universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Analysis of RNA chaperone activity of truncated ribosomal protein S12 from *E.coli* and of eukaryal and archaeal S12 orthologues

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer.nat.)

Verfasser:	Martin Alexander López
Studienrichtung /Studienzweig (lt. Studienblatt):	Molekulare Biologie
Betreuerin / Betreuer:	Dr. Katharina Semrad

Wien, am 8.9.2010

Table of contents

Zusammenfassung

RNA Moleküle sind sowohl strukturell als auch funktionell sehr flexibel. Zum Erlangen ihrer nativen Konformation benötigen sie daher die Hilfe von RNA Chaperonen. RNA Chaperone sind Proteine die missgefaltete RNA Strukturen reparieren können bzw. die Missfaltung von RNA Molekülen verhindern.

Die meisten Chaperone besitzen unstrukturierte Domänen oder sind in manchen Fällen komplett entfaltet. Theorien besagen, dass erst diese unstrukturierten Domänen die RNA Chaperon Aktivität der Proteine ermöglicht. Einige der bekanntesten und best konserviertesten RNA Chaperone sind ribosomale Proteine. Die ribosomalen Proteine werden einerseits für den Zusammenbau der Ribosomen benötigt, helfen, unter anderem, beim dekodieren der mRNA und besitzen auch extraribosomale Funktionen. Fast alle ribosomalen Proteine besitzen unstrukturierte Regionen. Eines der extremsten Beispiele ist das ribosomale Protein S12, das eine lange N-terminale unstrukturierte Domäne besitzt die sich durch die kleine ribosomale Untereinheit schlängelt. Weiters ist bekannt dass *Escherichia coli* S12 (*Eco* S12) ein RNA Chaperon ist.

In dieser Diplomarbeit wird untersucht ob *Eco* S12 auch dann noch RNA Chaperonaktivität besitzt wenn seine lange unstrukturierte Domäne entfernt wird. Weiters wird getestet ob auch S12 Proteine aus anderen Organismen Chaperonaktivität besitzen. Um dies zu untersuchen mußten die verschiedenen S12 Gene kloniert werden, die Proteine aufgereinigt werden und die Chaperonaktivität wurde mittels Hammerhead Assay getestet. Die erzielten Resultate lassen erkennen, dass die unstrukturierte Domäne für die RNA Chaperonaktivität essentiell ist und dass auch die getesteten eukaryotischen Proteine RNA Chaperonaktivität besitzen, nicht aber ein aus einem thermophilen Archaea isoliertes S12 Protein.

Weiters wird die RNA Chaperonaktivität des ribosomalen Proteins L19 mit einer *in vivo* Methode untersucht. Dabei wird dieser Versuch erstmals mit Fluoreszenz markierten Proben durchgeführt.

Zudem wurde noch die RNA Chaperonaktivität von rekombinanten L1 Proteinen *in vitro* und von L1 Orthologen *in vivo* untersucht.

Abstract

RNA molecules are structurally and functionally very flexible. As a result, they require the assistance of RNA chaperones to reach their ultimate folding state. RNA chaperones are proteins which can resolve misfolded RNA structures or which prevent the formation of misfolded structures.

Many RNA chaperones possess intrinsically unstructured domains or are even completely unstructured. Some theories claim that these unstructured domains are required for the RNA chaperone activity. Many ribosomal proteins possess strong RNA chaperone activity. In addition, they are highly conserved. Ribosomal proteins are required for the ribosome assembly, assist, among others, the decoding of the mRNA and also possess extraribosomal functions. The majority of the ribosomal proteins have intrinsically unstructured domains. One of the most impressive examples is the ribosomal protein S12. It has a long N-terminal extension which penetrates through the small ribosomal subunit. Additionally, *Escherichia coli* S12 (*Eco* S12) has RNA chaperone activity.

The aim of this diploma thesis is to discover whether a truncated *Eco* S12 protein missing its unstructured tail still has RNA chaperone activity. Further, it is tested whether S12 orthologues from other organisms also show RNA chaperone activity. To investigate this, the S12 genes were cloned, the proteins were isolated, and the chaperone activity was tested with a hammerhead cleavage assay. The results indicated that the intrinsically unstructured domain is required for the RNA chaperone activity of S12 and that the eukaryotic S12 proteins show RNA chaperone activity too. In contrast, S12 from a thermophilic archaea does not show any RNA chaperone activity.

Further, the RNA chaperone activity of the ribosomal protein L19 was analysed *in vivo.* For the first time this assay is performed with fluorescence labelled samples.

Moreover, the RNA chaperone activity of recombinant L1 proteins was tested *in vitro* and from L1 orthologues *in vivo*.

1. Introduction

1.1. RNA folding:

RNA is usually a single stranded nucleotide chain. RNA is very flexible and therefore can be formed into a plethora of complex three dimensional structures. In contrast to DNA, RNA consists of ribose sugars and contains the bases adenine, guanine, cytosine and uracil. Moreover, some RNAs can catalyze biological reactions (for review see (Griffiths et al. 2005)).

RNA molecules are structurally and functionally very flexible. As a result, RNA can misfold easily and become kinetically trapped in an inactive conformation. Such an inactive conformation can be extremely stable and can therefore prevent the formation of the native structure. Misfolded intermediates are even able to aggregate irreversibly. This effect is known as the RNA folding problem (Herschlag 1995; Schroeder et al. 2004). In addition, RNAs with different sequences can form similar structures and vice versa RNAs with comparable sequences can form diverse structures. Moreover, several RNAs are so called riboswitches. These are RNAs which undergo a conformational change by the binding of small metabolites (Mandal et al. 2004).

How do RNA molecules become folded correctly? The RNA structure consists of the secondary structure and the tertiary structure. The secondary structure arises from regions which are close to each other. Mostly, the secondary structures consist of A-form double helices. These secondary structures then form the tertiary structure by various interactions. In case of larger RNAs, several domains may fold independently from each other. Moreover, for the formation of the tertiary structure many RNAs require metal ions (Schroeder et al. 2004).

RNA molecules are, like DNA molecules, polyanionic. This negative charge is another problem for the correct folding of especially larger RNA molecules. Therefore, monovalent or divalent cations are required to neutralize areas of high negative potential. As a result, RNAs can not be folded in the absence of cations (Shcherbakova et al. 2008). Even the hammerhead ribozyme, that consists of only about 150 nucleotides requires metal ions for folding (Hammann et al. 2002).

In general, most RNA based reactions seem to take place together with proteins. These RNA binding proteins recognize specific RNA molecules and form specific, tight complexes with their target RNA. However, the way the proteins bind to the RNA and the

way these proteins help the RNA stabilizing their ultimate folding state can vary remarkably (Schroeder et al. 2004).

RNAs can reach there final folding state on multiple pathways. The folding landscape of RNAs (and proteins) is visualized as a funnel. This funnel is uneven and is therefore populated with various intermediates which are located in local energy minima. Without any guide an unfolded RNA would take an enormous amount of time to become folded correctly, at least *in vitro*. Anyway, it seems like kinetic folding pathways for large RNA molecules are very complex and full of kinetic traps which need to be resolved by for example partial unwinding. Figure 1 is taken from Shcherbakova et al. 2008 and shows a folding funnel of the *Tetrahymena* ribozyme (Shcherbakova et al. 2008).

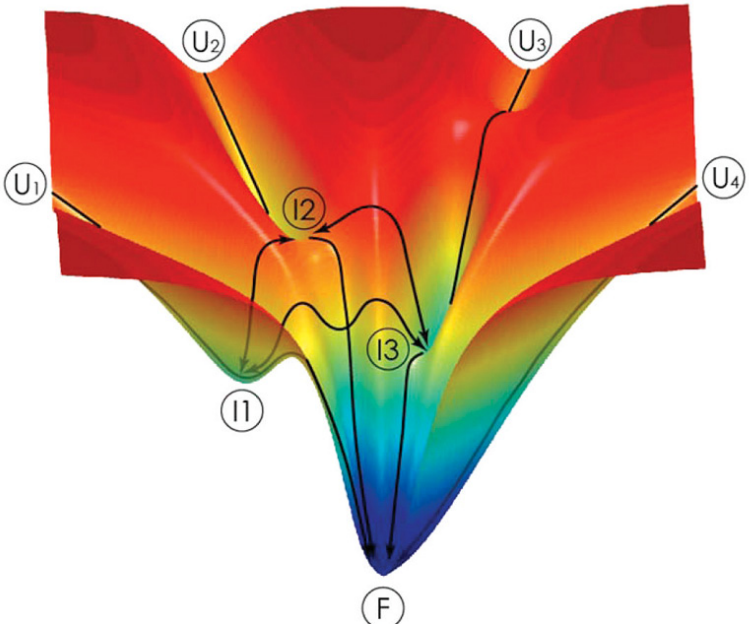


Figure 1: This figure shows a folding funnel of the *Tetrahymena* ribozyme with different intermediates and different starting points. It is taken from Shcherbakova et al. 2008 and illustrates how complex folding pathways can be.

Some of the most important guides for unfolded RNAs through the folding funnel are so called RNA chaperones. RNA chaperones are proteins which help the RNA to become folded correctly. These RNA chaperones can also rescue RNAs which become trapped in inactive conformations. So far all known RNA chaperones are multifunctional and can cause pleiotropic effect in case of a mutation (Schroeder et al. 2004).

Is the RNA folding problem only an *in vitro* problem? Of course, there are great differences between the RNA folding *in vivo* and *in vitro*. *In vivo*, for example, the RNA folding occurs co-transcriptionally. Nevertheless, at least one example for a folding inability is known *in vivo*, namely the pre-mRNA of the *thymidylate synthase* gene from the phage T4 which contains a group I intron. This intron can only be spliced efficiently during translation, indicating ribosomal chaperones are required for this process (Semrad et al. 1998; Schroeder et al. 2004). This pre-mRNA is used for the so far only

known assay to identify RNA chaperones *in vivo* (Clodi et al. 1999; Prenninger et al. 2006).

RNA chaperones are defined as proteins that aid in the process of RNA folding by preventing RNA misfolding and by resolving misfolded RNA structures. As a result, the RNA can perform its biological function. In contrast to the RNA binding proteins, RNA chaperones do not bind specifically the RNA. They are binding RNA folding intermediates in a nonspecific manner to assist folding. The specific RNA binding proteins do also help the RNA finding their native structure by stabilizing specific tertiary structures but they are not referred to as RNA chaperones. RNA chaperones have some similarities to protein chaperones, but in contrast to protein chaperones RNA chaperones do not require ATP consumption to resolve misfolded structures. The RNA chaperone activity is also pretty demanding as they have to assist the folding of many different unrelated RNAs in the intracellular milieu. RNA chaperones also facilitate the self splicing of group I introns by helping the intron to fold correctly. This was one of the first methods to identify RNA chaperones (Herschlag 1995; Schroeder et al. 2004; Tompa et al. 2004).

One classical example of a RNA chaperone is the A1 protein of heteronuclear ribonucleoprotein (hnRNP). This protein promotes the renaturation of complementary nucleic acid strands. It is one of the first discovered protein's, which is binding non-specifically to RNA and which has RNA chaperone activity. Moreover, it was demonstrated that especially its glycine rich C-terminal domain is required for the RNA chaperone activity (Herschlag et al. 1994). As this C-terminal domain is unstructured it has lead to the concept that unstructured, basic (positively charged) and nonspecific domains are required for the RNA chaperone activity of proteins or to accelerate these reactions (Tompa et al. 2004).

1.2. The role of intrinsically unstructured protein domains:

Until the early 90s of the 20th century the common concept was that a protein can only fulfill its function in its structured state. This theory was supported by various experiments and observations, like the first crystal structures of globular proteins and enzymes, or with experiments were denaturated proteins lost their function. Occasional counterexamples were largely ignored. However, newer discoveries demonstrated that intrinsically unstructured domains are common among proteins in living cells. Moreover, proteins with unstructured domains are essential for cellular functions and involved in

cell-signaling, regulation and control pathways (Tompa et al. 2004; Radivojac et al. 2007).

When is a domain defined as unstructured? Functional disordered segments can consist of only a few amino acids. They can also consist of rather large regions or protein ends. Some proteins, like α-casein are even considered to be completely disordered. There is no general definition of an unstructured domain but very short disordered regions are mostly ignored. The unstructured regions should be long enough to be simply characterized. A recent review defines the significant size of an unstructured region by containing more than 50 residues. However, for smaller proteins this is already pretty large (Dyson et al. 2005; Radivojac et al. 2007).

One of the most challenging aspects is identifying intrinsically unstructured domains. Several physicochemical methods have been used to identify intrinsically disordered proteins. Nowadays, many computer programs are available for the prediction of unstructured domains. These programs are based on a spectrum of different computational methods. For example, they analyse the different sequence biases between ordered und disordered regions. This method is based on various analyses which group different amino acid residues into an order promoting group or a disorder promoting group. Other programs rely on derived properties like a secondary structure prediction. In addition, some databases contain information about intrinsically disordered domains and their function (Dyson et al. 2005; Radivojac et al. 2007; Dunker et al. 2008).

