Investigating the mechanism of a novel glycine-dependent riboswitch and a putative non-coding regulatory RNA in *Streptococcus pyogenes*

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Rostock

vorgelegt von

Afsaneh Khani

aus dem Iran

geboren am 01.05.1980 in Shahrood

Gutachter:

Prof. Dr. Bernd Kreikemeyer

Universität Rostock

Universitätsmedizin Rostock

Institut für Medizinische Mikrobiologie, Virologie und Hygiene

Schillingallee 70

18057 Rostock

Prof. Dr. Hubert Bahl

Universität Rostock

Institut für Biowissenschaften

Abteilung Mikrobiologie

Albert-Einstein-Str. 3

18051 Rostock

Table of content

1   Introduction

1.1   Taxonomy of *Streptococcus pyogenes* (*S. pyogenes*)

S. pyogenes belongs to the genus *Streptococcus,* the family *Streptococcaceae* and to the order *Lactobacillales*. The order *Lactobacillales* belongs to the class Bacilli and to the phylum Firmicutes. Firmicutes included all gram-positive bacteria. Gram-positive bacteria are divided into two separate phyla: the Firmicutes (low G+C gram-positive bacteria) and the Actinobacteria (high G+C gram-positive bacteria) (Rocha, 2002; Patterson., 1996). Several members of the genus *Streptococcus* are pathogenic for humans and animals. Some species are proven or suspected to be zoonotic. Several members are commensal species. *S. pyogenes* is a human pathogen and has a worldwide distribution (Fulde & Valentin Weigand, 2013).

1.2   General properties of *Streptococcus pyogenes*

S. pyogenes is a fermentative, aerotolerant, catalase negative, beta hemolytic, nonmotile, non-spore-forming gram-positive coccus. *S. pyogenes* has a hyaluronic acid capsule. It occurs in chains or pairs (Patterson., 1996). It grows under aerobic and anaerobic conditions. *S. pyogenes* is a lactic acid bacterium and generates energy through acid-generating fermentation pathways. It converts carbohydrates to lactic acid by fermentation (Murdoch, 1998; Steele, et al., 1953; Pancholi & Caparon, 2016). *Streptococcus* is considered as a low G+C content genus (Gao, et al., 2014; Rocha, 2002).

1.3   Classification

Streptococci are classified based on hemolysis into three groups according to their appearance on blood agar: α-hemolytic, β-hemolytic and γ-hemolytic. α -hemolysis, which is observed as green zones around colonies, is caused by oxidation of the hemoglobin by bacterially produced H_2O_2. β -hemolysis appears as clear zones surrounding the colonies, and is due to the lysis of red blood cells by secreted hemolysin. Streptococci that display no zones or haloes around their colonies are called γ-hemolytic (Becker, 1916). They are also classified based on serologic specificity. In the early 20th century, Rebecca Lancefield described a serological classification of streptococci based on the carbohydrate composition of bacterial antigens found in the cell walls of β-hemolytic streptococci (Lancefield, 1933a)**.** The Lancefield serological grouping system for identification of streptococci is based on the immunological differences in their cell wall polysaccharides (group A, B, C, F and G) or lipoteichoic acids (group D)

(Cunningham, 2000). *S. pyogenes* is known as group <u>A</u> <u>s</u>treptococci (GAS) and appears as beta-hemolytic on blood agar.

S. pyogenes is further subdivided into different serotypes based on the highly variable N-terminal sequence of the cell surface M protein as described by Rebecca Lancefield (Lancefield, 1962b). Currently, there are over 190 types of M-protein genes (*emm*) identified in *S. pyogenes* (Maripuu, et al., 2008). This typing system is based on a sequence at the 5' end of *emm* that is present in all isolates. The targeted region of *emm* displays the high level of sequence polymorphism. There are four main subfamilies of *emm* genes which are distinguished by sequence variations in the 3´end, encoding the peptidoglycan-spanning domain. *emm* subfamily genes are arranged in the chromosome in 5 different patterns from A to E (McGregor & Spratt, 2004).

1.4   Clinical manifestation

Group A streptococci colonize the skin or throat and lead to several suppurative infections and non-suppurative sequelae. They have developed different virulence mechanisms to avoid host immune system (Cunningham, 2000). Acute *Streptococcus pyogenes* infections may take the form of pharyngitis, scarlet fever (rash), impetigo, cellulitis, or erysipelas. Invasive infections can result in life-threatening diseases like necrotizing fasciitis, myositis, bacteremia and streptococcal toxic shock syndrome (Cunningham, 2008; Stevens, 2000). Rheumatic fever, Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcal infections (PANDAS) and glomerulonephritis are autoimmune post infectious sequelae of *S. pyogenes* infections (Benedek, 2006; Cunningham, 2016; Swedo, et al., 1998). *S. pyogenes* is sensitive to β-lactam antibiotics. The number of strains resistant to macrolide antibiotics is rising (Logan, et al., 2012). Macrolides, chloramphenicol, teicoplanin, vancomycin, and levofloxacin are considered as alternative therapies of *S. pyogenes* infections (Camara, et al., 2013).

1.5   Virulence factors of *S. pyogenes*

Different virulence factors of group A streptococci have been studied. Streptococcal adhesins such as M protein and lipoteichoic acid are important factors which mediate attachment to the host epithelial cell (Courtney & Podbielski, 2009; Ryan & Juncosa, 2016). The M protein and the hyaluronic acid capsule are factors that inhibit phagocytosis (Fischetti, 2016). All *S. pyogenes* infections are established firstly by attachment of the bacterial organism to human epithelial cells, including those in the oral and nasal cavities

and the skin. Lipoteichoic acid (LTA) was the first streptococcal adhesin to be identified that reacted with fibronectin on the host cell (Hanski , et al., 1992). At least 24 other adhesins have been identified after LTA such as M protein, SpeB, FbaA, FbaB, protein F1 and protein F2, Cpa, Lmb, Sfbx, which bind to human extracellular matrix proteins, including collagen, fibronectin (Fn), and laminin (Courtney & Podbielski, 2009; Fischetti, 2016; Walker, et al., 2014; Fiedler, et al., 2010). Oligocomponent pilus structures represent a class of virulence factors with adhesive and matrix protein-binding activity. These extended surface structures play key roles in host cell and tissue adherence in *S. pyogenes* (Kreikemeyer, et al., 2011c). Streptodornase (DNase), streptokinase (plasminogen activation), and streptolysine S biosynthesis proteins (hemolysis) have been proved as virulence factors in *S. pyogenes* (Patterson., 1996; Winter & Berwheimer, 1963; Fiedler, et al., 2010).

Once streptococci have attached to the epithelium, different events can occur: in one pathway, the activation of nonspecific, innate, and specific host defense mechanisms or potential induction of bacterial "apoptosis" will happen. These phenomena lead to the limitation of infection and clearance of the bacteria from host cells. The other pathway is to prevent the primary immune responses of host and a successful colonization of the host's tissues (Courtney & Podbielski, 2009). Extracellular products, such as pyrogenic (erythrogenic) toxins are essential factors to destroy host cells and to colonize suitable sites in the host, resulting in the spreading of the bacteria to a new site in the host. Subsequently streptococci can proliferate and spread to adjacent host tissues. *S. pyogenes* possesses many virulence factors, which are mediators of infection and survival within the host (Figure 1-1), (Courtney & Podbielski, 2009; Podbielski & Kreikemeyer, 2004b).
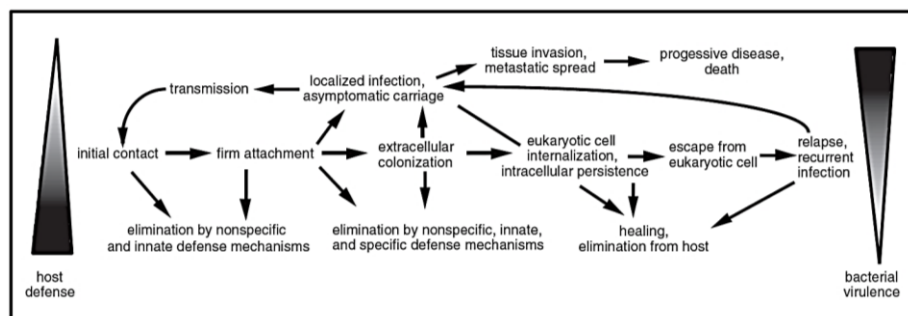
Figure 1-1. Model of the impact of adhesion, invasion, and intracellular persistence on the pathogenesis of streptococcal infections (Courtney & Podbielski, 2009).

1.6 Regulation of gene expression

During the infection process, expression of pathogenicity factors is precisely controlled by a network of protein and RNA regulators. Precise regulation is essential for responding to the host conditions for avoidance of the host immune system. These responses are mostly provided by two main types of regulators: approximately 30 stand-alone response regulators (RRs) and up to 13 two-component signal transduction systems (TCSs) (Fiedler, et al., 2010; Parkinson & Kofoid, 1992; Kreikemeyer, et al., 2003b).

During the first decade of 21th century, RNAs were discovered as important regulatory factors in eukaryotes as well as in prokaryotes. Among them, miRNAs, sRNAs and riboswitches have been investigated in all three domains of life (Patenge, et al., 2015). RNA-mediated regulation confers the ability of fast adaptation of cell growth in response to environmental stress. Bacterial sRNAs are composed of 30 to 500 nucleotides and influence gene expression (inhibition or activation) in response to environmental signals through different mechanisms at both transcriptional and translational level (Landt, et al., 2008).

Bacteria produce three types of non-coding regulatory RNAs:

(A) *Cis*-encoded antisense RNAs; a c*is*-encoded antisense RNA (as RNA) has only one target mRNA. It is encoded on the opposite strand of the target mRNA and its sequence is complementary to the respective target mRNA. Naturally, *cis*-acting antisense RNAs pair to the target RNA at specific regions of high complementarity over a small region of 6 to 8 nt to control biological function by regulating gene expression at the post-transcriptional level. Specific and fast recognition is attained with a short binding site. An asRNA-mediated regulation generally inhibits mRNA transcription and/or translation or induces their rapid mRNA degradation. In fewer cases, asRNAs activate expression of mRNAs (Cho & Kim, 2015) (Figure 1-2.A).

(B) *Trans*-acting small non-coding RNAs; *trans*-acting small non-coding RNAs (sRNAs) have multiple target mRNAs. They are encoded in intergenic regions on the chromosome in a distinct locus from the target mRNAs. As an outcome, its sequence is similar to targets but not identical and can pair to more than one target. After binding to target it controls translation or degradation of target mRNAs (Cho & Kim, 2015) (Figure 1-2.B).
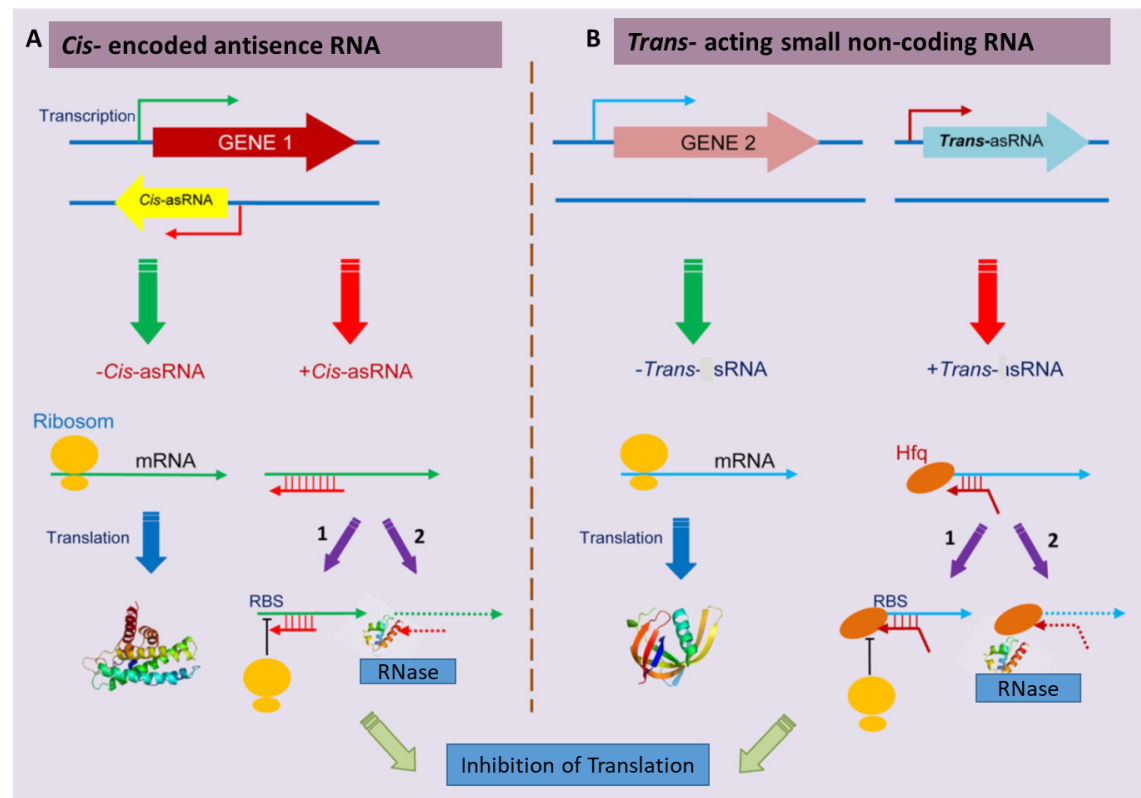
Figure 1-2. Schematic representation of gene regulation at translational level by *cis*-encoded antisense RNA and *Trans*-acting small non-coding in bacteria. A) *Cis*- encoded asRNA has a high complementarity to the target mRNA. B) *Trans*- acting small non-coding RNAs are encoded on a locus distinct from their target mRNAs. They are similar to their targets but not identical. In some cases, *trans*-acting small non-coding RNAs need Hfq to facilitate binding to the targets (Saberi, et al., 2016).

(C) *Cis*-acting 5′ element non-coding RNAs; since ribosomes can bind mRNA before the transcription of DNA to mRNA is complete, most *cis*-regulatory signals are found within the 5′ untranslated region (UTR) of mRNAs. These regulatory regions contain structures which sense environmental signals (riboswitches or thermosensors) or are potential binding sites for *trans*-acting regulators (non-coding RNAs or proteins). In riboswitches, binding of small metabolites leads to a structural change that influences transcription or translation of the downstream gene(s) in an operon (Cho & Kim, 2015) .

In comparison to transcriptional regulation mediated by protein–DNA interaction and to post-translational regulation exerted by protein–protein interaction, regulation by non-coding RNAs is profitable when fast responses to environmental signals are needed (Shimoni, et al., 2007). As non-coding RNAs are not translated, it is beneficial to cells to use sRNA-mediated regulation, so the energetic cost of their synthesis is lower than synthesis of regulatory proteins (Altuvia & Wagner, 2000). Regulation by non-coding RNAs at the translational level results in a quicker effect than at the transcriptional level since mRNA is the direct target (Cho & Kim, 2015; Sharma, et al., 2007). Some sRNA

molecules bind to proteins and affect transcription or translation. For example, CsrB or RsmA sRNA family inhibit RNA binding regulatory proteins (Babitzke & Romeo, 2007). The Csr system also modulates gene expression post-transcriptionally by affecting mRNA stability in *E. coli* (Esquerré, et al., 2016). sRNAs can protect bacteria against foreign and invasive DNAs and phages as a part of the CRISPR system which is known as immune system for bacteria (Jiang & Doudna, 2015).

1.6.1   Regulation by sRNAs in *S. pyogenes*

A few sRNAs with regulatory functions have been identified and described in *S. pyogenes*. Pel/sagA RNA (pleiotropic effect locus/streptolysin-associated gene A) and RivX RNA and FasX (fibronectin/fibrinogen-binding/hemolytic-activity/streptokinase-regulator-X) have been described and shown to play critical roles in the expression of virulence factors in *S. pyogenes* (Kreikemeyer, et al., 2001a; Mangold, et al., 2004; Ramirez-Peña, et al., 2010; Le Rhun, et al., 2016; Roberts & Scott, 2007).

Pleiotropic effect locus (*pel*) which contains *sagA,* the structural gene for streptolysin S, is an effector of virulence factor expression in *S. pyogenes*. Regulation by pel RNA occurs at both transcriptional (*emm, sic, nga*) and post-transcriptional (SpeB) levels (Li, et al., 1999; Mangold, et al., 2004).

RivX is an sRNA that has been studied in the M1T1 serotype. RivX is putatively involved in the regulation of the Mga (multiple gene activator) regulon that controls the expression of a number of virulence-associated factors (Roberts & Scott, 2007; Le Rhun, et al., 2016).

FasBCA is a growth phase-associated two component system (fibronecin/fibrinogen binding/hemolytic activity/streptokinase regulator) that has been studied in detail in *S. pyogenes*. The *fas* operon is transcribed as polycistronic message (*fasBCA*) and contains genes encoding two potential histidine protein kinases (FasB and FasC) and one response regulator (FasA). Downstream of *fasBCA*, a 300 nucleotide monocistronic transcript (*fasX*) is located that does not encode any peptide sequences (Kreikemeyer, et al., 2001a). FasX affects virulence factor expression using at least three distinct known mechanisms: (A) FasX positively regulates streptokinase gene expression by stabilizing the *ska* mRNA (Ramirez-Peña, et al., 2010). (B) FasX negatively regulates pilus expression and adherence in *S. pyogenes*. It base pairs to the 5′ end of mRNA of the pilus biosynthesis operon. RNA:RNA interaction destabilizes the mRNA and subsequently reduces the

ability of adherence to human keratinocytes (Liu, et al., 2012). (C) FasX also negatively controls *cpa* mRNA encoding a pilin protein and inhibits the translation of *cpa* and leads to the negative regulation of pilus expression (Danger, et al., 2015; Liu, et al., 2012).

1.6.2 Riboswitches as regulatory elements

Riboswitches were first discovered in 2002, when the first RNA-based intracellular sensors of vitamins were found (Mironov, et al., 2002; Nahvi, et al., 2002; Winkler, et al., 2004). They are present in all three domains of life. Riboswitches are regulatory non-coding RNA structures and are mostly found within the 5′ untranslated regions of mRNAs. They play a crucial role in adaptive responses through sensing environmental signals. Known riboswitches sense different biological signals such as amino acids and their derivatives, carbohydrates, coenzymes, thiamin pyrophosphate, nucleobases and their derivatives (Serganov & Nudler, 2013b), inorganic ligands including metals (Mg^{2+} cations) (Cromie, et al., 2006), and anions (Baker, et al., 2012).

1.6.3 Structure and function of riboswitches

Most riboswitches are composed of two distinct functional domains, the aptamer(s) and the expression platform. The aptamer is the sensor domain of the riboswitch and very conserved in sequence and structure across bacterial species, because it is involved in the recognition of the signal (Serganov & Nudler, 2013b). Riboswitch classes have evolved by alteration of their metabolite specificity, accommodating a different ligand or signaling compound. Different variants are known for riboswitch classes that recognize different signaling molecules (Weinberg, et al., 2017).

The second domain of a riboswitch is an expression platform that is located immediately downstream of the aptamer domain (Figure 1-3). Ligand binding to the aptamer domain induces a conformational change in an expression platform, resulting in transcriptional or translational regulation. Regulation sometimes leads to RNA processing or affects mRNA stability (Serganov & Nudler, 2013b).

Figure 1-3. Schematic representation of the arrangement of aptamer domain and expression platform in a riboswitch in an mRNA transcript. The aptamer domain and the expression platform overlap through a sequence that can base pair with either domain (yellow bar).

1.6.3.1 Transcriptionally acting riboswitch

The presence of a signal either induces a Rho-independent terminator or exposes Rho binding site and leads to premature transcription termination. Consequently, synthesis of the full-length mRNA will be inhibited (Serganov & Nudler, 2013b). If formation of an anti-terminator hairpin is promoted, RNA polymerase completes transcription of the gene. The Gly riboswitch in *B. subtilis* is one example of this type (Phan & Schumann, 2007).

1.6.3.2 Translationally acting riboswitch

Signal detection by the sensor domain of the riboswitch leads to either sequestering of the ribosome binding site (RBS or Shine-Dalgarno sequence: SD) and subsequently repression of ORF expression or facilitating the access of ribosome to the RBS by formation of SD antisequester hairpin and translation initiation (Serganov & Nudler, 2013b).

1.6.3.3 Induction of mRNA degradation

mRNA degradation as a mechanism for riboswitch function was observed in gram-positive bacteria in 3 different fashions.

(A) Ligand binding promotes cleavage by riboswitch-ribozyme. An open 5′-OH is created, which stimulates RNase J, and leads to degradation of mRNA (Bastet, et al., 2011).

(B) Self-splicing and simultaneous translation activation was determined as a riboswitch mechanism in *C. difficile*. Self-splicing occurs in the presence of ligand. Somehow the complete SD will be accessible by the ribosome and translation will be proceed. In the absence of ligand, SD can be removed from the mRNA resulting in repression of translation initiation (Chen, et al., 2011).

(C) In the absence of the metabolite, an anti-terminator structure forms and transcription proceeds. In the presence of a ligand, formation of a terminator structure inhibits transcription initiation. There is one hypothesis that truncated RNA will bind to the 5′-UTR of the corresponding mRNA and RNA:RNA interaction will reduce RNA stability. This kind of modulation was observed in *L. monocytogenes* (Serganov & Nudler, 2013b).

1.7 Glycine (Gly) riboswitch

Gly riboswitches commonly consist of two adjacent homologous aptamers, linked by a short linker region preceding Gly catabolism and efflux genes in different groups of bacteria (Barrick, et al., 2004; Mandal, et al., 2004; Ruff & Strobel, 2014; Phan & Schumann, 2007). They regulate the expression of downstream genes that are involved

in transport or degradation of Gly to balance the concentration of Gly in bacterial cells (Kazanov, et al., 2007). Studies have shown that two aptamers function cooperatively. It means Gly binding to one aptamer increases the affinity for Gly of the other aptamer (Mandal, et al., 2004; Butler, et al., 2011). However, a full-length derivative of the Gly riboswitch containing its extended 5′ leader, did not show cooperative binding (Sherman, et al., 2012).

In Gly riboswitches ligand binds to both aptamers in ligand-binding pockets which have nearly identical and conserved sequence features. Nucleotide positions near the ligand binding pocket at positions 32, 35, and 69 are a base for classifying Gly riboswitches to different variant groups. In one variant key nucleotides are G, G, U. The other variant is D, G, A that D identify every nucleotide but not C (Weinberg, et al., 2017) (Figure 1-4).
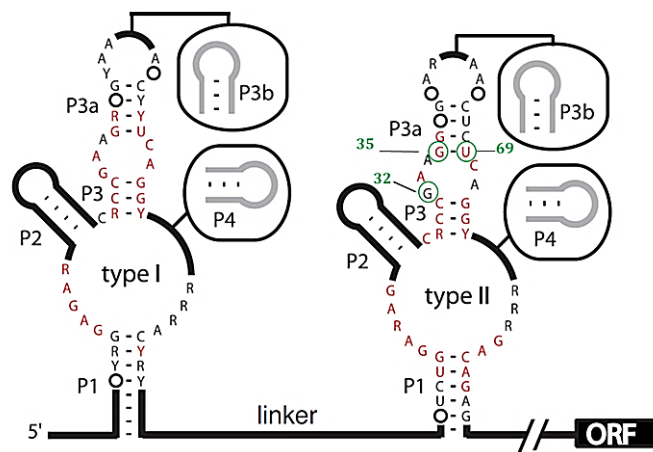
Figure 1-4. Consensus nucleotides in the Gly riboswitch *in B. subtilis*. Circles and thick lines in black show not conserved sequences. P1 through P4 identify common base-paired elements. Nucleotide positions at 32, 35, and 69 are defined in green (Mandal, et al., 2004).

1.7.1   Function and structure of the Gly riboswitch in *B. subtilis*

The Gly riboswitch in *B. subtilis* is located in the the 5´-UTR of three-gene operon *gcvT–gcvPA–gcvPB* (Mandal, et al., 2004). Computational analysis of atomic-resolution structures proved that the Gly riboswitch in *B. subtilis* belongs to the GGA variant (Weinberg, et al., 2017). Nucleotide positions revealing the variant are marked with green circle (Figure 1-4). The Gly riboswitch in *B. subtilis* is selectively induced by Gly. It is a switch for *gcvT* operon which codes for the Gly cleavage system proteins that are responsible for degradation of Gly as an energy source. Gly degradation leads to the formation of 5-10-methylene-tetrahydrofolate, ammonia and CO_2 (Kikuchi, 1973).

Frequently, the *gcvT* element is located adjacent to genes for putative Na^+/Alanine symporters (Barrick, et al., 2004).

The absence of Gly leads to the formation of a Rho-independent terminator and attenuated transcript. In presence of Gly and binding simultaneously to the aptamers, allosteric modulation of the secondary and tertiary structure of the 5′-UTR directs to the formation of an anti-terminator structure and transcription of the full-length mRNA transcript (Mandal, et al., 2004; Phan & Schumann, 2007) (Figure 1-5).
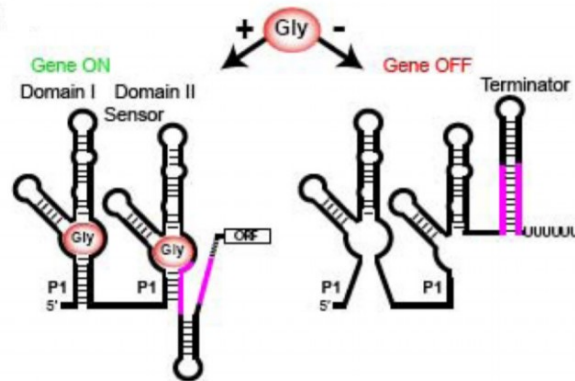
Figure 1-5. Mechanism of regulation of gene expression mediated by Gly riboswitch in *B. subtilis* (Serganov & Patel, 2009a).

The Gly riboswitch in *B. subtilis* is composed of two aptamer domains (I and II) that are connected by a short linker. The interaptamer linker forms helix P0 with the 5′ region of the first aptamer domain (Baird & Ferre-Damare, 2013). P0 folds into a kink-turn and supports partial prefolding of Gly-free aptamers. P0 kink-turn helps to organize more compact structure that enhances glycine binding and facilitates other interaptamer interactions (Baird & Ferre-Damare, 2013). P0 helix is connected to P1. Tertiary contact occurs between P3 of the second aptamer and P1 of the first aptamer and causes aptamer dimerization. Previous report showed that dimerization of t he two aptamers is dependent on Gly binding in at least one aptamer (Ruff & Strobel, 2014). Aptamer dimerization stabilizes the P1 stem of aptamer II, that controls the expression platform (Ruff & Strobel, 2014). Interaptamer interactions and dimerization can increase the glycine-binding affinities of the individual aptamers (Figure 1-6) (Ruff & Strobel, 2014; Baird & Ferre-Damare, 2013; Sherman, et al., 2012).

Figure 1-6. Schematic representation of the Gly riboswitch in *B. subtilis*. It is composed of aptamer (I) and aptamer (II). P1 through P3 represent common base paired elements. P0, K-turn sequence in *B. subtilis* is presented (Baird & Ferre-Damare, 2013).

1.7.2 Gly riboswitch in *S. pyogenes*

55 putative sRNAs were identified in *S. pyogenes* M49 (Patenge, et al., 2012). Among those, a putative Gly riboswitch (*ribogly*) was found, which is located within the 5´-UTR of a putative Na$^+$/Alanine symporter gene that is preceding a cation efflux system protein gene (Table 1).

Table 1. Characteristics of *ribogly* upstream and downstream genes in *S. pyogenes* M49.

Identification	*ribogly*	Upstram gene	Downstream genes	
		DNA helicase pcrA	Na$^+$/Alanine symporter	cation efflux system protein
Old locus tag	-	Spy49_1007c	Spy49_1008	Spy40_1009
Locus tag	-	Spy49_RS05030	Spy49_RS05035	Spy49_RS05040
Location	1000814-1000899	998180-1000498	1001027-1002349	1002469-1003704
Rfam	RF00504	-	-	-

Provided by: https://www.ncbi.nlm.nih.gov/nuccore/209558587

http://rfam.xfam.org/family/RF00504

The putative Gly riboswitch is annotated by NCBI (https://www.ncbi.nlm.nih.gov/nuccore/209558587) (Figure 1-7) , and Rfam database using Rfam web service (http://rfam.xfam.org/family/RF00504).

Figure 1-7. Representation of the location of the Na$^+$/Alanine symporter gene and the cation efflux system protein gene as well as the upstram gene (*pcrA*) from NCBI. *pcrA* is located on the opposite strand. The *ribogly* is depicted in green.

1.8   CRISPR system

Another family of sRNAs is involved in CRISPR systems (<u>C</u>lustered <u>R</u>egularly <u>I</u>nterspaced <u>S</u>hort <u>P</u>alindromic <u>R</u>epeats), which play a key role in bacterial defense against viruses, transposable elements and conjugative plasmids. CRISPR systems are found in approximately 90% of sequenced archaeal genome and approximately 40% of sequenced bacterial genome (Marraffini & Sontheimer, 2010). CRISPR systems are considered as RNA-guided adaptive immune systems which target and inactivate plasmids and bacteriophages. They consist of short foreign DNA sequences that are separated by repeat sequences adjacent to CRISPR-associated (*cas*) genes. ~30 base pair foreign DNA fragment from viruses or mobile elements integrates into the bacterial genome at CRISPR loci and served as a memory for the next encountering with invading viruses. Bacteriophages and plasmids are detected and removed by CRISPR system composed of Cas proteins and a library of small CRISPR RNAs (crRNAs) (Jiang & Doudna, 2015). The CRISPR-Cas complex targets non-self DNA through base pairing with the crRNA guide sequence, leading to Cas protein-mediated DNA cleavage and degradation (van der Oost, et al., 2014).

CRISPR-Cas systems differ in the number and sequences of the associated *cas* genes, the repeat sequence and the number of repeat-spacer units (Kunin, et al., 2007). CRISPR-Cas systems have been classified based on their *cas* gene content, in three main types, each containing different subtypes that demonstrate different mechanisms of targeting of the invasive nucleic acid (van der Oost, et al., 2014) (Appendix 7-1).

In general, the mechanism of the CRISPR system consists of three stages. The first stage is the acquisition step. An invading genetic element is fragmented and integrated as a new spacer at the leader end of the CRISPR array. The second stage is named expression step. The CRISPR locus is transcribed and pre-crRNA is processed into small crRNA by

CRISPR-associated (Cas6) and/or housekeeping ribonucleases (such as RNase III) and the mature crRNAs together with Cas proteins form a crRNP complex (CRISPR ribonucleoprotein). The third step is known as degradation step. The crRNP scans foreign nucleic acid for a complementary sequence and the target is degraded by Cas nucleases (Heler, et al., 2014; van der Oost, et al., 2014) (Figure 1-8).
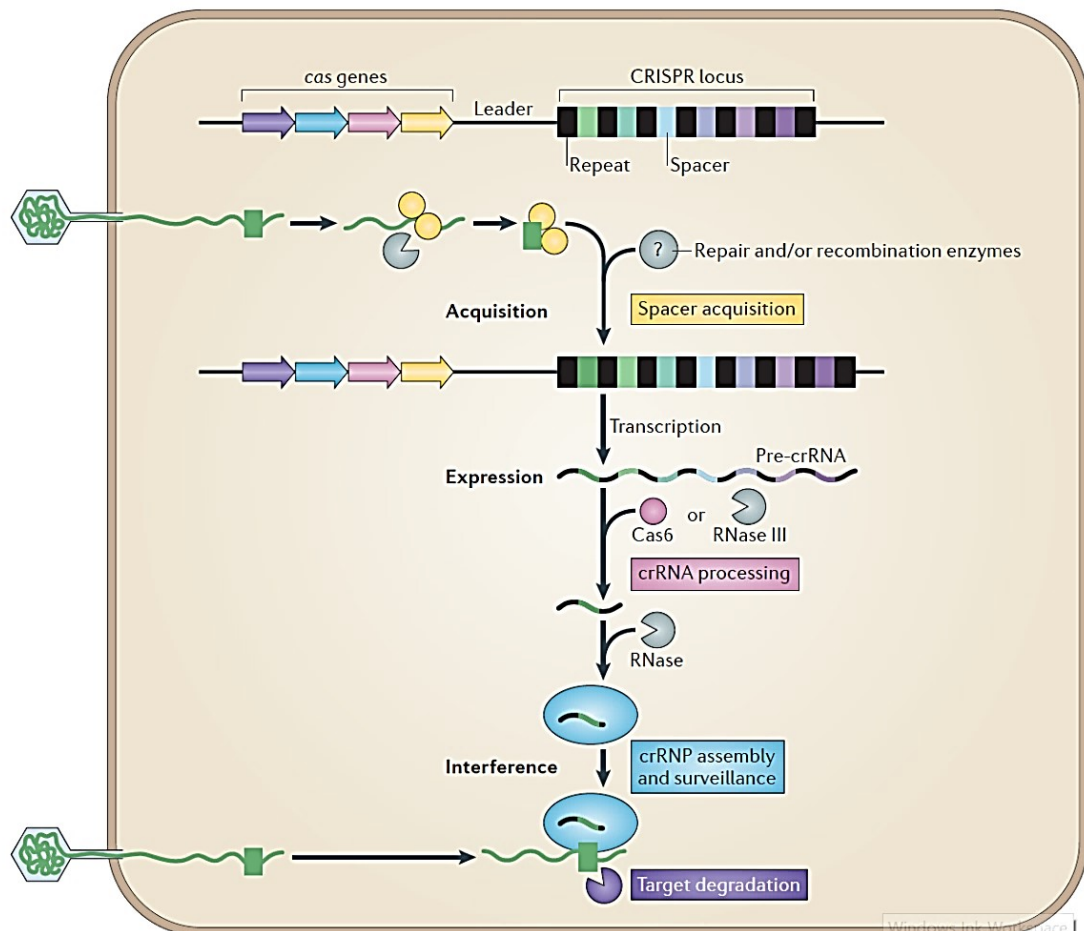
Figure 1-8. Schematic illustration of the function of CRISPR-Cas system which consists of three steps. The first step (yellow box) is considered as spacer acquisition. Short fragment of invading DNA is integrated into the genome at the end of CRISPR array. At the second step (pink box), when transcription of CRISPR array is complete, it is processed into small crRNAs. The combination of crRNAs and Cas proteins forms crRNP. During the third step, crRNP complex scans invading nucleic acid for a complementary sequence and Cas nucleases mediate degradation of the target (van der Oost, et al., 2014).

CRISPRs also are known to play a role in regulating self-gene expression by utilizing spacers that target self-genes. One hypothesis suggests that self-targeting is a consequence of autoimmunity (Stern, et al., 2010). The CRISPR immune system also helps the bacterium to bypass the human immune system. It was revealed that CRISPR can change the pathogenic behavior of *Campylobacter jejuni* (Louwen, et al., 2013).

