university of groningen

Riboswitch control: the regulatory function of prokaryotic RNAs

An opportunity in medicine, biotechnology and evolutionary biology

Bachelor Thesis by Jakob Viel

Supervisor: Prof. Dr. Jan Kok

TABLE OF CONTENTS

Abstract

The expression of certain genes is regulated by riboswitches. These are RNA sequences situated prior to the ribosome binding site in messenger RNA. When binding a specific ligand, the riboswitch undergoes a conformation change which influences gene expression by inhibiting or stimulating transcription or translation. In this paper, the mechanisms of action employed by riboswitches, and the applicability of riboswitches in biotechnology, medicine and understanding evolutionary history is explored. It was concluded that riboswitches have high potential when it comes to biotechnology, might be useful for the development of new antibiotics, and while not overwhelming, add to the case of the RNA world hypothesis.

Introduction

Introduction to RNA

The role of RNA in translation is well-established; in the process of converting DNA into protein, RNA plays the crucial role of an intermediate, since DNA itself cannot be converted directly into protein. The sequence encoding a protein is copied from the DNA in the form of RNA, which is then translated by ribosomes into the amino acid chain that subsequently folds into the protein (Alberts *et al.,* 2008). In translation, multiple types of RNA are involved. They will be described shortly in the following section.

RNAs in translation

In translation, multiple roles are performed by RNAs and RNA-protein complexes. Some RNAs are involved in carrying the protein encoding sequence, other RNAs read this sequence and direct the amino acids corresponding to it, while still others link these amino acids into a polypeptide (Madigan *et al., 2012*) **(Figure 1).**

Messenger RNAs (mRNA)

In transcription, messenger RNA is the nucleotide sequence that is copied from a DNA sequence encoding a protein. It is this RNA sequence that is translated into a protein.

Making a copy from the DNA in the form of RNA has many advantages, caused by corresponding molecular characteristics. DNA, holding the genetic information of the organism, is more stable because DNA's deoxyribose is less reactive than RNA's ribose. DNA also usually is double stranded, adding to its stability. This stability is crucial for conserving genetic information, but impedes the possibilities when it comes to retrieving it (Alberts *et al.,* 2008).

The use of "disposable" RNA copies from the DNA for protein synthesis makes it possible to regulate expression by transcription quantity. Also, the fact that RNA is single stranded and its ribose is more reactive, allows for better manipulation and controlled degradation of the sequence. These traits are used to regulate gene expression.

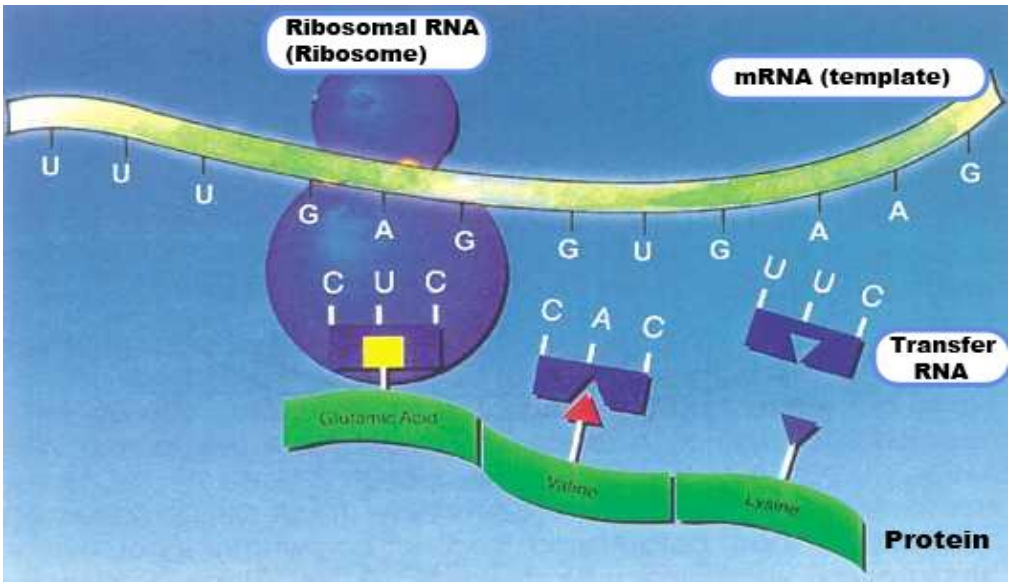

Figure 1 Translation: The basis of translation lies in three different RNAs; two ribozymes and a coding mRNA (Nourysolutions, 2008)

Ribosomal RNA (rRNA)

After mRNA has been synthesized, it needs to be translated into an amino acid sequence. This translation is facilitated by ribosomes. Ribosomes are rRNA-protein complexes in which the rRNA part fulfills mainly a catalytic function while the proteins have a mainly structural purpose **(Figure 2)**. The small ribosome subunit binds the mRNA, followed by assemblage of the ribosome, which catalyzes the formation of peptide bonds between the amino acids that the mRNA encodes. RNAs capable of catalyzing reactions like this are called ribozymes (Madigan *et al., 2012*).

The last step in polypeptide synthesis is finding the amino acids that are encoded in by the mRNA and directing them towards the ribosome, so they can be incorporated into the chain. This role is fulfilled by Transfer RNA.

Transfer RNA (tRNA)

tRNAs are intermediating molecules with a secondary structure, that can recognize both an mRNA sequence encoding an amino acid (codon), as well as the corresponding amino acid. For each codon, a three-nucleotide amino-acid-encoding sequence, there is a corresponding tRNA that recognizes the codon and brings with it the right amino acid **(Figure 3)**. When the tRNA can bind the mRNA, this means the tRNA's anti-codon matches the mRNA codon, and the amino acid that the tRNA carries is incorporated into the amino acid chain (Madigan *et al., 2012*).

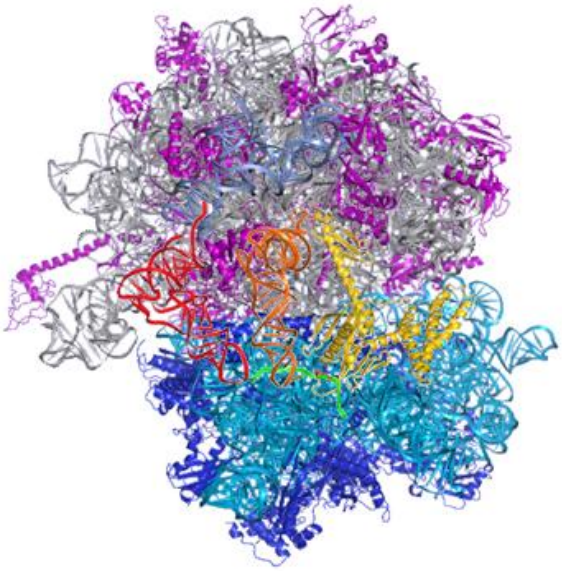

Figure 2 Ribosomal RNA in its folded state. The intermolecular interactions, resulting in a tertiary structure, are crucial for the function of ribosomes (Laurenberg *et al.,* 2008)

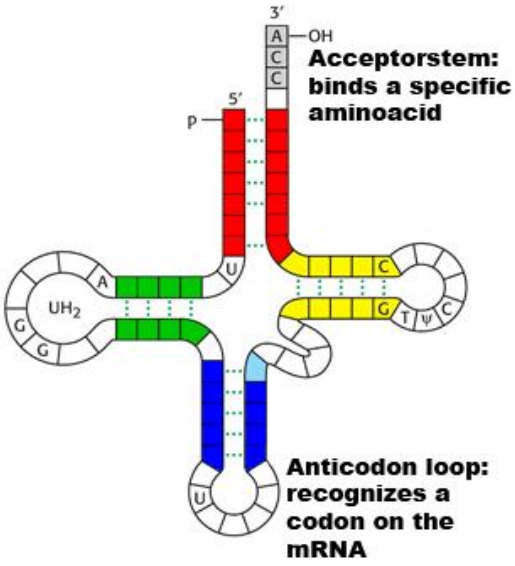

Figure 3 Transfer RNA has a secondary structure that is needed for recognition by e.g. the ribosome. The acceptorstem is "charged" with a specific amino acid corresponding to tRNA anticodon (Darling D., 2013)