Disorder prediction programs are already an important tool in science. These programs can be used for many purposes. For example such programs have been used to identify protein binding sites (Garner et al. 1999). Other groups used these programs to identify post translational modification sites, like phosphorylation, ubiquitination or methylation sites as such modification sites are often within disordered regions. Of course, these programs only help finding putative modification sites and do not substitute bench experiments. The identification of intrinsically disordered domains can also help crystallizing proteins, as assumed unstructured domains can be truncated to allow the crystallization of proteins (Radivojac et al. 2007).

Two different theories are present about the different structural forms in that intrinsically unstructured domains exist. One theory suggests a completely unstructured state and a molten globule state. The latter is a state in which a protein has already some secondary structures but a disordered tertiary structure. Together with the ordered

form these three states form the protein trinity hypothesis. Besides, the protein quartet hypothesis exists with an additional state the pre-molten globule. Both hypothesis define a protein as being in one of the three or four states although it seems clear that there are a number of different structural subtypes in each state (Dyson et al. 2005; Radivojac et al. 2007).

Intrinsically unstructured domains are conserved in all kingdoms of life, even in viruses. However, most of the longer conserved disordered regions are only present in eukaryotic and viral proteins indicating that long disordered regions are more prevalent in eukaryotes than in prokaryotes. Moreover, with the increasing complexity of an organism there is an increase of intrinsically unstructured proteins (Dyson et al. 2005; Dunker et al. 2008).

What are the functions of this disordered proteins or unstructured domains, respectively? Intrinsically unstructured domains are involved in regulation, cell-signaling and control pathways. In some experiments, databases have been used to identify disordered proteins and to categorize them into functional classes. These classes include molecular recognition, molecular assembly, protein modifications and entropic chain activities. In the first three cases the proteins undergo an unfolded to folded transition while in case of the entropic chain activity the function arises from the unfolded state. In most cases the unstructured areas are forming interaction with partner proteins and become folded. The recognition of partner proteins is one of the main functions of disordered structures (Radivojac et al. 2007). So what are the advantages of disordered regions in molecular recognition compared to folded domains?

Unfolded domains have a higher binding diversity as the disordered region and therefore can recognize various different binding partners. Or the other way round, many unfolded domains can recognize the same binding partner. In other words, many globular proteins can bind to one intrinsically unfolded binding partner, or numerous unstructured proteins bind to one structured partner. Accordingly, several hub proteins are entirely unstructured and other hub proteins are at least partially disordered. Hub proteins are defined as proteins which interact with more than 10 partners. However, some hub proteins like 14-3-3 are highly structured. Nevertheless, it was discovered that the binding partners of 14-3-3 are indeed proteins containing unstructured domains. Moreover, a single binding groove of 14-3-3 is associated with five different disordered sequences. Another example is β-catenin which is also interacting with several unstructured binding partners (Dyson et al. 2005; Patil et al. 2006; Dunker et al. 2008).

Other advantages of intrinsically disordered proteins are that they can bind their target with high specifity and low affinity. Additionally, they can form a larger interaction surface and therefore enhance the speed of an association. Moreover the dissociation can also be faster using an unzippering mechanism (Radivojac et al. 2007). Indeed without the intrinsically unstructured domains the proteins would have to be 2 to 3 times larger to form such extensive interfaces. Of course, this would lead to an increase of the cellular crowding or the cell would have to increase its size (Dyson et al. 2005).

Intrinsically unstructured proteins are common in many signaling pathways and in regulation. Therefore, they are also involved in diseases. For example cancer associated proteins show significantly more disorder than typical eukaryotic proteins. 79% of them contain predicted regions of disorder of 30 residues or even longer (Iakoucheva et al. 2002). Also the proteins of the high risk human papillomaviruses contain a significantly high amount of disorder. Moreover, many proteins associated with cardiovascular diseases are intrinsically unstructured. Some of them are completely disordered. Also notable is the high amount of intrinsically disordered regions among transcription factors. However, the eukaryotic transcription factors have far more intrinsically disordered regions compared to the prokaryotic ones. In addition, most of the alternative spliced RNAs are coding for intrinsically unstructured proteins (Romero et al. 2006; Radivojac et al. 2007; Dunker et al. 2008).

These are only a few known examples in which intrinsically unstructured proteins and protein domains are involved. Also many diseases are correlated with proteins which are disordered. Further, intrinsically disordered regions can also be sites of chromosomal translocations. Furthermore, another function of the intrinsically unstructured proteins is the binding to nucleic acids. Moreover, one of the most important functions is the chaperoning activity of these proteins to assist the folding of RNAs. These unfolded DNA binding and RNA folding proteins are the ones with the highest conservation between the three kingdoms of life. Furthermore, the frequency of disorder in RNA chaperones is much higher than in any other protein class. Especially many ribosomal proteins contain intrinsically unstructured extensions (Ramakrishnan et al. 2001; Tompa et al. 2004; Dunker et al. 2008).

The disordered domains are not only restricted to RNA chaperones, some protein chaperones contain unfolded regions as well. Protein chaperones are often devoid of long disordered regions but they contain many disordered scattered short segments. On the other hand, some protein chaperones are completely disordered. One example

would be the already mentioned α-casein. The chaperone activity of the RNA chaperones is often indispensable linked to the unstructured domains. The deletion of the disordered segment abolish the chaperone activity in most cases and sometimes reduces it (Tompa et al. 2004).

Why are the intrinsically unstructured domains so important for RNA and protein chaperones to fulfill their function? As already mentioned intrinsically unstructured domains have the advantage of a high flexibility and therefore bind to various unrelated partners rapidly. This ability comes in handy as chaperones need to recognize and bind many different misfolded substrates. The enhanced speed of the interactions can be useful to prevent the aggregation of misfolded intermediates. Furthermore, the disordered regions can uncouple specificity from binding strength which may be useful for the reversible interactions with the partner protein (Tompa et al. 2004).

Another contribution of intrinsically unstructured domains to the chaperone function may be their rapid but still transient binding to the substrate. The following conformational change of the chaperone could balance the thermodynamic costs of the substrate unfolding. Many repeated cycles of binding and release of the chaperone could help the substrate to be folded correctly. Of course, this is only one possible working model how some RNA chaperones might work. Other chaperones are known to stabilize their substrates. Another advantage of the intrinsically unstructured domains could be that while unfolding a substrate the different strands of the substrate are kept at a near distance. This again could speed up the folding of the RNA (Tompa et al. 2004).

An important group of RNA chaperones are ribosomal proteins. Almost half of the ribosomal proteins have globular bodies with long intrinsically unstructured extensions. The extensions can penetrate rather deeply into the ribosomal core (Timsit et al. 2009). Therefore, I will next present some information about the ribosome in general and about ribosomal proteins with RNA chaperone activity.

1.3. The ribosome and ribosomal proteins:

The ribosome, a universal dynamic ribonucleoprotein complex, catalyzes mRNA directed protein synthesis in every organism. It uses mRNA as a template and aminoacyl tRNAs as substrates. Ribosomes make up approximately 30% of the cell mass with up to 10^6 ribosomes in bacteria and mammalian cells. In a growing bacterial cell most ribosomes are active in translation and synthesize polypeptide chains at rates of about 20 amino acids per second (Kramer et al. 2009; Schmeing et al. 2009). The ribosome consists of

two different subunits. The prokaryotic subunits are the 50S subunit and the 30S subunit which assemble into a functional 70S particle. The ribosome consist of two third RNA and only one third proteins. In contrast, mitochondrial ribosomes are two thirds proteins and only one third RNA. As a result, mitochondrial ribosomes have larger proteins and consist of more organellar specific proteins (Ramakrishnan et al. 2001; Wilson et al. 2005).

The large subunit contains two different RNAs, the 23S rRNA which is about 2900 nucleotides long, and the 5S rRNA which is about 120 nucleotides long. Moreover, the large subunit is made up of 30 to 40 different proteins, depending on the species. The small subunit is composed of the ~1540 nucleotide long 16S rRNA and 20 proteins. Both subunits are working together. The small subunit is responsible for the decoding of the mRNA. Further, it provides the mRNA binding machinery, the path for the progression of the mRNA and most components that control the fidelity of the translation. The large subunit performs the peptide bond formation, the main task of the ribosome. Furthermore, it provides the protein exit tunnel. The three binding sites for the tRNAs, namely the A (aminoacyl), P (peptidyl) and the E (exit) sites, are localized on both subunits. The first tRNA binds to the mRNA at the P site while the A site remains empty. The tRNA carrying the next amino acid binds to the A site. Then, the peptide bond is formed and the tRNA which was at the A site before translocates to the P site, while the tRNA which was in the P site moves to the E site. The catalytic site of this peptidyl transferase reaction is located in the 23S rRNA, composed entirely of RNA and therefore performed by the rRNA solely. This region of the 23S rRNA is highly conserved (Ramakrishnan et al. 2001; Bashan et al. 2008; Schmeing et al. 2009).

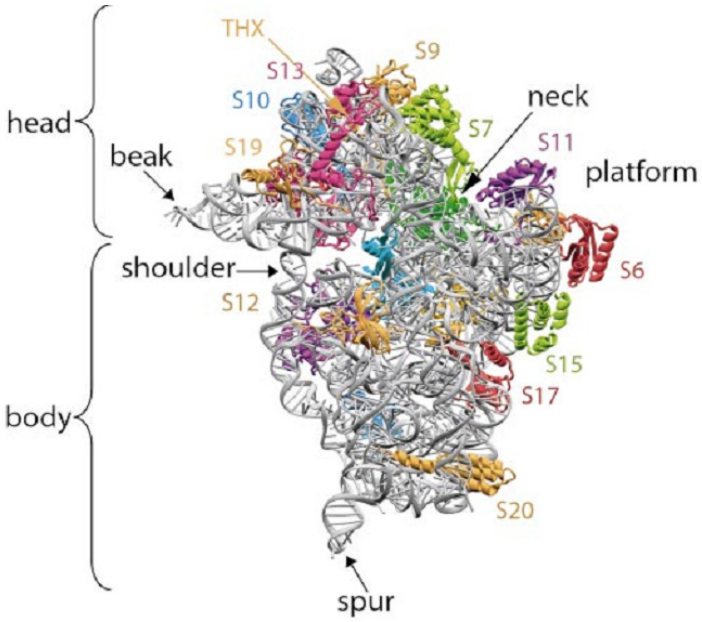


Figure 2: The 30S subunit seen from the interface side of the subunit. The 16S rRNA is shown in a grey ribbon stick presentation. The figure is taken from Brodersen et al. 2002.

The 23S rRNA is separated into six domains. However, the

domains are intricately interwoven in the 50S subunit. Each protein of the 50S subunit contacts in average 2.6 rRNA domains. The ribosomal protein L22 even interacts with all six domains demonstrating the complexity of the large subunit (Wilson et al. 2005).

The 16S rRNA of the small subunit is divided into four main domains. These domains are termed the 5' domain, the central domain, the 3' major domain and the 3' minor domain. In contrast to the 23S rRNA, each of these domains forms a distinct morphological component. Therefore, the shape of the small subunit is largely determined by the 16S rRNA. The 30S subunit is generally divided into an upper third called the head and two lower thirds know as the body. Figure 2 (taken from Brodersen et al. 2002) shows an overview of the 30S subunit in the canonical front view. The front is the side that faces and interacts with the 50S subunit while the side facing the cytoplasm is called the back (Wimberly et al. 2000; Brodersen et al. 2002).

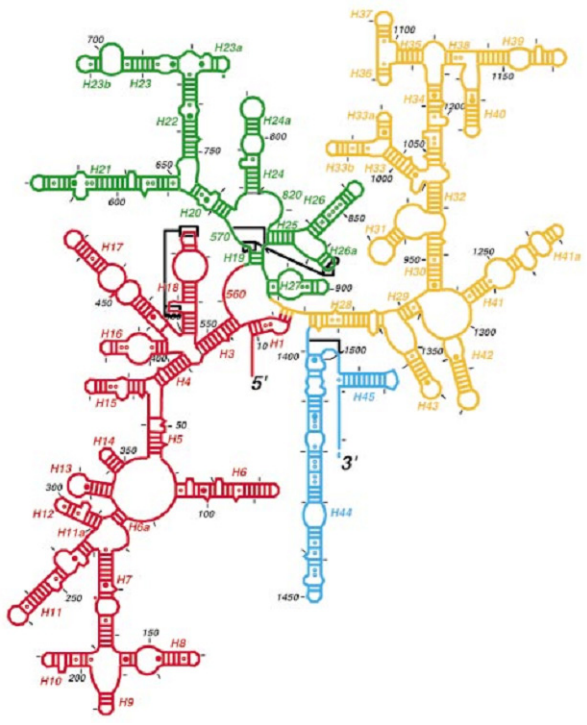


Figure 3: Secondary structure diagram of the 16S rRNA. The 5' domain is coloured in red, the central domain in green, the 3' major domain in orange and the 3' minor domain in cyan. The diagram is taken from Wimberly et al. 2000.

