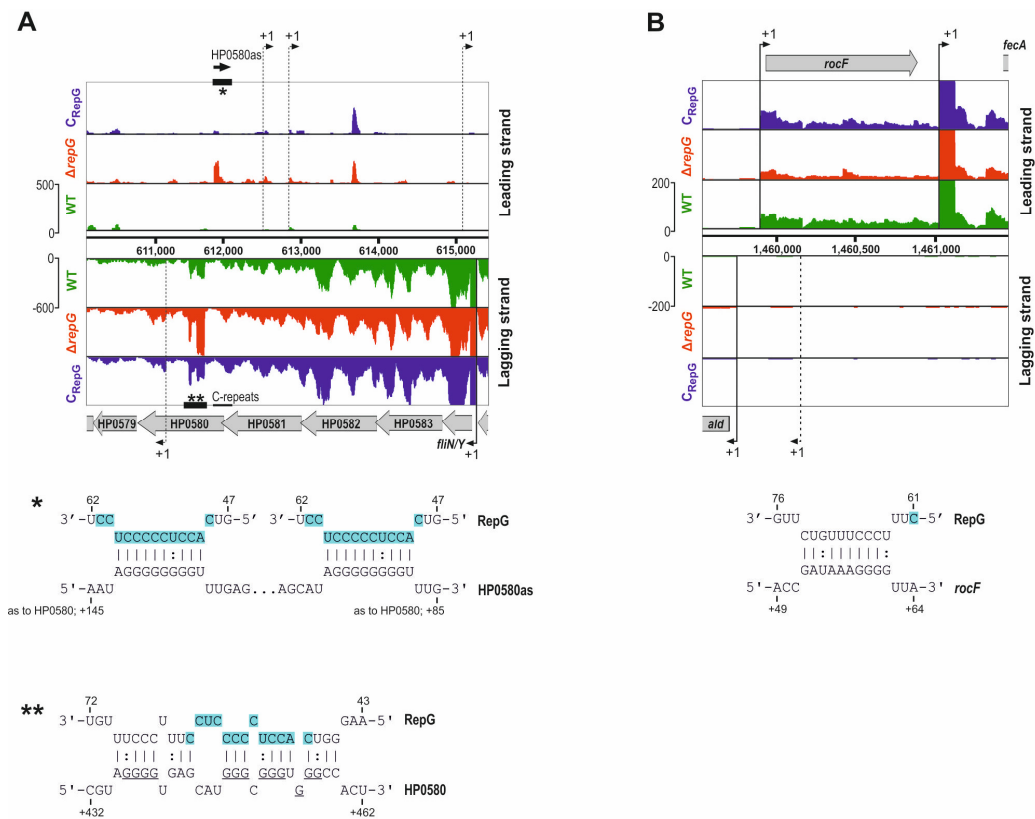
Figure 4.14: RepG-mediated regulation of the Kdo-hydrolase HP0580 and arginase RocF. (A-B) IGB screenshot showing cDNA reads from RNA-seq of WT, Δ*repG* and C$_{RepG}$ mapped to the *fliN/Y*-HP0579 operon and *rocF* region in *H. pylori* strain 26695. Gray and black arrows represent the annotated ORFs and TSS (+1; filled line – primary TSS, dotted line – antisense TSS), respectively. **(A)** The Kdo-hydrolase (HP0580-HP0579) is encoded in an operon with *fliN/Y* and HP0583-HP0581. Specific enrichment patterns for cDNA reads (thick black lines) in Δ*repG* greatly overlap with putative RepG-target mRNA interaction sites (denoted by */**). The 5' coding region of HP0580 contains two homopolymeric C-repeats. RNA-seq indicated expression of HP0580as, which was predicted to bind to the C/U-rich terminator loop of RepG (*). RepG also binds to the coding region of the HP0580 mRNA (weak, **). Numbers indicate position in the HP0580 coding region according to the annotated start codon. Repetitive G-rich sequences in the **-interaction are underlined. **(B)** A putative, predicted RepG–*rocF* mRNA interaction is shown below the IGB screenshot. Numbers indicate position according to the annotated *rocF* start codon.

The arginase RocF is part of the urea cycle and hydrolyzes L-arginine to L-ornithine and urea, which is subsequently used by the urease enzyme (Ikemoto *et al.*, 1990). Like *pylori* nitrogen metabolism and acid acclimatization (McGee *et al.*, 1999). It is also considered an oxidative stress-combating enzyme as RocF competes with the host's nitric oxide (NO) synthase for the common substrate L-arginine, thereby reducing the synthesis of NO from host cells and RNS-mediated oxidative stress for *Helicobacter* (Gobert *et al.*, 2001). Although RocF is not essential for colonization of the murine stomach, it is considered an important virulence factor in *H. pylori* (Baldari *et al.*, 2005). Posttranscriptional activation of

rocF through RepG might contribute to acid acclimatization/oxidative stress management in *H. pylori* (*e.g.* Figure 2.24 B).

Together with the Trx1 and TrxR, Trx2 catalyses the reduction of different antioxidant enzymes, such as peroxiredoxins (Wang *et al.*, 2006a). Peroxiredoxins require donation of electrons from reduced thioredoxins for hydrogen peroxide and lipid hydroperoxide reduction, thereby contributing to the membrane integrity of *H. pylori* under oxidative or nitrosative stress conditions (Wang *et al.*, 2006b). The exact function of *trx2* is still unknown. Deletion of either *trx1* or *trx2* was shown to result in an increased sensitivity to ROS/RNS (Comtois *et al.*, 2003), implicating a function of Trx2 in *H. pylori* oxidative stress management. In addition, proteome studies of the *H. pylori* secretome showed that Trx2 is constitutively secreted (Bumann *et al.*, 2002, Kim *et al.*, 2002). This raises the intriguing possibility that Trx2 might have an effect on gastric host cells. In agreement with this, *H. pylori* Trxs have been shown to be efficient reductants of human immunoglobulin (IgA) and mucins *in vitro* (Windle *et al.*, 2000). Thus, it is conceivable that Trx2 (and/or RepG-mediated activation of *trx2* expression) may contribute to *H. pylori* oxidative stress management and/or host colonization. In line with this, *repG* expression was shown to be regulated in response to elevated oxygen tensions (Figure 2.18). Interestingly, Trx1, but not Trx2, is assumed to function as a chaperon of the arginase RocF. Future studies will be required to investigate how and/or to which extent RepG-mediated activation of the *rocF* and *trx2* influences oxygen sensitivity/host colonization of *H. pylori*. In addition, it would be interesting to see whether or not Trx2 might also function as a protein chaperone.

Validation of RepG-mediated target gene regulation at the protein level. Work presented in this Chapter was mainly restricted to RepG-mediated changes in gene expression at the transcript level, not considering alterations in protein abundances. Changes in the RNA content are not necessarily associated with alterations in protein abundances. Indeed, mRNA and protein levels have been shown to be imperfectly correlated in yeast, mammalian cells and bacteria (Vogel & Marcotte, 2012), and posttranscriptional control can have a crucial role in modulating gene expression (Vogel, 2011). Although sRNA-mediated inhibition of translation is often coupled to target mRNA degradation, many enterobacterial sRNAs have been shown to regulate translation of target genes without affecting their transcript levels. For example, the Spot42 sRNA was shown to specifically block ribosome binding to the *galK* mRNA; however, no mRNA degradation occurs (Moller *et al.*, 2002). Conversely, sRNA-mediated destabilization of a target mRNA is not necessarily associated with the reduction of the overall protein abundance (*i.e.* RepG-mediated regulation of *tlpB* in *H. pylori* strain G27). In addition, the number of mRNAs positively affected by RepG is most likely underrepresented in the microarray and RNA-seq study. Since sRNAs often act as translational activators, *e.g. trx2* activation by RepG in *H. pylori*

strain G27, the effect on targeted mRNAs is often too subtle to be detected by northern blot or whole transcriptome analysis (Frohlich & Vogel, 2009).

A global protein-tagging strategy is hardly feasible as it cannot be excluded that FLAG-tagging, for instance, will affect protein localization, expression and sRNA-mediated regulation (*e.g.* RepG-mediated repression of HP1181, Figure 4.10). Moreover, several attempts to validate chosen RepG target genes by chromosomal FLAG-tagging and or GFP-reporter fusions failed as fusion proteins are often not stably expressed or incorrectly folded under the examined conditions (data not shown). Ribosome profiling might represent an alternative strategy that could be used for global expression profiling of the translatome. To date, this methodology was applied not only to study ribosome positions at mRNAs, but also to measure gene expression quantitatively and changes thereof at the level of actual protein synthesis in both eukaryotic and prokaryotic organisms (reviewed in Ingolia, 2014). The combination of both, RNA sequencing and ribosome profiling could help investigate RepG-mediated gene expression changes at the transcript and protein level. In addition, it might provide a useful tool to unveil whether G-rich sequences might serve as internal ribosome binding as well as stalling sites, or translational enhancers.

The identification and validation of direct target candidates will provide further insights into the physiological role of RepG in *Helicobacter*. Moreover, in depth analysis of chosen RepG-mRNA interactions will contribute to the understanding of the underlying molecular mechanism of RepG-mediated regulation.

5. Conclusion and perspective

The here presented study aimed to functionally characterize the first example of a *trans*-acting riboregulator, RepG sRNA, in *H. pylori*. Due to its strong abundance and high degree of conservation in diverse *H. pylori* strains, RepG represents an excellent candidate for studying the role of riboregulation through sRNAs in this human pathogen. Using microarray and RNA-seq, RepG-mediated gene expression changes were examined in *H. pylori* strains 26695 and G27 to identify its cellular interaction partners. In addition, the interaction between RepG and the discitronic *tlpB*-HP0102 mRNA was investigated in depth to understand the underlying molecular mechanism(s) of sRNA-mediated gene expression control in Epsilonproteobacteria. Finally, the role of RepG-mediated target gene (*tlpB* and HP0102) regulation in *H. pylori* virulence- and survival-associated phenotypes, including the colonization in mice, was analyzed to gain insights into the biological function of RepG and its cellular targets.

This thesis provides evidence that RepG is a multi-target riboregulator that mediates both activation and repression of more than 40 putative target mRNAs in *H. pylori* (Chapter 4). Alignment of RepG homologs in various *Helicobacter* species revealed a highly conserved C/U-rich region within the RepG terminator loop, which was subsequently shown to be required for direct interaction and regulation of selected target genes. Accordingly, bioinformatics-based predictions revealed that RepG preferentially interacts with homopolymeric G-repeats or (repetitive) G-rich sequences upstream of the RBS and/or within the coding region of its target mRNAs. In particular, the direct base-pairing between the C/U-rich terminator loop of RepG and a homopolymeric G-repeat in the *tlpB* mRNA leader was shown to be sufficient for RepG-mediated repression of the entire *tlpB*-HP0102 operon in *H. pylori* strain 26695. Contrary to the canonical pathway employed by other repressor sRNAs, *i.e.* inhibition of translation initiation by blocking ribosome entry, RepG binding to the *tlpB*-HP0102 mRNA far upstream of the *tlpB* RBS most likely influences target mRNA translation (initiation and/or elongation) by interference with a translational enhancer element, *i.e.* the homopolymeric G-repeat in the *tlpB* 5' UTR (Chapter 2). The presence of G-rich sequences – which might act as translational enhancers and/or protect mRNAs from degradation – in almost all putative target mRNAs indicates that RepG hijacks a conserved element to recognize and regulate multiple target mRNAs.

Although this study provided first insights into the molecular mechanism(s) of sRNA-mediated gene expression control in *Helicobacter*, the exact way in which RepG binding to G-rich sequences could interfere with the translational enhancer activity and/or lead to the recruitment of RNases remains elusive. It is still unclear whether or not

structural features such as G-quadruplexes, flanking regions of the RepG interaction site or its position within the target mRNA (5' UTR or coding region), might contribute to the specificity of sRNA binding and outcome of RepG-mediated gene expression control. Target regulation via C-rich sRNA loops and G-rich motifs in mRNAs has been described in several Gram-positive and -negative bacteria, including *Staphylococcus*, *Neisseria* and *Campylobacter,* and thus, might represent a more widespread and conserved feature/mechanism of sRNA-mediated posttranscriptional control. Accordingly, *H. pylori*, and in particular the interaction between RepG and the *tlpB*-HP0102 mRNA, could represent a model for studying how specificity of sRNA binding to simple sequence repeats is achieved in these human pathogens.

Phase variation by hypermutable SSRs represents a high-frequent and reversible mechanism of gene expression switching in diverse bacterial pathogens, including *H. pylori*. Work presented in this thesis demonstrated that the length of a homopolymeric G-repeat in the *tlpB* mRNA leader determines the outcome of RepG-mediated posttranscriptional control of the chemotaxis receptor TlpB and glycosyltransferase HP0102 (Chapters 2 and 3). The modulation of *tlpB*-HP0102 expression though length variation of an SSR represents a new paradigm in sRNA-mediated gene regulation and uniquely links posttranscriptional gene expression control with phenotypic variations though SSRs.

In *H. pylori*, length variation of SSRs have been shown to affect expression of genes involved in bacterial surface structures and thus, host recognition and adhesion. In line with this, the glycosyltransferase HP0102 was shown to be essential for O-chain and Lewis x antigen production in diverse *H. pylori* strains and therefore, represents an important component of the LPS biosynthesis pathway (Chapter 3). Moreover, mouse infection studies revelaed that this glycosyltranferase represents a novel colonization factor in *H. pylori*. In contrast to the digital ON/OFF switches of gene expression associated with SSRs in the coding region, RepG-mediated posttranscriptional control of the *tlpB*-HP0102 operon allows for a gradual modulation and thus, fine-tuning of LPS expression. Variations in *H. pylori* surface features might be required for the adaptation to a specific niche in the human host or during persistent infection.

The here investigated interaction of an sRNA with an SSR exemplifies how mutations within the bacterial genome can impact on posttranscriptional control mechanisms. However, it is still unclear when (how frequent and/or under which conditions) the length of the homopolymeric G-repeat in the *tlpB*-HP0102 mRNA might change in *H. pylori*. Considering that phase-variation by SSRs is a stochastic event, length variation in the G-repeat and thus, changes in the *tlpB*-HP0102 expression will only occur in some, but not all cells of a given population. Usually, the host milieu promotes selection for bacterial strains with particular characteristics (*e.g.* Lewis antigen expression) that facilitate

adaptation and survival in a given environment. Since this evolutionary pressure is likely to be absent or negligible under standard lab conditions, single-cell analysis, including single-cell sequencing (Saliba *et al.*, 2014) or FACS-based applications using GFP-reporter fusions, could be used to assess the switching frequency (*e.g.* 12 to 14Gs, repression to activation by RepG) and cell-to-cell variations within a *H. pylori* population. Moreover, sequencing of *H. pylori* isolates (clinical or after mouse infection) could provide further insights into how and/or to what extent the host milieu contributes to phase-viable expression of the *tlpB*-HP0102 operon. It still unknown whether or not the presence or expression levels of the RepG sRNA itself might provide a selective force towards a certain length of the homopolymeric G-repeat, or influence the rate of SSR polymorphism. If so, it would be interesting to investigate how and/or to which extent the abundant sRNA RepG might contribute to the genomic stability of *H. pylori*.

Length variations of poly-G tracts under selective environmental conditions and after passage through animals have also been observed in other human pathogens such as *C. jejuni* and *Neisseria meningitides* (Jerome *et al.*, 2011, Bayliss *et al.*, 2012, Alamro *et al.*, 2014). In principal, these G-stretches could also represent potential target sites of sRNAs, which have been recently identified in these pathogens (Dugar *et al.*, 2013, Pannekoek & van der Ende, 2012, Remmele *et al.*, 2014). Therefore, sRNA-mediated gene expression control though phase-variable SSRs of either *cis-* or *trans*-encoded target genes might be more widespread and represent a general theme of riboregulation in pathogenic bacteria.

Similar to other sRNAs, work presented in this thesis showed that RepG is global and versatile riboregulator, which seems to be involved in controlling gene expression in diverse stress responses (*e.g.* oxidative stress) and virulence pathways in *H. pylori*. Nearly one-fourth of the RepG target candidates have been previously shown or suggested to be important for colonization of the host and/or bacterial pathogenicity, implicating a central role for sRNAs in *Helicobacter* virulence control. Therefore, studying sRNA-mediated gene expression control in *Helicobacter* will not only provide new insights into its virulence mechanisms, but will also lead to the discovery of novel virulence-associated factors, such as the glycosyltransferase HP0102 (Chapter 3). Parallel sequencing of infection samples (*in vitro* or *in vivo)* by dual RNA-seq (Westermann *et al.*, 2012), which allows for the simultaneous analysis of gene expression changes in both the pathogen and the host, could be used to investigate the role of RepG during colonization and long-term (persistent) *H. pylori* infections. In addition, dual RNA-seq could be applied to identify additional sRNAs that are important for *H. pylori* virulence and pathogenicity.

One central question remaining is under which conditions exactly RepG-mediated posttranscriptional control occurs and/or is (most) important for *H. pylori*. Preliminary data suggested that *repG* expression is regulated in response to different iron availabilities, acidic

and oxidative stress conditions. However, transcriptional regulator(s) of RepG remain unknown. Additional phenotypic characterization of the *repG* deletion mutant (*e.g.* survival under acidic or oxidative stress conditions) could provide further insight into its physiological function. Considering that RepG is highly abundant in *Helicobacter* and its levels are often not limiting for target gene regulation under examined conditions, RepG might represent a housekeeping sRNA that prevents target mRNA overexpression, protects mRNAs from degradation or counteracts transcriptional responses at the posttranscriptional level.

Epsilonproteobacteria lack homologs of the RNA chaperone Hfq and RNase E, which both are key players in sRNA-mediated gene expression control in enterobacteria. Using a two-plasmid system, preliminary data showed that RepG is unable to repress a *tlpB*::*gfpmut3* reporter fusion in *E. coli.* Whether this is due to variations in RepG expression and stability, or a phylum-specific protein factor (Hfq-like chaperon) is unknown so far. The RepG-*tlpB* mRNA interaction provides an excellent case for identifying and/or studying novel protein factors important for sRNA activity in *Helicobacter*. Especially, global approaches, including RIP-seq and Tn-seq (deep sequencing of transposon-insertion site junctions) could be used to identify RNA-protein interaction partners and protein factors that may be required for sRNA-mediated posttranscriptional control in *H. pylori*, including a functional homolog of Hfq. The identification of RNA-binding proteins will help to improve the knowledge on sRNA-mediated gene expression control not only in Epsilonproteobacteria, but also in other bacteria which lack Hfq and/or RNase E.

In summary, this study demonstrates that *H. pylori* indeed uses riboregulation for gene expression control. Small RNA-mediated posttranscriptional control of gene expression might complement/supplement the relatively modest repertoire of *H. pylori* transcriptional regulators. The ongoing functional characterization of *cis-*and *trans*-acting sRNA candidates along with the identification of RNA-binding proteins will help to unveil their regulatory functions and roles in *H. pylori* virulence control and stress response. As shown in the current work, these studies could reveal novel mechanisms of posttranscriptional regulation independent of the RNA chaperone Hfq. Research on riboregulation in *H. pylori* will help to understand common themes of posttranscriptional regulation in other Epsilonproteo-bacteria, including the widespread and emerging pathogen *Campylobacter*. Overall, this study established *H. pylori* as a new model organism for studying general aspects of riboregulation in pathogenic bacteria lacking the major RNA chaperone Hfq.

6. Material and methods

6.1. Material

6.1.1. Instruments and devices

Table 6.1: Instruments and devices.

Instrument/device	Manufacturer
analytical balances TE64, TE601	Satorius
Bio-Link BLX 254 UV-Crosslinker	Peqlab
cell culture hood, HERASafe	Thermo Scientific
Celltron Shaker	Infors HT
centrifuge Eppendorf 5415C	Eppendorf
centrifuge Eppendorf 5424R	Eppendorf
centrifuge Eppendorf 5418R	Eppendorf
centrifuge Heraus Multifuge X3R	Thermo Scientific
CFX96 Real-Time System	Bio-Rad
eraser for imaging plates	GE Healthcare
gel documentation system Gel iX Imager	Intas
gel dryer Bio-Rad Model 583	Bio-Rad
heat block Eppendorf comfort	Eppendorf
horizontal electrophoresis systems PerfectBlue Mini S, M, L	Peqlab
hybridization oven HB-1000	UVP
imaging system Image Quant LAS 4000	GE Healthcare
incubator for bacterial plates (*H. pylori*), Galaxy 170R	Eppendorf
incubator for *H. pylori* liquid cultures, HERAcell 150i	Thermo Scientific
incubator for *E. coli*, HERAcell (Kendro)	Thermo Scientific
PCR engine, T3 thermocycler	Biometra
PhosphoImager Typhoon FLA 7000	GE Healthcare
photometer Ultrospec 3100 pro Cell Density Meter	GE Healthcare
power supplies peqPOWER E250, E300	Peqlab
Orbital Shaker NB-101S RC	N-BIOTEK
Retsch MM400 ball mill	Retsch
Rotamax 120	Heidolph
rotator – SB2 STUART	STUART
scanner for protein gels, HP Scanjet 7400c	HP
semi-dry electroblotter PerfectBlue SEDEC M	Peqlab
shaker 37 °C room, SM-30	Bühler
SORVALL centrifuge RC5B	Thieme Labortechnik
spectrophotometer NanoDrop 2000	Peqlab
tank electroblotter PerfectBlue Web S, M	Peqlab
thermal cycler MJ Mini	Bio-Rad
Thermo Mixing Block MB-102	BIOER
vaccum pump	KnF LAB
vertical electrophoresis systems PerfectBlue Twin S, ExW S, L	Peqlab
vertical sequencing gel system CBS SG-400-20	C.B.S. Scientific
Victor3 1420 multilabel counter	Perkin-Elmer
Vortex-Genie 2	Scientific Industries
waterbath, GFL	Hartenstein

6.1.2. Labware and consumables

Table 6.2: Glass/plastic ware and consumables.

Labware	Manufacturer
250 ml buckets for centrifuge	Thieme Labortechnik
Bio-spin disposable chromatography columns	BioRad
boxes (plastic), 20.5 x 20.5 cm or 9.5 x 20.5 cm	Hartenstein
boxes (metal), 10 x 21 cm	Hartenstein
cell culture flasks 25 cm^3	PAA, Corning
cell culture flasks 75 cm^3	PAA, Corning
Cellstar serological pipets (plastic) 5 ml, 10 ml, 25 ml, 50 ml	Greiner bio-one
chromatography columns	Biorad
cotton swaps	DELTALAB, Stein Labortechnik
cover slips	Hartenstein
dewar canister	Hartenstein
Erlenmeyer glass flasks 250 ml, 1l	DURAN, SIMAX
G-25, G-50 MicroSpin columns	GE Healthcare
Genomic-tip 100/G	Qiagen
Gilson pipets 10 µl, 20 µl, 200 µl, 1000 µl	Gilson
glass beads (0.1 mm) for cell lysis	Roth, Biospec
glass beads (2.85 – 3.35 mm) for plating of CFU/ml	Roth
glass bottles	Schott
glass test tubes and lids	Roth
hard-shell PCR plates 96-well WHT/WHT	Biorad
Hybond-XL membrane for nucleic acid transfer	GE Healthcare
imaging plates BAS-IP MS 2325, 2340	Fujifilm
imaging plates cassettes BAS 2325, 2340	Fujifilm
inoculation loops 10 µl	VWR
L-shape bacteriology loops	VWR
object slides	Hartenstein
PCR tubes 8 x 0.5 ml	Thermo Scientific
petri dishes	Corning
Phase Lock Gel (PLG)-tubes, 2 ml	5 Prime
Pipetboy accu-jet pro	BRAND
pipet tips	Sarstedt
PolyScreen PVDF Transfer Membrane	PerkinElmer
racks for PCR tubes / reaction tubes	Hartenstein
reaction tubes 1.5 ml, 2.0 ml	Sarstedt
reagent and centrifuge tubes 15 ml, 50 ml	Sarstedt
safe-lock tubes 1.5 ml, 2.0 ml	Eppendorf
spectrophotometer cuvettes	BRAND
sterile filters (0.20 µm pore size)	Sarstedt
tube holder 15 ml, 50 ml	Hartenstein
Whatman paper	ALBET LabScience
custom-made microarrays for *H. pylori*	Microarray Core Facility, MPI-IB, Berlin

6.1.3. Chemicals and reagents

Chemicals used in this study were purchased from Merck, Roth and Sigma.

Table 6.3: Chemicals, reagents, proteins and size markers.

Chemical/reagent/protein/size marker	Manufacturer
2 x gel loading buffer II (RNA)	Ambion and self-made
10 x RNA structure buffer	Ambion
1 x alkaline hydrolysis buffer	Ambion
2'2 dipyridyl (DPP, 99 %)	Sigma
4-aminobenzoic acid	AppliChem
acetic acid (100 %)	Roth
acetone	Roth
adenine	Sigma-Aldrich
agarose	Roth
albumin Fraktion V (BSA)	Roth
ammonia (NH_3, 25 %)	Roth
ampicillin sodium salt	Roth
amylose resin	New England Biolabs
Bacto Brain Heart Infusion	Becton, Dickinson and Company
BBL™ Brucella Broth	Becton, Dickinson and Company
chloramphenicol	Roth
cocarboxylase	Sigma
Cy3-dCTP	GE Healthcare
Cy5-dCTP	GE Healthcare
D(+)-glucose	Merck
DEPC-water	Roth
Difco-agar	BD
dimethyl sulfoxide (DMSO)	Roth
dithiothreitol (DTT)	Roth
donor horse serum (DHS)	Biochrom AG
ethanol	Roth
ethanol (absolute for analysis)	Merck
ethylenediaminetetraacetic acid disodium salt dehydrate (EDTA)	Roth
erythromycin	Biochemia
fetal calf/bovine serum (FCS/FBS)	Biochrom AG
formamide (99.5 %)	Roth
formaldehyde (37 %)	Roth
GC-agar Base	Oxoid
Gene Ruler 1 kb plus DNA ladder	Thermo Scientific
Gene Ruler 100 bp DNA ladder	Thermo Scientific
Genomic DNA buffer set	Qiagen
gentamicin sulfate	Roth
glycerol (99 %)	Sigma
glycine	Roth
GlycoBlue™	Ambion
guanine chloride	Sigma
heparin-agarose beads	Sigma
hydrochloric acid (HCl, 32 %)	Roth
hydrogen peroxide (H_2O_2, 30 %)	AppliChem
iron (III)-nitrate nonahydrate	AppliChem
iron (II)-sulfate	Roth
isopropanol	Roth
isopropyl-β-D-thiogalactopyranosid (IPTG)	Roth

Chemical/reagent/protein/size marker	Manufacturer
kanamycin sulfate	Roth
L-arginin-monohydrochloride	Sigma
L-cystine	Sigma
L-cysteine-hydrochloride monohydrate	Merck
L-glutamin	Merck
lead (II)-acetate	Roth, Fluka
maltose	Roth
magensium chloride	Roth
methanol	Roth
Midori Green	Nippon Genetics GmbH
milk powder (blotting grade)	Roth
nickel chloride (NiCl$_2$)	Roth
nicotinamide-adenine dinucleotide free acid	Sigma
nystatin	Sigma
PAGE Blue staining solution (Coomassie)	Thermo Scientific
PageRulerTM Plus Prestained Protein Ladder	Thermo Scientific
PBS	Gibco
phenol	Roth
polymyxin B	Sigma
pUC Marker Mix, 8	Thermo Scientific
random hexamers	Invitrogen
random primer/reaction buffer mix	BioPrime
rifampicin	Fluka
rifampicin (disks, 0.6 mm)	Oxoid
RNA Ladder High Range	Thermo Scientific
RNA Ladder Low Range	Thermo Scientific
Roti-Aqua-P/C/I	Roth
Roti-Hybri-Quick	Roth
Rotiphorese gel 40 (19:1)	Roth
Rotiphorese gel 40 (37.5:1)	Roth
sodium carbonate (Na$_2$CO$_3$)	Roth
sodium dodecyl sulfate (SDS), powder/pellets	Roth
sodium hydroxide (NaOH)	Roth
sodium periodate (Na$_5$IO$_6$)	Roth
sodium thiosulfate (Na$_2$S$_2$O$_3$)	Roth
trimethoprim	Sigma
Triton-X100	Sigma
TRIzol Reagent	Invitrogen
tRNA$^{f\text{-}Met}$	Sigma
Tween^{20}	Roth
uracil	Sigma
vancomycin sulfate	Roth
vitamin B12	Roth
Western Lightning chemoluminescence reagent	PerkinElmer, self-made
yeast RNA	Ambion
γ-^{32}P-ATP (^{32}P; 222TBq (6000Ci)/mmol 370MBq (10mCi)/ml)	Hartmann Analytic

6.1.4. Commercial kits

Table 6.4: Commercial kits.

Kit	Manufacturer
BioPrime DNA Labeling System	BioPrime
CycleReader™ DNA sequencing	Ambion
GeneJET™ Gel extraction	Fermentas
LightShift Chemiluminescent EMSA	Pearce
Masterpure DNA Purification	Epicentre
MEGAscript T7 *in-vitro* Transcription	Ambion
MinElute PCR Purification	Qiagen
NucleoSpin Plasmid	Macherey-Nagel
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel
Power SYBR GREEN RNA-to-CT™ 1-Step	Life Technologies
PureSystem	Cosmo Bio, New England Biolabs
QIAquick Gel Extraction	Qiagen
QIAquick PCR Purification	Qiagen
SequiTherm EXCEL™ II DNA Sequencing	Epicentre
SUPERscriptII Reverse Transcription	Invitrogen
SUPERscriptIII Reverse Transcription	Invitrogen

6.1.5. Enzymes

Table 6.5: Enzymes.

Enzymes	Manufacturer
30S ribosomal subunit in Tico buffer	K. Nierhaus, MPI for Molecular Genetics, Berlin
AffinityScript multi-temperature Reverse Transcriptase	Stratagene
Antarctic Phosphatase	New England Biolabs
AMV Reverse Transcriptase	New England Biolabs
Calf Intestinal Phosphatase (CIP, 10 u/µl)	New England Biolabs
Deoxyribonuclease (DNase I, 1 u/µl)	Thermo Scientific
DpnI (20 u /µl)	New England Biolabs
Klenow enzyme	BioPrime
Phusion High-Fidelity DNA polymerase (2 u/µl)	Thermo Scientific
Protease Inhibitor Cocktail (EDTA-free)	Roche
Proteinase K	Roth
Ribonuclease III (RNase III, 1.3 u/µl)	New England Biolabs
Ribonuclease A (RNase A)	Qiagen
Ribonuclease T1 (RNase T1, 1 u/µl)	Ambion
Shrimp Alkaline Phosphatase (SAP, 1 u/µl)	Thermo Scientific
SUPERaseIN RNase Inhibitor	Ambion
SUPERscriptII Reverse Transcriptase	Invitrogen
SUPERscriptIII Reverse Transcriptase	Invitrogen
T4 DNA Ligase (5 u/µl)	Thermo Scientific
T4 Polynucleotide Kinase (PNK, 10 u/µl)	Thermo Scientific
Taq DNA polymerase (5 u/µl)	New England Biolabs
lysozyme	Roth
diverse restriction enzymes (*NheI*, *XhoI*, *XbaI*, *ClaI* etc.)	New England Biolabs, Thermo Scientific
diverse chemicals (buffers & solutions)	Merck, Roth, Sigma

6.1.6. Antibodies

Table 6.6: Antibodies and antisera.

Antibody/antiserum	Origin	Dilution (3 % BSA/TBS-T)	Manufacturer
anti-CagA antiserum	rabbit	1:1,000	Rainer Haas, Max-von-Pettenkofer-Institut, Munich, Germany
anti-HP1043 antiserum	rabbit	1:1,000	Dagmar Beier, Biocenter, University of Würzburg, Würzburg, Germany
anti-TlpA22 antiserum	rabbit	1:2,000	Karen Ottemann, University of California, Santa Cruz, USA
monoclonal anti-CagA antibody (AK299)	mouse	1:5,000	Abcam
monoclonal anti-FLAG M2 antibody	mouse	1:1,000	Sigma-Aldrich
monoclonal anti-GFP antibody	mouse	1:1,000	Roche
monoclonal anti-Lewis x antibody	mouse	1:500 – 1:1,000	Calbiochem
monoclonal anti-S1 antibody	rabbit	1:10,000	Mathias Springer, IBPC Paris, France
polyclonal anti-GroEL antiserum	rabbit	1:10,000	Sigma-Aldrich
ECL anti-mouse IgG, HRP-conjugate	goat	1:10,000	GE Healthcare
ECL anti-rabbit IgG, HRP-conjugate	donkey	1:10,000	GE Healthcare

6.1.7. Oligonucleotides

Table 6.7: Synthetic oligonucleotides.
Sequences are given in 5' → 3' direction; P~ denotes a 5' monophosphate.

Name	Sequence 5' → 3'	Description
CSO-0003	GAAAGGAGGGGGAGGT	Northern blot probe for RepG
CSO-0017	gtttttTCTAGAGATCAGCCTGCCTTTAGG	RepG cloning
CSO-0018	gtttttCTCGAGCTTAGCGCTTAATGAAACGC	RepG cloning
CSO-0035	GGTGTTCTATAATTTCAGCTTGTATGCTTTATAACTATGGATTAAACACTTTT	Cloning of *tlpB*
CSO-0036	CTTTTTTAATAAACTCCCCCTGATTACTTATTAAATAATTTATAGCTATTGAAAAGAG	Cloning of *tlpB*
CSO-0037	TTTAATCCATAGTTATAAAGCATACAAGCTGAAATTATAGAACACCC	Cloning of *tlpB*
CSO-0038	GCTATAAATTATTTAATAAGTAATCAGGGGGAGTTTATTAAAAAAG	Cloning of *tlpB*
CSO-0039	gtttttCTCGAGTCTCAAAATCCGCTGAAATCT	Cloning of *tlpB*
CSO-0040	gtttttTCTAGATCAGTTGCAACCAGGAGATT	Cloning of *tlpB*
CSO-0045	AAGATGACGACGATAAATAGTAAATGCTTTATAACTATGGATTAAACACTTTT	Cloning of *tlpB*::3xFLAG
CSO-0046	TTTAATCCATAGTTATAAAGCATTTACTATTTATCGTCGTCATCTTTGT	Cloning of *tlpB*::3xFLAG
CSO-0050	GCATGAAAGATTCCTCAACCAC	Verification of *tlpB*::3xFLAG
CSO-0051	TGTCACTTATATTTACAAGTTCGCT	Verification of *tlpB* deletion
CSO-0053	ACTGCCAGGTTCGGAATGG	Northern blot probe for *Helicobacter mustelae* 5S RNA
CSO-0065	P~GACTACAAAGACCATGACGGT	Cloning of 3xFLAG (diverse proteins)
CSO-0074	gtttttGAATTCTATTCCCTCCAGGTACTAAAACA	Cloning of 3xFLAG (diverse proteins)
CSO-0075	TCCTTCACAAAGAAGGGG	Cloning of 3xFLAG (diverse proteins)
CSO-0076	TTCTTTTGTTTTAGGAGCGATATA	Cloning of *tlpB*::3xFLAG
CSO-0078	gtttttTCTAGACTATAAATGCTCGCTTCAAATTT	Cloning of 3xFLAG (diverse proteins)
CSO-0079	gtttttGAATTCAGACAAAAGACTAAGGATTAATCAGT	Cloning of *tlpB*::3xFLAG
CSO-0080	ACCGCTAAGATTGAAGGGTC	RepG-2SL cloning
CSO-0081	P~TTGTAATTATACAATAAAATGGTGATTTTTAG	RepG-2SL cloning

Name	Sequence 5' → 3'	Description
CSO-0083	gtttttATGCATCAAGCCCTTGATTATTGGTTG	Primer extension
CSO-0126	gttttttGCTAGCTGAAGAAAACATCATTTATCTCCT	Cloning of GFP fusion (*tlpB*)
CSO-0138	AAGATTGAAGGGTACGCCCTTTGTCTTGGCGG	RepG-ΔCU cloning
CSO-0139	CAAGACAAAGGGCGTACCCTTCAATCTTAGCGGTG	RepG-ΔCU cloning
CSO-0140	AGGGTCACCTCCGCCTCCTTTCCCTTTGTCTT	RepG-1xG* cloning
CSO-0141	AGGGAAAGGAGGCGGAGGTGACCCTTCAATCT	RepG-1xG* cloning
CSO-0142	TTGAAGGGTCACGTCCGCCTGCTTTCCCTTTGTCTTGGC	RepG-3xG cloning
CSO-0143	ACAAAGGGAAAGCAGGCGGACGTGACCCTTCAATCTTAGCG	RepG-3xG cloning
CSO-0146	gtttttATCGATGTATGCTCTTTAAGACCCAGC	Cloning of GFP fusion
CSO-0147	gtttttCATATGCTCGAATTCAGATCCACGTT	RepG cloning in *H. pylori* G27
CSO-0205	AATTACAACAGTACTGCGATGAGT	RepG cloning
CSO-0206	AATCTCACGCCAAGCATTT	RepG cloning
CSO-0207	AGTTCTGATTTCATGCCCTT	RepG cloning, verification
CSO-0208	gtttttCTCGAGGCGCGCACACTGAAGA	Cloning of *tlpB*::3xFLAG
CSO-0209	gtttttGAATTCTTACTTATTAAATAATTTATAGCTATTGAAAAGAG	Cloning of *tlpB*::3xFLAG
CSO-0210	gtttttGAATTCAGGGGGAGTTTATTAAAAAAG	Cloning of *tlpB*::3xFLAG
CSO-0211	gtttttTCTAGATCTCAAAATCCGCTGAAATCT	Cloning of *tlpB*::3xFLAG
CSO-0245	P~AGTTTTAAACAAATTCACTTGTTTGTC	Cloning of *tlpB*::3xFLAG
CSO-0263	GACTGATGTCATCAGCGGT	Cloning of *tlpB* in *H. pylori* X47-2AL
CSO-0277	TTTATCCACCACCACCATATAAA	Cloning of *tlpB*
CSO-0278	gtttttttaatacgactcactatagGTTCAAAGACATGAATTGATTACTC	*In-vitro* transcription
CSO-0284	gtttttGCTAGCAGCCACTTGAAGATTATTGATAAAT	Cloning GFP fusion
CSO-0291	gtttttCTCGAGTTAGGCATTTTATAATAAGTGTAGCCT	Cloning of *tlpB*
CSO-0292	gtttttGGATCCGCAGCTCCATCAGCAAAAG	Cloning of RepG
CSO-0293	gtttttGAATTcTCCAACGTCATCTCGTTCT	Cloning of *tlpB*-HP0102
CSO-0294	gtttttGGATCCTTTATTATTTTATCTTTAAGCCTAACTTAA	Cloning of *tlpB*
CSO-0295	gtttttGAATTCTAAAAAAATTTTATTTAACTTCACTCTCTT	Cloning of *tlpB*
CSO-0306	gtttttGGATCCGATCGGGCTTTTTTCAATATT	Cloning of *tlpB*
CSO-0308	gtttttGGATCCTGCTTTATAACTATGGATTAAACACTT	Cloning of *tlpB*
CSO-0309	gtttttGAATTCTTACTTATTAAATAATTTATAGCTATTGAAAAGA	Cloning of *tlpB*
CSO-0313	gtttttGGATCCTAGAGATCCGCCATATTGTGT	Cloning of HP0102/RepG
CSO-0314	CTCATTTTTCGGGGGGCGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGG	Cloning of *tlpB*; compensatory base-pair exchange
CSO-0315	TTCTAAATGCACCCCCGCCCCCCGAAAAATGAGTGGCACAAA	Cloning of *tlpB*; compensatory base-pair exchange
CSO-0316	CTCATTTTTCGGCGGGCGGGCGTGCATTTAGAAGCTAAACTCTAAAATTAGGG	Cloning of *tlpB*; compensatory base-pair exchange
CSO-0317	TTCTAAATGCACGCCCGCCCGCCGAAAAATGAGTGGCACAAA	Cloning of *tlpB*; compensatory base-pair exchange
CSO-0318	P~TGCATTTAGAAGCTAAACTCTAAAATTAG	Cloning of *tlpB*
CSO-0319	GAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0424	gtttttCATATGCAAGCCCTTGATTATTGGTTG	RepG cloning in *H. pylori* X47-2AL
CSO-0425	gtttttCATATGCAAGCCCTTGGTTATTGGTTG	RepG cloning in *H. pylori* G27
CSO-0426	gtttttATCGATCAAGCGGTTAAAAATAAGAACTAAAC	RepG cloning in *H. pylori* X47-2AL
CSO-0427	gtttttATCGATCAAGCGGTTAAAAACAAGGAGT	RepG cloning in *H. pylori* G27
CSO-0428	P~TTACTATTTATCGTCGTCATCTTT	Cloning of *tlpB*::3xFLAG
CSO-0429	TCAGGGGGAGTTTATTAAAAAAG	Cloning of *tlpB*::3xFLAG
CSO-0430	P~TTGTTTGTTCTTTTGTTTCGTTTT	Cloning of *tlpB*
CSO-0431	ATTCTTATTATACAACAATATCAAGCATT	Cloning of *tlpB*
CSO-0440	gttttttGTCGACGTTGGCGTATAACATAGTATCGA	Cloning GFP fusion
CSO-0441	gttttttGCGGCCGCGGGAGTTAACTGCAGGTCTG	Cloning GFP fusion
CSO-0442	gtttttGTCGACGTATGCTCTTTAAGACCCAGC	Cloning GFP fusion
CSO-0443	gtttttGCGGCCGCCTCGAATTCAGATCCACGTT	Cloning GFP fusion
CSO-0448	CCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0449	CCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0450	CCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0451	CCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0452	CCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*

Name	Sequence 5' → 3'	Description
CSO-0453	CCCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0454	CCCCCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0455	CCCCCCCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0456	CCCCCCCCCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0457	CCCCCCCCCCCCCCCCCCCGAAAAATGAGTGGCACAAAAC	Cloning of *tlpB*
CSO-0511	CACCATCAGGGTCTGGTCACTGTTTGTTCTTTTGTTTCGTT	MS2-tagging of *tlpB*
CSO-0523	CACCATCAGGGTCTGGTCACTGTTTGTTCTCTGTTTCGTTTTC	MS2-tagging of *tlpB*
CSO-0581	gttttttATCGATTTATTATTTTATCTTTAAGCCTAACTTAA	Cloning of GFP fusion (*tlpB*)
CSO-0590	gttttttATCGATGATCGGGCTTTTTTCAATAT	Cloning of GFP fusion
CSO-0591	gttttttGCTAGCTTTAAGAAGGAGATATACATATGAGTAA	Cloning of GFP fusion
CSO-0683	gttttttGCTAGCAGTAAAGGAGAAGAACTTTTCACTGGA	Cloning of GFP fusion
CSO-0867	AGCCATTTATGAACGCTACGGC	Quantitative RT-PCR (HP0102)
CSO-0868	CACCACAAACAAGCGGATGAT	Quantitative RT-PCR (HP0102)
CSO-0869	gttttttGGAtccAACCCTTTTTTAATAAACTCC	Cloning of HP0102
CSO-0870	gttttttCTCGAGTTATCAATGATATTGCCGATCAAAC	Cloning of HP0102
CSO-0871	gttttttGAAttcCATGCTTTTTAAAATACCCTAG	Cloning of *tlpB*-HP0102
CSO-0872	gttttttTCTAGATTTGGATACAACTATTCGCAAGC	Cloning of HP0102
CSO-0873	gtttttttCTAGAGGCATCAAATAAAACGAAA	Cloning in pJV572.1
CSO-0874	gttttttCTCGAGGTGAAGACGAAAGG	Cloning in pJV572.1
CSO-0882	P~CAAAACTAACGGCGCGACAA	Cloning of *dcuA*::3xFLAG
CSO-0883	gttttttCTCGAGATCGCTTTTGTTAGCCCGGT	Cloning of *dcuA*::3xFLAG
CSO-0884	gttttttGAATTCATTTTATCACCAACGATAAAAAGCT	Cloning of *dcuA*::3xFLAG
CSO-0885	gttttttTCTAGaAAGCAAAAGATGGGGGTTAAG	Cloning of *dcuA*::3xFLAG
CSO-0886	TCGCCCCCTAAAATCAAAG	Verification of *dcuA*::3xFLAG
CSO-0888	CGGTGATTGTGGAAGTGAGC	Quantitative RT-PCR (*dcuA*)
CSO-0889	TGCCGCTCATAAACACCACC	Quantitative RT-PCR (*dcuA*)
CSO-0890	P~CAATAACGCTTTTAGAGCGTCT	Cloning of *trx2*::3xFLAG
CSO-0891	gttttttCTCGAGAACGCTGGAGTGTTGGATTTAA	Cloning of *trx2*::3xFLAG
CSO-0895	TCGCTCATCAAGCGGTGGTG	Quantitative RT-PCR (*trx2*)
CSO-0896	GGCTAAATTTTCCATGATCGGC	Quantitative RT-PCR (*trx2*), Northern blot probe
CSO-0899	P~TTTTCTAAAGTTTTGCGCTAAGTG	Cloning of HP1181::3xFLAG
CSO-0900	gttttttCTCGAGAGCGAGCGTTATGGCTGAAA	Cloning of HP1181::3xFLAG
CSO-0901	gttttttGAATTCAACAATTAAGGATCAAAAATGGCC	Cloning of HP1181::3xFLAG
CSO-0902	gttttttTCTAGAATCAAAGGGCGGATCACCAA	Cloning of HP1181::3xFLAG
CSO-0903	ATCTTGGTGTATGTGCCTGG	Verification of HP1181::3xFLAG
CSO-0904	TCATTCAAGGCATGGGGGCT	Quantitative RT-PCR (HP1181)
CSO-0905	CTTTGGTGCGCTCTTCTTCTT	Quantitative RT-PCR (HP1181)
CSO-0906	GATTTGTGCCAAGAGCAAGGC	Quantitative RT-PCR (HP1182)
CSO-0907	AACGAGCTTTTTTCACGGCGTT	Quantitative RT-PCR (HP1182)
CSO-0913	ATTCAGCCCCGGTGATTAGG	Quantitative RT-PCR (*secD*)
CSO-0914	GCCCCACTCCTTAAAGCGAT	Quantitative RT-PCR (*secD*)
CSO-0924	TGATGCGATTGTAGTGGCTAG	Quantitative RT-PCR (*xseA*)
CSO-0925	CCAAATACAGAGCATCAGCGA	Quantitative RT-PCR (*xseA*)
CSO-0932	GGCATTGATGAAGCGGGTAG	Quantitative RT-PCR (*rnhB*)
CSO-0933	GGCTGAGCTTCTTGCTGTCT	Quantitative RT-PCR (*rnhB*)
CSO-0938	gttttttGCTAGCTTTTTTTTAAGGCAGTTATTGGTGC	Cloning of HP0102/RepG
CSO-0939	gttttttGCTAGCTCCAACGTCATCTCGTTCT	Cloning of HP0102/RepG
CSO-0940	gttttttGGATCCCAGCGATTTCTTCTAATTCTGT	Cloning of HP0102/RepG
CSO-0942	gttttttGAATTCTCAGGGGGAGTTTATTAAAAAAG	Cloning of *tlpB* in *H. pylori* X47-2AL
CSO-0956	P~GAACACATATTGATACCCAATGT	Cloning of HP1057::3xFLAG
CSO-0957	gttttttCTCGAGACTTAAAAAGGGCGAACGCT	Cloning of HP1057::3xFLAG
CSO-0958	gttttttGAATTCTCTTCAAAAGAGACTTGGTGG	Cloning of HP1057::3xFLAG
CSO-0959	gttttttTCTAGAGCGCTAATCCCGTTAGCA	Cloning of HP1057::3xFLAG
CSO-0960	TATGGTCTTTAATGGGGTTTTATG	Verification of HP1057-1055::3xFLAG
CSO-0961	TCGGTTATGCTGAATGGGGAT	Quantitative RT-PCR (HP1057)
CSO-0962	ACCCCCAAACCTAAACCCATA	Quantitative RT-PCR (HP1057)
CSO-0963	TGCGTTAGGGATATTTGGAGG	Quantitative RT-PCR (HP1056)
CSO-0964	AAGCGTTAGGCTGTGAATACC	Quantitative RT-PCR (HP1056)