Transfer-messenger RNA (tmRNA)

As said earlier, RNA is quite an unstable molecule. Its instability can cause an mRNA strand to break or fold onto itself causing the translating ribosome to jam. In prokaryotes, specialized RNA molecules called transfer-messenger RNAs (tmRNAs), are facilitated to resolve this blockage. They also tag the amino acid chain that is left over from the incomplete or faulty translation for proteolysis. This is achieved in the following way: the tmRNA binds to the ribosome at the normal mRNA binding site and pushes the faulty mRNA out of the ribosome as it is being translated. The incomplete amino acid chain remains bound to

the ribosome and is further elongated by the amino acids that are encoded by the tmRNA. The tmRNA encodes for a proteolysis tag causing the unfinished amino acid chain to be broken down when it is released from the ribosome, which is now no longer jammed and can be used again in translation (Madigan *et al., 2012*).

RNA as a regulator

The previously described functions exemplifying the diversity of RNA, may lead one to believe that RNA is capable of performing more complex regulatory functions, and this has indeed been proven to be the case. Binding and folding capabilities of (m)RNA sequences have been found to be used to regulate expression (Henkin *et al*., 2002). In the next section, a number of RNA regulatory systems will be described.

Bacterial small RNAs (sRNA)

Small RNAs are short RNA sequences that perform a number of regulatory functions in bacteria. They are usually under 250 nt in length and have a tertiary structure that determines their function. These non coding RNAs can, by binding, influence either protein function or mRNA stability and expression. They can, for example, by binding to an target enzyme increase substrate affinity (Vogel and Wager, 2007);(Cao *et al.,* 2010). Small RNAs have been found to perform a function in the regulation of amongst others: quorum sensing, various stress responses and virulence (Lenz *et al.,* 2004);(Ionescu *et al*., 2010).

Clustered Regulatory Interspace Short Palindromic Repeats (CRISPR)

CRISPR are proposed to be a basic prokaryote immune system that stores segments of encountered bacteriophage and plasmid nucleotide sequences. The system is thought to use these segments to recognize similar future threats (Barrangou *et al.,* 2007). CRISPR loci consist of repeat/spacer regions in which the spacers are unique DNA sequences resembling those of specific bacteriophages and other encountered DNA or RNA.

The repeat/spacer regions are usually surrounded by Cas (CRISPR-associated) genes. These genes code for proteins that are believed to bind RNA transcripts from the repeat/spacer regions. The repeat/spacer sequences are transcribed into CRISPR RNAs (crRNAs) which are incorporated into Cas proteins i.e. nucleases, where they function as a recognition site for foreign nucleotide sequences that has been encountered before.

If a DNA sequence is recognized, it is cleaved by the nuclease, and the cleaved nucleotide sequence may then be used to create new spacers, which would keep the defense mechanism "up to date" **(Figure 4)**. And Interestingly, since the CRISPR loci are part of the, inherited, chromosomal DNA, it is an evolutionary strategy resembling a Lamarckian mechanism (Gasianas *et al*., 2014). Lamarck thought that environmental influences would cause changes in the heritable material that would cause offspring to be better adapted to the environment.

However, when purposely engineered to be resistant to specific bacteriophages, the spacers seem not to cause bacteria to become resistant on their own. It is speculated that the spacers may need to be in a certain "genetic context" to be effective (Barrangou *et al.,* 2007).

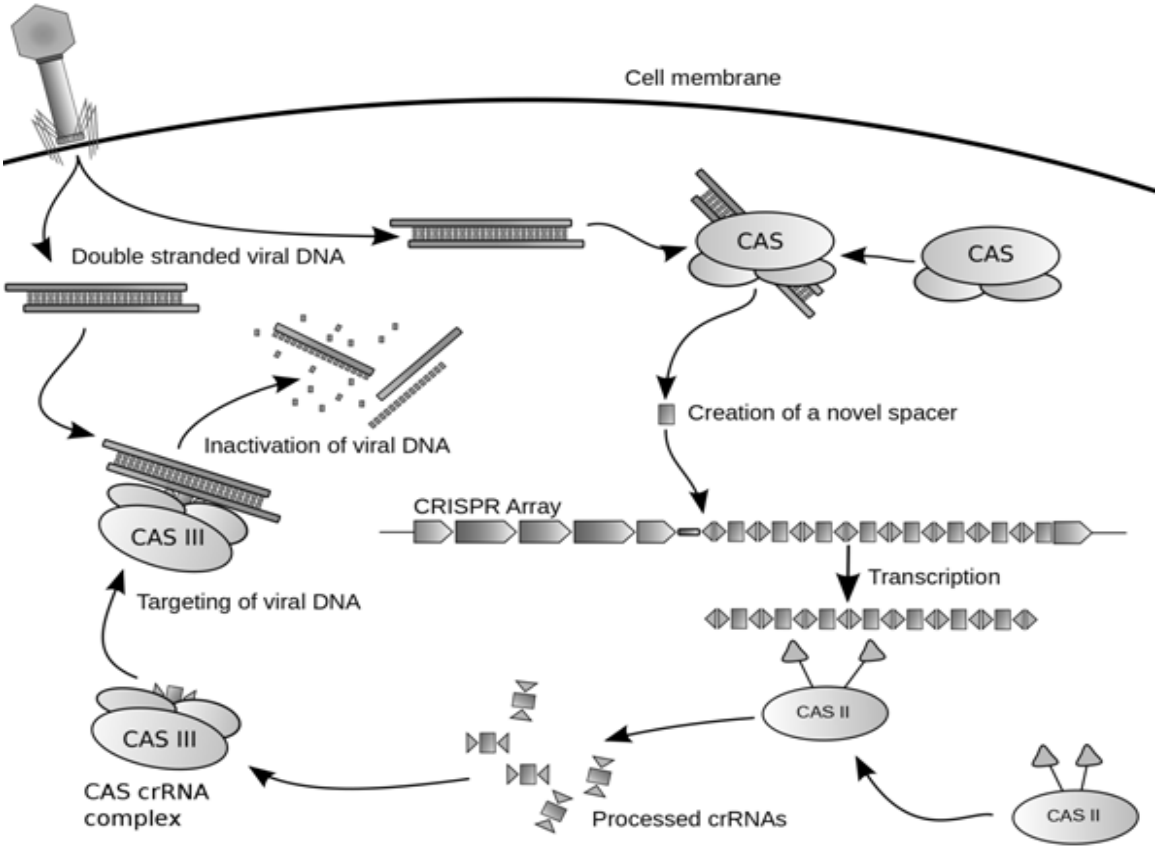

Figure 4 The CRISPR mechanism of action: CAS (CRISPR associated) proteins create a novel spacer from viral DNA, the spacer is inserted to the repeat/spacer region. The repeat/spacer region is transcribed and fragmented by CAS II proteins into CRISPR RNAs (crRNAs). These crRNAs form a complex with CAS III proteins; a nuclease that uses the single stranded bacterial sequence as an recognition site. In this way the CAS III proteins can recognize viruses the cell has encountered before and cleave its DNA. The cleaved DNA can be used to form new spacers, keeping the "immune system" up to date (Horvath and Barrangou, 2010).