The four domains of the 16S rRNA are building the major morphological features of the small subunit. As a result, the 5' major domain forms the bulk of the body, the central domain most of the platform and the 3' major domain makes up most of the head. The functionally important 3' minor domain is located at the subunit interface and is therefore a part of the body. This domain consists of only two helices. The 3' minor domain is considered to be very flexible and to be involved in complex processes that implicate conformational changes of the ribosome. Figure 3 shows the secondary structure diagram of the 16S rRNA (taken from Wimberly et al. 2000). The proteins of the 30S subunit are asymmetrically distributed. The majority of the proteins is located on the top sides and the back of the small subunit. The interface area between the

two subunits is almost devoid of proteins. The only exception is the ribosomal protein S12 which is located near the coding site of the 30S subunit (Wimberly et al. 2000; Brodersen et al. 2002).

The 16S rRNA structure consists of over 50 regular double helices. These helices are connected by irregular single stranded loops. Still, most of the 16S rRNA is helical or at least approximately helical. Therefore, the structure can be considered as a 3D formation of helical elements. The double stranded helical segments have been labeled helix 1 through helix 45. The packing of these elements is responsible for the overall fold of the four domains of the 16S rRNA. Three different types of helix-helix packing in the structure are known. They all use the wide and shallow minor groove for their interactions (Wimberly et al. 2000).

Although the 16S rRNA is responsible for the shape of the small subunit, the ribosomal proteins are required for the correct folding and for the assembly of the small subunit. The assembly of the ribosome involves the coordinated transcription, maturation and folding of the rRNA and of course the binding of the ribosomal proteins. It is assumed that the assembly of the 30S subunit occurs by an alternate series of changes in the RNA conformation and protein binding events. In this theory the folding of an rRNA area generates a new binding site for a protein which then assists the next RNA folding event. In the assembly map of the 30S subunit the proteins are divided into three groups: primary binding proteins, secondary binding proteins and tertiary binding proteins. The primary binding proteins are able to bind to the 16S rRNA directly. The secondary binding proteins already need the help of the primary binding proteins to bind to the rRNA and the tertiary binding proteins need even the binding of at least one secondary binding protein before they can bind to the ribosome. Even though the secondary structure of the 16S rRNA is stable, the formation of the tertiary structure requires the assembly of proteins (Brodersen et al. 2002; Williamson 2005).

The primary binding proteins are typically globular proteins. They assist the folding and fix the conformation of RNA helices in the local area. In addition, the primary binding proteins can also tie together different parts of the 30 subunit and therefore make inter-domain contacts. However, these inter-domain contacts can not lead to the final assembly as proteins containing long extensions still have to bind to the ribosome. These long extensions can often span long distances. Thus, it seems unlikely that the domains are fully packed against each other after the binding of the primary binding proteins. Of course, also protein-protein interactions are important for the assembly of

the 30S subunit. Especially the binding of some tertiary binding proteins require an interaction between ribosomal proteins (Brodersen et al. 2002).

What are the characteristics of the ribosomal proteins? The ribosomal proteins are rather small proteins with an average size of 25 to 300 amino acids. This is especially tiny compared to the large sizes of the rRNAs. The ribosomal proteins are ordered roughly by decreasing size. More precisely they are numbered according to their position on a 2D gel and therefore small basic proteins have the highest numbers. Further, the ribosomal proteins are RNA binding proteins. To this end, most of the ribosomal proteins are very basic with an isoelectric point of over 10. Many ribosomal proteins consist of at least one globular domain which is located on the ribosomal surface and long extended regions which penetrate into the interior of the ribosome. They are required to stabilize the tertiary structure of the rRNA. However, the stabilization is not the sole function of the ribosomal proteins. Some of them also perform functions during translation, for example tRNA translocation (Brodersen et al. 2002; Wilson et al. 2005; Bashan et al. 2008). However, it is almost impossible to attribute a specific function to a single ribosomal protein as there is a great cooperation between the ribosomal proteins in the ribosome and also between the proteins and the rRNA in the ribosome (Wilson et al. 2005).

The ribosomal proteins also allow the binding of other proteins, especially protein chaperones which help folding the nascent polypeptide during ongoing synthesis. Examples are the bacterial trigger factor or the signal recognition particle (SRP) (Bashan et al. 2008; Kramer et al. 2009). Moreover, in higher eukaryotes these protein chaperones are required for the ribosomal assembly (Karbstein 2010).

The ribosomal proteins are highly conserved in the three kingdoms of life. Around 30% of the *Escherichia coli* proteins have orthologue counterparts in higher eukaryotes and archaea. The archaeal ribosomes have additional 30% orthologues with the higher eukaryotes (Wilson et al. 2005).

Many proteins of the 30S subunit contain long extensions. These extensions are normally not defined in the structures of the isolated proteins as they are disordered in the absence of the rRNA. The amino acid composition of the extensions is distinctive and typically consists of various glycine residues to keep the extension flexible and to allow tight packing. Furthermore, many basic arginin and lysine residues are located in the extensions to ease the interaction with the rRNA. The extensions reach far into the RNA and allow the proteins to contact several RNA areas. In general the extensions are

narrow and therefore allow a close approach of different RNA segments. Especially the smaller proteins seem to be completely surrounded by RNA and have a big fraction of their surface buried in the RNA. The globular domains of the proteins of the 30S subunit are β-barrels and α helices packed against β-sheets. The globular domains are only found on the surface of the small subunit. The most impressive example for a protein with a globular domain and a long extension is S12 which has its globular domain on the interface side and an extension which penetrates through the ribosome to interact at the back side with S8 and S17 (Wimberly et al. 2000; Brodersen et al. 2002; Wilson et al. 2005).

Nevertheless, compared to other RNA binding proteins the ribosomal proteins make far fewer base-specific contacts with the RNA. As an alternative, they interact through salt bridges between their own positively charged residues and the negatively charged backbone of the RNA (Brodersen et al. 2002).

The proteins of the small subunit are often binding to junctions between helices. These bindings seem to be important for the assembly of the different domains of the 30S subunit and therefore for the correct tertiary structure of the 16S rRNA. However, for the assembly of the 30S subunit none of the primary binding proteins has a long extended region. In contrast, the primary binding proteins which are essential for the assembly are more globular proteins (Wimberly et al. 2000; Timsit et al. 2009).

The ribosomal proteins are unevenly scattered over the ribosome. In case of the small subunit almost all the proteins are located on the back and periphery of the subunit. Most proteins bind to the head domain of the 30S subunit. In contrast, the interface side is almost devoid of proteins. Only the protein S12 is located at this site and to a lesser extent S13 and S19. Also the functionally important 3' minor domain, containing the anti-Shine-Dalgarno sequence and the coding center of the A-site, has very few contacts to any protein. Again, S12 is responsible for one of the few contacts there. The distribution of the proteins is demonstrating that the ribosome is primarily an RNA machine (Brodersen et al. 2002; Wilson et al. 2005).

Some of the ribosomal proteins also have extraribosomal functions. How is an extraribosomal function defined? A ribosomal protein has to interact with a non ribosomal protein or RNA and this interaction has to have a physiological effect on the cell in the absence of the ribosome. One of the best known extraribosomal functions of ribosomal proteins is the feedback regulation of their own synthesis at the translational level. Ribosomal proteins can prevent their own translation by binding to the operon. As

a result, the translation of the whole polycistronic mRNA is stopped. For example the S10 operon encodes for 11 different ribosomal proteins. The regulation of this mechanism is rather easy. A newly translated protein binds to a newly transcribed rRNA. However, if there is no rRNA the protein binds to its own mRNA and represses the translation of more proteins to balance the production of rRNA and proteins. In eukaryotes this system is more complicated due to the lack of polycistronic mRNA (Wool 1996; Wilson et al. 2005; Warner et al. 2009).

The knock out of some genes coding for ribosomal proteins can lead to apoptosis or cell cycle arrest. L5, L11 and L23 for example can all bind to MDM2 and therefore regulate the level of p53. The overexpression of these proteins leads to an accumulation of p53 and to cell-cycle arrest or apoptosis. Moreover, MDM2 can degrade L26 by ubiquitination. L26 again can stimulate the translation of p53 by binding to its mRNA. In addition, haploinsufficiency of various ribosomal proteins can lead to tumors, anemia and even cancer in higher eukaryotes. Also in plants haploinsufficiency can lead to several defects. These are only some examples of extraribosomal function of proteins. However, also the reduced level of ribosomes due to the haploinsufficiency of ribosomal proteins is leading to the severe phenotypes. Still some ribosomal proteins are involved in controlling the cell cycle and a lower level is leading to these phenotypes as well (Warner et al. 2009).

Some ribosomal proteins are known to have RNA chaperone activity. In the following I will present the proteins I am working with in more detail.

1.4. Ribosomal proteins S12 and L1:

The protein I was working with is primarily S12. In two side projects I was also working with ribosomal protein L1 and L19 (The nomenclature is applied to *E.coli*.). These three ribosomal proteins have in common that they all are strong RNA chaperones. In fact, L19 and L1 seem to be the strongest RNA chaperones of the 50S subunit. In addition, S12 was one of the first identified RNA chaperones (Coetzee et al. 1994; Semrad et al. 1998).

The ribosomal protein S12 is the only protein of the small subunit which is located at the interface side of the 30S subunit. Additionally, S12 is one of the few proteins which interacts with the 3’ minor domain of the 16S rRNA. S12 interacts near the top of helix 44 with the rRNA. Of course, S12 also interacts with the 5’ domain of the 16S rRNA. However, not only the location of S12 is unusual also its structure is quite unique

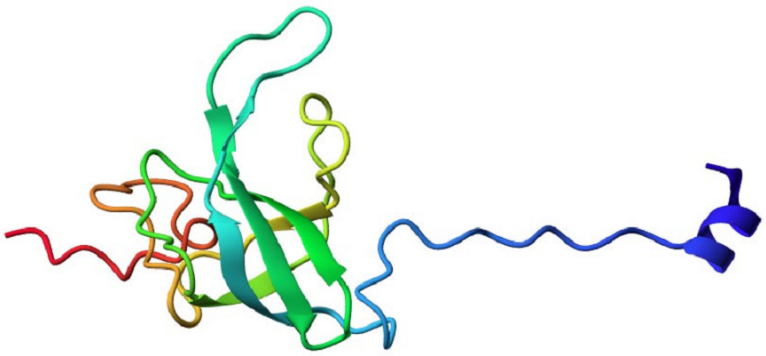


Figure 4: Crystal structure of the ribosomal protein S12. The N-terminal end is coloured in blue and the C-terminal end in red. This structure shows S12 from *Thermus thermophilus.* In *Eco* the short unstructured C-terminal end is absent. The picture is taken from Brodersen et al. 2002.

as one can see in figure 4 (taken from Brodersen et al. 2002). S12 consists of a core with a five-stranded β barrel formation which binds H44 and various helices of the 5' domain together. Furthermore, S12 contains an especially long extension that connects the N-terminal end with the core unit. As described earlier, this extension penetrates through the ribosome to the back of the ribosome, 50 Å away from the β-barrel. The N-terminal end consists of a two turn helix which is contacting the proteins S8 and S17. The interaction with these two proteins seems to hold the shoulder in the 5' domain and the platform of the central domain together (Brodersen et al. 2002; Wilson et al. 2005).

S12 together with S4 and S5 is responsible for the decoding and fidelity of the translation. The accuracy of ribosomal translation is not maximal and is depending on the growth rate of the cell. This is known because S12 mutants can increase the accuracy. Moreover, mutations in the protein S12, S4 and S5 have the strongest effects on fidelity. The amino acids which were usually mutated in S12 are lysine 42 and proline 90. Multiple S12 mutations at these two positions are known, the classical one however is the one isolated out of streptomycin resistant cells (K42N). The aminoglycoside antibiotic streptomycin increases the translational error rate. In the streptomycin resistant mutants the translational accuracy is increased and the elongation rate is decreased, suggesting that S12 increases the translational speed at the costs of fidelity. In contrast, mutations in S4 and S5 are leading to a strongly reduced accuracy of translation. This phenotype is known as the ribosomal ambiguity mutant (*ram*). Still, the *ram* phenotype is viable and is the revertant mutation to the hyperaccurate S12 mutation (Wilson et al. 2005).

What are the characteristics of the *ram* mutation? It is thought that the *ram* conformation stabilizes a closed A-site conformation characterized by a high affinity for aminoacyl-tRNAs. The binding of cognate but false tRNAs is stabilized and therefore the

error rate is increased. Streptomycin mimics these effects by binding to the A site and stabilizing the *ram* conformation. However, mutations in the S12 gene can reverse the effects of streptomycin and can consolidate the open A-site conformation. The affinity for near cognate tRNAs is reduced and the fidelity of the decoding is strongly increased (Holberger et al. 2009). The transition to the closed A-site conformation requires the disruption of various interactions between S4 and S5 and the formation of salt bridge interactions between S12 and the 16S rRNA (H44). Therefore, S12 mutants may destabilize the interaction between the protein and the 16S rRNA and thus lead to an open conformation. In contrast, S4 and S5 mutants can still disrupt their interactions and therefore consolidate the closed A-site conformation (Wilson et al. 2005).