Name	Sequence 5' → 3'	Description
CSO-0965	TTGGGACACGATTTTATGGGG	Quantitative RT-PCR (HP1055)
CSO-0966	TACAGGCTGCTTGATTGCATC	Quantitative RT-PCR (HP1055) 26695
CSO-0967	TACAGGTTGCTTGCTTGCATT	Quantitative RT-PCR (HP1055) G27
CSO-0970	gttttttGAATTCAGGCTAAAATGTTGAAATTTAAATATG	Cloning of HP1056::3xFLAG
CSO-0971	gttttttTCTAGACTAAAGGCATATCAAACAATAAATC	Cloning of HP1056::3xFLAG
CSO-0972	P~AAAAATAAACGCATAATTCACTAATAAAA	Cloning of HP1055::3xFLAG
CSO-0973	P~AAAAATGAACGCATAATTCACTAATAAAA	Cloning of HP1055::3xFLAG
CSO-0974	gttttttCTCGAGATTAACGCTCCTGTTTCTATGAT	Cloning of HP1055::3xFLAG
CSO-0975	gttttttGAATTCGGCTTGATCTTGGAGTTAAG	Cloning of HP1055::3xFLAG
CSO-0976	gttttttTCTAGAGACAAATTCTTATCCAAAGCTTC	Cloning of HP1055::3xFLAG
CSO-1022	TTTAATCCATAGTTATAAAGCATCATAGATAAAACTCATTATT ATTTTAAT	Cloning of *fur*
CSO-1023	GAATTTATCGCCAGAAGGCT	Cloning of *fur*
CSO-1024	AATAATAATGAGTTTTATCTATGATGCTTTATAACTATGGATT AAACAC	Cloning of *fur*
CSO-1025	GCTATAAATTATTTAATAAGTAACATGACATGAAAATGTTTGT GTG	Cloning of *fur*
CSO-1026	CACACAAACATTTTCATGTCATGTTACTTATTAAATAATTTAT AGCTATTGAAAA	Cloning of *fur*
CSO-1027	GCTGTAGAGTTGCCTGGA	Cloning of *fur*
CSO-1028	TCTTTAAGAGGGAGCGATGA	Verification of Δ*fur*
CSO-1173	GGTAGTGGTTTTTGTGTGATGG	Quantitative RT-PCR (6S RNA)
CSO-1174	CCAGATGACCGCTACTTTTACA	Quantitative RT-PCR (6S RNA)
CSO-1283	gttttttCTCGAGCCGATCCCTTAAAACGGCGC	Cloning of *trx2*
CSO-1284	gttttttTCTAGATAATACCACAAATTCGCTCATTTTTAGCTTAATT AAAG	Cloning 3xFLAG::*trx2*
CSO-1291	GTGTGCCATTTTTCTAACCATTCCATCGC	Quantitative RT-PCR (*cheV2*)
CSO-1292	CCCTCTTCATTAATGCCTTGTTTGTCCC	Quantitative RT-PCR (*cheV2*)
CSO-1333	TTGTTGTTTGTCGTAGGGGTG	Quantitative RT-PCR (HP0580)
CSO-1334	GGGCTTAAAGAGAGTTTTCGC	Quantitative RT-PCR (HP0580)
CSO-1335	CCGGCTCTATTTCGCAAGTG	Quantitative RT-PCR (*radA*)
CSO-1336	CTCCCTTCTTTAGTGATATGACC	Quantitative RT-PCR (*radA*)
CSO-1337	TATGCGGCGATTTTACGCTATAA	Quantitative RT-PCR (*rnhA*)
CSO-1338	TTCATTGAGCGCTCTTAATTCCA	Quantitative RT-PCR (*rnhA*)
CSO-1339	TCAAGCGCTTTTGAGGCTTTAAT	Quantitative RT-PCR (*rnc*)
CSO-1340	TCCAAACGCTTGTAAGCGCG	Quantitative RT-PCR (*rnc*)
CSO-1341	TTGGGGAATCGCATGGGGAT	Quantitative RT-PCR (HP0663)
CSO-1342	CGGCGCTTCATTTCATTTTCTAAT	Quantitative RT-PCR (HP0663)
CSO-1345	GCGGGCAAACCGCTTTGAAT	Quantitative RT-PCR (*carB*)
CSO-1346	CCTGCCTGTCTTCGCCTTTT	Quantitative RT-PCR (*carB*)
CSO-1359	gttttttCTCGAGTTTGGATACAACTATTCGCAAGC	Cloning of *tlpB*-HP0102
CSO-1364	gttttttGAATTCGCCTTGCCGGTAGGGAAAATGC	Cloning of 3xFLAG::*trx2*
CSO-1365	gttttttGGATCCCCCAACTTTGAAGTATAGCATAACTCG	Cloning of *trx2*
CSO-1546	P~CAATAAATAGTTGTAACTAATATAATAAATATTAGTGC	Cloning of HP1056::3xFLAG
CSO-1547	gttttttCTCGAGATGAATCTCTCTTGTTCCAACGAATG	Cloning of HP1056::3xFLAG
CSO-1737	gttttttTCTAGATTATCAATGATATTGCCGATCAAAC	Cloning of HP0102
CSO-1738	TGTCAAAGAGAGCAAGGATGC	Verification of ΔHP0102
CSO-1739	gttttttGGAtccAACACGCCGTGATCACAGAA	Cloning of HP0102
CSO-1740	gttttttCATATGTAAGATTTTTTATTTTATTTTTAAGCCTAAC	Cloning of HP0102
CSO-1741	gttttttATCGATAGTTAGGGATTGAAATACCCTTA	Cloning of HP0102
CSO-1742	P~TCAGGGGGAGTTTATCAAAAAAG	Cloning of HP0102
CSO-1743	AGCTAAAATTATAGAACACCCTTTT	Cloning of HP0102
CSO-1745	CATTTATCTCCTATAAATCATTTTTAAGT	Cloning of *tlpB* in *H. pylori* X47-2AL
CSO-1748	ATCTTTATAATCACCGTCATGGTCTTTGTAGTCCATGATAATCC TTATTTTTTATTTCCCCCAACTTT	Cloning 3xFLAG::*trx2*
CSO-1749	ATCTTTATAATCACCGTCATGGTCTTTGTAGTCCATGATAATCC TTGTTTTTTATTTCCCCCAACTTT	Cloning 3xFLAG::*trx2*
CSO-1795	P~CATGATATCGACTACAAAGATGACGACGATAAATCAGAAAT GATTAACGGGAAGAATTACGCA	Cloning 3xFLAG::*trx2*

Name	Sequence 5' → 3'	Description
CSO-1796	P~CATGATATCGACTACAAAGATGACGACGATAAATCAGAAATTATTAATGGGAAGAATTACGCA	Cloning 3xFLAG::*trx2*
CSO-1803	gttttttGCTAGCACACGCCGTGATCACAGAAA	Cloning of GFP fusion (HP0102)
CSO-1813	gttttttGGATCCTTTTATGGATAATTTTTAAAATCATTTG	Cloning of HP0102
CSO-1814	gttttttGCTAGCTATTCCCTCCAGGTACTAAAACA	Cloning of HP0102
CSO-1984	GGCTAACACGACCAGCATGA	Cloning of GFP fusion (HP0102)
CSO-1985	P~TAATCAGGGGGAGTTTATTAAAAAAG	Cloning of GFP fusion (HP0102)
CSO-2052	P~AGTTTATTAAAAAAGGGTTGGATTG	Cloning of GFP fusion (HP0102)
CSO-2053	TGATTAAGTTTTAAACAAATTCACTTGT	Cloning of GFP fusion (HP0102)
CSO-2055	P~TCAGGGGGAGTTTATTAAAAAAGG	Cloning of GFP fusion (HP0102)
CSO-2056	AGCTGAAATTATAGAACACCCTTTTT	Cloning of GFP fusion (HP0102)
CSO-2061	P~ATTTAAGTTTATTAAAAAAGGGTTGGATTG	Cloning of GFP fusion (HP0102)
CSONIH-0033	TCAAAGCCACTAGTAAGTCTTACTT	Verification oligo for insertion of *rpsL-erm* cassette
JVO-0155	CCGTATGTAGCATCACCTTC	Cloning of GFP fusion (Urban & Vogel, 2007)
JVO-0352	gtttttTCTAGATGCGCTCAATCAAGCT	Control-PCR DNase I digestion
JVO-0353	gttttttCTCGAGGCTTTGGAGATGATGAGTTAG	Control-PCR DNase I digestion
JVO-0485	TCGGAATGGTTAACTGGGTAGTTCCT	Northern blot probe for *Helicobacter pylori* 5S rRNA
JVO-1049	TAACATCACCATCTAATTCAAC	Northern blot probe for *gfpmut3* fusions
JVO-2134	AAACCATAAGGAATGGTTGGAT	Northern blot probe for RepG
JVO-4201	gtttttttttaatacgactcactataggTCGTACACCATCAGGGTAC	MS2-tagging
JVO-4202	GTTTTTTTTTAATACGACTCACTATAGG	MS2-tagging
JVO-4203	GTGACCAGACCCTGATGG	MS2-tagging
JVO-4204	CACCATCAGGGTCTGGTCACATCCAACCATTCCTTATGG	MS2-tagging of RepG
JVO-5069	CTTCACGCCCCTTGTAAATA	Verification of *repG* deletion mutant
JVO-5070	GATAAGGTTTAGCGATGTAATCGT	RepG cloning
JVO-5072	CGTTTCTTGACACGCTTAATT	RepG cloning
JVO-5125	gtttttttttaatacgactcactataggATCCAACCATTCCTTATGGTT	*In-vitro* transcription RepG
JVO-5126	AAAACAACCGCCAAGACA	*In-vitro* transcription RepG
JVO-5127	gtttttttttaatacgactcactataggTGTTTGTTCTTTTGTTTCGTT	*In-vitro* transcription *tlpB*
JVO-5142	GACTACAAAGACCATGACGG	Cloning of 3xFLAG (diverse proteins)
JVO-5143	TTACTATTTATCGTCGTCATCTTT	*In-vitro* transcription *tlpB*
JVO-5257	TATAGGTTTTCATTTTCTCCCAC	Verification of *repG* deletion mutant
JVO-5267	ACGGGGTGGTATTGTTTGAT	Quantitative RT-PCR (*tlpB*)
JVO-5268	AAGTGTAGCCTCCCCCTTTT	Quantitative RT-PCR (*tlpB*)
JVO-5702	GTATTTCACACCGGGTAAATCC**C**TAACC**C**TACCCCCACG	Antisense primer for amplification of *rpsL* fragment containing the Lys43Arg mutation (AAA129AGA) *str^R* allele in primer rpsL1 and the Lys88Arg mutation (AAG263AGG) *str^R* allele in rpsL2 in 26695, P12 and G27 (Dailidiene *et al.*, 2006)
JVO-5703	CTAGGGTTTATACGACTACCCCTA**G**AAAGCCTAACTCG	Sense primer for amplification of *rpsL* fragment containing the Lys43Arg mutation (AAA129AGA) *str^R* allele in primer rpsL1 and the Lys88Arg mutation (AAG263AGG) *str^R* allele in rpsL2 in 26695, P12 and G27 (Dailidiene *et al.*, 2006)
JVO-5704	AGAAGCCAGTATCGCTATGA	Sense oligo for verification of *rpsL* (HP1197) mutation in 26695, G27
JVO-5953	CCCTAAACCTAAAAGAGCGG	*In-vitro* transcription *tlpB*, toeprint
MO-1	TTTATGGATAATTTTTAAAATCATTTGCTAAAAAT	Sense oligo for amplification of the specific SC competitor DNA fragment used in EMSA
pZE-A	GTGCCACCTGACGTCTAAGA	Colony PCR and sequencing of pZE12-*luc* derived plasmids

continued on next page

Name	Sequence 5' → 3'	Description
pZE-*XbaI*	TCGTTTTATTTGATGCCTCTAGA	Colony PCR and sequencing of pZE-derived plasmids
***rpoA*-5**	TCATGAGTTTGACTCTTTAAGG	Antisense primer for the amplification of the *rpoA* competitor DNA fragment used in EMSA
***rpoA*-3**	GATTGACGATTTCTATCTGC	Sense primer for the amplification of the *rpoA* competitor DNA fragment used in EMSA
S5490-5	AAACCATAAGGAATGGTTGGATT	Antisense primer for the amplification of the full length *repG* promoter probe used in EMSA
S5490-3btn	Btn-AAAGTTATCGCTTGGTTAAAATCAA	Biotinylated sense oligo for the amplification of the full length *repG* promoter probe used in EMSA
S5490-3	AAAGTTATCGCTTGGTTAAAATCAA	Sense oligo for the amplification of the full length *repG* promoter probe used in EMSA
USC-1	AACGCTCAAAGGCATTTTTTACAACCAATAATCAAGGGCTTGATTTTAACCAAGCGAT	59mer antisense oligonucleotide to generate double stranded USC competitor fragment used in EMSA
USC-2	TATCGCTTGGTTAAAATCAAGCCCTTGATTATTGGTTGTAAAAATGCCTTTGAGCGTT	59mer sense oligonucleotide to generate double stranded USC competitor fragment used in EMSA

6.1.8. Plasmids

Plasmids constructed/used in this thesis are listed in Appendix, Table 13.3. Detailed descriptions for the construction of selected plasmids are given in Chapter 6.2 (Microbiological methods).

6.1.9. Bacterial strains

Bacterial strains constructed/used in this thesis are listed in Appendix, Table 13.4. Detailed descriptions for the generation of selected *Helicobacter pylori* strains are given in Chapter 6.2 (Microbiological methods).

6.1.10. Media, buffer and supplements

6.1.10.1. Media and stocks

Lennox Broth (LB) medium: 10 g tryptone or peptone
5 g yeast extract
5 g NaCl
ad 1 l H_2O

LB-agar plates: LB medium (see above)
1.5 % (w/v) Difco-agar

Brain Heart Infusion (BHI):		37 g	Bacto Brain Heart Infusion
		ad 1 l H_2O	
	after autoclaving, supplement with		
		10 % (v/v)	FBS (heat-inactivated)
		10 µg/ml	vancomycin
		5 µg/ml	trimethoprim
		1 µg/ml	nystatin

GC-agar plates:		36 g	GC-agar
		ad 1 l H_2O	
	after autoclaving, supplement with		
		10 % (v/v)	DHS (heat-inactivated)
		1 %	vitamin mix (see below)
		10 µg/ml	vancomycin
		5 µg/ml	trimethoprim
		1 µg/ml	nystatin

Vitamin mix for GC-agar plates:	**(A)**	200 g	D(+)-glucose
		20 g	L-glutamine
		52 g	L-cystein-hydrochloride monohydrate
		0.2 g	cocarboxylase
		0.04 g	iron (II)-nitrate nonahydrate
		0.006 g	thiamine hydrochloride
		0.026 g	4-aminobenzoic acid
		0.5 g	nicotinaminde-adenine dinucleotide free acid
		0.02 g	vitamin B12
		ad 1 l H_2O	

Vitamin mix for GC-agar plates:	**(B)**	2.2 g	L-cysteine
		2 g	adenine
		0.060	guanine chloride
		0.3 g	L-arginine monohydrochloride
		1 g	uracil
		ad 600 ml H_2O and add 30 ml 32 % HCl	
		mix solution A+B	
		ad 2 l H_2O	

Brucella Broth (BB):		28 g	BBL™ Brucella Broth
		ad 1 l H_2O	
	after autoclaving, supplement with		
		10 % (v/v)	FBS (heat-inactivated)
		10 µg/ml	vancomycin
		5 µg/ml	trimethoprim
		1 µg/ml	nystatin

BB soft-agar plates (0.4 %):

	28 g	BBL™ Brucella Broth
	4 g	Difco-agar
	ad 1 l H$_2$O	

after autoclaving, supplement with

	5 % (v/v)	FBS (heat-inactivated)
	10 µg/ml	vancomycin
	5 µg/ml	trimethoprim
	1 µg/ml	nystatin

SOB medium:

20 g tryptone
5 g yeast extract
0.5 g NaCl
ad 800 ml H$_2$O
add 10 ml 250 mM potassium chloride
adjust to pH 7.0 (using NaOH)
ad 1 l H$_2$O

SOC medium (transformation *E. coli*):

1 l SOB medium
add 5 ml magnesium chloride
add 20 ml 1 M glucose

Superbroth medium (competent *E. coli*):

35 g tryptone
30 g yeast extract
5 g NaCl
ad 1 l H$_2$O

6.1.10.2. Media supplements

Table 6.8: Antibiotics and media supplements.

H. pylori

Antibiotic	Solvent	Stock	Working concentration
vancomycin	H$_2$O	10 mg/ml	10 µg/ml
trimethoprim	DMF	2.5 mg/ml	5 µg/ml
nystatin	DMF	10 mg/ml	1 µg/ml
kanamycin	H$_2$O	20 mg/ml	20 µg/ml
chloramphenicol	100 % EtOH	8 mg/ml	8 µg/ml
erythromycin	DMF	10 mg/ml	10 µg/ml
gentamicin	H$_2$O	10 mg/ml	10 µg/ml
streptomycin	H$_2$O	10 mg/ml	10 µg/ml
rifampicin	DMSO	50 mg/ml	500 µg/ml
rifampicin	DMSO		10 µg/ml (commercial disks)
polymyxin B	H$_2$O	30.000 u/ml	300 u

Supplement	Solvent	Stock	Working concentration
FBS			heat-inactivated, 10 % (v/v)
DHS			heat-inactivated, 10 % (v/v)
IPTG	H$_2$O	1 M	1 mM/ml

E. coli

Antibiotic	Solvent	Stock	Working concentration
ampicillin	H$_2$O	100 mg/ml	100 µg/ml
kanamycin	H$_2$O	20 to 50 mg/ml	20 to 50 µg/ml
chloramphenicol	100 % EtOH	20 mg/ml	20 µg/ml

6.1.10.3. Buffers and solutions

30:1 ethanol/sodium acetate (pH 6.5): (for RNA precipitation)	30 parts of 100 % ethanol 1 part of 3 M sodium acetate (pH 6.5)	

10 x DNA loading dye:

	1.66 ml	1 M Tris-HCl (pH 7.5)
	12 ml	0.5 M EDTA (pH 8.0)
	0.05 g	bromophenol blue
	0.05 g	xylene cyanol
	60 ml	glycerol
	ad 100 ml H$_2$O	

2 x gel loading buffer II (RNA, GL II):

	0.13 % (w/v)	SDS
	18 µM	EDTA (pH 8.0)
	95 %	formamide
	0.025 % (w/v)	bromphenol blue
	0.025 % (w/v)	xylene cyanol

5 x native sample buffer (gel-shifts):

	50 %	glycerol
	0.02 %	bromophenol blue
	0.5 x	TBE buffer

5 x protein loading dye:

	10 g	SDS pellets
	31.3 ml	1 M Tris-HCl (pH 6.8)
	50 ml	glycerol
	add 2.5 ml of 2 % (w/v) bromophenol blue	
	ad 100 ml H$_2$O	

1 x protein loading dye:

	40 ml	5 x protein loading dye
	10 ml	2 M DTT
	ad 200 ml H$_2$O	

agarose gel electrophoresis solution:

	X % (w/v)	agarose in 1 x TAE/TBE buffer

AMV reaction buffer (1 x, primer extension):

	50 mM	Tris-HCl (pH 8.3)
	75 mM	potassium
	8 mM	magnesium
	10 mM	DTT

chemiluminescence solution A:	0.1 M	Tris-HCl (pH 8.6)
	0.025 % (w/v) luminol	
chemiluminescence solution B:	0.11 % (w/v)	*p*-coumaric acid (in DMSO)
colorless gel-loading solution (in-line probing):	10 M	urea
	1.5 mM	ETDA (pH 8.0)
Developing solution (silver-staining protein):	60 g	sodium carbonate
	4 mg	sodium thiosulfate (160 µl $Na_2S_2O_3$ stock/l)
	0.5 ml	formaldehyde (37 %)
	ad 1 l H_2O	
Developing solution (LPS silver-staining):	2.5 % (w/v)	sodium carbonate
	27 µl	formaldehyde (37 %)
	ad 100 ml H_2O	
EDTA, 0.5 M, pH 8.0:	186.1 g	EDTA
	ad 800 ml H_2O	
	adjust to pH 8.0 (using NaOH)	
	ad 1 l H_2O	
Elution buffer/buffer B (affinity chromatography):	Lysis buffer (buffer A)	
	12 mM	maltose
Fixing solution (1 x, silver-staining proteins):	50 % (v/v)	methanol
	12 % (v/v)	acetic acid
	0.5 ml	formaldehyde (37 %)
	ad 1 l H_2O	
Fixing solution (2 x, LPS silver-staining):	50 % (v/v)	isopropanol
	14 % (v/v)	acetic acid
	ad 1 l H_2O	
In-line probing buffer (1 x):	50 mM	Tris-HCl (pH 8.3)
	20 mM	magnesium chloride
	100 mM	potassium chloride
Lysis buffer/buffer A (affinity chromatography):	20 mM	Tris-HCl (pH 8.0)
	150 mM	potassium chloride
	1 mM	magnesium chloride
	1 mM	DTT (add prior to use)

150

NaAc buffer (in-line probing):	0.25 M	sodium citrate

adjust to pH 5.0 (at 23 °C)

Na$_2$CO$_3$ buffer (in-line probing):	0.5 M	sodium carbonate
	10 mM	EDTA (pH 8.0)

adjust to pH 9.0 (using NaOH)

Na$_5$IO$_6$ (oxidizing solution, LPS silver-staining):	0.7 g	sodium periodate
	50 ml	Fixing solution (2 x)
	ad 100 ml H$_2$O	

Na$_2$S$_2$O$_3$ pretreatment solution (silver-staining):	0.2 mg/ml	sodium thiosulfate
		(x 5 H$_2$O)
	(1.6 ml stock dissolved in 200 ml H$_2$O)	

Na$_2$S$_2$O$_3$ stock solution (silver-staining proteins):	25 mg/ml	sodium thiosulfate
		(x 5 H$_2$O)

PAA gel electrophoresis solution for western blots:

PAA gel for separation gel (10 ml)	10 %	12 %	15 %
1 M Tris "lower" buffer (pH 8.8)	3.75 ml	3.75 ml	3.75 ml
40 % PAA solution (37.5:1 acrylamide/bisacrylamide)	2.5 ml	3 ml	3.75 ml
H$_2$O	3.75 ml	3.25 ml	2.5 ml
10 % (w/v) SDS	100 µl	100 µl	100 µl
10 % (w/v) APS	75 µl	75 µl	75 µl
TEMED	7.5 µl	7.5 µl	7.5 µl

PAA gel for stacking gel (10 ml)	4 %
1 M Tris "upper" buffer (pH 6.8)	1.25 ml
40 % PAA solution (37.5:1 acrylamide/bisacrylamide)	1 ml
H$_2$O	7.5 ml
10 % (w/v) SDS	100 µl
10 % (w/v) APS	150 µl
TEMED	15 µl

PAA gel electrophoresis solution for northern blots and sequencing gels:

PAA gel (500 ml, stock solution, 7 M urea)	6 %	10 %	15 %
40 % PAA solution (19:1 acrylamide/bisacrylamide)	75 ml	125 ml	187.5 ml
urea	210 g	210 g	210 g
10 x TBE buffer	50 ml	50 ml	50 ml

1 gel (70 ml):	70 ml	stock solution
	700 µl	10 % (w/v) APS
	70 µl	TEMED

PAA gel electrophoresis solution for gel-shift assays (native PAGE):

1 gel (70 ml, 6 % PAA):	10.5 ml	40 % PAA sol. (19:1)
	3.5 ml	10 x TBE
	56 ml	H_2O
	700 µl	10 % (w/v) APS
	70 µl	TEMED

PBS (10 x stock):	80 g	sodium chloride
	2 g	potassium chloride
	17.7 g	disodium hydrogen phosphate
	2.72 g	monopotassium phosphate
	ad 800 ml H_2O	
	adjust to pH 7.4	
	ad 1 l H_2O	

Proteinase K solution (LPS):	20 mg/ml	Proteinase K

RNA elution buffer (*in-vitro* transcription):	0.1 M	sodium acetate
	0.1 %	SDS
	10 mM	EDTA (pH 8.0)

RNA structure buffer (10 x):	100 mM	Tris-HCl (pH 7.0)
	1 M	potassium chloride
	100 mM	magnesium chloride

SB 5 x – Mg (toeprint):	50 mM	Tris-acetate (pH 7.6)
	500 mM	potassium acetate
	5 mM	DTT

SB 1 x Mg10 (toeprint):	10 mM	Tris-acetate (pH 7.6)
	100 mM	potassium acetate
	1 mM	DTT
	10 mM	magnesium acetate

SB 1 x Mg60 (toeprint):	10 mM	Tris-acetate (pH 7.6)
	100 mM	potassium acetate
	1 mM	DTT
	60 mM	magnesium acetate

SDS running buffer (10 x stock):	30.275 g	Tris base
	144 g	glycin
	10 g	SDS
	ad 1 l H_2O	

Silver-staining solution (protein):	2 g	silver nitrate ($AgNO_3$)
	750 µl	formaldehyde (37 %)
	ad 1 l H_2O	
Silver-staining solution (LPS):	100 ml	H_2O
	0.2 ml	10 N NaOH
	1.4 ml	ammonia (25 %)
	1.8 ml	silver nitrate solution (0.4 g $AgNO_3$/2 ml H_2O)
	keep the order of incredients/sol.	
SSC (saline-sodium citrate) buffer (20 x stock):	173.5 g	sodium chloride
	88.2 g	sodium citrate
	ad 800 ml H_2O	
	adjust to pH 7.0 (using HCl)	
	ad 1 l H_2O	
Stains-All (*in-vitro* transcription):	30 ml	Stains-All stock
	90 ml	formamide
	ad 200 ml H_2O	
Stains-All stock (*in-vitro* transcription):	0.03 g dissolved in 30 ml formamide	
Stop mix (RNA preparation):	95 % (v/v)	ethanol (absolute)
	5 % (v/v)	phenol
Stop solution (silber-staining proteins):	1 % (w/v)	glycine
Stop solution (LPS silver-staining):	50 mM	EDTA
TAE buffer (50 x stock):	242 g	Tris base
	51.7 ml	acetic acid
	100 ml	0.5 M EDTA (pH 8.0)
	ad 1 l H_2O	
TBE buffer (10 x stock):	108 g	Tris base
	55 g	boric acid
	40 ml	0.5 M EDTA (pH 8.0)
	ad 1 l H_2O	
Tbf I buffer (competent *E. coli* TOP 10):	1.47 g	potassium acetate
	4.975 g	manganese (II)-chloride
	3.73 g	potassium chloride
	ad 400 ml H_2O	

Tbf I buffer (competent *E. coli* TOP 10, continued):	adjust pH to 5.8 (using CH_3COOH) add 75 ml glycerin ad 500 ml H_2O	
Tbf II buffer (competent *E. coli* TOP 10):	2 ml	1 M MOPS
	150 ml	0.1 M calcium chloride
	8 ml	250 mM potassium chloride
	30 ml	glycerin
	ad 200 ml H_2O	
TBS buffer (10 x stock):	24.11 g	Tris base
	87.66 g	NaCl
	adjust to pH 7.4 (using HCl) ad 1 l H_2O	
TBS-T buffer (1 x):	100 ml	10 x stock solution
	1 ml	Tween[20]
	ad 1 l H_2O	
TE buffer (1 x):	100 mM	Tris-HCl (pH 8.0)
	10 mM	EDTA (pH 8.0)
Toeprint stop solution	50 mM	Tris-HCl (pH 7.5)
	0.1 % (w/v)	SDS
	10 mM	EDTA (pH 8.0)
Transfer buffer (10 x stock):	30 g	Tris base
	144 g	glycin
	ad 1 l H_2O	
Transfer buffer (1 x):	100 ml	10 x stock solution
	200 ml	methanol
	ad 1 l H_2O	
Tris "lower buffer" solution:	1.5 M	Tris-HCl (pH 8.8)
	0.4% (w/v)	SDS
Tris "upper buffer" solution:	0.5 M	Tris-HCl (pH 6.8)
	0.4% (w/v)	SDS
Western development solution:	2 mL	chemiluminescence solution A
	200 µl	chemiluminescence solution B
	6 µL	3% (v/v) H_2O_2

6.1.10.4. Sterilization

All media and solutions used in this study were sterilized prior to use by autoclaving at 120 °C and 1 bar atmospheric pressure for 20 min. Heat-sensitive solutions were sterile filtered. Glassware was sterilized by heating to 80 °C for a minimum of three hours.

6.2. Microbiological methods

6.2.1. Growth conditions and phenotypic characterization

6.2.1.1. *Helicobacter*

***Helicobacter* standard growth conditions**. Unless stated otherwise, *Helicobacter* strains were grown on GC-agar plates supplemented with 10% donor horse serum (DHS), 1 % vitamin mix, 10 µg/ml vancoymcin, 5 µg/ml trimethoprim and 1 µg/ml nystatin. For transformant selection and growth of mutant strains, 20 µg/ml kanamycin, 8 µg/ml chloramphenicol, 10 µg/ ml gentamicin or 10 µg/ml erythromycin were added. For liquid cultures, 15 or 50 ml Brain Heart Infusion (BHI) medium supplemented with 10% FBS and 10 µg/ml vancoymcin, 5 µg/ml trimethoprim and 1 µg/ml nystatin were inoculated with *Helicobacter* from plate to a final OD_{600nm} of 0.02 – 0.05 and grown under agitation at 140 rpm in 25 cm^3 or 75 cm^3 cell culture flasks. Bacteria were grown at 37 °C in a HERAcell 150i incubator in a microaerobic environment (10 % CO_2, 5 % O_2, and 85 % N_2). *H. pylori* changes from spiral to coccoid morphology during extended growth on plate or in liquid culture.

Induction of acidic stress. *H. pylori* strains were grown in nutrient-rich BHI medium (pH ~ 7.0) to exponential growth phase (OD_{600nm} of ~ 0.7 – 1.0) and split into three independent subcultures. One subculture was left untreated, whereas the media of the other two were adjusted to pH 5.0 or pH 8.5 with 37 % HCl or 1 M NaOH, respectively. Growth was carried out under microaerobic conditions at 37 °C and 140 rpm.

Growth in the presence/absence of iron and nickel. For growth under different ion (iron and nickel) availabilities, *H. pylori* was grown to exponential growth phase (OD_{600nm} of ~ 0.7 – 1.0) in BHI medium in the absence or presence of 20 µM nickel chloride ($NiCl_2$), 100 µM iron (II)-sulfate (Fe(II)SO_4) and 60 µM of the high-affinity iron chelator 2,2' dipyridyl (DPP). Growth was carried out in a microaerobic environment.

Growth under long-term oxidative stress conditions. *H. pylori* is a capnophilic aerobe as its growth is stimulated by atmospheric oxygen levels in the presence of 10 % carbon dioxide (CO_2) (Park & Lee, 2013, Bury-Mone *et al.*, 2006). In order to induce long-term

oxidative stress, *H. pylori* was grown in BHI medium at 37 °C and 140 rpm in a HERAcell 150i incubator that mimics atmospheric oxygen conditions (20 % O_2, 10 % CO_2).

Induction of short-term oxidative stress using hydrogen peroxide (H_2O_2). For transient oxidative stress, *H. pylori* was grown in BHI medium to exponential phase (OD_{600nm} of ~ 0.8) and split into three independent subcultures. One subculture was left untreated (control), while the other two were shifted to oxidative stress conditions by addition of 1 mM or 10 mM hydrogen peroxide (H_2O_2), respectively. Growth was performed under microaerobic conditions.

Growth of conditional RNase J-deficient mutant. *H. pylori* strain B128 wildtype (carrying pILL-2157) and $P_i\Delta N$-*rnj* (conditional Δrnj mutant; Δrnj containing pHP135, *i.e.* pILL2157 with N-terminal truncation of *rnj* gene under control of an IPTG-inducible promoter - P_i) were grown on GC-agar plates (supplemented as described above) containing 8 µg/ml chloramphenicol and 8 µg/ml chloramphenicol, 20 µg/ml kanamycin as well as 1 mM/ml IPTG, respectively. For pre-cultures, 15 ml BHI supplemented with 10 % FBS, 10 µg/ml vancomycin, 5 µg/ml trimethoprim, 1 µg/ml nystatin, 8 µg/ml chloramphenicol and 1 mM/ml IPTG were inoculated with *Helicobacter* from plate to a final $OD_{600\ nm}$ of 0.02 – 0.05 and grown over night in 25 cm^3 cell culture flasks. *H. pylori* strains were harvested by centrifugation at 6,500 x g for 5 min and washed two to three times with 1 x phosphate-buffered saline (PBS) to remove residual IPTG. Finally, cells were resuspendend in 1 ml PBS and used for inoculation of 100 ml BHI (supplemented as mentioned above) without IPTG at an $OD_{600\ nm}$ of 0.05 to 0.1. The 100 ml cultures of the wildtype and $P_i\Delta N$-*rnj* mutant were split into two subcultures each (2 x 50 ml), one was left untreated and to the other 1 mM/ml IPTG was added. Growth in absence (-IPTG) or presence of inducer (+IPTG) was used to control expression of ΔN-*rnj*. Cells were grown microaerobically to early exponential growth phase (OD_{600nm} of ~ 0.4) at 37 °C and 140 rpm. For further details (strains, plasmids), see also (Redko *et al.*, 2013).

Survival under acidic stress conditions. To determine the cellular viability of *H. pylori* strains exposed to low pH, bacteria were grown in BHI medium to early exponential growth phase (OD_{600nm} of ~ 0.4) and split into two independent subcultures. While one subculture was left untreated (pH ~ 7.0), the media of the other one was adjusted to pH 5.0 with 37 % HCl. Colony forming units (CFU) per ml prior to and at different time points after pH adjustment (*e.g.* one and 24 hrs) were determined by serial dilutions. The survival (%) of each tested *H. pylori* strain was assayed by comparison of CFU/ml obtained under pH 5 to those of the untreated culture (for each time point). Survival assays were done in biological duplicates.

Growth/survival under high salt conditions. *H. pylori* strains were grown in liquid media (BHI) to exponential growth phase (OD_{600nm} of ∼ 0.7 – 1.0). Cells were adjusted to an OD_{600nm} of 1.0 in 0.5 ml BHI and serial dilutions of the indicated strains were spotted on GC-agar plates containing 85 mM (no stress, normal salt content), 200 mM (mild stress) or 260 mM sodium chloride (harsh stress). Except for variations in the sodium chloride concentration, GC-agar plates were supplemented as described above (DHS, vitamin mix and antibiotics). Plates were incubated for 3 to 5 days at 37 °C under microaerobic conditions.

Motility assay. *H. pylori* strains were inoculated from the GC-agar plates into 10 ml Brucella Broth supplemented with 10% FBS, 1% vitamin mix, 10 µg/ml vancoymcin, 5 µg/ml trimethoprim and 1 µg/ml nystatin and grown microaerobically while shaking at 140 rpm and 37 ℃ to an OD_{600nm} of about 0.3 to 0.4. Cells were harvested by centrifugation at 6,500 x g for 5 min and resuspended at an OD_{600} of 0.1 in Brucella Broth (BB). For each strain, 1.0 µl of bacterial suspension was inoculated into motility soft-agar plates (BB + 0.4 % agar, supplemented with 5 % FBS) poured one day prior to the experiment. Plates were incubated for approximately 5 to 7 days microaerobically at 37 °C. Three measurements of each motility halo were made for each inoculation, which were averaged to give the mean swimming/swarming distance for each strain on a plate. All strains were inoculated together on three to four replicate plates and the mean swimming/swarming distance +/- standard error on these plates was used to assess the motility of each strain. For comparison between the wildtype and mutant strains, the median swimming/swarming diameter of the wildtype was used as reference and set to 100 %. At least three biological replicates were used to assay the motility behavior of different *H. pylori* strains/mutants.

Autoagglutination assay. Autoagglutination was determined as described previously for *C. jejuni* (Golden & Acheson, 2002). Briefly, *H. pylori* strains were grown to exponential phase in nutrient-rich BHI medium. Cells were adjusted to an OD_{600nm} of 1.0 in 1 x PBS (pH 7.4). Two milliliters were placed into at least two replicate tubes and incubated under microaerobic conditions for 24 hrs at 37 °C without shaking. After indicated time points (0, 2.5, 4.5, 7, 9 and 24 hrs) the OD_{600nm} of the top 100 µl was measured. Measurements were normalized to the optical density of each strain at the time point zero, *i.e.* the OD_{600nm} at 0 min was set to 100 % and the percentage of OD_{600nm} remaining at indicated time points was plotted. Autoagglutination assays were performed at least two times for each strain.

Resistance/sensitivity to antibiotics. To assay the sensitivity of *H. pylori* to rifampicin and polymyxin B, disk diffusion assays were performed. In these assays, *H. pylori* strains grown over night on fresh GC-agar plates (exponential growth phase) were resuspended in 1 ml BHI. Cells were adjusted to OD_{600nm} of 0.1 and 100 µl of *H. pylori* suspensions were spread on GC-agar plates by glass beads. Disks impregnated with 10 µg/ml rifampicin or 300 units

polymyxin B were placed on top. Plates were incubated for 3 days under microaerobic conditions at 37 °C. Three measurements of each inhibition zone were made, which were averaged to give the diameter of the inhibition zone of each strain on a plate. For comparison between wild-type and mutant stains of a given *H. pylori* strain background, the inhibition zone diameter of the wildtype was used as reference and set to 100 %.

For polymyxin B sensitivity testing of *H. pylori* 26695 *tlpB* mRNA leader variants (ΔG, 6G to 16G), the inhibition zone diameter of the strain carrying the *tlpB* 6G leader, which showed the highest levels of *tlpB*-HP0102 expression and thus, smooth LPS production, was used as reference and set to 100 %. To assess whether or to which extent *repG* deletion affects polymyxin B sensitivity of the *tlpB* mRNA leader variants, the diameter of inhibition zone of Δ*repG* was compared to the respective wild-type background of each G-stretch mutant, *e.g.* Δ6G vs. 6G. Antibiotic sensitivity assays were performed in biological duplicates and/or triplicates.

6.2.1.2. *E. coli*

***E. coli* growth conditions.** Bacteria were grown on LB-agar plates and in LB at 37 °C, 220 rpm with normal aeration. Cultures were inoculated from a single colony of strains grown over night on plates at 37 °C, *e.g.* in 5 ml LB medium for plasmid preparation. When appropriate, media were supplement with antibiotics (Table 6.8).

6.2.2. Genetic manipulation

6.2.2.1. *Helicobacter*

Construction of *Helicobacter* mutant strains. Mutants were cloned by double-crossover homologous recombination and natural transformation of PCR-amplified constructs carrying either the *aphA*-3 kanamycin (Skouloubris *et al.*, 1998), *catGC* chloramphenicol (Boneca *et al.*, 2008), *rpsL-erm* erythromycin (Dailidiene *et al.*, 2006) or *aac(3)-IV* gentamicin resistance cassette (Bury-Mone *et al.*, 2003) flanked by about 500 bp homology regions up- and downstream of the respective genomic locus as described previously (Bury-Mone *et al.*, 2001). Briefly, *H. pylori* was grown from frozen stocks until passage two, then streaked in small circles on a fresh GC-agar plate and grown for 6-8 hrs at 37 °C under microaerobic conditions. For transformation, 500 ng - 1 µg purified PCR product was added to the cells. After incubation for 14-16 hrs at 37 °C, cells were passed onto selective plates with indicated antibiotics. Genomic DNA (gDNA) of transformants was isolated using

NucleoSpin Plasmid kit and mutants were validated by colony-PCR on gDNA and/or sequencing.

In general, *H. pylori* mutants, *i.e.* deletion and appropriate complementation, overexpression, GFP-reporter fusion and chromosomally 3xFLAG-tagging of a given gene, were cloned by transformation of constructs amplified from plasmids (details see Appendix, Table 13.3) or generated by overlap-extension-PCR (overlap-PCR). To elucidate these two strategies, chosen examples of diverse *H. pylori* mutants are described below; the underlying procedures were similar for all mutants (also in different strain backgrounds). Plasmids, gDNA and oligonucleotides used for the construction of all *H. pylori* mutants are described in detail in Appendix, Table 13.4. Sequences of *gfpmut3* reporter fusions as well as of diverse *repG* and *tlpB* mutants (G-stretch variants, compensatory base-pair exchanges) are listed in Tables 6.9, 6.10 and 6.12.

Complementation of the Δ*repG* mutant with different sRNA variants. To construct RepG complementation strains, the intergenic region of HP1043 and HP1044 including *repG* under control of its own promoter together with the *catGC* resistance cassette (Boneca *et al.*, 2008) was inserted into the *rdxA* gene locus, which is frequently used for complementation in *H. pylori* (Goodwin *et al.*, 1998). An *rdxA*(500 nt up)-*catGC-repG-rdxA*(500 nt down) complementation construct was amplified from gDNA of a RepG complementation strain (kindly provided by F. Darfeuille and J. Reignier, University of Bordeaux, France; for sequence details see Figure 6.1) using oligos CSO-0017 and CSO-0018. The amplified about 2.3 kbp PCR product was transformed into the Δ*repG* deletion mutant (JVS-7014, Sharma *et al.*, 2010). The obtained strain CSS-0046 (C$_{RepG}$) was verified by PCR using oligos CSO-0205 and CSO-0207. Furthermore, the *XhoI/XbaI*-digested about 2.3 kbp PCR fragment was also cloned into plasmid pJV752.1 (Sharma *et al.*, 2007), resulting in plasmid pSP39-3, which was used for further mutant generation of RepG.

The second stem loop mutant of RepG (SL 2, Figure 2.3 A) was constructed by PCR amplification of pSP39-3 using oligos CSO-0080 and CSO-0081. Upon *DpnI*-digestion, the PCR product was gel-purified, self-ligated and transformed into *E. coli* TOP 10 cells. Positive clones were selected on plates with 100 µg/ml ampicillin and confirmed by colony PCR using oligos pZE-A and CSO-0205. The resulting plasmid pSP42-1 was validated by sequencing with CSO-0206. Afterwards, a PCR product amplified from pSP42-1 using oligos CSO-0017 and CSO-0018 was used for complementation of the *H. pylori* Δ*repG* mutant, resulting in strain CSS-0747 (SL 2). To construct the other RepG variants, ΔCU, 3xG and 1xG*, overlap-extension-PCR was performed as previously described (Stingl *et al.*, 2007). First, PCR fragments were amplified from pSP39-3 using oligos CSO-0017 x CSO-0139 and CSO-0018 x CSO-0138 for RepG ΔCU, CSO-0017 x CSO-0143 and CSO-0018 x CSO-0142 for

RepG complementation construct: *rdxA*(500 nt up)–*catGC–repG–rdxA*(500 nt down)

```
        CSO-0017
tctagaGATCAGCCTGCCTTTAGGGTATGTTTTTGGGAGGATTGGGAATTTTTTAAACCAGGAGCTTGTGGG
AAGAATTGTCCCCAAAGACAGCCATTTAGGGCAAATCATAGGCATTATGGTGGATAATGAGTTGCGTTATCC
CAGCCAATTGATTGAAGCGTTTTTAGAGGGGGTTATCGTGTTTTTAATGGTAATGTGGGCTAAAAAACACAC
CAAAACGCATGGGTTGCTGATTGTGGTTTATGGTTTGGGGTATTCCTTGATGCGCTTTATTGCGGAATTTTA
CAGAGAGCCGGACAGCCAAATGGGGGTTTATTTTTTAAATTTGAGCATGGGGCAGATTTTAAGCTTATTTAT
GGTAATTGTTTCGTTAGGGATTTTATTGTATGCTACAAAAAATTCTAAAAAATAAAGGAAAATCAATGAAA
TTTTTGGATCAAGAAAAAGAAGACAATTATTAAACGAGCGCCATTCTTGCAAGATGTTTGATAGCCATTAT
GAGTTTTCTAGCACAGAATTAGAAGAAATCGCTGctagagatccgccatattgtgttgaaacaccgcccgga
acccgaTATAATccgcccttcaacagatccgagattttcaggagctaAGGAAGctaaaATGGAGAAAAAAT
CACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCAGTCAGTTGC
TCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCA
CAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATTCCGTATGGCAAT
GAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCCTTGTTACACCGTTTTCCATGAGCAAACTGAAAC
GTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTGGC
GTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCC
CTGGGTGAGTTTCACCAGTTTTGATTTAAACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACCAT
GGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTCAGGTTCATCATGCCGTTTGTGA
TGGCTTCCATGTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAtt
tttttaaggcagttattggtgcccttaaacgcctggttgctacgcctgaataagtgataataagcggatgaa
tggcagaaattcggatcttccatacctaccagttctgcgcctgcaggtcgataaaccgatacaattaaaggc
tcctttttggagcctttttttttggagattttcaacgtggatctgaattcgagatgcatcaagcccttgatta
ttggttgtaaaaaatgcctttgagcgtttttatggataatttttaaaatcatttgctaaaaatcaccatttt
attgTATAATtacaaATCCAACCATTCCTTATGGTTTGGTTGGCACCGCTAAGATTGAAGGGTCACCTCCCC
CTCCTTTCCCTTTGTCTTGGCGGTTGTTTTttaatccttgtttagttcttatttttaaccgcttgggtaccG
TATGCTCTTTAAGACCCAGCGAGTTGTTACCACACGGCCACTACATGCAAAATCTCTATCCGGAGTCTTATA
AAGTTAGAGTGATCCCCTCTTTTGCTCAAATGCTTGGCGTGAGATTCAACCACAGCATGCAAAGATTAGAAA
GCTATATTTTAGAGCAATGCTATATCGCTGTGGGGCAAATTTGCATGGGCGTGAGCTTAATGGGATTGGATA
GTTGCATTATTGGAGGCTTTGATCCTTTAAAGGTGGGCGAAGTTTTAGAAGAGCGTATCAATAAGCCTAAAA
TCGCATGCTTGATCGCTTTGGGCAAGAGGGTGGCAGAAGCGAGTCAAAAATCAAGAAAATCAAAAGTTGATG
CGATTACTTGGTTGTGATTAAACAAAATCAAAAACTTTTTAACTATAATCAAACCTAAATTAAAGTTCAAGG
AGTGGCATTTTGTTTAAAAGAATGGTTTTAATCGCTCTTTTAGGGGTGTTTTCAAGCGTTTCATTAAGCGCT
AAGctcgag
CSO-0018
```

rdxA 500 nt upstream: 1014605 – 1014074	
catGC cassette	
Intergenic region of HP1043 and HP1044 with ***repG*** under control of its native **promoter**	
rdxA 500 nt down: 1013930 – 1013423	

Figure 6.1: Sequence of the RepG (*H. pylori* strain 26695) complementation construct in the *rdxA* locus. The intergenic region of HP1043 and HP1044 (highlighted in light red) with the *repG* gene (red letters) was introduced into pSP39-3, which harbors 500 nt up- (light green) and downstream (dark green) of the *rdxA* locus as well as the *catGC* resistance cassette (blue). The -10 promoter regions and/or the RBS of *repG* as well as the *catGC* resistance cassette are highlighted in bold. The oligos CSO-0017 and CSO-0018, which were used for amplification of the *rdxA*(500 nt up)-*catGC-repG-rdxA*(500 nt down) complementation construct are underlined.