RNA thermometer

The RNA thermometer is an RNA sequence usually situated at the 5' untranslated region, the part of the mRNA preceding the ribosome binding site (Shine-Dalgarno sequence). At lower temperatures and thus lower thermodynamic energy, this part of the mRNA folds onto itself, forming a hairpin like structure. When folded, the ribosome binding site cannot be accessed by the ribosome, and in this way translation is prevented. When temperature increases, the hairpin melts, exposing the ribosome binding site and allowing translation **(Figure 5)** (Storz, 1999).

One way in which the RNA thermometer is applied is as a response in expression during cold shock and heat shock. Under high temperatures proteins can denature, causing them to lose their tertiary structure and thus their function. Heat shock proteins, including chaperones and proteases, are expressed in response to this insult. While expression of heat shock proteins is primarily regulated at

transcription level, an RNA thermometer is suspected to work as a fine-tuning mechanism at the translation level (Narberhaus *et al*., 2006).

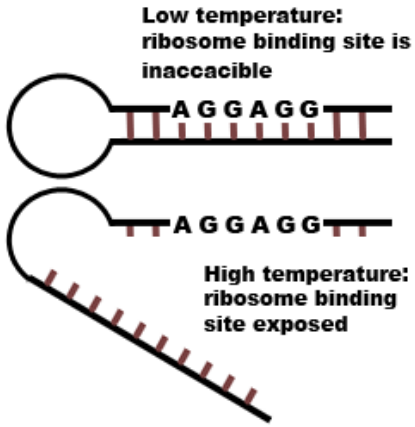

Figure 5 RNA thermometer. When a threshold temperature is reached, the ribosome binding site (Shine-Dalgarno) sequence is exposed.

Current RNA knowledge

In the last few decades, the broad role of RNA in processes other than translation has been uncovered. Non coding RNAs play a role in translation, enzymatic functions, processing functions and the previously described regulatory functions. This study focuses on RNA and its regulatory functions, and because the study of regulatory RNAs is such a broad research field, it is restricted to regulation by RNA in the form of riboswitches.

Research question

As a consequence of the restriction of this subject, the research question addressed to in this thesis is the following:

"*What is the possible impact of riboswitches on biotechnology, medicine and the view on the origin of life?*"

To tackle this question, the subject of this thesis is subdivided in sections, exploring riboswitches and their action mechanisms in general, the possibilities of manipulating their processes to our advantage and the speculative consequences resulting from this knowledge.

Riboswitches

What is a riboswitch

Riboswitches are translation regulating non-coding nucleotide sequences that most commonly precede the translated region of mRNAs, known as the 5' Untranslated Region (5' UTR) or leader sequence. They usually bind ions or molecules relevant to the transcript they carry through recognition by Van der Waals and hydrogen bonds. For example; an mRNA encoding the synthesis operon of a certain amino

acid may be inactivated by that specific amino acid or its derivatives, by binding to the RNA aptameric sequence, a sequence with high affinity for a target molecule, preceding the ribosome binding site (Edwards and Batey, 2010). In this way, there will be no synthesis of molecules that are sufficiently available in the cell. Regulation of translation by riboswitches is very diverse and the mechanism of action and regulatory functions differ per type of riboswitch or, for a homologous group of riboswitches. even per species.

The importance of non-coding RNAs, especially riboswitches, should spark the interest of specialists working in fields like antibiotics development or in biotechnological industries that are e.g. medicine-or-food-related. Regardless of the promising possibilities of manipulating their action mechanisms to our advantage, having knowledge of riboswitches in effect can be most useful in protocol optimization in biotechnology.

Before further exploring the application possibilities, riboswitch diversity will be illustrated in the next section. A number of confirmed riboswitch action mechanisms will be dealt with.

Riboswitch examples; principles of mechanism of action

B_{12} Riboswitch (B12-box)

The biosynthesis of vitamin B_{12} (cobalamin) by bacteria is regulated by the B12-box riboswitch. Both the expression of the outer membrane cobalamin transporter protein, called BtuB, and the cobalamin synthesis operon are regulated by vitamin B_{12} concentrations. In *Salmonella typhimurium,* both transcripts have large, 241-468 nt, leader sequences that, alongside unpreserved regions, contain a 25 nt conserved sequence known as the B_{12}-box (Franklund and Kadner, 1997). The B_{12}-box is involved in forming secondary structures with helices that are presumed to be crucial for metabolite binding.

The proposed mechanisms of regulation are as follows; in gram-negative bacteria regulation of B_{12} synthesis takes place through inhibition of translation. There is an unstable RNA sequence at the 5'UTR of the cobalamin biosynthesis transcript. This sequence can form one of two alternative hairpin structures depending on the cobalamin concentration in the cell. If there is sufficient cobalamin to bind the RNA sequence, a terminator hairpin will form that inhibits transcription. If there is insufficient cobalamin, the antiterminator hairpin is allowed to form, resulting in transcription of the operon **(Figure 6)** (Vitreschak *et al.,* 2003).

In most gram-positive bacteria, regulation is thought to happen at the transcription level; the presence and binding of cobalamin to the leader sequence causes a terminator loop to form and in this way premature termination of transcription.

Sam I, II, III, IV & V Riboswitch (S-box)

First known as the S-box due to its relation to sulphur metabolism, the S-adenosylmethionine (Sam) riboswitch family, which includes at least five classes, regulates the biosynthesis of methionine and cysteine (Grundy and Henking, 1998);(Edwards and Batey, 2010). The different family members have different mechanisms of action and are found in different organisms.

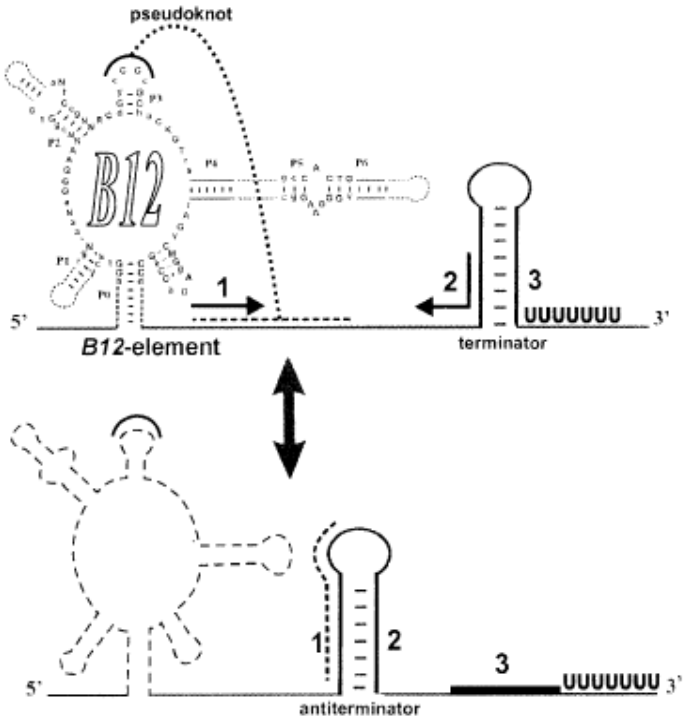

Figure 6 The proposed mechanism of action for the B_{12} riboswitch; When B_{12} binds to the riboswitch, a hairpin structure forms that prevents the antiterminator loop from forming, causing the terminator loop to form and inhibit transcription.(Vitreschak *et al.*, 2003)

Their proposed structures are often a consensus reached by bioinformatics, and Sam-riboswitch classification is not indisputable. It is notable that while Sam has analogs which very much resemble Sam, the riboswitches have a binding affinity for the binding of Sam that ranges from 100 to 1000 fold opposed to its analogs.

Sam-I and Sam-IV

The Sam-I riboswitch is thought to regulate transcription of the methionine and cysteine biosynthesis genes in gram-positive bacteria. When Sam is absent, natural folding of the 5'UTR occurs, and an antiterminator loop is allowed to form, resulting in complete transcription (Grundy and Henking, 1998). In the presence of Sam, the terminator loop is formed, leading to premature termination of transcription. Sam-IV is suspected to have a mechanism of action similar to Sam-I.