Mutations in the ribosomal protein L19 from the large subunit can also compensate the effects of S12 mutations. (Maisnier-Patin et al. 2007) This finding is indicating that there is a functional interaction between these two proteins although they are located on the different subunits. L19 is located at the subunit interface and is interacting with the rRNAs of both subunits. It even interacts with H44 of the 16S rRNA and forms a bridge to connect the two subunits (Harms et al. 2001). Besides, L19 has the strongest RNA chaperone activity of all the ribosomal proteins of the 50S subunit (Semrad et al. 2004). The amino acids at position 40 and 104 are highly conserved in L19 orthologues. A mutation of L19 at position 40 increases the fitness of cells with an S12 mutation and enhances their elongation rate. However, the fitness of wildtype cells is decreased. L19 mutations in wildtype cells show similar effects as *ram* mutants. To this end, L19 is required in the decoding step of protein synthesis. The L19 mutation phenotype can be rescued by mutations in L14, L19 itself and mutations in S12. Strange enough one of the compensating S12 mutations is within its long N-terminal extension (Maisnier-Patin et al. 2007).

Mutations in S12 are not only leading to hyper-accuracy and slower growth rates, they can also decrease the level of mRNA cleavage at the A-site and subsequently the tmRNA rescue. The cleavage of mRNA and the following recruitment of tmRNA are necessary to rescue stalled ribosomes. Nevertheless, there is no correlation known between the higher fidelity and the reduced ability to rescue stalled ribosomes (Holberger et al. 2009).

S12 is one of the first identified RNA chaperones. This protein has a small binding preference to unstructured RNAs (Coetzee et al. 1994). The reason for this seems to be that structured RNAs are less accessible for S12. Still, S12 is binding to a variety of

different RNAs in a non specific way besides binding to its target site on the ribospme. Moreover it was shown that S12 can assist the folding of RNAs and is not required anymore after the RNA is folded correctly. Further, S12 can promote splicing of kinetically trapped group I introns and promote hammerhead cleavage (Coetzee et al. 1994).

The other ribosomal protein we are working with is L1 from the 50S ribosomal subunit. L1 is highly conserved among all three domains of life. However, L1 does not seem to be essential for the translational function of the ribosome as *E.coli* strains lacking the protein are still viable although they have a strongly reduced activity in polypeptide synthesis (Subramanian et al. 1980). L1 constitutes the so called L1 stalk region with its binding to the rRNA. This stalk region is nearby the E-site of the ribosome. L1 seems to be involved in the tRNA translocation and the release of tRNAs from the ribosome (Harms et al. 2001).

It is thought that the movement of the stalk region is responsible for the release of the tRNA. The stalk can either be in an open or in a closed conformation. In the closed conformation the release of the tRNA would be inhibited while in the open conformation the tRNA can be released from the E-site. Moreover, the flexibility of the L1 stalk region is confirmed by the lack of order within this region in available crystal structures (Harms et al. 2001; Wilson et al. 2005).

The crystal structure of L1 consists of two globular domains connected by a hinge region. The L1 protein itself can be either in an open and or in a closed formation. Crystal structures from L1 protein orthologues showed that thermophilic archaeal L1 proteins, from *Methanococcus jannaschii* possess an open conformation, even in the absence of RNA, whereas the thermophilic bacterial L1 protein from *Thermus thermophilus* obtains a closed conformation in the absence of RNA and an open conformation when bound to RNA. In the closed conformation the two domains are shifted towards each other, while when bound to an rRNA they move away from each other (Nikonov et al. 1996; Nevskaya et al. 2000; Wilson et al. 2005).

L1 is also a translational repressor of its own mRNA. By binding to the L11 operon it inhibits further translation of the entire operon and therefore balances the amount of proteins and rRNA. The structure of the L1 protein bound to the mRNA is remarkably similar to the one bound to the rRNA (Wilson et al. 2005).

L1 isolated from *E.coli* has a strong RNA chaperone activity. It has the second strongest RNA chaperone activity of proteins of the large subunit from *E.coli*. The RNA

chaperone activity was also seen in eukaryotic L1 proteins, in other bacteria and in mesophilic archaea. In contrast, L1 proteins from thermophilic archaea do not show any RNA chaperone activity. It was shown for the thermophilic archaeal L1 protein from *Methanococcus jannaschii* that it even inhibits group I intron splicing (Ameres et al. 2007). Interestingly *Mja* L1 shares 70% identity with the mesophilic L1 protein from *Methanococcus vannielii* (*Mva* L1) that possess strong RNA chaperone activity. If the L1 proteins are prebound to RNA they do not have RNA chaperone activity any more. The different behavior of the L1 proteins from diverse organism can be due to a change in the L1 conformation (Semrad et al. 2004; Ameres et al. 2007).

Questions of this diploma thesis:

In this diploma thesis I address the question whether the long unstructured N-terminal domain of the *Eco* S12 protein is required for the RNA chaperone activity. Therefore, I cloned and purified a truncated S12 protein and tested its ability to fold RNA correctly. Moreover, I investigated whether S12 protein orthologues show RNA chaperone activity, too. So far only *Eco* S12 had been studied. Thus, I am analyzing the RNA chaperone activity of several eukaryotic and archaeal S12 proteins.

Analogue to the S12 project, I investigated in a short preliminary project the RNA chaperone activity of ribosomal protein L1 mutants from *M.vannielii* and *M.jannaschii in vitro* and the RNA chaperone activity of L1 orthologues *in vivo*.

Additionally, I was analysing whether L19 is showing RNA chaperone activity in an *in vivo* assay.

2. Analysis of RNA chaperone activity of truncated ribosomal protein S12 from *E.coli* and of eukaryal and archaeal S12 orthologues

The goal of this experiment is the analysis of the chaperone activity of the ribosomal protein S12. S12 is a protein of the small ribosomal subunit. It is located at the interface of the small subunit and is the only protein near the coding site of the small subunit. (Brodersen et al. 2002) It is already known that *Escherichia coli* S12 (*Eco* S12) has RNA chaperone activity (Coetzee et al. 1994). We addressed the question whether the long N-terminal extension is necessary for the RNA chaperone activity of the protein. Moreover, we wanted to know whether S12 protein orthologues show RNA chaperone activity too.

2.1. RESULTS

To analyse the RNA chaperone activity of S12 protein orthologues we chose seven different model organisms from which we cloned the S12 gene into the protein expression vectors pTWIN I and pTWIN II. In addition, we cloned a truncated *Eco* S12 protein missing the first 19 amino acids. The model organisms that we have chosen are the following: the bacteria *Escherichia coli*, the archaea *Methanococcus jannaschii* (*Mja* S12) and *Sulfolobus solfataricus* (*Sso* S12) and the eukaryotic organisms, *Caenorhabditis elegans* (*Cel* S12), *Arabidopsis thaliana* (*Ath* S12), *Saccharomyces cerevisiae* (*Sce* S12) and *Homo sapiens* (*Hsa* S12).

2.1.a. Cloning of S12 genes into pTWIN I and pTWIN II:

We were using the Impact TWIN system from NEB for the cloning of the S12 genes and for the purification of the proteins. The vectors pTWIN I and pTWIN II both contain two different Inteins. An Intein (**Int**ervening prot**ein**s) is a protein splicing element.

The genes can either be cloned with their N-terminal end or with their C-terminal end fused to an intein. Moreover, the choice of the vector depends on the amino acid sequence of the candidate protein as specific amino acid residues have significant effects on the cleavage behavior of the intein. As a result, we decided to clone all of our genes into pTWIN I with the exception of *Eco* S12 which we cloned into pTWIN II. The genes were cloned as C-terminal fusions except for *Eco* S12 truncated which was cloned as an N-terminal fusion, so that S12 truncated did not require to start with the AUG start codon. In addition, *Eco* S12 was cloned N-terminal as well.

The vectors pTWIN I and pTWIN II were digested with the restriction enzymes NdeI and SapI. In addition, pTWIN I was digested with PstI and SapI for the cloning of the truncated *Eco* S12 gene (See also material and methods).

The S12 genes of the seven different model organisms were amplified by PCR and digested with the corresponding enzymes (see material and methods). Then, the S12 genes were ligated into the vectors, transformed into XL1-Blue cells and the insertion of the S12 gene into the vector was controlled by colony PCR. The plasmid was isolated out of the positive colonies and sequenced. Thus, all the S12 sequences with the exception of *Sso* S12 were confirmed. In case of *Sso* S12 we used an old S12 sequence when we designed the primers. As a result, two non polar amino acids were changed into polar amino acids. Therefore, we designed new primers. Unfortunately we ran out of genomic *S.solfataricus* DNA and could not clone the gene without any mismatch any more.

Figure 5a) shows the PCR products of *Eco* S12, the truncated *Eco* S12 and *Hsa* S12. In addition, Figure 5b) shows the positive colony PCR results of *Eco* S12, *Ath* S12, *Hsa* S12, *Sce* S12 and *Mja* S12. The genes of *Cel* S12 and *Eco* truncated S12 were successfully cloned without a colony PCR.

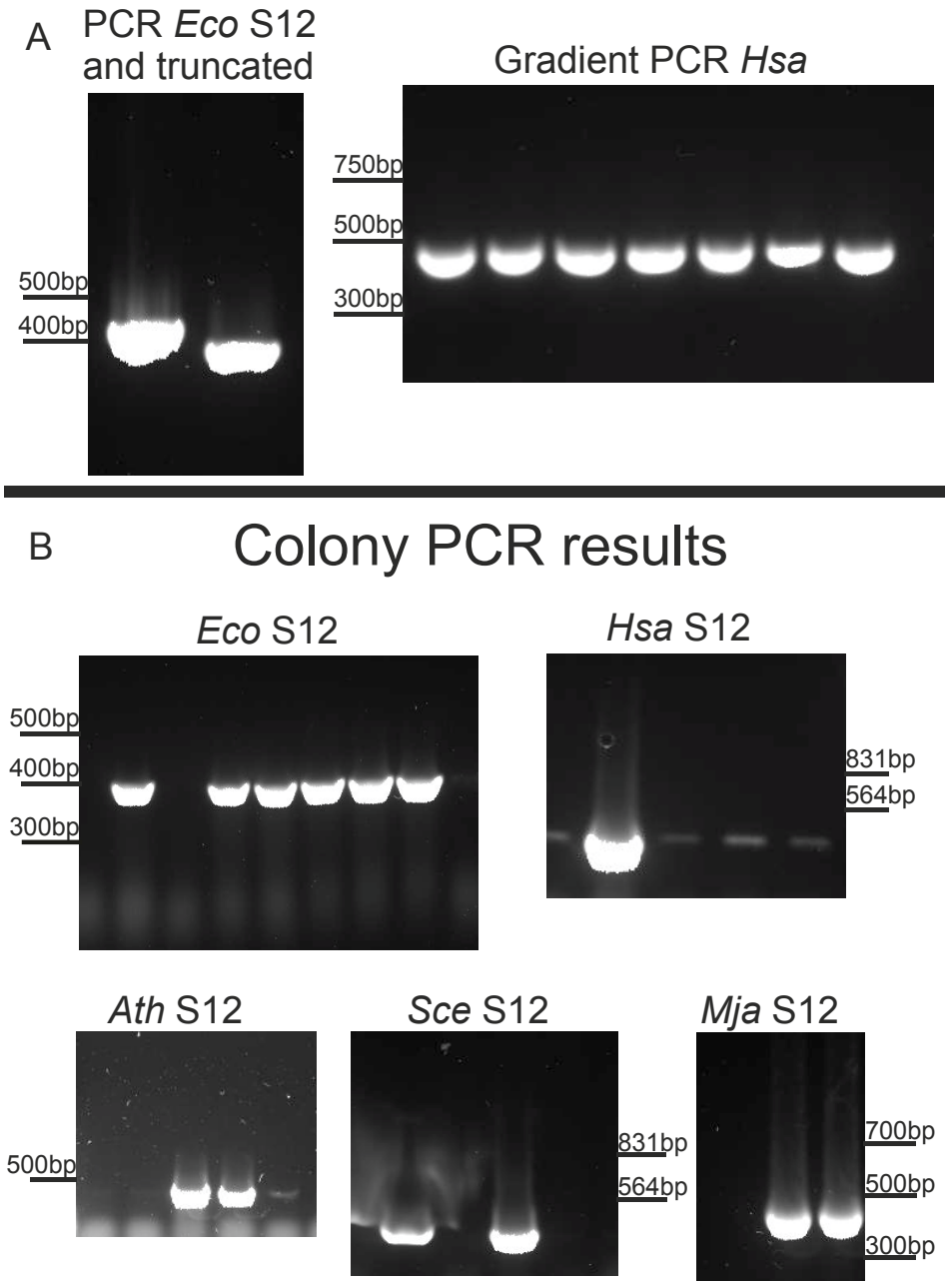


Figure 5: a) Here we see the PCR results from *Eco* S12 and *Eco* S12 truncated as well as a gradient PCR with *Hsa* S12. In **b)** the collected results of the colony PCRs are seen. In each of these cases multiple colonies were picked and analysed. Only the strong bands are really positive ones. Faint bands are considered to be the natural S12 of the cells.

2.1.b. Purification of S12 proteins using the TWIN system:

After the successful cloning of the S12 genes into the pTWIN vectors the proteins had to be purified. Our proteins were tagged with one of two different inteins. *Eco* S12 and *Eco* truncated S12 were tagged with *Ssp* DnaB intein. The *Ssp* DnaB intein was N-terminal fused to the target protein, which was especially necessary for *Eco* truncated S12 as it

did not have the N-terminal methionine. The cleavage of this intein was activated by a pH shift and a temperature shift.