1.9 CRISPR system in *S. pyogenes*

S. pyogenes strains usually have two CRISPR-Cas loci in their genomes: type II-B, which has been shown to be functional, and type I-C (Hatoum-Aslan & Marraffini, 2014). Beside these findings, there are some strains which lack CRISPR systems. These strains have more prophages than CRISPR harboring strains (Nozawa, et al., 2011).

In one study, (Patenge, et al., 2012) in *S. pyogenes* M49, one sRNA gene belonging to the CRISPR system was identified. It was named *moses16,* with ID number sRNASpy491206c (Raasch, et al., 2010). *Moses16* was expressed during growth in culture media as well as in blood. In another study in *S. pyogenes* serotype M1 it was silent (Perez, et al., 2009). In the present work, it was named *candidate34* (*cand34*). *cand34* is an sRNA annotated as CRISPR RNA and located adjacent to the *cas* genes annotated as CRISPR I-C associated proteins. The order of *cas2, cas1, cas4, cas7/csd2, cas8/csd1, cas5, cas3* is in accordance to CRISPR system I-C (Figure 1-9).
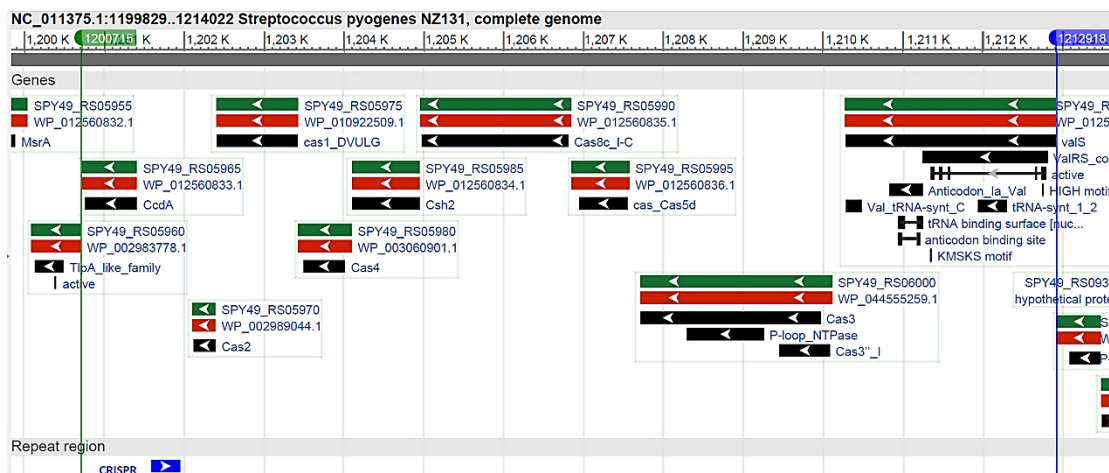
Figure 1-9. Pictorial representation of the location of CRISPR system I-C RNA and *cas* genes. https://www.ncbi.nlm.nih.gov/nuccore/209539788?report=graph.

Aim of this work

The aim of this study was the investigation of two potential regulatory RNAs in *S. pyogenes* serotype M49 strain 591. In Part I, functional characterization of a *cis*-regulatory RNA which is annotated as a putative riboswitch and is encoded upstream of an open reading frame annotated as Na$^+$/Alanine symporter (Spy49_1008) was performed. Clarifying the regulatory function of the *ribogly* is a prerequisite of its

potential use for recombinant gene expression purposes. Due to the mechanism of riboswitches with allosteric properties, sensing a specific metabolite and modulating the expression of downstream operons, production of recombinant proteins can be controlled using riboswitches as was established previously (Phan & Schumann, 2007).

Another potential application can be using riboswitch elements as targets for therapeutic agents. Effector-analogues can be used as artificial riboswitch ligands to restrict bacterial growth. (Soukup & Breaker, 2000). In Part II, the influence of a CRISPR element on the virulence of *S. pyogenes* was studied.

Objectives

Part I: Investigating the mechanism of a glycine-dependent riboswitch

- ❖ To creat and test a luciferase-based reporter gene assay in the presence and absence of glycine.
- ❖ To analyze the transcripts of the putative riboswitch and the corresponding operon by RT-PCR and Northern blotting to clarify the mechanism of the *cis*-regulatory RNA in response to the ligand and its regulatory effects on the $Na^{+/}$Alanine symporter gene transcript.
- ❖ To investigate the stability of the putative riboswitch and Na^+/Alanine symporter gene transcripts.
- ❖ To investigate the native promoter activity in the metabolite-dependent induction and influence on the respective operon, by fusing endogenous promoter to the reporter gene.
- ❖ To determine the operon structure of downstream gene(s) by RT-qPCR and Northern blotting.

Part II: Investigating the influence of the *cand34* on the virulence of *S. pyogenes* M49.

- ❖ To construct a deletion mutant of the *cand34* in *S. pyogenes* M49.
- ❖ To construct a complementation and an overexpression strain for comparison.
- ❖ To perform human blood survival assay with the different strains.
- ❖ To determine the adherence to and internalization into human keratinocytes of the different strains.
- ❖ To identify potential regulatory targets of the sRNA using bioinformatic target analyses and to evaluate predicted targets using RT-qPCR transcript analyses.

2 Material and Methods

2.1 Materials

Table 2. Chemicals and enzymes

Chemicals and enzymes	Manufacturer
Agar (Agar technical No.3)	Oxoid Ltd., Basingstoke, England
Agarose Standard	Carl Roth GmbH & Co. KG, Karlsruhe
Alanine	Sigma-Aldrich Chemie GmbH, Steinheim
Arginine-Hydrochloride	Sigma-Aldrich Chemie GmbH, Steinheim
Asparagine	Sigma-Aldrich Chemie GmbH, Steinheim
Biotin	Sigma-Aldrich Chemie GmbH, Steinheim
Calcium chloride-Hexahydrate	Merck KGaA, Darmstadt
Casein Hydrolysate	Merck KGaA, Darmstadt
Columbia Agar with 5% Sheep Blood	BD, Maryland, USA
10x Coral Load PCR Puffer	Qiagen GmbH, Hilden
Cysteine	Sigma Aldrich Chemie Gmbh, Steinheim
Cystine	Fluka Chemie AG, Buchs, Schweiz
DIG Easy Hyb	Roche Diagnostics GmbH, Mannheim
Dipotasium hydrogen phosphate	AppliChem GmbH, Darmstadt
Disodium hydrogen phosphat-Dihydrat	Merck KGaA, Darmstadt
D-Luciferin	Molecular Probes, Eugene, Oregon, USA
dNTP Mix	Qiagen GmbH, Hilden
Diethylpyrocarbonat	Sigma-Aldrich, St. Louis, MO, USA
DMEM (Dulbecco's Modified Eagle Medium), high glucose, GlutaMAXTM	GIBCO, Life Technologies, Paisley, Scotland
EDTA-Na$_2$ x 2 H$_2$O	Merck KGaA, Darmstadt
Ferrous sulfate heptahydrate (FeSO$_4 \cdot$ 7H$_2$O)	Fluka Chemie AG, Buchs, Schweiz
Iron(III) nitrate nonahydrate	Merck KGaA, Darmstadt
Acetic acid	AppliChem GmbH, Darmstadt
Erythromycin	Fluka Chemie AG, Buchs, Schweiz
Ethanol (absolute)	Zentralapotheke der Universitätsklinik Rostock
Ethidium bromide (1 % [w/v] in H$_2$O)	Carl Roth GmbH & Co.KG, Karlsruhe
10x FastDigest® Buffer	Thermo Fisher Scientific, Dreieich
10x FastDigest® Green Buffer	Thermo Fisher Scientific, Waltham, USA
FastDigest® Restriktion enzymes	Thermo Fisher Scientific, Waltham, USA

Table 2. Chemicals and enzymes (Continued)

Chemicals and enzymes	Manufacturer
Folic acid	Fluka Chemie AG, Buchs, Schweiz
Formaldehyde 37%	Baker, Deventer, Holland
Glucose	Merck KGaA, Darmstadt
Glutamine	Sigma-Aldrich Chemie GmbH, Steinheim
Glutamic acid	Sigma-Aldrich Chemie GmbH, Steinheim
Glycerin (99 %)	Carl Roth GmbH & Co.KG, Karlsruhe
Gly	Carl Roth GmbH & Co.KG, Karlsruhe
Glycyl-Glycin	Fluka Chemie AG, Buchs, Schweiz
Guanine-Hydrochloride-Monohydrate	Fluka Chemie AG, Buchs, Schweiz
Yeast extrakt	Oxoid Ltd., Basingstoke, England
Histidine-Hydrochloride-Monohydrate	Sigma-Aldrich Chemie GmbH, Steinheim
HaCat	DKFZ, Heidelberg (Boukamp, et al., 1988)
Hydroxyproline	Sigma-Aldrich Chemie GmbH, Steinheim
IRDye® 800CW Streptavidin	LI-COR Biotechnology GmbH, Germany
Isoleucine	Sigma-Aldrich Chemie GmbH, Steinheim
6x Loading Dye	Thermo Fisher Scientific, Waltham, USA
LB Broth Base	Invitrogen, Carlsbad, California, USA
Leucine	Sigma-Aldrich Chemie GmbH, Steinheim
Luciferin	Sigma-Aldrich, St. Louis, MO, USA
Lysine	Merck KGaA, Darmstadt
Magnesium sulfate heptahydrate	Merck KGaA, Darmstadt
Manganese(II) sulfate	Fluka Chemie AG, Buchs, Schweiz
Maxima SYBR Green/ROX qPCR Master Mix	Thermo Fisher Scientific, Waltham, USA
Millennium™ RNA Marker	Applied biosystem (Ambion)
Methionine	Sigma-Aldrich Chemie GmbH, Steinheim
MOPS	Fluka Chemie AG, Buchs, Schweiz
NADH	Sigma-Aldrich Chemie GmbH, Steinheim
Nicotinamide	Fluka Chemie AG, Buchs, Swiss
O´ GeneRuler™ 1 kb DNA Ladder	Thermo Fisher Scientific, Waltham, USA
Odyssey® Blocking Buffer	Li-cor Bioscience-GmbH, Bad Homburg
Penicillin/Streptomycin-stock solution (1%)	Invitrogen GIBCO, Fisher Scientific, Schwerte
Potassium chloride	Merck KGaA, Darmstadt
Potassium dihydrogen phosphate	Merck KGaA, Darmstadt
Potassium hydroxide	Merck KGaA, Darmstadt

Table 2. Chemicals and enzymes (Continued)

Chemicals and enzymes	Manufacturer
Primer	Eurogentec Germany GmbH, Cologne
Pantothenic acid	Fluka Chemie AG, Buchs, Swiss
Phenylalanine	Sigma-Aldrich Chemie GmbH, Steinheim
Phusion® DNA Polymerase	Thermo Fisher Scientific, Waltham, USA
5x Phusion® High-Fidelity PCR Buffer	Thermo Fisher Scientific, Waltham, USA
Proline	Sigma-Aldrich Chemie GmbH, Steinheim
Pyridoxal hydrochloride	Sigma-Aldrich Chemie GmbH, Steinheim
Pyridoxamine-Dihydrochloride	Sigma-Aldrich Chemie GmbH, Steinheim
Riboflavin	Fluka Chemie AG, Buchs, Swiss
RiboRuler Low Range RNA	Thermo Fisher Scientific, Waltham, USA
Rifampicin	Riemser, Greifswald
2X RNA loading dye	Thermo Fisher Scientific Inc.
Serine	Sigma-Aldrich Chemie GmbH, Steinheim
Sodium acetate	Merck KGaA, Darmstadt
Sodium chloride	Carl Roth GmbH & Co.KG, Karlsruhe
Sodium hydrogen carbonate	Merck KGaA, Darmstadt
Sodium hydroxide	Merck KGaA, Darmstadt
Sodium dihydrogen phosphate	Merck KGaA, Darmstadt
Sodium dodecyl sulfate	SERVA, Heidelberg, Germany
Spectinomycin dihydrochloride pentahydrate	MP Biomedicals LLC, Illkirch, France
10x T4 DNA Ligase Buffer	Thermo Fisher Scientific, Waltham, USA
T4 DNA Ligase	Thermo Fisher Scientific, Waltham, USA
Taq-Polymerase	Qiagen GmbH, Hilden
Thiamine-Hydrochloride	Fluka Chemie AG, Buchs, Swiss
Threonine	Sigma-Aldrich Chemie GmbH, Steinheim
Trypsin – EDTA (0,25%), phenol red	GIBCO by Life Technologies, Paisley, Scotland
Tryptophan	Sigma-Aldrich Chemie GmbH, Steinheim
Tris-HCl	Carl Roth GmbH & Co.KG, Karlsruhe
Todd Hewitt Broth	Oxoid Ltd., Basingstoke, England
Tween® 20	Serva, Heidelberg, Germany
Tyrosine	Sigma-Aldrich Chemie GmbH, Steinheim
Uracil	Sigma-Aldrich Chemie GmbH, Steinheim
Valine	Sigma-Aldrich Chemie GmbH, Steinheim

Table 3. Materials

Consumable supplies	Manufacturer
Cell culture flask (75 cm2)	Greiner Bio-One, Kremsmünster, Österreich
Cell culture 24 well-plate	Greiner Bio-one, Kremsmunster, Osterreich
Disposable Plastic Vacuum Filters 150 and 500 ml (0,22 µm)	Corning incorporated, Kaiserslautern
Disinfectants AHD 2000	Lysoform Dr. Rosemann GmbH, Berlin
Disinfectants Bacillol® AF	Bode Chemie, Hamburg
Disposable Inoculation Loops	Greiner Bio-One, Kremsmünster, Austria
Electroporation cuvettes	peqLab Biotechnologie GmbH, Erlangen
Falcon Tube Cellstar® (15 ml, 50 ml)	Greiner Bio-One, Kremsmünster, Austria
GENbox anaer (bag)	Biometra, Göttingen
Glass beads (0.10-0.11 mm)	Sartorius, Gottingen
Glass beads (2.85-3.3 mm)	Carl Roth GmbH & Co. Karlsruhe
Hybridization bottle	Thermo Fisher Scientific, Waltham, USA
MicroAmp® Fast 96-Well Reaction Plate	Thermo Fisher Scientific, Waltham, USA
Parafilm	Bemis Company, Wisconsin, USA
PCR Soft tubes (0.2 ml)	Biozym, Hessisch Oldendorf
Pipettes (10, 20, 100, 200, 1000 µl)	Gilson, France
Pipette tips (10, 20, 100, 200, 1000 µl)	Greiner Bio-One, Kremsmünster, Austria
Pipetting aids	Eppendorf, Hamburg
Plastik pipets Cellstar® (5, 10, 25 ml)	Greiner Bio-One, Kremsmünster, Austria
RNA Nano Chips	Agilent Technologies, Inc. Santa Clara, USA
Rotilabo®-Blotting Papers, Thickness 1.0 mm	Carl Roth GmbH & Co. KG, Karlsruhe
Roti®-Nylon 0.2, pore size 0.2 µm	Carl Roth GmbH & Co. KG, Karlsruhe
Round bottom test tubes (12 x 75 mm)	Roche Diagnostics GmbH, Mannheim
Safe-Lock tubes (0.5, 1.5, 2 ml)	Eppendorf AG, Hamburg
Sterilfix® Injection filter (0.2 µm)	Braun, Melsungen
UVpette tips	Biozym, Hessisch Oldendorf

Table 4. Devices

Devices	Manufacturer
Agilent G2565AA Scanner	Agilent Technologies, Inc. Santa Clara, USA
Anaerocult® Jar	Merck KGaA, Darmstadt
Analytical balance	BP 4100S, Sartorius, Göttingen
Biological safety cabinet class ll	Thermo Fisher Scientific, Waltham, USA
Centrifuge Biofuge pico	Heraeus, Hamburg
Centrifuge Biofuge *fresco*	Heraeus, Hamburg
Refrigerated centrifuge Type 5417R	Eppendorf, Hamburg
Centrifuge Varifuge 3.OR	Heraeus, Hamburg
Centrifuge Megafuge 1.0R	Heraeus, Hamburg
Centrifuge Sorvall Lynx 4000	Thermo Fisher Scientific, Waltham, USA
Eddy Jet Spiralplater1.22	IUL Instruments GmbH, Königswinter
Electroporator ECM 399	Harvard Apparatus, Holliston, USA
Freezer -80 °C Type Hera freeze	Heraeus, Hamburg
Fume hood	
Gel documentation system	Intas Science GmbH, Göttingen
GENbox anaer System	biomérieux Deutschland GmbH, Nürtingen
GS Gene Linker® UV Chamber	Bio-Rad Laboratories GmbH, California, USA
Horizontal electrophoresis system	Bio-Rad Laboratories GmbH, Munich
Hybridization oven	Biometra, Ma, USA
Incubator B6060	Heraeus, Hamburg
Luminometer Lumat LB 9501	Berthold Technologies GmbH Bad Wildbad
Microscope CKX41	Olympus Deutschland GmbH, Hamburg
Mini-Vacuum pump E1	Axon Lab AG, Baden-Dättwil
Odyssey® Imager	Li-cor Bioscience-GmbH, Bad Homburg
pH-Meter WTW Serie Type pH 720	WTW GmbH, Weilheim
Picodrop Microliter Spectrophotometer	Biozym, Oldendorf
Power supply Model 200/20	Bio-Rad-Laboratories GmbH, Munich
Power supply, Power Pac 3000 B	Bio-Rad-Laboratories GmbH, Munich
Qubit®2.0 Fluorometer	Life Technology (Invitrogen), Schwerte
Ribolyser	Peqlab Biotechnologie Gmbh, Erlangen
Spectrophotometer SmartSpec™ 3000	Bio-Rad-Laboratories GmbH, Munich
Thermocycler Type T3	Biometra GmbH, Göttingen
Thermomixer	Comfort Eppendorf, Hamburg
GS Gene Linker® UV Chamber	Bio-Rad-Laboratories GmbH, Munich
ViiA™ 7 Real-Time PCR System	Thermo Scientific, Waltham, USA

Table 4. Devices (Continued)

Devices	Manufacture
Vortex-Genie Touch Mixer	Scientific Industries, Bohemia, NY, USA
Water mantel incubator with regulated CO2 and O2 pressures	Heraeus, Hamburg

Table 5. Bacterial strains

Bacterial Strain	Plasmid	Resistance	Strain[*]	Source
E. coli DH5α	-	-		GibcoBRL Eggenstein
E. coli DH5α	pFW11-*luc*2	Spectinomycin	1985	IMIKRO (Podbielski, et al., 1999a)
E. coli DH5α	pFW11-*ribogly-luc*2	Spectinomycin	4863	IMIKRO (Popp, 2015)
S. pyogenes M49 strain 591**	-	-	4588	R. Lütticken (Aachen)
S. pyogenes M49/ pFW11-*ribogly-luc2*	pFW11-*ribogly-luc2*	Spectinomycin	4864	IMIKRO (Popp, 2015)

*Strain number of IMIKRO strain collection.

** In following text, short form; *S. pyogenes* M49 refers to this strain.

Table 6. Antibiotics

Antibiotics*	Solvent	Concentration for *S. pyogenes*	Concentration for *E. coli*
Spectinomycin	Ultra pure water	60 µg/ml	100 µg/ml
Erythromycin	Ethanol absolute	5 µg/ml	150 µg/ml
Rifampicin	Dimethyl sulfoxide (DMSO)	200 µg/ml	(not used at this study)

Penicillin/Streptomycin: Ready to use for cell culture 5000 U/ml

*Sterilization was performed with filtration through Sterifix ® Injection filter 0.2 µm.

Solutions containing Ethanol were not sterilized.

Table 7. Plasmid

Plasmid (Size)	Antibiotic resistance	Origin
pFW11-*luc*2 (4210 bp)	Spectinomycin	IMIKRO (Podbielski, et al., 1999a)
pFW11-*ribogly-luc2* (5107 bp)	Spectinomycin	IMIKRO (Popp, 2015)
pAT19 (6600 bp)	Erythromycin	(Trieu-Cuot, et al., 1991)

Table 8. Oligonucleotides and their sequences

	Oligonucleotide	Sequences 5´-3´
	Gly_luc_xho_rv	CCTCGAGCTCCTAAAAGTCATCGAAGACG
	Ribo.prom. BamHl-fwd	TCGGATCCATGCGCTTCATCAACG
	Ribo.prom. xhol-rev	CCTCGAGAATTAGGACTATCATACCTC
	Luc-check-rev	CCATGATAATAATTTTCTGGA
	RT_Luc_for	GAGACATAGCTTACTGGGACG
	RT_Luc_rev	TATCGACTCCAATTCAGCGGG
	5s-RNA-RT-fwd	AGCGACTACCTTATCTCACAG
	5s-RNA-RT-rev	GAGATACACCTGTACCCATG
	groES-M49-for	TCGGGTGTTCGCACTATTACA G
	groES-M49-rev	CATGCCCATTTTCAACTAAAACC
A	RT_*ribogly*_for	CTGGAGAGACCTTATTAGGC
	RT_*ribogly*_rev	GAGAGATTGAGCAGTATGCC
B	RT_AASymporter_for	AATCAGCCTTCACTCCGACAGC
	RT_AASymporter_rev	ACTCATTGGAGAAGACACCGCG
C	Efflux.fwd	GTCTCAACGAGCTAGGACC
	Efflux.rev	GGCAGGCTCAACGTGAATG
D	Ribo.J.fwd	TCAGGCAAAAGGACAGAAGG

Table 8. Oligonucleotides and their sequences (Continued)

Nucleotides	Sequences 5´-3´
E AA.J.rev	CGAGTGCTATCATCTTTCTCTCC
AA.Efflux.joint.fwd	CGCCAGTTGTTATTTTAGAAACC
AA.Ffflux.joint.rev	GATTGGCATCGGCTGGTTGGC
T7.*ribogly*_rev(2)	cttaatacgactcactataggGAGAGATTGAGCAGTATGCC
AAsymorter_fwd (4)	GGACGGGGATTTACCTTACC
T7. AASymporter_rev (4)	cttaatacgactcactataggCAGTTCCGACAGTAGCGGC
T7.5s-RNA-RT-fwd	cttaatacgactcactataggAGCGACTACCTTATCTCACAG
T7.Efflux.rev	cttaatacgactcactataggGGCAGGCTCAACGTGAATG
BamHI_cand34_5`flanking_fwd	AG<u>GATCC</u>TTCCAAGGTCGGAGTAAACG
SalI_cand34_5´flanking_rev	GC<u>GTCGAC</u>AGCCTTTTTATCTCCATTTCGC
Ncol_cand34-3´ flanking-fwd	G<u>ATCCATGG</u>CCTGACCTCATTAAAAGAGAG
PstI_cand34-3´ flanking-rev	GA<u>CTGCAG</u>AGCTGCCATGCTCATCTTGTCCT
Compl _M16/c34_Bam_fwd	A<u>GGATCC</u>AATCTCTTGTGCGAAGCTAGC
Compl _M16/c34_Sall_rev	GC<u>GTCGAC</u>CTTGAATTTGATACGCTATCAG
Spec-weg-FOR	ATATTGAATGGACTAATGAAA ATG
RT-cyto-fwd	GCATTGAACAAACCGTTGCG
RT-cyto-rev	GGTACTTACTGGGAGCCG
RT-cas2-fwd	CCAAGTGTTTCCACACGCC
RT-cas2-rev	CGTCATGTTGCCAAACTCTG
RT-ABC-fwd	GGGTTTACCCTTGAGCATTACC
RT-ABC-rev	GGATAAAGATAGCTCCAAAGGTCC
RT-NAD Synthetase-fwd	GCTTGATCACTTGTTTCCTCC
RT-NAD Synthetase-rev	GGCATTTCTGGAGGTCAGG

- Restriction sites are underlined. Nucleotide sequences of the T7 promoter are given in lower case letters.
- Primer sets A, B, C and D amplified fragments which are shown in (Figure 3-6).
- All primers were synthesized by "Eurogentec Deutschland GmbH" (Cologne).

Table 9. kits

Kit	Supplier
QIAquick® PCR Purification Kit (250)	Qiagen GmbH, Hilden
QIAprep® Spin Miniprep Kit (250)	Qiagen GmbH, Hilden
TURBO DNA-*free*™ Kit	Life Technologies (Ambion), Schwerte
SuperScript™ First-Strand Synthesis System for RT-PCR	Life Technologies (Invitrogen), Schwerte
Colony Fast-Screen™ Kit	Epicentre Biotechnologies, Madison, USA
MEGAscript® T7 Transcription Kit	Life Technologies (Invitrogen), Schwerte
RNA Clean & Concentrator™-5 kit	Zymo Research, USA
RNA Clean & Concentrator™-25 kit	Zymo Research, USA
Pierce™ RNA 3' End Biotinylation	Thermo Scientific, USA

2.2 Methods

2.2.1 Media preparation

Broth media were prepared by suspending the defined amount of component as following in 1 L of Ultra-pure water in screw capped bottles. Autoclave cycle conditions were 121°C (250 F), 15 psi for 15 min.

Solid media were prepared by adding 1.5% agar to broth media. Sterilized media were poured in Petri dishes. *S. pyogenes* was cultivated in THY medium and *E. coli* was grown in LB medium.

Addition of antibiotics to media was carried out after cooling the sterilized media to approximately 50°C. The respective concentrations of antibiotics are given in Table 6. After pouring the media in plates these were allowed to solidify. The plates were stored at 2-8°C in a plastic bag to avoid drying out.

◆ **LB-Medium (Lennox L Broth Base):**

LB-Broth-Base	20g
NaCl	4.5g

Sterilization with autoclave

◆ **THY-Medium (Todd Hewitt Yeast):**

Todd-Hewitt-Broth	36.4g
Yeast-Extract	5 g

Sterilization with autoclave

◆ **CDM**

Chemically Defined Medium (CDM) was used as described (van de Rijn & Kessler, 1980). Solutions were prepared as described in 2.2.1 and sterilized using 0.22 µm filters. CDM was prepared according to Table 10. pH 7,4 was adjusted with NaOH. The solution was diluted to the final volume of 1000 ml using Ultra pure water. CDM was sterilized by filtration through a 0.22 µm filter.

▪ **Solutions for CDM**

Phosphate Buffer: 100 ml

KH_2PO_4	2 g
K_2HPO_4	10 g

Ultra pure water to final volume

Asparagine: 200 ml

Asparagine	1 g

Ultra pure water to final volume

Amino Acid Mix: 400 ml

Was not heated higher than 57 °C

L-Aspartic Acid	2 g
L-Phenylalanine	2 g
L-Serine	2 g
L-Proline	4 g
L-Hydroxyproline	4 g
Gly	4 g
2.5 N NaOH	6.4 ml

Ultra pure water to final volume

L-Leucine: 100 ml

Was heated not higher than 57 °C

L-Leucine	1 g

Ultra pure water to final volume

L-Glutamic acid : 200 ml

L-Glutamic acid	6 g
2.5 N NaOH	19.2 ml

Ultra pure water to final volume

D/L-Alanine: 200 ml

D/L-Alanine	4 g

Ultra pure water to final volume

L-Isoleucine: 200 ml

Was heated not higher than 57 °C

L-Isoleucine   2 g

Ultra pure water to final volume

L-Methionine: 200 ml

Was heated not higher than 57 °C

L-Methionine   2 g

Ultra pure water to final volume

L-Threonine: 100 ml

Was heated not higher than 57 °C

L-Threonine   2 g

Ultra pure water to final volume

L-Argenine HCl: 200 ml

L-Argenine HCl   4.84 g

Ultra pure water to final volume

L-Histidine HCl H$_2$O: 200 ml

L-Histidine HCl H$_2$O   5.4 g

Ultra pure water to final volume

L-Thryptophan: 200 ml

L-Thryptophan   4 g

2N HCl   20 ml

Ultra pure water to final volume

Folic Acid: 250 ml

Folic Acid	5 mg
2.5 N NaOH	0.1 ml
Ultra pure water to final volume	

L-Valine: 100 ml

L-Valine	2 g
Ultra pure water to final volume	

L-Lysine: 100 ml

L-Lysine	2.76
Ultra pure water to final volume	

Riboflavin:100 ml

Riboflavin	8 mg
conc. Acetic Acid	0.1 ml

Vitamin Mix: 250 ml

Nicotinamide	11 mg
Thiamine HCl	20 mg
Aminobenzoic Acid	4 mg
Pantothenic Acid	40 mg
Ultra pure water to final volume	

Biotin: 200 ml

Storage in brown bottle

Solution 1:	Biotin	5 mg
	95% EtOH	1 ml
	0.01N HCl	49 ml
Solution 2:	Solution 1	20 ml
	0.01N HCl	180 ml

Pyridoxamine 2HCl: 100 ml

Pyridoxamine 2HCl	23 mg
Ultra pure water to final volume	

MgSO$_4$ x 7 H$_2$O: 50 ml

MgSO$_4$ x 7 H$_2$O	7 g
2 N HCl	1 ml
Ultra pure water to final volume	

AGU Mix: 500 ml

Adenine sulfate	435 mg
Guanine HCl H$_2$O	310 mg
Uracil	250 mg
2 N HCl	50 ml
Ultra pure water to final volume	

Salts: 100 ml

FeSO$_4$ x 7 H$_2$O	100 mg
CaCl$_2$ x 6 H$_2$O	200 mg
MnSO$_4$	100 mg

Fe(NO$_3$)$_2$ x 9 H$_2$O	20 mg
2 N HCL	1.2 ml

Ultra pure water to final volume

L-Glutamine: 200 ml

L-Glutamine	4 g

Ultra pure water to final volume

L-Tyrosine: 400 ml

L-Tyrosine	4 g
2.5 N NaOH	16 ml

Ultra pure water to final volume

Sodium Bicarbonate: 10 ml

Was prepared freshly before use and not heated higher than 57°C

Sodium Bicarbonate	1.25 g

Ultra pure water to final volume

Solution was sterilized using Sterifix ® Injection filter 0.2 µm.

L-Cystine: 400 ml

L-Cystine	4 g
2 N HCl	80 ml

Ultra pure water to final volume

L-Cysteine: 5 ml

L-Cysteine	2.5 mg

Ultra pure water to final volume

Solution was sterilized using Sterifix ® Injection filter 0.2 µm.

Table 10. Components for 1000 ml Chemically Defined Medium

Component	Amount
Glucose	10 g
Na_2HPO_4	7.35 g
NaH_2PO_4	3.2 g
$Na(CH_3COO)$	4.5 g
Phosphate Buffer	10 ml
Asparagine	20 ml
Amino Acid Mix	20 ml
L-Leucine	20 ml
L-Glutamic Acid	3.4 ml
D/L-Alanine	5 ml
L-Isoleucine	10 ml
L-Methionine	10 ml
L-Threonine	10 ml
L-Argenine HCl	5 ml
L-Histidine HCl H_2O	5 ml
L-Thryptophan	5 ml
L-Valine	5 ml
L-Lysine	5 ml
Riboflavin	25 ml
Vitamin Mix	25 ml
Biotin	20 ml
Folic Acid	40 ml
Pyridoxamine 2HCl	6.25 ml
L-Glutamine	10 ml
L-Cystine	5 ml
L-Tyrosine	10 ml
AGU Mix	44 ml
$MgSO_4$ x 7 H_2O	5 ml
Salts	5 ml
Pyridoxal HCL	1 mg
NADH	2.5 mg

Following components were added to the medium before use.

Sodium Bicarbonate	200 µl	for 10 ml CDM
L-Cysteine	100 µl	

2.2.2 Overnight culture

2.2.2.1 *E. coli* overnight culture

To prepare an overnight culture of *E. coli* strain a single colony from LB agar plate was removed using a sterile loop and transferred to a tube containing 10 ml LB broth. In plasmid carrying strains respected antibiotic was added to the media. The tube was loosely closed and incubated in shaking incubator at 37°C and 180 rpm (revolutions per min) overnight. After incubation the culture was used for next experiments (2.2.3).

2.2.2.2 *S. pyogenes* overnight culture

An overnight culture of *S. pyogenes* was prepared by transferring a single colony from THY agar plate to a tube containing THY broth using a sterile loop. To culture mutant strains carrying antibiotic resistance cassette respected antibiotic was added to the media. The tube was loosely closed and incubated at 37°C and 5% CO_2 overnight. The culture was used for the subsequent experiments (2.2.3).

2.2.3 Strain maintenancece

650 µl overnight culture of the respective strain was mixed with 350 µl sterile Glycerol (99%) and stored at -20°C.

2.2.4 Measuring optical density

To plot standardized growth curves of the respective bacterial strain (absorbance versus time) 1 ml of overnight culture was inoculated in 19 ml of broth media. Samples were collected at hourly intervals. The optical density at a wavelength of 600 nm (OD_{600}) was measured using a spectrophotometer and recorded.

2.2.5 Reporter gene assay

   ♦ **Luciferase assay**

Bacteria were grown overnight in THY. Cells were pelleted and washed with PBS. Washed cells were inoculated into 6 tubes containing CDM with different concentrations of Gly; 0 mM, 0.01 mM, 0.1 mM, 1 mM, 2.6 mM and 10 mM to $OD_{600} \approx 0.08$- 0.1. Tubes were incubated at 37°C, 5% CO_2. In one-hour intervals and once after 24 hours, the OD_{600}

of each sample was measured and luciferase activity was monitored. A 150 µl aliquot of culture was mixed with 150 µl 1 x assay buffer in a test tube (12x75 mm). 150 µl luciferin was added to the mixture. The tube was placed in the luminometer immediately and emitted light was measured for 15 seconds at room temperature.

Assay buffer

Glycyl-Glycin	4.1 g
$MgCl_2$	2.54 g
Ultra pure water	500 ml

pH 7.8 adjusted with $Mg(OH)_2$

D- Luciferin solution

D-Luciferin (33 mM)

KH_2PO_4 buffer (10mM), pH 6.5

After preparation, 50 µl aliquots were stored at -20 °C. For every experiment, D-Luciferin solution was diluted in ratio of 1:100 with sterilized Ultra pure water.

DNA manipulation and analysis

2.2.6 Chromosomal DNA isolation

Chromosomal DNA of *S. pyogenes* was obtained from 5 ml overnight culture in THY broth. Cells were harvested by centrifugation at 4000 rpm for 5 min. Pellets were resuspended in 1 ml 1X PBS in a 1.5 ml microcentrifuge tube containing beads (0.1-0.11 mm) and mechanical cell (lysis) disruption was achieved using a ribolyser (6000 upm, 40 s, 1-times). This step was followed by centrifugation at 13000 rpm for 10 min. 100 µl of supernatant was transferred to a 1.5 ml microcentrifuge tube. DNA was purified with the QIAquick® PCR Purification Kit according to the manufacturer´s instruction. DNA was stored at -20 °C.

2.2.7 Plasmid DNA isolation

Plasmid isolation was performed using the QIAprep® Spin Miniprep Kit according to the manufacturer´s instruction. Plasmid solutions were stored at -20°C.