RepG 3xG as well as CSO-0017 x CSO-0141 and CSO-0018 x CSO-0140 for RepG 1xG*. Next, the corresponding PCR fragments were mixed in an equimolar ratio and used as templates for overlap-extension-PCR reactions using CSO-0017 and CSO-0018. Afterwards, gel-purified PCR reactions were transformed into the *H. pylori* Δ*repG* mutant (JVS-7014). All *H. pylori* complementation mutants were verified by PCR using oligos JVO-5069 x JVO-5257 and CSO-0207 x CSO-205 on gDNA of the respective mutants as template. All clones were checked by sequencing with CSO-0206.

Construction of Δ*tlpB* and *tlpB*::3xFLAG strains. The *tlpB* gene (HP0103) was deleted from strain CSS-0004 (*H. pylori* 26695 wildtype) by insertion of the *rpsL-erm* cassette, which confers dominant streptomycin susceptibility and erythromycin resistance (Dailidiene *et al.*, 2006). A construct containing the *rpsL-erm* cassette flanked by 500 nt up- and downstream of the *tlpB* open reading frame was generated by overlap-PCR. PCR products corresponding to 500 nt upstream of *tlpB* (CSO-0040 x CSO-0037 on gDNA of CSS-0004), 500 nt downstream of the *tlpB* stop codon (CSO-0038 x CSO-0039 on gDNA of CSS-0004), and the *rpsL-erm* cassette (CSO-0035 x CSO-0036 on gDNA of *H. pylori* 26695 carrying chromosomal *rpsL-erm* cassette; kindly provided by D. Berg, University of California, San Diego, La Jolla, CA/Washington University Medical School, St. Louis, MO; Dailidiene *et al.*, 2006) were mixed in an equimolar ratio and used as templates for overlap-PCR with CSO-0040 x CSO-0039. Gel purified PCR product was used for natural transformation in *H. pylori* 26695. Positive erythromycin-resistant clones were checked by PCR on gDNA using CSO-0051 x CSONIH-0033, resulting in strain CSS-0163 (Δ*tlpB*). For the construction of a double deletion mutant Δ*tlpB*/Δ*repG*, the Δ*repG* deletion construct (*aphA-3* flanked by 500 nt up- and downstream of *repG*) was amplified by PCR using JVO-5070 x JVO-5072 on gDNA of JVS-7014 (Sharma *et al.*, 2010). The purified PCR product was transformed into strain CSS-0163 (Δ*tlpB*). The double deletion strain CSS-0164 (Δ*tlpB*/Δ*repG*) was verified by PCR using JVO-5069 x JVO-5257.

To construct a *tlpB*::3xFLAG-tagged strain (CSS-0190), a plasmid (pSP57-4) containing the 3xFLAG and the *rpsL-erm* cassette flanked by 500 nt up- and downstream of the *tlpB* stop codon was cloned into *E. coli.* First, 500 nt up- and downstream of the *tlpB* stop codon were amplified from gDNA of strain CSS-0004 using CSO-0208 x CSO-0211. The resulting PCR product was *XbaI/XhoI* digested and introduced into likewise digested pJV752-1, resulting in pSP55-4. Next, pSP55-4 was used as template for a PCR with CSO-0245 and CSO-0210 to fuse the 3xFLAG-tag to *tlpB* and introduce the *rpsL-erm* cassette using an *EcoRI* restriction site. In parallel, an overlap-PCR with PCR fragments of the 3xFLAG-tag (CSO-0065 x CSO-0046 on gDNA of JVS-7033, Rieder *et al.*, 2012) and the *rpsL-erm* cassette (CSO-0045 x CSO-0209 on gDNA of an *H. pylori* 26695 strain carrying the *rpsL-erm* cassette) was performed and the resulting 3xFLAG:*rpsL-erm* construct was digested

with *EcoRI*. Both *EcoRI*-digested PCR products (plasmid with *tlpB* up- and downstream region as well as 3xFLAG:*rpsL-erm*) were ligated and transformed into *E. coli*, resulting in pSP57-4. Insertion of the *rpsL-erm* cassette was verified by colony-PCR (pZE-A x CSONIH-0033) and in-frame fusion of *tlpB*::3xFLAG by sequencing with CSO-0208, respectively. A PCR product amplified from pSP57-4 with CSO-0208 x CSO-0211 was transformed into *H. pylori* 26695 (CSS-0004) and G27 (CSS-0010)*.* Positive erythromycin-resistant mutants were confirmed by PCR on gDNA (CSO-0050 x CSO-0046) and sequencing with CSO-0208, resulting in CSS-0190 (26695 *tlpB*::3xFLAG) and CSS-0196 (G27 *tlpB*::3xFLAG). Deletions of *repG* in CSS-0190 and CSS-0196 were constructed as described above, resulting in strains CSS-0215 (26695 *tlpB*::3xFLAG/Δ*repG*) and CSS-0197 (G27 *tlpB*::3xFLAG/Δ*repG*). Strains CSS-0215 and CSS-0197 were complemented with RepG from *H. pylori* 26695 in the *rdxA* locus by transformation of a PCR product amplified from gDNA of CSS-0046 (C$_{RepG}$) using CSO-0017 x CSO-0018. The resulting strains are CSS-0285 (26695 *tlpB*::3xFLAG/Δ*repG*/C$_{RepG}$) and CSS-0283 (G27 *tlpB*::3xFLAG/Δ*repG*/C$_{RepG}$).

Deletion of *repG* in diverse *Helicobacter* strains. To delete *repG* in diverse *H. pylori* strains, the sRNA deletion construct was amplified from JVS-7014 (Sharma *et al.*, 2010) using oligos JVO-5070 x JVO-5072 and transformed into *H. pylori* strains J99 (CSS-0001), B8 (CSS-0213), India7 (CSS-0099), Shi470 (CSS-0173), Lithuania75 (CSS-0101), Cuz20 (CSS-0097) and X47-2AL (CSS-0996). Deletion of *repG* was verified by PCR using JVO-5069 x JVO-5257 on gDNA, resulting in strains CSS-0732 (J99 Δ*repG*), CSS-0733 (B8 Δ*repG*), CSS-0734 (India7 Δ*repG*), CSS-0735 (Shi470 Δ*repG*), CSS-0736 (Lithuania75 Δ*repG*), CSS-0737 (Cuz20 Δ*repG*), and CSS-0997 (X47-2AL Δ*repG*), respectively.

Generation of compensatory base-pair exchanges in the *tlpB* 5' UTR. To either delete the G-repeat (ΔG) or introduce compensatory base-pair exchanges (3xC and 1xC*) in the *tlpB* 5' UTR, a backbone plasmid (pSP60-2) containing 500 nt up- and downstream of the *tlpB* transcriptional start site (TSS) as well as the *rpsL-erm* resistance cassette was constructed. First, 500 nt up- and downstream of the TSS of *tlpB* were amplified with CSO-0291 x CSO-0040 from gDNA of *H. pylori* 26695 (CSS-0004). The resulting PCR product was inserted into pJV752-1 using *XhoI*/*XbaI* restriction sites, which resulted in pSP58-5. Next, the *rpsL-erm* cassette was inserted upstream of the *tlpB* promoter (P$_{tlpB}$) by ligation of *EcoRI*/*BamHI*-digested PCR product based on pSP58-5 (CSO-0294 x CSO-0295) and a *rpsL-erm* cassette PCR product (CSO-0308 x CSO-0309 on gDNA of CSS-0163), which resulted in pSP60-2. Afterwards, site-directed mutagenesis by Quick-change-PCR was performed with pSP60-2 as template to delete the G-repeat (CSO-0318 x CSO-0319) or introduce triple (CSO-0316 x CSO-0317) or single (CSO-0314 x CSO-0315) G to C nucleotide exchanges, which finally resulted in pSP64-1 (ΔG), pSP66-4 (3xC) and pSP65-4 (1xC*), respectively. All

plasmids were checked by sequencing using CSO-0291. PCR fragments based on pSP60-2, pSP64-1, pSP65-4 and pSP66-4 were amplified with CSO-0291 x CSO-0040 and used for direct transformation in *H. pylori* 26695 (CSO-0004). Positive *H. pylori* clones were selected on erythromycin plates and confirmed by PCR using CSO-0051 x CSO-0308 and sequencing of the corresponding gDNA with CSO-0291 or CSONIH-0033, which led to CSS-0384 (P*tlpB*), CSS-0385 (*tlpB* ΔG), CSS-0386 (*tlpB* 3xC) and CSS-0387 (*tlpB* 1xC*). The *repG* gene was deleted from CSS-0384, CSS-0386 and CSS-0387 as described previously, resulting in CSS-0388 (P*tlpB*/Δ*repG*), CSS-0389 (*tlpB* 3xC/Δ*repG*) and CSS-0397 (*tlpB* 1xC*/Δ*repG*). These mutants were complemented with either wild-type RepG or RepG mutant sRNAs (3xG, 1xG*); for details see Appendix, Table 13.4.

Cloning of the *tlpB* promoter exchange. To exchange the *tlpB* promoter with the *cagA* promoter, the plasmid pSP60-2 was used. The *cagA* promoter region was amplified with CSO-0306 x CSO-0431 from gDNA of *H. pylori* 26695 (CSS-0004) and a PCR on pSP60-2 was performed with CSO-0430 x CSO-308. Both PCR products were *BamHI* digested and ligated, resulting in pSP91-3. The plasmid was verified by colony PCR with CSO-0306 x pZE-A and checked by sequencing with CSONIH-0033. PCR fragments were amplified with CSO-0291 x CSO-0040 from pSP91-3 and transformed into *H. pylori* 26695. Transformants were selected on erythromycin plates and verified by PCR with CSO-0051 x CSO-0308 on gDNA and sequencing with CSONIH-0033, resulting in strain CSS-0657 (P*cagA*). The *repG* gene was deleted in strain CSS-0657 as described previously. The obtained strain is CSS-0658 (P*cagA*/Δ*repG*).

Cloning of translational reporter fusions to *gfpmut3*. For the generation of a translational reporter fusion, we fused the regions corresponding to the promoters, 5' UTRs and a fraction of the N-terminal coding region of *tlpB* or *cagA* to *gfpmut3* (Carpenter *et al.*, 2007) and introduced it together with the *catGC* resistance cassette (Boneca *et al.*, 2008) into the *rdxA* locus of *H. pylori* G27. First, a transcriptional *ureA*::*gfpmut3* fusion was inserted into pSP39-3 using CSO-0440 x CSO-0441 on p463 (*ureA*::*gfpmut3* based on pTM117; kindly provided by D. S. Merrell, USU, Bethesda, MD, USA) and CSO-0442 x CSO-0443 on pSP39-3, respectively. The PCR products were digested with *SalI*/*NotI* and ligated, resulting in pPT3-1 (*ureA*::*gfpmut3*). The plasmid pPT3-1 served as backbone for the generation of the translational fusions of *tlpB* and *cagA* to *gfpmut3*. The *tlpB* promoter region and its 5' UTR including the first five amino acids of the *tlpB* coding region (regarding the annotated ATG) were amplified from gDNA of *H. pylori* 26695 (CSS-0004) using oligos CSO-0581 and CSO-0126. Similarly, the *cagA* promoter, its 5' UTR and the first 28 amino acids of the *cagA* coding region were amplified with CSO-0284 and CSO-0590 from gDNA of *H. pylori* G27 (CSS-0010). The purified PCR products were digested with *ClaI/NheI* and

ligated with a likewise digested PCR product, which was amplified from pPT3-1 using CSO-0146 and CSO-0683, resulting in pSP109-6 (*tlpB* 5th::*gfpmut3*) and pMA5-2 (*cagA* 28th::*gfpmut3*). These plasmids were checked by colony PCR using CSO-581 x pZE-*XbaI* for pSP109-6 and CSO-0590 x pZE-*XbaI* for pMA5-2 and in-frame fusions to *gfpmut3* were validated by sequencing with CSO-0206 and/or JVO-0155. PCR products amplified with CSO-0017 x CSO-0018 from pPT3-1, pMA5-2 and pSP109-6 were transformed into *H. pylori* G27 wildtype (CSS-0010) and/or Δ*repG* (CSS-0169) strains. Positive transformants were checked by PCR with CSO-0205 and CSO-0207 and the in-frame fusion of *tlpB* or *cagA* to *gfpmut3* was verified by sequencing with CSO-0206 and/or JVO-0155. The corresponding *H. pylori* G27 strains are CSS-0748 (*tlpB* 5th::*gfpmut3*), CSS-0751 (*tlpB* 5th::*gfpmut3*/Δ*repG*), CSS-0804 (*cagA* 28th::*gfpmut3*), CSS-0805 (*cagA* 28th::*gfpmut3*/Δ*repG*).

Detailed sequences for all constructed *gfpmut3* reporter fusions (*tlpB*, HP0102, *cagA*) are shown in Table 6.9.

Construction of markerless *tlpB*::3xFLAG* strains. Markerless *tlpB*::3xFLAG* strains were constructed as described in (Dailidiene *et al.*, 2006) by using a contra-selectable streptomycin susceptibility determinant. For this purpose, a streptomycin resistant *H. pylori* 26695 strain (26695 StrR, CSS-0024) was generated by introduction of two point mutations in the *rpsL* gene (K43R and K88R). A PCR product amplified from genomic DNA of *H. pylori* 26695 with JVO-5702 x JVO-5703 was used for mutagenesis. Transformants were selected on plates containing 10 μg/ml streptomycin and positive clones were checked by sequencing of gDNA using JVO-5704.

For the construction of a markerless *tlpB*::3xFLAG* strain, a PCR product amplified from pSP57-4 with CSO-0208 x CSO-0211 was transformed into *H. pylori* 26695 StrR (CSS-0024) and selected on erythromycin plates. The 3xFLAG-tagging of *tlpB* was confirmed as described above. Furthermore, pSP70-1 was constructed; a plasmid, which contains 500 nt up- and downstream of the *tlpB* stop codon including a 3xFLAG tag. To construct this plasmid, cycle-PCR with CSO-0428 x CSO-0429 on pSP57-4 was performed and the resulting PCR product was digested with *DpnI* and transformed into *E. coli*. Loss of the *rpsL-erm* resistance cassette was verified by PCR using pZE-A x CSONIH-0033 and in-frame fusion of *tlpB::*3xFLAG was checked by sequencing with CSO-0208. A PCR product amplified from pSP70-1 with CSO-0208 x CSO-0211 was used for the transformation and removal of the *rpsL-erm* resistance cassette from CSS-0461 (26695 StrR *tlpB*::3xFLAG). Positive *H. pylori* mutants were selected on plates containing 10 μg/ml streptomycin, and the removal of the *rpsL-erm* resistance cassette was checked by plating of streptomycin resistant mutants on plates containing 10 μg/ml erythromycin. The markerless *tlpB*::3xFLAG* tagged 26695

Table 6.9: RNA sequences of *H. pylori* *gfpmut3* reporter fusions. The homopolymeric G-repeat in the 5' UTR of *tlpB* mRNA and G^HP0102 in the *tlpB*-HP0102 IGR are underlined. The RBSs as well as start codons (AUG) and ORFs are shown in light green and gray letters, respectively. Mutations within the G^HP0102 are shown in red. *gfpmut3*-fusion constructs are shown as RNA as some of them were used in *in-vitro* translation assays. The *NheI* restriction site is marked in yellow.

Name	Sequence 5' → 3'	Strain number
tlpB 5th::*gfpmut3* mRNA	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUCUUCAGCUAGCAGU-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-0748/ CSS-0751
cagA 28th::*gfpmut3* mRNA	GUUCAAAGACAUGAAUUGAUUACUCAAGUGUGUAGCAGUUUUUAGCAGUCUUUGUAUACCAACAAGAUACCGAUAGGUAUGAAACUAGGUAUAGUAAGGAGAAACAAUGACUAACGAAACCAUUAACCAACAACCACAAACUGAAGCGGCUUUUUAACCCGCAGCAAUUUAUCAAUAAUCUUCAAGUGGCUGCUAGCAGU-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-0804/ CSS-0805
tlpB-HP0102 10th::*gfpmut3* mRNA	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUG-TLPB CODING REGION-UAAUCAGGGGGAGUUUAUUAAAAAAGGGUUGGAUUGUUAAAAGUUUCUGUGAUCACGGCGUGUGCUAGCAGU-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-2104/ CSS-2107
*tlpB*mini-HP0102 10th::*gfpmut3* mRNA	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCUAAUCAGGGGGAGUUUAUUAAAAAAGGGUUGGAUUGUUAAAAGUUUCUGUGAUCACGGCGUGUGCUAGCAGUAAAG-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-2116/ CSS-2119
*PtlpB*HP0102 10th::*gfpmut3* mRNA	UCAGGGGGAGUUUAUUAAAAAAGGGUUGGAUUGUUAAAAGUUUCUGUGAUCACGGCGUGUGCUAGCAGUAAAG-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-2138/ CSS-2141
tlpB-HP0102 10th::*gfpmut3* ΔG^HP0102 mRNA	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUG-TLPB CODING REGION-UAAUCAGUUUAUUAAAAAAGGGUUGGAUUGUUAAAAGUUUCUGUGAUCACGGCGUGUGCUAGCAGU-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-2150/ CSS-2153
tlpB-HP0102 10th::*gfpmut3* ATTTA^HP0102 mRNA	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUG-TLPB CODING REGION-UAAUCAAUUUAAGUUUAUUAAAAAAGGGUUGGAUUGUUAAAAGUUUCUGUGAUCACGGCGUGUGCUAGCAGU-AAAGGA-GFPMUT3-UAAAUGUCCAGACCUGCAGUUAACUCCCGCGGCCGCAAAAAAC	CSS-2156/ CSS-2159

strain (CSS-0464, 26695 Str^R *tlpB*::3xFLAG*) was checked by PCR on gDNA with CSO-0050 x CSO-0046 or CSO-0050 x CSONIH-0033 and sequencing with CSO-0208. RepG was deleted from CSS-0464 as described before, resulting in strain CSS-0467.

Variation of the G-repeat length in the *tlpB* 5' UTR of *H. pylori* 26695. *H. pylori* 26695 strains with varying G-repeat length of 6-16 guanines were generated by cycle-PCR on pSP64-1 with CSO-0318 and CSO-0448 to CSO-0457 (see Tables 6.7 and 6.10). The obtained PCR products were *DpnI*-digested, self-ligated and transformed into *E. coli*. The resulting plasmids pSP73-1 to pSP82-1 (Appendix, Table 13.3) were sequenced with CSONIH-0033. Afterwards, these plasmids and pSP60-2 (represents the 26695 WT 5' UTR of *tlpB* with a 12 nt-long G-repeat) as well as pSP64-1 (ΔG) were used as templates for PCR with CSO-0040

x CSO-0291. The PCR products with different G-repeat lengths were transformed into CSS-0464 (wildtype background) and CSO-0467 (Δ*repG*). Positive clones were checked by colony PCR on gDNA with CSO-0051 x CSO-0308 and sequencing with CSONIH-0033 x CSO-0277, resulting in strains CSS-0471 to CSS-0493 (Appendix, Table 13.4), respectively.

Table 6.10: Sequences of *tlpB* mRNA leader mutants of *H. pylori* strain 26695. The length-variable, homopolymeric G-repeat in the 5' UTR of the *tlpB* mRNA is shown in red and the RBS as well as start codon (ATG) are marked in light green. The gDNAs of the *tlpB* mRNA leader mutants were used for the generation of DNA templates in the T7 *in-vitro* transcription assays.

Name	Sequence 5' → 3'	Strain number
tlpB WT	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTT ATAGGAGATAAATG	CSS-0470/ CSS-0482
tlpB ΔG	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATAGGAGATAAA TG	CSS-0471/ CSS-0483
tlpB 6G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATAGGA GATAAATG	CSS-0472/ CSS-0484
tlpB 7G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATAGG AGATAAATG	CSS-0473/ CSS-0485
tlpB 8G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATAG GAGATAAATG	CSS-0474/ CSS-0486
tlpB 9G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATA GGAGATAAATG	CSS-0475/ CSS-0487
tlpB 10G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTAT AGGAGATAAATG	CSS-0476/ CSS-0488
tlpB 11G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTA TAGGAGATAAATG	CSS-0477/ CSS-0489
tlpB 13G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATT TATAGGAGATAAATG	CSS-0478/ CSS-0490
tlpB 14G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGAT TTATAGGAGATAAATG	CSS-0479/ CSS-0491
tlpB 15G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGA TTTATAGGAGATAAATG	CSS-0480/ CSS-0492
tlpB 16G	TGTTTGTTCTTTTGTTTCGTTTTCAAACAACCGGGTTTTAATTTTGTTTTGTGCCACTCATTTT TCGGGGGGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATG ATTTATAGGAGATAAATG	CSS-0481/ CSS-0493

6.2.2.2. *E. coli*

Transformation of chemically competent *E. coli* (plasmid construction). About 30 μl of chemically competent *E. coli* TOP 10 cells (self-made) were mixed with 5 μl of a ligation reaction, 10 – 50 ng of plasmid DNA and/or *DpnI*-digested PCR product amplified from plasmid by cycle-PCR. The mixture was pre-incubated for 45 min on ice. Subsequently cells were heat-shocked for 90 s at 42 °C and chilled for 5 min on ice. Then, 120 μl of SOC medium was added and cells were recovered for 60 min at 37 °C, 220 rpm. Recovered cells were plated on LB-agar supplemented with the appropriate selection antibiotics.

Chemically competent *E. coli* cells using magnesium chloride ($MgCl_2$). About 50 ml Superbroth medium supplemented with 10 mM $MgCl_2$ were inoculated with 350 μl of an *E. coli* TOP 10 over-night culture (5 ml LB) and grown until early exponential growth phase ($OD_{600\,nm}$ of 0.3 – 0.4). Bacteria were filled into pre-cooled tubes and centrifuged for 4 min at 1,100 x g and 4 °C. After resuspension in 15 ml cooled Tbf I buffer and incubation for 20 min on ice, bacteria were centrifuged for 8 min at 785 x g and 4 °C. Finally, bacteria were resuspended on ice in 900 μl TBf II buffer and 60 μl aliquots were snap-frozen in liquid nitrogen. Competent *E. coli* TOP 10 aliquots were stored at -80 °C until use.

6.2.3. Animal infection assays with mouse-adapted *H. pylori* strain X47-2AL

Mouse infection studies with *H. pylori* strain X47-2AL and NMRI Swiss mice were performed by the lab of PhD Hilde de Reuse, Institut Pasteur, Paris, France as published previously in Bury-Mone *et al.*, 2008. The mouse adapted *H. pylori* strain X47-2AL was originally isolated from domestic cats (Handt *et al.*, 1995), and has been successfully used in mouse infection studies to identify virulence factors (Bury-Mone *et al.*, 2008). Moreover its genome sequence was recently published (Veyrier *et al.*, 2013). For *in-vivo* infection studies, *H. pylori* X47-2AL wildtype and mutant strains were grown on blood-agar plates, harvested after 24 hrs and resuspended in nutrient-rich peptone broth. The bacteria were adjusted to 10^8 cells/ml. Aliquots of 100 μl bacterial suspension per strain/mutant were administered orogastrically to seven female NMRI Swiss mice (4 weeks old) as well as plated in parallel (input). In each experiment, five mice were inoculated with peptone broth as a negative control. Mice were sacrificed 4 weeks after inoculation. The colony forming units (CFU) per gram of stomach weight were calculated by serial dilutions and plating assays on blood-agar plates supplemented with the usual antibiotic-fungicide mixture in the presence of bacitracin (200 μg/ml) plus nalidixic acid (10 μg/ml). Cultures on plates were grown without additional antibiotics for the mice infected with the parental strains or on plates supplemented with the respective antibiotics for mutant strains.

6.3. Basic molecular biological methods (DNA techniques)

6.3.1. Determination of the concentration of nucleic acids

The concentrations of all nucleic acid solutions (DNA and RNA, undiluted and 1:10 dilution) were determined using NanoDrop2000.

6.3.2. Preparation of plasmid DNA (*E. coli*) and genomic DNA (*H. pylori*)

Plasmid DNA from *E. coli* and genomic DNA from *H. pylori* were extracted from bacterial cells using the NucleoSpin Plasmid kit according to the manufacturer's protocol.

6.3.3. Polymerase chain reaction

DNA fragments of interest were amplified *in vitro* by PCR using Phusion or *Taq* DNA polymerase and the synthetic DNA oligonucleotides listed in Table 6.7.

	Phusion DNA polymerase			*Taq* DNA polymerase		
Reagents for 50 µl reaction						
DNA template (plasmid, gDNA)	~ 50 – 100 ng			~ 100 ng		
reaction buffer	10 µl (5 x HF buffer)			5 µl (10 x Thermo buffer)		
dNTPs (100 µM)	1 µl			1 µl		
forward primer (100 µM)	0.3 µl			0.5 µl		
reverse primer (100 µM)	0.3 µl			0.5 µl		
polymerase	0.25 µl (2 u / µl)			0.5 – 1 µl (5 u / µl)		
water	ad 50 µl			ad 50 µl		
Program						
Step	**Time**	**Temperature**	**Cycles**	**Time**	**Temperature**	**Cycles**
initialization	5 min	98 ℃	1	10 min	95 ℃	1
denaturation	10 sec	98 ℃		1 min	95 ℃	
annealing	30 sec	T_m – 0.5 to 1 ℃	30-39	30 sec	T_m – 0.5 to 1 ℃	34
elongation	30 sec/kbp	72 ℃		1 min/kbp	72 ℃	
final elongation	10 min	72 ℃	1		72 ℃	1

For screening of bacterial transformants in *E. coli*, cells were picked from plates and directly inoculated into tubes to serve as template in colony-PCR. In contrast, *H. pylori* mutant were validated by colony-PCR on purified gDNA of wildtype (control) and mutant strains. Detailed information of oligos used in colony-PCR reactions are listed in Appendix, Table 13.4. In general, colony-PCR reactions and/or reactions for sequencing of plasmids/mutants were performed by *Taq* DNA polymerase, whereas the amplification PCR fragments used for *H. pylori* cloning (construction on plasmid or overlap-PCR) were carried

out using Phusion DNA polymerase as it possesses improved proof-reading characteristics. PCR products were purified using the NucleoSpin Gel and PCR Clean-up kit according to the manufacturers' instructions. Single-nucleotide exchanges, the deletion or addition of sequence stretches were introduced by amplification of the original plasmids and self-ligation of purified PCR products amplified by cycle-PCR.

6.3.4. Agarose gel electrophoresis

Agarose gels were used to separate DNA fragments. At a gel solution temperature of 50-60 °C, Midori Green DNA strain was added (4 µl/100 ml). Prior to loading, nine volumes of sample were mixed with one volume of 10 x DNA loading dye. Gels were run in 1 x TAE buffer at 100 to 150 V for about 30-60 min (according to fragment size).

6.3.5. Restriction digest, *DpnI* digestion and DNA ligation

PCR fragments amplified from plasmids were incubated with *DpnI* for 3 hrs at 37 °C to remove remnants of the template plasmid. All other restriction enzyme digests were performed in the buffers from New England Biolabs and Fermentas according to the conditions suggested by the manufacturers. Digested DNA fragments and linearized vectors (blunt or sticky-end) were ligated by T4 DNA Ligase either 1 h at 37 °C or over-night at 16 °C prior to transformation into chemically competent *E. coli*.

6.3.6. Sequencing of DNA

Determination of the G-repeat length in the *tlpB* leaders from sequential *H. pylori* isolates. The lengths of the G-repeat in the 5' UTR of the *tlpB* mRNA from different *H. pylori* strains that are listed in Appendix, Table 13.1 were extracted from NCBI database (if available) or from published literature. Furthermore, *tlpB* 5' UTR sequences of sequential *H. pylori* isolates with available 454 genome sequences from the study of Kennemann *et al.*, 2011 were re-sequenced using Sanger sequencing (PCR products amplified by Phusion DNA polymerase). The genomic DNA of these strains was kindly provided by Sebastian Suerbaum and Christine Josenhans, Medizinische Hochschule Hannover, Germany.

6.4. RNA techniques

6.4.1. RNA preparation

If not mentioned otherwise, *H. pylori* was grown in liquid culture to mid-exponential growth phase ($OD_{600\,nm}$ of ~ 0.7 – 1.0) and cells corresponding to an OD_{600nm} of 4 were harvested, mixed with 0.2 volumes of Stop mix (95 % ethanol/5 % phenol) and immediately shock-frozen in liquid nitrogen. Depending on the experiment, TRIzol method (rifampicin assay, MS2 affinity chromatography) or Hot Phenol method (northern blot, primer extension, RNA-sequencing, qRT-PCR etc.) were applied for RNA extraction.

TRIzol method. For RNA preparation, snap-frozen cells were thawed on ice and cells were collected by centrifugation (15 min, 4,000 rpm, 4°C). The supernatant was removed and the pellet was resuspended in 1 ml TRIzol. The sample was transferred to a phase lock tube and 400 µl chloroform was added. The mixture was vigorously shaken, incubated for 5 min at room temperature (RT) and centrifuged (15 min, 13,000 rpm, 4°C). The aqueous layer was transferred into a fresh tube and RNA was precipitated by addition of 450 µl isopropanol for 30 min at RT. Samples were centrifuged (30 min, 13,000 rpm, 4°C) and the supernatant was removed. RNA pellets were washed with 75 % ethanol and air-dried. RNA pellets were resuspended in sterile water.

Hot Phenol method. Total RNA was extracted as described previously (Blomberg *et al.*, 1990, Sharma *et al.*, 2010). Frozen bacterial cell pellets were thawed on ice, centrifuged for 10 min at 4,000 rpm at 4 °C and resuspended in 600 µl TE buffer (pH 8.0) containing 0.5 mg/ml lysozyme and 1 % (w/v) SDS. The suspension was mixed by inversion and incubated at 64 °C in a water bath for 2 min. The pH was equilibrated by addition of 66 µl of sodium acetate (pH 5.2), and samples were mixed with 750 µl phenol. Tubes were incubated at 64 °C for 6 min and frequently mixed (every 30 sec). Upon 1 to 5 min chilling on ice, samples were centrifuged (15 min, 13,000 rpm, 4 °C) to ensure phase separation. The aqueous layer was transferred to a phase lock tube, mixed with 750 µl chloroform and centrifuged (10 min, 13,000 rpm, 4 °C). RNA was precipitated from the aqueous layer by addition of two- to three volumes of a 30:1 ethanol:sodium acetate (pH 6.5) mix and incubation for at least 3 hrs at -20 °C. RNA pellets were washed with 75% ethanol and air-dried. RNA was resuspended in sterile water. For RNA analysis by denaturing PAGE (northern blot), RNA was diluted in water and RNA loading dye GL II. RNA quality was checked by TBE gel electrophoresis (1 µg RNA (in GL II) separated on 1 % TBE gel; run in 1 x TBE buffer).

6.4.2. Nortern blot analysis – denaturing PAA gel electrophoresis

For northern blot analysis, 5 – 10 μg total RNA and ^{32}P-labeled pUC Mix Marker 8 were separated on 6 % polyacrylamide (PAA) gels containing 7 M urea. Prior to loading, RNA samples (in GL II) and marker were denatured at 95 °C for 5 min and chilled on ice for 5 min. Gels were run in the presence of 1 x TBE at 300 V at room temperature, and were subsequently blotted to Hybond-XL membranes by electro-blotting (1 h 15 min, 50 V, 4 °C) in a tank electroblotter in 1 x TBE buffer. After blotting, total RNA was UV cross-linked (120 mJ) to the membrane and hybridized with 5' end-labeled (^{32}P) DNA oligonucelotides as described previously (Urban & Vogel, 2007). Briefly, membranes were pre-hybridized in 15 ml Roti-Hybri-Quick at 42 °C for 1 h prior to addition of gene-specific [32]P-5' end-labeled DNA oligonucleotide (Table 6.7). Hybridization with gene-specific northern blot probe was performed at least one hour at 42 °C. Following hybridization, membrane was washed in three subsequent steps using SSC buffer (20 min – 5 x SSC, 15 min – 1 x SSC and 15 min 0.5 x SSC). Membranes were sealed and exposed to imaging plates. Signals were determined on a Typhoon FLA 7000 PhosphoImager (Fuji) and band intensities quantified with AIDA software (Raytest, Germany).

6.4.3. DNase I treatment

To remove contaminations of genomic DNA, RNA samples for primer extension, qRT-PCR, microarray and RNA-seq analysis were digested with 1 u DNase I per 4 μg RNA for 45 min at 37 °C. RNA was purified by P:C:I extraction and precipitated with three volumes of 30:1 ethanol:sodium acetate (pH 6.5) mix. Complete removal of gDNA was verified by control-PCR using *Taq*-polymerase and JVO-0352 x JVO-0353 on DNase I-treated RNA sample. RNA quality was checked by TBE gel electrophoresis.

6.4.4. Quantitative real-time PCR

All qRT-PCR experiments were carried out in technical and biological triplicates on a CFX96 system (Biorad) using Power SYBR Green RNA-to-C_T™ 1-Step kit (Applied Biosystems) according to the manufactures' instructions. For each reaction (10 μl final volume), 2 μl of RNA sample (50 ng/reaction) was mixed with 0.1 μl of primer pairs (0.1 μM final), 0.08 μl of RT enzyme mix (125 x) and 5 μl Power SYBR Green RT-Mix (2 x). Reaction conditions were: 30 min 48 °C, 10 min 95 °C, and 50 cycles at 95 °C for 15 sec, 59 °C for 1 min, followed by a denaturing step at 95 °C for 15 sec and melting curve detection were performed by stepwise increase of the temperature from 59 °C to 95 °C (0.5 °C/15 sec). The oligo sets used for qRT-

PCR analysis are listed in Table 6.7 and their specificity had been confirmed by Primer-BLAST (NCBI). Please note that CSO-1173/CSO-1174 for 6S RNA served as internal standard. Fold-changes were determined by the $2^{(-\Delta\Delta CT)}$-method (Livak & Schmittgen, 2001).

6.4.5. Determination of RNA stability - rifampicin assay

To determine the half-life of the *tlpB* mRNA in wildtype, *ΔrepG* and RepG complementation (C$_{RepG}$) strains of *H. pylori* strains 26695 and G27, cells were grown to an OD$_{600nm}$ of 1.0 and treated with rifampicin (final concentration 500 μg/ml). RNA samples were harvested prior to and at indicated time points (0, 1, 2, 4, 8, 16, and 32 min) after rifampicin treatments. RNA decay was analyzed by quantitative RT-PCR analysis as described previously (see above; Sharma *et al.*, 2010, Papenfort *et al.*, 2006, Papenfort *et al.*, 2012).

6.4.6. Primer extension (for RepG)

About 10 μg DNase I treated RNA from *H. pylori* strains 26695 and G27 was used for primer extension. For the sequencing ladder, 1.5 pmol 5'end-labeled JVO-5126 and about 100 ng PCR product, which was amplified from gDNA of *H. pylori* strains 26695 and G27 using primer pairs JVO-5126 x CSO-0083, were used in sequencing reactions with the SequiTherm EXCEL™ II DNA Sequencing kit (according to the manufacturers' instructions). Primer extension was performed in 10 μl reactions. After heating (80 °C), 10 μg RNA and 0.5 pmol 5'end-labeled JVO-5126 were annealed by gradual cooling of the reaction to 42 °C. Reverse transcription was performed in 1 x AMV reaction buffer by addition of 10 mM dNTPs and AMV Reverse Transcriptase. After incubation for 1 h at 42 °C, the reaction was stopped by addition of 10 μl RNA loading buffer (GL II). 10 μl of the primer extension reactions and 2 μl of the ladder with 8 μl GL II were separated on 6 % PAA/7M urea sequencing gels and monitored by the PhosphoImager Typhoon FLA 7000.

6.4.7. Gel-mobility shifts, *in-vitro* structure probing, and in-line probing assays

***In-vitro* transcription and 5'-end labeling.** DNA templates that contain the T7 promoter sequence were generated by PCR using oligos and DNA templates listed in Table 6.11. T7 transcription was carried out using the MEGAscript® T7 kit (Ambion) according to the manufactures' instructions (input ∼ 200 ng DNA). Sequences of the resulting T7 transcripts are listed in Table 6.12. For *tlpB* mRNA leader variants (6G-16G) see also Table 6.10. Following P:C:I extraction and precipitation by three volumes of 30:1 ethanol:sodium

acetate (pH 6.5) mix, the correct size and integrity of the *in-vitro* transcribed RNAs were quality checked by denaturing PAGE. Gels were stained using Stains-All solution.

Table 6.11: Details of RNAs used for *in-vitro* work.

T7 RNA	DNA template plasmid/ gDNA of CSS...	Oligonucleotides	For mRNAs: 5' part to AUG [nt]	For mRNAs: 3' part to AUG [nt]	Size of the T7 transcripts [nt]
RepG	CSS-0004	JVO-5125 x JVO-5126			87
RepG ΔCU	CSS-0157	JVO-5125 x JVO-5126			58
RepG 3xG	CSS-0158	JVO-5125 x JVO-5126			87
RepG 1xG	CSS-0159	JVO-5125 x JVO-5126			87
***tlpB* leader**	CSS-0004	JVO-5127 x JVO-5953	-139	+78	217
***tlpB* ΔG leader**	CSS-0385	JVO-5127 x JVO-5953	-127	+78	205
***tlpB* 3xC leader**	CSS-0386	JVO-5127 x JVO-5953	-139	+78	217
***tlpB* 1xC* leader**	CSS-0385	JVO-5127 x JVO-5953	-139	+78	217
***tlpB* 6G leader**	CSS-0472	JVO-5127 x JVO-5953	-133	+78	211
***tlpB* 7G leader**	CSS-0473	JVO-5127 x JVO-5953	-134	+78	212
***tlpB* 8G leader**	CSS-0474	JVO-5127 x JVO-5953	-135	+78	213
***tlpB* 9G leader**	CSS-0475	JVO-5127 x JVO-5953	-136	+78	214
***tlpB* 10G leader**	CSS-0476	JVO-5127 x JVO-5953	-137	+78	215
***tlpB* 11G leader**	CSS-0477	JVO-5127 x JVO-5953	-138	+78	216
***tlpB* 13G leader**	CSS-0478	JVO-5127 x JVO-5953	-140	+78	218
***tlpB* 14G leader**	CSS-0479	JVO-5127 x JVO-5953	-141	+78	219
***tlpB* 15G leader**	CSS-0480	JVO-5127 x JVO-5953	-142	+78	220
***tlpB* 16G leader**	CSS-0481	JVO-5127 x JVO-5953	-143	+78	221
***tlpB* 5th::*gfpmut3* mRNA**	pSP109-6	JVO-5127 x CSO-0441	-139	+775	917
***cagA* 28th::*gfpmut3* mRNA**	pMA5-2	CSO-0278 x CSO-0441	-105	+844	949
***tlpB*::3xFLAG mRNA**	CSS-0464	JVO-5127 x JVO-5143	-139	+1768	1907
***tlpB* ΔG::3xFLAG mRNA**	CSS-0471	JVO-5127 x JVO-5143	-127	+1768	1895
***tlpB* 10G::3xFLAG mRNA**	CSS-0476	JVO-5127 x JVO-5143	-137	+1768	1905
***tlpB* 11G::3xFLAG mRNA**	CSS-0477	JVO-5127 x JVO-5143	-138	+1768	1906
***tlpB* 13G::3xFLAG mRNA**	CSS-0478	JVO-5127 x JVO-5143	-140	+1768	1908
***tlpB* 14G::3xFLAG mRNA**	CSS-0479	JVO-5127 x JVO-5143	-141	+1768	1909

Table 6.12: Sequences of T7 transcripts. The C/U-rich *tlpB* binding site of RepG and the G-repeat in the 5' UTR of *tlpB* mRNA are underlined. Mutations in the sRNA-mRNA interaction sites are indicated by different colors. The RBS and start codon (AUG) of the *tlpB* mRNA are marked in light green and the ORFs are shown in gray letters. The MS2-tag is highlighted in purple. Sequences of T7 RNAs generated from *gfpmut3* reporter fusion constructs are shown in Table 6.9.

Name	Sequence 5' → 3'
RepG WT	AUCCAACCAUUCCUUAUGGUUUGGUUGGCACCGCUAAGAUUGAAGGGUCACCUCCCCCUCCUUUCCCUUUGUCUUGGCGGUUGUUUU
RepG SL 2	ACCGCUAAGAUUGAAGGGUCACCUCCCCCUCCUUUCCCUUUGUCUUGGCGGUUGUUUU
RepG ΔCU	AUCCAACCAUUCCUUAUGGUUUGGUUGGCACCGCUAAGAUUGAAGGGUACGCCCUUUGUCUUGGCGGUUGUUUU
RepG 3xG	AUCCAACCAUUCCUUAUGGUUUGGUUGGCACCGCUAAGAUUGAAGGGUCACGUCCGCCUGCUUUCCCUUUGUCUUGGCGGUUGUUUU
RepG 1xG*	AUCCAACCAUUCCUUAUGGUUUGGUUGGCACCGCUAAGAUUGAAGGGUCACCUCCGCCUCCUUUCCCUUUGUCUUGGCGGUUGUUUU
tlpB WT leader	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
tlpB ΔG leader	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
tlpB 6G-16G leader	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUC−(G)_{6−16}−UGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
tlpB 3xC leader	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGCGGGCGGGCGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
tlpB 1xC* leader	UGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGCGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
MS2-RepG	GGUCGUACACCAUCAGGGUACGUUUUUCAGACACCAUCAGGGUCUGGUCACAUCCAACCAUUCCUUAUGGUUUGGUUGGCACCGCUAAGAUUGAAGGGUCACCUCCCCCUCCUUUCCCUUUGUCUUGGCGGUUGUUUU
MS2-*tlpB* (26695)	GGUCGUACACCAUCAGGGUACGUUUUUCAGACACCAUCAGGGUCUGGUCACUGUUUGUUCUUUUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUGUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUAGGUUUAGGG
MS2-*tlpB* (G27)	GUCGUACACCAUCAGGGUACGUUUUUCAGACACCAUCAGGGUCUGGUCACUGUUUGUUCUCUGUUUCGUUUUCAAACAACCGGGUUUUAAUUUUAUUUUGUGCCACUCAUUUUUCGGGGGGGGGGGGGUGCAUUUAGAAGCUAAACUCUAAAAUUAGGGUUUGACUUAAAAAUGAUUUAUUAGGAGAUAAAUGAUGUUUUUCUUCAAUGUUUGCUUCGUUGGGGACUCGUAUCAUGCUGGUCGUGUUAGCCGCUCUUUUUGGGUUUAGGG

For 5' end-labeling (³²P), 20 pmol RNA were dephosphorylated by Calf Intestinal Phosphatase (10 u) or Antarctic Phosphatase treatment (10 u) at 37 °C for 1 h. Following P:C:I extraction, the RNA was precipitated by three volumes of 30:1 ethanol:sodium acetate (pH 6.5) mix in the presence of 1 µl GlycoBlue™ at -20 °C for at least 3 hrs. Next, the RNA was 5' phosphorylated at 37 °C for 1 h by PNK (1 u) in the presence of 20 µCi ³²P-ATP. Unincorporated nucleotides were removed using Microspin G-50 or G-25 columns, and labeled RNA was separated by denaturing PAGE. Upon exposure of the gel to imaging plates and determination of signals by a PhosphoImager, RNA was cut and recovered from the gel

by elution at 4 °C over-night in RNA elution buffer. Labeled RNA was purified by P:C:I extraction, and quantified by NanoDrop measurement. For further details see also Sittka *et al.*, 2007 and Papenfort *et al.*, 2006. DNA oligonucleotides were ^{32}P-labeled as described previously in Urban & Vogel, 2007.

EMSA (electrophoretic mobility shift assay for RNA-RNA interactions). Gel-mobility shift assays were performed with about 0.04 pmol 5'-end labeled RNA (4 nM final concentration) and increasing amounts of unlabeled RNA in 10 µl reactions. After denaturation (1 min at 95 °C), labeled RNAs were cooled for 5 min on ice and 1 µg yeast RNA as well as 10 x RNA structure buffer were added. Increasing concentrations of unlabeled RNA were added to final concentrations of 8 nM, 16 nM, 62.5 nM, 125 nM, 250 nM, 500 nM and 1,000 nM. After incubation for 15 min at 37 °C, samples were immediately loaded after addition of 3 µl 5 x native loading dye to a native 6 % PAA gel. Gel electrophoresis was done in 0.5 x TBE buffer at 300 V. To avoid heating, the gel apparatus was connected to a water cooling system that maintained the gel at a constant temperature of 4 °C. Afterwards, gels were dried and analyzed using a PhosphoImager (FLA 7000 series) and AIDA software.

For gel-shift assays with ^{32}P-labeled RepG and *tlpB* mRNA leader variants, about 0.04 pmol ^{32}P-labeled RepG was incubated with 1,000 nM of unlabeled *tlpB* that either lacks the G-repeat (ΔG) or comprises different G-stretch length (6G-16G).