All eukaryotic S12 proteins were fused with the *Mxe* GyrA intein. The intein was C-terminally tagged to the candidate protein. This intein had a thiol inducible cleavage activity. Thus, the addition of DTT was required for the cleavage of the protein from the intein.

Both inteins contain a chitin binding domain. Therefore, chitin beads were required for the protein purification. Due to the chitin binding domain the inteins would still be bound to the column after the cleavage reaction was activated. Thus, the S12 proteins should be isolated in a rather pure form.

For the purification of the proteins the vectors, containing the S12 genes, were transformed into *E.coli* BL21 cells or *E.coli* Rosetta strains. The vectors containing the eukaryotic proteins and the archaeal protein were transformed into the Rosetta strain, because the Rosetta strain increases the expression of eukaryotic proteins which contain codons that hardly appear in *E.coli*. Rosetta contains an additional vector coding for rare tRNAs that are commonly used in eukaryotes and archaea.

After the transformation five colonies were picked and inoculated in 5ml of TBY-E medium. The cultures were experimentally induced for four hours or over night. Then, it was analysed which colony expressed the protein best and the purification was performed with this clone. Although the vectors containing the eukaryotic forms of S12 were transformed into *E.coli* Rosetta strains the induction was significantly decreased compared to the prokaryotic S12 proteins.

The proteins were isolated as described in material and methods. Nevertheless, some problems occurred at the beginning. The biggest problem was the appearance of various unknown bands in the aliquots taken during the purification. The dilemma was that it looked like the protein was hardly binding to the column; especially because a band appeared in the flowthrough aliquot in exactly the same height as the induced protein (Figure 6). As a result, we no longer analysed the isolation by coomassie staining but used western blots instead and visualized the protein bands with anti-intein antibodies. We observed that the band we have seen in the coomassie gels did not appear in the western blots. Thus, this band was not our recombinant protein. On the contrary, this band was some strongly expressed protein in the cell lysate that accidentally was in the same height as our overexpressed protein. To this end, our

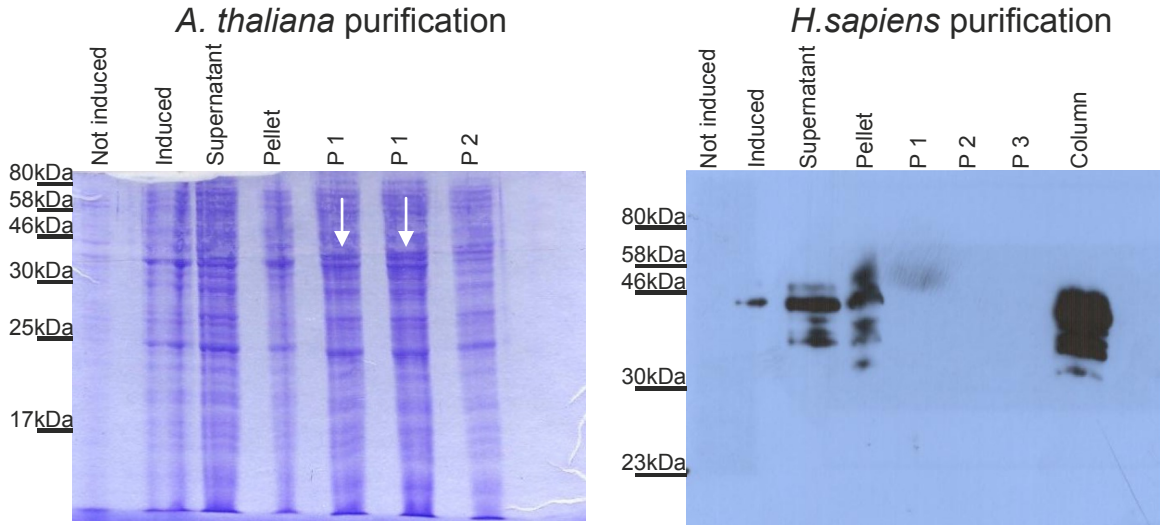


Figure 6: Here we see the comparison between two protein purifications, one analysed with coomassie staining the other with western blot. While several bands can be seen in the different aliquots in the coomassie gel especially in the height of ~ 40kDa (arrows) these bands do not appear in the western blot. In the aliquots P1, P2 and P3 no bands are visible in the western blot indicating that the bands seen in the coomassie gel are unspecific ones and not the induced fusion protein. Thus, further purifications were only studied with western blots (The abbreviations of the aliquots are explained in material and methods).

fusion protein was binding to the chitin beads. Figure 6 compares two typical protein purifications, one analysed with coomassie staining and one analysed with western blot.

Another problem we encountered was that in case of *Eco* truncated S12 the cleaved protein somehow bound to the column. Therefore, we added 0.1%Triton X-100 to the elution buffer. Afterward, we were able to elute this protein. This problem only occurred with the truncated S12 protein.

The elution of the eukaryotic proteins were quite challenging due to their lower induction level. Thus, we tested different induction settings and discovered that *Ath* S12 and *Cel* S12 were better induced when the induction time was limited to 6 hours, while *Sce* S12 and *Hsa* S12 preferred an over night induction (figure 7). Moreover, the *Sce S12* and *Hsa* S12 needed a higher induction temperature for an optimal purification.

Adapting the setting we were able to isolate all of the seven proteins. In case of the *Eco* S12 proteins the purification was additionally confirmed by a Western blot with an anti-S12 antibody (figure 8).

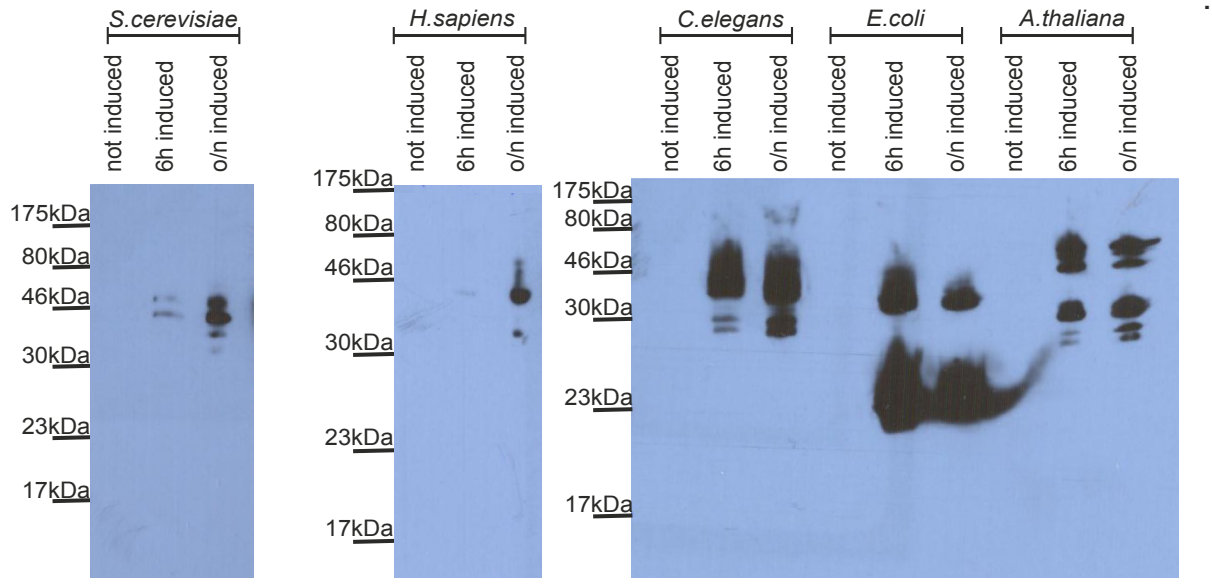


Figure 7: Here we compare the S12 protein expression after different induction times. *Eco* S12, *Cel* S12 and *Ath* S12 are stronger induced after 6 hours and are less degraded than after an over night induction. In contrast, *Sce* S12 and *Hsa* S12 are hardly induced after 6 hours.

2.1.c. Analysing the RNA chaperone activity of the isolated S12 proteins:

Once the proteins were isolated we started analysing their RNA chaperone activity. We decided to use hammerhead cleavage experiments for the analysis. The advantage of the hammerhead cleavage is that two different RNA chaperone activities can be measured: the RNA annealing and the strand displacement activity.

In case of a single turnover assay a limited amount of substrate is added to an excess of hammerhead ribozyme. Therefore, the putative RNA chaperones should help primary with the annealing of the substrate because during the first round of ribozyme substrate binding all substrate will already be bound by a ribozyme.

For the multiple turnover assay a limited amount of hammerhead ribozyme is mixed with an excess of substrate. Thus, in this assay the product release will be observed primarily. This is the case because product release is the rate limiting step and the annealing step of substrate to ribozyme is significantly faster so that the impact of annealing on the overall rate can be neglected.

Both assays are already well established with radioactive labeled substrates. Nevertheless, we decided to use fluorophore labeled substrates, more precisely Cy-5 labeled substrates. Therefore, the optimal settings still had to be established. Both assays worked best at 20°C and are described in material and methods.

We first compared the RNA chaperone activity of the *Eco* S12 protein with the truncated *Eco* S12 protein using the single turnover assay. As a negative control we used a sample without any protein. We could clearly see that the truncated protein was no longer able to promote the cleavage of the hammerhead substrate indicating that the long unstructured N-terminal extension of S12 is indeed necessary for the RNA chaperone activity of S12. The results of the truncated S12 protein were almost the same as the results of the negative control as can be seen in figure 9. In contrast, the *Eco* S12 protein promoted the RNA chaperone activity.

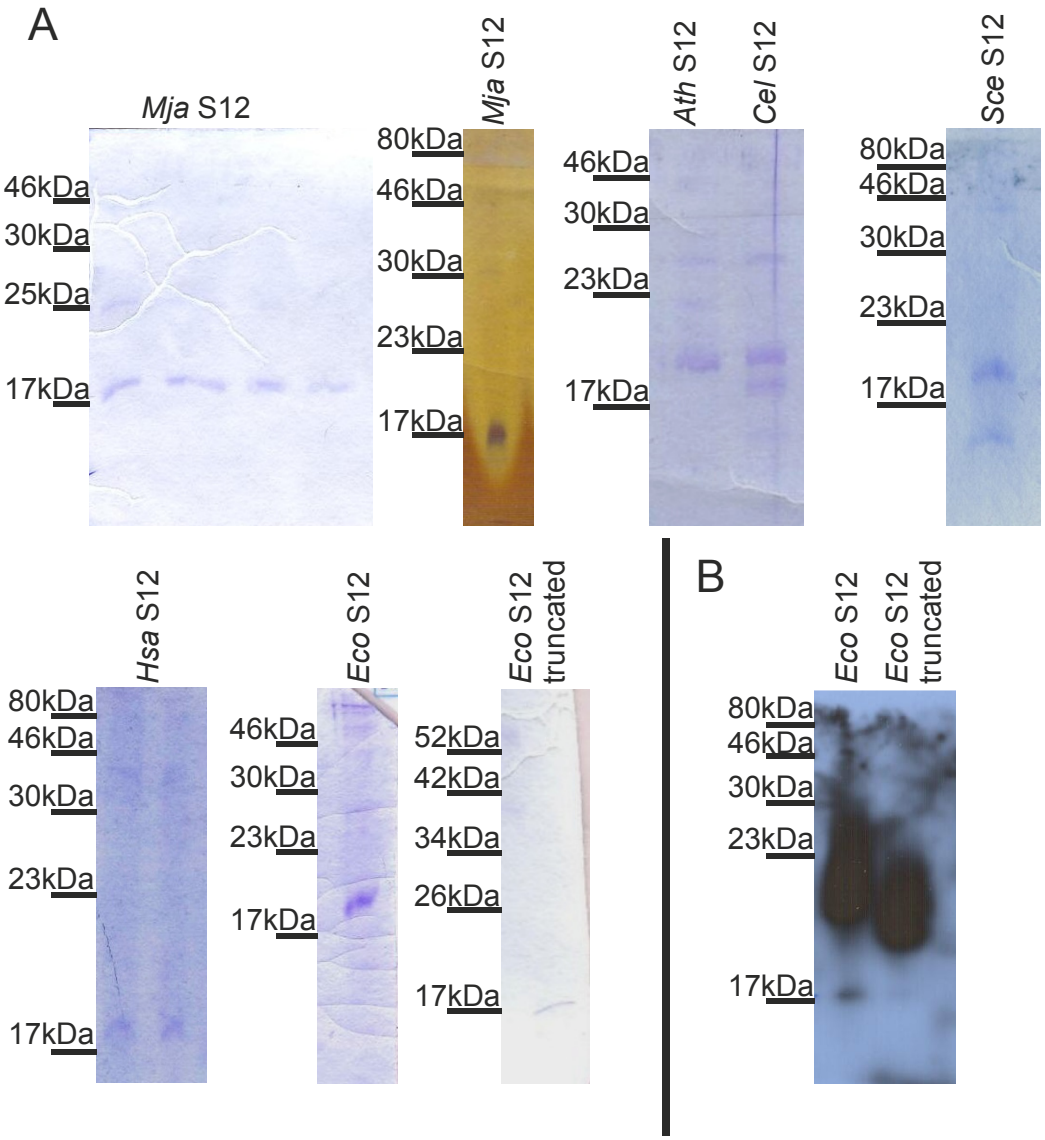


Figure 8 shows the collected results of the protein purification. In **a)** the 7 proteins are seen in gels coloured with either coomassie or silver staining. In case of, *Cel* S12 and *Sce* S12 some degradation can be seen as well. In **b)** a western blot using an *Eco* S12 antibody was performed to verify the identity of the bands.