Sam-II and Sam-V

The Sam-II riboswitch, that is only present in gram-negative bacteria, uses a machanism that is distincive from that of Sam-1 to regulate expression. The fact that this metabolite-recognizing RNA sequences are so different, indicates convergent evolution of the action mechanisms. The Sam-II riboswitch, which is only 70 nt long, is situated upstream of the methionine biosynthesis operon (Corbino *et al.,* 2005). Although subdivided into several differing action mechanisms, they are all found in a range of

proteobacteria, and in all cases a hairpin is formed, as is a pseudoknot. The pseudoknot surrounds the Sam molecule as an RNA triple helix structure (Lu *et al*., 2008).

For example; in the Sam-II riboswitch, a stem loop is formed when Sam binds to its aptameric sequence, which is situated 11 nt upstream of the start codon of the serine biosynthesis mRNA. This is believed to inhibit translation by making the ribosome binding site unavailable (Corbino *et al.,* 2005). Mg^{2+} stabilizes the structure which allows Sam to bind **(Figure 7)**.

While largely homologous to Sam-II, a group of riboswitches has been classified as Sam-V riboswitches (Meyer *et al*., 2009). Although it likely forms a ligand binding pocket that is similar to that of Sam-II, these mechanisms of action are proposed to be the product of convergent evolution (Poiate *et al*., 2009).

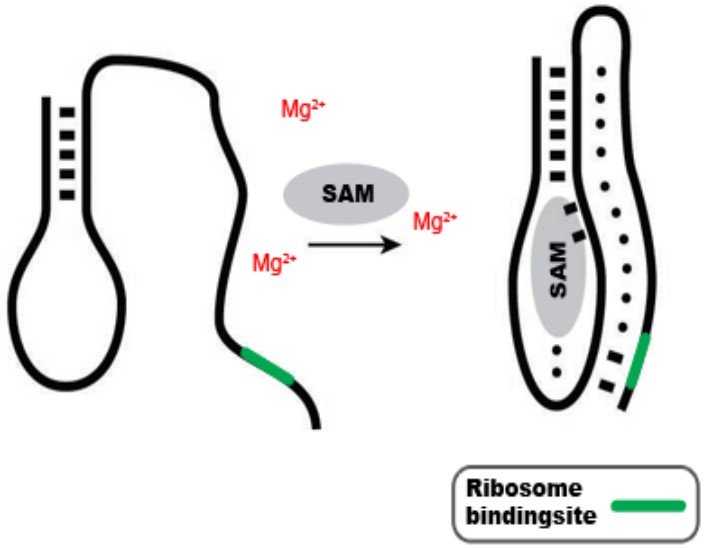

Figure 7 Binding of Sam to the Sam-II riboswitch causes a conformation change that makes the ribosome binding site unavailable (Corbino *et al.,* 2005).

Sam-III

The Sam-III riboswitch or S_{MK} box is found in lactic acid bacteria and regulates *metK* gene translation and in this way the synthesis of Sam. By binding of Sam to the aptameric sequence on the 5' UTR of the *metK* transcript, the 30S ribosomal subunit binding site is shielded and, thus, translation cannot occur **(Figure 8)** (Lu *et al*., 2008).

Thiamin Riboswitch (THI-box)

The thiamin riboswitch, which show large resemblance to the B_{12} riboswitch, is involved in the regulation of thiamin biosynthesis (Miranda-Ríos *et al.*, 2001). The transcript of the *Rhizobium etli thiCOGE* genes, which encode enzymes for thiamin synthesis, contains an 211 nt 5' UTR that includes the 38 nt Thi-box. The Thi-box is also crucial for high thiamin synthesis in the absence of thiamin; apparently it enhances translation. The proposed action mechanism for this riboswitch, which is found in both gram-positive and gram-negative bacteria is as follows; when thiamin is present in sufficient

concentrations, it binds to the riboswitch on its aptameric sequence, which causes a conformational change that subsequently masks the Shine-Dalgarno sequence and in this way prevents translation (Miranda-Ríos *et al.*, 2001).

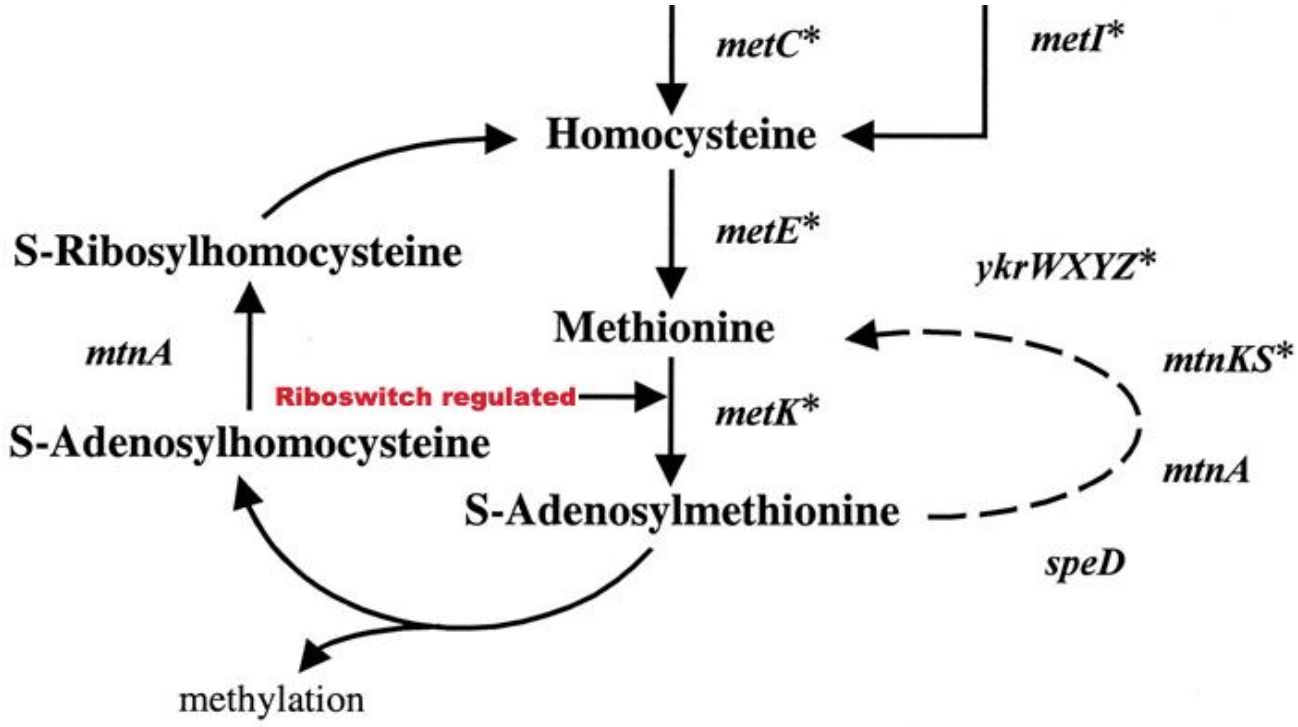

Figure 8 Part of the methionine pathway. Sam biosynthesis is controlled by the SAMIII riboswitch that regulates *metK* expression. (Auger *et al.,* 2002)

Flavin Mono Nucleotide (FMN) Riboswitch

The FMN riboswitch, which regulates the translation of FMN biosynthesis-related transcripts, is a strongly conserved leader sequence that forms five hairpin structures **(Figure 9)**. The hairpin structures fold into a circle in which they radiate outwards (five-way junction), leaving a cavity for FMN to bind. This riboswitch is present in the 5' UTR of the FMN biosynthesis operon (*rib-DEAHT*) mRNA of *Bacilus subtilis*. The FMN riboswitch undergoes a conformational change when binding FMN. This, again, causes the Shine-Dalgarno sequence on the *rib-DEAHT* transcript to be masked and inhibits translation (Winkler *et al*., 2002).