2.2.8 Primer design

Primers were designed considering: 1) the optimal length of 18-20 nucleotides, 2) GC content between 40 - 60 % and balanced distribution of A, T, G and C, 3) avoiding intra primer homology and inter primer homology, 4) adding the corresponding restriction site

to the primer followed by suggested extra base (s) at the beginning of the restriction site according to the restriction enzyme´s manufacturer. To calculate oligonucleotide melting temperature, self-complementarity, and hairpin formation estimation, OligoCalc (Kibbe, 2007) was used which is available at following link: http://biotools.nubic.northwestern.edu/OligoCalc.html.

2.2.9 PCR (Polymerase Chain Reaction)

PCR Components are given in Table 11. The Phusion polymerase was used for low error rates. It has a 3'-5'-Proofreading exonuclease activity (Thermofisher® Inc) (Mullis, et al., 1986).

Table 11. PCR reaction components

Component	Volume	Final Concentration
5x Phusion® High-Fidelity Buffer	4 µl	1 x
5x Phusion® High-Fidelity DNA-Polymerase	0.2 µl	0.02 U/µl
dNTPs (10 mM)	0.4 µl	200 µM
Forward primer (10 µM)	1 µl	0.5 µM
Reverse primer (10 µM)	1 µl	0.5 µM
Genomic DNA template	1 µl	(1 ng/ µl)
Aquabidest.	To final volume of 20 µl	

Thermocycling conditions

Step	Temperature	Time	
Initial denaturation	95°C	5 min	
Denaturation	95°C	10 s	
Annealing	Primer	30 s	25X
Extention	72°C	30 s/kb	
Final Extention	72°C	10 min	

2.2.10 Agarose gel electrophoresis

Agarose gels were prepared using 1% agarose (w/v) in TAE buffer. The agarose/buffer mixture was melted by heating in a microwave until the agarose was completely dissolved. The gel was poured into the gel casting tray with an appropriate comb to create wells. After setting the gel at room temperature and solidification, the comb was removed. The Casting tray was placed in a gel box with 1X TAE buffer and the gel was covered by the buffer. The marker O´ GeneRuler™ 1 kb DNA Ladder was loaded in one well. Samples were mixed with 0.2 vol, 6X loading dye and loaded in the wells. The lid of gel box was closed. Power supply settings were at 5 V/cm. After completing electrophoresis, the gel was immersed in 0.15 µl/ml ethidium bromide solution for 10 min. The gel was exposed to UV light and a photo was taken for documentation (Lee, et al., 2012).

<u>50 X TAE Buffer</u>

Tris-HCl	2.0 M
Accetic acid	1.0 M
EDTA-Na$_2$-2 H$_2$O	50 mM
pH	7.5

<u>6X loading dye</u>

Bromphenol blue	0.25% (w/v)
Cyanol	0.25% (w/v)
Glycerol	30% (v/v)
EDTA-Na$_2$-2 H$_2$O	50 mM

To estimate the size of double-stranded DNA on agarose gels, O´ GeneRuler™ 1 kb DNA Ladder from Thermo Fisher Scientific was used (Figure 2-1).
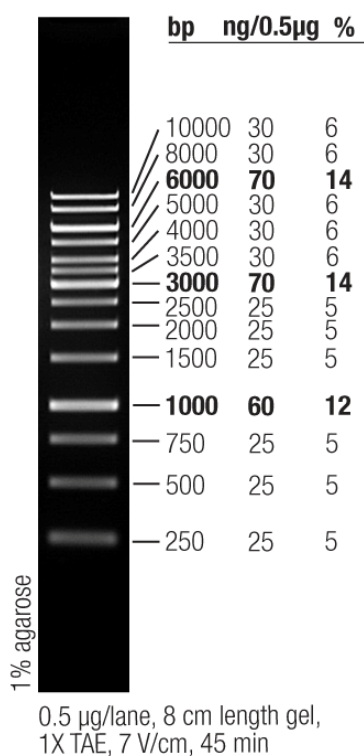
Figure 2-1. O´ GeneRuler™ 1 kb DNA Ladder (Thermo Fisher Scientific).

2.2.11   Molecular cloning

2.2.11.1 Restriction digest

Restriction digests were performed according to the instructions of the supplier.

Table 12. Restriction digest components

Components	Plasmid DNA	Purified PCR product
10X FastDigest Buffer	2 µl	2 µl
DNA	1000 ng	200 ng
FastDigest enzyme1*	1 µl	1 µl
FastDigest enzyme2*	1 µl	1 µl
Aquabidest.	to 20 µl	to 30 µl

*Thermo Scientific™ FastDigest™

Table 13. Restriction digest conditions

Restriction Enzyme	Restriction site	Incubation Temperature	Incubation Time(min)		Inactivation
			PCR Product	Plasmid DNA	
BamHI	G↓GATCC	37°C	5	5	80°C; 5 min
Nhel	G↓CTAGC	37°C	5	15	65°C; 5 min
Sall	G↓TCGAC	37°C	60	5	65°C; 10 min
XhoI	C↓TCGAG	37°C	10	10	80°C; 5 min
Pstl	CTGCA↓G	37°C	30	5	Purification*
Ncol	C↓CATGC	37°C	10	10	65°C; 15 min

*Enzyme inactivation was performed using QIAquick® PCR Purification Kit. Digests were performed with two restriction enzymes simultaneously in one reaction tube. Incubation time was set according to longer incubation time.

2.2.11.2 Ligation

Ligation was performed considering molecular sizes of the fragments. A molar ratio of 1:5 vector to insert according to DNA fragment sizes was calculated via NEBioCalculator (https://nebiocalculator.neb.com/#!/ligation). Reactions were set up in a microcentrifuge tube on ice. The T4 DNA Ligase Buffer was thawed and resuspended at room temperature. Reaction mixture was prepared (Table 14) and spun briefly. The tube was incubated at 16°C overnight. The product was used directly for transformation in *E. coli.*

Table 14. Components and their volumes for a ligation reaction

Component	Volume
T4 DNA Ligase Buffer (10X)	2 µl
Vector DNA	50 ng
Insert DNA	Molar ratio of 1:5 vector to insert according to fragment size
Aquabidest.	to 20 µl
T4 DNA Ligase	1 µl (0.1U/ µl)

2.2.11.3 Transformation of *E. coli* DH5α

2.2.11.3.1 Preparing chemically competent cells

E. coli DH5α competent cells were prepared for general cloning purpose via $CaCl_2$ method (Hanahan, 1983).

E. coli DH5α was streaked on LB plate and incubated at 37°C overnight. 5 ml LB medium was inoculated with a single colony and incubated at 37°C and 180 rpm overnight. 1 ml overnight culture was used to inoculate 100 ml LB medium and incubated at 37°C and 180 rpm until OD_{600} = 0.4. The Culture was chilled for 5 min on ice and swirled occasionally to ensure cooling. All bottles and solutions which were used in the process, were precooled. The cell suspension was centrifuged for 5 min at 4°C and 4000 rpm. The supernatant was decanted, and the pellets were carefully resuspended in 30 ml ice-cold buffer 1. The cell suspension was incubated for 90 min on ice. This step was followed by centrifugation for 5 min at 4°C and 4000 rpm .The supernatant was decanted, and the pellets were carefully resuspended in 4 ml ice-cold buffer 2. 100 µl aliquots were transferred in sterile, cold, 1.5 ml microfuge tubes as quick as possible and snap frozen in liquid nitrogen. The tubes (aliquots) were stored at -80°C.

Buffer1		**Buffer2**	
RbCl	100 mM	MOPS	10 mM
$MnCl_2$	50 mM	RbCl	10 mM
Potassium acetate	30 mM	$CaCl_2$	75 mM
$CaCl_2$	10 mM	Glycerol	15%
pH 5.8	Adjusted with HCl	pH 8	Adjusted with KOH

Buffers were sterilized by filtration using a 0.22 µm filter and were stored at room temperature.

2.2.11.3.2 Heat shock transformation

Chemically competent cells were thawed on ice. 10 µl ligation mixture was added to the thawed cells. Cells were incubated for 30 min on ice. This step was followed by 60 seconds incubation at 42°C. Tubes were transferred on ice for 2 min. 1 ml warm LB

medium was added and tubes were incubated at 37°C and 450 rpm for 1 hour in a Thermomixer (Comfort Eppendorf). Cells were centrifuged at room temperature for 5 min and 4000 rpm. The supernatant was decanted and 300 µl LB medium was added carefully. 50 µl, 100 µl, and 150 µl were plated on each LB-Agar medium plate which contained the appropriate antibiotics. Plates were incubated overnight at 37°C. To check for the presence of a plasmid in the transformants colony fast screening was performed. To evaluate the identity of the construct, plasmids were extracted (2.2.7) and cut with the appropriate restriction enzymes (2.2.11.1).

2.2.11.3.3 Colony fast screening

As a rapid procedure for identifying recombinant plasmids in DH5α colonies, colony fast screening was performed using Colony Fast-Screen™ Kit (epicentre®, Illumina company®). 10 µl of Restricti-Lyse™ Solution was distributed into a 0.5 ml microcentrifuge tube for each colony to be screened. Bacterial colonies were picked using a sterile loop and as much of the colony as possible was transferred to the microcentrifuge tube containing the lysis buffer. A sample of each colony was spotted onto a master plate. 2 µl of gel loading solution was added to each tube and mixed well using vortex vigorously (10-60 seconds). Tubes were incubated at 75°C for 15 min in a thermocycler. Each tube was centrifuged briefly. 15 µl reaction sample was analyzed by agarose gel electrophoresis.

2.2.11.4   Transformation of *S. pyogenes*

2.2.11.4.1 Preparing *S. pyogenes* electrocompetent cells

S. pyogenes was streaked on THY plate and incubated at 37°C and 5% CO_2 overnight. A single colony was inoculated in 3 ml THY broth/20mM Gly and incubated at 37°C and 5% CO_2 overnight. 500 µl of overnight culture was inoculated in 30 ml THY broth/20mM Gly at 37°C and 5% CO_2 to an OD_{600} of 0.2. The culture was chilled for 30 min on ice and swirled occasionally to ensure cooling. All bottles and the ultra pure water which were used in the process were precooled. Cells were collected by centrifugation at 4°C and 4000 rpm for 5 min. The supernatant was discarded, and cells were washed in 30 ml ice-cold ultra pure water. The cells were pelleted by centrifugation for 5 min at 4°C and 4000 rpm. The supernatant was decanted, and cells were resuspended in 30 ml ice-cold ultra pure water. Cells were collected by centrifugation at 4°C and 4000 rpm for 5 min. The supernatant was discarded and 500 µl ice-cold ultra pure water was added to the

pellet. Cell suspension was distributed in 160 µl aliquots in cold cuvettes appropriate for electroporation (Dunny, et al., 1991).

2.2.11.4.2 Electroporation

10-30 µg plasmid was added to an ice-cold cuvette containing 160 µl *S. pyogenes* suspension. The cuvette was placed in the electroporation device. Transformation was performed with 1750 V. 1 ml THY medium was added to the mixture immediately. Cells were incubated at 37°C and 5% CO_2 for 1 hour. 250 µl suspension were plated on each selective THY plate containing appropriate antibiotics. Plates were incubated in an anaerobic jar containing *S. pyogenes* pack at 37°C for 72 hours.

2.2.12 RNA analysing methods

2.2.12.1 RNA isolation

RNA was isolated from *S. pyogenes* M49. 700 µl of overnight culture was added to 10 ml of fresh CDM and incubated at 37°C and 5% CO_2 for 3 hours. Cells were pelleted by centrifugation at 4000 rpm at 4°C for 5 min. The supernatant was discarded. 1 ml TRIzol® was added to the pellets and incubated at room temperature for 5 min. Resuspended cells were transferred to the tube containing beads (0.1-0.11mm). Sample tube was processed in a ribolyser for 40 seconds at 6000 rpm. This step was followed by centrifugation at 13000 rpm at 4°C for 5 min. 750 µl of supernatant was transferred to a microfuge tube containing 600 µl acid phenol and 20 µl sodium acetate. Homogenization was achieved in a ribolyser at 4000 rpm for 2 x 10 seconds. The tube was centrifuged at 7500 rpm and 4°C for 15 min. The upper phase was transferred to a 2-ml tube without disturbing the interphase. This step was repeated to increase RNA purity. The upper phase was transferred to a 2-ml tube containing 1 ml ice-cold absolute ethanol. Sample was incubated at -20°C overnight. Centrifugation was performed at 13000 rpm for 60 min. To remove residual phenol 1 ml ethanol 70% in DEPC treated H_2O was added to the pellets. This process was followed by spinning at 13000 rpm for 5 min. Pellets air-dried at room temperature for 10 min and were resuspended in 50 µl DEPC treated H_2O. RNA concentration and purity were determined from a 1:10 diluted sample using a Picodrop Microliter spectrophotometer. The RNA sample was stored at -80°C.

2.2.12.2 DNase treatment

Genomic DNA contamination was removed from RNA using TURBO DNA-*free*™ Kit according to the kit manufacturer´s instruction. RNA concentration and purity were

determined using a Picodrop Microliter spectrophotometer. The RNA sample was stored at -80°C.

2.2.12.3 RNA quality assessment

RNA integrity was determined using Agilent G2565AA Scanner. Sample preparation was performed in accordance with the instructions of the manufacturer.

2.2.12.4 cDNA synthesis

cDNA was synthesized from the total RNA using "SuperScript™ First-Strand Synthesis System for RT-PCR" kit according to the manufacturer's instructions.

2.2.12.5 Real-time quantitative PCR (qPCR)

Real-time PCR was performed using a 96 well plate format in the ViiA™ 7 Real-Time PCR System. Each sample was prepared in duplicate as follows (Table 15). Two types of negative controls were included in each qPCR experiment: A no template control (NTC) and a no reverse transcriptase control (NRT) in which reverse transcriptase was omitted from the reverse transcription step. This control evaluates the presence of DNA contamination in an RNA sample.

Table 15. Components and their volume to establish one qPCR reaction

Component	Single rxn	Notes
SYBR® GreenER™ qPCR SuperMix Universal	10 µl	1X final conc.
Forward primer, 10 µM	1.5 µl	300 nM final conc.
Reverse primer, 10 µM	1.5 µl	300 nM final conc.
Template: Test well (20 ng of cDNA) NRT (20 ng of RNA)	4 µl	-
DEPC-treated water	3 µl	-
Final volume	20 µl	-

Thermocycling conditions

Step	Temperature (°C)	Time (min)	Cycles
Hold Stage	50	2:00	
	95	10:00	1
	95	0:15	
PCR-Stage	90	0:15	40
	60	1:00	
Melt Curve Stage	90	0:15	1
	60	1:00	

2.2.12.6 Comparative method or $\Delta\Delta C_T$ method of relative quantification

Data from qPCR were analyzed using relative quantification method. The $2^{-\Delta\Delta Ct}$ Method analyses the relative changes in gene expression from qPCR experiments (LIvak & Schmittgen, 2001). This method relies on direct comparison of Ct values between the sample and a calibrator (control group). The target and reference gene (housekeeping gene) were amplified from the same sample. 5S rRNA was used as an endogenous reference gene, which is present in constant levels in *S. pyogenes*.

The ΔCt value for each sample was achieved by determining the difference between the C_t value of the target gene and the C_t value of the reference gene for each unknown sample as well as for the calibrator.

- ΔCt (sample) = Ct target gene – Ct reference gene
- ΔCt (calibrator) = Ct target gene – Ct reference gene

In the next step, the sample was normalized to the calibrator sample. At the end, the relative quantification (RQ) of the target gene in the sample group in comparison to the calibrator was calculated.

- $\Delta\Delta C_T = \Delta Ct$ (sample) – ΔCt (calibrator)
- $RQ = 2^{-\Delta\Delta Ct}$

2.2.12.7 Confirmation of primer specificity and linearity

All real-time PCR primers were validated experimentally with two assays. First; Synthesis of a single amplicon was verified with a melting program which was run on the real-time PCR instrument at the end of the cycling program. The size of the RT–PCR product was analyzed by agarose gel electrophoresis. The second assay verified an amplification efficiency > 90 %. Serial dilution of known concentrations of DNA was performed and measured in duplicate (Pfaffl & Pfaffl, 2001). Ct values were plotted versus the initial amounts of the DNA concentrations. The data were fitted to a straight line and the slope was calculated. The efficiency was calculated through following formula:

$$Efficiency\ (\%) = 10^{-\frac{1}{Slope}} - 1 * 100$$

Efficiencies between 90% -100% were accepted (-3.6 ≥ slope ≥ -3.3).

2.2.12.8 RNA stability test

700 µl of an overnight culture was added to 10 ml of fresh CDM containing 0.1 mM or 10 mM Gly and incubated at 37°C, 5% CO_2 for 3 hours. 5 ml of sample was recovered, and *de novo* synthesis of RNA was stopped by the addition of 1 mg/ml rifampicin (Ramirez-Peña, et al., 2010). Samples were recovered after 1, 2, 5, and 10 min to monitor degradation of mRNA over time. Recovered samples were added to two volumes of RNAprotect, incubated at room temperature for 5 min. Samples were pelleted by centrifugation at 4000 rpm, 4°C for 5 min and quick frozen in liquid nitrogen. RNA was isolated from each sample as explained (2.2.12.1). Reverse Transcription was performed as described (2.2.12.4). qPCR was carried out as explained (2.2.12.5).

2.2.12.9 Northern blotting

2.2.12.9.1 RNA probe generation

   ♦   **Primer design considerations**

To detect mRNA, "antisense" probes were designed, i.e. complementary to the target mRNA. T7 RNA polymerase promoter site (see below) was added to the 3' end of the reverse primer.

cttaatacgactcactatagg

♦ ***In-vitro* transcription**

In-vitro transcription of gene of interest was performed using MEGAscript® Kit according to the kit manufacturer´s instruction. The DNA template was created by PCR as described (2.2.7). The PCR product was purified using PCR purification kit in accordance with the instructions of the manufacturer. The purified PCR product was used as T7-DNA template for *in-vitro* transcription.

♦ **RNA recovery**

The product of *in-vitro* transcription was purified using the RNA Clean & Concentrator™-5 kit according to the kit manufacturer´s instruction. The quantity of purified RNA was measured using Qubit®2.0 Fluorometer. The integrity of purified RNA was assessed using the Agilent G2565AA Scanner and an Agilent RNA 6000 Nano kit according to the kit manufacturer´s instruction.

♦ **Biotin labeling of the 3' terminal end of RNA probe**

Labeling reaction was performed using the Pierce™ RNA 3' End Biotinylation Kit according to the kit manufacturer 's instruction.

♦ **Biotin labeled probe purification**

Biotin labeled RNA probe was purified using the RNA Clean & Concentrator™-5 kit in accordance with the instructions of the manufacturer. Quality of purified RNA was assessed utilizing using the Agilent G2565AA Scanner and an Agilent RNA 6000 Nano kit according to the kit manufacturer´s instruction. Concentration was measured using Qubit®2.0 Fluorometer.

2.2.12.9.2   Denaturing agarose gel electrophoresis

♦ **Formaldehyde-agarose gel preparation**

Agarose (1.5%) was heated in 1X MOPS until dissolved. The gel was cooled to 60˚C. All subsequent steps were performed in a chemical hood. 2.2 M formaldehyde 37% was added to the solution. After gentle mixing, the gel was cast into the gel casting tray with an appropriate comb. After solidification, the comb was removed.

♦ **Sample preparation**

Total RNA was isolated as explained (2.2.12.1). Isolated RNA quality was tested using Agilent G2565AA Scanner. RNA integrity number 9-10 was accepted.

Samples were prepared by mixing a sufficient amount of RNA with the same volume of 2X RNA gel loading dye (Thermo Fisher Scientific Inc.) in a RNase free microfuge tube. Samples were incubated at 70°C for 10 min and immediately chilled on ice.

♦ **RNA marker preparation**

Millennium™ RNA Marker was thawed on ice and mixed well by pipetting. 3 µl of 2X RNA gel loading dye was mixed with 3 µl marker in an RNase free microfuge tube. The tube was incubated at 70°C for 10 min. It was chilled immediately on ice.

♦ **Pre-run**

The gel was pre-run submerged in 1x formaldehyde gel-running buffer for 30 min at 5V/cm in a chemical hood.

♦ **Main run**

Samples and ladder were loaded in to the wells. The gel was run in 1x formaldehyde gel-running buffer for 180 min at 5V/cm.

♦ **Buffers and solutions**

5X MOPS

Morpholinopropansulfonic acid (MOPS)	0.1 M
Sodium Acetate	40 mM
EDTA DEPC treated	5 mM
DEPC treated water	800 ml

pH 7.0 adjusted using 2 N NaOH

The solution was sterilized by filtration through a 0.2-micron Millipore filter, was stored at room temperature and protected from light.

1x formaldehyde gel-running buffer

1 x MOPS

Formaldehyde 37%   2.2 M

♦ **Visualizing the Formaldehyde/Agarose gel on a UV transilluminator**

Transparent ruler was aligned with the gel. The gel was photographed. The distance of migration from the loading well was measured.

2.2.12.9.3 Blotting

♦ **Installation of the Northern blotting stack**

20 Extra thick blot papers (PROTEAN®XL siz) were placed. 3 strips of Whatman 1 mm paper were cut to the width and length of the gel. Strips were placed over the stack. A strip of pre-wet Whatman paper with 10 X SSC was placed over the strips. Prewet Roti®-Nylon 0.2, pore size 0.2 µm was added. A pipette was rolled over the membrane to remove bubbles. Gel was placed on the membrane. A pipette was rolled over the gel. 3 prewet strips of Whatman 1 mm paper were placed over the gel. A bridge (a long stripe the width of the gel) of Whatman 1mm paper was prewet in 10X SSC. Both ends of the bridge were soaked in two containers located one on each side of the prepared stack. A piece of light plastic was placed on the top of the stack to prevent the bridge from drying. Transfer was allowed for 4 hours. Well positions were marked in ink. Upper part of the blot was marked by cutting.

20X SSC

NaCl	3.0M
Sodium citrate	0.3M

Adjusting pH 7.0

Ultra-pure water to 1 L

DEPC-treated with 1ml/L incubated overnight at room temperature.

The solution was sterilized by autoclave and stored at room temperature.

♦ **Immobilization**

Membrane was exposed to UV150 mJoule/cm^2 to immobilize RNA on the membrane using UV-cross linker.

2.2.12.9.4 Hybridisation

- ♦ **Step 1**

Prehybridization was performed by adding 5ml/100 cm^2 membrane of DIG Easy Hyb buffer to the membrane in a hybridization tube and incubated for one hour at 50°C while rotating constantly in a hybridization oven.

- ♦ **Step 2**

RNA probe (100 ng/ml hybridization solution) was added to DIG Easy Hyb buffer (3.5 ml/100cm^2 membrane). Prehybridization solution was poured off. RNA probe/DIG Easy Hyb buffer was immediately added to the membrane and incubated overnight at 50°C while rotating constantly in the hybridization oven.

2.2.12.9.5 Stringency wash

Hybridization solution was removed. The blot was washed twice in the hybridization tube with 50 ml of 2X Wash solution for 15 min per wash at room temperature. 100 ml of 0.5X Wash solution was heated to 65°C in a water bath. 50 ml of it was added to the membrane in the hybridization tube. The tube was placed in hybridization oven, which has been heated to 65°C, and incubated for 15 min. This step was repeated once.

2.2.12.9.6 Biotin detection

- ♦ **Blocking**

1 mL of 10% SDS was added to 9 mL Odyssey® Blocking Buffer (PBS) to a final concentration of 1% SDS. In a container, the blot was covered with the solution with gently shaking at room temperature for one hour.

- ♦ **Streptavidin incubation**

IRDye® IRDye 800CW Streptavidin was diluted with Odyssey Blocking Buffer (PBS) and 1% SDS to a concentration of 1:3000. The blocking buffer was removed and the blot was covered with the diluted IRDye Streptavidin solution. 5 mL of buffer/10 cm^2 of membrane was used. The container was incubated for one hour at room temperature with gentle agitation protected from light.

♦ **Washing**

The blot was washed 3 times in 1X PBS-Tween (0.1% Tween® 20) with shaking for 5 min each at room temperature and protected from light. The blot was rinsed with 1X PBS with shaking for 5 min at room temperature in darkness.

♦ **Scanning of the blot with the Odyssey® Imager**

The blot was scanned on an Odyssey® Imager according to the Odyssey Operator's Manual.

2.2.13   Eukaryotic cell culture

To study the adherence and internalization of *S. pyogenes* to human keratinocytes, the HaCaT cell line (human adult skin keratinocytes) was used.

2.2.13.1   HaCat cell line cultivation

HaCat cells were cultured in cell culture flask (75 cm^2, Bio-Greiner) containing DMEM (Dulbecco's modified Eagle's medium) supplemented with 10% (v/v) FCS (fetal calf serum) at 37°C and 5% CO_2. Media was renewed three times a week. For this purpose, DMEM medium was discarded from the cell culture flask by pipette and the bottle was subsequently washed using 2 ml of 0.25% trypsin/EDTA. A new 2 ml trypsin/EDTA solution was added to the flask and incubated at 37°C and 5% CO_2 for 10 min. It was checked for detaching cells from the inner surface of the flask. 8 ml of DMEM/FCS solution was added to the suspension. Afterward live cells were counted. 25 μl of cell suspension and 25 μl of trypan blue were placed on a cell-counting chamber and live cells were counted under the microscope (CKX41, Olympus). For following experiments monolayer cell line in 24 well-plate was needed. To create a monolayer of HaCat cells 2ml of 3.5×10^5 cells/ml suspension was added in a cell culture flask containing 15 ml of fresh DMEM/FCS and incubated at 37°C and 5% CO_2 for two to three days.

2.2.13.2 Preparing HaCat cell culture for adherence and internalization assay

A 3.5×10^5 HaCat cells/ml suspension was aliquoted 1 ml into the 24 well-plate (Greiner Bio one) with replicate number 3 for each bacterial strain and incubated at 37°C and 5% CO_2 overnight for following experiments.

2.2.14 Virulence-Relevant assays

2.2.14.1 Adherence and internalization assay

Adherence to and internalization into human keratinocytes cells was carried out using HaCaT cells (Nakata, et al., 2009). The assays were performed in 24-well plates (Greiner bio-one). Three wells were assigned for the adherence and three wells for the internalization assay for each strain. Cell culture plates were prepared as described (2.2.13.2). After 12-18 hours of incubation, the vitality and confluence of the cells was verified using the microscope (CKX41, Olympus).

Overnight culture (2.2.2.2) from the bacterial strain were transferred to non-selective THY medium (1:20). With an OD_{600} of 0.3, the sample were centrifuged at 4000 rpm for 10 min. The supernatant was discarded, and the pellet was washed with 5 ml PBS. The pellet was resuspended in DMEM and adjusted to an OD_{600} of 0.5. 1:10 (OD600 = 0.05) dilution was performed.

1 ml bacterial suspension in DMEM solution (OD_{600} = 0.05) were added to each well containing prepared confluent HaCaT cells in 24 well-plate (2.2.13.2). Three wells were assigned per strain. One well was used per strain for the B0 sample (bacterial solution without HaCaT cells). The 24-well plate was incubated for 2 hours at 37°C, 5% CO_2. The B0 control was collected. A serial dilution of 10^{-3} to 10^{-6} was made and 100 µl plated on non-selective THY agar plates using Eddy Jet (IUL Instruments). Subsequently, the plate was used for adherence or for the internalization assay:

Adherence assay:

The supernatants of the sample (test) were discarded. The wells carefully were washed with PBS. 200 µl of 0.25% trypsin-EDTA (1x) was added to each well and incubated 10 min at 37°C ,5% CO_2. The HaCaT cells from the wells were collected in a 2-ml reaction tube. The wells were washed with a total 600 µl of 1xPBS and added to the 2ml tube.

This was followed by centrifugation for 2 min at 10,000 rpm. The supernatant was discarded and 1ml ultra pure water was added to the tube. The cells were lysed by constant shaking for 10 min. Serial dilution of the suspension was made and 100 µl of 10^{-3} to 10^{-6} were plated on non-selective THY agar plates using Eddy Jet (IUL Instruments) and incubated at 37°C, 5% CO_2 for 24 hours.

Internalization assay:

The supernatant of each well was discarded and replaced with 1 ml of fresh DMEM and 10 µl of Penicillin/Streptomycin (5000 U/ml) solution per well. A further incubation was

carried out for 2 hours at 37°C and 5% CO_2. The HaCaT cells were treated as the adherence assay. Serial dilution of 10^{-1} to 10^{-4} was prepared. 100 µl of each sample were plated on none-selective THY agar plates using Eddy Jet (IUL Instruments) and incubated at 37°C, 5% CO_2 for 24 hours. Analyzing the adherence and internalization was carried out through the ratio of colony-forming units (CFU) in comparison to the B0-sample.

2.2.14.2 Blood survival assay

To study the survival of bacteria in human blood 1 ml overnight culture (2.2.2) of the bacterial strain was inoculated into 19 ml of THY broth. Culture was incubated at 37°C and 5% CO_2 to early exponential growth phase ($OD_{600} \approx 0.3$). A dilution of 1: 10,000 was performed in PBS. 100 µl of dilution was plated on the THY agar plate as a control using Eddy Jet (control). The test was preceded by inoculating 20 µl of prepared dilution to 480 µl of heparinized blood. The experiment was carried out in triplicate using blood from three independent donors. Inoculated blood was incubated for 3 hours at 37°C and 800 rpm. 100 µl of samples was plated on THY agar plate using Eddy Jet (IUL Instruments). Plates were incubated at 37°C and 5% CO_2 for 24 hours. Viable colony forming bacteria (CFU) after 3 hours incubation in blood considering the initial quantity of bacteria (control) gave the multiplication factor as a measure of survival ability of the bacteria (Nakata, et al., 2009).

2.2.14.3 Oxidative stress

500 µl overnight culture (2.2.2) of the bacterial strain was inoculated into 9 ml of THY broth. The culture was incubated at 37°C and 5% CO_2 to early exponential growth phase ($OD_{600} \approx 0.5$). Hydrogen peroxide solution was prepared by adding 1µl H_2O_2 to 99 µl DEPC treated water (1:100 dilution). 3 tubes containing 20 ml of bacterial suspension were added to (control), 0.5 and 1 mM (tests) H_2O_2 respectively. Tubes were incubated at 37°C and 5% CO_2 for 2hours. 1ml of each tube was centrifuged at 4500rpm, for 10 min. Pellets were washed in 1 ml PBS and resuspended in 1ml PBS. Serial dilution was prepared from each sample and100 µl of 10^{-4}, 10^{-5} and10^{-6} dilutions were plated in THY agar plates. CFUs were counted after 24 hours.

3 Results

♦ Part I

3.1 Investigating Gly as a possible inducer for a putative Gly riboswitch

Among 22 *cis*-regulatory RNAs in *S. pyogenes* M49 (Patenge, et al., 2012) a putative Gly riboswitch (*ribogly*) was identified which is located in the 5′ -UTR of the putative Na$^+$/Alanine symporter gene that is preceding a putative cation efflux system protein gene (Figure 1-7).

The secondary structure of ribogly was predicted with the "RNAfold web server" (Gruber, et al., 2008). The secondary structure in dot-bracket notation was imported into VARNA. VARNA is a tool for the automated drawing, visualization and annotation of the secondary structure of RNA designed as companion software for web servers and databases (Darty, et al., 2009). The predicted ribogly secondary structure in *S. pyogenes* is shown in Figure 3-1.
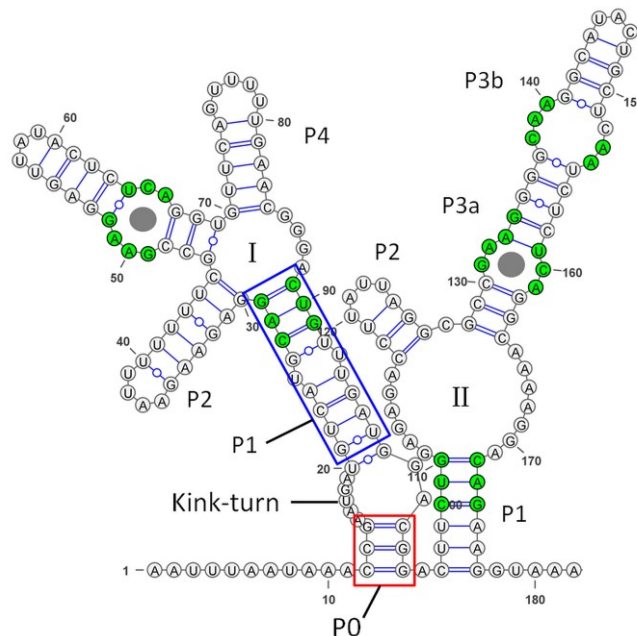
Figure 3-1. Secondary structure of *ribogly* from *S. pyogenes* M49 was predicted by RNAfold and illustrated using VARNA GUI (Darty, et al., 2009). It is composed of aptamer (I) and aptamer (II). Conserved sequences are represented in green (Esquiaqui, et al., 2014; Mandal, et al., 2004; Ruff, et al., 2016). Glycine binding sites are indicated by grey circles. P0 and P1 are boxed in red and blue respectively, following the observations from the glycine riboswitch in *V. cholerae* (Esquiaqui, et al., 2014). P1 through P4 represent common base paired elements.

The putative Gly riboswitch (*ribogly*) in *S. pyogenes* M49 was fused to a luciferase gene as a reporter. *S. pyogenes* M49 was transformed with the plasmid carrying the construct PFW11-*ribogly-luc2* (Figure 3-2.A) (Popp, 2015).
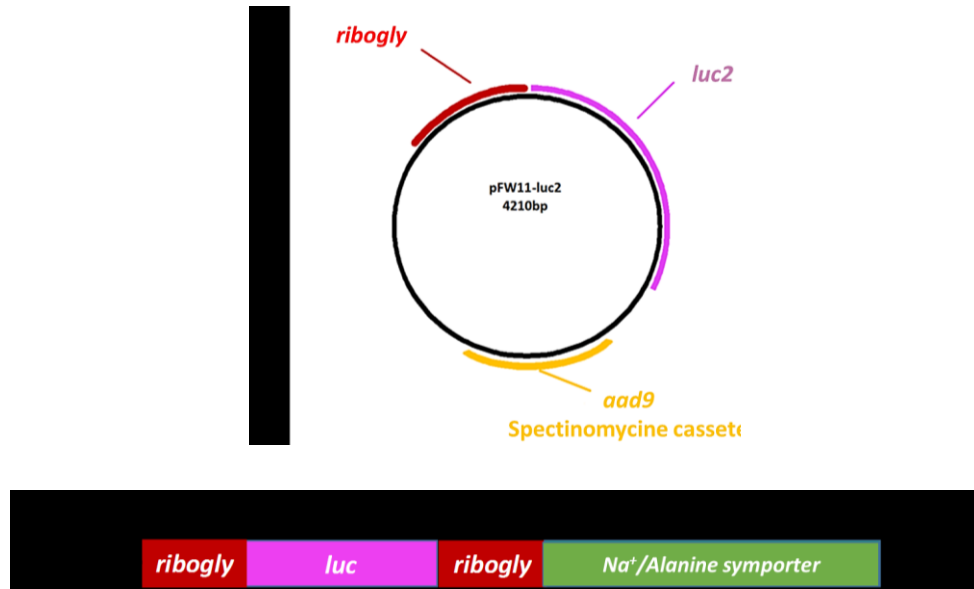

Figure 3-2. Schematic illustration of A) pFW11-*ribogly-luc2* B) the location of *ribogly* in the genome of *S. pyogenes* M49 after transformation with pFW11-*ribogly-luc2*.