Structure probing. For structure probing and footprinting assays, about 0.1 pmol 5' end-labeled RNA was subjected in absence or presence of unlabeled target mRNA or sRNA to RNase T1, lead (II)-acetate or RNase III treatment in 10 µl reactions as previously described (Sharma *et al.*, 2007). In brief, about 0.1 pmol 5' end-labeled RNA was denatured for 1 min at 95 °C and chilled on ice for 5 min. Next, 1 µg yeast competitor RNA and 10 x RNA structure buffer were added (provided together with RNase T1, Ambion). Unlabeled RepG/*tlpB* mRNA leader wild-type or mutant RNAs were added at 10- or 100-fold excess (see figure legends). After incubation for 10 min at 37 °C, 2 µl RNase T1 (0.01 u/µl) or 2 µl freshly prepared lead (II)-acetate solution (25 mM) were added and reactions were incubated for 3 min or 90 sec, respectively. RNase III cleavage reactions were performed for 6 min at 37 °C in 1 x RNA structure buffer containing 1 mM DTT and 1.3 u/µl enzyme. For RNase T1 ladders, about 0.2 pmol labeled RNA were denatured in 1 x RNA structure buffer for 1 min at 95 °C and afterwards incubated with 0.1 u/µl RNase T1 for 5 min. OH ladders were generated by the incubation of about 0.2 pmol labeled RNA in 1 x alkaline hydrolysis buffer for 5 min at 95 °C. All reactions were stopped by addition of 12 µl RNA loading buffer GL II on ice. Ladders and samples were denatured 3 min at 95 °C and separated on pre-warmed 6-15 % PAA/7M urea sequencing gels. Gels were run in the presence of 1 x TBE

buffer at 40 W at room temperature. Afterwards, gels were dried and analyzed using the PhosphoImager Typhoon FLA 7000 and AIDA software.

In-line probing. For in-line probing assays (Regulski & Breaker, 2008), about 0.2 pmol labeled RNA (20 nM final concentration) was incubated in absence or presence of 20 nM or 200 nM unlabeled sRNA or mRNA for 40 hrs at room temperature in 1x in-line probing buffer. For RNase T1 ladders, about 0.2 pmol labeled RNA was incubated in 0.25 M sodium citrate buffer (pH 5.0 at 23 °C) with 1 u/µl RNase T1 for 5 min at 55 °C. For alkaline ladders, about 0.2 pmol labeled RNA was denatured for 5 min at 95 °C in Na_2CO_3 buffer. All reactions were stopped by adding 10 µl colorless gel-loading solution on ice. Cleavage products were analyzed on 6-10 % PAA gels under denaturing conditions and visualized as described above.

6.4.8. *In-vitro* toeprint assays with 30S ribosomal subunit

Preparation of sequencing ladder. Sequencing ladders were generated with CycleReader™ DNA Sequencing kit according to the manufacturer's protocol on the same DNA template used for T7 transcription and with the same 5' end-labeled primer (JVO-5953) as in the toeprint reactions.

In-vitro **30S toeprint experiments** were carried out as previously described (Hartz et al 1988; Udekwu 2005) with few modifications. For annealing of the primer, an unlabeled mRNA fragment (0.2 pmol) and 0.5 pmol 5' end-labeled oligonucleotide (JVO-5953, binds about 78-nt downstream of AUG in *tlpB*) were denatured in the presence of 0.8 µl SB 5 x – Mg in a total volume of 3 µl at 90 °C for 1 min and adjacently chilled on ice for 5 min. Next, 1 µl dNTPs (5 mM each) and 1 µl SB 1 x Mg60 were added, and samples were shifted to 37 °C. Reactions were incubated for 5-10 min when either unlabeled sRNA (100, 200 and 1,000 nM final, see figure legends) or water were added. Following this, samples were mixed with 2 pmol purified 30S ribosomal subunit (or SB 1 x Mg10 for the control). After 5 min, 10 pmol uncharged tRNA[fMet] (or SB 1 x Mg10 for the control) was added to the samples and reactions were incubated for additional 15 min before reverse transcription was initiated by addition of 100 u SuperScriptII Reverse Transcriptase. Following cDNA synthesis for 20 min, 10 µl reactions were stopped with 100 µl Toeprint stop solution. Following P:C:I extraction, RNA was digested by alkaline hydrolysis in the presence of potassium hydroxide at 90 °C for 5 min, and cDNA was precipitated in the presence of 1 µg GlycoBlue™ and 3 M acetic acid by addition of three volumes of 30:1 ethanol:sodium acetate (pH 6.5) mix. After precipitation and washing, samples were resuspended directly in 10 µl GL II. cDNA samples were denatured prior to loading (95 °C, 2 minutes) and separated in the presence of gene specific

sequencing ladders by denaturing PAGE on 6-10 % sequencing gels at constant power of 40 W. Gels were dried and signals were determined as described above.

6.4.9. Identification of RNA-protein complexes using MS2-tagged RNAs

Affinity chromatography of 5'-end MS2-tagged RepG and *tlpB* mRNA leaders (26695 and G27) was performed according to Rieder *et al.*, 2012. The DNA templates used for *in-vitro* transcription of MS2-tagged RNAs were generated by overlap-PCR of two products: (I) 5' MS2-tagged PCR fragment containing the T7 promotor and the MS2-tag (PCR with JVO-4201/4203 on a template containing the MS2 tag, *e.g.* pRR05), and (II) PCR fragment containing part of the MS2 tag and RNA of interest (*e.g.* RepG, PCR with JVO-4204 x JVO-5126 on genomic DNA of *H. pylori* 26695). PCR products were mixed in equimolar ratios and overlap-PCR was performed with JVO-4202 and RNA specific oligonucleotide, *e.g.* JVO-5126. Oligonucleotides and template DNAs are listed in Table 6.13. T7-RNA sequences are shown in Table 6.12.

Table 6.13: Construction of T7 DNA templates for *in-vitro* transcribed RNAs used for affinity chromatography.

In-vitro synthesized tagged RNA	Oligonucleotides for amplification of the MS-tag/ template	Oligonucleotides for amplification of the RNA/template	Oligonucleotides for amplification of T7 DNA template of tagged RNA via overlap-PCR	Size of the T7 transcripts [nt]
RepG-MS2	JVO-4201 x JVO-4203/ pRR05	JVO-4204 x JVO-5126/ gDNA of CSS-0004	JVO-4202 x JVO-5126	138
MS2-*tlpB* (26695)	JVO-4201 x JVO-4203/ pRR05	CSO-0511 x JVO-5953/ gDNA of CSS-0004	JVO-4202 x JVO-5953	268
MS2-*tlpB* (G27)	JVO-4201 x JVO-4203/ pRR05	CSO-0523 x JVO-5953/ gDNA of CSS-0010	JVO-4202 x JVO-5953	270

The MS2-aptamer specifically binds to the MS2 coat protein (MS2), which is fused to a bait protein, namely the maltose binding protein (MBP), and thus, can be immobilized to a matrix of a column. The MS2-MBP was purified as described previously in Said *et al.*, 2009 and references therein. Briefly, MS2-MBP was expressed in *E. coli* and purified in two steps: (I) over an amylose column in 20 mM HEPES (pH 7.9), 200 mM KCl, 1 mM EDTA, and (II) over a heparin column using a KCl gradient in the same buffer. Purity of the MS2-MBP was verified by SDS-PAGE and Coomassie staining.

For affinity purification of RNA binding proteins, *H. pylori* strain 26695 was grown in liquid culture to an exponential growth phase (OD_{600nm} of $\sim$ 0.7 – 1.0). Bacteria

corresponding to an OD_{600nm} of 50 were chilled 20 min on ice and harvested by centrifugation (30 min, 2,900 x g, 4 °C). The supernatant was discarded and cells were resuspended in 2 ml lysis buffer (buffer A). Following an additional centrifugation step (5 min, 11,200 x g, 4 °C), cells were resuspended in 600 µl buffer A. An equal volume (~ 750 µl) of glass beads (0.1 mm) was then added to the cell suspension. Cells were lysed using a Retsch MM400 ball mill (30 kHz, 10 min) in blocks pre-cooled to 4 °C, and centrifuged for 10 min at 15,200 x g and 4°C to remove cell debris. The supernatant (cell lysate) was transferred to a new 1.5 ml reaction tube and 300 pmol of *in-vitro* transcribed RNA (MS2-RepG, MS2-*tlpB* mRNA leaders or MS2-tag alone) was added. Cell lysate/RNA mixture was incubated for 1-2 hrs at 4 °C by rotation (rotator – SB2 STUART). All flowing steps of the affinity purification were performed at 4 °C. For preparation of the affinity column, 100 µl amylose resin was applied to Bio-spin disposable chromatography columns. The prepared column was washed two times with the lysis buffer (buffer A). Next, 600 pmol MS2-MBP diluted in 1 ml buffer A was immobilized on the amylose resin, and the column was washed again with 2 ml of buffer A. Subsequently, the mixture of lysate and *in-vitro* transcribed RNA was loaded onto the column, followed by three washes with 2 ml buffer A. For RNA analysis, aliquots equivalent to 17 pmol *in-vitro* transcribed RNA of the lysate, the flow-through and the wash fractions were mixed with elution buffer (buffer B = buffer A + 12 mM maltose) to a final volume of 900 µl and stored on ice until RNA extraction. For protein analysis, aliquots equivalent to 1.0 OD_{600nm} were mixed with protein loading buffer to a final volume of 100 µl (0.01 OD_{600nm}/µl). Finally, RNA and proteins were eluted from the column with 900 µl (2 x 450 µl) of buffer B. Eluted RNA and those of the lysate, flow-through and wash fractions was extracted with P:C:I followed by precipitation of the aqueous phase with three volumes of 30:1 ethanol:sodium acetate (pH 6.5) mix. RNA was resolved in 50 µl water and stored at -20 °C. For protein isolation, the organic phase was subjected to acetone precipitation at - 20 °C over-night. The pellet was washed twice with acetone and air-dried. The pellet was resolved in 20 µl 1 x protein loading buffer, resulting in about 0.5 OD_{600nm}/µl. RNA samples were used for northern blot analysis. Proteins were denatured for 2 min at 98 °C, and separated by SDS-PAGE using 12-15% polyacrylamide gels (see below). Proteins were visualized by silver-staining (see below).

6.5. Protein techniques

6.5.1. Preparation of whole protein fractions

If not mentioned otherwise, cells corresponding to an OD_{600nm} of 1 from *H. pylori* grown to mid-exponential growth phase were collected by centrifugation at 16,100 x g at 4 °C for

2 min. Cell pellets were dissolved in 100 μL of 1 x protein loading buffer to a final concentration of 0.01 OD_{600nm}/μl. After denaturing at 95 °C for 8 min, protein samples were stored at -20 °C.

6.5.2. One-dimensional SDS-PAGE, Coomassie and silver-staining

After boiling for 2 minutes at 98 °C, whole protein samples corresponding to 0.1 OD_{600nm} were separated according to their electrophoretic mobility by 10-15 % (v/v) one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). For whole protein analysis using Coosmassie staining, SDS-PAA gels were incubated carefully shaking in PAGE Blue staining solution over-night according to the manufacturers' instructions. Background staining of the gel matrix was removed by serial washing steps with water.

For protein detection by silver-staining, about 0.05 to 3.5 OD_{600nm} (see figure legends) were separated on one-dimensional SDS-PAGE and stained according to Jungblut & Seifert, 1990. In brief, gels were fixed over-night in 50 % methanol, 12 % acetic acid and 0.5 ml/l formaldehyde (37 %). Following three washing steps (2 x 20 min with 50 % EtOH, 1 x 20 min 30 % EtOH), gels were incubated for exactly 1 min in 200 ml 0.2 mg/ml sodium thiosulfate ($Na_2S_2O_3$, pretreatment solution). Gels were rinsed three times for 20 sec with water and incubated for 20 min with the silver solution. After 2 x 20 sec washing steps with water, protein staining was developed by incubation of gels in developing solution (Na_2CO_3, $Na_2S_2O_3$, formaldehyde (37 %)) until protein bands were clearly visible. Finally, silver-staining was stopped by addition of 1 % glycine (Stop solution).

6.5.3. One-dimensional SDS-PAGE and western blot

For western blot analysis, protein samples corresponding to an OD_{600nm} of 0.005, 0.01 and 0.1 (see figure legends) were separated by 12-15 % (v/v) SDS-PAGE and transferred to a PVDF membrane. If necessary, PVDF membranes were activated by incubation in methanol (90 s), H_2O (5 min), and transfer buffer (5 min). Gels were blotted for 1 h 30 min at 2 mA/cm² membrane in a semi-dry blotter onto PVDF membrane in transfer buffer. After rinsing in 1 x in TBS-T buffer, membranes were blocked for 1 h with 10 % (w/v) milk powder/TBS-T and incubated over-night with primary antibody at 4 °C. Afterwards, membranes were washed with TBS-T, followed by 1 h incubation with secondary antibody linked to horseradish peroxidase. After additional washing steps, chemiluminescence was detected using ECL-reagent and Image Quant LAS 4000. Details about the used antisera and antibodies are listed in Table 6.6.

6.5.4. *In-vitro* translation assay

Translation reactions were carried out with PureSystem according to the manufactures' instructions. In brief, 1 pmol of *in-vitro* transcribed mRNAs (*tlpB* 5th::*gfpmut3*, *cagA* 28th::*gfpmut3*, *tlpB*::3xFLAG, *tlpB* ΔG::3xFLAG, *tlpB* 10G::3xFLAG, *tlpB* 11G::3xFLAG, *tlpB* 13G::3xFLAG and *tlpB* 14G::3xFLAG) were denatured in the absence or presence of 1, 10, 50 and 100 pmol of RepG or RepG mutants (ΔCU, 3xG, 1xG; see also figure legends) for 1 min at 95 °C and chilled for 5 min on ice. The mRNA and sRNA were pre-incubated for 10 min at 37 °C before addition of PureSystem mix and translation was performed at 37 °C for 30 min. Reactions were stopped by addition of 60 µl acetone, chilled for 15 min on ice and proteins were collected by centrifugation for 10 min at 10,000 x g and 4 °C. *In-vitro* translated TlpB and CagA were quantified by western blot analysis using monoclonal anti-FLAG or anti-GFP and anti-mouse IgG antibodies (see Table 6.6). The ribosomal protein S1 served as a loading control and was detected by an S1 antibody, and anti-rabbit secondary antibody.

6.6. Additional biochemical techniques

6.6.1. EMSA for DNA-protein interaction – P$_{repG}$ and HP1043

EMSA was performed using the LightShift Chemiluminescent EMSA kit according to the manufacturers' instructions. The 145 bp biotin-labeled DNA probe (P$_{repG}$*) was prepared by PCR with primer pair S5490-5btn/S5490-3 and gel purification of the PCR product using the GeneJET™ Gel extraction kit. The unlabeled competitor DNA fragments SC (83 bp) and *rpoA* (206 bp) were PCR-amplified with primer pairs S5490-5/MO-1 and *rpoA*-5/*rpoA*-3, respectively, while competitor fragment USC was obtained by annealing of the complementary 59mer oligonucleotides USC-1 and USC-2. Chromosomal DNA of *H. pylori* 26695 was used as template for all PCR amplifications.

6.6.2. LPS-staining and Lewis x antigen analysis

H. pylori cells were grown to exponential phase in BHI medium. Cells corresponding to an OD$_{600nm}$ of 1.0 were harvested by centrifugation (2 min, 16,100 x g, 4 °C) and resuspended in 100 µl 1 x protein loading dye. Protein samples were boiled for 10 min at 98 °C. After cooling down to room temperature, 30 µl of the protein sample were discarded. Exact 6 µl of 5 x protein leading dye and 24 µl Proteinase K solution (20 mg/ml) were added (final volume 100 µl) and protein digestion was performed for 1 h 30 min at 60 °C. Finally, Proteinase K was heat-inactivated by boiling of the samples for 10 min at 98 °C. LPS samples were stored at -20 °C.

180

Lipopolysaccharide structures from *H. pylori* were visualized by silver-staining according to (Moran *et al.*, 2002). Briefly, about 10 to 12 µl Proteinase K-digested (LPS) samples were loaded to 15 % SDS-PAGE and separated by gel electrophoresis for about 3 hrs at 150 to 300 V in 1 x SDS running buffer. After rinsing with water, gels were fixed over-night in 25 % isopropanol and 7 % acetic acid (1 x fixing solution) and incubated in 100 ml sodium periodate solution (oxidizing solution, 0.7 % (w/v) Na_5IO_6 in 1 x fixing solution) for 15 min. Following washing steps with water (3 x 30 min), the silver staining solution (0.35 % ammonia, 0.02 N NaOH, 0.4 % (w/v) silver nitrate) was applied with vigorous agitation for 10 min. Gels were subsequently washed three times for 10 min in water and developed using a solution containing 2.5 % (w/v) sodium carbonate and 0.01 % (v/v) formaldehyde (37 %) (developing solution). Upon completion, 50 mM EDTA was used to stop the development.

For western blot analysis of Lewis x antigens, LPS samples (10 µl/well) were separated on 15 % SDS-PAGE and transferred to a PVDF membrane as described above. Immunoblotting was performed using monoclonal anti-Lewis x antibody (1:1,000 in 3 % BSA/TBS-T) and anti-mouse IgG conjugated to horseradish peroxidase (1:10,000 in 3 % BSA/TBS-T, Table 6.6). Signals were detected using ECL-reagent and Image Quant LAS 4000.

6.7. Transcriptome analyses

6.7.1. Microarray

Whole transcriptome analysis by microarrays was performed in collaboration with the Microarray Core Facility headed by Dr. H. Mollenkopf, Max Planck Institute for Infection Biology (MPI-IB), Berlin, Germany. Data were obtained from custom-made oligonucleotide arrays which comprise about 1,600 open reading frames, intergenic regions as well as regulatory RNA elements such as non-coding RNAs and 5' UTRs of *H. pylori* strain 26695 (provided by the Microarray Core Facility of the MPI-IB).

The experimental design of the microarray study involved the use of genomic DNA of *H. pylori* strain 26695 as the co-hybridized control for one channel on all microarrays (Cy5-labeled gDNA *vs.* Cy3-labeled cDNA (RNA) in the ratio 1:9). This method has the advantage of allowing the direct comparison of multiple samples with each other. Genomic DNA was isolated from *H. pylori* strain 26695 grown in liquid culture to mid-exponential growth phase (OD_{600nm} of 0.7) using Qiagen genomic DNA isolation buffer sets and Genomic-tip 100/G columns according to the manufacturers' instructions. In brief, cells corresponding to an OD_{600nm} of 15 were harvested by centrifugation (15 min, 2,900 x g, 4 °C) and pellets were shock frozen in liquid nitrogen. Cell pellets were thawed on ice and resuspended in 3.5 ml

buffer B1 containing 100 mg/ml RNase A, 25 mg/ml Proteinase K and 100 mg/ml lysozyme. After incubation for at least 30 min at 37 °C, 1.2 ml buffer B2 was added and the cell lysate was incubated for additional 30 min at 50 °C. Lysis samples were applied to Qiagen Genomic-tip 100/G columns equilibrated with 4 ml QBT buffer and subsequently washed with 2 x 7.5 ml QC buffer. Genomic DNA was eluted from the column in 5 ml QF buffer and precipitated by isopropanol. Following centrifugation (> 10,000 x g for 15 min, 4 °C), DNA was washed by 70 % EtOH, air-dried and resuspended in 250 µl water. Quantity and quality of the gDNA were checked by Nanodrop and analytical 1 x TAE gel.

RNA samples were collected in duplicates from *H. pylori* strain 26695 wildtype, Δ*repG* and complementation (C$_{RepG}$) grown in liquid culture to exponential growth phase (OD$_{600nm}$ of ~ 1.0). RNA was extracted by Hot Phenol method, digested with DNase I and quality-checked as described above. About 10 µg of DNase I-digested RNA was mixed with 5 µg random hexamers in a total volume of 9.4 µl, incubated at 70 °C for 5 min and subsequently chilled on ice for 10 min. About 4.6 µl reaction mix (2 µl of 10 x RT buffer, 2.0 µl DTT and 0.5 µl of 50 x dNTPs) together with 2 µl Cy3-dCTPs and 4 µl AffinityScript multi-temperature Reverse Transcriptase was incubated for 10 min at 25 °C, followed by an incubation step over-night at 42 °C. RNA was hydrolyzed by incubation of samples for 10 min at 70 °C in the presence of 0.1 M NaOH. Next, the solutions were neutralized by the addition of 15 µl of 0.1 M HCl and purified using the QIAquick PCR Purification kit following the manufacturers' instructions.

For labeling of the reference DNA, 20 µl of 2.5 x Random primer/reaction buffer mix (BioPrime) together with 2 µg of gDNA (in a final volume of 21 µl) was incubated at 95 °C for 5 min and subsequently placed on ice for 5 min. On ice, 5 µl 10 x dNTP mix (1.2 mM each dATP, dGTP, dTTP, 0.6 mM dCTP, 10 mM Tris pH 8.0, 1 mM EDTA), 3 µl Cy5-dCTP and 1 µl Klenow enzyme were added to the sample (total reaction volume 50 µl). The reaction mixture was incubated at 37 °C over-night and protected from light. Labeled DNA samples were purified by the QIAquick PCR Purification kit.

Each Cy3-labeled cDNA sample was mixed with Cy5-labeled chromosomal DNA. Hybridization and raw data generation were performed by the Microarray Core Facility of the MPI-IB, Berlin, Germany. In general, samples were hybridized to the arrays over-night at 65 °C according to the manufacturers' instructions. After hybridization, slides were washed and scanned by a G2565CA high resolution laser microarray scanner (Agilent Technologies). Raw microarray image data were analyzed with the Image Analysis/Feature Extraction software G2567AA (Version A.10.5.1, Agilent Technologies). For each array feature both 'Median-Signals' were background corrected by appropriate 'BG-Median-Signal' subtraction while signals < 10 were adjusted to 10. To compensate for unequal dye incorporation, data centering to zero was performed for each single microarray on one slide. Microarray data

were analyzed by Kai Papenfort using GeneSpring 7.3 (Agilent). Genes were considered to be significantly differentially expressed if they displayed $\geq$ 2-fold changes in both replicates and a p-value of $\leq$ 0.1 (*i.e.* significantly regulated between $\Delta repG$ vs. WT and C_{RepG}, but not between C_{RepG} vs. WT).

6.7.2. RNA-sequencing

RNA-seq analyses were applied for whole transcriptome profiling of the wildtype, $\Delta repG$ and RepG complementation (C_{RepG}) of *H. pylori* strains 26695 and G27. To allow for an unbiased comparison between both microarray and RNA-seq, the same RNA samples as mentioned above were converted into cDNA libraries and subjected to high-throughput sequencing by Illumina using the Genome Analyzer IIx. Libraries for Illumina sequencing of cDNA were constructed by *vertis* Biotechnology AG, Germany as described previously (Sharma *et al.*, 2010). Computational analysis of raw sequencing data were performed by automated RNA-seq processing pipeline READemption (Forstner *et al.*, 2014). In brief, after quality trimming of raw sequencing data (Illumnia reads in FASTQ were trimmed with a cut-off phred score of 20 by fast_quality_trimmer from FASTX toolkit), poly(A)-sequences (from cDNA library construction) and sequencing reads that are shorter than 20 nt were removed. Remaining cDNA reads were mapped to the reference genome sequences of *H. pylori* strains 26695 (NC_000915) and G27 (NC_011333 and NC_011334), respectively. Coverage plots in bam or wiggle format representing the number of aligned reads per nucleotide were generated based on the aligned cDNA reads and visualized in the Integrated Genome Browser (Nicol *et al.*, 2009). Each graph was normalized to the total number of reads which could be aligned from the respective library. To restore the original data range and prevent rounding of small error to zero by genome browsers, each graph was then multiplied by the minimum number of mapped reads calculated over all libraries.

6.8. Bioinformatics-based analyses

6.8.1. Gene-wise expression quantification and differential gene expression analysis

Based on annotation files from *H. pylori* strains 26695 and G27 (accession numbers see above), reads overlapping with at least 10 nt of a gene/CDS, rRNA, tRNA and/or sRNA were taken into account for differential gene expression analysis, *e.g.* $\Delta repG$ vs. wildtype. Also, reads that overlap in sense or anti-sense orientation with the annotated gene and/or RNA was well as reads overlapping with several annotations were considered and counted separately. The pairwise expression comparison (read counts of *e.g.* $\Delta repG$ vs. wildtype) was

performed on these gene quantifications with the READemption subcommand 'deseq', which relies on DESeq (version 1.0, see also Anders & Huber, 2010). Genes with a fold-change of at least 2.0 and p-values of ≤ 0.1 were considered to be significantly differentially expressed among two sequencing libraries (*i.e.* significantly regulated between $\Delta repG$ vs. WT and C_{RepG}, but not between C_{RepG} vs. WT).

6.8.2. Peak detection and RepG binding motif analysis

In order to automatically define cDNA reads enriched in the $\Delta repG$ libraries of *H. pylori* strain 26695, a peak calling algorithm was developed based on a sliding window approach (PEAK achu; T. Bischler & C. M. Sharma, unpublished). In this software, normalized wiggle files of the $\Delta repG$ and wild-type (WT) library were used as input to determine sites/sequence regions showing a continuous enrichment of cDNA reads in $\Delta repG$ compared to the wildtype. The identification of these so-called cDNA peaks is based on four parameters: (I) a minimum required fold-change ($\geq$ 3-fold, $\Delta repG$ vs. WT) for the enrichment, (II) a factor (1.5) multiplied by the 90th percentile of the wiggle graph, which reflects the minimum of required expression, (III) a window size in nt (20 nt) for which the minimum fold-change and expression values (I, II) were calculated in a sliding window approach, and (IV) a step size (5 nt) that defines the steps by which the window is moved along the genome. Peak detection was performed separately on leading and lagging strand of each replicon. Peak regions were available as annotations for manual inspection in the IGB or as sequence (see Appendix, Table 13.5).

   For the prediction of a consensus motif (putative RepG binding motif) based on sequences corresponding to peak regions found within putative RepG target mRNAs (RNA-seq) using MEME (Bailey *et al.*, 2009), the following parameters were applied: (I) single motif distribution in sequence: zero or one, (II) number of different motifs: 10, (III) minimum motif width: 6, and (IV) maximum motif width: 50. The peak sequences (70 peaks within 40 RepG target mRNAs) are shown in Table 13.5.

6.8.3. Alignments

Sequence alignments were calculated using MultAlin (http://multalin.toulouse.inra.fr/ multalin/multalin.html).

6.8.4. Predictions of RNA-RNA interactions

Potential interactions between two RNA molecules were predicted using RNAhybrid (Rehmsmeier *et al.*, 2004, http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/). Furthermore, genome-wide prediction of potential RepG target mRNA was performed by TargetRNA (Tjaden *et al.*, 2006) and CopraRNA (Wright *et al.*, 2013) using webpage interface: rna.informatik.uni-freiburg.de/CopraRNA/input.jsp.

6.8.5. RNA secondary structure predictions

Structure probing data were combined with bioinformatics-based RNA secondary structure prediction using RNAstructure (Mathews, 2006) to solve the complex structure of *repG* and *tlpB* mRNA leader.

6.8.6. Melting temperature of oligos

Melting temperatures of oligonucleotides were calculated according to Kibbe, 2007 using 'Oligo Calc', http://www.basic.northwestern.edu/biotools/oligocalc.html.

6.8.7. Software

Table 6.14: Software used in this study.

Software	Manufacturer
Adobe Acrobate Pro	Adobe Systems
Abode Photoshop CS5	Adobe Systems
AIDA	Raytest
Chromas LITE	Technelysium
CoralDraw X6	Corel
DESeq	Anders & Huber, 2010
Endnote	Thomas Reuters
FIMO	Grant *et al.*, 2011
Gene Spring 7	Aglient
Integrated genome browser	Affymetrix
PEAKachu	T. Bischler & Dr. C. M. Sharma, unpublished
Microsoft Windows and Office	Windows
MEME	Bailey *et al.*, 2009
READemption	Forstner *et al.,* 2014

7. References

Aberg, A., P. Gideonsson, A. Vallstrom, A. Olofsson, C. Ohman, L. Rakhimova, T. Boren, L. Engstrand, K. Brannstrom & A. Arnqvist, (2014) A repetitive DNA element regulates expression of the Helicobacter pylori sialic acid binding adhesin by a rheostat-like mechanism. *PLoS pathogens* **10**: e1004234.

Afonyushkin, T., B. Vecerek, I. Moll, U. Blasi & V.R. Kaberdin, (2005) Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic acids research* **33**: 1678-1689.

Agriesti, F., D. Roncarati, F. Musiani, C. Del Campo, M. Iurlaro, F. Sparla, S. Ciurli, A. Danielli & V. Scarlato, (2014) FeON-FeOFF: the Helicobacter pylori Fur regulator commutates iron-responsive transcription by discriminative readout of opposed DNA grooves. *Nucleic acids research* **42**: 3138-3151.

Aihara, E., C. Closson, A.L. Matthis, M.A. Schumacher, A.C. Engevik, Y. Zavros, K.M. Ottemann & M.H. Montrose, (2014) Motility and chemotaxis mediate the preferential colonization of gastric injury sites by Helicobacter pylori. *PLoS pathogens* **10**: e1004275.

Akada, J.K., K. Ogura, D. Dailidiene, G. Dailide, J.M. Cheverud & D.E. Berg, (2003) Helicobacter pylori tissue tropism: mouse-colonizing strains can target different gastric niches. *Microbiology* **149**: 1901-1909.

Alamro, M., F.A. Bidmos, H. Chan, N.J. Oldfield, E. Newton, X. Bai, J. Aidley, R. Care, C. Mattick, D.P. Turner, K.R. Neal, D.A. Ala'aldeen, I. Feavers, R. Borrow & C.D. Bayliss, (2014) Phase variation mediates reductions in expression of surface proteins during persistent meningococcal carriage. *Infection and immunity* **82**: 2472-2484.

Alm, R.A., L.S. Ling, D.T. Moir, B.L. King, E.D. Brown, P.C. Doig, D.R. Smith, B. Noonan, B.C. Guild, B.L. deJonge, G. Carmel, P.J. Tummino, A. Caruso, M. Uria-Nickelsen, D.M. Mills, C. Ives, R. Gibson, D. Merberg, S.D. Mills, Q. Jiang, D.E. Taylor, G.F. Vovis & T.J. Trust, (1999) Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. *Nature* **397**: 176-180.

Altman, E., N. Smirnova, J. Li, A. Aubry & S.M. Logan, (2003) Occurrence of a nontypable Helicobacter pylori strain lacking Lewis blood group O antigens and DD-heptoglycan: evidence for the role of the core alpha1,6-glucan chain in colonization. *Glycobiology* **13**: 777-783.

Anders, S. & W. Huber, (2010) Differential expression analysis for sequence count data. *Genome biology* **11**: R106.

Appelmelk, B.J., S.L. Martin, M.A. Monteiro, C.A. Clayton, A.A. McColm, P. Zheng, T. Verboom, J.J. Maaskant, D.H. van den Eijnden, C.H. Hokke, M.B. Perry, C.M. Vandenbroucke-Grauls & J.G. Kusters, (1999) Phase variation in Helicobacter pylori lipopolysaccharide due to changes in the lengths of poly(C) tracts in alpha3-fucosyltransferase genes. *Infection and immunity* **67**: 5361-5366.

Appelmelk, B.J., M.C. Martino, E. Veenhof, M.A. Monteiro, J.J. Maaskant, R. Negrini, F. Lindh, M. Perry, G. Del Giudice & C.M. Vandenbroucke-Grauls, (2000) Phase variation in H type I and Lewis a epitopes of Helicobacter pylori lipopolysaccharide. *Infection and immunity* **68**: 5928-5932.

Appelmelk, B.J. & C. Vandenbroucke-Grauls, (2001) Lipopolysaccharide Lewis Antigens. In: Helicobacter pylori: Physiology and Genetics. H.L.T. Mobley, G.L. Mendz & S.L. Hazell (eds). Washington (DC), pp.

Appelmelk, B.J. & C.M. Vandenbroucke-Grauls, (2000) H pylori and Lewis antigens. *Gut* **47**: 10-11.

Argaman, L. & S. Altuvia, (2000) fhlA repression by OxyS RNA: kissing complex formation at two sites results in a stable antisense-target RNA complex. *Journal of molecular biology* **300**: 1101-1112.

Attia, A.S. & E.J. Hansen, (2006) A conserved tetranucleotide repeat is necessary for wild-type expression of the Moraxella catarrhalis UspA2 protein. *Journal of bacteriology* **188**: 7840-7852.

Austin, C.M. & R.J. Maier, (2013) Aconitase-mediated posttranscriptional regulation of Helicobacter pylori peptidoglycan deacetylase. *Journal of bacteriology* **195**: 5316-5322.

Avasthi, T.S., S.H. Devi, T.D. Taylor, N. Kumar, R. Baddam, S. Kondo, Y. Suzuki, H. Lamouliatte, F. Megraud & N. Ahmed, (2011) Genomes of two chronological isolates (Helicobacter pylori 2017 and 2018) of the West African Helicobacter pylori strain 908 obtained from a single patient. *Journal of bacteriology* **193**: 3385-3386.

Babitzke, P. & T. Romeo, (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Current opinion in microbiology* **10**: 156-163.

Backert, S., N. Tegtmeyer & M. Selbach, (2010) The versatility of Helicobacter pylori CagA effector protein functions: The master key hypothesis. *Helicobacter* **15**: 163-176.

Backofen, R. & W.R. Hess, (2010) Computational prediction of sRNAs and their targets in bacteria. *RNA biology* **7**: 33-42.

Bailey, T.L., M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li & W.S. Noble, (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic acids research* **37**: W202-208.

Balasubramanian, D. & C.K. Vanderpool, (2013) New developments in post-transcriptional regulation of operons by small RNAs. *RNA biology* **10**: 337-341.

Baldari, C.T., A. Lanzavecchia & J.L. Telford, (2005) Immune subversion by Helicobacter pylori. *Trends in immunology* **26**: 199-207.

Baldwin, D.N., B. Shepherd, P. Kraemer, M.K. Hall, L.K. Sycuro, D.M. Pinto-Santini & N.R. Salama, (2007) Identification of Helicobacter pylori genes that contribute to stomach colonization. *Infection and immunity* **75**: 1005-1016.

Baltrus, D.A., M.R. Amieva, A. Covacci, T.M. Lowe, D.S. Merrell, K.M. Ottemann, M. Stein, N.R. Salama & K. Guillemin, (2009) The complete genome sequence of Helicobacter pylori strain G27. *J Bacteriol* **191**: 447-448.

Bandyra, K.J., M. Bouvier, A.J. Carpousis & B.F. Luisi, (2013) The social fabric of the RNA degradosome. *Biochimica et biophysica acta* **1829**: 514-522.

Baranov, P.V., A.W. Hammer, J. Zhou, R.F. Gesteland & J.F. Atkins, (2005) Transcriptional slippage in bacteria: distribution in sequenced genomes and utilization in IS element gene expression. *Genome biology* **6**: R25.

Barnard, F.M., M.F. Loughlin, H.P. Fainberg, M.P. Messenger, D.W. Ussery, P. Williams & P.J. Jenks, (2004) Global regulation of virulence and the stress response by CsrA in the highly adapted human gastric pathogen Helicobacter pylori. *Molecular microbiology* **51**: 15-32.

Barquist, L., G.C. Langridge, D.J. Turner, M.D. Phan, A.K. Turner, A. Bateman, J. Parkhill, J. Wain & P.P. Gardner, (2013) A comparison of dense transposon insertion libraries in the Salmonella serovars Typhi and Typhimurium. *Nucleic acids research* **41**: 4549-4564.

Barrangou, R. & L.A. Marraffini, (2014) CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Molecular cell* **54**: 234-244.

Barrick, J.E., N. Sudarsan, Z. Weinberg, W.L. Ruzzo & R.R. Breaker, (2005) 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *Rna* **11**: 774-784.

Bartel, D.P., (2009) MicroRNAs: target recognition and regulatory functions. *Cell* **136**: 215-233.

Bauer, B., E. Pang, C. Holland, M. Kessler, S. Bartfeld & T.F. Meyer, (2012) The Helicobacter pylori virulence effector CagA abrogates human beta-defensin 3 expression via inactivation of EGFR signaling. *Cell host & microbe* **11**: 576-586.

Bauer, S., M. Endres, M. Lange, T. Schmidt, C. Schumbrutzki, A. Sickmann & D. Beier, (2013) Novel function assignment to a member of the essential HP1043 response regulator family of epsilon-proteobacteria. *Microbiology* **159**: 880-889.

Bayliss, C.D., (2009) Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. *FEMS microbiology reviews* **33**: 504-520.

Bayliss, C.D., F.A. Bidmos, A. Anjum, V.T. Manchev, R.L. Richards, J.P. Grossier, K.G. Wooldridge, J.M. Ketley, P.A. Barrow, M.A. Jones & M.V. Tretyakov, (2012) Phase variable genes of Campylobacter jejuni exhibit high mutation rates and specific mutational patterns but mutability is not the major determinant of population structure during host colonization. *Nucleic acids research* **40**: 5876-5889.

Behrens, W., T. Schweinitzer, J. Bal, M. Dorsch, A. Bleich, F. Kops, B. Brenneke, X. Didelot, S. Suerbaum & C. Josenhans, (2013) Role of energy sensor TlpD of Helicobacter pylori in gerbil colonization and genome analyses after adaptation in the gerbil. *Infection and immunity* **81**: 3534-3551.

Beier, D. & R. Frank, (2000) Molecular characterization of two-component systems of Helicobacter pylori. *Journal of bacteriology* **182**: 2068-2076.

Beisel, C.L. & G. Storz, (2010) Base pairing small RNAs and their roles in global regulatory networks. *FEMS microbiology reviews* **34**: 866-882.

Bergman, M., G. Del Prete, Y. van Kooyk & B. Appelmelk, (2006) Helicobacter pylori phase variation, immune modulation and gastric autoimmunity. *Nature reviews. Microbiology* **4**: 151-159.

Blattner, F.R., G. Plunkett, 3rd, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, N.W. Davis, H.A. Kirkpatrick, M.A. Goeden, D.J. Rose, B. Mau & Y. Shao, (1997) The complete genome sequence of Escherichia coli K-12. *Science* **277**: 1453-1462.

Blomberg, P., E.G. Wagner & K. Nordstrom, (1990) Control of replication of plasmid R1: the duplex between the antisense RNA, CopA, and its target, CopT, is processed specifically in vivo and in vitro by RNase III. *The EMBO journal* **9**: 2331-2340.

Bochman, M.L., K. Paeschke & V.A. Zakian, (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nature reviews. Genetics* **13**: 770-780.

Boisset, S., T. Geissmann, E. Huntzinger, P. Fechter, N. Bendridi, M. Possedko, C. Chevalier, A.C. Helfer, Y. Benito, A. Jacquier, C. Gaspin, F. Vandenesch & P. Romby, (2007) Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & development* **21**: 1353-1366.

Boneca, I.G., C. Ecobichon, C. Chaput, A. Mathieu, S. Guadagnini, M.C. Prevost, F. Colland, A. Labigne & H. de Reuse, (2008) Development of inducible systems to engineer conditional mutants of essential genes of Helicobacter pylori. *Appl Environ Microbiol* **74**: 2095-2102.

Bossi, L., A. Schwartz, B. Guillemardet, M. Boudvillain & N. Figueroa-Bossi, (2012) A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes & development* **26**: 1864-1873.

Brantl, S., (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current opinion in microbiology* **10**: 102-109.

Brantl, S., (2012) Acting antisense: plasmid- and chromosome-encoded sRNAs from Gram-positive bacteria. *Future microbiology* **7**: 853-871.

Bucker, R., M. Azevedo-Vethacke, C. Groll, D. Garten, C. Josenhans, S. Suerbaum & S. Schreiber, (2012) Helicobacter pylori colonization critically depends on postprandial gastric conditions. *Scientific reports* **2**: 994.

Bugaut, A. & S. Balasubramanian, (2012) 5'-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic acids research* **40**: 4727-4741.

Bumann, D., S. Aksu, M. Wendland, K. Janek, U. Zimny-Arndt, N. Sabarth, T.F. Meyer & P.R. Jungblut, (2002) Proteome analysis of secreted proteins of the gastric pathogen Helicobacter pylori. *Infection and immunity* **70**: 3396-3403.

Bury-Mone, S., N.O. Kaakoush, C. Asencio, F. Megraud, M. Thibonnier, H. De Reuse & G.L. Mendz, (2006) Is Helicobacter pylori a true microaerophile? *Helicobacter* **11**: 296-303.

Bury-Mone, S., G.L. Mendz, G.E. Ball, M. Thibonnier, K. Stingl, C. Ecobichon, P. Ave, M. Huerre, A. Labigne, J.M. Thiberge & H. De Reuse, (2008) Roles of alpha and beta carbonic anhydrases of Helicobacter pylori in the urease-dependent response to acidity and in colonization of the murine gastric mucosa. *Infection and immunity* **76**: 497-509.

Bury-Mone, S., S. Skouloubris, C. Dauga, J.M. Thiberge, D. Dailidiene, D.E. Berg, A. Labigne & H. De Reuse, (2003) Presence of active aliphatic amidases in Helicobacter species able to colonize the stomach. *Infection and immunity* **71**: 5613-5622.

Bury-Mone, S., S. Skouloubris, A. Labigne & H. De Reuse, (2001) The Helicobacter pylori UreI protein: role in adaptation to acidity and identification of residues essential for its activity and for acid activation. *Mol Microbiol* **42**: 1021-1034.

Bury-Mone, S., J.M. Thiberge, M. Contreras, A. Maitournam, A. Labigne & H. De Reuse, (2004) Responsiveness to acidity via metal ion regulators mediates virulence in the gastric pathogen Helicobacter pylori. *Molecular microbiology* **53**: 623-638.

Cahoon, L.A. & H.S. Seifert, (2009) An alternative DNA structure is necessary for pilin antigenic variation in Neisseria gonorrhoeae. *Science* **325**: 764-767.

Cahoon, L.A. & H.S. Seifert, (2013) Transcription of a cis-acting, noncoding, small RNA is required for pilin antigenic variation in Neisseria gonorrhoeae. *PLoS pathogens* **9**: e1003074.

Caldelari, I., Y. Chao, P. Romby & J. Vogel, (2013) RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harbor perspectives in medicine* **3**: a010298.

Calderon, I.L., E.H. Morales, B. Collao, P.F. Calderon, C.A. Chahuan, L.G. Acuna, F. Gil & C.P. Saavedra, (2014) Role of Salmonella Typhimurium small RNAs RyhB-1 and RyhB-2 in the oxidative stress response. *Research in microbiology* **165**: 30-40.

Carpenter, B.M., J.J. Gilbreath, O.Q. Pich, A.M. McKelvey, E.L. Maynard, Z.Z. Li & D.S. Merrell, (2013) Identification and characterization of novel Helicobacter pylori apo-fur-regulated target genes. *Journal of bacteriology* **195**: 5526-5539.

Carpenter, B.M., T.K. McDaniel, J.M. Whitmire, H. Gancz, S. Guidotti, S. Censini & D.S. Merrell, (2007) Expanding the Helicobacter pylori genetic toolbox: modification of an endogenous plasmid for use as a transcriptional reporter and complementation vector. *Applied and environmental microbiology* **73**: 7506-7514.

Carpousis, A.J., B.F. Luisi & K.J. McDowall, (2009) Endonucleolytic initiation of mRNA decay in Escherichia coli. *Progress in molecular biology and translational science* **85**: 91-135.

Cavanagh, A.T. & K.M. Wassarman, (2014) 6S RNA, a global regulator of transcription in Escherichia coli, Bacillus subtilis, and beyond. *Annual review of microbiology* **68**: 45-60.

Cellini, L., (2014) Helicobacter pylori: a chameleon-like approach to life. *World journal of gastroenterology : WJG* **20**: 5575-5582.

Chabelskaya, S., V. Bordeau & B. Felden, (2014) Dual RNA regulatory control of a Staphylococcus aureus virulence factor. *Nucleic acids research* **42**: 4847-4858.

Chao, Y., K. Papenfort, R. Reinhardt, C.M. Sharma & J. Vogel, (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *The EMBO journal* **31**: 4005-4019.

Chao, Y. & J. Vogel, (2010) The role of Hfq in bacterial pathogens. *Current opinion in microbiology* **13**: 24-33.

Chatr-aryamontri, A., A. Ceol, L.M. Palazzi, G. Nardelli, M.V. Schneider, L. Castagnoli & G. Cesareni, (2007) MINT: the Molecular INTeraction database. *Nucleic acids research* **35**: D572-574.

Chaudhuri, R.R., E. Morgan, S.E. Peters, S.J. Pleasance, D.L. Hudson, H.M. Davies, J. Wang, P.M. van Diemen, A.M. Buckley, A.J. Bowen, G.D. Pullinger, D.J. Turner, G.C. Langridge, A.K. Turner, J. Parkhill, I.G. Charles, D.J. Maskell & M.P. Stevens, (2013) Comprehensive assignment of roles for Salmonella typhimurium genes in intestinal colonization of food-producing animals. *PLoS genetics* **9**: e1003456.

Chen, H.D. & E.A. Groisman, (2013) The biology of the PmrA/PmrB two-component system: the major regulator of lipopolysaccharide modifications. *Annual review of microbiology* **67**: 83-112.

Chmiela, M., E. Miszczyk & K. Rudnicka, (2014) Structural modifications of Helicobacter pylori lipopolysaccharide: an idea for how to live in peace. *World journal of gastroenterology : WJG* **20**: 9882-9897.

Comtois, S.L., M.D. Gidley & D.J. Kelly, (2003) Role of the thioredoxin system and the thiol-peroxidases Tpx and Bcp in mediating resistance to oxidative and nitrosative stress in Helicobacter pylori. *Microbiology* **149**: 121-129.

Condon, C., (2010) What is the role of RNase J in mRNA turnover? *RNA biology* **7**: 316-321.

Coornaert, A., A. Lu, P. Mandin, M. Springer, S. Gottesman & M. Guillier, (2010) MicA sRNA links the PhoP regulon to cell envelope stress. *Molecular microbiology* **76**: 467-479.

Corcoran, C.P., D. Podkaminski, K. Papenfort, J.H. Urban, J.C. Hinton & J. Vogel, (2012a) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Molecular microbiology* **84**: 428-445.

Corcoran, C.P., R. Rieder, D. Podkaminski, B. Hofmann & J. Vogel, (2012b) Use of aptamer tagging to identify in vivo protein binding partners of small regulatory RNAs. *Methods in molecular biology* **905**: 177-200.