The FMN riboswitch is, contrary to the riboswitches described thus far, confirmed to form a butterfly-like structure in which the FMN is enveloped (Serganov *et al*., 2009). Its 180° rotation symmetry is also a unique feature that, up to now, has mainly been found in ribosomal RNAs.

The ligand binding specificity of the FMN riboswitch is also unusually low. When the cellular concentration of FAD is increased 17-fold, FAD is able to bind to the aptameric sequence instead of

FMN. The plasticity of the ligand binding site makes the FMN riboswitch a probable target for antibiotics.

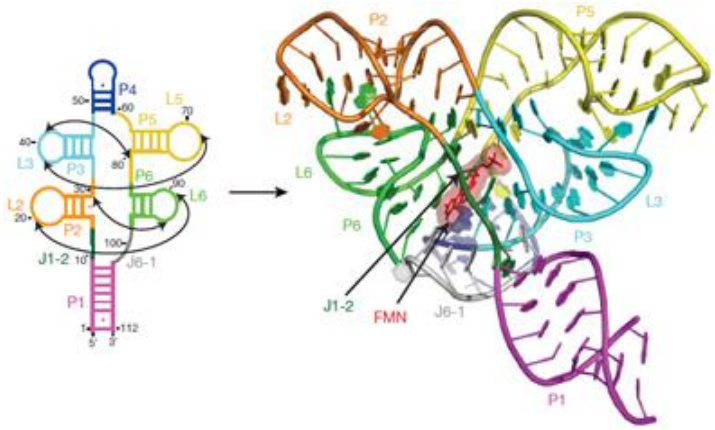

Figure 9 The FMN riboswitch secondary and tertiary butterfly-like structure with FMN bound in the middle (Serganov *et al*., 2009).

Lysine riboswitch

Lysine biosynthesis is also regulated by a riboswitch. Like the FMN riboswitch, the lysine riboswitch or LYS element also contains five hairpin structures that form a cavity where lysine can bind (Rodionov *et al*., 2003). The riboswitch is present in gram-positive and gram-negative bacteria, where it regulates transcription of various lysine biosynthesis operons by termination after it after the binding of lysine. Transcription termination happens in 40% of the transcription attempts in the absence of lysine and in up to 80% of transcription attempts in a saturated lysine concentration. Thus, the lysine riboswitch can suppress transcription by two-fold when sufficient lysine is present in the cell (Garst *et al*., 2012).

The concentration of lysine needed to bind a certain number of riboswitches is much higher than would be expected when taking the equilibrium constant (K_d) into account. The equilibrium constant is a measured number that is unique for every reaction between two different molecules. Using this number the ratio between two products forming, when adding certain concentrations of reagent, can be calculated. The lysine riboswitch does not bind to its ligand as would be expected when using the equilibrium constant, and is thus said to be under kinetic control, opposed to thermodynamic control **(Figure 10)** (Garst *et al*., 2012).

Kinetic versus thermodynamic control

When any of two possible products can form from a reaction, in which one product is more stable than the other, but requires more activation energy, thermodynamics and kinetics determine what product will be formed. In the case of the lysine riboswitch the two products are the 5' UTR aptameric sequence bound to the ligand, and the unbound secondary structure of the 5' UTR i.e. an antiterminator loop.

During transcription, the 5' UTR with the aptameric sequence is synthesized first. Until the RNA is further synthesized, it cannot form its normal secondary structure (antiterminator loop). This gives the ligand of said riboswitch, lysine, time to bind to the aptameric sequence. Transcription has been

measured to even pause for up to 10 seconds to allow for a ligand to bind, after synthesis of the aptameric sequence (Edwards and Batey, 2010). As the lysine riboswitch needs a relatively high concentration of its ligand to bind, this would indicate that it is under kinetic control. The guanine, SAM-I, FMN and tetrahydrofolate are also suspected to be under kinetic control (Garst *et al*., 2012).

However, in a recent study the lysine riboswitch was shown to be influenced by the nucleoside triphosphate (NTP) concentration in the cell. It was found that when NTP concentrations are low, the riboswitch is practically under thermodynamic control. This means that in poor conditions, transcription repression is more prevalent, which would lead to preservation of metabolites when there is little of those available (Garst *et al*., 2012). It must be noted that the latter was found in *in vitro* experiments, which perhaps does not do right to its complexity *in vivo*.

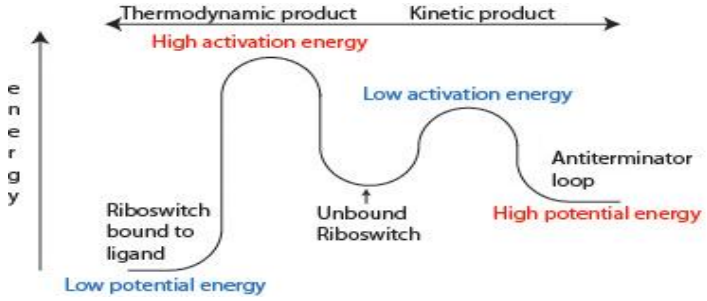

Figure 10 Thermodynamic and kinetic product of the lysine riboswitch in the presence of its ligand. When the thermodynamic product is more abundant there is thermodynamic control and vice versa.

Glycine Riboswitch

The glycine riboswitch, that amongst others regulates the glycin cleavage-related operon *gcvT* in *Bacillus subtilis*, is unique in that it has two ligand-binding aptameric sequences. Both sequences can bind a glycine molecule, causing the riboswitch to function as a "two-state genetic switch"(Mandal *et al*., 2004). Butler *and co-workers* (2011) propose that binding of the first glycine brings about a conformation change that exposes the second aptameric sequence, bringing it in a state of increased affinity for a second molecule of glycine to bind.

Unlike most riboswitches, the glycine riboswitch is believed to activate translation when both ligands have bound. When the *gcvT* operon is translated, the cleavage system will be established, causing glycine to be degraded. Since glycine is needed for many processes such as the citric acid cycle and protein synthesis, a two-state switch makes sense. It could give glycine the time to be used for its many purposes before being degraded.

Purine Riboswitch

For purine biosynthesis control, two riboswitches have been described: the so called G(guanine)-riboswitch and A(adenosine)-riboswitch. These riboswitches regulate purine concentrations in prokaryotic cells in different manners (Lescoute and Westhof, 2005);(Ling *et al*., 2009)

When a purine binds to the aptameric sequence of the G-riboswitch, it causes Mg^{+} to bind the riboswitch, which induces a conformation change resulting in a terminator helix, preventing

transcription of guanine biosynthesis genes downstream of the riboswitch (Brenner *et al.,* 2010). When a purine binds to the A-riboswitch, an anti-terminator helix is formed that exposes the ribosome binding site of an adenosine deanimase transcript. This allows for translation of the deaminase and breakdown of adenosine (Reining *et al*., 2013)

The G- and A-riboswitches are respectively situated at the 5' UTR of bacterial mRNA of guanine or adenosine biosynthesis genes. They are structured as a three-way junction that binds purine and its derivates as well as pyrimidine derivates (Ling *et al.,* 2009). It was the first class of riboswitches that was successfully altered, i. e. engineered, to accept other ligands (Dixon *et al.,* 2010). This was an important step in utilizing riboswitches for e.g. health and food related biotechnology.