As first step in this study, the results of the former study were verified. Growth behavior of *S. pyogenes* M49/pFW11-*ribogly-luc2* was monitored in CDM deficient in Gly and in CDM with different concentrations of Gly (0.01, 0.1, 1, 2.6,10 mM). Growth of *S. pyogenes* M49/pFW11-*ribogly- luc2* in CDM in the absence of Gly and in 0.01 mM concentration of Gly was poor. Growth behavior was improved at higher concentrations of Gly (Figure 3-3).
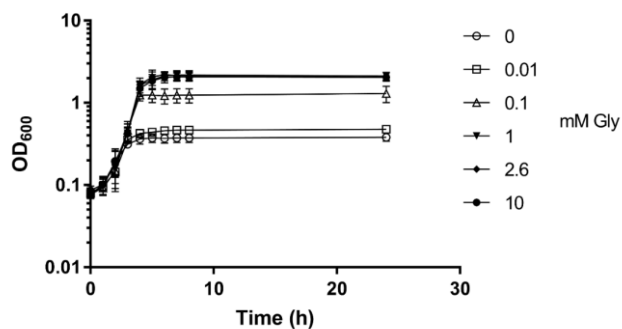

Figure 3-3. Growth curves of *S. pyogenes* M49/pFW11-*ribogly-luc2* in CDM containing different concentrations of Gly. In the absence of Gly and at 0.01 mM Gly, growth of *S. pyogenes* M49/pFW11-*ribogly-luc2* was poor (n=3). The data are presented as the mean values ± standard deviation.

Gly was reported as an essential amino acid (Slade, et al., 1951). In contrast, Levering and colleagues observed that repeated re-inoculation of cultures grown in the absence of glycine or serine, led to adaptation to the omissions and significant growth of the bacteria in the absence of these amino acids (Levering, et al., 2016). Glycine and serine can be converted into one another by glycine hydroxymethyltransferase (EC 2.1.2.1). According to the biochemical and metabolic database, KEGG (Kanehisaa & Goto, 2000), a glycine hydroxylmethyltransferase gene (Spy49_0898) is annotated in *S. pyogenes*. However, to date the expression of the glycine hydroxymethyltransferase gene and the function of its product has not been studied in *S. pyogenes*.

For the expression analysis of the luciferase gene, emission of bioluminescence was measured using the Luminometer at the same time points as the $OD_{600\,nm}$ measurements. Luciferin served as a substrate. To compensate the differences in growth behavior in different Gly concentrations, light emission data detected by luminometer were divided by the corresponding OD_{600} to determine the relative light units per OD_{600} (RLU/ OD_{600}). After 3 hours of growth in CDM containing 0.01 mM Gly, the highest induction of ribogly was observed with 600,000 RLU/OD_{600} compared to 8000 RLU/OD_{600} following growth in CDM containing 10 mM Gly. Overall, luminescence of bacteria grown in CDM containing 0, 0.01, or 0.1 mM Gly was always increased compared to luminescence of bacteria grown in CDM containing 1, 2.6 or 10 mM Gly. Wild type *S. pyogenes* M49 was used as a negative control (Figure 3-4). ribogly RNA was specific for Gly and did not respond to neither serine nor alanine (Popp, 2015).

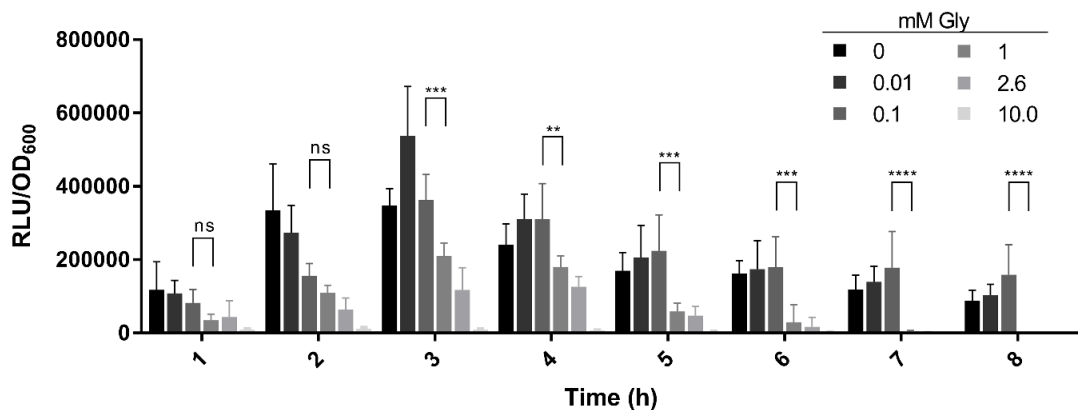Figure 3-4. Relative light units (RLU/OD_{600}) were measured in *S. pyogenes* M49/pFW11-*ribogly*-luc2 in CDM in the absence of Gly and in CDM containing increasing Gly concentrations. (Two-way ANOVA multiple comparisons test: <u>n</u>ot <u>s</u>ignificant $P \geq 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $P \leq 0.0001$, n=5). The data are presented as the mean values ± standard deviation.

3.1.1 Transcription analysis of *luc2* in *S. pyogenes* M49/pFW11-*ribogly-luc2*

S. pyogenes M49/pFW11-*ribogly-luc2* were grown in CDM containing three different concentrations of Gly (0.1 mM, 2.6 mM, and 10 mM Gly) for 3 hours. Total RNA from each sample was isolated as explained (2.2.12.1). DNase treatment and cDNA synthesis were performed as described in 2.2.12.2 and 2.2.12.4 respectively. Transcript analysis was conducted via RT-qPCR as explained (2.2.12.5) to determine the transcript level of *luc2* in *S. pyogenes* M49/pFW11-*ribogly-luc2* and its results were compared to the results of the luciferase assay. All primers used for RT-qPCR are listed in Table 8. 5S rRNA served as internal control. Relative expression was calculated as described in 2.2.12.6. Transcript level of bacteria grown in CDM complemented with 0.1 mM Gly served as calibrator.

In comparison to the transcript level of *luc2* at 0.1 mM Gly, the transcript level decreased dramatically at high concentrations of Gly (2.6 mM and 10mM) (Figure 3-5). These results show that the reporter gene assay reflects the transcriptional regulation mediated by *ribogly.*
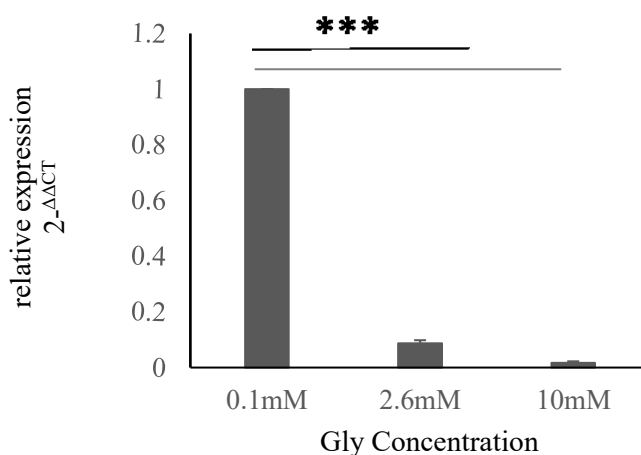
Figure 3-5. Transcript level of *luc2* in *S. pyogenes* M49/pFW11-*ribogly-luc2* grown in CDM complemented with different Gly concentrations (t-test; *** *p ≤ 0.001*, n=3). The data are presented as the mean values ± standard deviation.

3.2 Transcription analysis via RT-qPCR

In *B. subtilis,* binding of Gly to the Gly riboswitch leads to the formation of an anti-terminator structure and synthesis of the full-length transcript (Phan & Schumann, 2007).

To investigate whether *ribogly* in *S. pyogenes* M49 functions similarly, transcription analyses were performed by RT-qPCR. The transcriptional level of *ribogly* and the Na^+/Alanine symporter gene were determined separately. Furthermore, the putative cation efflux system gene, which is located immediately downstream of the Na^+/Alanine symporter gene, was analyzed to investigate whether both genes are transcribed to a bi-cistronic message. Transcription analyses were performed using total RNA of wild-type *S. pyogenes* M49.

3 sets of primers were designed according to the arrangement of *ribogly* and downstream genes. To detect attenuation products and the full-length transcript two pairs of primers were needed. One set was designed to amplify the putative riboswitch according to the Rfam prediction (Figure 3-6. A). This amplicon is part of the truncated as well as of the full-length transcript, because it is located in the 5′of the putative terminator structure. The second primer pair was designed corresponding to the symporter gene and it is specific for the full-length transcript (Figure 3-6.B). A third primer pair was selected within the putative cation efflux system protein gene to determine any response to the ligand (Figure 3-6. C). All primers used for RT-qPCR are listed in Table 8.
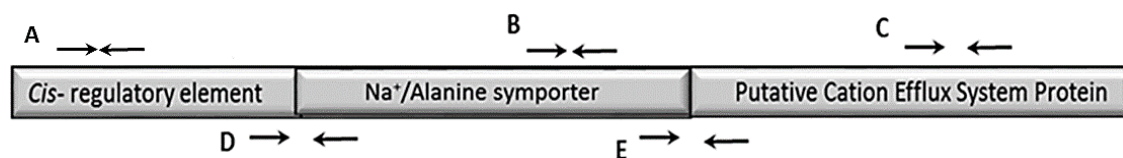
Figure 3-6. Schematic representation of the location of primers to study the transcription attenuation model by transcription analyses of *ribogly*, the Na^+/Alanine symporter gene, and the putative cation efflux system protein gene.

GAS strains were grown in CDM at the same condition as described in 3.1.1. Total RNA was extracted as explained (2.2.12.1). Transcript level of target genes were calculated as described (2.2.12.6) compared to 5S rRNA (Figure 3-7). Transcript abundance of *ribogly* remained static at different concentrations of Gly (Figure 3-7). The Na^+/Alanine symporter and cation efflux system protein transcript levels decreased drastically at 2.6 mM and 10 mM Gly (Figure 3-7).
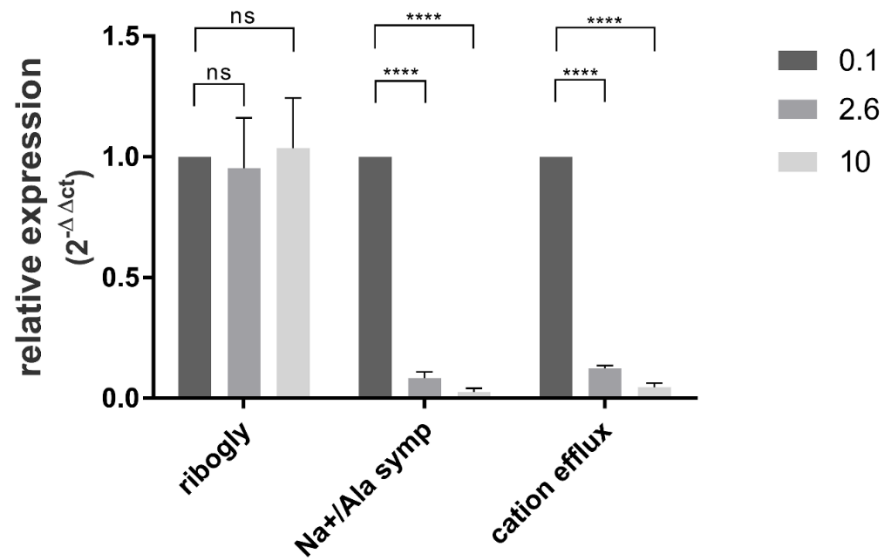
Figure 3-7. Transcript analysis of ribogly, Na^+/Alanine symporter gene, and cation efflux system gene. Data are presented as target genes transcript-levels (that are assigned in the graph with their name) relative to the amount of transcript at 0.1 mM Gly concentration compared to 5S rRNA. (Two-way ANOVA, multiple comparisons; <u>not</u> <u>significant</u> $P \geq 0.05$, $****\, p \leq 0.0001$, n=4). The data are presented as the mean values ± standard deviation.

To verify transcription attenuation, the transcript level of Na^+/Alanine symporter gene was calculated compared to ribogly as described (2.2.12.6) (Figure 3-8). This result shows that the short transcript predominates in all concentrations of Gly and the ratio of the full-length transcript to the truncated product dramatically decreased upon addition of Gly.
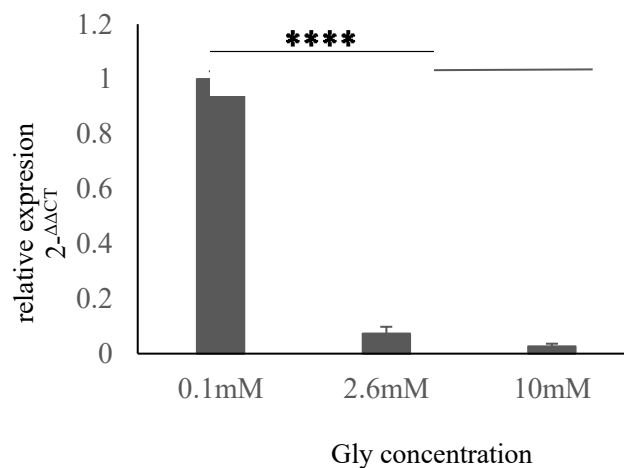

Figure 3-8. Transcript analysis of the Na^+/Alanine symporter gene. Data are presented as target gene transcript-level relative to the amount of transcript at 0. 1 mM Gly compared to ribogly (t-test; $****\, p \leq 0.0001$, n=4). The data are presented as the mean values ± standard deviation.

3.2.1 Analysis of the operon transcript

Transcript levels of the Na$^+$/Alanine symporter gene and the putative cation efflux system protein gene were repressed in the presence of high Gly concentrations (Figure 3-7). Following experiment was performed to clarify whether the cation efflux system protein gene is transcribed as a part of a bi-cistronic mRNA together with the Na$^+$/Alanine symporter gene. Two sets of primers to amplify gene junctions were designed: one at the junction between the putative riboswitch and the Na$^+$/Alanine symporter gene as a control for experiment (Figure 3-6. D) and another at the junction between the Na$^+$/Alanine symporter gene and the putative cation efflux system protein gene (Figure 3-6. E). cDNA synthesis was performed and followed by PCR as described in 2.2.12.4 and 2.2.9 respectively. Detection of the PCR product suggested co-transcription of the two genes (Figure 3-9). NRT control (no reverse transcriptase control) was included in the test as a control using primer set E. Sequencing data of the PCR product support the existence of a bi-cistronic message (Appendix 7-9). All primers used for this step are listed in Table 8.
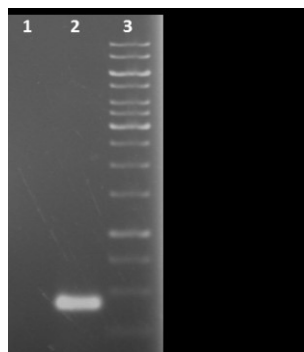
Figure 3-9. PCR followed by agarose gel electrophoresis to verify whether the cation efflux system protein gene is transcribed as a part of a bi-cistronic mRNA together with the Na$^+$/Alanine symporter gene. Lane 1; NRT control using RNA sample isolated from *S. pyogenes* M49 at 0.1 mM Gly and primer set E. (location of the primer was shown in Figure 3-6). Lane 2; PCR product was amplified from cDNA from *S. pyogenes* M49 at 0.1 mM Gly using primer set E. Lane 3; O´ GeneRuler™ 1 kb DNA Ladder.

To prove that the putative riboswitch and Na$^+$/Alanine symporter are transcribed from the same operon, RT-PCR was carried out using primer pair D (Figure 3-6). The forward primer is specific for a part of the putative riboswitch, located on the aptamer domain, and the reverse primer is specific for the Na$^+$/Alanine symporter gene. Amplification of

the product determined that riboswitch and the Na$^+$/Alanine symporter gene are transcribed from the same operon. No product could be detected at a high concentration of Gly (10 mM) (Figure 3-10).
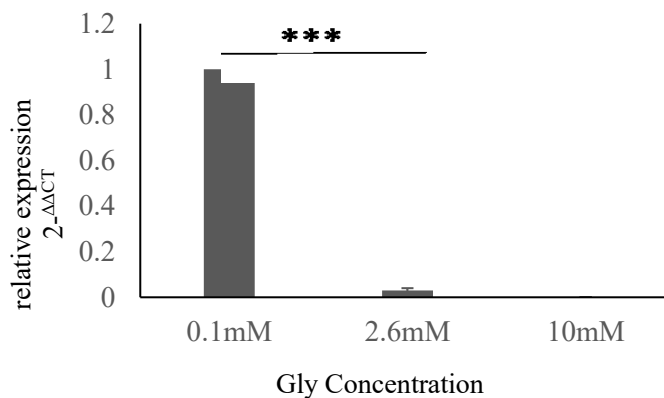
Figure 3-10. Relative expression of the Na$^+$/Alanine symporter gene (using primer pair D shown in Figure 3-6) compared to ribogly (t- test; *** $p \leq 0.001$, n=3). The data are presented as the mean values ± standard deviation.

3.3 Transcription analysis via Northern blotting

To confirm the results achieved by RT-qPCR (3.2) and to prove the hypothesis of transcription attenuation caused by Gly riboswitch at high Gly concentrations (2.6mM and 10mM), Northern blot analyses were performed. The total RNA was extracted as explained (3.1.1). 6 μg of each sample was used for the experiment. Northern blotting was performed as explained (2.2.12.9). Blots were probed by antisense RNA specific for ribogly and were stripped and re-probed with Na$^+$/Alanine symporter-specific probe (Figure 3-11). At low concentration of Gly (0.1 mM) a 3 kb band was detected by both probes. This size corresponds to the size of a co-transcript of the Na$^+$/Alanine symporter gene and the cation efflux system protein gene. In low amount of Gly (0.1mM), both probes detected a sharp band about 600 bp. This apparent size corresponds the product of a termination at the putative terminator TERM 125 within Na$^+$/Alanine symporter gene (TransTermHP v2.07) (Kingsford, et al., 2007). A highly aboundant 200 bp band was detectable by the ribogly probe in the presense of high Gly concentrations (2.6 and 10 mM). This could arise from premature transcription termination or internal processing. At high concentrations of Gly no Na$^+$/Alanine symporter gene transcript was detected by ribogly probe. A 300 bp band was detected only using the ribogly probe. These results fully confirm the data of *luc* reporter assay (Figure 3-4) and RT-qPCR (Figure 3-7). Gly

riboswitch locus and transcript sizes as well as the location of the predicted terminators are illusterated in Figure 3-12.
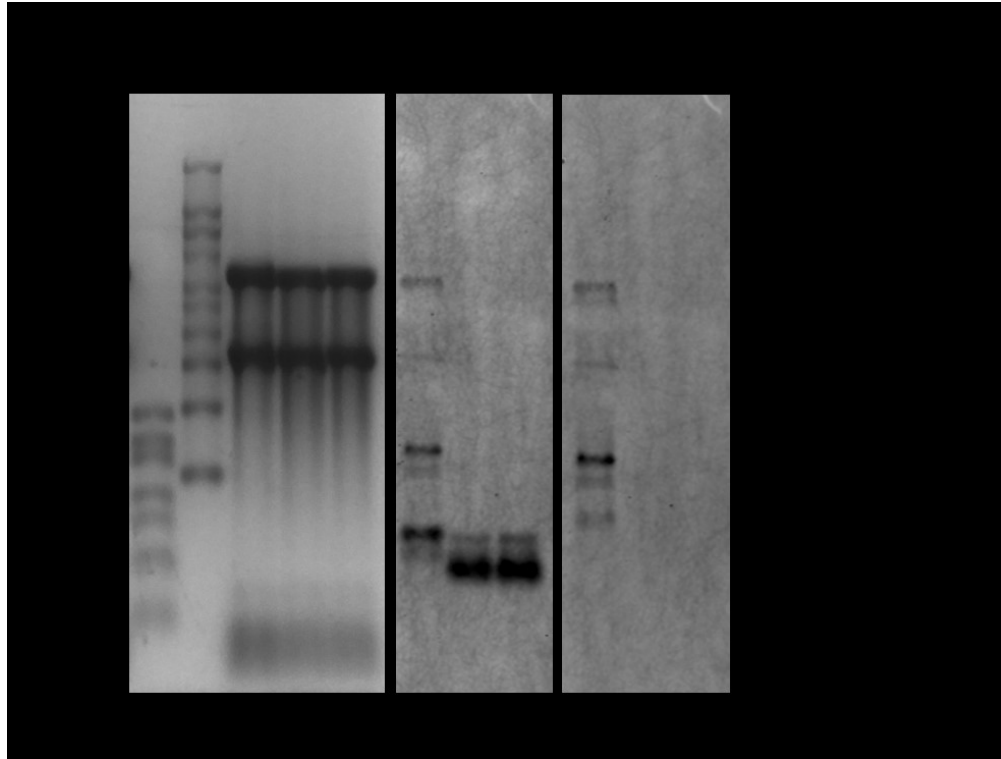
Figure 3-11. Northern blot analyses. Total RNA was isolated from *S. pyogenes* M49 grown in CDM complemented with 0.1mM, 2.6mM and 10 mM Gly concentrations. 6 μg of RNA were loaded per well. The Northern blot was probed with antisense RNA complementary to the Na$^+$/Alanine symporter and to the ribogly. Transcript size was determined by comparison with RNA size marker M1; Thermo Scientific RiboRuler Low range RNA Ladder and M2 Ambion Millenium ^{TM} RNA Marker.
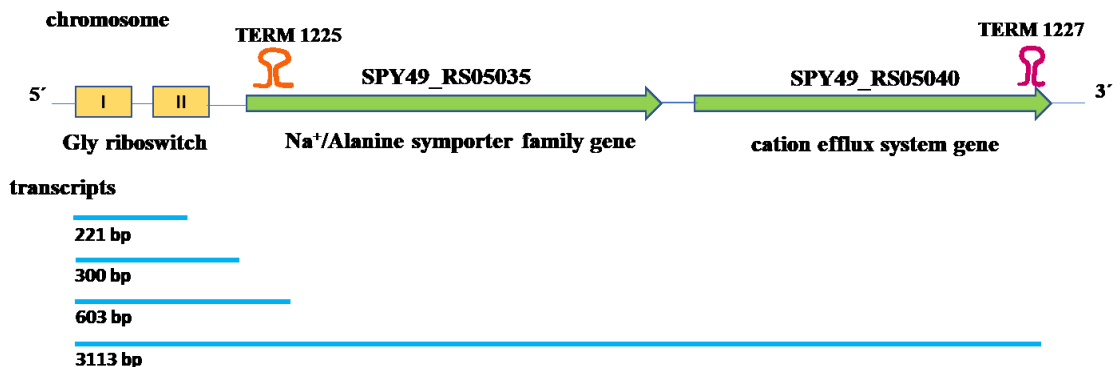
Figure 3-12. Schematic representation of the Gly riboswitch locus. Gly riboswitch aptamers I and II are boxed in yellow. Genes are depicted as green arrows, pointing into the direction of transcription. Terminators 1225 and 1227 predicted by TransTermHP (Kingsford, et al., 2007) are drawn as stem-loop structures. Transcripts detected by Northern blotting are shown as blue lines.

3.4 Transcript stability

Decrease of the Na^+/Alanine symporter mRNA level in the presence of 10 mM Gly suggests the formation of truncated product caused by transcription termination. Another known mechanism of riboswitches for regulation of gene expression in gram-positive bacteria is induction of mRNA degradation (Carona, et al., 2012). To investigate the stability of *ribogly* and Na^+/Alanine symporter transcript under repressing and non-repressing conditions, following experiment was performed.

S. pyogenes M49 was grown in CDM complemented with two different concentrations of Gly (0.1 mM and 10 mM). After 3 hours of growth, one sample from each concentration was taken (time point = 0) and cultures were treated with rifampicin to inhibit synthesis of RNA. Samples were collected after 1, 2, 5 and 10 min following the addition of rifampicin. Procedure was explained in detail in 2.2.12.8. RT-qPCR was conducted using primer pairs indicated in Figure 3-6. 5S rRNA and *groES* mRNA served as controls. Samples at time point zero were considered as calibrators.

5S rRNA was stable over time under both conditions. The *groES* mRNA transcript was degraded at similar rates in the presence of 0.1 mM and 10 mM Gly (half-life 1.8 min and 1.9 min, respectively). The *ribogly* transcript was highly unstable in 0.1 mM and 10 mM of Gly (half-life 0.4 min under both conditions). The stability of the Na^+/Alanine symporter mRNA transcript was drastically reduced at 10 mM Gly (half-life 0.5 min). It was highly stable in 0.1 mM Gly (half-life 4.2 min) (Figure 3-13). Half-life of the respective transcripts has been determined by non-linear regression (least square) curve-fitting using GraphPad Prism.

A

B
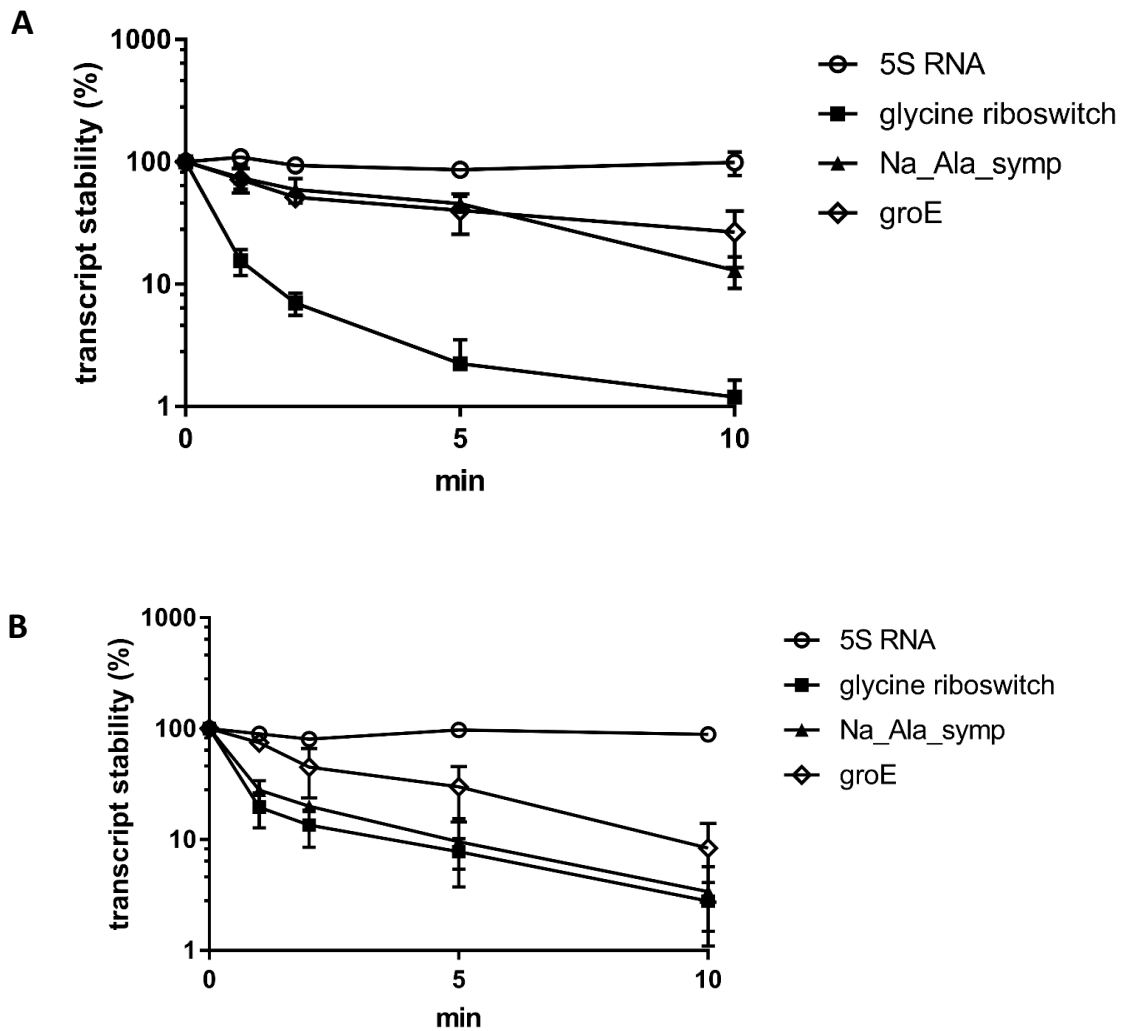
Figure 3-13. Transcript stability of ribogly, Na$^+$/Alanine symporter (tests), 5S rRNA and *groES* (controls) under low (0.1 mM) and high (10 mM) Gly concentrations via RT-qPCR. Total RNA was isolated after rifampicine treatment. Data are presented as percent transcript level relative to time point zero (n=3). The data are presented as the mean values ± standard deviation.

3.5 Investigation of the Na$^+$/Alanine symporter gene promoter activity

To investigate whether the glycine-dependent repression is influenced by the native promoter, a 588 bp fragment carrying the promoter region 5' of *ribogly*, was fused to the luciferase reporter gene. Repression at different concentrations of Gly was assessed via RT-qPCR through measuring the transcript level of *luc2*.

3.5.1 Construction of *S. pyogenes* M49 /pFW11- promoter*ribogly*luc2

Endogenous promoter was predicted via online tools; PePPER (de Jong, et al., 2012) and BDGP (Adams, et al., 2000). The predicted promoter region (Appendix 7-6) was amplified using designed primer pair (Table 8) and fused to *luc2* in pFW11-*luc2* plasmid (Figure 3-14). The resulting plasmid was tested by restriction digest (Figure 3-15). *S. pyogenes* M49 was transformed with pFW11- promoter*ribogly*luc2. The recombinant strain was verified by sequencing (Appendix 7-8).
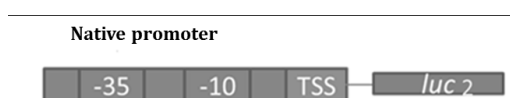
Figure 3-14. Schematic representation of pFW11- promoter*ribogly*luc2 construct.
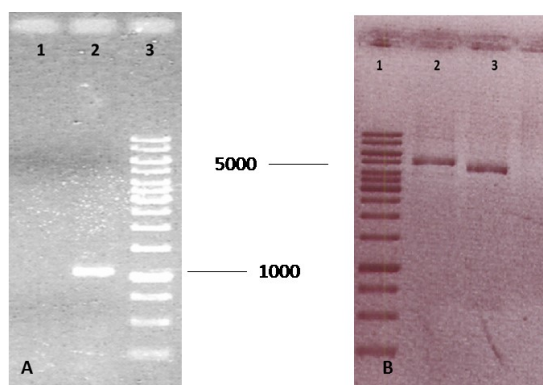
Figure 3-15. A: Verification of *S. pyogenes* M49/pFW11- promoter*ribogly*luc2. Lane 1; *S. pyogenes* M49 control. Lane 2; PCR fragment amplified with primer set Ribo.prom. BamHl-fwd and luc-check-rev using transformant chromosomal DNA as template. Lane 3; O´ GeneRuler™ 1 kb DNA Ladder. B: Verification of the recombinant pFW11-native prompter-*luc2* plasmid through restriction digest. Lane 1; O´ GeneRuler™ 1 kb DNA Ladder. Lane 2; Recombinant plasmid after BamHI restriction digest. Lane 3; Recombinant plasmid after BamHI and XhoI restriction digest created pFW11–*luc2* and promoter*ribogly*.

3.5.2 Transcript analysis of *luc2* in *S. pyogenes* M49/ pFW11- promoter*ribogly*luc2

The transcript-level of *luc2* in *S. pyogenes* M49/ pFW11- promoter*ribogly*luc2 was assessed via RT-qPCR, following growth in CDM with varying concentrations of Gly as explaind

(3.1.1). Transcript levels of *luc2* showed no Gly-dependent variation. The promoter in this construct did not respond to the different concentrations of Gly in the growth medium (Figure 3-16).
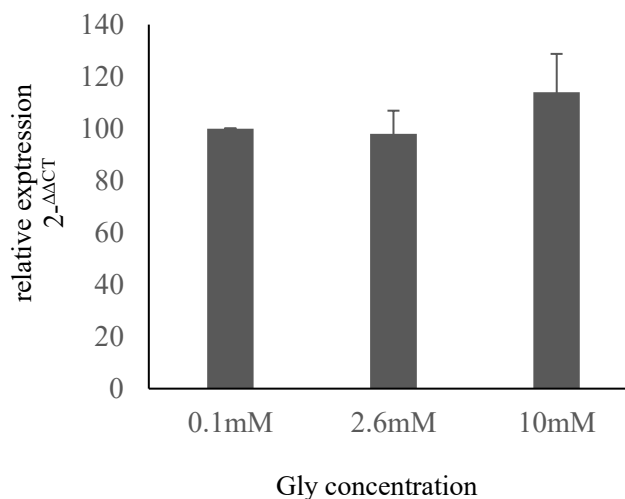
Figure 3-16. Transcript analysis of *luc2* in *S. pyogenes* M49/ pFW11- promoter$_{ribogly}$*luc2* (n=3). The data are presented as the mean values ± standard deviation.

The transcript level of *luc2* in the absence of the riboswitch aptamers remains consistently (Figure 3-16). The transcript-level of *luc2* in *S pyogenes* M49/ pFW11-promoter$_{ribogly}$*luc2* construct was compared to the transcript level of *luc2* in *S. pyogenes* M49/pFW11-*ribogly-luc2* transformant at non-repressing level of Gly (Figure 3-17). Results showed *luc2* transcript level in the absence of riboswitch aptamers is significantly higher than in the presence of the riboswitch aptamers upstream of the native promoter.

This outcome is explained by an influence of *ribogly* on the promoter. Results indicate that Gly-dependent expression of the genes located downstream of *ribogly* is not caused by an inducible promoter.
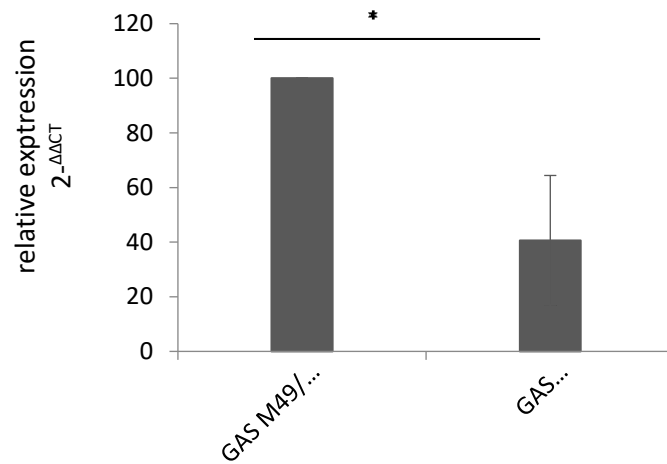
Figure 3-17. Comparing transcript level of *luc2* in *S. pyogenes* M49/pFW11- promoter$_{ribogly}$*luc2* and *S. pyogenes* M49/pFW11-*ribogly-luc2* (t-test; * *p ≤ 0.05*, n=3). The data are presented as the mean values ± standard deviation.