Cover, T.L. & M.J. Blaser, (2009) Helicobacter pylori in health and disease. *Gastroenterology* **136**: 1863-1873.

Creecy, J.P. & T. Conway, (2015) Quantitative bacterial transcriptomics with RNA-seq. *Current opinion in microbiology* **23C**: 133-140.

Croucher, N.J. & N.R. Thomson, (2010) Studying bacterial transcriptomes using RNA-seq. *Current opinion in microbiology* **13**: 619-624.

Croxen, M.A., P.B. Ernst & P.S. Hoffman, (2007) Antisense RNA modulation of alkyl hydroperoxide reductase levels in Helicobacter pylori correlates with organic peroxide toxicity but not infectivity. *Journal of bacteriology* **189**: 3359-3368.

Croxen, M.A., G. Sisson, R. Melano & P.S. Hoffman, (2006) The Helicobacter pylori chemotaxis receptor TlpB (HP0103) is required for pH taxis and for colonization of the gastric mucosa. *Journal of bacteriology* **188**: 2656-2665.

Cullen, T.W., D.K. Giles, L.N. Wolf, C. Ecobichon, I.G. Boneca & M.S. Trent, (2011) Helicobacter pylori versus the host: remodeling of the bacterial outer membrane is required for survival in the gastric mucosa. *PLoS pathogens* **7**: e1002454.

Dailidiene, D., G. Dailide, D. Kersulyte & D.E. Berg, (2006) Contraselectable streptomycin susceptibility determinant for genetic manipulation and analysis of Helicobacter pylori. *Applied and environmental microbiology* **72**: 5908-5914.

Dairi, T., (2009) An alternative menaquinone biosynthetic pathway operating in microorganisms: an attractive target for drug discovery to pathogenic Helicobacter and Chlamydia strains. *The Journal of antibiotics* **62**: 347-352.

Danielli, A., D. Roncarati, I. Delany, V. Chiarini, R. Rappuoli & V. Scarlato, (2006) In vivo dissection of the Helicobacter pylori Fur regulatory circuit by genome-wide location analysis. *Journal of bacteriology* **188**: 4654-4662.

Danielli, A. & V. Scarlato, (2010) Regulatory circuits in Helicobacter pylori : network motifs and regulators involved in metal-dependent responses. *FEMS microbiology reviews* **34**: 738-752.

Danne, C., S. Dubrac, P. Trieu-Cuot & S. Dramsi, (2014) Single cell stochastic regulation of pilus phase variation by an attenuation-like mechanism. *PLoS pathogens* **10**: e1003860.

Darfeuille, F., C. Unoson, J. Vogel & E.G. Wagner, (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Molecular cell* **26**: 381-392.

de Vries, N., D. Duinsbergen, E.J. Kuipers, R.G. Pot, P. Wiesenekker, C.W. Penn, A.H. van Vliet, C.M. Vandenbroucke-Grauls & J.G. Kusters, (2002) Transcriptional phase variation of a type III restriction-modification system in Helicobacter pylori. *Journal of bacteriology* **184**: 6615-6623.

de Vries, N., A.H.M. van Vliet & J.G. Kusters, (2001) Gene Regulation.

Delany, I., G. Spohn, R. Rappuoli & V. Scarlato, (2002) Growth phase-dependent regulation of target gene promoters for binding of the essential orphan response regulator HP1043 of Helicobacter pylori. *Journal of bacteriology* **184**: 4800-4810.

Desnoyers, G., M.P. Bouchard & E. Masse, (2013) New insights into small RNA-dependent translational regulation in prokaryotes. *Trends in genetics : TIG* **29**: 92-98.

Desnoyers, G. & E. Masse, (2012) Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes & development* **26**: 726-739.

Desnoyers, G., A. Morissette, K. Prevost & E. Masse, (2009) Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA. *The EMBO journal* **28**: 1551-1561.

Devi, S.H., T.D. Taylor, T.S. Avasthi, S. Kondo, Y. Suzuki, F. Megraud & N. Ahmed, (2010) Genome of Helicobacter pylori strain 908. *Journal of bacteriology* **192**: 6488-6489.

Ding, Y., Y. Tang, C.K. Kwok, Y. Zhang, P.C. Bevilacqua & S.M. Assmann, (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* **505**: 696-700.

Doma, M.K. & R. Parker, (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**: 561-564.

Donczew, R., L. Makowski, P. Jaworski, M. Bezulska, M. Nowaczyk, J. Zakrzewska-Czerwinska & A. Zawilak-Pawlik, (2015) The atypical response regulator HP1021 controls formation of the Helicobacter pylori replication initiation complex. *Molecular microbiology* **95**: 297-312.

Dorleans, A., I. Li de la Sierra-Gallay, J. Piton, L. Zig, L. Gilet, H. Putzer & C. Condon, (2011) Molecular basis for the recognition and cleavage of RNA by the bifunctional 5'-3' exo/endoribonuclease RNase J. *Structure* **19**: 1252-1261.

Douillard, F.P., K.A. Ryan, D.L. Caly, J. Hinds, A.A. Witney, S.E. Husain & P.W. O'Toole, (2008) Posttranscriptional regulation of flagellin synthesis in Helicobacter pylori by the RpoN chaperone HP0958. *Journal of bacteriology* **190**: 7975-7984.

Dugar, G., A. Herbig, K.U. Forstner, N. Heidrich, R. Reinhardt, K. Nieselt & C.M. Sharma, (2013) High-resolution transcriptome maps reveal strain-specific regulatory features of multiple Campylobacter jejuni isolates. *PLoS genetics* **9**: e1003495.

Edwards, N.J., M.A. Monteiro, G. Faller, E.J. Walsh, A.P. Moran, I.S. Roberts & N.J. High, (2000) Lewis X structures in the O antigen side-chain promote adhesion of Helicobacter pylori to the gastric epithelium. *Molecular microbiology* **35**: 1530-1539.

Egbert, R.G. & E. Klavins, (2012) Fine-tuning gene networks using simple sequence repeats. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 16817-16822.

Emerson, J.E., C.B. Reynolds, R.P. Fagan, H.A. Shaw, D. Goulding & N.F. Fairweather, (2009) A novel genetic switch controls phase variable expression of CwpV, a Clostridium difficile cell wall protein. *Molecular microbiology* **74**: 541-556.

Endoh, T., Y. Kawasaki & N. Sugimoto, (2013) Translational halt during elongation caused by G-quadruplex formed by mRNA. *Methods* **64**: 73-78.

Eppinger, M., C. Baar, B. Linz, G. Raddatz, C. Lanz, H. Keller, G. Morelli, H. Gressmann, M. Achtman & S.C. Schuster, (2006) Who ate whom? Adaptive Helicobacter genomic changes that accompanied a host jump from early humans to large felines. *PLoS Genet* **2**: e120.

Ernst, F.D., S. Bereswill, B. Waidner, J. Stoof, U. Mader, J.G. Kusters, E.J. Kuipers, M. Kist, A.H. van Vliet & G. Homuth, (2005) Transcriptional profiling of Helicobacter pylori Fur- and iron-regulated gene expression. *Microbiology* **151**: 533-546.

Farnbacher, M., T. Jahns, D. Willrodt, R. Daniel, R. Haas, A. Goesmann, S. Kurtz & G. Rieder, (2010) Sequencing, annotation, and comparative genome analysis of the gerbil-adapted Helicobacter pylori strain B8. *BMC Genomics* **11**: 335.

Finn, R.D., A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L. Sonnhammer, J. Tate & M. Punta, (2014) Pfam: the protein families database. *Nucleic acids research* **42**: D222-230.

Fischer, W., S. Prassl & R. Haas, (2009) Virulence mechanisms and persistence strategies of the human gastric pathogen Helicobacter pylori. *Current topics in microbiology and immunology* **337**: 129-171.

Fischer, W., L. Windhager, S. Rohrer, M. Zeiller, A. Karnholz, R. Hoffmann, R. Zimmer & R. Haas, (2010) Strain-specific genes of Helicobacter pylori: genome evolution driven by a novel type IV secretion system and genomic island transfer. *Nucleic Acids Res* **38**: 6089-6101.

Forstner, K.U., J. Vogel & C.M. Sharma, (2014) READemption-a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics* **30**: 3421-3423.

Fowler, M., R.J. Thomas, J. Atherton, I.S. Roberts & N.J. High, (2006) Galectin-3 binds to Helicobacter pylori O-antigen: it is upregulated and rapidly secreted by gastric epithelial cells in response to H. pylori adhesion. *Cellular microbiology* **8**: 44-54.

Franceschini, A., D. Szklarczyk, S. Frankild, M. Kuhn, M. Simonovic, A. Roth, J. Lin, P. Minguez, P. Bork, C. von Mering & L.J. Jensen, (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research* **41**: D808-815.

Frohlich, K.S., K. Papenfort, A.A. Berger & J. Vogel, (2012) A conserved RpoS-dependent small RNA controls the synthesis of major porin OmpD. *Nucleic acids research* **40**: 3623-3640.

Frohlich, K.S. & J. Vogel, (2009) Activation of gene expression by small RNA. *Current opinion in microbiology* **12**: 674-682.

Gauntlett, J.C., H.O. Nilsson, A. Fulurija, B.J. Marshall & M. Benghezal, (2014) Phase-variable restriction/modification systems are required for Helicobacter pylori colonization. *Gut pathogens* **6**: 35.

Geissmann, T., C. Chevalier, M.J. Cros, S. Boisset, P. Fechter, C. Noirot, J. Schrenzel, P. Francois, F. Vandenesch, C. Gaspin & P. Romby, (2009) A search for small noncoding RNAs in Staphylococcus aureus reveals a conserved sequence motif for regulation. *Nucleic acids research* **37**: 7239-7257.

Georg, J. & W.R. Hess, (2011) cis-antisense RNA, another level of gene regulation in bacteria. *Microbiology and molecular biology reviews : MMBR* **75**: 286-300.

George, J.T., P.K. Boughan, H. Karageorgiou & M. Bajaj-Elliott, (2003) Host anti-microbial response to Helicobacter pylori infection. *Molecular immunology* **40**: 451-456.

Giedroc, D.P. & P.V. Cornish, (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus research* **139**: 193-208.

Git, A., H. Dvinge, M. Salmon-Divon, M. Osborne, C. Kutter, J. Hadfield, P. Bertone & C. Caldas, (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *Rna* **16**: 991-1006.

Gobert, A.P., D.J. McGee, M. Akhtar, G.L. Mendz, J.C. Newton, Y. Cheng, H.L. Mobley & K.T. Wilson, (2001) Helicobacter pylori arginase inhibits nitric oxide production by eukaryotic cells: a strategy for bacterial survival. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 13844-13849.

Golden, N.J. & D.W. Acheson, (2002) Identification of motility and autoagglutination Campylobacter jejuni mutants by random transposon mutagenesis. *Infection and immunity* **70**: 1761-1771.

Goodwin, A., D. Kersulyte, G. Sisson, S.J. Veldhuyzen van Zanten, D.E. Berg & P.S. Hoffman, (1998) Metronidazole resistance in Helicobacter pylori is due to null mutations in a

gene (rdxA) that encodes an oxygen-insensitive NADPH nitroreductase. *Mol Microbiol* **28**: 383-393.

Gopel, Y., M.A. Khan & B. Gorke, (2014) Menage a trois: post-transcriptional control of the key enzyme for cell envelope synthesis by a base-pairing small RNA, an RNase adaptor protein, and a small RNA mimic. *RNA biology* **11**: 433-442.

Grant, C.E., T.L. Bailey & W.S. Noble, (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017-1018.

Gu, W., T. Zhou & C.O. Wilke, (2010) A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS computational biology* **6**: e1000664.

Gueguen, E., N.M. Wills, J.F. Atkins & E. Cascales, (2014) Transcriptional frameshifting rescues Citrobacter rodentium type VI secretion by the production of two length variants from the prematurely interrupted tssM gene. *PLoS genetics* **10**: e1004869.

Guerry, P., C.P. Ewing, M. Schirm, M. Lorenzo, J. Kelly, D. Pattarini, G. Majam, P. Thibault & S. Logan, (2006) Changes in flagellin glycosylation affect Campylobacter autoagglutination and virulence. *Molecular microbiology* **60**: 299-311.

Guo, M.S., T.B. Updegrove, E.B. Gogol, S.A. Shabalina, C.A. Gross & G. Storz, (2014) MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes & development* **28**: 1620-1634.

Hajnsdorf, E. & I.V. Boni, (2012) Multiple activities of RNA-binding proteins S1 and Hfq. *Biochimie* **94**: 1544-1553.

Haley, K.P. & J.A. Gaddy, (2015) Helicobacter pylori: Genomic Insight into the Host-Pathogen Interaction. *International journal of genomics* **2015**: 386905.

Hammerle, H., F. Amman, B. Vecerek, J. Stulke, I. Hofacker & U. Blasi, (2014) Impact of Hfq on the Bacillus subtilis transcriptome. *PloS one* **9**: e98661.

Handt, L.K., J.G. Fox, I.H. Stalis, R. Rufo, G. Lee, J. Linn, X. Li & H. Kleanthous, (1995) Characterization of feline Helicobacter pylori strains and associated gastritis in a colony of domestic cats. *Journal of clinical microbiology* **33**: 2280-2289.

Heidrich, N., I. Moll & S. Brantl, (2007) In vitro analysis of the interaction between the small RNA SR1 and its primary target ahrC mRNA. *Nucleic acids research* **35**: 4331-4346.

Heinrichs, D.E., J.A. Yethon, P.A. Amor & C. Whitfield, (1998) The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of Escherichia coli. The R1 core-specific beta-glucosyltransferase provides a novel attachment site for O-polysaccharides. *The Journal of biological chemistry* **273**: 29497-29505.

Holst, O., A.J. Ulmer, H. Brade, H.D. Flad & E.T. Rietschel, (1996) Biochemistry and cell biology of bacterial endotoxins. *FEMS immunology and medical microbiology* **16**: 83-104.

Hug, I., M.R. Couturier, M.M. Rooker, D.E. Taylor, M. Stein & M.F. Feldman, (2010) Helicobacter pylori lipopolysaccharide is synthesized via a novel pathway with an evolutionary connection to protein N-glycosylation. *PLoS pathogens* **6**: e1000819.

Ikemoto, M., M. Tabata, T. Miyake, T. Kono, M. Mori, M. Totani & T. Murachi, (1990) Expression of human liver arginase in Escherichia coli. Purification and properties of the product. *The Biochemical journal* **270**: 697-703.

Ingolia, N.T., (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nature reviews. Genetics* **15**: 205-213.

Ingolia, N.T., S. Ghaemmaghami, J.R. Newman & J.S. Weissman, (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**: 218-223.

Janssens, S. & R. Beyaert, (2003) Role of Toll-like receptors in pathogen recognition. *Clinical microbiology reviews* **16**: 637-646.

Jeon, B., K. Itoh, N. Misawa & S. Ryu, (2003) Effects of quorum sensing on flaA transcription and autoagglutination in Campylobacter jejuni. *Microbiology and immunology* **47**: 833-839.

Jerome, J.P., J.A. Bell, A.E. Plovanich-Jones, J.E. Barrick, C.T. Brown & L.S. Mansfield, (2011) Standing genetic variation in contingency loci drives the rapid adaptation of Campylobacter jejuni to a novel host. *PloS one* **6**: e16399.

Jimenez-Pearson, M.A., I. Delany, V. Scarlato & D. Beier, (2005) Phosphate flow in the chemotactic response system of Helicobacter pylori. *Microbiology* **151**: 3299-3311.

Josenhans, C., D. Beier, B. Linz, T.F. Meyer & S. Suerbaum, (2007) Pathogenomics of helicobacter. *International journal of medical microbiology : IJMM* **297**: 589-600.

Josenhans, C., K.A. Eaton, T. Thevenot & S. Suerbaum, (2000) Switching of flagellar motility in Helicobacter pylori by reversible length variation of a short homopolymeric sequence repeat in fliP, a gene encoding a basal body protein. *Infection and immunity* **68**: 4598-4603.

Jungblut, P.R. & R. Seifert, (1990) Analysis by high-resolution two-dimensional electrophoresis of differentiation-dependent alterations in cytosolic protein pattern of HL-60 leukemic cells. *Journal of biochemical and biophysical methods* **21**: 47-58.

Kaatz, G.W., S.M. Seo & C.A. Ruble, (1993) Efflux-mediated fluoroquinolone resistance in Staphylococcus aureus. *Antimicrobial agents and chemotherapy* **37**: 1086-1094.

Kennemann, L., B. Brenneke, S. Andres, L. Engstrand, T.F. Meyer, T. Aebischer, C. Josenhans & S. Suerbaum, (2012) In vivo sequence variation in HopZ, a phase-variable outer membrane protein of Helicobacter pylori. *Infection and immunity* **80**: 4364-4373.

Kennemann, L., X. Didelot, T. Aebischer, S. Kuhn, B. Drescher, M. Droege, R. Reinhardt, P. Correa, T.F. Meyer, C. Josenhans, D. Falush & S. Suerbaum, (2011) Helicobacter pylori genome evolution during human infection. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 5033-5038.

Kersulyte, D., A. Kalia, R.H. Gilman, M. Mendez, P. Herrera, L. Cabrera, B. Velapatino, J. Balqui, F. Paredes Puente de la Vega, C.A. Rodriguez Ulloa, J. Cok, C.C. Hooper, G. Dailide, S. Tamma & D.E. Berg, (2010) Helicobacter pylori from Peruvian amerindians: traces of human migrations in strains from remote Amazon, and genome sequence of an Amerind strain. *PLoS One* **5**: e15076.

Kibbe, W.A., (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic acids research* **35**: W43-46.

Kim, N., D.L. Weeks, J.M. Shin, D.R. Scott, M.K. Young & G. Sachs, (2002) Proteins released by Helicobacter pylori in vitro. *Journal of bacteriology* **184**: 6155-6162.

Knutton, S., R.K. Shaw, R.P. Anantha, M.S. Donnenberg & A.A. Zorgani, (1999) The type IV bundle-forming pilus of enteropathogenic Escherichia coli undergoes dramatic alterations in structure associated with bacterial adherence, aggregation and dispersal. *Molecular microbiology* **33**: 499-509.

Komar, A.A., (2009) A pause for thought along the co-translational folding pathway. *Trends in biochemical sciences* **34**: 16-24.

Konig, J., K. Zarnack, N.M. Luscombe & J. Ule, (2011) Protein-RNA interactions: new genomic technologies and perspectives. *Nature reviews. Genetics* **13**: 77-83.

Kroger, C., A. Colgan, S. Srikumar, K. Handler, S.K. Sivasankaran, D.L. Hammarlof, R. Canals, J.E. Grissom, T. Conway, K. Hokamp & J.C. Hinton, (2013) An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium. *Cell host & microbe* **14**: 683-695.

Kudva, R., K. Denks, P. Kuhn, A. Vogt, M. Muller & H.G. Koch, (2013) Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. *Research in microbiology* **164**: 505-534.

Lafontaine, E.R., N.J. Wagner & E.J. Hansen, (2001) Expression of the Moraxella catarrhalis UspA1 protein undergoes phase variation and is regulated at the transcriptional level. *Journal of bacteriology* **183**: 1540-1551.

Lalaouna, D., M. Simoneau-Roy, D. Lafontaine & E. Masse, (2013) Regulatory RNAs and target mRNA decay in prokaryotes. *Biochimica et biophysica acta* **1829**: 742-747.

Langdon, R., J.E. Craig, M. Goldrick, R. Houldsworth & N.J. High, (2005) Analysis of the role of HP0208, a phase-variable open reading frame, and its homologues HP1416 and

HP0159 in the biosynthesis of Helicobacter pylori lipopolysaccharide. *Journal of medical microbiology* **54**: 697-706.

Leduc, D., J. Gallaud, K. Stingl & H. de Reuse, (2010) Coupled amino acid deamidase-transport systems essential for Helicobacter pylori colonization. *Infection and immunity* **78**: 2782-2792.

Lehnik-Habrink, M., R.J. Lewis, U. Mader & J. Stulke, (2012) RNA degradation in Bacillus subtilis: an interplay of essential endo- and exoribonucleases. *Molecular microbiology* **84**: 1005-1017.

Lertsethtakarn, P., K.M. Ottemann & D.R. Hendrixson, (2011) Motility and chemotaxis in Campylobacter and Helicobacter. *Annual review of microbiology* **65**: 389-410.

Livak, K.J. & T.D. Schmittgen, (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402-408.

Logan, S.M., E. Altman, O. Mykytczuk, J.R. Brisson, V. Chandan, M.J. Schur, F. St Michael, A. Masson, S. Leclerc, K. Hiratsuka, N. Smirnova, J. Li, Y. Wu & W.W. Wakarchuk, (2005) Novel biosynthetic functions of lipopolysaccharide rfaJ homologs from Helicobacter pylori. *Glycobiology* **15**: 721-733.

Logan, S.M., J.W. Conlan, M.A. Monteiro, W.W. Wakarchuk & E. Altman, (2000) Functional genomics of Helicobacter pylori: identification of a beta-1,4 galactosyltransferase and generation of mutants with altered lipopolysaccharide. *Molecular microbiology* **35**: 1156-1167.

Loh, J.T., V.J. Torres & T.L. Cover, (2007) Regulation of Helicobacter pylori cagA expression in response to salt. *Cancer research* **67**: 4709-4715.

Lowenthal, A.C., C. Simon, A.S. Fair, K. Mehmood, K. Terry, S. Anastasia & K.M. Ottemann, (2009) A fixed-time diffusion analysis method determines that the three cheV genes of Helicobacter pylori differentially affect motility. *Microbiology* **155**: 1181-1191.

Lu, J. & A. Holmgren, (2014) The thioredoxin antioxidant system. *Free radical biology & medicine* **66**: 75-87.

Lundin, A., C. Nilsson, M. Gerhard, D.I. Andersson, M. Krabbe & L. Engstrand, (2003) The NudA protein in the gastric pathogen Helicobacter pylori is an ubiquitous and constitutively expressed dinucleoside polyphosphate hydrolase. *The Journal of biological chemistry* **278**: 12574-12578.

Mandin, P. & S. Gottesman, (2009) A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Molecular microbiology* **72**: 551-565.

Mandin, P. & S. Gottesman, (2010) Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *The EMBO journal* **29**: 3094-3107.

Marshall, B.J. & J.R. Warren, (1984) Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* **1**: 1311-1315.

Masse, E., C.K. Vanderpool & S. Gottesman, (2005) Effect of RyhB small RNA on global iron use in Escherichia coli. *Journal of bacteriology* **187**: 6962-6971.

Mathews, D.H., (2006) RNA secondary structure analysis using RNAstructure. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* **Chapter 12**: Unit 12 16.

McClain, M.S., S.S. Duncan, J.A. Gaddy & T.L. Cover, (2013) Control of gene expression in Helicobacter pylori using the Tet repressor. *Journal of microbiological methods* **95**: 336-341.

McClain, M.S., C.L. Shaffer, D.A. Israel, R.M. Peek, Jr. & T.L. Cover, (2009) Genome sequence analysis of Helicobacter pylori strains associated with gastric ulceration and gastric cancer. *BMC genomics* **10**: 3.

McDaniel, T.K., K.C. Dewalt, N.R. Salama & S. Falkow, (2001) New approaches for validation of lethal phenotypes and genetic reversion in Helicobacter pylori. *Helicobacter* **6**: 15-23.

McGee, D.J., M.L. Langford, E.L. Watson, J.E. Carter, Y.T. Chen & K.M. Ottemann, (2005) Colonization and inflammation deficiencies in Mongolian gerbils infected by Helicobacter pylori chemotaxis mutants. *Infection and immunity* **73**: 1820-1827.

McGee, D.J., F.J. Radcliff, G.L. Mendz, R.L. Ferrero & H.L. Mobley, (1999) Helicobacter pylori rocF is required for arginase activity and acid protection in vitro but is not essential for colonization of mice or for urease activity. *Journal of bacteriology* **181**: 7314-7322.

McGowan, C.C., A. Necheva, S.A. Thompson, T.L. Cover & M.J. Blaser, (1998) Acid-induced expression of an LPS-associated gene in Helicobacter pylori. *Molecular microbiology* **30**: 19-31.

Merrell, D.S., M.L. Goodrich, G. Otto, L.S. Tompkins & S. Falkow, (2003a) pH-regulated gene expression of the gastric pathogen Helicobacter pylori. *Infection and immunity* **71**: 3529-3539.

Merrell, D.S., L.J. Thompson, C.C. Kim, H. Mitchell, L.S. Tompkins, A. Lee & S. Falkow, (2003b) Growth phase-dependent response of Helicobacter pylori to iron starvation. *Infection and immunity* **71**: 6510-6525.

Mey, A.R., S.A. Craig & S.M. Payne, (2005) Characterization of Vibrio cholerae RyhB: the RyhB regulon and role of ryhB in biofilm formation. *Infection and immunity* **73**: 5706-5719.

Michaux, C., N. Verneuil, A. Hartke & J.C. Giard, (2014) Physiological roles of small RNA molecules. *Microbiology* **160**: 1007-1019.

Millevoi, S., H. Moine & S. Vagner, (2012) G-quadruplexes in RNA biology. *Wiley interdisciplinary reviews. RNA* **3**: 495-507.

Misawa, N. & M.J. Blaser, (2000) Detection and characterization of autoagglutination activity by Campylobacter jejuni. *Infection and immunity* **68**: 6168-6175.

Mitarai, N., A.M. Andersson, S. Krishna, S. Semsey & K. Sneppen, (2007) Efficient degradation and expression prioritization with small RNAs. *Physical biology* **4**: 164-171.

Mitobe, J., I. Yanagihara, K. Ohnishi, S. Yamamoto, M. Ohnishi, A. Ishihama & H. Watanabe, (2011) RodZ regulates the post-transcriptional processing of the Shigella sonnei type III secretion system. *EMBO reports* **12**: 911-916.

Miyakoshi, M., Y. Chao & J. Vogel, (2015) Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Current opinion in microbiology* **24C**: 132-139.

Moller, T., T. Franch, C. Udesen, K. Gerdes & P. Valentin-Hansen, (2002) Spot 42 RNA mediates discoordinate expression of the E. coli galactose operon. *Genes & development* **16**: 1696-1706.

Monack, D.M., (2013) Helicobacter and salmonella persistent infection strategies. *Cold Spring Harbor perspectives in medicine* **3**: a010348.

Monteiro, M.A., P. Zheng, B. Ho, S. Yokota, K. Amano, Z. Pan, D.E. Berg, K.H. Chan, L.L. MacLean & M.B. Perry, (2000) Expression of histo-blood group antigens by lipopolysaccharides of Helicobacter pylori strains from asian hosts: the propensity to express type 1 blood-group antigens. *Glycobiology* **10**: 701-713.

Moon, K. & S. Gottesman, (2009) A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides. *Molecular microbiology* **74**: 1314-1330.

Moon, K., D.A. Six, H.J. Lee, C.R. Raetz & S. Gottesman, (2013) Complex transcriptional and post-transcriptional regulation of an enzyme for lipopolysaccharide modification. *Molecular microbiology* **89**: 52-64.

Moran, A.P., (1995) Cell surface characteristics of Helicobacter pylori. *FEMS immunology and medical microbiology* **10**: 271-280.

Moran, A.P., (2001) Molecular Structure, Biosynthesis, and Pathogenic Roles of Lipopolysaccharides. In: Helicobacter pylori: Physiology and Genetics. H.L.T. Mobley, G.L. Mendz & S.L. Hazell (eds). Washington (DC), pp.

Moran, A.P., (2007) Lipopolysaccharide in bacterial chronic infection: insights from Helicobacter pylori lipopolysaccharide and lipid A. *International journal of medical microbiology : IJMM* **297**: 307-319.

Moran, A.P., (2008) Relevance of fucosylation and Lewis antigen expression in the bacterial gastroduodenal pathogen Helicobacter pylori. *Carbohydrate research* **343**: 1952-1965.

Moran, A.P., I.M. Helander & T.U. Kosunen, (1992) Compositional analysis of Helicobacter pylori rough-form lipopolysaccharides. *Journal of bacteriology* **174**: 1370-1377.

Moran, A.P., Y.A. Knirel, S.N. Senchenkova, G. Widmalm, S.O. Hynes & P.E. Jansson, (2002) Phenotypic variation in molecular mimicry between Helicobacter pylori lipopolysaccharides and human gastric epithelial cell surface glycoforms. Acid-induced phase variation in Lewis(x) and Lewis(y) expression by H. Pylori lipopolysaccharides. *The Journal of biological chemistry* **277**: 5785-5795.

Moran, A.P., E. Sturegard, H. Sjunnesson, T. Wadstrom & S.O. Hynes, (2000) The relationship between O-chain expression and colonisation ability of Helicobacter pylori in a mouse model. *FEMS immunology and medical microbiology* **29**: 263-270.

Morita, T., Y. Mochizuki & H. Aiba, (2006) Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 4858-4863.

Morrison, S., A. Ward, C.J. Hoyle & P.J. Henderson, (2003) Cloning, expression, purification and properties of a putative multidrug resistance efflux protein from Helicobacter pylori. *International journal of antimicrobial agents* **22**: 242-249.

Moxon, R., C. Bayliss & D. Hood, (2006) Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annual review of genetics* **40**: 307-333.

Muller, S., M. Pflock, J. Schar, S. Kennard & D. Beier, (2007) Regulation of expression of atypical orphan response regulators of Helicobacter pylori. *Microbiological research* **162**: 1-14.

Na, D., S.M. Yoo, H. Chung, H. Park, J.H. Park & S.Y. Lee, (2013) Metabolic engineering of Escherichia coli using synthetic small regulatory RNAs. *Nature biotechnology* **31**: 170-174.

Nardone, G., A. Rocco & P. Malfertheiner, (2004) Review article: helicobacter pylori and molecular events in precancerous gastric lesions. *Alimentary pharmacology & therapeutics* **20**: 261-270.

Nicol, J.W., G.A. Helt, S.G. Blanchard, Jr., A. Raja & A.E. Loraine, (2009) The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**: 2730-2731.

Nilsson, C., A. Skoglund, A.P. Moran, H. Annuk, L. Engstrand & S. Normark, (2006) An enzymatic ruler modulates Lewis antigen glycosylation of Helicobacter pylori LPS during persistent infection. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 2863-2868.

Novick, R.P., H.F. Ross, S.J. Projan, J. Kornblum, B. Kreiswirth & S. Moghazeh, (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *The EMBO journal* **12**: 3967-3975.

Obana, N., Y. Shirahama, K. Abe & K. Nakamura, (2010) Stabilization of Clostridium perfringens collagenase mRNA by VR-RNA-dependent cleavage in 5' leader sequence. *Molecular microbiology* **77**: 1416-1428.

Oberto, J., (2013) SyntTax: a web server linking synteny to prokaryotic taxonomy. *BMC bioinformatics* **14**: 4.

Oglesby-Sherrouse, A.G. & E.R. Murphy, (2013) Iron-responsive bacterial small RNAs: variations on a theme. *Metallomics : integrated biometal science* **5**: 276-286.

Oleastro, M. & A. Menard, (2013) The Role of Helicobacter pylori Outer Membrane Proteins in Adherence and Pathogenesis. *Biology* **2**: 1110-1134.

Oleastro, M., L. Monteiro, P. Lehours, F. Megraud & A. Menard, (2006) Identification of markers for Helicobacter pylori strains isolated from children with peptic ulcer

disease by suppressive subtractive hybridization. *Infection and immunity* **74**: 4064-4074.

Oleastro, M., A. Santos, R. Cordeiro, B. Nunes, F. Megraud & A. Menard, (2010) Clinical relevance and diversity of two homologous genes encoding glycosyltransferases in Helicobacter pylori. *Journal of clinical microbiology* **48**: 2885-2891.

Olekhnovich, I.N., S. Vitko, O. Chertihin, R. Hontecillas, M. Viladomiu, J. Bassaganya-Riera & P.S. Hoffman, (2013) Mutations to essential orphan response regulator HP1043 of Helicobacter pylori result in growth-stage regulatory defects. *Infection and immunity* **81**: 1439-1449.

Olekhnovich, I.N., S. Vitko, M. Valliere & P.S. Hoffman, (2014) Response to metronidazole and oxidative stress is mediated through homeostatic regulator HsrA (HP1043) in Helicobacter pylori. *Journal of bacteriology* **196**: 729-739.

Opdyke, J.A., E.M. Fozo, M.R. Hemm & G. Storz, (2011) RNase III participates in GadY-dependent cleavage of the gadX-gadW mRNA. *Journal of molecular biology* **406**: 29-43.

Opdyke, J.A., J.G. Kang & G. Storz, (2004) GadY, a small-RNA regulator of acid response genes in Escherichia coli. *Journal of bacteriology* **186**: 6698-6705.

Overgaard, M., J. Johansen, J. Moller-Jensen & P. Valentin-Hansen, (2009) Switching off small RNA regulation with trap-mRNA. *Molecular microbiology* **73**: 790-800.

Palframan, S.L., T. Kwok & K. Gabriel, (2012) Vacuolating cytotoxin A (VacA), a key toxin for Helicobacter pylori pathogenesis. *Frontiers in cellular and infection microbiology* **2**: 92.

Pandey, S.P., B.K. Minesinger, J. Kumar & G.C. Walker, (2011) A highly conserved protein of unknown function in Sinorhizobium meliloti affects sRNA regulation similar to Hfq. *Nucleic acids research* **39**: 4691-4708.

Pannekoek, Y. & A. van der Ende, (2012) Identification and functional characterization of sRNAs in Neisseria meningitidis. *Methods in molecular biology* **799**: 73-89.

Papenfort, K., M. Bouvier, F. Mika, C.M. Sharma & J. Vogel, (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 20435-20440.

Papenfort, K., V. Pfeiffer, F. Mika, S. Lucchini, J.C. Hinton & J. Vogel, (2006) SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Molecular microbiology* **62**: 1674-1688.

Papenfort, K., D. Podkaminski, J.C. Hinton & J. Vogel, (2012) The ancestral SgrS RNA discriminates horizontally acquired Salmonella mRNAs through a single G-U wobble pair. *Proc Natl Acad Sci U S A* **109**: E757-764.

Papenfort, K., N. Said, T. Welsink, S. Lucchini, J.C. Hinton & J. Vogel, (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Molecular microbiology* **74**: 139-158.

Papenfort, K., Y. Sun, M. Miyakoshi, C.K. Vanderpool & J. Vogel, (2013) Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* **153**: 426-437.

Papenfort, K. & J. Vogel, (2010) Regulatory RNA in bacterial pathogens. *Cell host & microbe* **8**: 116-127.

Papenfort, K. & J. Vogel, (2014) Small RNA functions in carbon metabolism and virulence of enteric pathogens. *Frontiers in cellular and infection microbiology* **4**: 91.

Park, S.A. & N.G. Lee, (2013) Global regulation of gene expression in the human gastric pathogen Helicobacter pylori in response to aerobic oxygen tension under a high carbon dioxide level. *Journal of microbiology and biotechnology* **23**: 451-458.

Peer, A. & H. Margalit, (2011) Accessibility and evolutionary conservation mark bacterial small-rna target-binding regions. *Journal of bacteriology* **193**: 1690-1701.

Pernitzsch, S.R., S.M. Tirier, D. Beier & C.M. Sharma, (2014) A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis

receptor in Helicobacter pylori. *Proceedings of the National Academy of Sciences of the United States of America* **111**: E501-510.

Pfeiffer, V., K. Papenfort, S. Lucchini, J.C. Hinton & J. Vogel, (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nature structural & molecular biology* **16**: 840-846.

Pfeiffer, V., A. Sittka, R. Tomer, K. Tedin, V. Brinkmann & J. Vogel, (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the Salmonella core genome. *Molecular microbiology* **66**: 1174-1191.

Pflock, M., M. Bathon, J. Schar, S. Muller, H. Mollenkopf, T.F. Meyer & D. Beier, (2007) The orphan response regulator HP1021 of Helicobacter pylori regulates transcription of a gene cluster presumably involved in acetone metabolism. *Journal of bacteriology* **189**: 2339-2349.

Pflock, M., S. Kennard, N. Finsterer & D. Beier, (2006a) Acid-responsive gene regulation in the human pathogen Helicobacter pylori. *Journal of biotechnology* **126**: 52-60.

Pflock, M., N. Finsterer, B. Joseph, H. Mollenkopf, T.F. Meyer & D. Beier, (2006b) Characterization of the ArsRS regulon of Helicobacter pylori, involved in acid adaptation. *Journal of bacteriology* **188**: 3449-3462.

Pich, O.Q., B.M. Carpenter, J.J. Gilbreath & D.S. Merrell, (2012) Detailed analysis of Helicobacter pylori Fur-regulated promoters reveals a Fur box core sequence and novel Fur-regulated genes. *Molecular microbiology* **84**: 921-941.

Pich, O.Q. & D.S. Merrell, (2013) The ferric uptake regulator of Helicobacter pylori: a critical player in the battle for iron and colonization of the stomach. *Future microbiology* **8**: 725-738.

Pohl, M.A., J. Romero-Gallo, J.L. Guruge, D.B. Tse, J.I. Gordon & M.J. Blaser, (2009) Host-dependent Lewis (Le) antigen expression in Helicobacter pylori cells recovered from Leb-transgenic mice. *The Journal of experimental medicine* **206**: 3061-3072.

Prevost, K., G. Desnoyers, J.F. Jacques, F. Lavoie & E. Masse, (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes & development* **25**: 385-396.

Prieto, C. & J. De Las Rivas, (2006) APID: Agile Protein Interaction DataAnalyzer. *Nucleic acids research* **34**: W298-302.

Rabhi, M., O. Espeli, A. Schwartz, B. Cayrol, A.R. Rahmouni, V. Arluison & M. Boudvillain, (2011) The Sm-like RNA chaperone Hfq mediates transcription antitermination at Rho-dependent terminators. *The EMBO journal* **30**: 2805-2816.

Rader, B.A., C. Wreden, K.G. Hicks, E.G. Sweeney, K.M. Ottemann & K. Guillemin, (2011) Helicobacter pylori perceives the quorum-sensing molecule AI-2 as a chemorepellent via the chemoreceptor TlpB. *Microbiology* **157**: 2445-2455.

Raetz, C.R. & C. Whitfield, (2002) Lipopolysaccharide endotoxins. *Annual review of biochemistry* **71**: 635-700.

Raghavan, R., E.A. Groisman & H. Ochman, (2011) Genome-wide detection of novel regulatory RNAs in E. coli. *Genome research* **21**: 1487-1497.

Ramirez-Pena, E., J. Trevino, Z. Liu, N. Perez & P. Sumby, (2010) The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. *Molecular microbiology* **78**: 1332-1347.

Redko, Y., S. Aubert, A. Stachowicz, P. Lenormand, A. Namane, F. Darfeuille, M. Thibonnier & H. De Reuse, (2013) A minimal bacterial RNase J-based degradosome is associated with translating ribosomes. *Nucleic acids research* **41**: 288-301.

Reeves, E.P., T. Ali, P. Leonard, S. Hearty, R. O'Kennedy, F.E. May, B.R. Westley, C. Josenhans, M. Rust, S. Suerbaum, A. Smith, B. Drumm & M. Clyne, (2008) Helicobacter pylori lipopolysaccharide interacts with TFF1 in a pH-dependent manner. *Gastroenterology* **135**: 2043-2054, 2054 e2041-2042.

Regulski, E.E. & R.R. Breaker, (2008) In-line probing analysis of riboswitches. *Methods in molecular biology* **419**: 53-67.

Rehmsmeier, M., P. Steffen, M. Hochsmann & R. Giegerich, (2004) Fast and effective prediction of microRNA/target duplexes. *Rna* **10**: 1507-1517.

Remmele, C.W., Y. Xian, M. Albrecht, M. Faulstich, M. Fraunholz, E. Heinrichs, M.T. Dittrich, T. Muller, R. Reinhardt & T. Rudel, (2014) Transcriptional landscape and essential genes of Neisseria gonorrhoeae. *Nucleic acids research* **42**: 10579-10595.

Repar, J., N. Briski, M. Buljubasic, K. Zahradka & D. Zahradka, (2013) Exonuclease VII is involved in "reckless" DNA degradation in UV-irradiated Escherichia coli. *Mutation research* **750**: 96-104.

Rice, J.B., D. Balasubramanian & C.K. Vanderpool, (2012) Small RNA binding-site multiplicity involved in translational regulation of a polycistronic mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **109**: E2691-2698.

Rice, J.B. & C.K. Vanderpool, (2011) The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. *Nucleic acids research* **39**: 3806-3819.

Rieder, R., R. Reinhardt, C. Sharma & J. Vogel, (2012) Experimental tools to identify RNA-protein interactions in Helicobacter pylori. *RNA biology* **9**: 520-531.

Roggenkamp, A., S. Schubert, C.A. Jacobi & J. Heesemann, (1995) Dissection of the Yersinia enterocolitica virulence plasmid pYVO8 into an operating unit and virulence gene modules. *FEMS microbiology letters* **134**: 69-73.

Rouleau, S.G., J.D. Beaudoin, M. Bisaillon & J.P. Perreault, (2015) Small antisense oligonucleotides against G-quadruplexes: specific mRNA translational switches. *Nucleic acids research* **43**: 595-606.

Sachs, G., D.L. Weeks, K. Melchers & D.R. Scott, (2003) The gastric biology of Helicobacter pylori. *Annual review of physiology* **65**: 349-369.

Said, N., R. Rieder, R. Hurwitz, J. Deckert, H. Urlaub & J. Vogel, (2009) In vivo expression and purification of aptamer-tagged small RNA regulators. *Nucleic acids research* **37**: e133.

Salama, N.R., M.L. Hartung & A. Muller, (2013) Life in the human stomach: persistence strategies of the bacterial pathogen Helicobacter pylori. *Nature reviews. Microbiology* **11**: 385-399.

Salaun, L., S. Ayraud & N.J. Saunders, (2005) Phase variation mediated niche adaptation during prolonged experimental murine infection with Helicobacter pylori. *Microbiology* **151**: 917-923.

Salaun, L., B. Linz, S. Suerbaum & N.J. Saunders, (2004) The diversity within an expanded and redefined repertoire of phase-variable genes in Helicobacter pylori. *Microbiology* **150**: 817-830.

Saliba, A.E., A.J. Westermann, S.A. Gorski & J. Vogel, (2014) Single-cell RNA-seq: advances and future challenges. *Nucleic acids research* **42**: 8845-8860.

Salvail, H. & E. Masse, (2012) Regulating iron storage and metabolism with RNA: an overview of posttranscriptional controls of intracellular iron homeostasis. *Wiley interdisciplinary reviews. RNA* **3**: 26-36.

Saramago, M., C. Barria, R.F. Dos Santos, I.J. Silva, V. Pobre, S. Domingues, J.M. Andrade, S.C. Viegas & C.M. Arraiano, (2014) The role of RNases in the regulation of small RNAs. *Current opinion in microbiology* **18**: 105-115.

Saunders, N.J., J.F. Peden, D.W. Hood & E.R. Moxon, (1998) Simple sequence repeats in the Helicobacter pylori genome. *Molecular microbiology* **27**: 1091-1098.

Schar, J., A. Sickmann & D. Beier, (2005) Phosphorylation-independent activity of atypical response regulators of Helicobacter pylori. *Journal of bacteriology* **187**: 3100-3109.

Schmidt, M. & N. Delihas, (1995) micF RNA is a substrate for RNase E. *FEMS microbiology letters* **133**: 209-213.

Schmidtke, C., U. Abendroth, J. Brock, J. Serrania, A. Becker & U. Bonas, (2013) Small RNA sX13: a multifaceted regulator of virulence in the plant pathogen Xanthomonas. *PLoS pathogens* **9**: e1003626.

Schreiber, S., M. Konradt, C. Groll, P. Scheid, G. Hanauer, H.O. Werling, C. Josenhans & S. Suerbaum, (2004) The spatial orientation of Helicobacter pylori in the gastric mucus. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 5024-5029.

Sesto, N., O. Wurtzel, C. Archambaud, R. Sorek & P. Cossart, (2013) The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nature reviews. Microbiology* **11**: 75-82.

Sharma, C.M., F. Darfeuille, T.H. Plantinga & J. Vogel, (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes & development* **21**: 2804-2817.

Sharma, C.M., S. Hoffmann, F. Darfeuille, J. Reignier, S. Findeiss, A. Sittka, S. Chabas, K. Reiche, J. Hackermuller, R. Reinhardt, P.F. Stadler & J. Vogel, (2010) The primary transcriptome of the major human pathogen Helicobacter pylori. *Nature* **464**: 250-255.

Sharma, C.M., K. Papenfort, S.R. Pernitzsch, H.J. Mollenkopf, J.C. Hinton & J. Vogel, (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Molecular microbiology* **81**: 1144-1165.

Sharma, C.M. & J. Vogel, (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. *Current opinion in microbiology* **12**: 536-546.

Sievers, S., E.M. Sternkopf Lillebaek, K. Jacobsen, A. Lund, M.S. Mollerup, P.K. Nielsen & B.H. Kallipolitis, (2014) A multicopy sRNA of Listeria monocytogenes regulates expression of the virulence adhesin LapB. *Nucleic acids research* **42**: 9383-9398.

Sittka, A., S. Lucchini, K. Papenfort, C.M. Sharma, K. Rolle, T.T. Binnewies, J.C. Hinton & J. Vogel, (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS genetics* **4**: e1000163.

Sittka, A., V. Pfeiffer, K. Tedin & J. Vogel, (2007) The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol Microbiol* **63**: 193-217.

Skoglund, A., H.K. Backhed, C. Nilsson, B. Bjorkholm, S. Normark & L. Engstrand, (2009) A changing gastric environment leads to adaptation of lipopolysaccharide variants in Helicobacter pylori populations during colonization. *PloS one* **4**: e5885.

Skouloubris, S., J.M. Thiberge, A. Labigne & H. De Reuse, (1998) The Helicobacter pylori UreI protein is not involved in urease activity but is essential for bacterial survival in vivo. *Infect Immun* **66**: 4517-4521.

Skurnik, M., A. Peippo & E. Ervela, (2000) Characterization of the O-antigen gene clusters of Yersinia pseudotuberculosis and the cryptic O-antigen gene cluster of Yersinia pestis shows that the plague bacillus is most closely related to and has evolved from Y. pseudotuberculosis serotype O:1b. *Molecular microbiology* **37**: 316-330.

Smolka, A.J. & S. Backert, (2012) How Helicobacter pylori infection controls gastric acid secretion. *Journal of gastroenterology* **47**: 609-618.

Solnick, J.V., L.M. Hansen, N.R. Salama, J.K. Boonjakuakul & M. Syvanen, (2004) Modification of Helicobacter pylori outer membrane protein expression during experimental infection of rhesus macaques. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 2106-2111.

Song, Y. & N.J. Sargentini, (1996) Escherichia coli DNA repair genes radA and sms are the same gene. *Journal of bacteriology* **178**: 5045-5048.