Cyclic di-guanosine monophosphate (di-GMP) Riboswitch

The cyclic di-GMP riboswitch is unique, in the sense that it probably does not regulate cyclic di-GMP levels in the cell. Cyclic di-GMP is a signal molecule involved in expression control, but the mechanism of action used to be unknown. Sudarsan *and co-workers* (2008) found cyclic di-GMP riboswitches that were associated with the regulation of o. a. pilus formation, flagellum biosynthesis and virulence expression. This suggested cyclic di-GMP regulates expression by being a ligand for these riboswitches. The seemingly broad function of cyclic di-GMP riboswitches in prokaryotes might make this a useful riboswitch for future antibiotics and controlled gene expression.(Fujita *et al.*, 2011)

Organisms like *Geobacter uraniumreducens* have been found to contain up to 30 cyclic di-GMP riboswitches, and the riboswitch has even been found in the lysis module of a prophage (Sudarsan *et al.,* 2008). One could speculate that a bacteriophage monitoring the amount of cyclic di-GMP in its host cell could induce its lytic cycle when conditions in the cell are sub-optimal, resulting in evolutionary advantage over bacteriophages lacking this riboswitch.

GlmS Glucosamine- 6- phosphate (*GlcN6P*) catalytic Riboswitch

GlmS enzymes, encoded by the *glmS* gene, are involved in the biosynthesis of GlcN6P, a precursor of amino-sugars. When there is a sufficient amount of GlcN6P in the cell, it will bind to the glmS riboswitch causing expression regulation in the form of biosynthesis reduction of GlcN6P (Collins *et al*., 2007).

The *glmS* riboswitch is situated in the 5'UTR of the *glmS* transcript. It is found in all sequenced gram-positive bacteria and has a unique mechanism of action. When binding its ligand, instead of undergoing a conformational change, it induces self cleavage of the *glmS* mRNA, leading to repression of *glmS*-related gene expression. The riboswitch is also activated by compounds that are structurally similar to GlcN6P, making it a possible target for future antibiotics, even more so since the riboswitch is also found in pathogenic bacteria like *Bacillus anthracis* and *Bacillus cereus* (Tinsley *et al.,* 2007). In fact, the riboswitch is so susceptible to other ligands partially activating it, that GlcN6P has been hypothesized to be, as paraphrased by Tinsley *and co-workers* (2007), "a coenzyme for general acid-base catalysis self cleavage rather than an allosteric activator".

Tetrahydrofolate Riboswitch

Folate (vitamin B_9) is the precursor of a number of important cofactors, one of which is tetrahydrofolate. A riboswitch, primarily found in Fermicutes, regulates biosynthesis and transport of folate in the cell,

using folate derivatives as ligands. This riboswitch, containing a pseudo-knot and a three-way junction, supposedly controls precursor levels by measuring the concentration of this precursor its derivatives, like folate (Trausch *et al.,* 2011). Like most other riboswitches it is present in the 5'UTR of mRNA encoding biosynthesis genes relevant to the riboswitch ligand.

Interestingly, the tetrahydrofolate riboswitch is unique in that it probably binds two ligands in the same structured domain, although tetrahydrofolate riboswitches that have just one ligand binding site have also been found **(Figure 11)**. Above we have seen that the lysine riboswitch also binds two ligands, but this is not in the same structured domain.

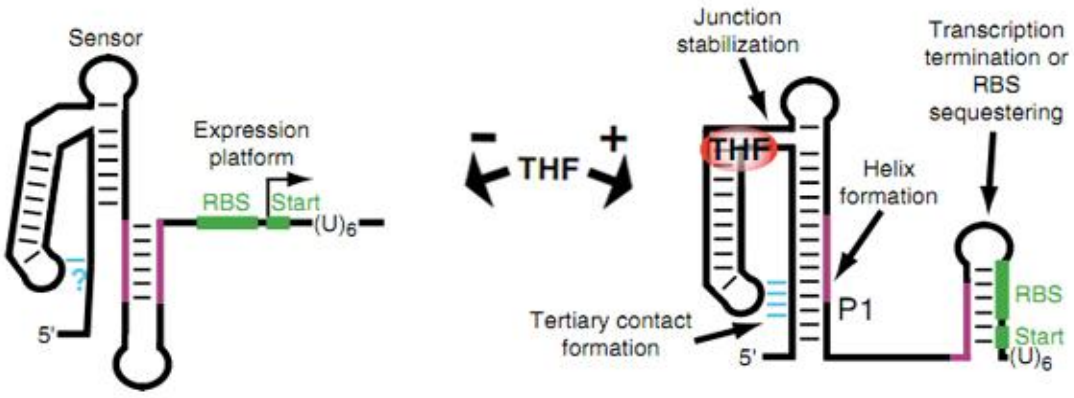

Figure 11 The proposed action mechanism of a one-ligand-binding tetrahydrofolate riboswitch (Huang *et al*., 2011).

Interpretation

It is evident that unravelling the action mechanisms of a great number of riboswitches has unravelled new possibilities and insights when it comes to regulation of transcription and translation. These mechanisms are points of departure for creating tools that can be used to stimulate and inhibit expression on the transcriptional as well as on the translational level. Alongside determining these mechanisms, some interesting discoveries have been made in the process, e.g. the mechanism of action by which cyclic di-GMP regulates expression. The discovery of riboswitches is also being considered to be evidence for the RNA world hypothesis, which will be discussed later.

From the idea to alter riboswitches, or even to create new ones from existing knowledge about RNA folding, emerged a sub-field that will be referred to here as riboswitch engineering. In the next section, the state-of-the-art and future perspective of this research field will be paraphrased.

Riboswitch Engineering

In 1998, Werstuck and Green already inserted an aptameric sequence into a luciferase reporter gene in mammalian cells. This sequence was transcribed in the 5' UTR of the mRNA, and by adding its aptamer expression could be inhibited by ten-fold because of the aptamer-ligand complex blocking the ribosome binding site (Bauer and Suess, 2006).

One could ask, if an aptameric sequence that binds a molecule, and in that way forms a bulky obstacle for a ribosome to bind, could be really called a riboswitch. While this is of course the basic idea of a riboswitch, an action mechanism that can be reversed and that corresponds to applied aptamer dosage in a way that is predictable by the K_d, would be of much greater use in biotechnology. In 2003, Hanson *and co-workers* designed such a reversible and dose dependent riboswitch for tetracycline, which has been successfully used to reduce translation in yeast.

However, the mechanisms described above cannot be used in prokaryotes, since the distance between the ribosome binding site and the start codon in bacteria cannot exceed 13 nucleotides. Therefore, insertion of an aptameric sequence at this position in these organisms would completely disable translation(Bauer and Suess, 2006).

For riboswitches to work in bacteria, a better strategy would be to introduce a terminator or anti-terminator-loop forming RNA sequence in the 5'UTR of a gene; similar to the situation in e.g. the B_{12}-box. Suess *and co-workers* (2004) successfully tested such an engineered device, in which an eight-fold expression increase was induced by adding of the aptamer ligand (Bauer and Suess, 2006). Another type of engineered riboswitches works with an anti-sense based mechanism; part of the 5'UTR is complementary to the ribosome binding site and blocks it by binding to it, either in the absence or the presence of a ligand. Thus this riboswitch can be used to either induce or inhibit gene expression **(Figure 12)**(Bauer and Suess, 2006).

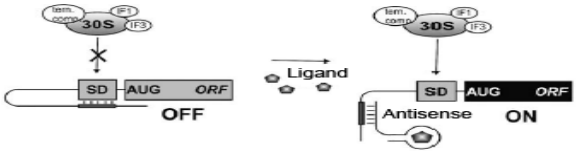

Figure 12 Example of an antisense mechanism that allows translation when the ligand binds its aptameric sequence (Bauer and Suess, 2006)

Engineered riboswitches do in general not disturb the natural cellular environment, which means they are non-toxic. They are easy to introduce into the genome of a cell and have already be used to control gene expression in prokaryotes and eukaryotes. The latter could lead one to believe that riboswitch engineering is feasible and its future seems very promising.