♦ **Part II**

3.6 Construction of *S. pyogenes* M49 *cand34* deletion mutant

cand34 is an sRNA annotated by NCBI as CRISPR RNA belonging to CRISPR I-C and is located adjacent to CRISPR associated protein genes (*cas* genes). *ccdA* is located upstream of *cand34,* which is annotated as cytochrome C biogenesis protein gene. Pictorial representation of CRISPR I-C and upstream gene is provided in 1.9 .

To study the function of *cand34* in *S. pyogenes* M49 strain 591, a *cand34* deletion mutant was constructed by replacing *cand34* in *S. pyogenes* M49 chromosome with the spectinomycin cassette from pFW11 vector (Appendix 7-3). 5′ (F1) and 3′ (F2) flanking regions of *cand34* were amplified by PCR (2.2.9). Related primer pairs are listed in Table 8. The experiment was followed by restriction digests by BamHl/Sall restriction enzymes (for amplified F1 fragment) and Ncol/Pstl restriction enzymes (for amplified F2 fragment) 2.2.11.1). Ligation was performed as described (2.2.11.2). To prove that F1 and F2 fragments are flanked the spectomycin cassette in PFW11 plasmid, restriction digest tests were performed 2.2.11.1). Verification of the recombinant plasmid is shown

in Figure 3-18. *S. pyogenes* M49 was transformed with the recombinant plasmid (2.2.11.4). *cand34* deletion mutant (*Δcand34)* was constructed by homologous recombination. Chromosomal *cand34* was replaced by a spectinomycin cassette from pFW11-F1-F2. *Δcand34* was selected on a selective THY agar plate (containing spectinomycin) and verified by test PCR (2.2.9) (Figure 3-19). To prove that F1 and F2 were inserted properly, PCR was performed using related primer pairs to amplify F1-*aad9* and *aad9*-F2 (Table 8). Additional test PCRs are described in the legend of Figure 3-19. The construct was also confirmed by Sanger sequencing (Appendix 7-10, Appendix 7-11).
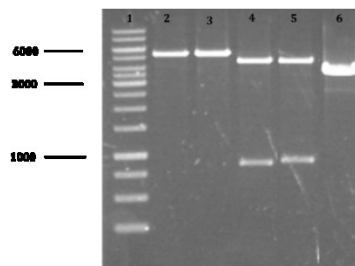
Figure 3-18. Representative agarose gel electrophoresis of the restriction digest products of the recombinant pFW11-F1-F2 plasmid. Lane 1; O´ GeneRuler™ 1 kb DNA Ladder. Lane 2; pFW11-F1-F2 construct digested by BamHl. Lane 3; pFW11-F1-F2 construct digested by Pstl. Lane 4; pFW11-F1-F2 construct digested by BamHl/Sall and results in pFW11-F2 and F1 fragment. Lane 5; pFW11-F1-F2 construct digested by Ncol/Pstl and results in pFW11-F1 and F2. Lane 6; pFW11digested by BamHl.
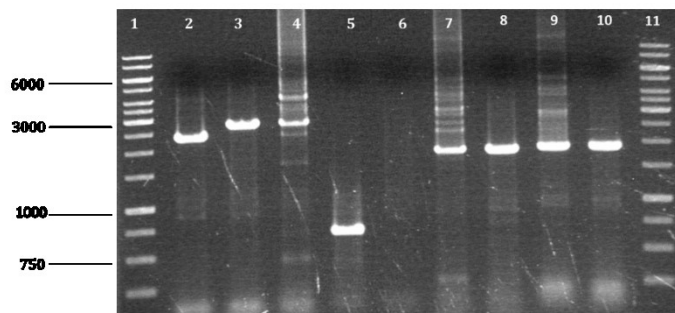
Figure 3-19. Verification of *cand34* deletion in GAS M49 *Δcand34* by PCR. Lane1&11: O´ GeneRuler™ 1 kb DNA Ladder. Lane2; amplified F1-*cand34*-F2. lane3&4; amplified F1-*aad9*-F2 fragment from PFW11-F1-F2 construct and *Δcand34*. Lane5&6; amplified *cand34* in WT and *Δcand34* respectively. Lane7&8; amplified F1-aad9 fragment from PFW11-F1-F2 construct and *Δcand34* respectively. Lane9&10; amplified F2-aad9 fragment from PFW11-F1-F2 construct and *Δcand34* respectively.

A complementation strain of *cand34* was constructed. The complete *cand34* sequence was predicted via online tool PePPER (de Jong, et al., 2012) and inserted into the shuttle vector pAT19 (Appendix 7-4) as explained (2.2.11). *Δcand34* was transformed with

pAT19_*cand34*. Verification of pAT19-*cand34* construct via restriction digest test is shown in Figure 3-20. Verification via sequencing is presented in Appendix 7-12. If the mutant phenotype was caused by the lack of *cand34* expression, ectopic expression of *cand34* in the complementation strain should restore the WT (wild-type) phenotype. In the case of mutations which are unrelated to the *cand34* deletion, the phenotype will not be restored. An overexpression strain (M49/pAT19_*cand34)* was constructed via transforming pAT19_*cand34* into the wild-type strain.
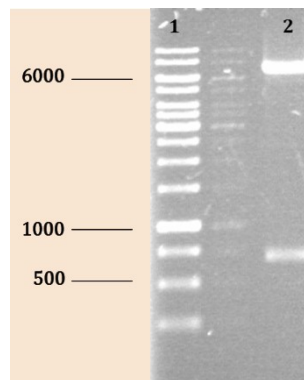
Figure 3-20. Verification of the pAT19-*cand34* construct through restriction digest. Lane 1; O´ GeneRuler™ 1 kb DNA Ladder. Lane 2; Recombinant plasmid after BamHI/SalI restriction digest created linear pAT19 and *cand34* fragment.

3.6.1   Transcript analysis of *cand34* and its upstream and downstream genes

Transcript analysis of *cand34* and its upstream and downstream genes (*ccdA* and *cas2* respectively) was performed to verify if deletion of *cand34* caused any changes in the expression of upstream and downstream genes. The location of genes was shown in Figure 1-9. *ccdA* is annotated as Cytochrome C biogenesis protein and its function in *S. pyogenes* has not been studied yet. As mentioned previously, *cas2* is a CRISPR associated protein. The expression of *cand34* and *ccdA* and *cas2* was verified via RT-qPCR 2.2.12.5) in WT, Δ*cand34*, complementation strain (Δ*cand34::cand34*) and the overexpression strain (M49/pAT19_*cand34)*. 5S rRNA served as internal control. Transcript level of target genes were calculated as described 2.2.12.6). In the deletion mutant strain, the *cand34* transcript was absent. Transcript-level of *cand34* in the complementation strain and in the overexpression strain was drastically higher compared to the WT. Results are shown in Figure 3-21.A. No transcript level of *ccdA* (the upstream gene of *cand34*) was detected in the WT. In *Δcand34* and *Δcand34::cand34,* a high level

of *ccdA* transcript was observed (Figure 3-21.B). The transcript level of *cas2* showed no significant difference between the tested strains (Figure 3-21.C).
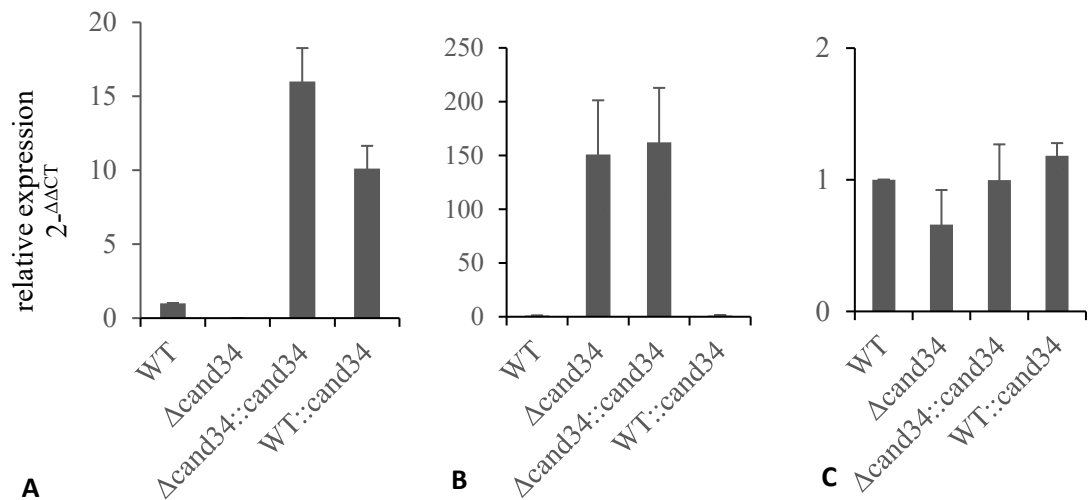
Figure 3-21. Transcript analysis of *cand34* (A)*, ccdA* (B), and *cas2* (C) in WT, Δ*cand34*, *Δcand34::cand34,* and M49/pAT19_*cand34* (n=3). The data are presented as the mean values ± standard deviation. Data are presented as target genes transcript levels relative to the amount of transcript level in WT compared to 5S rRNA.

3.6.2   Output from CRISPR finder and blast on the NCBI

As mentioned (1.9) *cand34* is annotated as CRISPR RNA in CRISPR I-C (Dvulg or CASS1) in *S. pyogenes*. Analyses of *cand34* sequence with CRISPR web server (http://crispr.i2bc.paris-saclay.fr/) (Grissa, et al., 2007) proved its repeat-spacer structure as a CRISPR RNA (Figure 3-22). Blast output results for 5 spacers in *cand34* shown in Table 16 revealed that spacers matched streptococcal phages and prophage regions in different *S. pyogenes* strains. In *S. pyogenes* NZ131 spacers are involved in CRISPR system and none of them matched chromosomal gene sequences in this microorganism. This outcome means autoimmunity by self-targeting as was reported previously (Stern, et al., 2010) did not happened in *S. pyogenes*.

CRISPR id : tmp_1_Crispr_1

- CRISPR start position : 109 ---------- CRISPR end position : 473 ---------- CRISPR length : 364
- DR consensus : ATTTCAATCCACTCACCCATGAAGGGTGAGAC
- DR length : 32 Number of spacers : 5

```
109  CTTTCAATCCACTCACCCATGAAGGGTGAGAC  AAGATGACACTGGTACTGCACATGTCGATTTAAAA  175
176  ATTTCAATCCACTCACCCATGAAGGGTGAGAC  AGATCTTAAAATGGATTCTTCTTCAGATATTTTTG  242
243  ATTTCAATCCACTCACCCATGAAGGGTGAGAC  GATAACAGTTGCTTTAGTCGATAAGTCGATTAGCG  309
310  ATTTCAATCCACTCACCCATGAAGGGTGAGAC  ATTATGTTTTGCCACATGAGAAAGTAAAAAATGGA  376
377  ATTTCAATCCACTCACCCATGAAGGGTGAGAC  TCCCTTATAATCGACAAAAAGCGCCGATTGATT     441
442  ATTTCAATCCACTCACCCACGAAGGGTGAGAC                                        473
```

Figure 3-22. Output of CRISPR web server to indicate repeat-spacer structure of CRISPR RNA.

Table 16. Blast results of *cand34* spacers

Spacer1	*S. pyogenes* NZ131
	*Streptococcus phage phi*3396
Spacer2	*S. pyogenes* NZ131
	S. equi 4042
Spacer3	*S. pyogenes* NZ131
	S. pyogenes strain M23ND
	Streptococcus dysgalactica subsp. equisimilis GGS
Spacer4	*S. pyogenes* NZ131
	S. pyogenes AP53
	S. pyogenes NGAS743
	S. pyogenes AP1
	S. pyogenes NCTC8198
	S. pyogenes HSC5
	S. pyogenes Alab49
	S. pyogenes Manfredo
	S. pyogenes MGAS10394
	Streptococcus pyogenes prophage phiRamid paratox (prx_sda1)

Table 16. Blast results of *cand34* spacers (Continued)

Spacer5	*S. pyogenes* NZ131
	S. pyogenes MGAS11027
	Streptococcus phage P9
	S. pyogenes MGAS10394
	S. pyogenes STAB14018
	S. pyogenes AP1
	S. pyogenes NCTC8198
	S. pyogenes ATCC 19615
	S. pyogenes MGAS1882
	S. pyogenes NGAS322
	S. pyogenes NGAS743
	Vibrio parahaemolyticus strain FDAARGOS_191
	S. pyogenes Alab49
	Vibrio parahaemolyticus strain FORC_008
	Vibrio parahaemolyticus strain ATCC 17802
	Vibrio alginolyticus NBRC 15630 ATCC 17749
	Lactobacillus paracllinoides strain TMW1-1995

3.6.3   Characterization of *cand34* deletion mutant

To investigate the function of *cand34*, growth behavior of *cand34* deletion mutant and complementation strain and overexpressed strain in THY broth and their virulence-relevant properties were analyzed.

3.6.3.1   Investigating the growth behavior

To investigate whether the growth behavior of *S. pyogenes* M49 is affected by deletion of *cand34* growth of WT, *Δcand34*, *Δcand34::cand34* and M49/pAT19_*cand34* strains were monitored in THY broth at 37°C and 5% CO_2. The optical density was measured at 600_{nm} every hour to 24 hours. A similar growth behavior of all strains was observed (Figure 3-23).
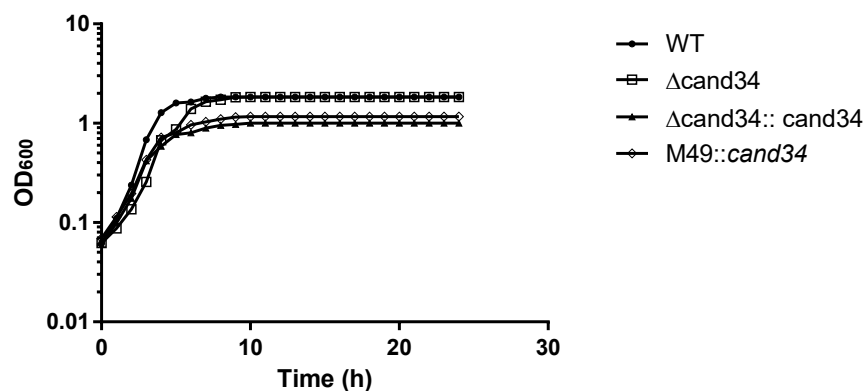
Figure 3-23. Growth curve of WT, *Δcand34*, *Δcand34*:: *cand34* and M49/pAT19_*cand34* in THY-Medium.

The growth rate (µ) and the doubling time of the strains were determined from the linear range of the exponential growth phase and are shown in Table 17.

$$\text{Growth rate (µ)} = \frac{\ln (X2) - \ln (X1)}{t2 - t1}$$

$$\text{Doubling time} = \frac{\ln 2}{\mu}$$

Table 17. Growth rate and doubling time of WT and recombinant *cand34* strains (n=3).

Strain	Growth rate(h^{-1})	Doubling time (h)
WT	0.72±0.057	0.96±0.057
Δcand34	0.81±0.041	0.87±0.045
Δcand34:: *cand34*	0.62±0.2	1.1±0.038
M49/pAT19_*cand34*	0.62±0.067	1.01±0.10

3.6.3.2 The influence of *cand34* on adherence to and internalization into human keratinocytes

To establish infection, the *Streptococcus* must firstly attach to the epithelium of the host cells. Interaction with the receptors on the cell suface leads to the streptococcal attachment to the epithelial cells. Internalization provides important protection against the immune response of the host as well as against treatment with antibiotics (Courtney & Podbielski, 2009). The influence of *cand34* on adherence and internalization of *S. pyogenes* M49 was studied. As *S. pyogenes* M49 was isolated from skin, the human HaCaT keratinocyte cell line was used as described (2.2.14.1).

A
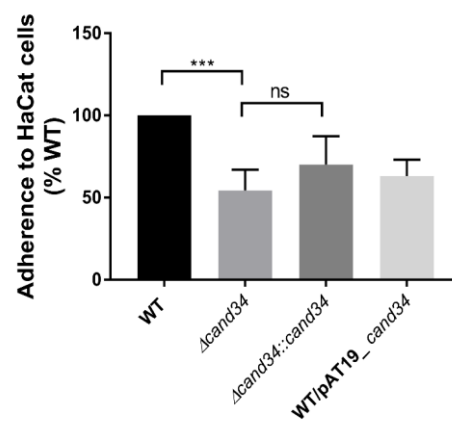
B
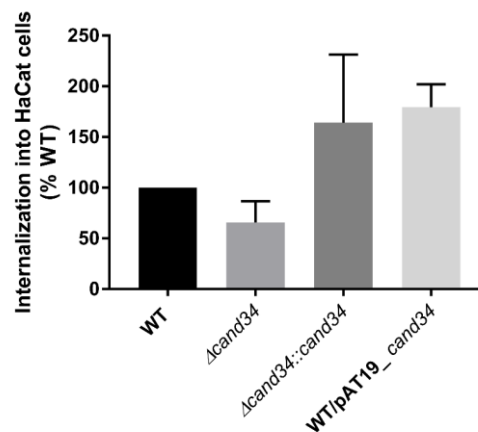
Figure 3-24. The effect of *cand34* on adherence and internalization of *S. pyogenes* on and into HaCaT cells. A: relative adherence of recombinant strains compared to WT (Kruskal Wallis test, multiple comparisons; not significant $P \geq 0.05$, *** $p \leq 0.001$, n=7). B: relative internalization of recombinant strains compared to the WT (n=5). The data are presented as the mean values $\pm$ standard deviation.

The relative adherence of *S. pyogenes* M49 and the recombinant strains to HaCaT cells was compared to the WT (Figure 3-24.A). Significant differences in the adherence behavior was determined between WT and *cand34* deletion mutant. No significant difference was observed between deletion mutant and complementation. This means complementation strain could not restore the WT phenotype. The internalization of *S. pyogenes* M49 and its recombinant derivatives into HaCaT cells was compared (Figure 3-24.B). No significant difference was observed between the WT and other strains.

3.6.3.3   The survival of *S. pyogenes* M49 and *Δcand34* in human blood

Another relevant factor in the virulence of *S. pyogenes* is the ability of the microorganism to survive in human blood.

To study the influence of *cand34* on survival of *S. pyogenes* in human blood, WT, *Δcand34*, *Δcand34::cand34* and M49/pAT19_*cand34* were incubated in heparinized human blood (2.2.14.2) and the survival rate was assessed. The survival rates of the recombinant strains were plotted against the multiplication factor (Figure 3-25). Results indicate that the *Δcand34* mutant has a significantly lower survival rate than the WT. *Δcand34::cand34* and M49/pAT19_*cand34* show no significant difference compare to the WT.
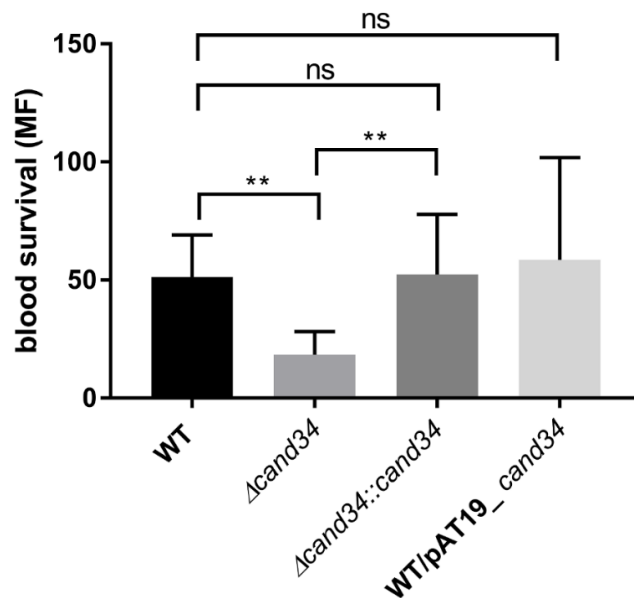
Figure 3-25. Growth of WT, *Δcand34*, *Δcand34::cand34* and M49/pAT19_*cand34* strains in human blood. *Δcand34* showed reduced growth compared to the WT. No significant difference was observed between WT and complementation strain and overexpression strain (Kruskal-Wallis test, multiple comparisons; not significant $P \geq 0.05$, ** $p \leq 0.01$, n =7). The data are presented as the mean values ± standard deviation.

3.6.4   The role of *cand34* in surviving under oxidative stress condition

To investigate the influence of *cand34* in surviving of *S. pyogenes* M49 under oxidative stress conditions, the following test was performed as explained (2.2.14.3). WT and *Δcand34* were grown to the early exponential growth phase ($OD_{600} \approx 0.5$) and different amounts of H_2O_2 were added to the cultures (in the absence of H_2O_2 as control and 0.5 and 1 mM as tests). Bacteria were incubated at 37°C and 5% CO_2 for 2 hours. Grown cells were thoroughly washed with PBS, serial dilutions were plated, and CFU were determined after 24h. Survival of the *cand34* deletion mutant strain was significantly lower than the WT at 0.5 mM of H_2O_2 (Figure 3-26).
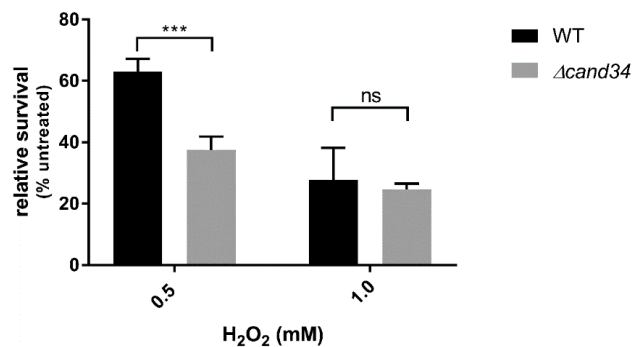
Figure 3-26. Relative survival of *S. pyogenes* M49 and *Δcand34* under oxidative stress conditions. (Two-way ANOVA, <u>n</u>ot <u>s</u>ignificant *P ≥ 0.05, *** p ≤ 0.001*. n =3). The data are presented as the mean values ± standard deviation.

3.6.5   Target prediction of sRNA cand34

In silico analysis can be used to predict sRNA target genes. Possible targets of *cand34* were predicted using the online tools TargetRNA2 (Kery & Feldman, 2014), CobraRNA (Wright, et al., 2013) and RNA predator (Zhao, et al., 2008). Predicted targets via online tools are given in Appendix 7-13, Appendix 7-14, Appendix 7-15.

4   Discussion

♦ Part I

This study was performed to characterize the function of the putative Gly riboswitch (ribogly) in *S. pyogenes* M49 strain 591. The expression of the riboswitch region in this microorganism was detected by intergenic tiling array (Patenge, et al., 2012). Gly riboswitch in GAS M49 is located in the 5′-UTR of the Na^+/Alanine symporter gene preceding the cation efflux system protein gene (Figure 1-7).

Riboswitches sense specific ligands. Known bacterial riboswitches control the expression of the genes that are involved in the respective ligand's transport or biosynthesis. Riboswitches regulate gene expression by different mechanisms which influence transcription termination, translation initiation, RNA stability and splicing. (Serganov & Nudler, 2013b; Mandal, et al., 2004).

Luciferase reporter assays revealed that the expression of the Na^+/Alanine symporter gene was induced in the absence of Gly or in the presence of low amount of Gly and was repressed in the presence of high Gly concentrations (Figure 3-4). RT-qPCR experiments revealed that upon increasing Gly concentrations, the full-length transcript level decreased drastically while ribogly transcript level remained constant (Figure 3-7). Northern blot analyses proved Na^+/Alanine symporter gene repression in the presence of high Gly concentrations and suggested premature transcription termination (Figure 3-11). One explanation for these results is premature transcription termination. In transcriptional regulation by riboswitches, conformational change leads to the formation of either a stem-loop structure analogous to Rho-independent termination or anti-terminator hairpin structure. The formation of a terminator leads to a transcription termination and short transcript will be derived. In *S. pyogenes* M49, in the presence of high concentrations of Gly transcription termination occurs and formation of the terminator inhibits elongation by RNA polymerase and only truncated transcript (ribogly) is detectable. In the absence of Gly or at low concentrations of Gly, anti-terminator hairpin structure is formed, and a full-length transcript will be produced. This kind of termination is similar to what happens in Rho-independent termination. It seems ribogly regulates transcription via controlling the formation of an intrinsic terminator.

One example of this mechanism was observed in *B. subtilis*. Gly riboswitch in *B. subtilis* is composed of two similar aptamer domains; Type I and Type II followed by a single

expression platform (*gcvT* operon) (Phan & Schumann, 2007) (Mandal, et al., 2004). In *S. pyogenes* one aptamer (type II) similar to the *B. subtilis* aptamers was predicted upstream of the Na^+/Alanine symporter gene (Figure 4-1). A Gly riboswitch in *V. cholera* preceding VC1422 gene (Na^+/Alanine symporter gene) also exhibits tandem architecture with two homologous aptamer domains followed by Na^+/Alanine symporter gene. The ligand-binding pocket sequence in the Gly riboswitch in both *B. subtilis* and *V. cholera* is similar. Type I and type II aptamers are identified based on the variations in the sequences which flank conserved cores. Sequences that are represented in black and red are conserved ($\geq$ 80% and $\geq$ 95%) and circles and black tick lines show less conserved sequences (Figure 4-1.A.B) (Mandal, et al., 2004).

The secondary structure of the Gly riboswitch in *S. pyogenes* M49 was predicted by RNAfold on the RNA web server (Gruber, et al., 2008) and was illustrated via VARNA (Darty, et al., 2009) (Figure 4-1.C). It is composed of two aptamer domains; type I and type II. Type I is homologous to type I aptamer in *V. cholera* (a gram-negative bacterium) and type II aptamer is homologous to type II in both *V. cholera* and *B. subtilis* (a gram-positive bacterium).

Evolution is one explanation for this secondary structure. To adapt new niches, RNA evolved to function as a genetic information storage and catalyze chemical reactions (Gilbert, 1986; Serganov & Nudler, 2013b). Natural selection leads to alteration in RNA sequences into specific metabolite-binding domains to adapt to environmental stresses and signals (e.g. nutritional depletion) (Serganov & Nudler, 2013b). Both aptamers in the Gly riboswitch in *S. pyogenes* possess conserved nucleotides in the ligand-binding pocket, which were previously described (Esquiaqui, et al., 2014; Mandal, et al., 2004; Weinberg, et al., 2017). Conserved sequences form a precise sensor for a ligand and never change through evolution (Breaker, 2011).
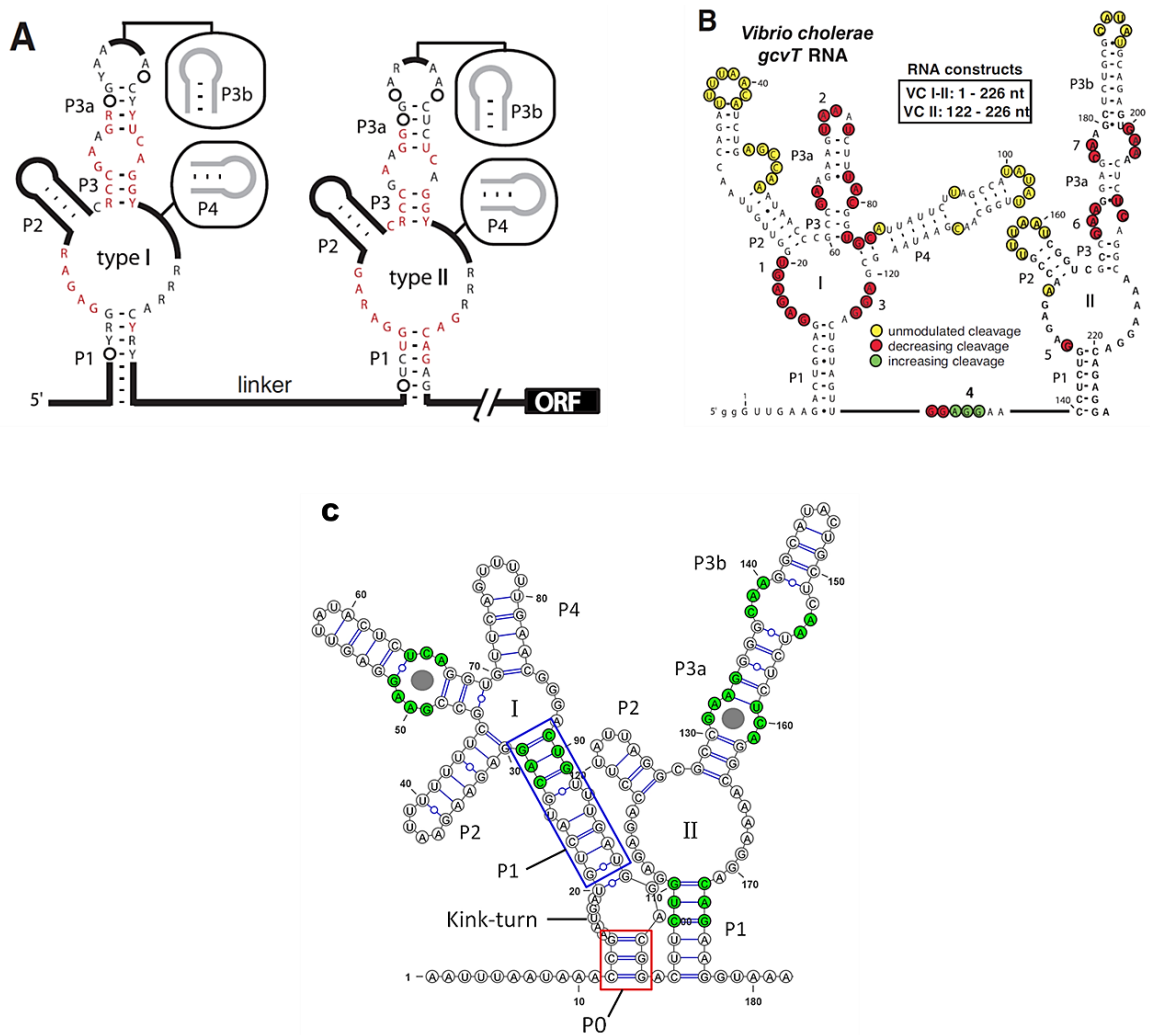
Figure 4-1. The secondary structure schematics of Gly riboswitch in A) *B. subtilis*, B) *V. cholerae*, C) *S. pyogenes.* Type I aptamer of Gly riboswitch in *S. pyogenes* is homologous to type I aptamer of Gly riboswitch in *V. cholera* (a gram-negative bacterium) and type II aptamer is homologous to type II in both *B. subtilis* and *V. cholera* (a gram-positive bacterium). Gly riboswitch secondary structure was explained in detail in 3.1.

In *S. pyogenes* M49 *ribogly* is preceding the Na$^+$/Alanine symporter gene and induction is detectable at low Gly concentration (0.01 mM). In *B. subtilis* Gly riboswitch is preceding *gcvT* operon and induction happens at high concentration of Gly (10 mM). *gcvT* operon encodes the Gly cleavage system proteins and catalyzes the degradation of Gly (to 5,10-methylene-tetrahydrofolate, CO_2, ammonia and NADH) for use of Gly as an energy source (Kikuchi, 1973; Mandal, et al., 2004). In the presence of high Gly concentration and binding to the aptamer domains, an anti-terminator structure is formed (Figure 4-2). Transcription is continued and expression of *gcvT* operon leads to the synthesis of the enzymes for Gly degradation. In absence of Gly, the transcription process is terminated prematurely and expression of *gcvT* operon is inhibited (Phan & Schumann, 2007). The fact that putative Na$^+$/Alanine symporter functions as an amino acid

transporter and *gcvT* operon functions for Gly degradation is consistent with the observations of their transcript levels in various amount of Gly. In *B. subtilis* when Gly concentration reaches to high level, a degradation system helps to adjust to the optimum condition. Na^+/Alanine symporter is a member of Na^+/Alanine symporter super family (SAF) and was studied in some microorganisms like *Alteromonas haloplanktis* (Reizer, et al., 1993). In this microorganism, the symporter transports alanine and glycine, Na^+-dependently. It has not been studied in *S. pyogenes* to date. The fact that Na^+/Alanine symporter transcription is drastically limited in high Gly concentrations and increased in low amounts, suggests that it is involved in Gly transport in *S. pyogenes* too.
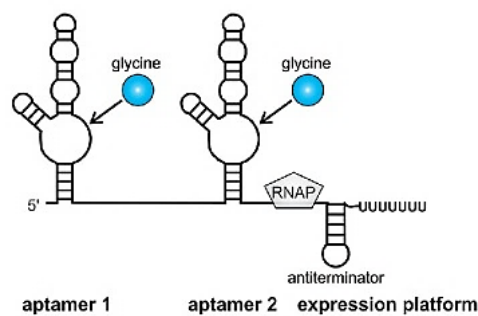
Figure 4-2. Formation of an anti-terminator structure in Gly riboswitch in *B. subtilis*. Two aptamers control a single expression platform.

An additional possibility for the continuous presence of ribogly in different concentration of Gly in contrast to the full-length transcript could be the high stability of its transcript. However, results of transcript stability test do not support this hypothesis. The ribogly transcript stability assay determined that short transcript (ribogly) is unstable in different concentrations of Gly and trace amounts were detected after one minute of stopping *de novo* synthesis of RNA in non-repressing (0.01 mM) and repressing (10 mM) level of Gly concentration. This fact implies that constant presence of ribogly is not due to the high stability of its transcript.

The Na^+/Alanine symporter transcript stability was rather high in non-repressing level (0.01 mM) of Gly and low in repressing levels (10 mM). ribogly is not stable under both conditions. These observations suggest that processing of the full-length transcript happens, and the processed Na^+/Alanine symporter gene mRNA is more stable than the full-length transcript. A model of the full-length transcript processing in repressing and non-repressing conditions is provided in Figure 4-3.

The hypothesis is that at low Gly concentrations the full-length transcript undergoes processing. Ribogly and the coding transcripts (Na$^+$/Alanine symporter and Efflux mRNA) are separated. Northern blot analysis of the coding transcript revealed a 300 nucleotides band that according to this model represents the outcome of processing. Subsequently ribogly is degraded. In this model, ribogly could be a target of a ribonuclease. It is also possible that riboswitch RNA has auto-catalytic function as was observed in glmS riboswitch-ribozyme in *B. subtilis* (Collins, et al., 2007). A low concentration of Gly in the medium on one side induces the formation of an anti-terminator structure and on the other side exposes a cleavage site downstream of the anti-terminator. Separation leads to the stabilization of the Na$^+$/Alanine symporter and Efflux mRNA.

This observation proves that neither high stability nor processing of the full-length transcript, explains constant presence of ribogly (short transcript) but its continuous synthesis. A similar mechanism has been observed in *B. subtilis,* where in the absence as well as in the presence of Gly, the stability of riboswitch RNA was low (Phan & Schumann, 2007). In contrast to the situation in *S. pyogenes*, the full-length transcript was more stable in high Gly concentrations, compared to Gly deficient conditions (Figure 3-13).