Soper, T., P. Mandin, N. Majdalani, S. Gottesman & S.A. Woodson, (2010) Positive regulation by small RNAs and the role of Hfq. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 9602-9607.

Sorek, R. & P. Cossart, (2010) Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nature reviews. Genetics* **11**: 9-16.

Spohn, G. & V. Scarlato, (2001) Motility, Chemotaxis, and Flagella. In: Helicobacter pylori: Physiology and Genetics. H.L.T. Mobley, G.L. Mendz & S.L. Hazell (eds). Washington (DC), pp.

Srikhanta, Y.N., K.L. Fox & M.P. Jennings, (2010) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nature reviews. Microbiology* **8**: 196-206.

Stead, C.M., J. Zhao, C.R. Raetz & M.S. Trent, (2010) Removal of the outer Kdo from Helicobacter pylori lipopolysaccharide and its impact on the bacterial surface. *Molecular microbiology* **78**: 837-852.

Stefanovic, S., G.J. Bassell & M.R. Mihailescu, (2015) G quadruplex RNA structures in PSD-95 mRNA: potential regulators of miR-125a seed binding site accessibility. *Rna* **21**: 48-60.

Stingl, K., S. Brandt, E.M. Uhlemann, R. Schmid, K. Altendorf, C. Zeilinger, C. Ecobichon, A. Labigne, E.P. Bakker & H. de Reuse, (2007) Channel-mediated potassium uptake in Helicobacter pylori is essential for gastric colonization. *EMBO J* **26**: 232-241.

Stingl, K. & H. De Reuse, (2005) Staying alive overdosed: how does Helicobacter pylori control urease activity? *International journal of medical microbiology : IJMM* **295**: 307-315.

Storz, G., J. Vogel & K.M. Wassarman, (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Molecular cell* **43**: 880-891.

Stougaard, P., S. Molin & K. Nordstrom, (1981) RNAs involved in copy-number control and incompatibility of plasmid R1. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 6008-6012.

Suerbaum, S. & C. Josenhans, (2007) Helicobacter pylori evolution and phenotypic diversification in a changing host. *Nature reviews. Microbiology* **5**: 441-452.

Thibonnier, M., S. Aubert, C. Ecobichon & H. De Reuse, (2010) Study of the functionality of the Helicobacter pylori trans-translation components SmpB and SsrA in an heterologous system. *BMC microbiology* **10**: 91.

Thibonnier, M., J.M. Thiberge & H. De Reuse, (2008) Trans-translation in Helicobacter pylori: essentiality of ribosome rescue and requirement of protein tagging for stress resistance and competence. *PloS one* **3**: e3810.

Thomason, M.K. & G. Storz, (2010) Bacterial antisense RNAs: how many are there, and what are they doing? *Annual review of genetics* **44**: 167-188.

Tjaden, B., S.S. Goodwin, J.A. Opdyke, M. Guillier, D.X. Fu, S. Gottesman & G. Storz, (2006) Target prediction for small, noncoding RNAs in bacteria. *Nucleic acids research* **34**: 2791-2802.

Tomb, J.F., O. White, A.R. Kerlavage, R.A. Clayton, G.G. Sutton, R.D. Fleischmann, K.A. Ketchum, H.P. Klenk, S. Gill, B.A. Dougherty, K. Nelson, J. Quackenbush, L. Zhou, E.F. Kirkness, S. Peterson, B. Loftus, D. Richardson, R. Dodson, H.G. Khalak, A. Glodek, K. McKenney, L.M. Fitzegerald, N. Lee, M.D. Adams, E.K. Hickey, D.E. Berg, J.D. Gocayne, T.R. Utterback, J.D. Peterson, J.M. Kelley, M.D. Cotton, J.M. Weidman, C. Fujii, C. Bowman, L. Watthey, E. Wallin, W.S. Hayes, M. Borodovsky, P.D. Karp, H.O. Smith, C.M. Fraser & J.C. Venter, (1997) The complete genome sequence of the gastric pathogen Helicobacter pylori. *Nature* **388**: 539-547.

Tomizawa, J., T. Itoh, G. Selzer & T. Som, (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 1421-1425.

Tree, J.J., S. Granneman, S.P. McAteer, D. Tollervey & D.L. Gally, (2014) Identification of bacteriophage-encoded anti-sRNAs in pathogenic Escherichia coli. *Molecular cell* **55**: 199-213.

Trotochaud, A.E. & K.M. Wassarman, (2005) A highly conserved 6S RNA structure is required for regulation of transcription. *Nature structural & molecular biology* **12**: 313-319.

Urban, J.H. & J. Vogel, (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. *Nucleic acids research* **35**: 1018-1037.

Urbanowski, M.L., L.T. Stauffer & G.V. Stauffer, (2000) The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in Escherichia coli. *Molecular microbiology* **37**: 856-868.

Valkonen, K.H., T. Wadstrom & A.P. Moran, (1997) Identification of the N-acetylneuraminyllactose-specific laminin-binding protein of Helicobacter pylori. *Infection and immunity* **65**: 916-923.

van der Woude, M.W., (2011) Phase variation: how to create and coordinate population diversity. *Current opinion in microbiology* **14**: 205-211.

van der Woude, M.W. & A.J. Baumler, (2004) Phase and antigenic variation in bacteria. *Clinical microbiology reviews* **17**: 581-611, table of contents.

van Vliet, A.H., E.J. Kuipers, J. Stoof, S.W. Poppelaars & J.G. Kusters, (2004) Acid-responsive gene induction of ammonia-producing enzymes in Helicobacter pylori is mediated via a metal-responsive repressor cascade. *Infection and immunity* **72**: 766-773.

Vanderpool, C.K., D. Balasubramanian & C.R. Lloyd, (2011) Dual-function RNA regulators in bacteria. *Biochimie* **93**: 1943-1949.

Vanderpool, C.K. & S. Gottesman, (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Molecular microbiology* **54**: 1076-1089.

Vecerek, B., I. Moll & U. Blasi, (2007) Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *The EMBO journal* **26**: 965-975.

Veyrier, F.J., C. Ecobichon & I.G. Boneca, (2013) Draft Genome Sequence of Strain X47-2AL, a Feline Helicobacter pylori Isolate. *Genome announcements* **1**.

Viegas, S.C., V. Pfeiffer, A. Sittka, I.J. Silva, J. Vogel & C.M. Arraiano, (2007) Characterization of the role of ribonucleases in Salmonella small RNA decay. *Nucleic acids research* **35**: 7651-7664.

Vinuesa, P., B.L. Reuhs, C. Breton & D. Werner, (1999) Identification of a plasmid-borne locus in Rhizobium etli KIM5s involved in lipopolysaccharide O-chain biosynthesis and nodulation of Phaseolus vulgaris. *Journal of bacteriology* **181**: 5606-5614.

Vogel, C., (2011) Translation's coming of age. *Molecular systems biology* **7**: 498.

Vogel, C. & E.M. Marcotte, (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature reviews. Genetics* **13**: 227-232.

Vogel, J., L. Argaman, E.G. Wagner & S. Altuvia, (2004) The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. *Current biology : CB* **14**: 2271-2276.

Vogel, J. & B.F. Luisi, (2011) Hfq and its constellation of RNA. *Nature reviews. Microbiology* **9**: 578-589.

Wachter, A., (2014) Gene regulation by structured mRNA elements. *Trends in genetics : TIG* **30**: 172-181.

Wade, J.T. & D.C. Grainger, (2014) Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nature reviews. Microbiology* **12**: 647-653.

Wagner, E.G., (2013) Cycling of RNAs on Hfq. *RNA biology* **10**: 619-626.

Wagner, L.A., R.B. Weiss, R. Driscoll, D.S. Dunn & R.F. Gesteland, (1990) Transcriptional slippage occurs during elongation at runs of adenine or thymine in Escherichia coli. *Nucleic acids research* **18**: 3529-3535.

Wang, G., P. Alamuri & R.J. Maier, (2006a) The diverse antioxidant systems of Helicobacter pylori. *Molecular microbiology* **61**: 847-860.

Wang, G., Y. Hong, M.K. Johnson & R.J. Maier, (2006b) Lipid peroxidation as a source of oxidative damage in Helicobacter pylori: protective roles of peroxiredoxins. *Biochimica et biophysica acta* **1760**: 1596-1603.

Wang, G., D.A. Rasko, R. Sherburne & D.E. Taylor, (1999) Molecular genetic basis for the variable expression of Lewis Y antigen in Helicobacter pylori: analysis of the alpha (1,2) fucosyltransferase gene. *Molecular microbiology* **31**: 1265-1274.

Wang, Z., M. Gerstein & M. Snyder, (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* **10**: 57-63.

Wassarman, K.M., A. Zhang & G. Storz, (1999) Small RNAs in Escherichia coli. *Trends in microbiology* **7**: 37-45.

Waters, L.S. & G. Storz, (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615-628.

Wen, Y., J. Feng, D.R. Scott, E.A. Marcus & G. Sachs, (2011) A cis-encoded antisense small RNA regulated by the HP0165-HP0166 two-component system controls expression of ureB in Helicobacter pylori. *Journal of bacteriology* **193**: 40-51.

Wen, Y., E.A. Marcus, U. Matrubutham, M.A. Gleeson, D.R. Scott & G. Sachs, (2003) Acid-adaptive genes of Helicobacter pylori. *Infection and immunity* **71**: 5921-5939.

Westermann, A.J., S.A. Gorski & J. Vogel, (2012) Dual RNA-seq of pathogen and host. *Nature reviews. Microbiology* **10**: 618-630.

Wieland, M. & J.S. Hartig, (2009) Investigation of mRNA quadruplex formation in Escherichia coli. *Nature protocols* **4**: 1632-1640.

Williams, S.M., Y.T. Chen, T.M. Andermann, J.E. Carter, D.J. McGee & K.M. Ottemann, (2007) Helicobacter pylori chemotaxis modulates inflammation and bacterium-gastric epithelium interactions in infected mice. *Infection and immunity* **75**: 3747-3757.

Windle, H.J., A. Fox, D. Ni Eidhin & D. Kelleher, (2000) The thioredoxin system of Helicobacter pylori. *The Journal of biological chemistry* **275**: 5081-5089.

Wright, P.R., J. Georg, M. Mann, D.A. Sorescu, A.S. Richter, S. Lott, R. Kleinkauf, W.R. Hess & R. Backofen, (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic acids research* **42**: W119-123.

Wright, P.R., A.S. Richter, K. Papenfort, M. Mann, J. Vogel, W.R. Hess, R. Backofen & J. Georg, (2013) Comparative genomics boosts target prediction for bacterial small RNAs. *Proceedings of the National Academy of Sciences of the United States of America* **110**: E3487-3496.

Xenarios, I., L. Salwinski, X.J. Duan, P. Higney, S.M. Kim & D. Eisenberg, (2002) DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic acids research* **30**: 303-305.

Xiao, B., W. Li, G. Guo, B. Li, Z. Liu, K. Jia, Y. Guo, X. Mao & Q. Zou, (2009a) Identification of small noncoding RNAs in Helicobacter pylori by a bioinformatics-based approach. *Current microbiology* **58**: 258-263.

Xiao, B., W. Li, G. Guo, B.S. Li, Z. Liu, B. Tang, X.H. Mao & Q.M. Zou, (2009b) Screening and identification of natural antisense transcripts in Helicobacter pylori by a novel approach based on RNase I protection assay. *Molecular biology reports* **36**: 1853-1858.

Yamaoka, Y., (2008) Increasing evidence of the role of Helicobacter pylori SabA in the pathogenesis of gastroduodenal disease. *Journal of infection in developing countries* **2**: 174-181.

Yamaoka, Y., M. Kita, T. Kodama, S. Imamura, T. Ohno, N. Sawai, A. Ishimaru, J. Imanishi & D.Y. Graham, (2002) Helicobacter pylori infection in mice: Role of outer membrane proteins in colonization and inflammation. *Gastroenterology* **123**: 1992-2004.

Yan, Y., S. Su, X. Meng, X. Ji, Y. Qu, Z. Liu, X. Wang, Y. Cui, Z. Deng, D. Zhou, W. Jiang, R. Yang & Y. Han, (2013) Determination of sRNA expressions by RNA-seq in Yersinia pestis grown in vitro and during infection. *PloS one* **8**: e74495.

Yu, H., N.M. Luscombe, H.X. Lu, X. Zhu, Y. Xia, J.D. Han, N. Bertin, S. Chung, M. Vidal & M. Gerstein, (2004) Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome research* **14**: 1107-1118.

Zhang, Y., S. Xie, H. Xu & L. Qu, (2015) CLIP: viewing the RNA world from an RNA-protein interactome perspective. *Science China. Life sciences* **58**: 75-88.

Zhou, K., A. Aertsen & C.W. Michiels, (2014) The role of variable DNA tandem repeats in bacterial adaptation. *FEMS microbiology reviews* **38**: 119-141.

8. List of Figures

9. List of Tables

10. Curriculum vitae

x

11.　List of publications

Publications associated to the present work and beyond:

Mueller SA, **Pernitzsch SR**, Hannge SB, Uetz P, von Bergen M, Sharma CM, Kalkhof S (2015). Stable isotope labeling by amino acids in cell culture based proteomics reveals differences in protein abundances between spiral and coccoid forms of the gastric pathogen *Helicobacter pylori. J Proteomics.* Accepted.
(Research article)

Pernitzsch SR, Tirier SM, Beier D, Sharma CM (2014). A variable homopolymeric G-repeat defines sRNA-mediated posttranscriptional regulation of a chemotaxis receptor in *Helicobacter pylori. PNAS,* 111(4):E501-10.
(Research article)

Müller SA, Findeiß S, **Pernitzsch SR**, Wissenbach DK, Stadler PF, Hofacker IL, von Bergen M, Kalkhof S. (2013). Identification of new protein coding sequences and signal peptidase cleavage sites of *Helicobacter pylori* strain 26695 by proteogenomics. *J Proteomics,* 86:27-42.
(Research article)

Pernitzsch SR, Sharma CM (2012). Transcriptome complexity and riboregulation in the human pathogen *Helicobacter pylori. Frontiers in Cellular and Infection Microbiology,* 2:14.
(Perspective article)

Manuscripts related to this PhD thesis which are in preparation and/or submitted:

Pernitzsch SR, Darfeuille F, Sharma CM (2015). The primary transcriptome and non-coding RNA repertoire of *Helicobacter pylori*. Book chapter, *Helicobacter pylori Research*: From Bench to Bedside. Submitted.
(Book chapter)

Pernitzsch SR, Aul B, Robbe-Saule M, De Reuse H, Sharma CM (2015). Small RNA-mediated gradual expression control of a novel LPS biosynthesis and colonization factor in *Helicobacter pylori.* Manuscript in preparation.
(Research article)

Pernitzsch SR, Aul B, Förstner KU, Bischler T, Sharma CM (2015). RepG, a conserved and abundant riboregulator that controls gene expression by antisense base-pairing to G-rich sequences in *Helicobacter pylori*. Manuscript in preparation.
(Research article)

Previous studies that were published prior to the beginning of, and/or are unrelated to, this PhD thesis:

Jäger D*, **Pernitzsch SR***, Richter A, Backofen R, Sharma CM, Schmitz RA (2012). An archaeal sRNA targeting *cis*- and *trans*-encoded mRNAs via two distinct domains. *Nucleic Acids Res.,* 40(21):10964-79
*Equally contributing authors.
(Research article)

Sharma CM, Papenfort K, **Pernitzsch SR**, Mollenkopf HJ, Hinton JC, Vogel J (2011). Pervasive posttranscriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol. Microbiol.,* 81(5):1144-1165.
(Research article)

12. Acknowledgements

At the end of my thesis, I would like to thank those people who contributed, in whatever form, to this thesis and made it an unforgettable experience to me.

First and foremost, I express my sincere gratitude to my mentor Dr. Cynthia M. Sharma for giving me the opportunity to perform this work in her group, embedded in an excellent and inspiring scientific environment, and for her continuous support during the last years. I enjoyed being a part of her team from the beginning on and thank her for the instructive guidance and great effort she put into my scientific training, and for encouraging me to visit and present my work at national and international conferences.

I am thankful to Prof. Dr. Dagmar Beier, Dr. Tobias Ölschläger and Prof. Dr. Jörg Vogel for being part of my PhD thesis committee and for their precious feedback on my work as well as fruitful discussions during our annual meetings. Also, I thank Prof. Dr. Thomas Dandekar for chairing the doctoral committee and the defense of this thesis.

Thanks to PhD Hilde De Reuse and Prof. Dr. Dagmar Beier for the fruitful collaboration on the HP0102 and HP1043 project, respectively.

I express my deepest gratitude to all present and former colleagues of the "Sharma-Gang" – you guys turn work into fun. Especially, I would like to thank Anika Lins and Belinda Aul for their excellent technical assistance over the last years. I am grateful to my talented student interns, Nina Kapica and Elisabeth Schönwetter, for their assistance to my projects. Also, thanks to Stephan M. Tirier, who literally worked day and night with me on the revision of the RepG–*tlpB* manuscript. I am very indebted to Mona Alzheimer, Gaurav Dugar, PhD Sarah Svensson, Patrick Tan and Dr. Marcus Resch for their continuous help, intellectual input, and great working atmosphere – a spirit of friendly relationships.

Thanks to Hilde Merkert for providing me a "PhD writing place" in the 2nd floor and her small visits. Also, thanks to Barbara Plaschke, Dr. Dimitri Podkaminski, PhD Yanjie Chao, Dr. Kathrin Fröhlich, Dr. Kai Papenfort and Dr. Renate Rieder, who were a great help in the lab (especially in the beginning) and provide a fantastic working atmosphere.

Thanks to the "Bioinformatics-Gang" – especially to Dr. Konrad U. Förstner and Thorsten Bischler – for always having an open door and answering my questions.

I am very grateful to all proof-readers of this thesis – especially Mona Alzheimer, PhD Sarah Svensson, Gaurav Dugar, Hilde Merkert and Dr. Alexander Westermann – for their critical comments and helpful discussions.

13.   Appendices

13.1.   Appendix to Chapter 2
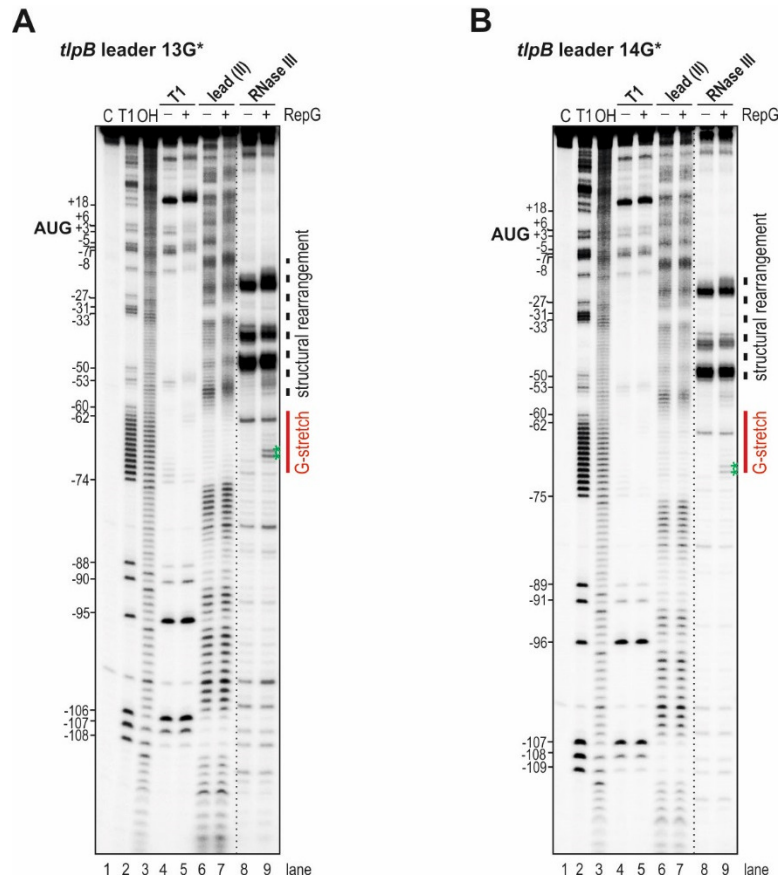
Figure 13.1: RepG-mediated structural rearrangements within *tlpB* mRNA leaders that contain a 13-nt or 14-nt long homopolymeric G-repeat. About 0.1 pmol ^{32}P-labeled *tlpB* mRNA leader variants, which either contain a 13-nt (A) or 14-nt long (B) G-repeat, were treated with RNase T1, lead (II)-acetate and RNase III in the absence or presence of 1000 nM RepG. RNase III cleavage sites in the G-repeat and structural rearrangements in the *tlpB* 5' UTR are indicated by green stars and black bars, respectively. Untreated RNA (lane C), partially alkali- (lane OH) or RNase T1- (lane T1) digested *tlpB* mRNA leader variants served as ladders in structure mapping experiments.

Table 13.1: G-repeat length in the *tlpB* 5' UTRs of sequential *H. pylori* strains/isolates from human or after re-isolation from animals. Lines frame *H. pylori* isolates that were obtained from the same patient or that were re-isolated after mice/gerbil infection studies. The lengths of the G-repeats in the *tlpB* 5' UTRs of isolates from Narino, Colombia, from the study of Kennemann *et al.*, 2011 were re-sequenced by Sanger sequencing.

Strain	NCBI Acc. No	Time scale	G-repeat length	G-repeat variation	Origin/Comment	Reference
908	NC_017357		17G–TGGTTTT–17G		West African duodenal ulcer disease patient in France	(Devi *et al.*, 2010)
2017	NC_017374	10 yrs	15G	yes	Re-isolate from antrum	(Avasthi *et al.*, 2011)
2018	NC_017381	10 yrs	13G	yes	Re-isolate from corpus	
NQ367	NZ_CADL00000000		15/16G°		Isolates from Narino, Colombia	
NQ1671	NZ_CADM00000000	3 yrs	13G	yes		(Kennemann *et al.*, 2011)
NQ4191	NZ_CADN00000000	16 yrs	13G	yes		
NQ392	NZ_CADI00000000		13G°		Isolates from Narino, Colombia	
NQ1707	NZ_CADJ00000000	3 yrs	14G	yes		(Kennemann *et al.*, 2011)
NQ4060	NZ_CADK00000000	16 yrs	15/16G°	yes		
NQ315	NZ_CADE00000000		12G		Isolates from Narino, Colombia	(Kennemann *et al.*, 2011)
NQ1712	NZ_CADF00000000	3 yrs	13G	yes		
NQ352	NZ_CADG00000000		12G°		Isolates from Narino, Colombia	(Kennemann *et al.*, 2011)
NQ1701	NZ_CADH00000000	3 yrs	14G°	no		
Hp141			12G-TGC#		Women with gastritis in Poitiers, France	(Salaun *et al.*, 2005)
Hp141*		150 days	10G-C#	yes	Re-isolate from female C57BL/6 inbred mice	
Hp145			10G		Women with prepyloric ulcer in Poitiers, France	(Salaun *et al.*, 2005)
Hp145*		150 days	10G	no	Re-isolate from female C57BL/6 inbred mice	
HP87			13G		Original human isolate	(Behrens *et al.*, 2013)
HP87 P7*			16/17/18G	yes	Gerbil adapted strain	
HP87 P7 *tlpD*			16/17/18G		*tlpD* mutant of gerbil adapted strain	(Behrens *et al.*, 2013)
HP87 P7 *tlpD* RI		6 weeks	16/17/18G	no	Re-isolate from gerbil antrum	

° The G-repeat length determined by Sanger sequencing differed from the genome sequence determined by 454-sequencing.

* *H. pylori* isolates that were re-isolated from C57BL/6 inbred mice or gerbils.

Additional nucleotide variations that were identified in the flanking region of the homopolymeric G-repeat.

Table 13.2: Location of simple sequence repeats in *H. pylori* and *C. jejuni*. Simple sequence repeats (SSR) in *H. pylori* strain 26695 and *C. jejuni* NCTC 11168 were extracted from previous studies (Fouts *et al.*, 2005, Saunders *et al.*, 1998, Alm *et al.*, 1999, Tomb *et al.*, 1997, Parkhill *et al.*, 2000, Salaun *et al.*, 2004) and classified regarding their location within promoter regions, 5' untranslated regions (5' UTR) or coding region (CDS) of annotated genes. Transcriptional start sites (TSS) from global transcriptome studies of *H. pylori* strain 26695 (Sharma *et al.*, 2010) and *C. jejuni* NCTC 11168 (Dugar *et al.*, 2013) were used for the definition of promoters and 5' UTRs. Gene Cj0565 has two alternative promoters (primary* and secondary# TSS) which lead to two possible locations for the SSR.

H. pylori 26695			*C. jejuni* NCTC 11168		
Promoter	**5' UTR**	**CDS**	**Promoter**	**5' UTR**	**CDS**
HP0009 (A14)	HP0103 (G12)	HP0009 (CT11)	Cj0565* (G12)	Cj0565# (G12)	Cj0031-32 (G10)
HP0025 (T15)	HP0208 (A11)	HP0051 (AG5)	Cj0628-29 (T5)	Cj0618 (G9)	Cj0045c (C10)
HP0227 (T14)	HP0211 (T7)	HP0058 (C15)		Cj0628-29 (G10)	Cj0046 (G11)
HP0228 (A14)	HP0585 (A8)	HP0093-94 (C14)		Cj0676 (G9)	Cj0170 (G9)
HP0349 (T15)	HP0876 (T16)	HP0143 (A7)		Cj1321 (G10)	Cj0208 (G6)
HP0350 (A15)	HP1400 (A16)	HP0208 (AG11)			Cj0279 (G6)
HP0547 (A14)		HP0211 (AT5/A7)			Cj0275 (G8)
HP0629 (T15)		HP0211 (A7)			Cj0348 (G6)
HP0651 (A7)		HP0217 (G12)			Cj0506 (G6)
HP0722 (T15)		HP0298 (T9)			Cj0617 (G9)
HP0725 (T14)		HP0335 (G9)			Cj0628-29 (A5)
HP0733 (T13)		HP0379 (C13)			Cj0684 (G6)
HP0896 (A14)		HP0381 (G7)			Cj0685c (C9)
HP0912 (T13)		HP0464 (C15)			Cj0735 (G6)
HP1342 (A14)		HP0499 (G8)			Cj0959c (G5)
		HP0580 (C8)			Cj1061c (G6)
		HP0586 (A8)			Cj1130c (G6)
		HP0619 (C13)			Cj1139c (G8)
		HP0638 (CT6)			Cj1144-45 (G8)
		HP0642 (G6)			Cj1184c (G6)
		HP0651 (C13)			Cj1238 (G6)
		HP0655 (G8)			Cj1295 (G9)
		HP0657 (G7)			Cj1296-97 (G9)
		HP0684-85 (C9)			Cj1305c (G9)
		HP0687 (G8)			Cj1306c (G9)
		HP0722 (CT8)			Cj1310c (G9)
		HP0725 (CT6)			Cj1318 (G11)
		HP0744 (AG9)			Cj1325-26 (G10)
		HP0752 (G6)			Cj1335-36 (G9)
		HP0753 (G7)			Cj1342c (G9)
		HP0767 (G11)			Cj1370 (G6)
		HP0839 (G7)			Cj1420c (G9)
		HP0855 (GA5)			Cj1421c (G9)
		HP0896 (CT11)			Cj1422c (G9)
		HP0908 (C8)			Cj1426c (G10)
		HP0919 (G9)			Cj1429 (G10)
		HP1206 (A10)			Cj1437c (G9)
		HP1353-54 (C15)			Cj1443c (G5)
		HP1366 (A6)			Cj1643 (G6)
		HP1369-70 (G10)			Cj1677-78 (T7)
		HP1433 (C6)			
		HP1417 (GA9)			
		HP1471 (G14)			
		HP1522 (G12)			

*considering primary TSS
#considering secondary TSS

13.2. Appendix to Chapter 4
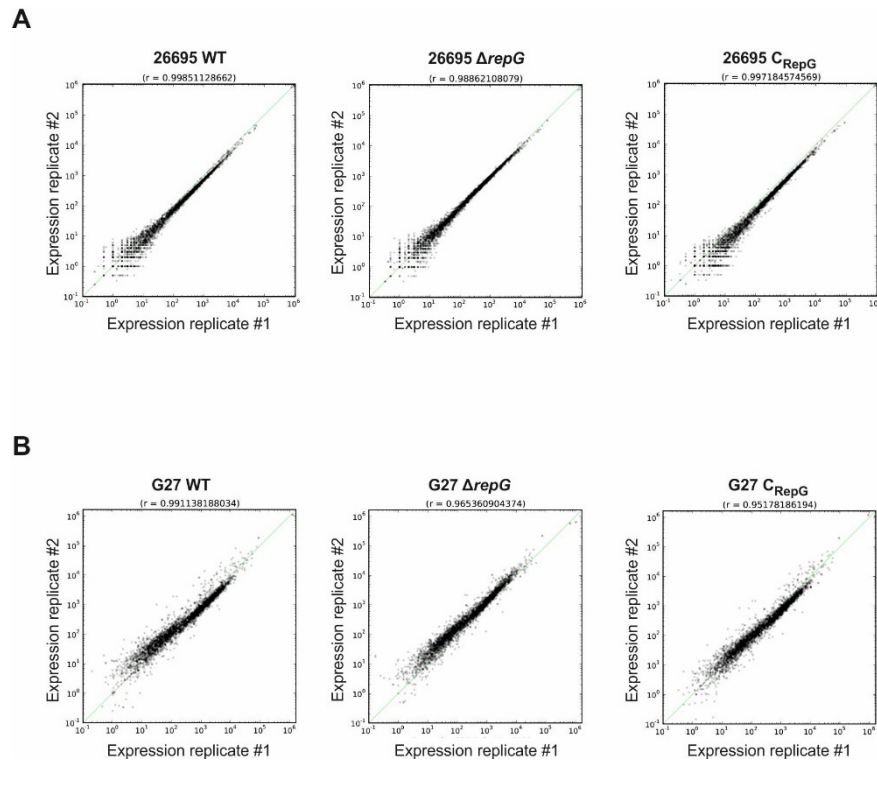
Figure 13.2: Correlations between biological replicates of *H. pylori* 26695 and G27 wildtype and mutant strains. Scatter plots compare expression values between two biological replicates (#1, #2) of *H. pylori* 26695 **(A)** and G27 **(B)** wildtype (WT), *repG* deletion (Δ*repG*) and sRNA complementation (C$_{RepG}$) mutants. Normalized read counts per gene are plotted between biological replicates. The green line represents a theoretical 1:1 fit. The statistical relationship between the replicates is indicated by the Pearson correlation coefficient r; value *r* = 1 means a perfect positive correlation.
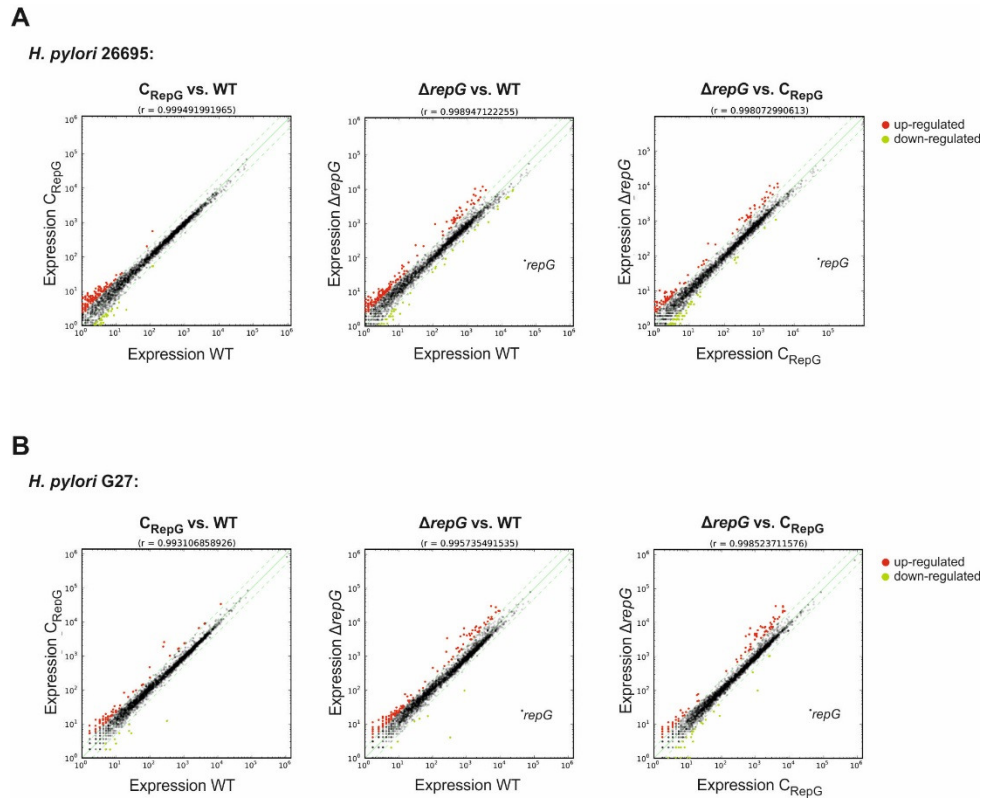
Figure 13.3: Global profiling of RepG-mediated expression changes in *H. pylori* strains 26695 and G27. Scatter plots compare expression values (DE-seq) of wildtype (WT), *repG* deletion mutant (Δ*repG*) and its complementation (C_RepG) in *H. pylori* strains 26695 **(A)** and G27 **(B)**. Normalized read counts per gene were calculated as the mean of two biological replicates and are plotted between wild-type and mutant strains. The green lines give two-fold changes in either direction. Genes which are up- and down-regulated upon *repG* deletion are indicated by red and green dots, respectively.
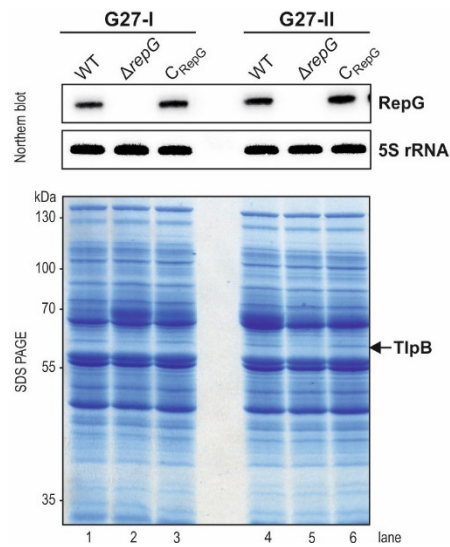
Figure 13.4: *H. pylori* **G27 wildtype, Δ*repG* and sRNA complementation strains used in the RNA-seq study.** Cells were grown to exponential growth phase (OD_{600nm} of ~ 1.0). RNA and protein samples were analyzed on northern blot and SDS-PAGE, respectively. RepG was probed with ^{32}P-labeled CSO-0003 and 5S rRNA with JVO-0485. I/II – biological replicates.
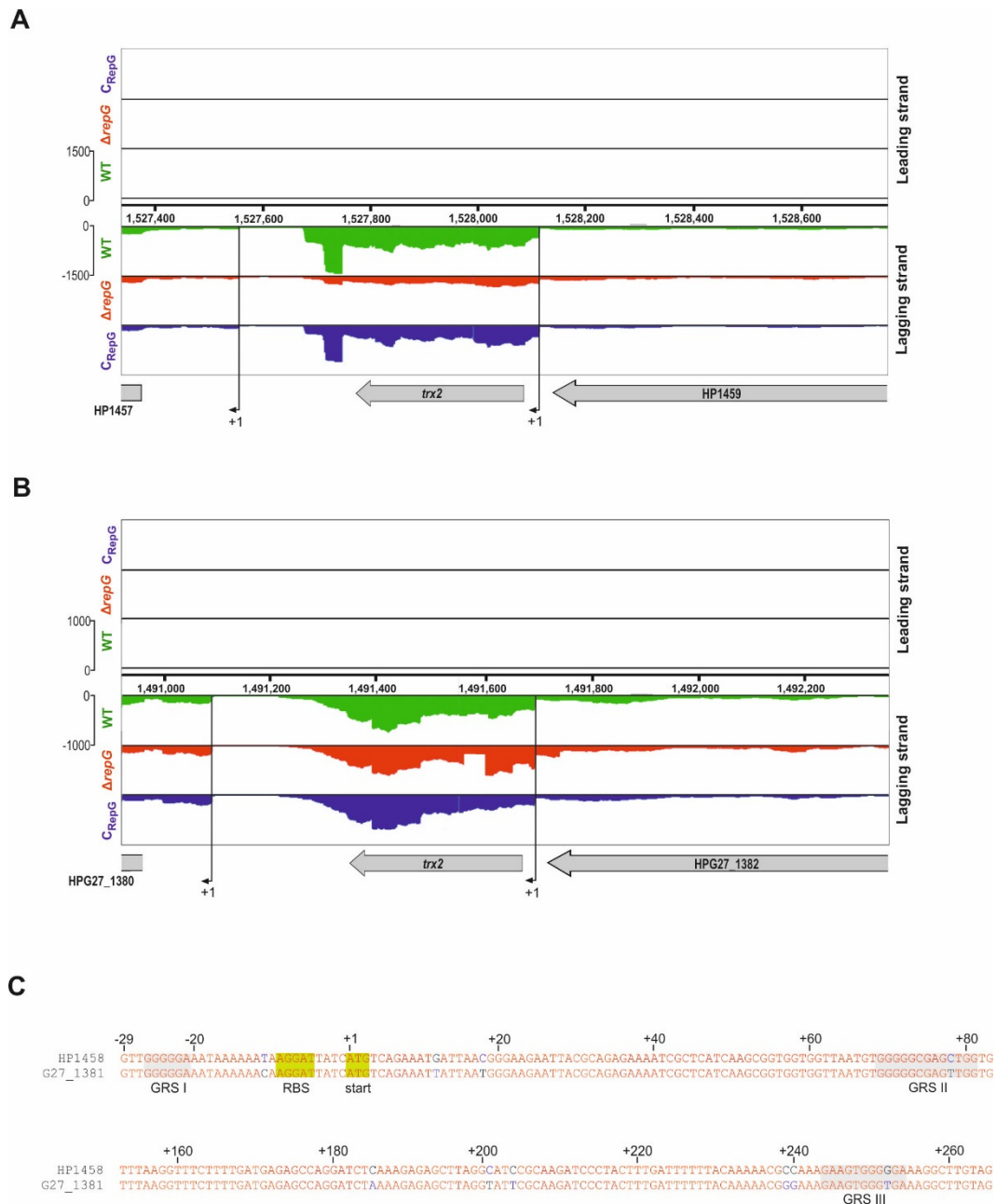
Figure 13.5: RepG activates expression of *trx2* in *H. pylori* strain 26695, but not G27. (A-B) The cDNA reads from the RNA-seq analysis of the wildtype (WT), Δ*repG* and complementation (C_RepG) mutants were mapped to the *trx2* gene in the chromosome of *H. pylori* strains 26695 **(A)** and G27 **(B)**. Gray and black arrows represent the annotated ORFs and published TSS (+1), respectively. **(C)** Sequence alignment showing the predicted RepG binding sites (GRS I-III) in the *trx2* 5' UTR and coding region of *H. pylori* strains 26695 and G27. Numbers indicate positions in the coding region with respect to the annotated start codon in 26695.
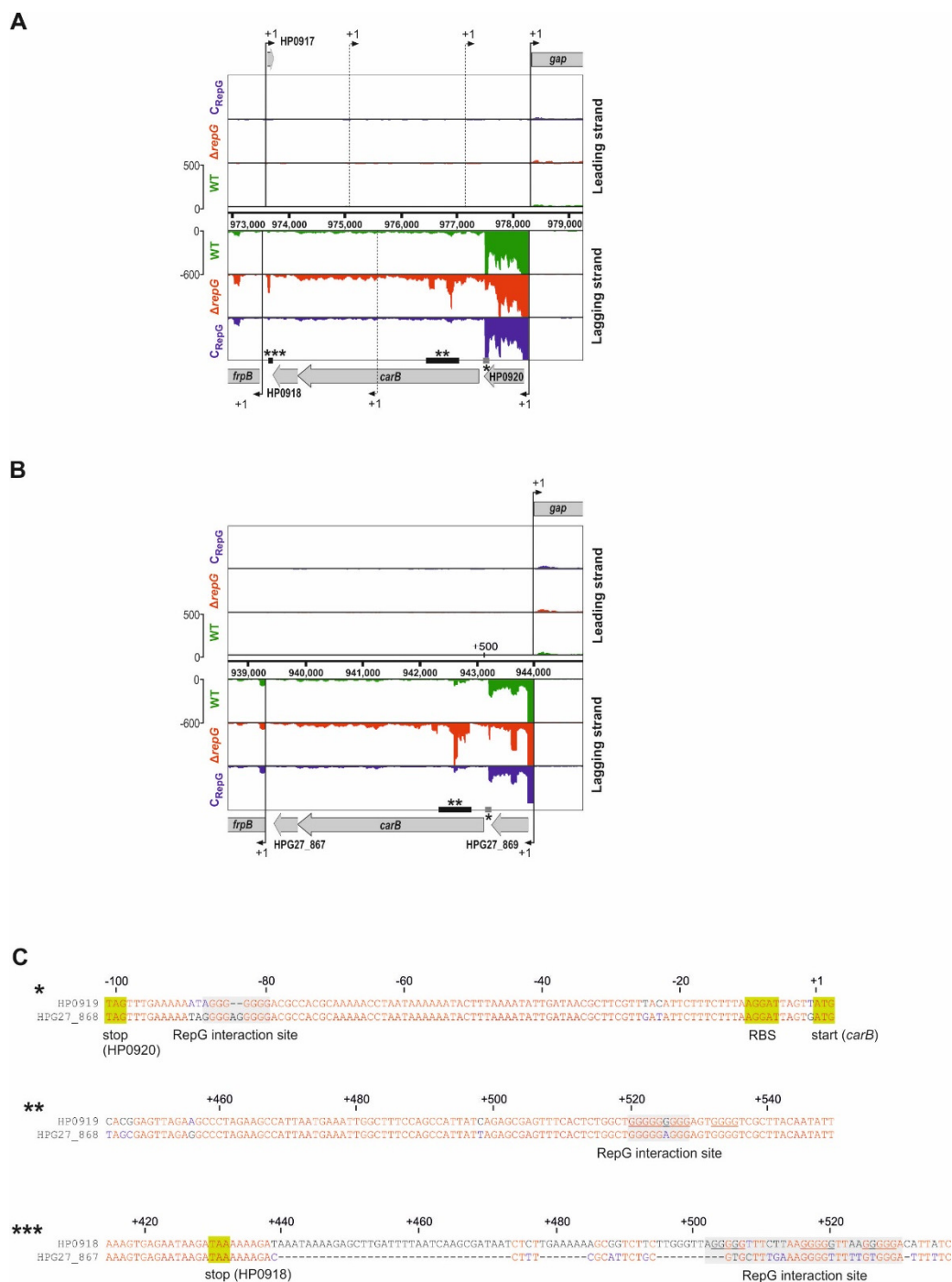
Figure 13.6: RepG represses expression of *carB* independent from the strain background. **(A-B)** The cDNA reads from the RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$ were mapped to the regions flanking *carB* (HP0919) in *H. pylori* strains 26695 **(A)** and G27 **(B)**. The *carB* gene is encoded in an operon with two hypothetical proteins, HP0920 and HP0918. Sequencing reads covering predicted RepG interaction sites accumulated in the intergenic region between HP0920 and *carB*, as well as in the coding region of the carbomoyl-phosphate synthetase (peaks, thick black lines) in Δ*repG* (denoted by */**). The *** cDNA peak /putative 3' UTR of HP0918 is not conserved between both strains; it even seem to be absent in strain G27. **(C)** Sequence alignment of the IGR between HP0920 and *carB* (*top*), the coding region of *carB* (*middle*) and the putative 3' UTR of HP0918 (*bottom*) of *H. pylori* strains 26695 and G27. Numbers indicate positions upstream or within the coding region with respect to the annotated *carB* start codon of 26695 (ATG, green, * and **) and HP0918 translation start (***). The stop codon of HP0918 is shown in green.
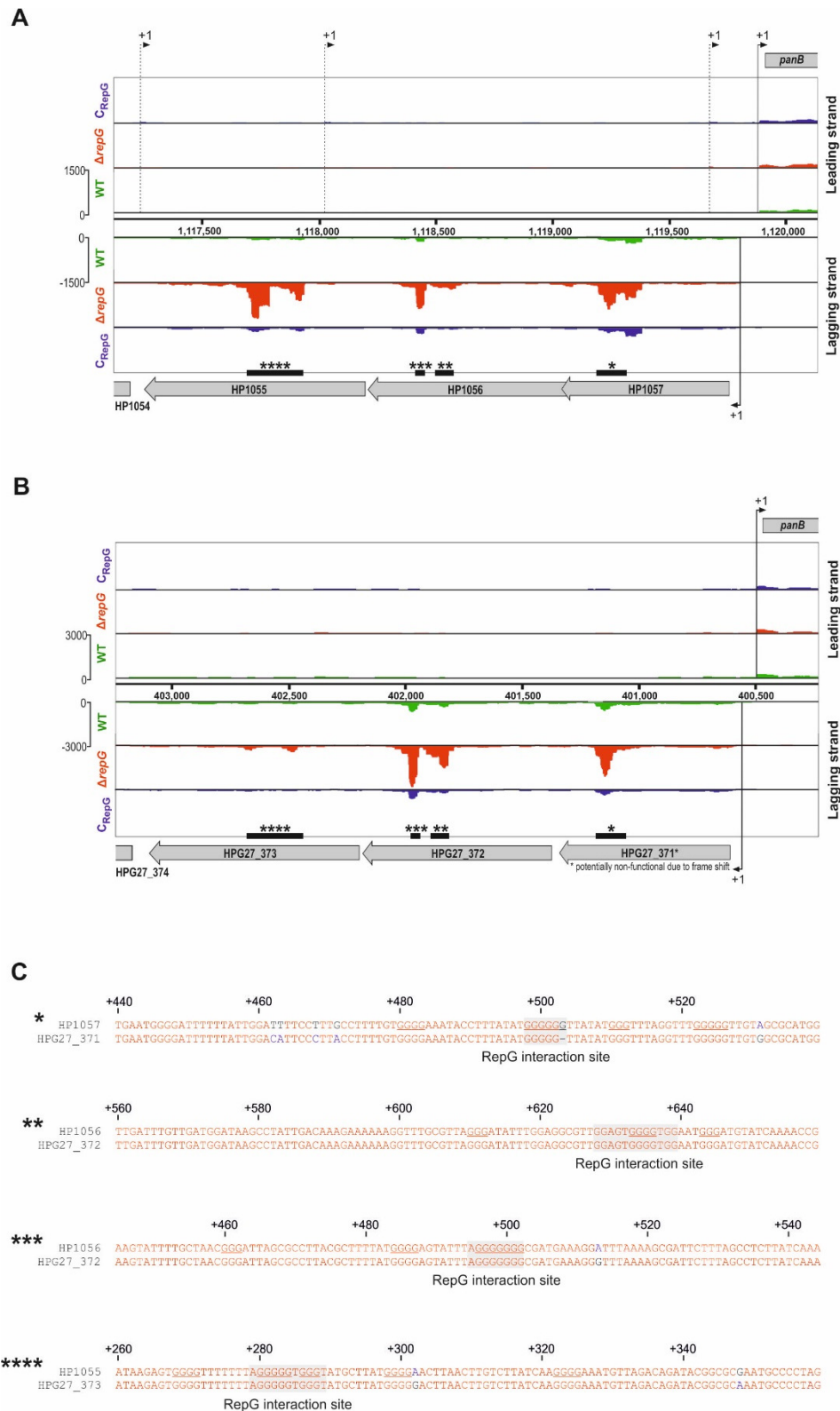
Figure 13.7: RepG controls mRNA levels of the three outer membrane proteins HP1057-HP1055. IGB screenshot showing cDNA reads from the RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$, which were mapped to the HP1057-HP1055 region in *H. pylori* strains 26695 **(A)** and G27 **(B)**. **(C)** Sequence alignment of the predicted RepG binding sites (black, thick black lines) in HP1057 (*top*), HP1056 (*middle**/***) and HP1055 (*bottom*) in *H. pylori* strains 26695 and G27. Numbers indicate positions in the coding region with respect to the annotated start codon of each gene in 26695.
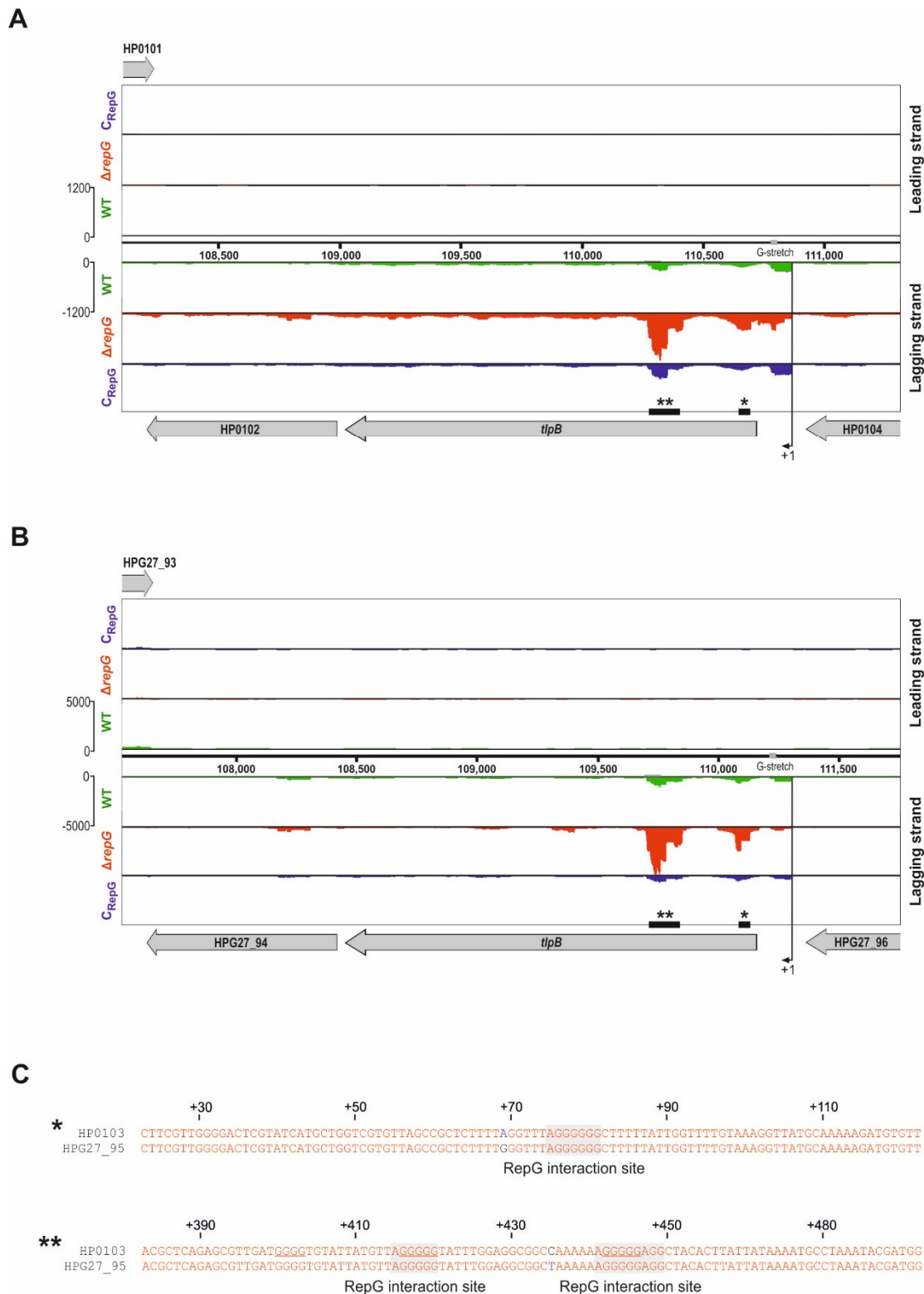
Figure 13.8: The *tlpB* mRNA levels are elevated in *H. pylori* strains 26695 and G27 upon *repG* deletion. IGB screenshot showing cDNA reads from the RNA-seq analysis of WT, Δ*repG* and C$_{RepG}$, which were mapped to the *tlpB*-HP0102 region in *H. pylori* strains 26695 **(A)** and G27 **(B)**. Specific enrichment of sequencing reads covering predicted RepG interaction sites in the *tlpB* coding region (excluding the homopolymeric G-repeat in 5' UTR) are highlighted by thick black bars. **(C)** Sequence alignment of the *tlpB* coding region of *H. pylori* strain 26695 and G27. Numbers indicate position in the coding region with respect to the annotated start codon.