Riboswitches in medicine

Riboswitches are present in a wide range of organisms, including gram-positive and gram-negative bacteria, of which some are pathogens. Because of the rapid emergence of antibiotics resistance, the decreasing number of antibiotics that are still effective there is an increasing need for new antibiotics targets. Because a lot of riboswitch-regulated genes in prokaryotes are absent in eukaryotes (e.g. vitamin B12 biosynthesis genes), riboswitches could be a promising new target for antibiotics (Blount and Breaker, 2006).

It has in fact been found that the natural antibiotic roseoflavin binds to the FMN riboswitch, and in this way inhibits the expression of the FMN biosynthesis genes in the same way that its natural ligand does (Weigand *et al.,* 2009). For a number of riboswitches unique to prokaryotes like the riboswitch

regulating the production of the essential amino acid lysine, a ligand has been found that inhibits biosynthesis of this amino acid. Unless it is toxic for other reasons, such a riboswitch inhibiting compound would not be toxic to patients. However, as was the case with roseoflavin, resistance emerges quickly in the form of mutations in the riboswitch sequence (Weigand and Suess, *2009*). On the other hand, expression of secondary metabolites that have a function in antibiotics resistance might also be regulated by riboswitches, which in their turn are potential new antibiotics targets. A new antibiotic inhibiting such a riboswitch could be administered to a patient together with the corresponding antibiotic, to which the pathogen causing the infection would normally be resistant.

Riboswitches as evidence for the RNA world hypothesis

The RNA world hypothesis is one of many proposed ideas about the emergence of life as we know it, i.e. built up of the three macromolecules of life; DNA, RNA and proteins. DNA only seems to have function in information storage and replication thereof, while proteins mainly fulfil catalytic and structural functions. RNA has an intermediary function, but can also performs DNA- and protein-like functions to a certain extent. RNAs are known to perform enzymatic, storage and regulatory functions to such extent that it has been hypothesized that life could well have emerged without DNA and proteins in the so-called RNA world (Copley *et al*., 2007).

Riboswitches have been considered for evidence for this hypothesis since they are a form of regulation that could have been applicable in an RNA world; they indicate that expression regulation on an RNA basis is possible. Some riboswitches employ particularly complex mechanisms of action, which are also widespread. Because of this observation, it is unlikely that these strikingly similar action mechanisms evolved independently: evolutionary relatedness seems very plausible (Breaker, 2012). Although it was never implied that these mechanisms have stayed unaltered throughout evolution and it is likely that new riboswitches have also developed over time, the principle of how riboswitches work is hypothesized to be a remnant from the RNA world.

Conclusion and Discussion

In this paper the mechanisms of action and possible future applications of several riboswitches have been reviewed. Although attempts at targeting and engineering riboswitches have been successful, they don't appear to be used to the full extent yet; for example in biotechnology.

It seems evident that knowledge of riboswitches, i.e. mapping their existence, diversity and action mechanisms, is useful in designing and optimizing protocols and growth conditions in biotechnology. Using riboswitches could be taken to the next level by riboswitch engineering. Since this tool has been used successfully in fundamental research, it could be used today or in the near future in e.g. the industrialized production of enzymes by genetically modified microorganisms; it could be used to keep production of secondary metabolites high, by purposely disabling the riboswitch giving negative feedback to this production. Also, an upregulating riboswitch could be built into the 5'UTR of a desired proteins' mRNA, supplying an additional layer of expression stimulation, together with a conventional inducing method. For example: in *Lactococcus lactis*, which is an expression host for a wide range of proteins, the nisin-controlled gene expression system (NICE) could be used to initiate expression of the

target gene on the transcription level (de Ruyter *et al*, 1996). A riboswitch on the resulting mRNA could then be used to enhance transcription or translation of the target gene.

Additionally, the expression of other secondary metabolites that lower the production of the intended metabolite could be inhibited by targeting corresponding riboswitches. While concrete examples of riboswitch appliances so far still lack, Klauser and Hartig (2013) underline their potential.

In medicine, riboswitches could be used as antibiotic targets, but they also might have therapeutic uses. While Mulbacher *and co-workers* (2010) describe development of ribozyme-based therapy for e.g. cancer and prion disease treatment, one could hypothesise similar employment of riboswitches. Diseases that are caused by protein or prion plaque accumulation could be stopped from progressing if an inhibiting riboswitch was found that down regulates gene expression of the plaque forming proteins. One could also speculate that down regulated apoptosis genes in cancer cells could have their expression restored by a potential upregulating riboswitch.

In the case of antibiotics that target riboswitches, new resistances will appear fast and in fact already have (Weigand and Suess, *2009*). Also, the possibility to use riboswitch-based medicine would rely on the degree of riboswitch presence in genes of interest and would not be generalizable.

As for the RNA world hypothesis; riboswitch regulation seems to be just another mechanism that does not make the RNA world hypothesis impossible, but it is also not the breakthrough evidence for the existence of an RNA world. However, it has to be noted that accumulation of such plausibilities strengthens the case of an RNA world. Additionally, models have been proposed in which the emergence of RNA molecules from smaller molecules is described (Copley *et al*., 2007). These models further build up the appeal for an RNA world preceding our DNA world.

Taking all of the above into consideration the main research question will now be answered: the possibilities of riboswitches appear to be of importance to the following subjects in descending order; biotechnology, medicine and lastly evolutionary history. Since microorganisms used in biotechnology are usually substantially engineered and controlled in an industrial setting, riboswitches can be inserted, deleted and altered on the basis of rational thinking and molecular tools. In medicine, only riboswitches that are already present and known can be targeted. Mapping riboswitches related to a disease of interest would at the least require time and money, while success is not guaranteed. Finally, in regard to evolutionary history, riboswitches do to some extent add to the accumulating knowledge from which one day a concrete theory might be distilled, but up to now they seem to add nothing definitive.

References

Literature

Alberts B., Johnson A., Lewis J., Raff M., Roberts R. and Walter P. (2008) "Moleculair biology of the cell, fifth edition". New York: Garland Science.

Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D. A. and Horvath P. (2007) CRISPR provides acquired restiance against viruses in prokaryotes. *Science*, Vol. 315:1709-1712.

Bauer G. and Suess B. (2006). Engineered riboswitches as novel tools in molecular biology. *Journal of Biotechnology*, Vol. 124, 4-11.

Blount F. and Breaker R. R. (2006). Riboswitches as antibacterial drug targets. *Nature Biotechnology*, 24:12, 1558-1564.

Breaker R. R. (2012). Riboswitches and the RNA world. *Cold Spring Harbor Perspectives in Biology*, 4:a003566.

Brenner M. D., Scanlan M. S., Nahas M. K., Ha T. and Silverman S. K. (2010). Multivector fluorescence analysis of the *xpt* guanine riboswitch aptamer domain and the conformational role of guanine. *Biochemistry,* 49(8), 1596-1605.

Butler E. B., Xiong Y. and Strobel S. A. (2011). Structural Basis of Cooperative Ligand Binding by the Glycine Riboswitch. *Chemistry & Biology*, 18, 293-298.

Cao Y., Wu J. and Liu, Q. (2010). sRNA TarBase: A comprehensive database of bacterial sRNA targets verified by experiments. *RNA,* 2010 16: 2051-2057.

Collins J. A., Irnov I., Baker S. and Winkler W. C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. *Genes and Development,* 21: 3356-3368.

Copley S. D., Smith E. S. and Morowitz H. J. (2007) The origin of the RNA world: Co-evolution of genes and metabolism. *Bioorganic Chemistry*, 35:6, 430-443.