A similar process has been reported for the *Bacillus* threonyl/tRNA synthetase gene (*thrS*). Expression of *thrS* is regulated by a T-box that responds to the limitation of threonine (Grundy, et al., 1993). A low concentration of threonine leads to a tRNA-mediated anti-termination and results in expression of *thrS*. An additional control system for *thrS* expression is processing of the full-length transcript. In the absence or in the presence of limited amounts of threonine, the leader region of mRNA is processed between the T-box and the Rho-independent transcription terminator. The *thrS* mRNA is detected in cleaved form and processed transcript is stable (Condon, et al., 1996; Carona, et al., 2012).

The same process was observed in Mg^{2+} riboswitch in *Salmonella enterica* serovar Typhimurium (Spinelli, et al., 2008). In the presence of high Mg^{2+} concentrations, besides transcription termination, RNA degradation by RNase E is an additional mechanism controlling *mgtA* expression that encodes Mg^{2+} transporter.

In gram-positive bacteria, the *glmS* riboswitch-ribozyme in the 5′-UTR of the gene encoding glucosamine-6-phosphate (GlcN6P) synthetase, acts based on mRNA destabilization (Collins, et al., 2007). Its mechanism is based on metabolite sensing and

subsequently mRNA cleavage (Winkler, et al., 2004). Binding of GlcN6P to the riboswitch leads to an autocatalytic site-specific cleavage of *glmS* mRNA and creats 5´-OH on the fragment that stimulates degradation of the downstream gene by RNase J (Collins, et al., 2007). RNase J1 and RNase J2 are essential ribonucleases for growth of *S. pyogenes* (Bugrysheva & Scott, 2010). To investigate the mechanism of processing in Gly riboswitch in *S. pyogenes* these two ribonucleases have high potential as a subject for further study.
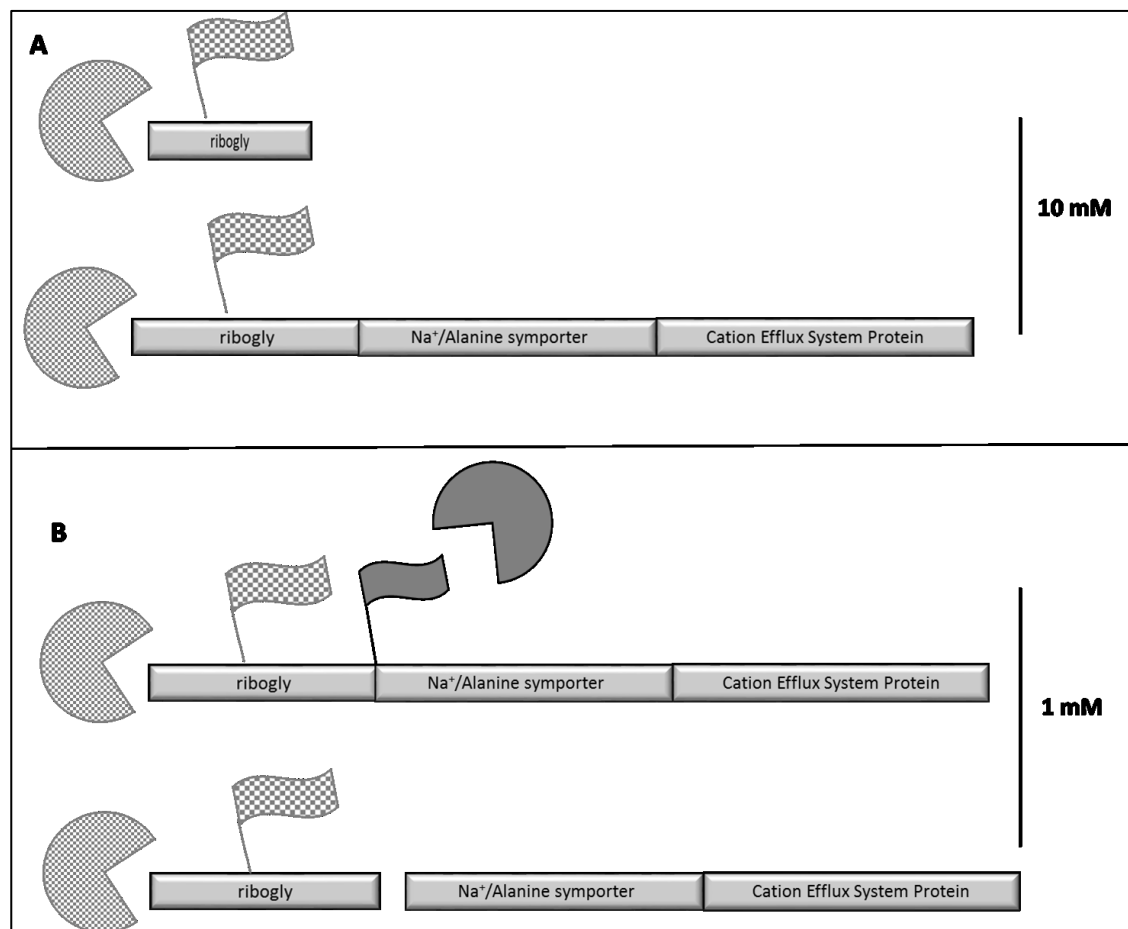
Figure 4-3. Model of the full-length transcript processing in low amount of Gly which leads to the removal of ribogly transcript. A) Degradation of the ribogly and the full-length transcript at 10 mM. B) Stabilization of the symporter and the efflux mRNA by processing: the ribogly is removed and degraded (Figure 3-13).

♦ **Part II**

In the second part, sRNA *cand34* in *S. pyogenes* M49 strain 591 was studied. It was identified as CRISPR RNA in CRISPR system I-C. As described (1.8) CRISPR RNA is composed of short foreign DNA fragments (spacers) that are separated by repeat sequences. In one study, spacer analysis of a type II-A CRISPR-Cas system in isolates of *S. agalactica* determined that two spacers targeted two regions on its own chromosome (Stern, et al., 2010). Nowaza et. al suggested that the limited presence of CRISPR-systems in *S. pyogenes* strains have allowed the integration of virulence genes by phages into different *S. pyogenes*. This theory suggests that their ability to cause diseases can be related to the integrated prophages (Beres, et al., 2002; Beres & Musser, 2007). The full genome sequences of several strains of *S. pyogenes* genome showed the prophage regions possess variable virulence genes and extend pathogenic capacities (Fischetti, 2007).

Inactivating the type II CRISPRCas system in *Campylobacter jejuni* revealed that the ability of this microorganism to infect the human peripheral nervous system (in Guillain-Barré Syndrome) is drastically decreased (Louwen, et al., 2013). To study the potential involvement of CRISPR system in pathogenicity, they inactivated *csn1* in the system, which effectively reduced virulence in *Campylobacter jejuni* isolates. This finding revealed a link between viral defence and virulence of *C. jejuni* and Guillain-Barré Syndrome.

Δcand34 in *S. pyogenes* M49 was constructed to study a potential regulatory function and possible role of *cand34* in virulence and metabolism. A quantification of *cand34* transcript levels in exponential phase via RT-qPCR in WT and the complementation strain showed drastic increase in the complementation strain. This effect can be caused by the high copy number of pAT19, which was used as shuttle vector in this experiment (Trieu-Cuot, et al., 1991).

Investigating upstream and downstream genes transcript levels after deletion of *cand34* showed that *ccdA,* that is annotated as cytochrome C biogenesis protein gene was silent in WT but showed high-level of transcription in *Δcand34*. Measuring its transcript level in complementation strain revealed that *cand34* in shuttle vector is not able to fully restore the transcript level of *ccdA* to the WT level. This phenomenon can be described by Primer sequences for amplifying flanking regions of *cand34* to construct deletion mutant that were designed using NZ131 genome and not *S. pyogenes* M49 (because whole genome of *S. pyogenes* M49 is not sequenced to date). It is possible that at this region, difference in sequences causes interruption in deletion of *cand34*. This hypothesis

can explain the disability of *cand34* in the complementation strain in restoring the transcript level of *ccdA*. Since the complementation strain does not complement *ccdA*, *cand34* overexpression strain was used for the following experiments. Increased transcript level of *cand34* in the overexpression strain could be used to study the effect of *cand34* on virulence of *S. pyogenes* M49.

A growth test was performed in THY to test comparability of the recombinant strains with WT. Disturbed growth behavior in constructed strains can influence the outcome of other experiments and distort their results.

Because no significant difference was observed in growth behavior in strains they could be used for subsequent experiments.

Survival of Δ*cand34* in human blood was significantly reduced compared to the WT. In the complementation strain, growth could be restored to the WT level. As discussed above, this phenotype can be attributed to *cand34*. Further study is needed to determine the *cand34* transcript level in the WT growing in human blood.

Investigating *cand34* function in oxidative stress condition revealed that the *cand34* deletion mutant has less ability to survive in this condition. As the complementation strain was excluded from oxidative stress experiment because of small colony phenotype, it is not clear whether *cand34* is responsible for decreased oxidative stress tolerance. As mentioned, *ccdA* was silent in WT, while in deletion mutant strain its transcript level dramatically increased and complementation strain was not able to return its transcript level to the WT. It might be proper a solution to construct *ccdA* deletion mutant and include it in the experiment beside *cand34*.

ccdA has been also predicted in silico as a target for *cand34* (Appendix 7-13, Appendix 7-15) Prediction of its product interaction with other proteins (Szklarczyk, et al., 2015) showed that *ccdA* probably interacts with *msrA* (Figure 4-4).

In silico target prediction (Appendix 7-15) suggests *msrA* as one possible target for *cand34*. *msrA* codes for methionine sulfoxide reductase which protects against oxidative stress. It has an important role as a repair enzyme for proteins that have been inactivated by oxidation. It catalyzes the oxidation-reduction of methionine sulfoxide to methionine (Cabreiro, et al., 2006). Regarding the observation in the oxidative stress assay (Figure 3-26), the transcript level of *msrA* in strains needs to be determined. The location of *msrA* in *S. pyogenes* chromosome is presented in Figure 4-5.
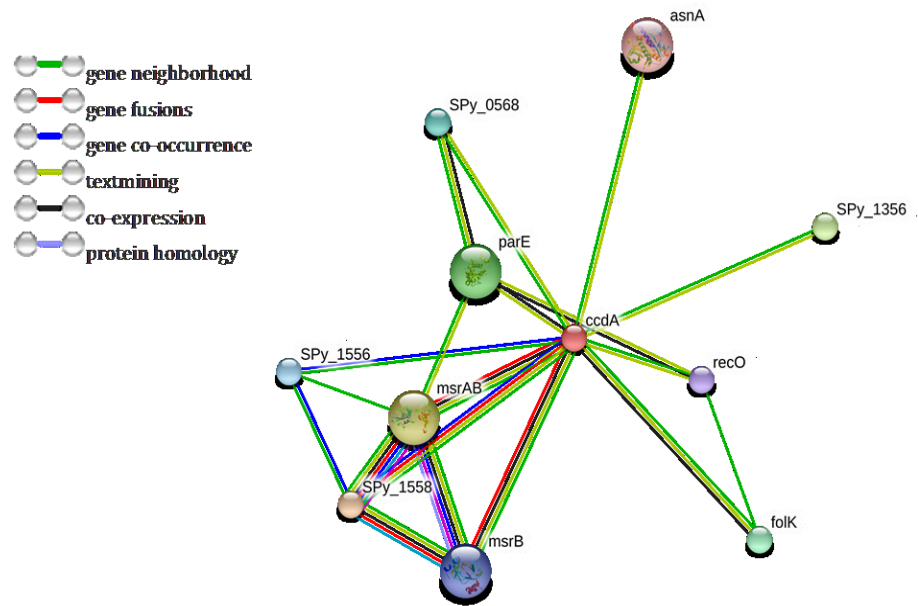
Figure 4-4. Prediction of *ccdA* interaction(s) with other genes. Among predicted genes *msrA* is seen which protect against oxidative stress.
https://string-db.org/cgi/input.pl?UserId=XnJ6BFyikDqU&sessionId=7lHSAPzFFJxe
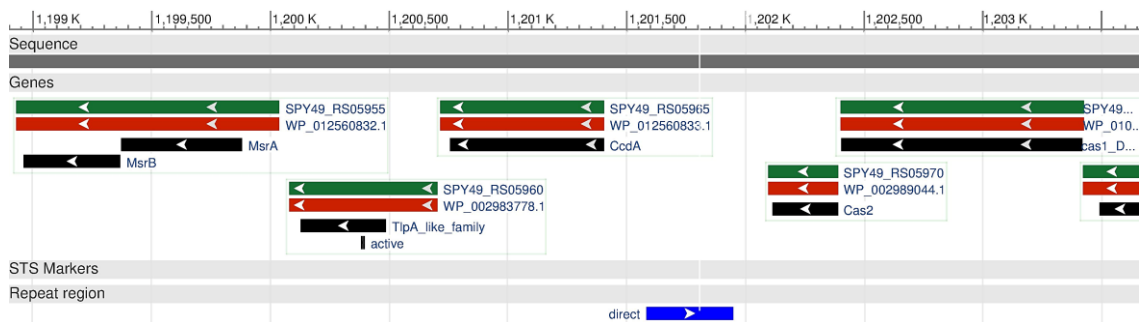
Figure 4-5. Pictorial representation o the location of CRISPR RNA, *ccdA* and *msrA*.
https://www.ncbi.nlm.nih.gov/nuccore/NC_011375.1?report=graph#

Further studies are needed to investigate the effect of *cand34* on *msrA* transcript level via RT-qPCR under different conditions of oxidative stress. For a more precise experiment, pAT19 as shuttle vector in complementation strain should be replaced to decrease the high copy number of plasmid and restore the transcript level of *cand34* to the WT level. The presence of *ccdA* deletion mutant in this experiment will be helpful.

5   Summary

Part I:

Riboswitches are non-coding RNAs that are induced by metabolites, ions or other environmental conditions and are mostly found within the 5′-UTR of mRNAs. They regulate the expression of downstream genes that are involved in transport or catabolism of the respective inducer. In this work, a putative glycine riboswitch (*ribogly*) located within the 5′-UTR of Na$^+$/Alanine symporter gene was studied. Ribogly secondary structure is composed of two aptamer domains. Type I is similar to the type I aptamer domain in Gly riboswitch in *Vibrio cholerae* (a gram-negative bacterium) and type II is similar to type II in *Bacillus subtilis* (a gram-positive bacterium) and in *Vibrio cholerae* (Mandal, et al., 2004). Using a reporter gene system, it could be shown that *ribogly* is a glycine-dependent riboswitch. Reverse transcription followed by quantitative PCR revealed that in different concentrations of glycine ribogly is consistently transcribed while the transcription of the full-length mRNA decreased drastically at high amounts of glycine. Visualizing transcripts by Northern blotting suggests that in the absence of glycine or in the presence of a low glycine concentration, ribogly forms an anti-terminator allowing synthesis of the full-length mRNA. In high glycine concentrations, a terminator is formed leading to transcription attenuation. Stopping *de novo* synthesis of RNA revealed that the full-length transcript is processed. Transcript stability assays showed that continuous presence of the ribogly is not caused by high stability of ribogly. Instead, the truncated ribogly transcript is synthesized continuously. Fusing the endogenous promoter of Na$^+$/Alanine symporter gene to the reporter gene demonstrated that induction behavior in *ribogly* is not affected by the native promoter.

Part II:

In this part, a putative regulatory sRNA (*cand34*) in *S. pyogenes* M49 was characterized through virulence assays. A deletion mutant of *cand34* was constructed. *ccdA* (upstream gene of *cand34*) which was silent in WT, was transcribed in Δ*cand34*. The ectopic expression of cand34 was not able to reduce *ccdA* transcript level to the WT. One possible explanation is the damage of a regulatory element in the 5′ region of *ccdA* that has been caused by the deletion of *cand34*. *cand34* deletion mutant showed a reduced growth in human blood than WT.

6 References

Adams, M. D., Celniker , S. E., Holt, R. A. & Evans , C., 2000. The Genome Sequence of Drosophila melanogaster. *Science,* Volume 287, pp. 2185-2195.

Altuvia, S. & Wagner, G., 2000. Switching on and off with RNA. *PNAS,* Volume 97, pp. 9824-9826.

Babitzke , P. & Romeo, T., 2007. CsrB sRNA Family: Sequestration of RNA-binding Regulatory Proteins. *Curr Opin Microbiol,* Volume 10, pp. 156-163.

Baird, N. J. & Ferre-Damare, A. R., 2013. Modulation of quaternary structure and enhancement of ligand binding by the K-turn of tandem glycine riboswitches. *RNA,* Volume 19, pp. 167-176.

Baker, J. L. et al., 2012. Widespread genetic switches and toxicity resistance proteins for fluoride. *Science,* Volume 335, pp. 233-235.

Barrick, J. E. et al., 2004. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *PNAS,* Volume 101, pp. 6421-6426.

Bastet, L., Dubé, A., Massé, E. & Lafontaine, D. A., 2011. New insights into riboswitch regulation mechanisms. *Molecular Microbiology,* Volume 80, pp. 1148-1154.

Becker, W. C., 1916. The nessecity of a standard blood agar plate for the determination of hemolysis by Streptococci. *The Journal of Infectious Diseases,* Volume 19, pp. 754-759.

Benedek, D. C., 2006. The history of bacteriologic concepts of rheumatic fever and rheumatoid arthritis. *Semin Arthritis Rheum,* Volume 36, pp. 109-123.

Beres, S. B. & Musser, . J. M., 2007. Contribution of exogenous genetic elements to the group A Streptococcus metagenome. *PLoS One,* Volume 8, p. e800.

Beres, S. et al., 2002. Genome sequence of a serotype M3 strain of group A Streptococcus: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc Natl Acad Sci U S A,* Volume 99, pp. 10078-10083.

Boukamp, P. et al., 1988. Normal Keratinization in a Spontaneously Immortalized Aneuploid Human Keratinocyte Cell Line. *The Journal of Cell Biology,* Volume 106, pp. 761-771.

Breaker, R., 2011. Prospects for Riboswitch Discovery and Analysis. *Mol Cell,* Volume 43, pp. 867-879.

Bugrysheva, J. & Scott, J., 2010. The ribonucleases J1 and J2 are essential for growth and have independent roles in mRNA decay in Streptococcus pyogenes. *Molecular Microbiology,* Volume 75, pp. 731-743.

Butler, E. B., Xiong, Y., Wang, J. & Strobel, S. A., 2011. Structural Basis of Cooperative Ligand Binding by the Glycine Riboswitch. *Chem Biol,* Volume 18, pp. 293-298.

Cabreiro, F., Picot, E. R., Friguet, B. & Petropoulos, I., 2006. Methionine sulfoxide reductases: relevance to aging and protection against oxidative stress. *Ann N Y Acad Sci,* Volume 1067, pp. 37-44.

Camara, M., Dieng, A. & Bouh Boye, C. S., 2013. Antibiotic Susceptibility of Streptococcus Pyogenes Isolated from Respiratory Tract Infections in Dakar, Senegal. *Microbiol Insights,* Volume 6, pp. 71-75.

Carona, M. P. et al., 2012. Dual-acting riboswitch control of translation initiation and mRNA decay. *Proc Natl Acad Sci,* Volume 109, pp. E3444-53.

Chen, A. G., Sudarsan, N. & Breaker, R. R., 2011. Mechanism for gene control by a natural allosteric group I ribozyme. *RNA,* Volume 17, pp. 1967-1972.

Cho, K. . H. & Kim, J.-H., 2015. Cis-encoded non-coding antisense RNAs in streptococci and other low GC Gram (+) bacterial pathogens. *Front Genet,* 6(110).

Collins, J. A., Irnov, I., Baker, S. & Winkler, W. C., 2007. Mechanism of mRNA destabilization by the glmS ribozyme. *Genes Dev,* Volume 21, pp. 3356-3368.

Condon, C., Putzer , H. & Grunb-Manago, M., 1996. Processing of the leader mRNA plays a major role in the induction of thrS expression following threonine starvation in Bacillus subtilis. *Proc Natl Acad Sci U S A,* Volume 93, pp. 6992-6997.

Courtney, H. S. & Podbielski, A., 2009. *chapter 8: Group A streptoccoal invation of host cells.* s.l.:s.n.

Cromie, M. J., Shi, Y., Latifi, T. & Groisman, E. A., 2006. An RNA sensor for intracellular Mg(2+). *Cell,* Volume 125, pp. 71-84.

Cunningham, M. M. W., 2000. Pathogenesis of Group A Streptococcal Infections. *Clin Microbiol Rev.,* Volume 13, pp. 470-511.

Cunningham, M. W., 2008. *Hot Topics in Infection and Immunity in Children IV: Pathogenesis of group A streptococcal infections and their sequelae.* New york: Springer, New York, NY.

Cunningham, M. W., 2016. *Streptococcus pyogenes: Post-Streptococcal Autoimmune Sequelae: Rheumatic Fever and Beyond.* Oklahoma City (OK): University of Oklahoma Health Sciences Center.

Danger, J. L., Makthal, N., Kumaraswami, M. & Sumbya, . P., 2015. The FasX Small Regulatory RNA Negatively Regulates the Expression of Two Fibronectin-Binding Proteins in Group A Streptococcus. *Journal of Bacteriology,* Volume 197(23), pp. 3720-3730.

Darty, K., Denise, A. & Ponty, Y., 2009. VARNA: Interactive drawing and editing of the RNA secondary structure. *APPLICATIONS NOTE,* Volume 25, pp. 1974-1975.

de Jong, A. et al., 2012. PePPER: a web server for prediction of prokaryote promoter elements and regulons. *BMC Genomics,* 13(229).

Dunny, G. M., Lee, L. N. & Leblanc, D. J., 1991. Improved electroporation and cloning vector system for gram-positive bacteria. *Appl Environ Microbiol,* Volume 57, pp. 1194 -1201.

Esquerré, T. et al., 2016. The Csr system regulates genome-wide mRNA stability and transcription and thus gene expression in Escherichia coli. *Scientific Reports,* p. DOI: 10.1038/srep25057.

Esquiaqui, J. M., Sherman, E. M., Ye, J.-D. & Fanucci, G. E., 2014. Site-Directed Spin-Labeling Strategies and Electron Paramagnetic Resonance Spectroscopy for Large Riboswitches. *Methods in Enzymology,* Volume 549, pp. 287-309.

Fiedler, T., Sugareva, V., Patenge, N. & Kreikemeyer, B., 2010. Insights into Streptococcus pyogenes pathogenesis from transcriptome studies. *Future Microbiol,* Volume 5, pp. 1675-1694.

Fischetti, V., 2007. In vivo acquisition of prophage in Streptococcus pyogenes. *Trends Microbiol,* Volume 15, pp. 297-300.

Fischetti, V. A., 2016. *M Protein and Other Surface Proteins on Streptococci.* Oklahoma City (OK): The University of Oklahoma Health Sciences Center.

Fulde, M. & Valentin Weigand, P., 2013. Epidemiology and pathogenicity of zoonotic streptococci. In: G. S. Chhatwal, ed. *Host-Pathogen Interactions in Streptococcal Diseases.* s.l.:Springer, Berlin, Heidelberg, pp. 49-81.

Gao, X. Y. et al., 2014. Comparative Genomics of the Bacterial Genus Streptococcus Illuminates Evolutionary Implications of Species Groups. *PLoS One,* Volume 9, p. e101229.

Gilbert, W., 1986. Origin of life: The RNA world. *Nature,* Volume 319, p. 618.

Grissa, I., Vergnaud, G. & Pourcel, C., 2007. CRISPR Finder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res,* Volume 35.

Gruber, A. et al., 2008. The Vienna RNA Websuite. *Nucleic Acids Res,* Volume 36, pp. 70-74.

Grundy, F. H. T., Grundy, F. J. & Henkin, T. M., 1993. tRNA as a positive regulator of transcription antitermination in B. subtilis. *Cell,* Volume 13, pp. 475-82.

Hanahan, D., 1983. Studies on transformation of Escherichia coli with plasmids. *J Mol Biol,* Volume 166, pp. 557-80.

Hanski , E., Horwitz, P. A. & Capron, M. G., 1992. Expression of protein F, the fibronectin-binding protein of Streptococcus pyogenes JRS4, in heterologous Streptococcal and Enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect Immun,* Volume 12, pp. 119-125.

Hatoum-Aslan, A. & Marraffini, L. A., 2014. Impact of CRISPR immunity on the emergence and virulence of bacterial pathogens. *Curr Opin Microbiol,* pp. 82-90.

Heler, R., Marraffini, L. A. & Bikard, . D., 2014. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. *Molecular Microbiology,* Volume 93, pp. 1-9.

Jiang, F. & Doudna, J. . A., 2015. The structural biology of CRISPR-Cas systems. *Current Opinion in Structural Biology,* Volume 30, pp. 100-111.

Kanehisaa, M. & Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes.. *Nucleic Acids Res,* 28(1), pp. 27-30.

Kazanov, M., Vitreschak, A. & Gelfand, M., 2007. Abundance and functional diversity of riboswitches in microbial communities. *BMC Genomics,* Volume 8, pp. 347-356.

Kery, M. . B. & Feldman, M., 2014. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids,* Volume 42, pp. 124-9.

Kibbe, W., 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res,* Volume 35, pp. 43-46.

Kikuchi, G., 1973. The Glycin cleavage system:composition, reaction mechanism and physiological significans. *Molecilar & Cellular Biochemistry,* 1(2), pp. 169-187.

Kingsford, C. L., Ayanbule, . K. & Salzberg, S. L., 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biology,* 8(2).

Kreikemeyer, B. et al., 2001a. Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (Fas) with homologies to two-component-type regulators requires a small RNA molecule. *Molecular Microbiology,* Volume 39, pp. 392-406.

Kreikemeyer, B. et al., 2011c. Genomic organization, structure, regulation and pathogenic role of pilus constituents in major pathogenic Streptococci and Enterococci. *Med Microbiol,* pp. 240- 251.

Kreikemeyer, B., McIver, K. S. & Podbielski, . A., 2003b. Virulence factor regulation and regulatory networks in Streptococcus pyogenes and their impact on pathogen-host interactions. *Trends Microbiol,* Volume 11, pp. 224-32.

Kunin, V., Sorek, R. & Hugenholtz, P., 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol,* 8 (R61).

Lancefield, R. C., 1933a. A serological differentiation of human and other groups of hemolytic Streotococci. *J Exp Med,* Volume 57, pp. 571-595.

Lancefield, R. C., 1962b. Current knowledge of type-specific M antigens of group A streptococci. *J. Immunol,* Volume 89, pp. 307-313.

Landt, S. G. et al., 2008. Small non-coding RNAs in Caulobacter crescentus. *Mol Microbiol,* Volume 68, pp. 600-614.

Le Rhun, A. et al., 2016. RNA sequencing uncovers antisense RNAs and novel small RNAs in Streptococcus pyogenes. *RNA BIOLOGY,* 13(2), pp. 177-195.

Lee, P. Y., Costumbrado, J., Hsu, C. Y. & Kim, Y. H., 2012. Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments,* Volume 62, p. e3923.

Levering, J. et al., 2016. Genome-scale reconstruction of the Streptococcus pyogenes M49 metabolic network reveals growth requirements and indicates potential drug targets. *Journal of Biotechnology,* Volume 232, pp. 25-37.

Liu, Z., Treviño, J., Ramirez-Pe, E. & Sumby, P., 2012. The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen group A Streptococcus. *Mol Microbiol,* Volume 86, pp. 140-154.

LIvak, K. J. & Schmittgen, . T. D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* Volume 4, pp. 402-408.

Li, Z. et al., 1999. Identification of pel, aStreptococcus pyogenes Locus That Affects both Surface and Secreted Proteins. *Journal of Bacteriology,* 181(19), pp. 6019-6027.

Logan, L. K., McAuley, . J. B. & Shulman, S. T., 2012. Macrolide Treatment Failure in Streptococcal Pharyngitis Resulting in Acute Rheumatic Fever. *PEDIATRICS,* Volume 129, pp. 798-802.

Louwen, R. et al., 2013. A novel link between Campylobacter jejuni bacteriophage defence, virulence and Guillain-Barré syndrome. *Eur J Clin Microbiol Infect Dis,* Volume 32, pp. 207-226.

Makarova, K. S. et al., 2011. Evolution and classification of the CRISPR–Cas systems. *NATURE REVIEWS,* Volume 9, pp. 467-477.

Mandal, M. et al., 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science,* Volume 306, pp. 275-279.

Mangold, M. et al., 2004. Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. *Molecular Microbiology,* Volume 53, pp. 1515-1527.

References

Maripuu, L., Eriksson, . A. & Norgren, M., 2008. Superantigen gene profile diversity among clinical group A streptococcal isolates.. *FEMS Immunol Med Microbiol,* Volume 54, pp. 236-244.

Marraffini, L. A. & Sontheimer, E., 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet,* Volume 11, pp. 181-190.

McGregor, K. & Spratt, B. G., 2004. Multilocus Sequence Typing of Streptococcus pyogenes Representing Most Known emm Types and Distinctions among Subpopulation Genetic Structures. *Jornal of Bacteriology,* Volume 186, pp. 4285-4294.

Mironov, A. S. et al., 2002. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell,* Volume 111, pp. 747-56.

Mullis, K. et al., 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol,* Volume 51, pp. 263-273.

Murdoch, A. D., 1998. Gram-Positive Anaerobic Cocci. *Clin Microbiol Rev,* Volume 1, pp. 81-120.

Nahvi, A. et al., 2002. Genetic control by a metabolite binding mRNA. *Chem Biol,* Volume 9, pp. 1043-1049.

Nakata, M. et al., 2009. Mode of expression and functional characterization of FCT-3 pilus region-encoded proteins in Streptococcus pyogenes serotype M49. *Infect. Immun,* Volume 77, pp. 32-44.

Nozawa, T. et al., 2011. CRISPR Inhibition of Prophage Acquisition in Streptococcus pyogenes. *PLoS One,* p. e19543.

Pancholi, V. & Caparon, M., 2016. Streptococcus pyogenes Metabolism. In: J. J. Ferretti , D. L. Stevens & V. A. Fischetti, eds. *Streptococcus pyogenes : Basic Biology to Clinical Manifestations.* Oklahoma City (OK): University of Oklahoma Health Sciences Center.

Parkinson, J. S. & Kofoid, E. C., 1992. Communication modules in bacterial signaling proteins. *Annu Rev Genet,* Volume 26, pp. 71-112.

Patenge, N. et al., 2012. Identification of novel growth phase- and mediadependent small non-coding RNAs in Streptococcus pyogenes M49 using intergenic tiling arrays. *BMC Genomics,* 13(550), pp. 550-564.

Patenge, N., Pappesch, R., Khani, A. & Kreikemeyer, B., 2015. Genome-wide analyses of small non-coding RNAs in streptococci. *Frontiers in Genetics,* Volume 6, p. 189.

Patterson., M. J., 1996. Streptococcus. In: S. Baron , ed. *Medical Microbiology. 4th edition.* Galveston (TX): University of Texas Medical Branch at Galveston.

Perez, N. et al., 2009. A genome-wide analysis of small regulatory RNAs in the human pathogen group A Streptococcus. *PLoS One,* 4(11), p. e7668.

Pfaffl, M. & Pfaffl, M. W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res,* Volume 29, pp. 2002-2007.

Phan, T. T. P. & Schumann, W., 2007. Development of a glycine-inducible expression system for Bacillus subtilis. *Journal of Biotechnology,* Volume 128, pp. 486-499.

Podbielski, A. & Kreikemeyer, B., 2004b. Cell density — dependent regulation: basic principles and effects on the virulence of Gram-positive cocci. *International Journal of Infectious Diseases,* Volume 8, pp. 81-95.

Podbielski, A., Woischnik, M., Leonard, B. & Schmidt, K.-H., 1999a. Characterization of nra, a global negative regulator gene in group A streptococci. *Molecular Microbiology,* 31(4), pp. 1051-1064.

Popp, N., 2015. Identifizierung und Funktionsnachweis von cis-regulatorischen Elementen in Streptococcus pyogenes Serotyp M49 (Master thesis).

Raasch, P. et al., 2010. Non-coding RNA detection methods combined to improve usability, reproducibility and precision. *BMC Bioinformatics,* 11(491).

Ramirez-Peña, E. et al., 2010. The group A Streptococcus small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the ska mRNA transcript. *Mol Microbiol.,* Volume 78, pp. 1332-1347.

Reizer, J., Reizer, A. & Saier, M. H., 1993. A functional superfamily of sodium/solute symporters. *Biochimica et Biophysica Acta,* Volume 29, pp. 133-166.

Roberts, S. A. & Scott, J. R., 2007. RivR and the small RNA RivX: the missing links between the CovR regulatory cascade and the Mga regulon. *Mol Microbiol,* 66(6), pp. 1506-1522.

Rocha, E. P., 2002. Is there a role for replication fork asymmetry in the distribution of genes in bacterial genomes. *TRENDS in Microbiology,* Volume 10, pp. 393-395.

Ruff, K. M., Muhammad, A., Mccown, P. J. & Ronald, B. R., 2016. Singlet glycine riboswitches bind ligand as well.. *RNA,* Volume 22, pp. 1-11.

Ruff, K. M. & Strobel, S. A., 2014. Ligand binding by the tandem glycine riboswitch depends on aptamer dimerization but not double ligand occupancy. *RNA,* Volume 20, pp. 1775-1788.

Ryan, P. A. & Juncosa, B. P., 2016. Group A Streptococcal Adherence. In: J. J. Ferretti , D. L. Stevens & V. A. Fischetti , eds. *Basic Biology to Clinical Manifestations.* Oklahoma City (OK): University of Oklahoma Health Sciences Center.

Saberi, F. et al., 2016. Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. *Cellular & Molecular Biology Letters,* 21(6).

Serganov, A. & Nudler, E., 2013b. A Decade of Riboswitches. *Cell,* Volume 152, pp. 17-24.

Serganov, A. & Patel, D. J., 2009a. Amino acid recognition and gene regulation by riboswitches. *Biochim Biophys Acta,* Volume 1789, pp. 592-611.

Sharma, C. M., Darfeuille, F., Plantinga, T. H. & Vogel, J., 2007. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *GENES & DEVELOPMENT,* Volume 21, pp. 2804-2817.

Sherman, E. M., Esquiaqui, J., Elsayed , G. & Ye, J.-D., 2012. An energetically beneficial leader–linker interaction abolishes ligand-binding cooperativity in glycine riboswitches. *RNA,* Volume 18, pp. 496-507.

Shimoni, Y. et al., 2007. Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol,* 3(138).

Slade, H. D., Grace, K. A. & Slamp, W. C., 1951. The amino acid nutrition of group Ahemolytic Streptococci, with reference to the effect of glutathione on thecystine requirement. *Bacteriol,* Volume 62, pp. 669-675.

Soukup, G. A. & Breaker, . R. R., 2000. Allosteric nucleic acid catalysts. *Current Opinion in Structural Biology,* Volume 10, pp. 318-325.