Further, we analysed the RNA chaperone activity of the eukaryotic S12 proteins and of the archaeal *Mja* S12 protein under single turnover conditions. We observed an increased cleavage of the substrate when the eukaryotic S12 proteins were added to the assay compared to the negative control. This result shows that also all the eukaryotic S12 proteins that we have tested possess RNA annealing activity. The RNA annealing activity of the 4 eukaryotic proteins was very similar and comparable to the RNA chaperone activity of *Eco* S12. In contrast, the *Mja* S12 protein did not show any RNA chaperone activity but it did not inhibit the cleavage of the substrate either. It behaved more or less like the negative control and the truncated *Eco* S12 protein.

In addition, we performed another control using an isolated intein. Although the intein tag of our proteins should not been eluted we wanted to rule out the possibility that some accidentally eluted intein is responsible for the observed RNA chaperone activity. Both of the inteins that we had tested (*Ssp* DnaB intein and *Mxe* GyrA intein) did not show any RNA chaperone activity.

The figures 9 and 10 show the collected results of the single turnover assay. The reaction rate in the absence of any protein was set to 1 (k_{rel}). The obtained reaction rates in the presence of the respective proteins were then calculated in relation to the reaction rate in the absence of a protein. The graph in figure 9c) compares the individual reaction rates from every single assay. As one can see the RNA chaperone activity of the proteins seems to be similar in this analysis. All the proteins which show no RNA chaperone activity behave like the negative control. The graph showing the absolute results (figure 10) is indicating that *Eco* S12 and *Sce* S12 have the strongest RNA chaperone activity. This graph demonstrates the RNA annealing activity of the eukaryotic S12 proteins and the *Eco* S12 protein, while the truncated S12 protein and *Mja* S12 as well as the intein do not show any RNA chaperone activity. With the exception of the intein the assays were performed more than ten times with every single protein.

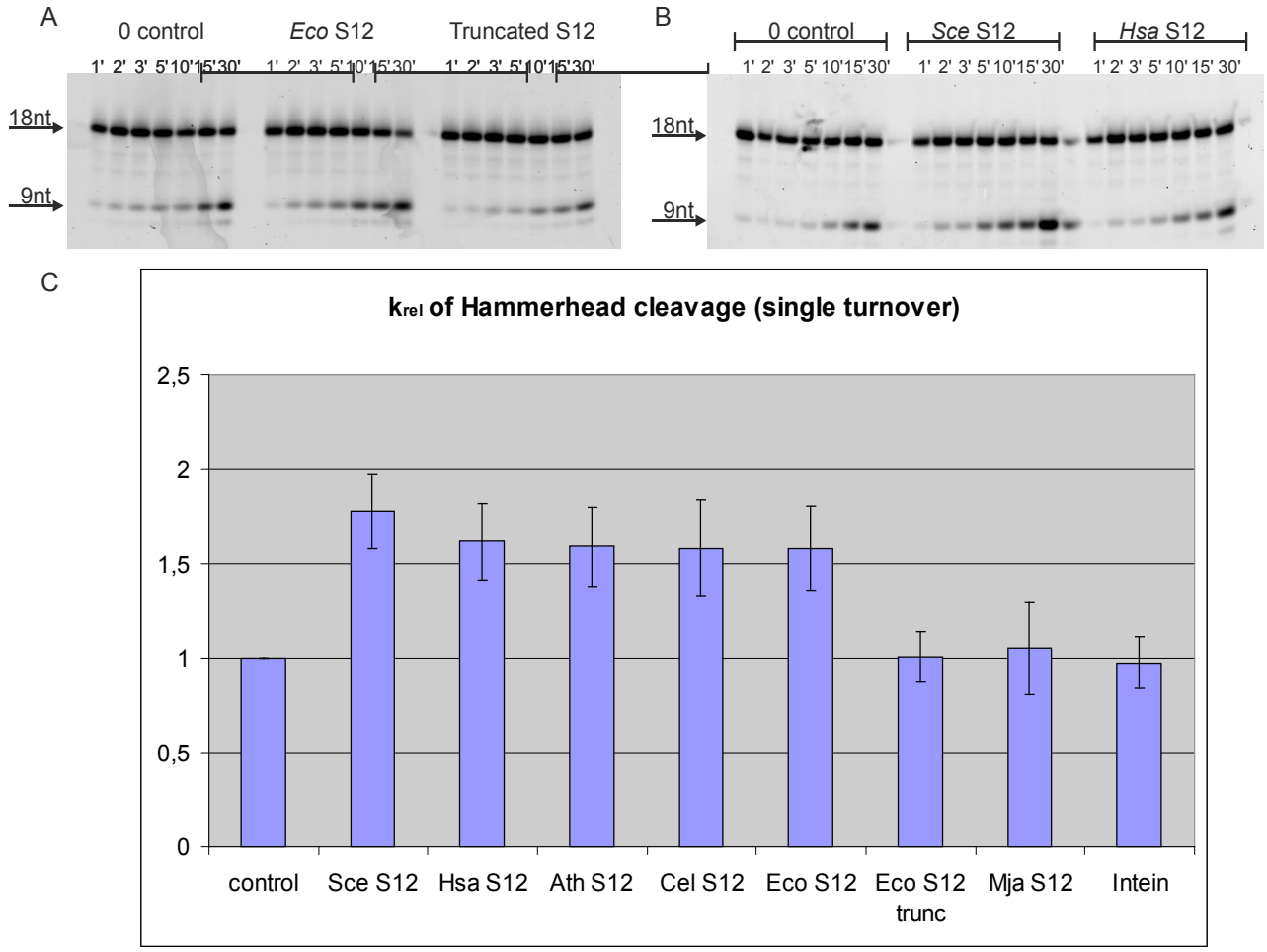
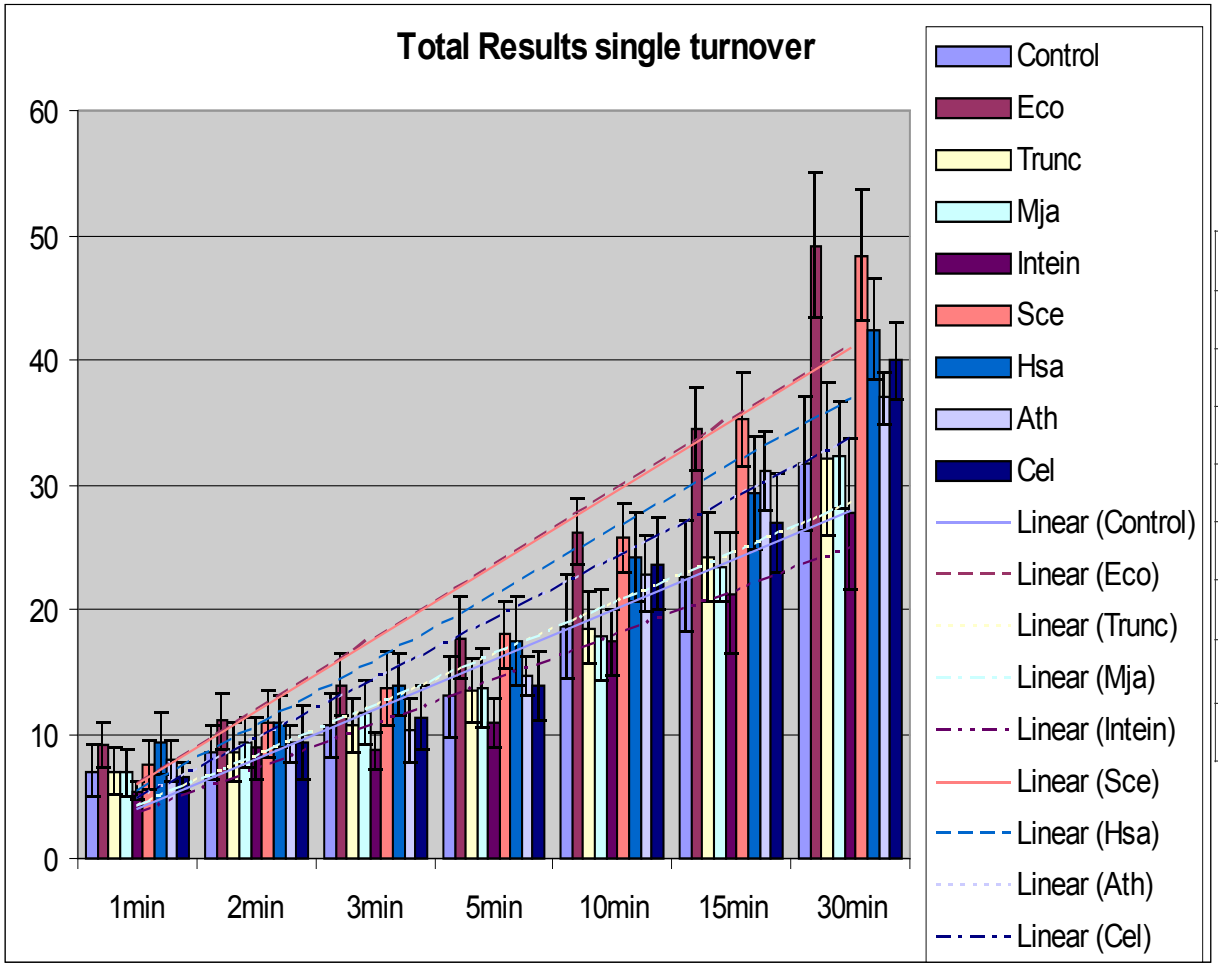


Figure 9: a) and **b)** show two different hammerhead assays. In the first one the cleavage of the substrate in the *Eco* S12 containing sample is far stronger than the ones without any protein or with the truncated S12 protein. Moreover, the full length substrate becomes fainter only in the *Eco* S12 sample. In **b)** one can see that both eukaryotic proteins, *Sce* S12 and *Hsa* S12, promote the cleavage of the substrate.
The graph in **c)** presents the collected relative results of the single turnover assay. The relative result compares the ratio of the trendlines from every sample to the trendline of the control, whereas the control is referred as 1. Every sample was analysed more than ten times. Only the intein was tested just 8 times. All the eukaryotic S12 proteins and *Eco* S12 have RNA annealing activity, while the truncated *Eco* S12 missing its unstructured domain has no longer any chaperone activity. *Mja* S12 is not stimulating the cleavage of the substrate either.



Control	3,9896
Eco S12	5,8948
Trunc	4,0842
Mja S12	4,0819
Intein	3,5752
Sce S12	5,8607
Hsa S12	5,2583
Ath S12	4,8214
Cel S12	4,8071

Figure 10: This figure shows the total collected result of the single turnover assay. Every sample was measured at least 10 times. The total results confirm the relative results. Again all the eukaryotic S12 proteins have RNA chaperone activity whereas *Sce* S12 seems to have the strongest activity. Only *Eco* S12 has an even greater chaperone activity. The results of the control, the truncated S12 protein, and *Mja* S12 are more or less equal. In addition, the intein is not promoting RNA annealing activity either. In this figure *A.th* S12 and *Cel* S12 seem to have the weakest RNA chaperone activity of the eukaryotic proteins. However, in some assay these two proteins had an even bigger RNA chaperone activities than *Sce* S12.

In our next step we tried to confirm our data with the multiple turnover assay. We started again by comparing the RNA chaperone activity of *Eco* S12 and *Eco* truncated S12. Again, we observed an increased cleavage of the substrate when *Eco* S12 was added to the assay while the addition of the N-terminal deleted protein did not amplify the cleavage. This result confirmed the outcomes from the single turnover assay and was once more indicating that the long unstructured N-terminal extension of *Eco* S12 is required for its RNA chaperone activity.

The results with the eukaryotic S12 proteins in the multiple turnover assay confirmed the results from the single turnover assay too as again an increased cleavage of the substrate was observed. Again, the RNA chaperone activity of the 4 proteins was very similar. In contrast, *Mja* S12 did not show any RNA chaperone activity. Moreover, intein did not stimulate the substrate cleavage either.

The graphs in Figure 12 show again the compared individual gradients from every assay and the total results. Every assay was performed at least four times.

The results of both assays confirmed that the long unstructured N-terminal extension of S12 is required for the RNA chaperone activity of the protein and that also the eukaryotic S12 proteins indeed show strand displacement activity. The S12 protein from the thermophilic archaea *M.jannaschii* did not possess any RNA chaperone activity. However, it did not inhibit the cleavage of the substrate either as was observed previously for *Mja* L1.