Corbino K. A., Barrick J. E., Lim J., Welz R., Tucker B. J., Puskarz I., Mandal M., Rudnick N. D. and Breaker R. R. (2005). Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biology*, 6:R70.

De Ruyter, P. G. G. A., Kuipers, O. P and de Vos, W. M. (1996). Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Applied and Environmental Microbiology,* 62, 3662-3667.

Dixon N., Duncan J. N., Geerlings T., Dunstan M. S., McCarthy J. E. G., Leys D. and Micklefield J. (2010). Reengineering orthogonally selective riboswitches. *PNAS*, 107:7, 2830-2835.

Edwards, A. L. and Batey, R. T. (2010). Riboswitches: A common RNA regulatory Element. *Nature Education,* 3(9):9.

Franklund C. V. and Kadner R. J. (1997). Multiple transcribed elements control expression of the *Escherichia coli btuB* Gene. *Journal of Bacteriology,* 179:12.

Fujita, Y., Tanaka, T., Furuta, H. and Yoshiya, I. (2012) Functional roles of tetraloop/receptor interacting module in a cyclic di-GMP riboswitch. *Journal of Bioscience and Bioengineering*. 113:2, 141-145.

Garst A. D., Porter E. B. and Batey R. T. (2012). Insights into the regulatory landscape of the lysine riboswitch. *Journal of Molecular Biology*. 423, 17-33.

Gasiunas G., Sinkunas T., Siksnys V. (2014) Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences*. Vol. 71:449-465.

Gelfand M., Mirinov A. A., Jomantas J., Kozlov Y. I. and Perumov D. A. (1999) A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends in Genetics*, Vol. 15, 439-442.

Grundy F. J. and Henkin T. M. (1998). The S box regulon: a new global transcription termination control system for methionine and cysteine biosynthesis genes in Gram-positive bacteria. *Molecular Microbiology*, 30(4), 737-749.

Henkin T. M. and Yanofsky C. (2002) Regulation by transcription attenuation in bacteria: how RNA provides instructions for transcription termination/antitermination decisions. *BioEssays* 24:8,700-707.

Huang L. H., Ishibe-Murakami S., Patel D. J. and Serganov A. (2011). Long-range pseudoknot interactions dictate the regulatory response in the tetrahydrofolate riboswitch. *PNAS,* 108:36, 14801-14806.

Ionescu D., Voss B., Oren A., Hess W. R. and Muro-Pastor A. M. (2010). Heterocyst-specific transcription of NsiR1, a non-coding RNA encoded in a tandem array of direct repeats in cyanobacteria. *Journal of Molecular Biology*, 398:2, 177-188.

Klauser B. and Hartig J. S. (2013). An engineered small RNA-mediated genetic switch based on a ribozyme expression platform. *Nucleic Acids Research*, 41: 10.

Lescoute A. and Westhof E. (2005). Riboswitch structures: purine ligands replace tertiary contacts. *Chemistry & Biology*, Vol 12. 10-13.

Lenz D. H., Mok K. C., Lilley B. N., Kulkarni R. V., Wingreen N. S. and Bassler B. L. (2004). The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio* harveyi and *Vibrio* cholerae. *Cell Press*, 118:1, 69-82.

Ling B., Wang Z., Zhang R., Meng X., Liu Y., Zhang C. and Liu C. (2009) Theoretical studies on the interaction of modified pyrimidines and purines with purine riboswitch. *Journal of Molecular Graphics and Modelling*, Vol 28, 37-45.

Lu C., Smith A. M., Fuchs R. T., Ding F., Rajashankar K., Henkin T. M. and Ke A. Crystal structures of the SAM-III/S_{MK} riboswitch reveal the SAM-dependant translation inhibition mechanism. *Nature Structural & Molecular Biology,* 15(10): 1076-1083

Madigan M., Martinko J., Stahl D. and Clark D. (2012). "Brock Biology of Microorganisms, thirteenth edition ". San Fransisco: Pearson Education.

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

Meyer M. M., Ames T. D., Smith D. P., Weinberg Z., Schwalbach M .S., Giovannoni S. J. and Breaker R. R. (2009). Identification of candidate structured RNAs in the marine organism 'Candidatus pelagibacter ubique'. *BMC Genomics,* 10:268.

Miranda-Ríos J., Navarro M., Soberón M. (2001). A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *PNAS*, 98:17 9736-9741.

Mulhbacher J, St-Pierre P and Lafontaine D. A. (2010) Therapeutic applications of ribozymes and riboswitches. *Current opinion in Pharmacology, 10:5, 551-556.*

Narberhaus F., Waldminghaus T. and Chowdhury, S. (2006) RNA thermometers. *FEMS Microbiology Reviews*. Vol. 30, 3-16.

Poiata E., Meyer M. M., Ames T. D. and Breaker R.R. (2009) A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA*. 2009, 15: 2046-2056.

Reining A., Nozinovic S., Schepckow K., Buhr F., Fürtig B and Schwalbe H. (2013). Three-state mechanism couples ligand and temperature sensing in riboswitches. *Nature,* 499, 355-359.

Rodionov D. A., Vitreschak A. G., Mironov, A. A. and Gelfand M. S. (2003). Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? *Nucleic Acids Research*. 31:23, 6748-6757.

Serganov A., Huang L. and Patel D. (2009). Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature*, Vol 458, 233-237.

Storz G. (1999) An RNA thermometer. *Genes & Developmenti,* 13:633-636.

Storz G., Vogel J. and Wassarman K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Molecular Cell*, Vol. 43, 880-891.

Sudarsan N., Weinberg E. R., Moy R. H., Kim J. N., Link K. H. and Breaker R. R. (2008). Riboswitches in eubacteria sense the second messenger cyclic di-gmp. *Science*, 321: 411-414.

Tinsley R. A., Furchak J. R. W. and Walter N. G. (2007) Trans-acting *glmS* catalytic riboswitch: Locked and loaded. *RNA* 13: 468-477.

Trausch J. J., Ceres P., Reyes F. E. and Batey R. T. (2011) The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer. *Cell Press*, Structure 19, 1413-1423.

Verhounig A., Karcher D. and Bock R. (2010) Inducible gene expression from the plastid genome by a synthetic riboswitch. *PNAS*, 107:14, 6204-6209.

Vitreschak A. G., Rodionov D. A., Mironov, A. A. and Gelfand M. S. (2003). Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. *RNA society* 2003 9: 1084-1097.

Vogel J. and Wager E. G. H. (2007). Target identification of small noncoding RNAs in bacteria. *Current Opinion in Microbiology*, 10:3, 262-270.

Weigand J. E. and Suess B. (2009) Aptamers and riboswitches: perspectives in biotechnology. *Applied Microbiology and Biotechnology.* 85:229-236.

Winkler W. C., Cohen-Chalamish S. and Breaker R. (2002). An mRNA structure that controls gene expression by binding FMN. *PNAS*, 99:25, 15908-15913.

Images

Auger S., Danchin A. and Martin-Verstraete I. (2002) Global Expression Profile of Bacillus subtilis Grown in the Presence of Sulfate or Methionine. *Journal of Bacteriology*, 184:18, 5179-5186.

Darlin D. (2013) Encyclopedia of science: transfer RNA. Retrieved April 13, 2014 from www.daviddarling.info/encyclopedia/T/tRNA.html

Laurenberg M., Asahara H., Korostelev S., Trakhanov S. and Noller H. F. (2008) Structural basis for translation termination on the 70S ribosome. *Nature*, 454: 852-857.

Nourysolutions (2008) Inside the libraries (III) : Data Activities. Retrieved April 13, 2014 from salamanca.nourysolutions.com/road/

Horvath P. and Barrangou R. (2010). CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science*, 327 (5962): 167–170
NB: Image was adapted by James Atmos in 2010. Retrieved May 12, 2014 from http://en.wikipedia.org/wiki/File:Crispr.png