Spinelli, S. V., Pontel, L. B., V´escovi, E. G. & Soncini, F. C., 2008. Regulation of magnesium homeostasis in Salmonella :Mg21 targets the mgtA transcript for degradation by RNase E. *FEMS Microbiol Lett,* Volume 280, pp. 226-234.

Steele, R. H., White, A. G. & Pierce, W. A., 1953. The fermentation of Galactose by Strepto coccus pyogenes. *J Bacteriol,* Volume 67, pp. 86-89.

Stern, A. et al., 2010. Self-targeting by CRISPR: gene regulation or autoimmunity?. *Trends Genet,* 26(8), pp. 335-340.

Stevens, D. L., 2000. Streptococcal toxic shock syndrome associated with necrotizing fasciitis. *Annu Rev Med,* Volume 51, pp. 271-288.

Swedo, S. E. et al., 1998. Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. *Am J Psychiatry,* 155(2), pp. 264-71.

Szklarczyk, D. et al., 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res,* Volume 43, pp. 447-452.

Trieu-Cuot, P., Carlier, C., Poyart-Salmeron, C. & Courvalin, P., 1991. Shuttle vectors containing a multiple cloning site and a la&a gene for conjugal transfer of DNA from Escherichia coli to Gram-positive bacteria. *Gene,* Volume 102, pp. 99-104.

van de Rijn, . I. & Kessler, R. E., 1980. Growth characteristics of group A streptococci in a new chemically defined medium. *Infect Immun,* Volume 27, pp. 444-448.

van der Oost, J., Westra, . E. R., Jackson, R. N. & Wiedenheft, B., 2014. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nature Reviews Microbiology,* Volume 12, pp. 479-92.

Vincze, T., Posfai, J. & Roberts, R. J., 2003. NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic acid research,* Volume 31, pp. 3688-3691.

Walker, M. J. et al., 2014. Disease Manifestations and Pathogenic Mechanisms of Group A Streptococcus. *Clin Microbiol Rev,* Volume 27, pp. 264-301.

Weinberg, Z. et al., 2017. Bioinformatic analysis of riboswitch structures uncovers variant classes with altered ligand specificity. *PNAS,* Volume 114, pp. 2077-2085.

Winkler, W., Nahvi, A. & Breaker, R. R., 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature,* Volume 428, pp. 281-286.

Winter, J. E. & Berwheimer, A. W., 1963. The deoxyribonucleases of Streptococcus pyogenes. *The journal of biological chemistry,* Volume 239, pp. 215-222.

Wright, P. R. et al., 2013. Comparative genomics boosts target prediction for bacterial small RNAs. *PNAS,* Volume 110, pp. 3487-3496.

Zhao, Y. et al., 2008. Construction of two mathematical models for prediction of bacterial sRNA targets. *Biochem Biophys Res Commun,* Volume 372, pp. 346-50.

7 Appendix

Appendix 7-1. Representation of 3 major types and 10 subtypes of known CRISPR system (Makarova, et al., 2011).
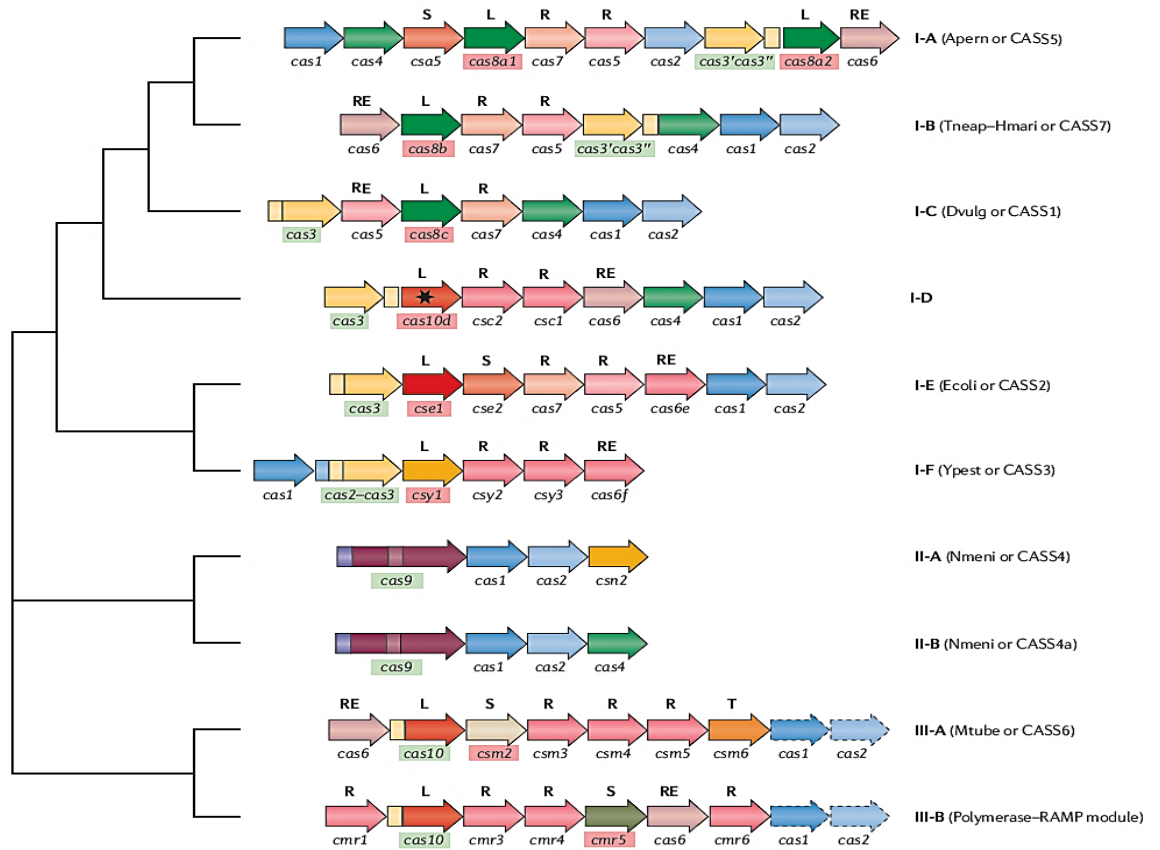
Appendix 7-2. A) Schematic of pFW11-*luc2* plasmid. B) Restriction enzyme map in pFW11 plasmid (Vincze, et al., 2003).
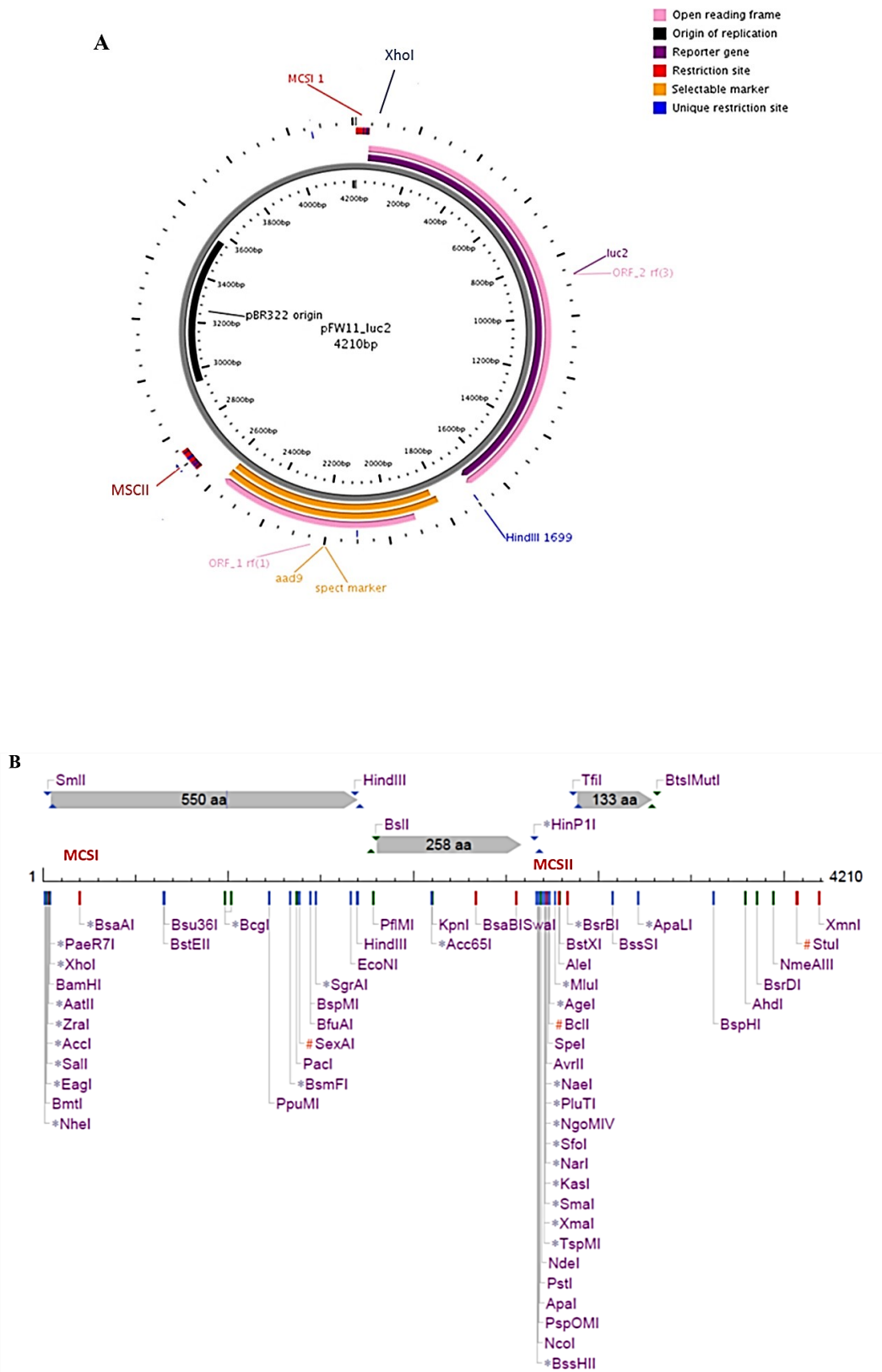
ID   pFW11_*luc2* (4210 bp)

DE   MCSI: POS.1-44; LUC: POS.45-1694; *aad9*: POS.1828-2577;

MCSII: POS.2666-2741

ID   pFW11_*luc2* (4210 bp)

DE   MCSI: POS.1-44; luc: POS.45-1694; aad9: POS.1828-2577;

MCSII: POS.2666-2741

MSCI: TGTACAATTGCTAGCGTACGGCCGTCGACGTCCGGATCCTCGAG**ATG**GAAGACGCCAAAAAC

SQ   SEQUENCE   4210 BP; 1287 A; 866 C; 955 G; 1102 T; 0 OTHER;

```
TGTACAATTG CTAGCGTACG GCCGTCGACG TCCGGATCCT CGAGATGGAA GACGCCAAAA
ACATAAAGAA AGGCCCGGTG CCATTCTATC CGCTAGAGGA TGGAACCGCT GGAGAGCAAC
TGCATAAGGC TATGAAGAGA TACGCCCTGG TTCCTGGAAC AATTGCTTTT ACAGATGCAC
ATATCGAGGT GAACATCACG TACGCGGAAT ACTTCGAAAT GTCCGTTCGG TTGGCAGAAG
CTATGAAACG ATATGGGCTG AATACAAATC ACAGAATCGT CGTATGCAGT GAAAACTCTC
TTCAATTCTT TATGCCGGTG TTGGGCGCGT TATTTATCGG AGTTGCAGTT GCGCCCGCGA
ACGACATTTA TAATGAACGT GAATTGCTCA ACAGTATGAA CATTTCGCAG CCTACCGTAG
TGTTTGTTTC CAAAAAGGGG TTGCAAAAAA TTTTGAACGT GCAAAAAAAA TTACCAATAA
TCCAGAAAAT TATTATCATG GATTCTAAAA CGGATTACCA GGGATTTCAG TCGATGTACA
CGTTCGTCAC ATCTCATCTA CCTCCCGGTT TTAATGAATA CGATTTTGTA CCAGAGTCCT
TTGATCGTGA CAAAACAATT GCACTGATAA TGAACTCCTC TGGATCTACT GGGTTACCTA
AGGGTGTGGC CCTTCCGCAT AGAACTGCCT GCGTCAGATT CTCGCATGCC AGAGATCCTA
TTTTTGGCAA TCAAATCATT CCGGATACTG CGATTTTAAG TGTTGTTCCA TTCCATCACG
GTTTTGGAAT GTTTACTACA CTCGGATATT TGATATGTGG ATTTCGAGTC GTCTTAATGT
ATAGATTTGA AGAAGAGCTG TTTTTACGAT CCCTTCAGGA TTACAAAATT CAAAGTGCGT
TGCTAGTACC AACCCTATTT TCATTCTTCG CCAAAAGCAC TCTGATTGAC AAATACGATT
TATCTAATTT ACACGAAATT GCTTCTGGGG GCGCACCTCT TTCGAAAGAA GTCGGGGAAG
CGGTTGCAAA ACGCTTCCAT CTTCCAGGGA TACGACAAGG ATATGGGCTC ACTGAGACTA
CATCAGCTAT TCTGATTACA CCCGAGGGGG ATGATAAACC GGGCGCGGTC GGTAAAGTTG
TTCCATTTTT TGAAGCGAAG GTTGTGGATC TGGATACCGG GAAAACGCTG GGCGTTAATC
AGAGAGGCGA ATTATGTGTC AGAGGACCTA TGATTATGTC CGGTTATGTA AACAATCCGG
AAGCGACCAA CGCCTTGATG GACAAGGATG GATGGCTACA TTCTGGAGAC ATAGCTTACT
GGGACGAAGA CGAACACTGC TGCATAGTTG ACCGCTTGAA GTCTTTAATT AAATACAAAG
GAGACCAGGT GGCCCCCGCT GAATTGGAGT CGATATTGTT ACAACACCCC AACATCTTCG
```

ACGCGGGCGT GGCAGGTCTT CCCGACGATG ACGCCGGTGA ACTTCCCGCC GCCGTTGTTG

TTTTGGAGCA CGGAAAGACG ATGACGGAAA AAGAGATCGT GGATTACGTC GCCAGTCAAG

TAACAACCGC CAAAAAGTTG CGCGGAGGAG TTGTGTTTGT GGACGAAGTA CCGAAAGGTC

TTACCGGAAA ACGCGACGCA AGAAAAATCA GAGAGATCCT CATAAAGGCC AAGAAGGGCG

GAAAGTCCAA ATTGTAAAAG CTTGATTTTC GTTCGTGAAT ACATGTTATA ATAACTATAA

CTAATAACGT AACGTAACGT AAAATTTGTT TGATTTGTCC AAAATATGGT ATAATAGGTA

CTAATCAAAA TAGTGAGGAG GATATATTTG AATACATACG AACAAATTAA TAAAGTGAAA

AAAATACTTC GGAAACATTT AAAAAATAAC CTTATTGGTA CTTACATGTT TGGATCAGGA

GTTGAGAGTG GACTAAAACC AAATAGTGAT CTTGACTTTT TAGTCGTCGT ATCTGAACCA

TTGACAGATC AAAGTAAAGA AATACTTATA CAAAAAATTA GACCTATTTC AAAAAAAATA

GGAGATAAAA GCAACTTACG ATATATTGAA TTAACAATTA TTATTCAGCA AGAAATGGTA

CCGTGGAATC ATCCTCCCAA ACAAGAATTT ATTTATGGAG AATGGTTACA AGAGCTTTAT

GAACAAGGAT ACATTCCTCA GAAGGAATTA AATTCAGATT TAACCATAAT GCTTTACCAA

GCAAAACGAA AAAATAAAAG AATATACGGA AATTATGACT TAGAGGAATT ACTACCTGAT

ATTCCATTTT CTGATGTGAG AAGAGCCATT ATGGATTCGT CAGAGGAATT AATAGATAAT

TATCAGGATG ATGAAACCAA CTCTATATTA ACTTTATGCC GTATGATTTT AACTATGGAC

ACGGGTAAAA TCATACCAAA AGATATTGCG GGAAATGCAG TGGCTGAATC TTCTCCATTA

GAACATAGGG AGAGAATTTT GTTAGCAGTT CGTAGTTATC TTGGAGAGAA TATTGAATGG

ACTAATGAAA ATGTAAATTT AACTATAAAC TATTTAAATA ACAGATTAAA AAAATTATAA

AAAAATTGAA AAAATGGTGG AAACACTTTT TTCAATTTTT TTGTTTTATT ATTTAATATT

TGGGAAATAT TCATTCTAAT TGGTAAGCGC GCCATGGGCC CTGCAGCATG CATATGCATC

CGGAGTTCCC GGGCGCCGGC CTAGGACTAG TGATCACCGG TCGCTCTTCC GCTTCCTCGC

TCACTGACGC GTAAAAAAGG CCCACAAAAG TGGGCCTTTT TTCTCGCTGC GCTCGGTCGT

TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC

AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA

AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA

TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC

CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC

CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG

TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA

CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC

GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC

AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT TTGGTATCTG

CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA

AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA

```
AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA

CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT

AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG

TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT

AGTTGCCTGA CTCCCCGTCG TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC

CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT CCAGATTTAT CAGCAATAAA

CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCGATA AACCCAGCGA ACCATTTGAG

GTGATAGGTA AGATTATACC GAGGTATGAA AACGAGAATT GGACCTTTAC AGAATTACTC

TATGAAGCGC CATATTTAAA AAGCTACCAA GAAAAAAAGC GGCCAAAAG GCCTGGTGCC

ACCAGGCCAA AAAGGCCGCT TTTTTTCGAA GAGGATGAAG AGGATGAGGA GGCAGATTGC

CTTGAATATA TTGACAATAC TGATAAGATA ATATATCTTT TATATAGAAG ATTTTCGTTC

GTGAATACAT
```

Appendix 7-3. pFW11 plasmid.
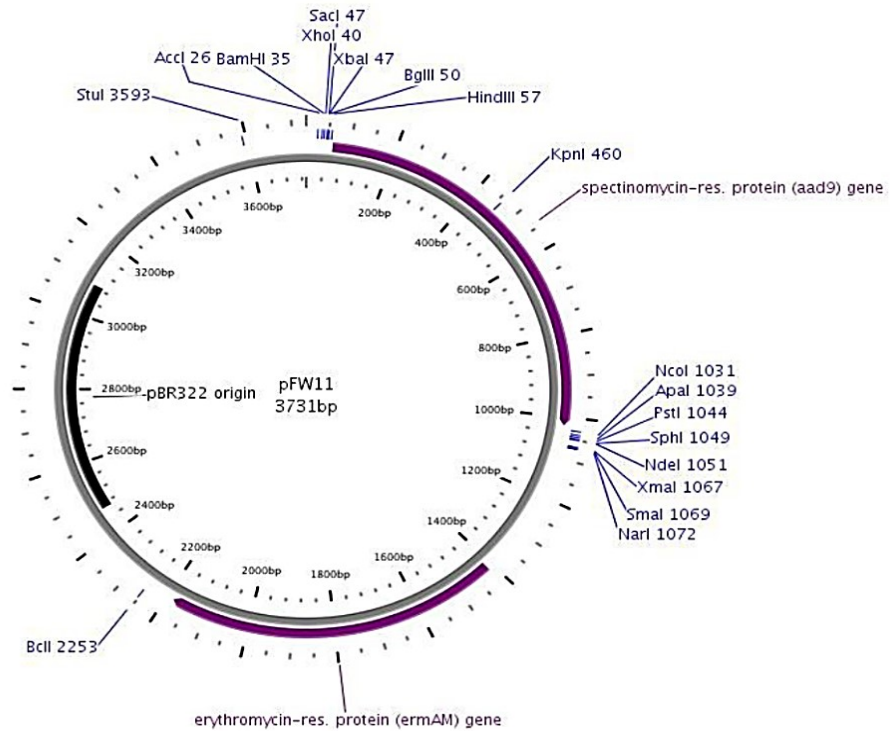
Appendix 7-4. Schematic of pAT18/19 plasmid.
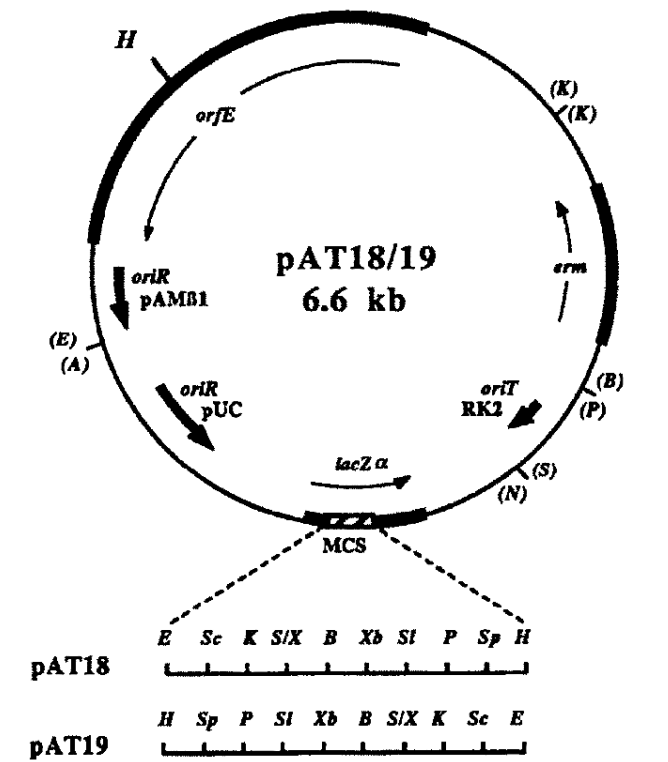
Appendix 7-5. luciferase gene sequence and the locations of designed primers.

Luc-check-rev	CCA TGA TAA TAA TTT TCT GGA
RT_Luc_for	GAG ACA TAG CTT ACT GGG ACG
RT_Luc_rev	TAT CGA CTC CAA TTC AGC GGG

```
ATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGTGCCATTCTATCCGCTAGAGGATGGAACCGCTGGAG
AGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATAT
CGAGGTGAACATCACGTACGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATAT
GGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGG
GCGCGTTATTTATCGGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAG
TATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGCAA
AAAAAATTACCAATAATCCAGAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGA
TGTACACGTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTCCTTTGA
TCGTGACAAAACAATTGCACTGATAATGAACTCCTCTGGATCTACTGGGTTACCTAAGGGTGTGGCCCTT
CCGCATAGAACTGCCTGCGTCAGATTCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAATCATTCCGG
ATACTGCGATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGAT
ATGTGGATTTCGAGTCGTCTTAATGTATAGATTTGAAGAAGAGCTGTTTTTACGATCCCTTCAGGATTAC
AAAATTCAAAGTGCGTTGCTAGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGACAAAT
```

```
ACGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAAAGAAGTCGGGGAAGCGGT
TGCAAAACGCTTCCATCTTCCAGGGATACGACAAGGATATGGGCTCACTGAGACTACATCAGCTATTCTG
ATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTCCATTTTTTGAAGCGAAGGTTG
TGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAGAGAGGCGAATTATGTGTCAGAGGACCTATGAT
TATGTCCGGTTATGTAAACAATCCGGAAGCGACCAACGCCTTGATGGACAAGGATGGATGGCTACATTCT
GGAGACATAGCTTACTGGGACGAAGACGAACACTGCTGCATAGTTGACCGCTTGAAGTCTTTAATTAAAT
ACAAAGGAGACCAGGTGGCCCCCGCTGAATTGGAGTCGATATTGTTACAACACCCCAACATCTTCGACGC
GGGCGTGGCAGGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGGAGCACGGA
AAGACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCCAAAAAGTTGCGCG
GAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAAAACGCGACGCAAGAAAAATCAGAGA
GATCCTCATAAAGGCCAAGAAGGGCGGAAAGTCCAAATTG
```

Appendix 7-6. *pcrA, ribogly and* Na[+]/Alanine symporter sequences and the position of predicted elements and designed primers:

TTA: Start of the predicted first aptamer

ACT: Start of the predicted second aptamer

AAG: End of the second aptamer

First Terminator

Second Terminator

T: Predicted transcriptional start site

TTTTGGTTGACTTCATTCGTTTTTTGAGGTATGATAGTCC: predicted promoter

Sequences and locations of designed primers within *ribogly* and Na[+]/Alanine symporter

RT_*ribogly*_for	CTGGAGAGACCTTATTAGGC
RT_*ribogly*_rev	GAGAGATTGAGCAGTATGCC
T7.*ribogly*_rev	cttaatacgactcactatagg GAGAGATTGAGCAGTATGCC
GlyLuc_Xho_rev	CCT CGAGCT CCT AAA AGT CAT CGA AGA CG
Ribo.prom. BamHl-fwd	TCGGATCCATGCGCTTCATCAACG
Ribo.prom. xhol-rev	CCTCGAGAATTAGGACTATCATACCTC
RT_AASymporter_for	AATCAGCCTTCACTCCGACAGC
RT_AASymporter-rev	ACTCATTGGAGAAGACACCGCG
AAsymporter.fwd4	GGACGGGGATTTACCTTACC
T7.AASymporter.rev4	cttaatacgactcactataggCAGTTCCGACAGTAGCGGC

>CP000829.1:1000027-1002349 Streptococcus pyogenes NZ131, complete
genome
CCTCGTATCCCTTTTCATCAAGCAAATCATTTTTAGCGTTTGAAATAGTTCCTAAAATAGAACGTTCATT
CCACTTTTTAGGGTCTATATTTAATTGTTTTAAAATGCGCTTCATCAACGTTCGTTGCTCACCTGGATCC
ACGATTGTGAAGTTTCGGTTGTAACCAATGTGATCAGCCTCACGACGCAGAATCCGTACACACATGGAGT
GAAAGGTAGCAATTAAGGTATCTTTAGTAGCTGGATTTAAGGCCAAGGCCCTTTCTTTCATTTCACGCGC
AGCCTTATTAGTAAAGGTAATAGCTAAAATATTCCAAGGATTGACAAATTTTTCATCAATTAAATAGGCA
ATACGATGGGTCAGTACCCTTGTTTTACCAGAACCAGCCCCCGCCATAATCAATAAGGGGCCTTCAGTTG
TTTGAACAGCTTGTGCCTGCCGATCATTCATTCCATTTAATAAAGGATTCATGCTTTCTCCTTAATCTCG
TTCAGTGACATCTCTTTTATCAACCTTAAAATTATATCATGTTTCCAAGAAAAATAGCAAAGAAGACCCC
AAAAGTAAAGAGATTCCTTTTTTCTAGTTTAGGATCAGATATTTAAAAATAATAGCGATTTCTTAAAGCA
TTCCTTATTTTTTTTGCTA**TTTTGGTTGACTTCATTCGTTTTTTGAGGTATGATAGTCCTAATT**TAATAAA
CCGAATGAT**GTCATGCAGGAGAAGAATTT**TTTTCGCCGAAGGAGTTATACTCTCAGGTGTTCAGTTTTTG
AACGGGACTGTTTGATGGACGGACTT**CTGGAGAGACCTTATTAGGC**GCCGAAGGGGCAAGGCATACTGCT
CAATCTCTCAGGCAAAAGGACAGAAGGTAAAATACAAACACCATTAAGAACAGTCTTAGTCTTTTTTGTG
TTTGCTGTTTTATCATTGCTTCAGAAGTTGTCTCAAAGAAAGAGATAGCTTTTTTCTTTTGGCGTCTTCG
ATGACTTTTAGGAGAGAAAGAATGATAGCACTCGTTAAATTAATTGATAACCTTGTTTGGGGACCGCCCC
TCTTAATTTTATTGGTTGGGACGGGGATTTACCTTACCAGTCGTTTAGGATTAATTCAAATCTTAAAACT
ACCAAGAGCCTTTAAACTCATTTTTTCAGATGACGAAGGACATGGAGATATTTCATCCTTTGCTGCTCTT
GCAACTGCCCTTGCCGCTACTGTCGGAACTGGTAACATTGTTGGGGTTGCCACTGCTATCAAGTCTGGTG
GTCCTGGAGCGCTCTTTTGGATGTGGGTTGCCGCTTTTTTTGGAATGGCCACTAAGTACGCTGAAGGGGT
ACTAGCGATTAAATACCGTACAAAGGATGCCAACGGTCATATTTCTGGTGGTCCTATGTACTATATTGTT
AATGGCATGGGGACAAAATGGAAACCCTTAGCCATTCTTTTTGCGGGTTCAGGGATCCTAGTTGCCCTAT
TTGGAATAGGAACCTTTGCTCAAGTGAACTCGATCACCTCGTCTTTAGGTCATAGCTTTGGCTTATCCCC
ACAAATGGTCAGTATTGTTTTAGCTATCTTTGTGGCTGCCATTATTTTTGGTGGCATTCATTCCATTTCC
AAAGTGGCTGAAAAAGTCGTTCCTTTTATGGCTATTTTTTATATTCTATCTAGTTTAGCAGTTATTTTTA
GTCACTACCAGCAGCTACTTCCTGTTATTCGTTTAGTCTTTCAATCAGCCTTCACTCCGACAGCTGCTAT
TGGGGGCTTTGCTGGAAGTCTCATGAAGGATGCCATTCAAAAAGGAATTGCACGCGGTGTCTTCTCCAAT
GAGTCTGGACTAGGATCTGCACCTATTGCTGCAGCAGCAGCTAAAACCAATGAACCCGTGGAGCAGGGAC
TGATCTCCATGACAGGAACTTTTATTGATACGATTATTATCTGTACCCTAACCGGCCTATCTATTTTAGT
AACAGGCCAATGGACTGGGCAGCTCGAAGGGGCTCCTCTGACCCAATCTGCCTTTGCCACTGTGTTTGGA
AATTTAGGGACTTTTGGCCTAACCTTTTCCCTAGTTCTCTTCGCCTTTACAACGATTTTAGGATGGAGTT
ATTATGGAGAACGCTGCTTTGAATTCCTCTTTGGCATTACACATCTCACCTACTTCCGTATTGTCTTTAT
CTTAATGGTAGGCCTAGGAGGCTTCCTTAAGTTAGAACTAATTTGGGTTTTGGCTGATATTGTCAATGGT
TTAATGGCACTTCCTAACCTGATTGCCCTCCTAGCTCTCT**CGCCAGTTGTTATTTTAGAAACC**AAGCATT
ACTTTATTAAGTAA

Appendix 7-7. Cation efflux system protein (spy49-1009) sequence and the locations of designed primers.

Ffflux.joint.rev	**GATTGGCATCGGCTGGTTGGC**
Efflux.fwd	GTCTCAACGAGCTAGGACC
Efflux.rev	GGCAGGCTCAACGTGAATG

```
>CP000829.1:1002469-1003704 Streptococcus pyogenes NZ131
ATGACACAGGATCCGATTGCAAATTTAAAAACTAGCCAGAAAGGGCCCAATCGTTAGTATTATTGTCTACT
TGTCGCTTAGTGTTGCCAAATTATTAGCTGGTTATTTGTTAAATGCAAGTTCCCTTATTGCTGATGGATT
TAACAATTTATCGGATATTGTGGGAAATGTAGCCCTGCTTATTGGTCTTCACTTAGCTAGCCAACCAGCC
GATGCCAATCATAAATTTGGTCATTGGAAAATTGAAGACTTATCCAGCCTTGTCACTTCTTTTATTATGT
TTCTTGTAGGTTTCCAAGTACTGATTCACACAATTAAAAGTATCTTTAGTGGTCAGCAAGTTGATATTGA
CCCTCTTGGGGCTATTGTCGGTATCGTTTCAGCTTTTGTTATGTTAGGGGTTTATGTCTTTAACAAACGT
CTTTCCAAACGTGTAAAATCCAGTGCCTTAGTCGCTGCTTCTAAGGATAATCTAGCTGATGCTGTTACTT
CTATCGGAACATCAATTGCTATTATAGCAGCTTCTTTGCATTTACCAGTTATCGATCATATAGCTGCTAT
GATCATTACGTTCTTTATTCTTAAAACAGCTTTTGATATCTTTATGGAAAGTTCGTTTAGTTTATCTGAT
GGATTTGATAGCCGTCATTTGAAAAAATCCGAAAAAGCCATTTTAGAAATCCCTAAAATTGTCGCCGTTA
AGTCTCAACGAGCTAGGACCTATGGTAGCAATGTCTATCTTGATATTGTACTTGAAATGAATCCTGATCT
TTCAGTCTATGAAAGTCACTCTATTACAGAGAAAGTGGAGCAGTTATTGAGTGACCAATTTTCTATTTAT
GACATTAACATTCACGTTGAGCCTGCCATGATTCCCGAAGAAGAGATTTTTGATAATGTCGCCAAAAAGC
TCTACCGCTACGAAAAATTAATTTGGAGTAAGGTTCCTGACTATGACCACTACATTGCTAAGTCTTTCCA
ACTGATTGATGCGAATGGCCAAACAGTTAACTATGAACAATTTTGGAACCAAGAAATTTATTATCCAAGT
AACTTCAACCATTTTCAGATTGAATCCATTAGTCAAAAAACGATGTTGGTAACTTACCAATTAAATGGCA
ATCAACGTACCAGTATTTGGAGGCGTCATGAATCTTGGTCCTTACTCTTCCACCAAATTACCCCTATCGC
TAAGAGACAATTACATCACACACTATCGTATTGTAAAAATGTAA
```

Appendix 7-8. Sequencing result of *S. pyogenes* M49/pFW11- promoter*ribogly*-*luc2* transformant. luc-check.rev primer.