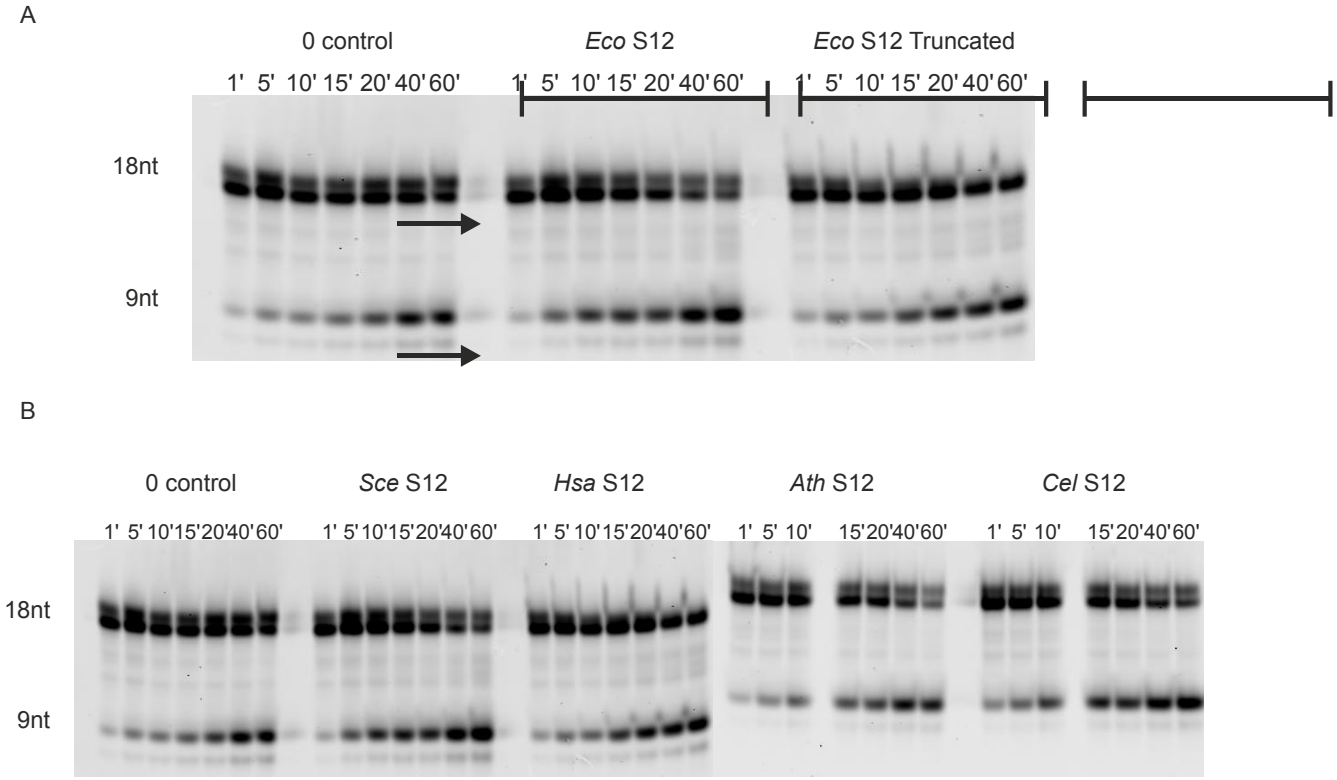
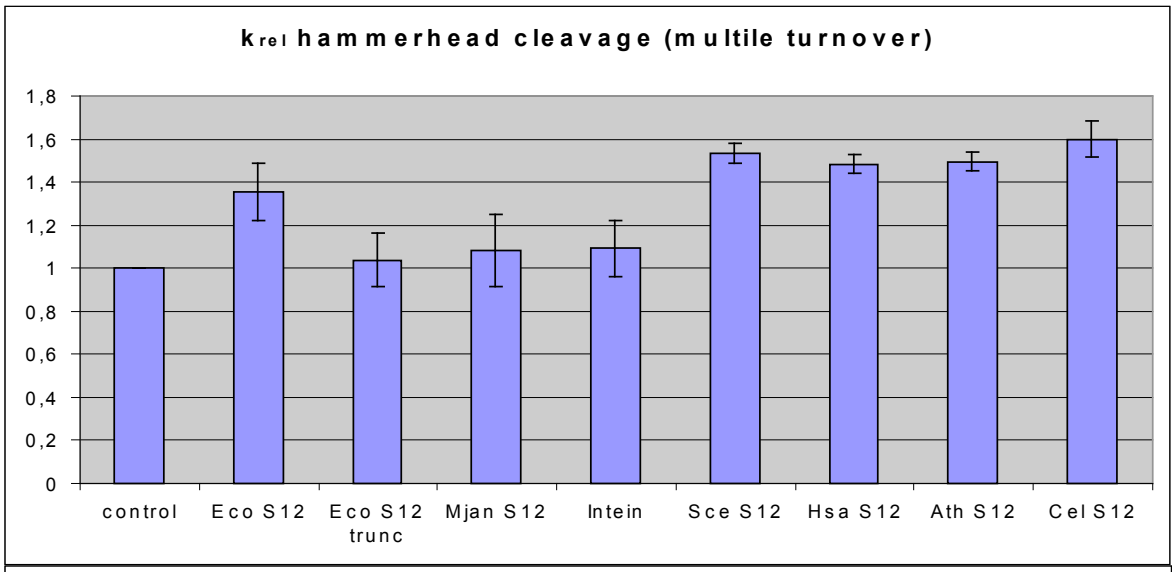
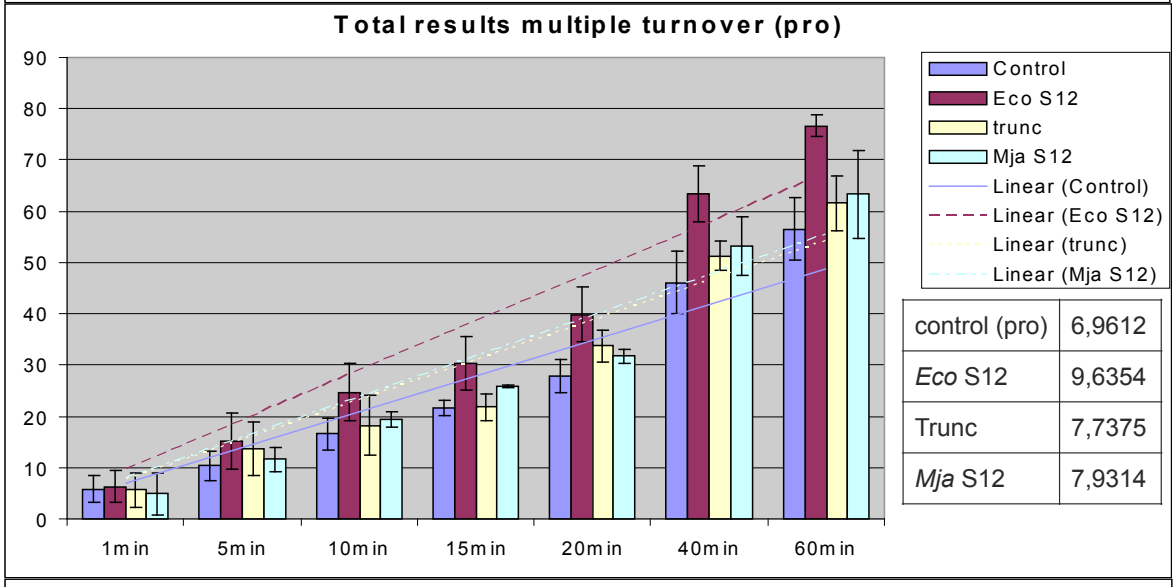
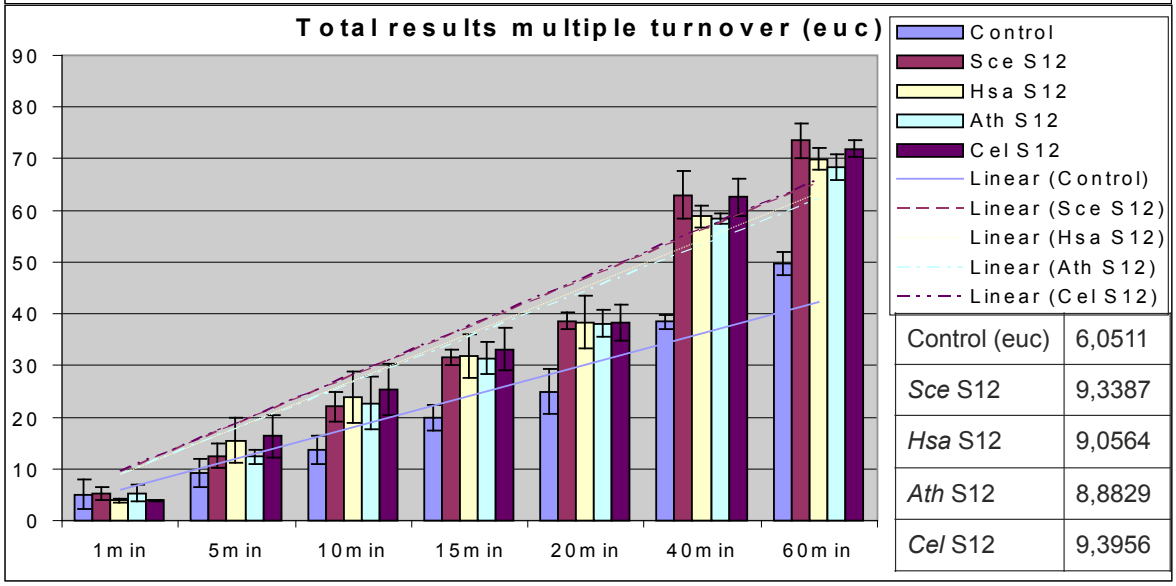


Figure 11: Here two individual multiple turnover assays are seen. In **a)** an assay with *Eco* S12 and the truncated S12 protein is performed. The *Eco* S12 sample is the only one where the uncleaved product becomes significantly weaker. In **b)** one can see an assay with all the eukaryotic proteins. In all cases the uncleaved substrate becomes weaker compared to the zero control.





control (pro)	6,9612
Eco S12	9,6354
Trunc	7,7375
Mja S12	7,9314



Control (euc)	6,0511
Sce S12	9,3387
Hsa S12	9,0564
Ath S12	8,8829
Cel S12	9,3956

Figure 12: The first graph shows the relative reaction rates (k_{rel}) of the hammerhead cleavage reaction under multiple turnover conditions. Like in the single turnover assay all the eukaryotic S12 proteins and *Eco* S12 show RNA chaperone activity. Only *Mja* S12, the truncated S12 and the intein do not stimulate the hammerhead cleavage. The assay was performed at least four times. In case of *Eco* S12 and the truncated S12 the assay was performed 7 times.
The other two graphs show the absolute results of the multiple turnover assay. In the first one, only *Eco* S12 is stimulating the cleavage of the substrate significantly. The third graph shows the results of the eukaryotic proteins. All of them seem to have more or less the same RNA chaperone activity.

2.2. DISCUSSION

The ribosomal protein S12 is part of the small subunit. It is the only protein nearby the coding site of the 30S subunit (Brodersen et al. 2002). S12 is not important for the early steps of the ribosome assembly but it is responsible for the accuracy of the translation. Thus, mutations in S12 can lead to streptomycin dependence because of an increased proofreading selection of tRNA (Ruusala et al. 1984; Bilgin et al. 1992). Moreover *Eco* S12 was shown to possess RNA chaperone activity (Coetzee et al. 1994). We were interested whether the long unstructured N-terminal domain of *Eco* S12 is required for the RNA chaperone activity and whether S12 protein orthologues have RNA chaperone activity too.

To answer these two questions we cloned six orthologue S12 genes and a truncated *Eco* S12 gene without its unstructured domain into adequate vectors and purified all of the seven proteins. Afterwards, we analysed the RNA chaperone activity of these proteins by hammerhead cleavage experiments.

We observed that the S12 protein missing the intrinsically disordered N-terminal domain had no longer any RNA chaperone activity. The assay with the truncated protein behaved like the assay where only protein storage buffer was added. The results were the same in the single turnover assay and in the multiple turnover assay. These results were indicating that the unstructured extension of the S12 protein is indeed required for both measured RNA chaperone activities of the protein. Without the intrinsically unstructured domain the annealing activity was no longer visible in the single turnover assay, nor was the duplex unwinding activity in the multiple turnover assay. Moreover the results confirm that in the case of the ribosomal protein S12 the theory of Peter Tompa that unstructured domains are required for the chaperone activity is valid (Tompa et al. 2004).

We also tested the RNA chaperone activity of the S12 protein from the thermophilic archaea *M.jannaschii*. We were especially interested in the results of this protein as

similar experiments with the ribosomal protein L1 revealed that *Mja* L1 had, in contrast to other organisms, no RNA chaperone activity. On the contrary, *Mja* L1 inhibited group I intron splicing (Ameres et al. 2007). Therefore, we addressed the question whether *Mja* S12 behaved like its L1 counterpart. The results of the hammerhead cleavage experiments indicated that *Mja* S12 had no RNA chaperone activity. It did not possess RNA annealing activity nor strand displacement activity. However, it was not inhibiting the hammerhead cleavage either. On the other hand, the usage of a disorder prediction program revealed that *Mja* S12 had hardly any disorder at the N-terminal region (data not shown). This again confirms the importance of intrinsically unstructured areas for the RNA chaperone activity of proteins.