13.3. Appendix to Chapter 6

Table 13.3: Plasmids. Please note that detailed descriptions for the construction of selected plasmids are given in Chapter 6.2., Microbiological methods.

Name	Description/Generation	Origin/ marker	Reference
p463	**GFP-expression vector carrying the promoter of *ureA* fused to the promoterless *gfpmut3* gene**	oriV/ pHP666/ Kan^R	D. S. Merrell, USU, Bethesda, USA
pBA1-1	**Intermediary plasmid for construction of pBA7, deletion of HP0102 in diverse *H. pylori* strains, 500 nt upstream of HP0102 stop codon introduced in pSP60-2.** Ligation of *BamHI/XhoI*-digested Phu-PCR amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0869/0870 and likewise digested pSP60-2.	p15A/ Amp^R/Erm^R	This study B. Aul
pBA3-1	**Intermediary plasmid for construction of pBA9, complementation of *repG* deletion with RepG in *H. pylori* strain X47-2AL using *aac(3)-IV* cassette.** Exchange of *catGC* against *aac(3)-IV* cassette in pSP39-3. Ligation of *BamHI/NheI*-digested Phu-PCR products amplified with CSO-0292/0939 on pUC1813apra and CSO-0940/0938 on pSP39-3 (*DpnI* digested).	p15A/ Amp^R/Gen^R	This study B. Aul
pBA4-2	**Plasmid for complementation of *repG* deletion with RepG in *H. pylori* strain X47-2AL using P$_{catGC}$*aac(3)-IV* cassette.** Exchange of *catGC* against P$_{catGC}$*aac(3)-IV* cassette in pSP39-3. Ligation of *BamHI/NheI*-digested PCR-products amplified with CSO-0313 /0939 on pSS22-1 (pUC1813apra derivate) and CSO-0940/0938 on pSP39-3 (*DpnI* digested).	P15A/ Amp^R/Gen^R	This study B. Aul
pBA5-4	**Intermediary plasmid for construction of pBA13, *tlpB* deletion in *H. pylori* strain X47-2AL.** Ligation of *XbaI/XhoI*-digested PCR-products amplified with CSO-0039/0040 on gDNA of *H. pylori* strain X47-2AL (CSS-0996) and CSO-0873/0874 on pSP39-3 (*DpnI* digested).	p15A/ Amp^R	This study B. Aul
pBA7-4	**Plasmid for deletion of HP0102 in diverse *H. pylori* strains, 500 nt up- and downstream of HP0102 stop codon, based on pBA1-1.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0871/0872 and with CSO-0309/0873 on pBA1-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study B. Aul
pBA9-5	**Plasmid for complementation of *repG* deletion with RepG in *H. pylori* strain X47-2AL using *aac(3)-IV* cassette, based on pBA3-1.** Ligation of *ClaI/NdeI*-digested Phu-PCR products amplified with CSO-0424/0426 on gDNA of *H. pylori* strain X47-2AL (CSS-0996) and with CSO-0146/0147 on pBA3-1 (*DpnI* digested).	p15A/ Amp^R/Gen^R	This study B. Aul
pBA13-5	**Plasmid for deletion of *tlpB* in *H. pylori* strain X47-2AL using *aac(3)-IV* cassette.** Ligation of *EcoRI* -digested PCR products of CSO-1745/0942 on pBA5-4 (*DpnI* digested) and CSO-0263/0293 on pUC1813apra.	p15A/ Amp^R/Gen^R	This study B. Aul
pPT3-1	**Plasmid for transcriptional fusion of the *ureA* promoter to *gfpmut3*.** Plasmid with P$_{ureA}$*gfpmut3* from p463 for cloning in *rdxA* locus of *H. pylori* strain 26695 and G27. Ligation of *SalI/NotI*-digested Phu-PCR products amplified from p463 with CSO-0440/0441 and from pSP39-3 with CSO-0442/0443.	p15A/ Amp^R/Cm^R	P. Tan
pGG17.1	**Plasmid for 3xFLAG-tagging of HP1423 based on pJV752.1, 500 nt up- and downstream of HP1423 stop codon, 3xFLAG, *aphA-3* cassette.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from pGG4 (*DpnI* treated) with CSO-0074/0075 and with CSO-0078/0079 from gDNA of *H. pylori* strain 26695 (CSS-0004).	p15A/ Amp^R/ Kan^R	G. Golferi
pJV752.1	**Cloning vector, pZE12-*luc* with modified p15A origin**	p15A/ Amp^R	(Sharma *et al.*, 2007)
pMA5-2	**Plasmid for translational fusion of the *cagA* 5' UTR including the 28th amino acid to *gfpmut*.** Ligation of *ClaI/NheI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain G27 (CSS-0010) with CSO-0284/0590 and with CSO-146/0683 on pPT3-1 (*DpnI* digested).	p15A/ Amp^R/Cm^R	M. Alzheimer
pSP39-3	**Plasmid for complementation of *repG* deletion with RepG in *H. pylori* strain 26695.** Ligation of *XbaI/XhoI*-digested pJV752.1 and Phu-PCR product amplified from gDNA of *H. pylori* 26695 C$_{RepG}$ (Jerémy Regnier & Fabien Darfeuille, Bordeaux, France) with CSO-0017/0018.	P15A/ Amp^R/Cm^R	This study

Name	Description/Generation	Origin/ marker	Reference
pSP42-1	**Plasmid for complementation of *repG* deletion mutant with RepG SL 2 (30-87 nt) in *H. pylori* strain 26695.** Ligation of cycle-PCR product (*DpnI* digested) which was amplified from pSP39-3 using CSO-0081/0080.	p15A/ AmpR/CmR	This study
pSP55-4	**Intermediary plasmid for construction of pSP57, 3xFLAG-tagging of HP0103 (*tlpB*) based on pJV752.1, 500 nt up- and downstream of *tlpB* stop codon.** Ligation of *XbaI/XhoI*-digested Phu-PCR product amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0208/0211 and likewise digested pJV752.1.	p15A/ AmpR	This study
pSP57-4	**Plasmid for 3xFLAG-tagging of HP0103 (*tlpB*) based on pSP55-4, 500 nt up- and downstream of *tlpB* stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI*-digested Phu-PCR product amplified from pSP55-4 with CSO-0210/0245 (*DpnI* digested and CIP treated) and overlap-PCR product of 3xFLAG:*rpsL-erm* (PCR I: CSO-0065/0046 on gDNA JVS-7033 for 3xFLAG and PCR II: CSO-0045/0209 on gDNA of *H. pylori* 26695 *ΔtlpB* (CSS-0163) for *rpsL-erm* cassette).	P15A/ AmpR/ErmR	This study
pSP58-5	**Intermediary plasmid for construction of pSP60.** Ligation of *XhoI/XbaI*-digested Phu-PCR product amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) with CSO-0291/0040, and likewise digested pJV-752.1.	p15A/ AmpR	This study
pSP60-2	**Backbone plasmid for deletion or nucleotide exchange in G-stretch in 5' UTR of *tlpB*.** Ligation of *EcoRI /BamHI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 *ΔtlpB* (CSS-0163) with CSO-0308/0309 and with CSO-0294/0295 on pSP58-5 (*DpnI* digested).	p15A/ AmpR/ErmR	This study
pSP64-1	**Plasmid for deletion of G-stretch in *tlpB* 5' UTR, based on pSP60-2.** Cycle-PCR (Phu) with CSO-0318/0319 on pSP60-2 (*DpnI* digested, direct transformation of PCR in *E. coli* TOP 10).	p15A/ AmpR/ErmR	This study
pSP65-4	**Plasmid for compensatory base pair exchange in G-stretch in *tlpB* 5' UTR; based on pSP60-2, 1xC*.** Cycle PCR (Phu) with CSO-0314/0315 on pSP60-2 (*DpnI* digested, direct transformation of PCR in *E. coli* TOP 10).	p15A/ AmpR/ErmR	This study
pSP66-4	**Plasmid for compensatory base pair exchange in G-stretch in *tlpB* 5' UTR; based on pSP60-2, 3xC.** Cycle PCR (Phu) with CSO-0316/0317 on pSP60-2 (*DpnI* digested, direct transformation of PCR in *E. coli* TOP 10).	p15A/ AmpR/ErmR	This study
pSP70-1	**Plasmid for *tlpB*::3xFLAG based on pSP57-4, for markerless exchange.** Ligation of cycle-PCR (Phu) amplified with CSO-0428/0429 on pSP57-4 (*DpnI* digested, direct transformation of PCR in *E. coli* TOP 10).	p15A/ AmpR	This study
pSP73-1	**Plasmid for variation of G-stretch length, 6G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0448 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP74-1	**Plasmid for variation of G-stretch length, 7G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0449 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP75-1	**Plasmid for variation of G-stretch length, 8G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0450 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP76-1	**Plasmid for variation of G-stretch length, 9G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0451 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP77-1	**Plasmid for variation of G-stretch length, 10G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0452 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP78-1	**Plasmid for variation of G-stretch length, 11G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0453 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP79-4	**Plasmid for variation of G-stretch length, 13G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0454 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP80-1	**Plasmid for variation of G-stretch length, 14G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0455 on pSP64-1.	p15A/ AmpR/ErmR	This study
pSP81-5	**Plasmid for variation of G-stretch length, 15G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0456 on pSP64-1.	p15A/ AmpR/ErmR	This study

Name	Description/Generation	Origin/ marker	Reference
pSP82-1	**Plasmid for variation of G-stretch length, 16G.** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-0318/0457 on pSP64-1.	p15A/ Amp^R/Erm^R	This study
pSP84-1	**Plasmid for complementation of *repG* deletion with RepG in *H. pylori* strain G27.** Ligation of *NdeI/ClaI* digested Phu-PCR products amplified from gDNA of *H. pylori* strain G27 (CSS-0010) with CSO-0425/0427 (RepG of G27) and from pSP39-3 with CSO-0146/0147 (*DpnI* digested).	p15A/ Amp^R/Cm^R	This study
pSP91-3	**Plasmid for promoter exchange (*cagA* instead of *tlpB*.** Ligation of *BamHI*-digested Phu-PCR products amplified with CSO-306/0431 on gDNA of *H. pylori* strain 26695 (CSS-0004) and from pSP60-2 using CSO-430/0308 (*DpnI* treated).	p15A/ Amp^R/Erm^R	This study
pSP109-6	**Plasmid for translational fusion of the *tlpB* 5' UTR including 5th amino acid to *gfpmut3*.** Ligation of *ClaI/NheI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) with CSO-0581/0126 and with CSO-146/0683 on pT3-1 (*DpnI* digested).	p15A/ Amp^R/Cm^R	This study
pSP112-6	**Intermediary plasmid for construction of pSP155, 3xFLAG-tagging of HP0724 (*dcuA*) in *H. pylori* strain 26695, 500 nt upstream of *dcuA* stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0882/0883 and JVO-5142/0874 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP113-1	**Intermediary plasmid for construction of pSP117, C/N-terminal 3xFLAG-tagging of HP1458 (*trx2*) in *H. pylori* strain 26695, 500 nt upstream of *trx2* stop codon, 3xFLAG, *rpsL-erm* cassette** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0890/0891 and JVO-5142 x CSO-0874 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP127-3	**Plasmid for deletion of the *tlpB*-HP0102 operon in *H. pylori* strain X47-2AL using *aac(3)-IV* cassette.** Ligation of *EcoRI/XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* X47-2AL (CSS-0996) with CSO-0871/1359 (500 nt downstream of HP0102) and with CSO-0874/0293 on pBA13-5 (*DpnI* digested).	p15A/ Amp^R/Gen^R	This study
pSP128-5	**Intermediary plasmid for construction of pSP146/pSP147, N-terminal 3xFLAG-tagging of HP1458 (*trx2*) in *H. pylori* strains 26695 and G27, 500 nt upstream of *trx2* start codon introduced in pSP113-1, *rpsL-erm* cassette** Ligation of *BamHI/XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-1283/1365 and with CSO-0874/0308 on pSP113-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP134-1	**Intermediary plasmid for construction of pSP141, 3xFLAG-tagging of HP1055 in *H. pylori* strain 26695, 500 nt downstream of HP1055 (26695) stop codon, 3xFLAG, *rpsL-erm* cassette** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-975/0976 and with CSO-0309/0873 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP135-1	**Intermediary plasmid for construction of pSP142, 3xFLAG-tagging of HP1055 in *H. pylori* strain G27, 500 nt downstream of HP1055 (G27) stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain G27 (CSS-0010) using CSO-975/0976 and with CSO-0309/0873 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP136-1	**Intermediary plasmid for construction of pSP173, 3xFLAG-tagging of HP1056 in *H. pylori* strains 26695 and G27, 500 nt downstream of HP1056 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0970/0971 and with CSO-0309/0873 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP137-1	**Intermediary plasmid for construction of pSP154, 3xFLAG-tagging of HP1057 in *H. pylori* strains 26695 and G27, 500 nt downstream of HP1057 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-958/0959 and with CSO-0309/0873 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP139-1	**Intermediary plasmid for construction of pSP144, 3xFLAG-tagging of HP1181 in *H. pylori* strains 26695 and G27, 500 nt downstream of HP1181 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0901/0902 and with CSO-0309/0873 on pSP57-4 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study

Name	Description/Generation	Origin/ marker	Reference
pSP141-1	**Plasmid for 3xFLAG-tagging of HP1055 in *H. pylori* strain 26695, based on pSP134-1, 500 nt up- and downstream of HP1055 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0974/0972 and with JVO-5142/0874 on p134-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP142-1	**Plasmid for 3xFLAG-tagging of HP1055 in *H. pylori* strain G27, based on pSP135-1, 500 nt up- and downstream of HP1055 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain G27 (CSS-0010) with CSO-0974/0973 and with JVO-5142/0874 on pSP135-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP144-1	**Plasmid for 3xFLAG-tagging of HP1181 in *H. pylori* strains 26695 and G27, based on pSP139-1, 500 nt up- and downstream of HP1181 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0899/0900 and with JVO-5142/0874 on pSP139-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP146-3	**Intermediary plasmid for construction of pSP187, N-terminal 3xFLAG-tagging of HP1458 (*trx2*) in *H. pylori* strain 26695, 500 nt downstream of *trx2* start codon introduced in pSP128-5, *rpsL-erm* cassette** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-1364/1284 and CSO-0309 /0873 on pSP128-5 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP147-3	**Intermediary plasmid construction of pSP188, N-terminal 3xFLAG-tagging of HP1458 (*trx2*) in *H. pylori* strain G27, 500 nt downstream of *trx2* start codon introduced in pSP128-5, *rpsL-erm* cassette** Ligation of *EcoRI/XbaI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain G27 (CSS-0010) using CSO-1364/1284 and CSO-0309/873 on pSP128-5 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP154-2	**Plasmid for 3xFLAG-tagging of HP1057 in *H. pylori* strains 26695 and G27, based on pSP137-1, 500 nt up- and downstream of HP1057 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0956/0957 and with JVO-5142/0874 on pSP137-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP155-1	**Plasmid for 3xFLAG-tagging of HP0724 (*dcuA*) in *H. pylori* strain 26695, based on pSP112-6, 500 nt up- and downstream of *dcuA* stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *EcoRI/XbaI*-digested PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-0884/0885 and with CSO-0309/0873 on pSP112-6 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP173-6	**Plasmid for 3xFLAG-tagging of HP1056 in *H. pylori* strains 26695 and G27, based on pSP136-1, 500 nt up- and downstream of HP1056 stop codon, 3xFLAG, *rpsL-erm* cassette.** Ligation of *XhoI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain 26695 (CSS-0004) using CSO-1546/1547 and JVO-5142 x CSO-0874 on pSP136-1 (*DpnI* digested).	p15A/ Amp^R/Erm^R	This study
pSP186-2	**Plasmid for construction of HP0102 deletion mutant in *H. pylori* strain X47-2AL using *aac(3)-IV* cassette.** Ligation of *XbaI/BamHI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain X47-2AL (CSS-0996) using CSO-1737/1739 and with CSO-0873/0313 on pSP127-3 (*DpnI* digested).	p15A/ Amp^R/Gen^R	This study
pSP187-8	**Plasmid for construction of N-terminal 3xFLAG-tag of HP1458 (*trx2*) in *H. pylori* strain 26695, based on pSP146-3, 500 nt up- and downstream of *trx2* start codon, 3xFLAG, *rpsL-erm* cassette** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-1748/1795 on pSP146-3.	p15A/ Amp^R/Erm^R	This study
pSP188-2	**Plasmid construction of N-terminal 3xFLAG-tag of HP1458 (*trx2*) in *H. pylori* strain G27, based on pSP147-3, 500 nt up- and downstream of *trx2* start codon, 3xFLAG, *rpsL-erm* cassette** Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-1749/1796 on pSP147-3.	p15A/ Amp^R/Erm^R	This study
pSP189-4	**Intermediary plasmid for construction of pSP190-1, complementation of *H. pylori* strain X47-2AL Δ*tlpB*-HP0102 or ΔHP0102 mutants with the *tlpB*-HP0102 operon, based on pBA4-2.** Ligation of *ClaI/NdeI*-digested Phu-PCR products amplified from gDNA of *H. pylori* strain X47-2AL (CSS-0996) using CSO-1740/1741 and with CSO-0146/0147 on pBA4-2 (*DpnI* digested).	p15A/ Amp^R/Gen^R	This study

Name	Description/Generation	Origin/marker	Reference
pSP190-1	**Plasmid for construction of complementation of *H. pylori* strain X47-2AL Δ*tlpB*-HP0102 or ΔHP0102 mutants with the *tlpB*-HP0102 operon using *aphA-3* resistance cassette, based on pSP189-4.** Ligation of *BamHI/NheI*-digested Phu-PCR products amplified from gDNA of *H. pylori* 26695 Δ*repG* (JVS-7014) using CSO-1813/1814 and with CSO-0940/0938 on pSP189-4 (*DpnI* digested).	p15A/ Amp^R/Kan^R	This study
pSP192-1	**Plasmid for construction of complementation of *H. pylori* strain X47-2AL Δ*tlpB*-HP0102 or ΔHP0102 mutants with HP0102 using *aphA-3* resistance cassette, based on pSP190-1.** P$_{tlpB}$(IGR between *tlpB* and HP0102)-HP0102 in *rdxA* locus. Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-1743/1742 on pSP190-1.	p15A/ Amp^R/Kan^R	This study
pSP195-6	**Plasmid for fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including 10th amino acid to *gfpmut3*, under control of the native promoter (P$_{tlpB}$), including *tlpB*, *tlpB*-HP0102 operon from *H. pylori* strain 26695.** *tlpB*-HP0102 10th::*gfpmut3* in *rdxA* locus of *H. pylori* strain G27. Ligation of *ClaI/NheI*-digested Phu-PCRs amplified with CSO-0146/0683 on pMA5-2 (*DpnI* digested) and with CSO-0581/1803 on gDNA of *H. pylori* strain 26695 (CSS-0004).	p15A/ Amp^R/Cm^R	This study
pSP197-3	**Plasmid for fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including 10th amino acid to *gfpmut3*, under control of the native promoter (P$_{tlpB}$), with *tlpB* mini gene (21 aa), *tlpB*-HP0102 from *H. pylori* strain 26695, based on pSP195-6.** *tlpB*$_{mini}$-HP0102 10th::*gfpmut3* in *rdxA* locus of *H. pylori* strain G27. Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-1984/1985 on pSP196-5.	p15A/ Amp^R/Cm^R	This study
pSP198-4	**Plasmid for fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including 10th amino acid to *gfpmut3*, under control of the native promoter (P$_{tlpB}$), without *tlpB*, HP0102 from *H. pylori* strain 26695, based on pSP195-6.** P$_{tlpB}$HP0102 10th::*gfpmut3* in *rdxA* locus of *H. pylori* strain G27. Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-2056/2055 on pSP195-6.	p15A/ Amp^R/Cm^R	This study
pSP200-2	**Plasmid for fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including 10th amino acid to *gfpmut3*, under control of the native promoter (P$_{tlpB}$), including *tlpB*, *tlpB*-HP0102 operon from *H. pylori* strain 26695, ΔG^{HP0102}, based on pSP195-6.** *tlpB*-HP0102 10th::*gfpmut3* ΔG^{HP0102} in *rdxA* locus of *H. pylori* strain G27. Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-2052/2053 on pSP195-6.	p15A/ Amp^R/Cm^R	This study
pSP201-1	**Plasmid for fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including 10th amino acid to *gfpmut3*, under control of the native promoter (P$_{tlpB}$), including *tlpB*, *tlpB*-HP0102 operon from *H. pylori* strain 26695, ATTTA^{HP0102}, based on pSP195-6.** *tlpB*-HP0102 10th::*gfpmut3* ATTTA^{HP0102} in *rdxA* locus of *H. pylori* strain G27. Direct transformation (in *E. coli* TOP 10) of *DpnI*-digested Phu-PCR amplified with CSO-2053/2061 on pSP195-6.	p15A/ Amp^R/Cm^R	This study
pSS22-1	**Plasmid carrying gentamicin cassette under control of the *catGC* cassette (P$_{catGC}$*aac(3)-IV*)**	p15A/ Amp^R/Gen^R	S. Stahl / C. Sharma
pTM117	**Cloning vector carrying an *H. pylori* origin of replication, *aphA-3* resistance cassette and a multiple cloning site upstream of the promoterless *gfpmut3* gene**	oriV/pHP66 6/Kan^R	(Carpenter *et al.*, 2007)
pUC1813apra	**Plasmid carrying apramycin/gentamicin resistance cassette (*aac(3)-IV*)**	pMB1/ Amp^R/Gen^R	(Bury-Mone *et al.*, 2003)
pRR05	**Plasmid containing dimeric MS2-tag**	Amp^R	R. Rieder/ J. Vogel
pHP135	**Plasmid carrying N-terminal truncated *rnj* gene (P$_i$-ΔN-*rnj*)**	Cm^R	(Redko *et al.*, 2013)
pILL2157	***E. coli*/*H. pylori* shuttle vector**	Cm^R	(Boneca *et al.*, 2008)

Table 13.4: Bacterial strains. Please note that detailed descriptions for the construction of selected *H. pylori* mutants are given in Chapter 6.2., Microbiologiocal methods.

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
WT/26695	26695		Wildtype (NCBI Acc-no. NC_000915), kindly provided by T. F. Meyer (MPI-IB, Berlin, Germany); Ref. (Tomb *et al.*, 1997)	CSS-0004				
Δ*repG*	26695	*repG::aphA-3*	Deletion of *repG* in CSS-0004 (26695), details for construction are published in (Sharma *et al.*, 2010)	JVS-7014	Kan^R			
C$_{RepG}$	26695	*repG::aphA-3, rdxA::repG:catGC*	Complementation of Δ*repG* (JVS-7014) with wild-type sRNA; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0046	Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP39-3 in JVS-7014		
SL 2	26695	*repG::aphA-3, rdxA::repG-SL 2:catGC*	Complementation of Δ*repG* (JVS-7014) with mutant RepG SL 2; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0747	Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP42-1 in JVS-7014		
ΔCU	26695	*repG::aphA-3, rdxA::repG-ΔCU:catGC*	Complementation of Δ*repG* (JVS-7014) with mutant RepG ΔCU; overlap-PCR; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0157	Kan^R Cm^R	CSO-0017/0139 on pSP39-3		CSO-0018/0138 on pSP39-3
3xG	26695	*repG::aphA-3, rdxA::repG-3xG:catGC*	Complementation of Δ*repG* (JVS-7014) with mutant RepG 3xG; overlap-PCR; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0158	Kan^R Cm^R	CSO-0017/0143 on pSP39-3		CSO-0018/0142 on pSP39-3
1xG*	26695	*repG::aphA-3, rdxA::repG-1xG*:catGC*	Complementation of Δ*repG* (JVS-7014) with mutant RepG 1xG*; overlap-PCR; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0159	Kan^R Cm^R	CSO-0017/0141 on pSP39-3		CSO-0018/0140 on pSP39-3
Δ*tlpB*	26695	*tlpB::rpsL-erm*	Deletion of HP0103 (*tlpB*) in CSS-0004, overlap-PCR; verification by colony-PCR using CSO-0051 and CSONIH-0033 on gDNA	CSO-0163	Erm^R	CSO-0040/0037 on gDNA of CSS-0004	CSO-0035/0036 on gDNA of *H. pylori* with *rpsL-erm* cassette (Dailidiene *et al.*, 2006)	CSO-0038/0039 on gDNA of CSS-0004
Δ*tlpB*/Δ*repG*	26695	*tlpB::rpsL-erm, repG::aphA-3*	Double deletion of *tlpB* and *repG*; verification by colony-PCR using JVO-5069/5257 on gDNA	CSO-0164	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0163		
tlpB::3xFLAG	26695	*tlpB::tlpB-3xFLAG:rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *tlpB* in the chromosome; verification by colony-PCR using CSO-0050/0046 on gDNA and sequencing with CSO-0208	CSO-0190	Erm^R	Transformation of PCR product amplified with CSO-0208/0211 on pSP57-4 in CSS-0004		
tlpB::3xFLAG/ Δ*repG*	26695	*tlpB::tlpB-3xFLAG:rpsL-erm, repG::aphA-3*	Deletion of *repG* in *H. pylori* carrying C-terminal 3xFLAG-tagging of *tlpB* (CSS-0190) at its native locus; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0215	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSO-0190		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
tlpB::3xFLAG/ ΔrepG/C$_{RepG}$	26695	*tlpB::tlpB*-3xFLAG:*rpsL-erm, repG::aphA-3, rdxA::repG:catGC*	Complementation of CSS-0215 with wild-type RepG; verification by colony-PCR using CSO-0205/-0207 on gDNA and sequencing with CSO-0206	CSS-0285	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0046 in CSS-0215		
P$_{tlpB}$	26695	*tlpB::tlpB* WT:*rpsL-erm*	Introduction of the *rpsL-erm* resistance cassette upstream of the *tlpB* promoter; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0384	Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP60-2 in CSS-0004		
P$_{tlpB}$/ΔrepG	26695	*tlpB::tlpB* WT:*rpsL-erm, repG::aphA-3*	Deletion of *repG* in *H. pylori* carrying *rpsL-erm* resistance cassette upstream of the *tlpB* promoter; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0388	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0384		
P$_{tlpB}$/ ΔrepG/C$_{RepG}$	26695	*tlpB::tlpB* WT:*rpsL-erm, repG::aphA-3, rdxA::repG:catGC*	Complementation of CSS-0388 with wildtype RepG; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0390	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP39-3 in CSS-388		
P$_{tlpB}$/ ΔrepG/3xG	26695	*tlpB::tlpB* WT:*rpsL-erm, repG::aphA-3, rdxA::repG-3xG:catGC*	Complementation of CSS-0388 with RepG mutant 3xG; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0392	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0158 in CSS-388		
P$_{tlpB}$/ ΔrepG/1xG*	26695	*tlpB::tlpB* WT:*rpsL-erm, repG::aphA-3, rdxA::repG-1xG*:catGC*	Complementation of CSS-0388 with RepG mutant 1xG*; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0391	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0159 in CSS-388		
tlpB ΔG	26695	*tlpB::tlpB* ΔG:*rpsL-erm*	Deletion of the G-repeat in the *tlpB* mRNA leader; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0385	Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP64-1 in CSS-0004		
tlpB 3xC	26695	*tlpB::tlpB* 3xC:*rpsL-erm*	Triple nucleotide exchange (3xG → 3xC) in the G-repeat in the *tlpB* mRNA leader; compensatory base-pair exchange to RepG mutant 3xG; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0386	Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP66-4 in CSS-0004		
tlpB 3xC/ ΔrepG	26695	*tlpB::tlpB* 3xC:*rpsL-erm, repG::aphA-3*	Deletion of *repG* in CSS-0386; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0389	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0386		
tlpB 3xC/ ΔrepG/C$_{RepG}$	26695	*tlpB::tlpB* 3xC:*rpsL-erm, repG::aphA-3, rdxA::repG:catGC*	Complementation of CSS-389 with wild-type RepG; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS0-393	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP39-3 in CSS-0389		
tlpB 3xC/ ΔrepG/3xG	26695	*tlpB::tlpB* 3xC:*rpsL-erm, repG::aphA-3, rdxA::repG-3xG:catGC*	Complementation of CSS-389 with RepG mutant 3xG; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0394	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0158 in CSS-0389		
tlpB 1xC*	26695	*tlpB::tlpB* 1xC*:*rpsL-erm*	Single nucleotide exchange (1xG* → 1xC*) in the G-repeat in the *tlpB* mRNA leader; compensatory base-pair exchange to RepG mutant 1xG*; verification by colony-PCR using CSO-0051/CSO-0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0387	Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP65-4 in CSS-0004		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
tlpB* 1xC*/ Δ*repG	26695	*tlpB::tlpB* 1xC*:*rpsL-erm, repG::aphA-3*	Deletion of *repG* in CSS-0387; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0397	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0387		
***tlpB* ΔG/ Δ*repG*/C$_{RepG}$**	26695	*tlpB::tlpB* ΔG:*rpsL-erm, repG::aphA-3, rdxA::repG:catGC*	Complementation of CSS-397; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0400	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP39-3 in CSS-0397		
tlpB* ΔG/ Δ*repG*/1xG	26695	*tlpB::tlpB* ΔG:*rpsL-erm, repG::aphA-3, rdxA::repG-*1xG*:*catGC*	Complementation of CSS-397; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS0-401	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0159 in CSS-0397		
P$_{cagA}$	26695	*tlpB::P$_{cagA}$-tlpB:rpsL-erm*	*tlpB* under control of the *cagA* (P$_{cagA}$) promoter; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033	CSS-0657	Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP91-3 in CSS-0004		
P$_{cagA}$/Δ*repG*	26695	*tlpB::P$_{cagA}$-tlpB:rpsL-erm, repG::aphA-3*	Deletion of *repG* in CSS-0657; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0658	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0657		
tlpB* 5th::*gfpmut3	G27	*rdxA::tlpB* 5th-*gfpmut3:catGC*	Translational fusion of the *tlpB* 5' UTR including the 5th amino acid to *gfpmut3*; *tlpB* from *H. pylori* 26695 introduced in G27 (CSS-0010); verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-0748	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP109-6 in CSS-0010		
tlpB* 5th::*gfpmut3*/ Δ*repG	G27	*repG::aphA-3, rdxA::tlpB* 5th-*gfpmut3:catGC*	Translation fusion of the *tlpB* 5' UTR including the 5th amino acid to *gfpmut3*; *tlpB* from *H. pylori* 26695 introduced in *H. pylori* G27 Δ*repG* (CSS-0169); verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-0751	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP109-6 in CSS-0169		
cagA* 28th::*gfpmut3	G27	*rdxA::cagA* 28th-*gfpmut3:catGC*	Translation fusion of the *cagA* 5' UTR including the 28th amino acid to *gfpmut3*; *cagA* from *H. pylori* G27 (CSS-0010); verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-0804	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pMA5-2 in CSS-0010		
cagA* 28th::*gfpmut3*/ Δ*repG	G27	*repG::aphA-3, rdxA::cagA* 28th-*gfpmut3:catGC*	Translation fusion of the *cagA* 5' UTR including the 28th amino acid to *gfpmut3*; *cagA* from *H. pylori* 26695 introduced in *H. pylori* G27 Δ*repG* (CSS-0169); verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/ JVO-0155	CSS-0805	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pMA5-2 in CSS-0169		
26695 Str^R	26695 Str^R	*rpsL-str^R*	Streptomycin-resistant 26695 based on CSS-0004; verification by sequencing with JVO-5704	CSS-0024	Str^R	Transformation of PCR product amplified with JVO-5702/5703 on gDNA of CSS-0004 in CSS-0004		
26695^R *tlpB*::3xFLAG	26695 Str^R	*rpsL-str^R, tlpB::tlpB-*3xFLAG:*rpsL-erm*	*H. pylori* carrying *tlpB*::3xFLAG; background strain for markerless exchange; verification by colony-PCR using CSO-0050/0046 on gDNA and sequencing with CSO-0208	CSS-0461	Str^R Erm^R	Transformation of PCR product amplified with CSO-0208/0211 on pSP57-4 in CSS-0024		