Streptococcus pyogenes NZ131, complete genome
Sequence ID: <u>CP000829.1</u>Length: 1815785Number of Matches: 1
Related Information
Range 1: 1000135 to 1000719

Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
1075 bits(582)	0.0	584/585(99%)	0/585(0%)	Plus/Minus

```
Query   430    AATTAGGACTATCATACCTCAAAAAACGAATGAAGTCAACCAAAATAGCaaaaaaaTAAG    489
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000719 AATTAGGACTATCATACCTCAAAAAACGAATGAAGTCAACCAAAATAGCAAAAAAATAAG   1000660

Query   490    GAATGCTTTAAGAAATCGCTATTATTTTTAAATATCTGATCCTAAACTAGAAAAAAGGAA    549
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000659 GAATGCTTTAAGAAATCGCTATTATTTTTAAATATCTGATCCTAAACTAGAAAAAAGGAA   1000600

Query   550    TCTCTTTACTTTTGGGGTCTTCTTTGCTATTTTTCTTGGAAACATGATATAATTTTAAGG    609
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000599 TCTCTTTACTTTTGGGGTCTTCTTTGCTATTTTTCTTGGAAACATGATATAATTTTAAGG   1000540

Query   610    TTGATAAAAGAGATGTCACTGAACGAGATTAAGGAGAAAGCATGAATCCTTTATTAAATG    669
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000539 TTGATAAAAGAGATGTCACTGAACGAGATTAAGGAGAAAGCATGAATCCTTTATTAAATG   1000480

Query   670    GAATGAATGATCGGCAGGCACAAGCTGTTCAAACAACTGAAGGCCCCTTATTGATTATGG    729
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000479 GAATGAATGATCGGCAGGCACAAGCTGTTCAAACAACTGAAGGCCCCTTATTGATTATGG   1000420

Query   730    CGGGGGCTGGTTCTGGTAAAACAAGGGTACTGACCCATCGTATTGCCTATTTAATTGATG    789
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct   1000419 CGGGGGCTGGTTCTGGTAAAACAAGGGTACTGACCCATCGTATTGCCTATTTAATTGATG   1000360

Query   790    AAAAATTTGTCAATCCTTGGAATATTTTAGCTATTACCTTTACTAATAAGGCTGCGCGTG    849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  1000359  AAAAATTTGTCAATCCTTGGAATATTTTAGCTATTACCTTTACTAATAAGGCTGCGCGTG  1000300

Query  850      AAATGAAAGAAAGGGCCTTGGCCTTAAATCCAGCTACTAAAGATACCTTAATTGCTACCT  909
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1000299  AAATGAAAGAAAGGGCCTTGGCCTTAAATCCAGCTACTAAAGATACCTTAATTGCTACCT  1000240

Query  910      TTCACTCCATGTGTGTACGGATTCTGCGTCGTGAGGCTGATCACATTGGTTACAACCGAA  969
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1000239  TTCACTCCATGTGTGTACGGATTCTGCGTCGTGAGGCTGATCACATTGGTTACAACCGAA  1000180

Query  970      ACTTCACAATCGTGGATCCAGGTGAGCAACGAACGTTGATGACGC  1014
                ||||||||||||||||||||||||||||||||||||||||||| ||
Sbjct  1000179  ACTTCACAATCGTGGATCCAGGTGAGCAACGAACGTTGATGAAGC  100013
```

Appendix 7-9. Sequencing result of Na$^+$/Alanine symporter and cation efflux system junction fragment. AA.Efflux.joint.fwd primer.

Streptococcus pyogenes NZ131, complete genome
Sequence ID: CP000829.1Length: 1815785Number of Matches: 1
Related Information
Range 1: 1002342 to 1002688

Alignment statistics for match #1

Score   Expect   Identities   Gaps   Strand

630 bits(341) 6e-177 346/348(99%) 2/348(0%) Plus/Plus

```
Query  8        TTATAGT-AACATGACGTTAGTACTACCCAGTTATACCTAGTCTTCTTCTGCCAAGTTCA  66
                ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002342  TTA-AGTAAACATGACGTTAGTACTACCCAGTTATACCTAGTCTTCTTCTGCCAAGTTCA  1002400

Query  67       AGACGCTAGGCTTTTTTCTTAAGCCCCTTTTTATGGTATAATATAAAGTTGAAATCAAAG  126
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002401  AGACGCTAGGCTTTTTTCTTAAGCCCCTTTTTATGGTATAATATAAAGTTGAAATCAAAG  1002460

Query  127      GAGTTATTATGACACAGGATCCGATTGCAAATTTAAAACTAGCCAGAAAGGGCCCAATCG  186
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002461  GAGTTATTATGACACAGGATCCGATTGCAAATTTAAAACTAGCCAGAAAGGGCCCAATCG  1002520

Query  187      TTAGTATTATTGTCTACTTGTCGCTTAGTGTTGCCAAATTATTAGCTGGTTATTTGTTAA  246
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002521  TTAGTATTATTGTCTACTTGTCGCTTAGTGTTGCCAAATTATTAGCTGGTTATTTGTTAA  1002580

Query  247      ATGCAAGTTCCCTTATTGCTGATGGATTTAACAATTTATCGGATATTGTGGGAAATGTAG  306
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002581  ATGCAAGTTCCCTTATTGCTGATGGATTTAACAATTTATCGGATATTGTGGGAAATGTAG  1002640

Query  307      CCCTGCTTATTGGTCTTCACTTAGCTAGCCAACCAGCCGATGCCAATC  354
                ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1002641  CCCTGCTTATTGGTCTTCACTTAGCTAGCCAACCAGCCGATGCCAATC  1002688
```

Appendix 7-10. Sequencing result of 3´flanking region of *cand34* in WTΔ*cand34* using Spec-weg-for primer.

Streptococcus pyogenes NZ131, complete genome
Sequence ID: CP000829.1Length: 1815785Number of Matches: 1
Range 1: 1200583 to 1201475

Score   Expect   Identities   Gaps   Strand

1633 bits(884)   0.0   890/894(99%) 1/894(0%) Plus/Minus

```
Query  119      TGCGCCACTAGCTTTCGGTAAACTAAGGGTAACTTAAGAGAGGAGTTATTTATGACACTA  178
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201475  TGCGCCACTAGCTTTCGGTAAACTAAGGGTAACTTAAGAGAGGAGTTATTTATGACACTA  1201416

Query  179      TCACTTGTATTAATGGTATCGGTTTTTGGGGCAGGCTTGTTATCATTTTTCTCTCCGTGT  238
                |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201415  TCACTTGTAT-AATGGTATCGGTTTTTGGGGCAGGCTTGTTATCATTTTTCTCTCCGTGT  1201357
```

```
Query    239     ATCTTTCCGGTCCTTCCTGTCTATTTAGGCATTTTACTGGATGCAGATGATTCAAAGACG    298
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201356 ATCTTTCCGGTCCTTCCTGTCTATTTAGGCATTTTACTGGATGCAGATGATTCAAAGACG    1201297

Query    299     ATTACTATATTTGGTAAAAAACTCTATTGGTATGGCATTGTCAAAACCTTAGCCTTTATT    358
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201296 ATTACTATATTTGGTAAAAAACTCTATTGGTATGGCATTGTCAAAACCTTAGCCTTTATT    1201237

Query    359     TTTGGTCTATCTACTATTTTTGTGATTTTAGGTTACGGGGCTGGTTTTTTAGGAAATATC    418
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201236 TTTGGTCTATCTACTATTTTTGTGATTTTAGGTTACGGGGCTGGTTTTTTAGGAAATATC    1201177

Query    419     CTTTATGCGGTATGGTTTCGGTACTTACTGGGAGCCGTGGTCATTATCTTGGGGATTCAT    478
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201176 CTTTATGCGGTATGGTTTCGGTACTTACTGGGAGCCGTGGTCATTATCTTGGGGATTCAT    1201117

Query    479     CAGATGGGCCTCATTACCATTAAGAGCTTACAATTTCAAAAATCACTGACTTTCCATAAT    538
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201116 CAGATGGGCCTCATTACCATTAAGAGCTTACAATTTCAAAAATCACTGACTTTCCATAAT    1201057

Query    539     AACAAGAATCGCAACGGTTTGTTCAATGCATTTATCCTTGGTTTAACCTTTAGCTTTGGT    598
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1201056 AACAAGAATCGCAACGGTTTGTTCAATGCATTTATCCTTGGTTTAACCTTTAGCTTTGGT    1200997

Query    599     TGGACGCCTTGTGTGGGACCTGTTTTGAGTTCTGTTCTAGCTTTGGTGGCTTCAGGGGGA    658
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1200996 TGGACGCCTTGTGTGGGACCTGTTTTGAGTTCTGTTCTAGCTTTGGTGGCTTCAGGGGGA    1200937

Query    659     AATGGTGCCTGGCAAGGTGGCGTTTTGATGATTATTTATACTCTTGGATTGGGCATTCCT    718
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1200936 AATGGTGCCTGGCAAGGTGGCGTTTTGATGATTATTTATACTCTTGGATTGGGCATTCCT    1200877

Query    719     TTCCTGCTTATCTCTTTTGCGTCAGGCATTGTTTTGAAACAGTTTAACAAGCTCAAACCC    778
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1200876 TTCCTGCTTATCTCTTTTGCGTCAGGCATTGTTTTGAAACAGTTTAACAAGCTCAAACCC    1200817

Query    779     CACATGCTTTTACTGAANAAAGTAGGAGGCGTTCTGATTATAGTCATGGGAATCTTGCTT    838
                 |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Sbjct    1200816 CACATGCTTTTACTGAAAAAAGTAGGAGGCGTTCTGATTATAGTCATGGGAATCTTGCTT    1200757

Query    839     ATGACAGGAACCTTAAATAACTTAGCACAACTTTTTGGATAAAGGAGAAACACAATGAAN    898
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1200756 ATGACAGGAACCTTAAATAACTTAGCACAACTTTTTGGATAAAGGAGAAACACAATGAAA    1200697

Query    899     AAAGGACTATTAGTAACAACTGGTTTGGCTTGTCTCGGGCTACTAACTGCTTGCTCAACC    958
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1200696 AAAGGACTATTAGTAACAACTGGTTTGGCTTGTCTCGGGCTACTAACTGCTTGCTCAACC    1200637

Query    959     CAAGACAATATGGCTAANAAGGAAATAACTCAGGACAAGATGAGCATGGCAGCT    1012
                 ||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct    1200636 CAAGACAATATGGCTAAAAAGGAAATAACTCAGGACAAGATGAGCATGGCAGCT    1200583
```

Appendix 7-11. Sequencing result of 5´flanking region of *cand34* in WTΔ*cand34* using MCSI-for primer.

Streptococcus pyogenes NZ131, complete genome

Sequence ID: CP000829.1Length: 1815785Number of Matches: 1

 Range 1: 1202004 to 1202863

Score	Expect	Identities	Gaps	Strand
1585 bits(858)	0.0	859/860(99%)	0/860(0%)	Plus/Minus

```
Query    59      CTTCCAAGGTCGGAGTAAACGACCACCCTTAGATTGTGTTAATGCCCTCTTGTCTTTTGG    118
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1202863 CTTCCAAGGTCGGAGTAAACGACCACCCTTAGATTGTGTTAATGCCCTCTTGTCTTTTGG    1202804

Query    119     TTACAGTTTACTGACCTTTGAATGTCAATCTGCCTTGGAAGCTGTCGGATTAGACAGTTA    178
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1202803 TTACAGTTTACTGACCTTTGAATGTCAATCTGCCTTGGAAGCTGTCGGATTAGACAGTTA    1202744

Query    179     CGTTGGTTTCTTTCACACGGATCGTCCTGGGCGTGCTAGTTTAGCGCTTGATTTAGTTGA    238
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct    1202743 CGTTGGTTTCTTTCACACGGATCGTCCTGGGCGTGCTAGTTTAGCGCTTGATTTAGTTGA    1202684
```

```
Query  239       AGAGTTCCGCTCATATATTGTAGATCGTTTTGTCTTTTCATTAATTAATAAAGGACAACT  298
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202683   AGAGTTCCGCTCATATATTGTAGATCGTTTTGTCTTTTCATTAATTAATAAAGGACAACT  1202624

Query  299       TCAGAAAAAACACTTTGAGGTTAAAGAAAATGGTAGTATTTTATTGACGGAAAATGGCAG  358
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202623   TCAGAAAAAACACTTTGAGGTTAAAGAAAATGGTAGTATTTTATTGACGGAAAATGGCAG  1202564

Query  359       AGCTATTTTTATTGATTTGTGGCAGAAGCGTAAGCATACTGAGGTAGAACATCCTTTTAC  418
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202563   AGCTATTTTTATTGATTTGTGGCAGAAGCGTAAGCATACTGAGGTAGAACATCCTTTTAC  1202504

Query  419       AAAAGAAAAAGTAAAACTTATGTTATTACCCTATGTACAAGCGCAGCTTTTAGCTAAGGC  478
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202503   AAAAGAAAAAGTAAAACTTATGTTATTACCCTATGTACAAGCGCAGCTTTTAGCTAAGGC  1202444

Query  479       CATACGAGGAGATTTAGAAAGCTATCCACCTTTTATGGTTTAGGAGATGTTATATGATGG  538
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202443   CATACGAGGAGATTTAGAAAGCTATCCACCTTTTATGGTTTAGGAGATGTTATATGATGG  1202384

Query  539       TTTTAGTCACTTATGATGTAAATACGGAAACACCTGCTGGTAGAAAAAGATTGCGTCATG  598
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202383   TTTTAGTCACTTATGATGTAAATACGGAAACACCTGCTGGTAGAAAAAGATTGCGTCATG  1202324

Query  599       TTGCCAAACTCTGTGTGGACTATGGGCAACGTGTTCAAAATTCTGTTTTTGAATGTTCTG  658
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202323   TTGCCAAACTCTGTGTGGACTATGGGCAACGTGTTCAAAATTCTGTTTTTGAATGTTCTG  1202264

Query  659       TGACACCCGCAGAATTTGTGGATATAAAGCACCGCTTAACACAAATCATTGATGAGAAAA  718
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202263   TGACACCCGCAGAATTTGTGGATATAAAGCACCGCTTAACACAAATCATTGATGAGAAAA  1202204

Query  719       CTGATAGTATTCGCTTTTATTTATTGGGGAAAAATTGGCAGAGGCGTGTGGAAACACTTG  778
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202203   CTGATAGTATTCGCTTTTATTTATTGGGGAAAAATTGGCAGAGGCGTGTGGAAACACTTG  1202144

Query  779       GTCGCTCAGACAGCTATGACCCAGATAAAGGTGTCTTATTATTGTAAAAATCTCTTGTGC  838
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202143   GTCGCTCAGACAGCTATGACCCAGATAAAGGTGTCTTATTATTGTAAAAATCTCTTGTGC  1202084

Query  839       GAAGCTAGCTTTCACAGAAACACCTTGCTTGCTCGCGCAAAAATAACTTAAAAAAGAAGC  898
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202083   GAAGCTAGCTTTCACAGAAACACCTTGCTTGCTCGCGCAAAAATAACTTAAAAAAGAAGC  1202024

Query  899       GAAATGGAGATAANAAGGCT  918
                 ||||||||||||| ||||||
Sbjct  1202023   GAAATGGAGATAAAAAGGCT  1202004
```

Appendix 7-12. Sequencing result of *cand34* in pAT19-*cand34*.

A: using Compl _M16/c34_Bam_fwd primer.

Streptococcus pyogenes NZ131, complete genome
Sequence ID: CP000829.1Length: 1815785Number of Matches: 12
Range 1: 1201501 to 1201949

Score	**Expect**	**Identities**	**Gaps**	**Strand**
708 bits(383)	0.0	429/451(95%)	3/451(0%)	Plus/Plus

```
Query  27        CTTGAATTTGATACGCTATCAGATAAATGGTTGACTAAAGGCATGTAAAAACCCGACCTC  86
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201501   CTTGAATTTGATACGCTATCAGATAAATGGTTGACTAAAGGCATGTAAAAACCCGACCTC  1201560

Query  87        TCTTTTAATGAGGTCAGGTTCATCTTTCAATCCACTCACCCATGAAGGGTGAGACAAGAT  146
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201561   TCTTTTAATGAGGTCAGGTTCATCTTTCAATCCACTCACCCATGAAGGGTGAGACAAGAT  1201620

Query  147       GACACTGGTACTGCACATGTCGATTTAAAAAATTTCAATCCACTCACCCATGAAGGGTGAG  206
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1201621   GACACTGGTACTGCACATGTCGATTTAAAAAATTTCAATCCACTCACCCATGAAGGGTGAG  1201680

Query  207       ACAGATCTTAAAATGGATTCTTCTTCAGATATTTTTGATTTCAATCCACTCACCCATGAA  266
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201681 ACAGATCTTAAAATGGATTCTTCTTCAGATATTTTTGATTTCAATCCACTCACCCATGAA 1201740

Query   267 GGGTGAGACGATAACAGTTGCTTTAGTCGATAAGTCGATTAGCGATTTCAATCCACTCAC 326
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201741 GGGTGAGACGATAACAGTTGCTTTAGTCGATAAGTCGATTAGCGATTTCAATCCACTCAC 1201800

Query   327 CCATGAAGGGTGAGACATTATGTTTTGCCACATGAGAAAGTAAAAAATGGAATTTCAATC 386
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201801 CCATGAAGGGTGAGACATTATGTTTTGCCACATGAGAAAGTAAAAAATGGAATTTCAATC 1201860

Query   387 CACTCACCCATGAAGGGTGAGACGCTGTGACATTGCGGGATGTAATCAAAG-TAAAAAAT 445
             |||||||||||||||||||||| |  |  |  | ||  |   || |  | |   |   ||
Sbjct 1201861 CACTCACCCATGAAGGGTGAGACTCCCTTATAAT-CGACAAA-AAGCGCCGATTGATTAT 1201918

Query   446 TTCAATCCACTCACCCATGAAGGGTGAGACT  476
             ||||||||||||||||| |||||||||||||
Sbjct 1201919 TTCAATCCACTCACCCACGAAGGGTGAGACT  1201949
```

B: using Compl _M16/c34_Sall_rev primer

Streptococcus pyogenes NZ131, complete genome
Sequence ID: <u>CP000829.1</u>Length: 1815785
Related Information
Range 1: 1201501 to 1201949

Score	**Expect**	**Identities**	**Gaps**	**Strand**
708 bits(383)	0.0	429/451(95%)	3/451(0%)	Plus/Minus

```
Query   301 AGTCTCACCCTTCATGGGTGAGTGGATTGAAATTTTTTA-CTTTGATTACATCCCGCAAT 359
             ||||||||||||| |||||||||||||||||||||  |  |  |  | ||    | || |
Sbjct 1201949 AGTCTCACCCTTCGTGGGTGAGTGGATTGAAATAATCAATCGGCGCTT-TTTGTCG-ATT 1201892

Query   360 GTCACAGCGTCTCACCCTTCATGGGTGAGTGGATTGAAATTCCATTTTTTACTTTCTCAT 419
              | |   | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201891 ATAAGGGAGTCTCACCCTTCATGGGTGAGTGGATTGAAATTCCATTTTTTACTTTCTCAT 1201832

Query   420 GTGGCAAAACATAATGTCTCACCCTTCATGGGTGAGTGGATTGAAATCGCTAATCGACTT 479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201831 GTGGCAAAACATAATGTCTCACCCTTCATGGGTGAGTGGATTGAAATCGCTAATCGACTT 1201772

Query   480 ATCGACTAAAGCAACTGTTATCGTCTCACCCTTCATGGGTGAGTGGATTGAAATCAAAAA 539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201771 ATCGACTAAAGCAACTGTTATCGTCTCACCCTTCATGGGTGAGTGGATTGAAATCAAAAA 1201712

Query   540 TATCTGAAGAAGAATCCATTTTAAGATCTGTCTCACCCTTCATGGGTGAGTGGATTGAAA 599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201711 TATCTGAAGAAGAATCCATTTTAAGATCTGTCTCACCCTTCATGGGTGAGTGGATTGAAA 1201652

Query   600 TTTTTAAATCGACATGTGCAGTACCAGTGTCATCTTGTCTCACCCTTCATGGGTGAGTGG 659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201651 TTTTTAAATCGACATGTGCAGTACCAGTGTCATCTTGTCTCACCCTTCATGGGTGAGTGG 1201592

Query   660 ATTGAAAGATGAACCTGACCTCATTAAAAGAGAGGTCGGGTTTTTACATGCCTTTAGTCA 719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct 1201591 ATTGAAAGATGAACCTGACCTCATTAAAAGAGAGGTCGGGTTTTTACATGCCTTTAGTCA 1201532

Query   720 ACCATTTATCTGATAGCGTATCAAATTCAAG  750
             |||||||||||||||||||||||||||||||
Sbjct 1201531 ACCATTTATCTGATAGCGTATCAAATTCAAG  1201501
```

Appendix 7-13. Target prediction of sRNA *can34* using the online tool CobraRNA (Wright, et al., 2013)**.**

Rank	CopraRNA p-value	Locus Tag	Gene Name	Energy [kcal/mol]	PositionmRNA	Position sRNA	Annotation
1.	1.557e-10	spy49_rs05965	N/A	-24.93	1 -- 20	69 -- 88	cytochrome C biogenesis protein CcdA
2.	0.0008724	spy49_rs06360	N/A	-15.54	136 -- 151	233 -- 248	NAD(+) synthetase
3.	0.0009724	spy49_rs07985	N/A	-15.59	260 -- 290	97 -- 126	peptide deformylase
4.	0.001523	spy49_rs04125	glmM	-15.25	232 -- 247	232 -- 248	phosphoglucosamine mutase
5.	0.00157	spy49_rs04330	N/A	-15.69	143 -- 153	32 -- 42	spermidine/purescine ABC transporter permease
6.	0.002337	spy49_rs02545	N/A	-14.76	244 -- 254	238 -- 248	PTS beta-glucoside transporter subunit EIIBCA
7.	0.002864	spy49_rs06845	N/A	-14.30	187 -- 221	115 -- 144	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C
8.	0.004982	spy49_rs05985	N/A	-13.33	16 -- 68	98 -- 156	type I-C CRISPR-associated protein Cas7/Csd2
9.	0.005282	spy49_rs00965	N/A	-12.99	172 -- 193	119 -- 140	L-xylulose 5-phosphate 3-epimerase
10.	0.007091	spy49_rs01155	N/A	-13.21	42 -- 54	75 -- 87	DNA repair protein RadA

Appendix 7-14. Target prediction of the sRNA *cand34* using the online tool TargetRNA2 (Kery & Feldman, 2014).

Rank	p-value	Energy [kcal/mol]	Position mRNA	Position sRNA	gene	Synonym	Annotation
1.	0.005	-12.65	(-70) - (-61)	54-45	pyrC	Spy49_0718	dihydroorotase
2.	0.006	-12.43	1-13	12-1	-	Spy49_1488c	Small terminase
3.	0.014	-10.97	(-1) - 12	56-43	rpmF	Spy49_1770	50S ribosomal protein L32
4.	0.015	-10.85	(-5) - 7	47- 36	-	Spy49_1556c	DAK2 domain-containing protein
5.	0.025	-9.85	(-34) - (-19)	279- 264	rpsE	Spy49_0065	30S ribosomal protein S5
6.	0.028	-9.61	(-48) - (-34)	36- 22	exoA	Spy49_0334C	3-exo-deoxyribonuclease
7.	0.033	-9.28	(-76) - (-68)	47- 39	hutG	Spy49_1731	Formimidoylglutamase
8.	0.036	-9.08	1-10	42-33	-	Spy49_0953c	hypothetical protein
9.	0.039	-8.89	(-9) - 6	15-2	-	Spy49_0018	glycerol-3-phosphate acyltransferase PlsX
10.	0.044	-8.62	(-70) - (-57)	270- 256		Spy49_0946	DNA-binding protein

Appendix 7-15. Target prediction of the sRNA *cand34* using the online tool RNApredator (Zhao, et al., 2008).

Rank	Energy [kJ/mol]	z-Score	mRNA [Start]	mRNA [End]	sRNA	Gene Annotation	Accession
1.	-29.12	-9.34	-153	-124	1-30	Cytochrome c type biogenesis protein Spy49	NC_011375
2.	-16.66	-3.75	-187	-172	32-47	Aspartate carbamoyltransferase Spy49	NC_011375
3.	-16.16	-3.52	223	243	31-51	D lactate dehydrogenase Spy49	NC_011375
4.	-16.07	-3.48	19	36	233-250	NAD synthetase Spy49	NC_011375
5.	-15.58	-3.26	638	659	119-140	Hexulose 6 phosphate synthase Spy49	NC_011375
6.	-15.41	-3.19	-58	-48	32-42	Spermidine Putrescine ABC transporter permease component potC Spy49	NC_011375
7.	-15.25	-3.12	733	743	32-42	Spermidine Putrescine ABC transporter permease component potB Spy49	NC_011375
8.	-15.15	-3.07	1133	1148	32-47	Putative uracil permease Spy49	NC_011375
9.	-14.99	-3	1249	1265	35-53	Putative cationic amino acid transporter protein Spy49	NC_011375
10.	-14.99	-3	53	81	116-144	Aspartyl tRNA(Asn) amidotransferase subunit C Spy49	NC_011375
11.	-14.90	-2.96	809	829	299-325	Putative protoporphyrinogen oxidase Spy49	NC_011375
12.	-14.83	-2.93	943	964	300-323	hypothetical protein Spy49 1048c Spy49	NC_011375
13.	-14.67	-2.86	384	401	244-261	hypothetical protein Spy49 1765c Spy49	NC_011375
14.	-14.42	-2.74	1382	1400	233-251	Putative salivaricin A modification enzyme; amino acid dehydration Spy49	NC_011375
15.	-14.41	-2.74	342	355	265-278	LSU ribosomal protein L18p (L5e) Spy49	NC_011375
16.	-14.30	-2.69	-186	-175	35-47	Seryl tRNA synthetase Spy49	NC_011375
17.	-14.25	-2.67	102	125	32-53	Putative Holliday junction resolvase Spy49	NC_011375
18.	-14.08	-2.59	-192	-173	232-251	hypothetical protein Spy49 1390c Spy49	NC_011375
19.	-14.00	-2.55	836	847	35-46	Two component system histidine kinase Spy49	NC_011375
20.	-13.89	-2.51	117	139	32-53	Glycerol uptake facilitator protein Spy49	NC_011375
21.	-13.88	-2.5	976	991	38-53	GTP binding and nucleic acid binding protein YchF Spy49	NC_011375
22.	-13.85	-2.49	813	828	238-253	DNA polymerase III delta prime subunit Spy49	NC_011375
23.	-13.83	-2.48	1218	1239	232-255	hypothetical protein Spy49 1238 Spy49	NC_011375
24.	-13.76	-2.45	753	768	38-52	Penicillin binding protein 7 precursor (PBP 7) (D alanyl D alanine endopeptidase) Spy49	NC_011375
25.	-13.69	-2.42	-199	-177	301-324	Transcriptional regulator; DeoR family Spy49	NC_011375
26.	-13.62	-2.38	954	968	38-52	Putative esterase Spy49	NC_011375
27.	-13.52	-2.34	1686	1705	32-53	Putative heavy metal transporting ATPase Spy49	NC_011375
28.	-13.47	-2.32	2343	2355	241-253	DNA mismatch repair protein MutS Spy49	NC_011375
29.	-13.44	-2.3	-106	-90	35-53	Histidine ammonia lyase Spy49	NC_011375
30.	-13.43	-2.3	292	310	32-51	hypothetical protein Spy49 0961c Spy49	NC_011375
31.	-13.43	-2.3	67	76	299-308	Putative chromosome segregation SMC Spy49	NC_011375
32.	-13.41	-2.29	-123	-109	236-250	ComX2 alternate sigma factor Spy49	NC_011375
33.	-13.40	-2.29	466	483	35-51	Peptide methionine sulfoxide reductase msrA Spy49	NC_011375
34.	-13.39	-2.28	716	733	116-133	Putative positive transcriptional regulator Spy49	NC_011375
35.	-13.27	-2.23	628	646	35-52	hypothetical protein Spy49 1801c Spy49	NC_011375

36.	-13.22	-2.21	29	43	37-51	Putative copper transporting ATPase Spy49	NC_011375
37.	-13.21	-2.2	739	753	34-47	hypothetical protein Spy49 1266c Spy49	NC_011375
38.	-13.21	-2.2	160	179	32-51	hypothetical protein Spy49 0160 Spy49	NC_011375
39.	-13.13	-2.16	-15	-7	300-308	UDP N acetylglucosamine 1 carboxyvinyltransferase Spy49	NC_011375
40.	-13.12	-2.16	-33	-20	265-278	SSU ribosomal protein S5p (S2e) Spy49	NC_011375
41.	-13.12	-2.16	117	127	43-53	Immunogenic secreted protein Spy49	NC_011375
42.	-13.03	-2.12	-136	-118	300-318	hypothetical protein Spy49 0347 Spy49	NC_011375
43.	-13.03	-2.12	28	44	231-248	Phosphoglucosamine mutase Spy49	NC_011375
44.	-12.99	-2.1	841	862	31-50	Probable regulatory protein Spy49	NC_011375
45.	-12.99	-2.1	1115	1125	299-309	Putative rhamnosyltransferase Spy49	NC_011375
46.	-12.96	-2.09	43	53	238-248	Beta glucoside permease IIABC component Spy49	NC_011375
47.	-12.94	-2.08	518	539	30-49	Antirepressor Spy49	NC_011375
48.	-12.88	-2.05	2089	2101	235-247	Extracellular hyaluronate lyase Spy49	NC_011375
49.	-12.86	-2.04	426	439	33-46	Probable dipeptidase Spy49	NC_011375
50.	-12.81	-2.02	-29	-8	119-140	Hexulose 6 phosphate isomerase Spy49	NC_011375
51.	-12.75	-1.99	98	115	37-53	Putative ATP dependent exonuclease; subunit A Spy49	NC_011375
52.	-12.71	-1.98	305	318	313-325	Putative peptidoglycan branched peptide synthesis protein; serine/alanine adding enzyme Spy49	NC_011375
53.	-12.67	-1.96	864	875	35-47	Aerobic glycerol 3 phosphate dehydrogenase Spy49	NC_011375
54.	-12.62	-1.94	-58	-34	235-259	Putative secreted protein Streptputative secreted protein Spy49	NC_011375
55.	-12.60	-1.93	112	121	1-10	hypothetical protein Spy49 0777 Spy49	NC_011375
56.	-12.57	-1.91	393	411	32-52	Transcriptional regulator; GntR family Spy49	NC_011375
57.	-12.52	-1.89	56	66	238-248	hypothetical protein Spy49 1218c Spy49	NC_011375
58.	-12.51	-1.89	-83	-62	232-255	Alpha mannosidase Spy49	NC_011375
59.	-12.48	-1.87	238	249	243-254	Putative efflux protein Spy49	NC_011375
60.	-12.46	-1.86	1026	1043	35-51	Amidophosphoribosyltransferase Spy49	NC_011375
61.	-12.45	-1.86	-14	6	299-325	Sua5/YciO/YrdC family protein Spy49	NC_011375
62.	-12.43	-1.85	341	355	241-255	hypothetical protein Spy49 0801 Spy49	NC_011375
63.	-12.43	-1.85	263	283	299-325	Iron sulfur cluster assembly protein SufD Spy49	NC_011375
64.	-12.42	-1.85	2285	2304	299-325	Putative DNA mismatch repair protein Spy49	NC_011375
65.	-12.40	-1.84	-31	-17	233-248	hypothetical protein Spy49 0352c Spy49	NC_011375
66.	-12.31	-1.8	159	175	243-259	hypothetical protein Spy49 1435c Spy49	NC_011375
67.	-12.30	-1.79	40	49	1-10	hypothetical protein Spy49 0756 Spy49	NC_011375
68.	-12.28	-1.78	267	283	33-51	Ribonuclease HII Spy49	NC_011375
69.	-12.28	-1.78	49	56	300-307	hypothetical protein Spy49 1497c Spy49	NC_011375
70.	-12.27	-1.78	358	373	32-47	hypothetical protein Spy49 1229c Spy49	NC_011375

71.	-12.24	-1.77	642	653	315-326	Transmembrane histidine kinase CsrS Spy49	NC_011375
72.	-12.23	-1.76	132	154	65-88	Putative transcriptional regulator Spy49	NC_011375
73.	-12.18	-1.74	674	701	300-325	hypothetical protein Spy49 0740 Spy49	NC_011375
74.	-12.16	-1.73	360	369	1-10	Lactate oxidase Spy49	NC_011375
75.	-12.15	-1.73	115	127	34-46	Putative PTS dependent galactosamine IID component Spy49	NC_011375
76.	-12.13	-1.72	-169	-161	32-40	Low molecular weight protein tyrosine phosphatase Spy49	NC_011375
77.	-12.12	-1.71	963	974	67-78	Two component sensor histidine kinase Spy49	NC_011375
78.	-12.08	-1.69	381	402	301-322	hypothetical protein Spy49 0642 Spy49	NC_011375
79.	-12.08	-1.69	273	281	300-308	ComX1 Alternate Sigma Factor Spy49	NC_011375
80.	-12.07	-1.69	352	363	66-77	Glyoxalase family protein Spy49	NC_011375
81.	-12.05	-1.68	884	902	32-47	hypothetical protein Spy49 0623 Spy49	NC_011375
82.	-12.04	-1.68	421	430	244-253	Endonuclease III Spy49	NC_011375
83.	-12.01	-1.66	256	275	35-53	SSU ribosomal protein S2p (SAe) Spy49	NC_011375
84.	-12.00	-1.66	493	507	244-258	Protein dltB Spy49	NC_011375
85.	-11.97	-1.64	96	117	31-51	Putative permease Spy49	NC_011375
86.	-11.92	-1.62	431	452	30-53	Protein dltD precursor Spy49	NC_011375
87.	-11.91	-1.62	440	448	300-308	Recombination protein Spy49	NC_011375
88.	-11.89	-1.61	632	643	242-253	Alcohol dehydrogenase Spy49	NC_011375
89.	-11.88	-1.6	183	208	301-325	hypothetical protein Spy49 1380c Spy49	NC_011375
90.	-11.87	-1.6	35	48	36-48	Regulatory protein spx Spy49	NC_011375
91.	-11.82	-1.58	1187	1207	303-324	Putative ABC transporter Spy49	NC_011375
92.	-11.80	-1.57	-174	-164	241-251	Beta glucosidase Spy49	NC_011375
93.	-11.79	-1.56	1694	1700	300-306	Putative ABC transporter; ATP binding protein Spy49	NC_011375
94.	-11.78	-1.56	190	202	238-250	Acetyl CoA acetyltransferase Spy49	NC_011375
95.	-11.78	-1.56	685	705	238-258	Laminin binding surface protein Spy49	NC_011375
96.	-11.77	-1.55	545	559	244-257	Sugar ABC transporter; permease protein Spy49	NC_011375
97.	-11.77	-1.55	709	717	32-40	Putative multi drug resistance efflux pump Spy49	NC_011375
98.	-11.77	-1.55	44	66	32-51	Ribonuclease Z Spy49	NC_011375
99.	-11.75	-1.55	111	120	239-248	hypothetical protein Spy49 1343c Spy49	NC_011375
100.	-11.75	-1.55	56	73	238-258	Zinc binding protein adcA precursor Spy49	NC_011375

Abreviations

*aad*9	resistance gene for spectinomycin
Ala	Alanine
bp	base pair
°C	Celcius / centigrade
CaCl2	Calcium Chloride
CFU	Colony Forming Unit
CO2	carbondioxide
cDNA	Conmlementary DNA
DMEM	Dulbecco's Modified Eagle's Medium
DNA	deoxyribonucleic acid
dNTP	dideoxynucleoside triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetraacetic Acid
FCS	Fetal Calf Serum
GAS	Group A *Streptococcus*
Gly	Glycine
h	hour
LB	Luria Bertani
luc	luciferase
Mg	Magnesium
min	minute
mM	millimolar
n	nano (10^{-9})
NaCl	sodium chloride
NaOH	sodium hydroxide
nm	nometer
OD	Optical Density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
pH	power of Hydrogen
RLU	Relative Light Unit
RNA	ribonucleic acid
rpm	revolutions per minute
RT-qPCR	Revers Transcriptase- quantitative PCR
S	Second
S. pyogenes	*Streptococcus pyogenes*
Ser	Serine
THB	Todd Hewitt Broth
THY	Todd Hewitt Yeast
Tween 20	polyoxyethylene sorbital monolaurate
UV	Ultraviolet
WT	Wild Type
μ	micro (10-6)

Acknowledgment