Trivial name used in thesis	H. pylori strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
26695^R tlpB::3xFLAG*	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*	Markerless *tlpB*::3xFLAG tagged background strain; verification by colony-PCR using CSO-0050/0046 and/or CSO-0050 x CSONIH-0033 on gDNA and sequencing with CSO-0208	CSS-0464	Str^R	Transformation of PCR product amplified with CSO-0208/0211 on pSP70-1 in CSS-0461		
***tlpB*::3xFLAG*/ *tlpB* ΔG**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* ΔG::*tlpB:rpsL-erm*	Deletion of the G-repeat in the *tlpB* mRNA leader in CSS-0464; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0471	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP64-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 6G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 6G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 6G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0472	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP73-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 7G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 7G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 7G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0473	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP74-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 8G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 8G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 8G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0474	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP75-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 9G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 9G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 9G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0475	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP76-1 in CSS-0464		
***tlpB*-3xFLAG*/ *tlpB* 10G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 10G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 10G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0476	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP77-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 11G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 11G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 11G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0477	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP78-1 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* WT (12G)**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 12G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 12G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0470	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP60-2 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 13G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 13G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 13G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0478	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP79-4 in CSS-0464		
***tlpB*::3xFLAG*/ *tlpB* 14G**	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 14G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 14G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0479	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP80-1 in CSS-0464		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
tlpB::3xFLAG*/ *tlpB* 15G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 15G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 15G; verification by colony-PCR using CSO-0051 and CSONIH-0033 on gDNA and sequencing with CSO-0051/CSO-0277	CSS-0480	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP81-5 in CSS-0464		
tlpB::3xFLAG*/ *tlpB* 16G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *tlpB* 16G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0464, 16G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0481	Str^R Erm^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP82-1 in CSS-0464		
26695^R *tlpB*::3xFLAG*/ Δ*repG*	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*	Deletion of *repG* in CSS-0464, verification by colony-PCR using JVO-5069 and JVO-5257 on gDNA	CSS-0467	Str^R Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0386		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* ΔG	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* ΔG::*tlpB:rpsL-erm*	Deletion of the G-repeat in the *tlpB* mRNA leader in CSS-0467; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0483	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP64-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 6G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 6G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 6G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0484	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP73-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 7G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 7G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 7G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0485	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP74-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 8G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 8G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 8G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0486	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP75-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 9G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 9G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 9G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0487	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP76-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 10G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 10G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 10G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0488	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP77-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* 11G	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 11G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 11G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0489	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP78-1 in CSS-0464		
tlpB::3xFLAG*/ Δ*repG*/*tlpB* WT (12G)	26695 Str^R	*rpsL-str^R*, *tlpB::tlpB*-3xFLAG*, *repG::aphA-3*, *tlpB* 12G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 12G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0482	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP60-2 in CSS-0464		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
***tlpB*::3xFLAG*/ Δ*repG*/*tlpB* 13G**	26695 Str^R	*rpsL-str^R*, *tlpB*::*tlpB*-3xFLAG*, *repG*::*aphA-3*, *tlpB* 13G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 13G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/ CSO-0277	CSS-0490	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP79-4 in CSS-0464		
***tlpB*::3xFLAG*/ Δ*repG*/*tlpB* 14G**	26695 Str^R	*rpsL-str^R*, *tlpB*::*tlpB*-3xFLAG*, *repG*::*aphA-3*, *tlpB* 14G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 14G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0491	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP80-1 in CSS-0464		
***tlpB*::3xFLAG*/ Δ*repG*/*tlpB* 15G**	26695 Str^R	*rpsL-str^R*, *tlpB*::*tlpB*-3xFLAG*, *repG*::*aphA-3*, *tlpB* 15G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 15G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0492	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP81-5 in CSS-0464		
***tlpB*::3xFLAG*/ Δ*repG*/*tlpB* 16G**	26695 Str^R	*rpsL-str^R*, *tlpB*::*tlpB*-3xFLAG*, *repG*::*aphA-3*, *tlpB* 16G::*tlpB:rpsL-erm*	Variation of the G-repeat length in the *tlpB* mRNA leader in CSS-0467, 16G; verification by colony-PCR using CSO-0051/0308 on gDNA and sequencing with CSONIH-0033/CSO-0277	CSS-0493	Str^R Erm^R Kan^R	Transformation of PCR product amplified with CSO-0291/0040 on pSP82-1 in CSS-0464		
ΔHP0102	26695	HP0102::*rpsL-erm*	Deletion of HP0102 in CSS-0004 (26695); verification by colony-PCR using CSO-0051 and CSONIH-33 on gDNA	CSS-1000	Erm^R	Transformation of PCR product amplified with CSO-0870/0872 on pBA7-4 in CSS-0004		
tlpB*-HP0102 10th::*gfpmut3	G27	*rdxA*::*tlpB*-HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon from *H. pylori* 26695 introduced in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2104	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP195-6 in CSS-0010		
tlpB*-HP0102 10th::*gfpmut3*/ Δ*repG	G27	*repG*::*aphA-3*, *rdxA*::*tlpB*-HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon from *H. pylori* 26695 introduced in CSS-0169; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2107	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP195-6 in CSS-0169		
tlpB*$_{mini}$-HP0102 10th::*gfpmut3	G27	*rdxA*::HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; *tlpB* mini gene (21 aa), *tlpB* and HP0102 from *H. pylori* 26695 introduced in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2116	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP197-3 in CSS-0010		
tlpB*$_{mini}$-HP0102 10th::*gfpmut3*/ Δ*repG	G27	*repG*::*aphA-3*, *rdxA*::HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; *tlpB* mini gene (21 aa), *tlpB* and HP0102 from *H. pylori* 26695 introduced in CSS-0169; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2119	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP197-3 in CSS-0169		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
P*tlpB*HP0102 10th::*gfpmut3*	G27	*rdxA::*P*tlpB*HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; under control of the *tlpB* promoter, HP0102 from *H. pylori* 26695 introduced in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2138	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP198-4 in CSS-0010		
P*tlpB*HP0102 10th::*gfpmut3*/ Δ*repG*	G27	*repG::aphA-3, rdxA::*P*tlpB*HP0102 10th-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; under control of the *tlpB* promoter, HP0102 from *H. pylori* 26695 introduced in CSS-0169; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2141	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP198-4 in CSS-0169		
tlpB-HP0102 10th::*gfpmut3* ΔG^{HP0102}	G27	*rdxA::tlpB*-HP0102 10th ΔG^{HP0102}-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon lacking G^{HP0102} from *H. pylori* 26695 introduced in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2150	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP200-2 in CSS-0010		
tlpB-HP0102 10th::*gfpmut3* ΔG^{HP0102}/ Δ*repG*	G27	*repG::aphA-3, rdxA::tlpB*-HP0102 10th ΔG^{HP0102}-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon lacking G^{HP0102} from *H. pylori* 26695 introduced in CSS-0169; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2153	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP200-2 in CSS-0169		
tlpB-HP0102 10th::*gfpmut3* ATTTA^{HP0102}	G27	*rdxA::tlpB*-HP0102 10th ATTTA^{HP0102}-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon carrying ATTTA instead of G^{HP0102} from *H. pylori* 26695 introduced in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/JVO-0155	CSS-2156	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP201-1 in CSS-0010		
tlpB-HP0102 10th::*gfpmut3* ATTTA^{HP0102}/ Δ*repG*	G27	*repG::aphA-3, rdxA::tlpB*-HP0102 10th ATTTA^{HP0102}-*gfpmut3:catGC*	Translation fusion of the HP0102 5' UTR (IGR *tlpB* and HP0102) including the 10th amino acid to *gfpmut3*; including *tlpB*, *tlpB*-HP0102 operon carrying ATTTA instead of G^{HP0102} from *H. pylori* 26695 introduced in CSS-0169; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206/ JVO-0155	CSS-2159	Cm^R Kan^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP201-1 in CSS-0169		
Δ*fur*	26695	*fur::rpsL-erm*	Deletion of *fur* in CSS-0004; overlap-PCR; verification by colony-PCR using CSO-1028 x CSONIH-33 on gDNA	CSS-1779	Erm^R	CSO-1023/1022 on gDNA of CSS-0004	CSO-1024/1026 on gDNA of CSS-0163	CSO-1025/1027 on gDNA of CSS-0004
Δ*fur*/Δ*repG*	26695	*fur::rpsL-erm, repG::aphA-3*	Deletion of *repG* in CSS-1779, verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-1781	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-1779		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
ΔrpsA	26695	*rpsA*::*aphA-3*	Deletion of ribosomal protein S1 (*rpsA*; HP0399) in CSS-0004; kindly provided by R. Rieder and J. Vogel, IMIB, Würzburg, Germany, Ref. (Rieder *et al.*, 2012)	JVS-7060	Kan[R]			
HP0143-OE	26695	*rdxA*::*P_{cagA}*HP1043:*catGC*	Overexpression of HP1043 under control of the *cagA* promoter in CSS-0004; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-2651	Cm[R]	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0214		
tlpB::3xFLAG/ HP1043-OE	26695	*tlpB*::*tlpB*-3xFLAG:*rpsL-erm*, *rdxA*::*P_{cagA}*HP1043:*catGC*	Overexpression of HP1043 under control of the *cagA* promoter in CSS-0190; verification by colony-PCR using CSO-0205/0207 on gDNA, sequencing with CSO-0206	CSS-2653	Erm[R] Cm[R]	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0214		
tlpB::3xFLAG/ HP0143-OE/ ΔrepG	26695	*tlpB*::*tlpB*-3xFLAG:*rpsL-erm*, *rdxA*::*P_{cagA}*HP1043:*catGC*, *repG*::*aphA-3*	Deletion of *repG* in CSS-2653, verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-2655	Erm[R] Cm[R] Kan[R]	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-2653		
HP1334::3x-FLAG	26695	HP1334::HP1334-3xFLAG:*aphA-3*	3xFLAG-tagging of HP1334 in *H. pylori* strain 26695 (CSS-0004), Ref. (Rieder *et al.*, 2012)	JVS-7033	Kan[R]			
WT/26695 (Dagmar Beier)	26695		Wildtype (NCBI Acc-no. NC_000915), kindly provided by Prof. Dr. D. Beier, Biocenter, University of Würzburg, Würzburg, Germany), Ref. (Tomb *et al.*, 1997)	CSS-0038				
ΔarsS	26695	*arsS*::*aphA-3*	Deletion of the sensor kinase *arsS* (HP0165) from the ArsRS two component system, kindly provided by Prof. Dr. D. Beier, Biocenter, University of Würzburg, Würzburg, Germany, Ref. (Pflock *et al.*, 2006)	CSS-0042	Kan[R]			
RepG complementation	26695	*repG*::*aphA-3*, *rdxA*::*repG:catGC*	Complementation of *repG* deficient mutant with RepG in *rdxA* under its native promoter; gDNA of the complementation strain was kindly provided by F. Darfeuille, University of Bordeaux, France, unpublished		Kan[R] Cm[R]			
HP1181::3xFLAG	26695	HP1181::HP1181-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1181 in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0903/CSONIH-0033 on gDNA and sequencing with CSO-0900/CSONIH-0033	CSS-1805	Erm[R]	Transformation of PCR product amplified with CSO-0900/0902 on pSP144-1 in CSS-0004		
HP1181::3xFLAG /ΔrepG	26695	*repG*::*aphA-3*, HP1181::HP1181-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1181 in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0903/CSONIH-0033 on gDNA and sequencing with CSO-0900/CSONIH-0033	CSS-1807	Erm[R] Kan[R]	Transformation of PCR product amplified with CSO-0900/0902 on pSP144-1 in JVS-7014		
HP1055::3xFLAG	26695	HP1055::HP1055-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1055 in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA, sequencing with CSO-0974/CSONIH-0033	CSS-1793	Erm[R]	Transformation of PCR product amplified with CSO-0974/0976 on pSP141-1 in CSS-0004		
HP1055::3xFLAG /ΔrepG	26695	*repG*::*aphA-3*, HP1055::HP1055-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1055 in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0974/CSONIH-0033	CSS-1795	Erm[R] Kan[R]	Transformation of PCR product amplified with CSO-0974/0976 on pSP141-1 in JVS-7014		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
HP1056::3xFLAG	26695	HP1056::HP1056-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1056 in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-1547 and/or CSONIH-0033	CSS-2366	Erm^R	Transformation of PCR product amplified with CSO-0970/1547 on pSP173-6 in CSS-0004		
HP1056::3xFLAG /ΔrepG	26695	*repG*::*aphA-3*, HP1056::HP1056-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1056 in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-1457/CSONIH-0033	CSS-2369	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0970/1547 on pSP173-6 in JVS-7014		
HP1057::3xFLAG	26695	HP1057::HP1057-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1057 in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0957/CSONIH-0033	CSS-1827	Erm^R	Transformation of PCR product amplified with CSO-0957/0959 on pSP154-2 in CSS-0004		
HP1057::3xFLAG /ΔrepG	26695	*repG*::*aphA-3*, HP1057::HP1057-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HP1057 in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0957/CSONIH-0033	CSS-1830	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0957/0959 on pSP154-2 in JVS-7014		
dcuA::3xFLAG	26695	*dcuA*::*dcuA-*3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *dcuA* (HP0724) in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0886/CSONIH-0033 on gDNA and sequencing with CSO-0883/CSONIH-0033	CSS-2665	Erm^R	Transformation of PCR product amplified with CSO-0883/0885 on pSP155-1 in CSS-0004		
dcuA::3xFLAG/ ΔrepG	26695	*repG*::*aphA-3*, *dcuA*::*dcuA-*3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *dcuA* (HP0724) in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0886/CSONIH-0033 on gDNA and sequencing with CSO-0883/CSONIH-0033	CSS-2667	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0883/0885 on pSP155-1 in JVS-7014		
3xFLAG::trx2	26695	*trx2*::3xFLAG-*trx2*:*rpsL-erm*	Construction of N-terminal 3xFLAG-tagging of *trx2* (HP1458) at its native locus in the *H. pylori* chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0894 and CSONIH-0033 on gDNA and sequencing with CSO-1284/1364	CSS-2059	Erm^R	Transformation of PCR product amplified with CSO-1283/1284 on pSP187-8 in CSS-0004		
3xFLAG::trx2/ ΔrepG	26695	*repG*::*aphA-3*, *trx2*::3xFLAG-*trx2*:*rpsL-erm*	Construction of N-terminal 3xFLAG-tagging of *trx2* (HP1458) at its native locus in the *H. pylori* chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0894 and CSONIH-0033 on gDNA and sequencing with CSO-1284/1364	CSS-2062	Erm^R Kan^R	Transformation of PCR product amplified with CSO-1283/1284 on pSP187-8 in JVS-7014		
G27	G27		Wildtype (NCBI Acc-no. NC_011333), kindly provided by T. F. Meyer (MPI-IB, Germany); Ref. (Baltrus *et al.*, 2009)	CSS-0010				
G27 ΔrepG	G27	*repG*::*aphA-3*	Deletion of *repG* in CSS-0010;verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0169	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0010		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
G27 C_{RepG}	G27	*repG*::*aphA-3, rdxA::repG:catGC*	Complementation of Δ*repG* (CSS-0169) with wild-type RepG derived from G27; verification by colony-PCR using CSO-0205/0207 on gDNA and seq. with CSO-0206	CSS-0499	Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP84-1 in CSS-0169		
G27 Δ*tlpB*	G27	*tlpB::rpsL-erm*	Deletion of *tlpB* (G27_95) in CSS-0010; verification by colony-PCR using CSO-0051/CSONIH-0033 on gDNA	CSO-0167	Erm^R	Transformation of PCR product amplified with CSO-0039/0040 on gDNA of CSS-0163 in CSS-0010		
G27 ΔHP0102	G27	HPG27_94::*rpsL-erm*	Deletion of HPG27_94 (HP0102) in CSS-0010; verification by colony-PCR using CSO-0051 and CSONIH-0033 on gDNA	CSS-1007	Erm^R	Transformation of PCR product amplified with CSO-0870/0872 on pBA7-4 in CSS-0010		
G27 *tlpB*::3xFLAG	G27	*tlpB::tlpB*-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *tlpB* in the *H. pylori* G27 (CSS-0010) chromosome; verification by colony-PCR using CSO-0050/0046 on gDNA and sequencing with CSO-0208	CSS-0196	Erm^R	Transformation of PCR product amplified with CSO-0208/0211 on pSP57-4 in CSS-0010		
G27 *tlpB*::3xFLAG/ Δ*repG*	G27	*tlpB::tlpB*-3xFLAG:*rpsL-erm, repG*::*aphA-3*	Deletion of *repG* in CSS-0196; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0197	Erm^R Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0196		
G27 *tlpB*::3xFLAG/ Δ*repG*/C_{RepG}	G27	*tlpB::tlpB*-3xFLAG:*rpsL-erm, repG*::*aphA-3, rdxA::repG:catGC*	Complementation of CSS-0197 with wild-type RepG derived from 26695; verification by colony-PCR using CSO-0205/0207 on gDNA and seq. with CSO-0206	CSS-0283	Erm^R Kan^R Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of CSS-0046 in CSS-0197		
HPG27_385-OE	G27	*rdxA::P*_{cagA}HPG27_385:*catGC*	Overexpression of HPG27_385 (HP1043 homolog, under control of the *cagA* promoter) in CSS-0010; verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-0214	Cm^R	Transformation of PCR product amplified with CSO-0017/0018 on gDNA of *H. pylori* carrying HPG27_385 in *rdxA* locus (Prof. Dr. D. Beier, Biocenter, University of Würzburg, Germany)		
G27 HP1181::3xFLAG	G27	HPG27_1124::HPG27_1124-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_1124 (HP1181) in the chromosome, cloning in CSS-0010; verification by colony-PCR using CSO-0903/CSONIH-0033 on gDNA and sequencing with CSO-0903/CSONIH-0033	CSS-1813	Erm^R	Transformation of PCR product amplified with CSO-0900/0902 on pSP144-1 in CSS-0010		
G27 HP1181::3xFLAG /Δ*repG*	G27	*repG*::*aphA-3,* HPG27_1124::HPG27_1124-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_1124 (HP1181) in the chromosome, cloning in CSS-0169; verification by colony-PCR using CSO-0903/CSONIH-0033 on gDNA and sequencing with CSO-0903/CSONIH-0033	CSS-1815	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0900/0902 on pSP144-1 in CSS-00169		
G27 HP1055::3xFLAG	G27	HPG27_0371::HPG27_0371-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0371 (HP1055) in the chromosome, cloning in CSS-0010; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0974/CSONIH-0033	CSS-1797	Erm^R	Transformation of PCR product amplified with CSO-0974/0976 on pSP142-1 in CSS-0010		
G27 HP1055::3xFLAG /Δ*repG*	G27	*repG*::*aphA-3,* HPG27_0371::HPG27_0371-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0371 (HP1055) in the chromosome, cloning in CSS-0169; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0974/CSONIH-0033	CSS-1799	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0974/0976 on pSP142-1 in CSS-0169		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
G27 HP1056::3xFLAG	G27	HPG27_0372:: HPG27_0372-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0372 (HP1056) in the chromosome, cloning in CSS-0010; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-1547/CSONIH-0033	CSS-2657	Erm^R	Transformation of PCR product amplified with CSO-0970/1547 on pSP173-6 in CSS-0010		
G27 HP1056::3xFLAG /ΔrepG	G27	*repG*::*aphA-3*, HPG27_0372:: HPG27_0372-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0372 (HP1056) in the chromosome, cloning in CSS-0169; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-1457/CSONIH-0033	CSS-2660	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0970/1547 on pSP173-6 in CSS_0169		
G27 HP1057::3xFLAG	G27	HPG27_0373:: HPG27_0373-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0373 (HP1057) in the chromosome, cloning in CSS-0010; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0957/CSONIH-0033	CSS-2661	Erm^R	Transformation of PCR product amplified with CSO-0957/0959 on pSP154-2 in CSS-0010		
G27 HP1057::3xFLAG /ΔrepG	G27	*repG*::*aphA-3*, HPG27_0373:: HPG27_0373-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of HPG27_0373 (HP1057) in the chromosome, cloning in CSS-0169; verification by colony-PCR using CSO-0960/CSONIH-0033 on gDNA and sequencing with CSO-0957/CSONIH-0033	CSS-2663	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0957/0959 on pSP154-2 in CSS-0169		
G27 *dcuA*::3xFLAG	G27	*dcuA*::*dcuA*-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *dcuA* (HPG27_78) in the chromosome, cloning in CSS-0004; verification by colony-PCR using CSO-0886/CSONIH-0033 on gDNA and seq.with CSO-0883/CSONIH-0033	CSS-1064	Erm^R	Transformation of PCR product amplified with CSO-0883/0885 on pSP155-1 in CSS-0004		
G27 *dcuA*::3xFLAG/ ΔrepG	G27	*repG*::*aphA-3*, *dcuA*::*dcuA*-3xFLAG:*rpsL-erm*	Construction of C-terminal 3xFLAG-tagging of *dcuA* (HPG27_78) in the chromosome, cloning in JVS-7014; verification by colony-PCR using CSO-0886/CSONIH-0033 on gDNA and sequencing with CSO-0883/CSONIH-0033	CSS-1066	Erm^R Kan^R	Transformation of PCR product amplified with CSO-0883/0885 on pSP155-1 in JVS-7014		
G27 3xFLAG::*trx2*	G27	*trx2*::3xFLAG-*trx2*: *rpsL-erm*	Construction of N-terminal 3xFLAG-tagging of *trx2* (HPG27_1381) at its native locus in the *H. pylori* chromosome, cloning in CSS-0010; verification by colony-PCR using CSO-0894 and CSONIH-0033 on gDNA and sequencing with CSO-1284/1364	CSS-2031	Erm^R	Transformation of PCR product amplified with CSO-1283/1284 on pSP188-2 in CSS-0010		
G27 3xFLAG::*trx2*/ ΔrepG	G27	*repG*::*aphA-3*, *trx2*::3xFLAG-*trx2*: *rpsL-erm*	Construction of N-terminal 3xFLAG-tagging of *trx2* (HPG27_1381) at its native locus in the *H. pylori* chromosome, cloning in CSS-0169; verification by colony-PCR using CSO-0894 and CSONIH-0033 on gDNA and sequencing with CSO-1284/1364	CSS-2034	Erm^R Kan^R	Transformation of PCR product amplified with CSO-1283/1284 on pSP188-2 in CSS-0169		
J99	J99		Wildtype (NCBI Acc-no. NC_000921), kindly provided by T. F. Meyer (MPI-IB, Germany); Ref. (Alm *et al.*, 1999)	CSS-0001				
J99 ΔrepG	J99	*repG*::*aphA-3*	Deletion of *repG* in CSS-0001; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0732	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0001		

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
J99 ΔHP0102	J99	jhp0094::*rpsL-erm*	Deletion of jhp0094 (HP0102) in CSS-0001; verification by colony-PCR using CSO-0051/CSONIH-0033 on gDNA	CSS-1019	Erm^R	Transformation of PCR product amplified with CSO-0870/0872 on pBA7-4 in CSS-0001		
X47-2AL	X47-2AL		Wildtype (REF 1386083), kindly provided by H. De Reuse, Institut Pasteur, Paris, France, Ref. (Handt *et al.*, 1995)	CSS-0996				
X47-2AL Δ*repG*	X47-2AL	*repG*::*aphA-3*	Deletion of *repG* in CSS-0996; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0997	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0996		
X47-2AL C$_{RepG}$	X47-2AL	*repG*::*aphA-3*, *rdxA*::*repG:aac(3)-IV*	Complementation *H. pylori* X47-2AL Δ*repG* (CSS-0997); verification by colony-PCR using CSO-0205/0207 on gDNA and sequencing with CSO-0206	CSS-1038	Kan^R Gen^R	Transformation of PCR product amplified with CSO-0017/0018 on pBA9-5 in CSS-0997		
X47-2AL Δ*tlpB*	X47-2AL	*tlpB*::*aac(3)-IV*	Deletion of *tlpB* in CSS-0996; verification by colony-PCR using CSO-0051/0293 on gDNA	CSS-1123	Gen^R	Transformation of PCR product amplified with CSO-0039/0040 on pBA13-5 in CSS-0996		
X47-2AL Δ*tlpB*/ Δ*repG*	X47-2AL	*tlpB*::*aac(3)-IV*, *repG*::*aphA-3*	Deletion of *repG* in CSS-1123; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-1769	Kan^R Gen^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-1123		
X47-2AL Δ*tlpB*-HP0102	X47-2AL	*tlpB*-HP0102::*aac(3)-IV*	Deletion of the *tlpB*-HP0102 operon in CSS-0996; verification by colony-PCR with CSO-0051/0293 (gDNA)	CSS-1743	Gen^R	Transformation of PCR product amplified withCSO-0040/1359 on pSP127-3 in CSS-0996		
X47-2AL Δ*tlpB*-HP0102/ Δ*repG*	X47-2AL	*tlpB*-HP0102::*aac(3)-IV*, *repG*::*aphA-3*	Deletion of *repG* in CSS-1743;verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-1773	Kan^R Gen^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-1743		
X47-2AL Δ*tlpB*-HP0102/ *tlpB*-HP0102	X47-2AL	*tlpB*-HP0102::*aac(3)-IV*, *rdxA*::*tlpB*-HP0102:*aphA-3*	Complementation of CSS-1743 with the *tlpB*-HP0102 operon in the *rdxA* locus; verification by colony-PCR using CSO-1813/0207 on gDNA and sequencing with CSO-0206/0581	CSS-2046	Kan^R Gen^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP190-1 in CSS-1743		
X47-2AL Δ*tlpB*-HP0102/ HP0102	X47-2AL	*tlpB*-HP0102::*aac(3)-IV*, *rdxA*::HP0102:*aphA-3*	Complementation of CSS-1743 with HP0102 in the *rdxA* locus; verification by colony-PCR using CSO-1813/0207 on gDNA and sequencing with CSO-0206/0581	CSS-2080	Kan^R Gen^R	Transformation of PCR product amplified withCSO-0017/0018 on pSP192-1 in CSS-1743		
X47-2AL ΔHP0102	X47-2AL	HP0102::*aac(3)-IV*	Deletion of HP0102 in CSS-0996; verification by colony-PCR using CSO-1738/0293 on gDNA	CSS-2019	Gen^R	Transformation of PCR product amplified with CSO-1737/1359 on pSP186-2 in CSS-0996		
X47-2AL ΔHP0102/ Δ*repG*	X47-2AL	HP0102::*aac(3)-IV*, *repG*::*aphA-3*	Deletion of HP0102 in CSS-0997; verification by colony-PCR using CSO-1738 and CSO-293 on gDNA	CSS-2022	Kan^R Gen^R	Transformation of PCR product amplified withCSO-1737/1359 on pSP186-2 in CSS-0997		
X47-2AL ΔHP0102/ HP0102	X47-2AL	HP0102::*aac(3)-IV*, *rdxA*::HP0102:*aphA-3*	Complementation of CSS-2019 with HP0102 in the *rdxA* locus; verification by colony-PCR using CSO-1813/0207 on gDNA and sequencing with CSO-0206/0581	CSS-2087	Kan^R Gen^R	Transformation of PCR product amplified with CSO-0017/0018 on pSP192-1 in CSS-2019		
P12	P12		Wildtype (NCBI Acc-no. NC_011498), kindly provided by T. F. Meyer (MPI-IB, Germany); Ref. (Fischer *et al.*, 2010)	CSS-0003				
PeCan4	PeCan4		Wildtype (NCBI Acc-no. NC_014555), kindly provided by D. E. Berg, Washington University, St. Louis, MO/University of California, San Diego, La Jolla, CA	CSS-0096				
Cuz20	Cuz20		Wildtype (NCBI Acc-no. NC_017358), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA	CSS-0097				

Trivial name used in thesis	*H. pylori* strain	Mutation	Description	Strain number	Marker	Oligos used for upstream region	Oligos for cassette	Oligos used for downstream region
Cuz20 Δ*repG*	Cuz20	*repG*::*aphA-3*	Deletion of *repG* in CSS-0097; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0737	Kan^R	Transformation of PCR product amplified with JVO-507/5072 on gDNA of JVS-7014 in CSS-0097		
Sat464	Sat464		Wildtype (NCBI Acc-no. NC_017359), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA	CSS-0098				
India7	India7		Wildtype (NCBI Acc-no. NC_017372), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA	CSS-0099				
India7 Δ*repG*	India7	*repG*::*aphA-3*	Deletion of *repG* in CSS-0099; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0734	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0099		
India7 Δ*tlpB*	India7	*tlpB*::*rpsL-erm*	Deletion of *tlpB* (HPIN_00495) in CSS-0099; verification by colony-PCR using CSO-0051/CSONIH-0033 on gDNA	CSS-2669	Erm^R	Transformation of PCR product amplified with CSO-0039/0040 on gDNA (CSS-0163) in CSS-0099		
Lithuania75	Lithuania 75		Wildtype (NCBI Acc-no. NC_017362), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA	CSS-0101				
Lithuania Δ*repG*	Lithuania 75	*repG*::*aphA-3*	Deletion of *repG* in CSS-0101; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0736	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0101		
Shi470	Shi470		Wildtype (NCBI Acc-no. NC_010698), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA; Ref. (Kersulyte *et al.*, 2010)	CSS-0173				
Shi470 Δ*repG*	Shi470	*repG*::*aphA-3*	Deletion of *repG* in CSS-0173; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0735	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0173		
SJM180	SJM180		Wildtype (NCBI Acc-no. NC_014560), kindly provided by D. E. Berg, Washington University, St. Louis, MO/ University of California, San Diego, La Jolla, CA	CSS-0174				
B8	B8		Wildtype (NCBI Acc-no. NC_014256), kindly provided by R. Haas, Max-von-Pettenkofer-Institut, Munich, Germany; Ref. (Farnbacher *et al.*, 2010)	CSS-0213				
B8 Δ*repG*	B8	*repG*::*aphA-3*	Deletion of *repG* in CSS-0213; verification by colony-PCR using JVO-5069/5257 on gDNA	CSS-0733	Kan^R	Transformation of PCR product amplified with JVO-5070/5072 on gDNA of JVS-7014 in CSS-0213		
B128 'WT'	B128	pILL2157:*catGC*	Wild-type like *H. pylori* B128, carrying pILL2157; kindly provided by H. De Reuse, Insitut Pasteur, Paris, France, Ref. (Redko *et al.*, 2013)		Cm^R			
B128 P$_i$ΔN-*rnj*	B128	*rnj*::*aphA-3* pHP135:*catGC*	RNase J-deficient mutant; conditional Δ*rnj* mutant; Δ*rnj* containing pHP135 = pILL2157 with N-terminal truncation of *rnj* gene under control of an IPTG-inducible promoter – Pi; kindly provided by H. De Reuse, Insitut Pasteur, Paris, France, Ref. (Redko *et al.*, 2013)		Kan^R Cm^R			
Hmu			Wildtype of *Helicobacter mustelae*, kindly provided by T. F. Meyer (MPI-IB, Germany)	CSS-0007				
Hac			Wildtype of *Helicobacter acinonychis* (NCBI Acc-no. NC_008229), kindly provided by T. F. Meyer (MPI-IB, Germany); Ref. (Eppinger *et al.*, 2006)	CSS-0008				

Trivial name used in thesis	Organism	Description	Strain number	Marker
TOP 10	***E. coli***	*mcr*A F*(mrr-hsd*RMS-*mcr*BC) F8*0lac*Z F M15 F *lac*X74 *deo*R *rec*A1 *ara*D139 F *(ara-leu)*7697 *gal*U *gal*K *rps*L *end*A1 *nup*G from Invitrogen	CSO-0296	
MS2-MBP	***E. coli* BL21 (DE3) 'Rosetta'**	*E. coli* expression strain carrying pMS2-MBP	JVS-3373	Cm^R Amp^R

Table 13.5: Sequences of cDNA peaks used for definition of a RepG binding motif. Using a peak calling algorithm, sequences corresponding to the enriched cDNA peaks in the RNA-seq library of *H. pylori* 26695 Δ*repG* deletion mutant compared to the wildtype were extracted. Only peaks within putative RepG target mRNAs were considered for motif prediction.

No.	Start	End	Strand	Gene ID	Name	Peak_sequence
1	35956	36075	+	HP0036	HP0036	TTTGCTATCAAATGTTAGGGCTAGGGGTAGGGGGGAATGGCTTTATAGAAACGAAATTTATCAAGCGCTTTTTAAACCAGCAAGAGCCTTATTATGGGGATATTGGGGTGCGTTTAGAAG
2	205311	205440	+	HP0199	HP0199	AATGCACCTAAAGCGTTTTCTTTAGAGCATGAGGGAGTGGTTTTAGAGGGCGAAGTTGTGCGGGTGGGGGCGAAATTGTTTCGTTTGGAAGCGTGCCTTAAGGGTGAATTGATGCTTATTTGCGATACAA
3	205971	206080	+	HP0201	HP0201	CGGTTTTAGTGGGGGATAAAGACAAAGCAACCCCTTTTATTTCTAAAGAGTTAGCCAGCAAAGTGGAAATGATCCACACGCAAGATTACATCAAGATGGAAGAAGCCGCC
4	206151	206290	+	HP0201	HP0201	CTTTGATTTCAGCGGGGCATAGCGGAGCGACTATGGGTTTAGCCACCTTGCGTTTAGGGCGTATCAAGGGGGTTGAAAGGCCTGCTATTTGCACTTTGATGCCTAGCGTTGGCAAACGCCCTAGCGTGCTGTTAGACGCA
5	206321	206640	+	HP0201	HP0201	TTGATTGATTTTGCTCTTATGGGGTATGAATACGCTAAAAGCGTGTTGCATTATGATAGCCCTAAGGTGGGTCTTTTGAGTAATGGCGAAGAAGATATTAAAGGGAACATGCTCGTTAAAGAAACGCATAAAATGCTGAAAGCTTATGACTTCTTTTATGGCAATGTGGAGGGGAGCGATATCTTCAAAGGGGTTGTGGATGTGGTAGTTTGCGATGGCTTTATGGGGAATGTGGTCTTAAAGACAGACAGGGGTCGAACTAGCGCAATAGGCTCTATTTTTAAAGATGAAATTAAAAGCTCTTTTAAATCTAAAATGGG
6	231221	231375	+	HP0223	HP0223	AAGCGAATTGGATATTGTTTTGGGTGGGGGGATCGCTAAAGGGGGGCTGTATTTAGTGGGGGGGAGTCCTGGGGTGGGGAAATCCACTCTGCTTTTAAAAGTGGCTTCTGGCTTAGCCAAAAACCAGCAGAAGGTTTTGTATGTGAGCGGGGAAG
7	268576	268715	+	HP0259	*xseA*	GTATCAACACCTTAATGCAAGGGGAGGGCTGCGTTCAAAGCGTGGTGGAAAGCATCGTTTATGCGGATAGTTTTCATGCACACAAAAAACGCTTTTGATGCGATTGTAGTGGCTAGGGGTGGGGGGAGCATGGAGGATTTG
8	320826	320940	+	HP0303	*obgE*	TTATCATCGCTTCGGGTAAGGGGGGGCCTGGAATGGTGAGTTTTAGGCGAGAAAAATTTGTCATCAAAGGAGGCCCTGATGGGGGCGATGGAGGCGATGGAGGCGATGTGTATTT
9	321026	321105	+	HP0303	*obgE*	CGAAATTGCGCGGGCAAAAAGGGCGAAGACAAGATCATTGTCGTGCCACCAGGAACGCAGGTTTTTGTAGGTGATGAGTT
10	321111	321230	+	HP0303	*obgE*	TTGATTTAGTGGAACCTAAAGAAAGGGTGTTAGCCTTAAAAGGGGGCAAGGGGGGGGTTAGGGAATGCACATTTTAAAAGCGCGACTAAACAACAACCCACTTACGCGCAAAAAGGCTTAG
11	325521	325585	+	HP0308	HP0308	TTCGCCCTGGGCTTGGAAATGTCTTTCAATGTGTGGCGTTTAGGAGGGTATGGGGTTTTACTAGG
12	525986	526180	+	HP0499	HP0499	CTAATGGCGTGGGGGGGCGCGCAATGTTACCAACCTTTTAATAAAGAAGGCAATCCTGAAAACCAGTTTCCAGGACAACCTGTAATCGTTAAAGATTATAATGGGCAAAAAGATGTGCGCTGGGGGGGGGTGTCGTTCGGTGAGCGCGGGGCAACGCCCTGTGTTTCGTTTGGTGTGGGAAAAGGGAGGCCTAAAAA
13	531946	531990	+	HP0506	HP0506	TATACTAATAAAATACAGATTTGATTGAAGGCATTAAACAACTGC
14	661256	661275	+	HP0616	HP0616	ACGGGGAGGTTACAGAGATT
15	710716	710805	+	HP0663	HP0663	AGGCTCACGACTTTTGGGGAATCGCATGGGGATGTGATAGGGGGGGGTATTAGACGGCATGCCTAGCGGGATTAAAATAGACTATGCGCTA
16	710911	711135	+	HP0663	HP0663	AGCACAGGGACTCCTATAGGGTTTTTAATCCACAACCAAAGGGCTAGGAGCAAGGATTACGATAACATTAAAAACCTTTTTAGGCCTAGCCATGCGGATTTCACTTATTTTCATAAATACGGCATTAGGGATTTTAGGGGTGGGGGGAGGAGTTCGGCCAGAGAGAGTGCTATAAGAGTGGCTGCTGGGGCGTTTGCTAAAATGCTTTTAAGAGAAATCGGTATT
17	711396	711460	+	HP0663	HP0663	AGCGATGATGGGGCTTAATGGGGTGAAAGCGGTTGAAATAGGCAAGGGGGTAGAAAGCTCTTTAT
18	711481	711590	+	HP0663	HP0663	GATTTAATGGATCAAAAGGGGTTTTTGAGCAATCGTAGCGGAGGGGTTTTAGGGGGCATGAGCAATGGGGAAGAAATCATTGTTAGAGTGCATTTCAAACCCACGCCAAG
19	777586	777770	+	HP0724	*dcuA*	AATTTAAAAAGGGTTAATGGTTGATGCCTTTTTCCAAATTGCAGTGTTACTTTTTTCGCTTTTTTTAGGGGCAAGGCTAGGGGGCTTGGGAGTGGGCTATGCGGGGGGGCTTGGGCGTGCTTATTTTATGCTTATTTTTGGGGCTAAATCCGGGCAAAATCCCTT

No.	Start	End	Strand	Gene ID	Name	Peak_sequence
20	1249521	1249645	+	HP1181	HP1181	TCATCGTTGCGGTTTTTGGGGCTTTTTATTGTTTTGCCGGTCATTAGTTTGTATGCGGATAGTTTCCATTCAAGCAGTCCCTTACTCGTGGGGTTGGCTGTGGGCGGAGCGTATCTTACGCAAAT
21	1249766	1249905	+	HP1181	HP1181	TTGGCTCGTTATAGGGCGCTTCATTCAAGGCATGGGGGCTTTAGGGGGGGTTATTAGTGCGATGGTGGCGGATGAAGTGAAAGAAGAAGAGCGCACCAAAGCCATGGCCATCATGGGAGCGTTTATTTTCATTAGCTTCA
22	1249916	1249965	+	HP1181	HP1181	GGCGATTGGCCCTGGGGTTGTAGCGTTTTTTGGGGGGGGGCAAAATGGCTCT
23	1250496	1250615	+	HP1181	HP1181	ACTTTTGGCTATTTAGGGAGCTTTGTTGGGGGGCGTGAGCGGGGGGGTTGAGCTACCATCATTTAGGCGTTTCTAACACAAGCTTGATCGTTGTAGCTTTAGGGCTTATTTGGGGGCTATCG
24	1364461	1364620	+	HP1288	HP1288	GAGTTTGGCAGATTTTGGGGGGGAGCTGTTGGTGGTGCAATTGGGGGTGGTGTTGGTGGTGCAGTGGGGGGGAGCTGTTGGTGGTCCTGCGGGTGGTTGGGCTGGCAGATTAGTTGGTGGTTCTGTGGGGAGAGAGTTTGGTCGGGAAATAGGCGATAGGG
25	1444921	1445045	+	HP1380	HP1380	CAGGCATTATTGGTTTAGGGCTTATGGGGGGGAGTTTAGGGCTAGCCTTGCAAGAATGGGGGCGTTTTAAAAGCGTTATAGGCTATGATCATAACGCTTTGCATGCTAAATTGGCTTTGACTTTG
26	1538496	1538560	+	HP1467	HP1467	GAGTTATAAGCATTATTTGGGGGCGTTTGGGGGAGCTAGGGGGGCTTTAGTCGTTTATACGGACA
27	1538586	1538625	+	HP1467	HP1467	TTCGGTGGTTTCAGGGGGGCTTAGCAATTAATGGGGGGGTT
28	1657056	1657165	+	HP1577	HP1577	TAGGCTATTCGGCTATGGCAGGAGCGTTAGGGGCTGGAGGTTTGGGGGATTTAGCCATTAGGATTGGCTATCAAAGTTATAGGGGCGATGTGCTTTTTTATGCGGTGGTT
29	46711	46795	-	HP0048	HP0048	CTTTATAGGCATTGTCTGGGATGGGAGTGGGGCTTATGAAAATAAGATTTATGGGCGGAGTGTTTTGTGGGGGATTTGGAACGC
30	108776	108820	-	HP0102	HP0102	GTAATGAGTGAAAAAGATGAGGGCATTTATGACGCTATGAATAAG
31	108846	108885	-	HP0102	HP0102	TAGACGGGGCTAGCACGGATAGCACTTTAGAAATCATTCA
32	110221	110445	-	HP0103	tlpB	GCATCAACGACGATAAGGGCATGATTTATATGGTGGTGGTGGATAAAAACGGGGTGGTATTGTTTGATCCGGTCAATCCTAAAACCGTAGGCCAATCAGGGCTTGACGCTCAGAGCGTTGATGGGGTGTATTATGTTAGGGGGTATTTGGAGGCGGCCAAAAAAGGGGGAGGCTACACTTATTATAAAATGCCTAAATACGATGGAGGCGTACCGGAGAAAAAAT
33	110576	110790	-	HP0103	tlpB	GGGGGGGGGGGGTGCATTTAGAAGCTAAACTCTAAAATTAGGGTTTGACTTAAAAATGATTTATAGGAGATAAATGATGTTTTCTTCAATGTTTGCTTCGTTGGGGACTCGTATCATGCTGGTCGTGTTAGCCGCTCTTTTAGGTTTAGGGGGGCTTTTTATTGGTTTTGTAAAGGTTATGCAAAAAGATGTGTTAGCGCAACTCATGGAGCATTTAGAA
34	221541	221625	-	HP0213	gidA	AAAACCAAAACGGGCGCTTTGGGGAAAACGCTTCTAATTCTTTAGCCTTGAATTTAAGGGAGCTTGGCTTTAAGGTGGATCGCCT
35	221811	222115	-	HP0213	gidA	GAAAGTGATATTTTAGTGGTTGGTGGGGGGCATGCAGGCATTGAAGCGAGCTTGATTGCGGCTAAAATGGGGGCTAGGGTGCATTTAATCACCATGCTCATAGACACGATCGGTTTAGCGAGCTGTAATCCGGCGATTGGGGGCTTGGGTAAAGGGCATTTGACTAAAGAAGTGGATGTTTTAGGGGGGGGCTATGGGGATTATTACGGATCATAGCGGTTTGCAATATCGTGTGTTAAACGCTTCTAAAGGGCCGGCGGTTAGGGGGACTAGAGCGCAAATTGATATGGACACTTACCGCATTTT
36	247701	247735	-	HP0238	HP0238	TTGGGGCGAAGATGAGGGATTTTGAATTGATTGGG
37	248056	248185	-	HP0238	HP0238	ATGAAAAGGATTTTCATGCGGTGGGCGTGGATTTAAAAGGGTTTGAAAATCTTGTTTATGCGGATATTGTCCAGGTTAAAGAGAGCGATCGTTGCCCTAATTGTCAAGGAGCGTTGAAATACCATAAGAG
38	727441	727480	-	HP0677	HP0677	TCACTTTAGGGATTGGTGGGGGGGATGCTCATGGTGCCTTT
39	727661	727720	-	HP0677	HP0677	ATTTTTCTTTAGGCTTGTTGATAGGGGCAGGGGGGCTGATAGGGGCGAGTTTTAGCGGAT
40	727761	727890	-	HP0677	HP0677	TTTTACTGGCATTCTATCAGGGATTTTTGGCATTGGTGGGGGGTTGATCATTGTCCCTATCATGCTCGCAACCGGGCATTCTTTTGAAGAATCCATTGGGATTTCCATTTTGCAAATGGCGCTTTCATCG
41	772756	772870	-	HP0718	HP0718	ATTTCTTTGTGTGCGGCGGTGGGGGCGCAATCCTTGTTTATTGTGGAAAGGGGGATGGCTAGGAATTATGTGTTTTTGATTTGCGCTTTGTGCTTTATGTGCGATATTGTGCTAA

No.	Start	End	Strand	Gene ID	Name	Peak_sequence
42	830756	830960	-	HP0777	*pyrH*	TGGGGAAGCGTTAGCTGGGGACAACCAGTTTGGGATTGACATTCATGTGTTAGATCACATCGCTAAAGAGATCAAAAGTTTAGTGGAAAACGATATTGAAGTGGGTATTGTGATTGGTGGAGGCAATATTATTAGGGGGGGTTAGCGCGGCTCAAGGGGGGATTATTAGGCGCACCAGTGGGGATTATATGGGCATGTTAGCCACC
43	881016	881080	-	HP0829	HP0829	TAAAAGCTATAGGGGCATGGGCAGCATTGGGGCTATGACTAAAGGGAGCTCTGATAGGTATTTTC
44	881191	881370	-	HP0829	HP0829	GATTTGATTAGCGCGGGAGCAGACGCTATTAAAGTGGGTATTGGGCCAGGAAGCATTTGCACCACTAGGATTGTGGCTGGGGTGGGAATGCCCCAAGTGAGCGCGATTGATAATTGCGTAGAAGTGGCGTCTAAATTTGATATTCCTGTGATTGCAGATGGAGGGATCCGCTATTCAGGC
45	881461	881610	-	HP0829	HP0829	AAACGCATTGAATACCCTGAGGCCAATAAAGATGATTTTGGGAGGTTGAGAGTGGGGCGGCTATTGGAGTGGGGCAGTTGGATAGGGCTGAGATGTTAGTTAAAGCGGGGGTGGATGCACTGGTGCTAGACAGCGCACATGGGCATTCA
46	976491	976600	-	HP0919	*carB*	TGGCGTGGATACGGGCGGGAGTAATGTGCAATTTGCGATCCACCCCGAGACTTTAAGAATGGTCGTGATTGAAATGAACCCACGGGTGAGCCGCAGCTCCGCATTAGCTT
47	976796	976950	-	HP0919	*carB*	ATTAATGAAATTGGCTTTCCAGCCATTATCAGAGCGAGTTTCACTCTGGCTGGGGGGGGGAGTGGGGTCGCTTACAATATTGAAGAGTTTCAAGAATTGGCTAAAAACGCCCTGGACGCTTCGCCCATTAATGAAATTTTGATTGAAGAGTCCTT
48	976981	977005	-	HP0919	*carB*	GATGGATTTGCCTAAAGGGCGTTAT
49	977061	977120	-	HP0919	*carB*	TGCAAATGCACCAAAAGGGCATGTTAGAAGGCGTGGAGCTTTTGGGGGCTAAGATTGAAG
50	1010831	1010895	-	HP0950	HP0950	TATCGCAGAGCCAGGGGCGATGATAGGCTTTGCGGGGCCTAGGGTGATTAAGCAAACTATAGGGG
51	1014656	1014720	-	HP0955	HP0955	TTGTAGGCATTCGTGGGATGAGCTATCATGGGGGGGTTGGTGGGGTTTTTGATCGCTTCGTATCTT
52	1014721	1014840	-	HP0955	HP0955	CGGAGCTTGGCATTGTGCTAGGGGCAAGGATAGGATACATTCTTATTTATGAGCCTAATTCTGGCTATTATTTGACGCATTTTTGGCAAATCTTTAACCCTTTTGATAGCCATGGGAATT
53	1019551	1019785	-	HP0961	*gpsA*	TAGCCATTGCTGGGGGGGTTTGTGATGGCTTGAAATTAGGCAATAGCGCTAAAGCGAGTTTATTGTCTAGAGGTTTGGTGGAAATGCAACGCTTTGGGGCGTTCTTTGGGGGCAAGACAGAGACTTTTTTAGGGCTTTCTGGGGCTGGGGATTTGTTTTTAACCGCTAATTCTATTTTATCTAGGAATTATCGTGTGGGTTTAGGGCTAGCCCAAAACAAGCCTTTAGAGGTGGT
54	1020196	1020310	-	HP0961	*gpsA*	TATTTGGTGGCGGGGCGTGGGGGAGGGCTTTAGCCTTTGCTTTTGGAGAAAAGAATGAAGTCAAAATCATTTCAAGGCGGGATTTAAACGAGCCGTTAAAAAAGCTCAATGACGC
55	1086276	1086410	-	HP1023	HP1023	AAGGGTGCGTTTAGACGCGATGGTGGCTTTTGGGGATTACCAAGAATATTCTTTAATGAGCAGTTTTAGGGTTTGGACTTATAGGAGTTTGTCTTTTGATATGGGTGGGGGGTATGTGTATGCTTACAATTCTAA
56	1117641	1117945	-	HP1055	HP1055	CGATGATAAGAGTGGGGTTTTTTTAGGGGGTGGGTATGCTTATGGGGAACTTAACTTGTCTTATCAAGGGGAAATGTTAGACAGATACGGCGCGAATGCCCCTAGCGCGTTTAAAAACAATATCAATATTAACGCTCCTGTTTCTATGATTAGCGCTAAATTTGGGTATCAAAAATACTTTGTGTCTTATTTTGGGACACGATTTTATGGGGATTTATTGCTTGGGGGTGGGGCATTAAAAGAGGATGCAATCAAGCAGCCTGTAGGCTCGTTTATTTATGTTTTAGGGGCTGTCAATACCGATT
57	1118356	1118465	-	HP1056	HP1056	AGGTTTGCGTTAGGGATATTTGGAGGCGTTGGAGTGGGGTGGAATGGGATGTATCAAAATTTAAAAGAGGTTAAAGGGTATTCACAGCCTAACGCTTTTGGATTAGTGCT
58	1118486	1118595	-	HP1056	HP1056	CGCCTTACGCTTTTATGGGGAGTATTTAGGGGGGGGCGATGAAAGGATTTAAAAGCGATTCTTTAGCCTCTTATCAAACCGCAAGCTTGAACATTGATTTGTTGATGGATA
59	1119056	1119085	-	HP1056	HP1056	TTGGTGGGGAGCTATGGCAAACATTGGGTA
60	1119181	1119330	-	HP1057	HP1057	GTTATGCTGAATGGGGATTTTTTATTGGATTTTCCTTTGCCTTTTGTGGGGAAATACCTTTATATGGGGGGGTTATATGGGTTTAGGTTTGGGGGTTGTAGCGCATGGGGTGAATTACACGGCGGAATGGGGGATGTCTTTTAACGCAGGA
61	1194136	1194250	-	HP1131	*atpC*	AAAATTAGTGTGGTAGTTCCTGAGGGGGAAGTCTATACAGGAGAGGTTAAAAGCGTTGTGTTGCCAGGAGTTGAAGGGGAATTTGGGGTGCTTTATGGGCATAGCAACATGATCA

No.	Start	End	Strand	Gene ID	Name	Peak_sequence
62	1195111	1195220	-	HP1132	HP1132	GGTGGGGCTGGCGTAGGCAAAACGGTGATCATTATGGAGCTTATCCATAATGTGGCTTATAAGCATAACGGGTATTCGGTGT TTGCAGGTGTGGGGGAGCGCACCAGAGA
63	1383826	1383960	-	HP1323	*rnhB*	ATGAAACATGTGAGTAGGGATTTTGATACCGGTTGGGTTGCGTATCAATGACTCTAGGCATTGATGAAGCGGGTAGGGGGTG TTTGGCCGGTTCGCTTTTTGTGGCTGGGGTGGCGTGTAATGAAAAAACAGCCT
64	1458931	1459030	-	HP1398	HP1398	TAAACGGGGCGGTGGGTCTAGTGGGAGCGGTGTTAGTTACAGCGAGCCAAACCCCTAAAGTGATCTTAAGAAAGCATTTAAA ATACATGCAAAAACAAGG
65	1459116	1459300	-	HP1398	HP1398	TAGTCCAGCATTATTTACTGGCTTTAGAGCATGTTAAGGGGTTTATGGGTAAGGGGGTCATGCTAGGGGGTTTATCGGGTGC GCAAAGGGCTAAAATTGTCGTAGTTGGGGGCGGTGTGGTTGGCATGGAGAGCGCGAAAGTCTTAAGCCAAATGGGGGCTAAA GTAACGATTTTAGAATTAGAC
66	1539116	1539385	-	HP1468	HP1468	GTAGGGGATAATTTGGGGGTGAAGCCGGCTAATGAATACCTTTTTATCGTGTTTTGTGCGCCTGTGGGGGCGTATTTTAAGG GGGGTATAGAAAAAGGGGGGGGCTAGGTTTATCACTACGATTTTTGATAGGGCCGCGCCTAAAGGCACCGGTGGGGTGAAAGT GGGAGGGAATTACGCTGCAAGCCTGTTAGCCCATAAAATGGCCACAGAGCAAGGCTATGATGATTGCATTTATTTAGACCCT ACTACGCACACTAAAATTGAAGAA
67	1629516	1629800	-	HP1550	*secD*	CCAAGTTTAGGGAAAGACAGCGTTAAAACTTCCATTATCGCTCTAGTTGGGGGCTTTATTTTAGTGATGGGCTTTATGGTGC TTTATTACTCTATGGCGGGGGGTGATCGCTTGTTTGGCGTTAGTGGTCAATCTTTTTTTTGATTGTGGCGGTCATGGCGATTTT TGGAGCGACGCTGACTTTACCGGGAATGGCGGGGGATTGTTTTAACCGTGGGGATTGCCGTGGATGCTAATATCATCATCAAC GAGCGCATTAGAGAAGCTCTTAAGAGAGAATGAGGGCATC
68	1629811	1630035	-	HP1550	*secD*	TGCGCAAGGGGCTAAGATTTTTGGGGATTTCTCAGGTGCGAATGTGGGCAAACGCATGGCGATTGTTTTAGACAATAAGGTC TATTCAGCCCCGGTGATTAGGGAGCGTATCGGTGGGGGGGAGCGGGCAGATTAGCGGGAATTTTAGCGTGGCTCAAGCGAGCG ATTTAGCGATCGCTTTAAGGAGTGGGGCGATGAGCGCTCCCATTCAGGTTTTAGAAAAAAG
69	1630041	1630155	-	HP1550	*secD*	TGTGGAAATGGGGGTAAAATCTTGCTCAAAGCGATCCCCATTTTAGATGGCGAAATGCTTACAGATGCGAAAGTGGTGTAT GACCAAAACAACCAGCCGGTGGTGAGCTTCACG
70	1630666	1630800	-	HP1550	*secD*	TATTGGCGCGCTTCTTTTAGGGGTAGGGTTTTCTGTGCCTTCTTTACTAGAAACTAAAGGCCCTAAAATCACTTTAGGTTTG GATTTAAGGGGGGGGGTTGAACATGCTTTTAGGGGTACAAACCGATGAGGCTTT