Moreover, we performed the hammerhead assay with four different eukaryotic S12 proteins. All of these proteins seemed to be RNA chaperones. In the single turnover assay the relative reaction rates of the hammerhead cleavage of *A.thaliana*, *C.elegans* and *H.sapiens* was very similar (figure 9) and more or less as strong as the RNA chaperone activity of *E.coli*. This result demonstrates that these S12 orthologues possess similar RNA annealing activity. Only *Sce* S12 showed higher RNA annealing activity. Nevertheless, if the absolute results were compared both *Eco* S12 and *Sce* S12 had the strongest RNA chaperone activity (figure 10). In the multiple turnover assays the RNA chaperone activities of all the eukaryotic proteins were similar. All of them possess strand unwinding activity. Moreover it seems that the RNA chaperone activity of S12 is phylogenetically conserved between prokaryotes and eukaryotes. Similar experiments were already performed with ribosomal protein L1 orthologues. Also in the case of L1 proteins the RNA chaperone activity seemed to be phylogenetically conserved with the exception of the thermophilic archaea (Ameres et al. 2007).

3. *In vivo* analysis of the RNA chaperone activity of the ribosomal protein L19

The ribosomal protein L19 is located in the large subunit at the subunit interface. It interacts with the proteins L14 and L3 (Harms et al. 2001). Previous trans-splicing experiments showed that L19 seems to have the highest RNA chaperone activity of all ribosomal proteins from the large subunit (Semrad et al. 2004). Of course, these experiments were done *in vitro*. Therefore, we tried to confirm this result also with an *in vivo* experiment. Hence, we performed the previously described poisoned primer assay (Prenninger et al. 2006). However, we did not use a radiolabeled oligo nucleotide primer. Instead, we used a Cy-5 fluorescence labeled NBS-2 primer. To this end, this experiment was also designed to establish the poisoned primer assay with fluorescence labeled oliginucleotides.

3.1. RESULTS

The gene coding for L19 was previously cloned into the pSU20 vector. As a positive control we used the known RNA chaperone StpA (Zhang et al. 1995). This protein was also previously cloned into the pSU20 vector.

All the vectors were transformed into C600 *E.coli* cells and a bunch of colonies were picked and inoculated in TBY-E media containing the required antibiotics. The assay was performed as described in material and methods (More detailed described in (Prenninger et al. 2006)).

Figure 13a) shows a typical growth curve of the assay. Interestingly, the culture overexpressing the L19 protein was always the last culture to reach the optimal OD_{600} value. This was indicating that the overexpression of the ribosomal protein may be toxic for the cells as the overexpression of some ribosomal proteins is poisonous.

The RNA extraction and the poisoned primer assay were performed as described in material and methods. Figure 13b) shows a typical polyacrylamidgel. Splicing occurred in the samples with the positive controls, namely the sample containing StpA and the sample containing the wildtype *thymidylate synthase* gene. No splicing occurred in the negative control and in the sample overexpressing L19.

However, the fluorescence signals were relatively low compared to radioactive experiments. Further, the background signals were too strong to analyse the splicing

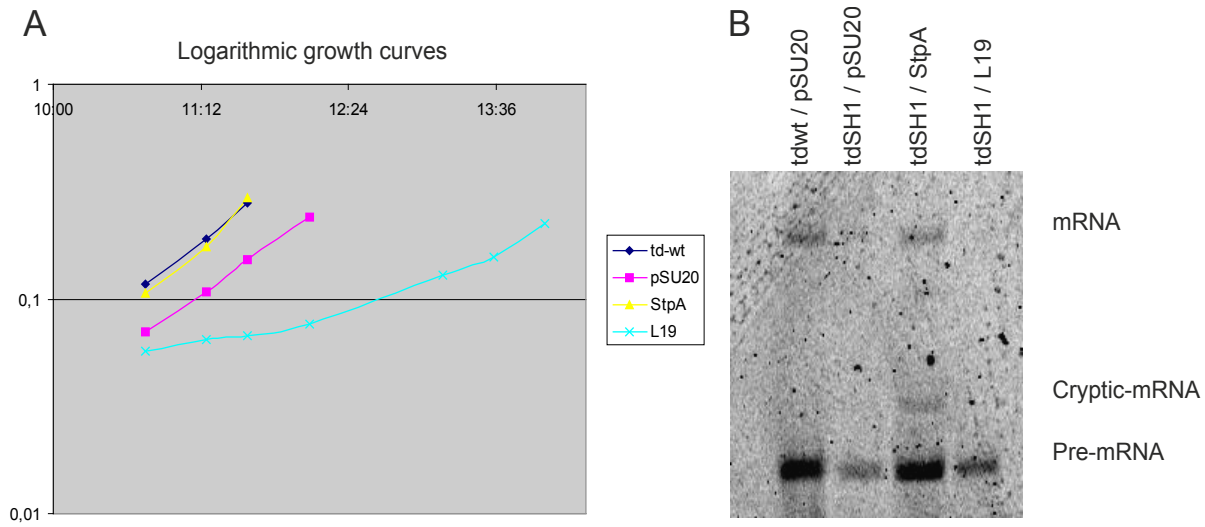


Figure 13: The logarithmic diagram shows the growth curves of the 4 *E.coli* strains. The growth of the strain overexpressing L19 is strongly impaired compared to the ones expressing the wildtype *thymidylate synthase* (*td*) gene, the mutated one, or StpA, respectively. In **b)** a typical poisoned primer gel is seen. Only the wildtype strain and the StpA expressing strain are able to splice the *td* gene.

amount properly. Hence, the assay itself is working with fluorescence but the signal is too weak for accurate analysis. In addition L19 did not show any RNA chaperone activity in the assay although it showed extremely high RNA chaperone activity in the trans-splicing experiments. This may be due to the toxicity of overexpressed ribosomal proteins.

3.2. DISCUSSION

The ribosomal protein L19 showed the strongest RNA chaperone activity of all the ribosomal proteins of the large subunit. Here, we tried to confirm this result additionally with an assay capable of measuring RNA chaperone activity *in vivo*. We tried to perform the assay with a fluorescence labeled primer.

Although our results did not show any RNA chaperone activity of the L19 protein *in vivo* it does not mean that the RNA chaperone activity of L19 in the trans-splicing assay was an *in vitro* artifact. On the contrary, the *in vivo* assay has some limits which become evident here. First of all, some proteins are toxic when overexpressed in *E.coli* and the splicing can not be examined. In our case the cultures overexpressing L19 are growing much slower than all the other cultures indicating that overexpression of L19 may be toxic for the cells.

Another limitation of this assay is that the induction of the possible RNA chaperone can not be accurately adjusted. Therefore, the optimal protein concentration for the RNA chaperone activity could be much lower.

We performed this assay for the first time with a fluorescence labeled NBS-2 primer. This has the advantage that we do not have to work with the more dangerous radioactive samples. Still, we did not know at the beginning if the fluorescence signal would be strong enough for detection.

In our results the primer signal was rather strong and could easily be detected. Nevertheless, the actual mRNAs and pre-mRNA signals were pretty weak. Furthermore, the background signals were incredibly high and made a correct analysis of the results impossible.

We also tried to increase the amount of RNA and to decrease the amount of Cy-5 labeled oligo primer to get better results. However, the changes in the protocol did not lead to better results; on the contrary, it seemed that the results became even worse. Still, more changes can be tried for future experiments. Nevertheless, it seems that for this assay the usage of radioactive labeled primers is better because the signal is stronger.

4. Analysis of RNA chaperone activity *in vitro* of ribosomal protein L1 domains and *in vivo* analysis of L1 orthologues

The ribosomal protein L1 of the large ribosomal subunit is highly conserved among the three domains of life. It consists of two globular domains which are connected by a hinge region (Nikonov et al. 1996).

Recent experiments showed that the ribosomal protein L1 had one of the strongest RNA chaperone activities of the ribosomal proteins of the large subunit (Semrad et al. 2004). Further experiments revealed that the RNA chaperone activity of L1 is conserved in all three domains of life. However, L1 proteins isolated from thermophilic archaea showed no RNA chaperone activity. On the contrary, *Mja* L1 even inhibited group I intron splicing (Ameres et al. 2007).

We were now testing whether one of the two globular L1 domains is responsible for the RNA chaperone activity alone or if both domains together were needed. We also wanted to know whether one of the two domains of the thermophilic archaea *M.jannaschii* is especially inhibiting the RNA chaperone activity. Therefore, we were working with L1 fusion proteins containing one domain from *Mja* L1, and the other domain from *Mva* L1. *Mva* L1 was shown to possess strong RNA chaperone activity *in vivo.* We used two different fusion proteins, one containing domain I from *M.jannaschii* and domain II from *M.vannielii* while the second one was the other way round. In addition, we tested the RNA chaperone activity of L1 protein orthologues *in vivo*.

4.1. RESULTS

4.1.a. Trans-splicing experiments with L1 fusion proteins:

We were working with two different L1 fusion proteins, which we called *Mvajava* L1 and *Mjavaja* L1. *Mvajava* consists of *Mva* domain I and *Mja* domain II while *Mjavaja* consists of *Mja* domain I and *Mva* domain II. Both proteins were previously cloned and purified.

We were testing the RNA chaperone activity of these two proteins with trans-splicing assays using the pre-mRNA of the *thymidylate synthase* gene. The procedure is described in material and methods.

The assay was performed at 55°C without the proteins and at 37°C with and without the proteins. Previous experiments already demonstrated that at 55°C the splicing occurred even in the absence of RNA chaperones while at lower temperatures RNA chaperones were required for the splicing (Semrad et al. 2004). Moreover, we used

Mja L1 and *Mva* L1 to compare their RNA chaperone activity with the ones of the fusion proteins. We observed that both *Mjavaja* L1 and *Mvajava* L1 were hardly increasing splicing of the group 1 intron. Figure 14 shows results from two trans-splicing experiments.

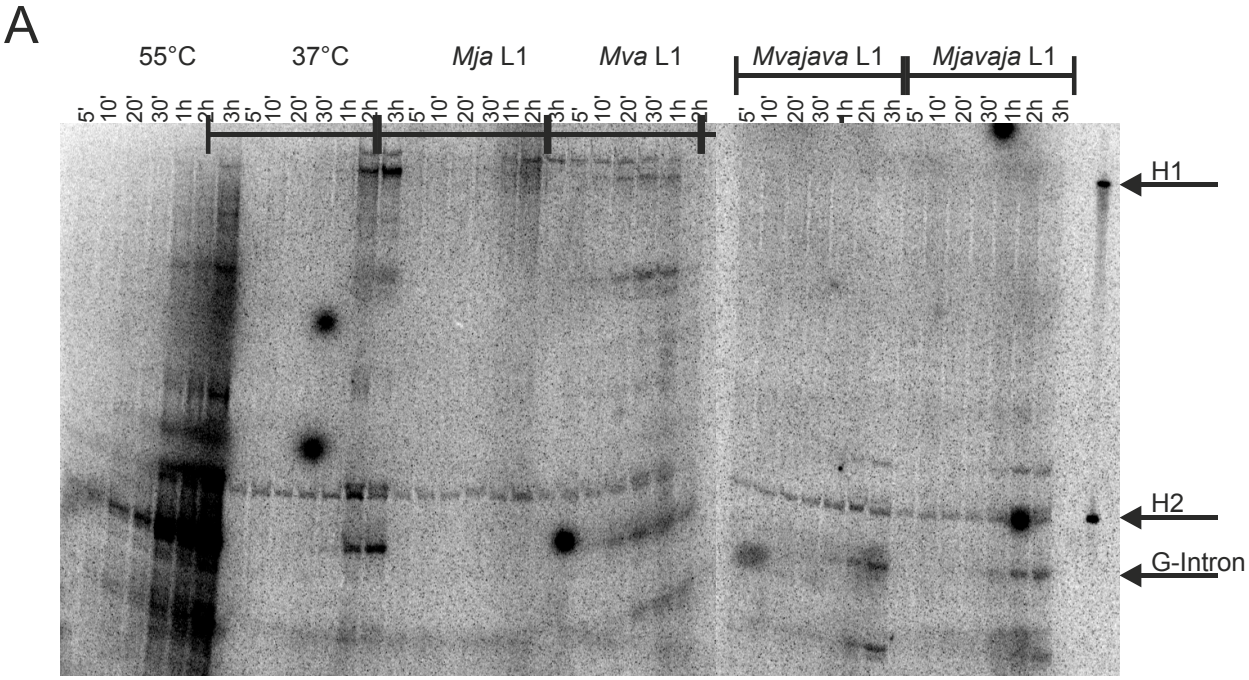


Figure 14: In the trans-splicing assay in **a)** the control at 55°C splices. In the presence of *Mva* L1 trans-splicing occurs even at 37°C. Both *Mvajava* L1 and *Mjavaja* L1 slightly (if at all) promote splicing of the intron at 37°C. At 37°C splicing is significantly impaired. The assay in **b)** confirms the result. *Mva* L1 again promotes splicing at 37°C. *Mvajava* L1 slightly (if at all) promotes splicing; however, the effect of both mutant L1 proteins is extremely low on trans-splicing.

4.1.b. In vivo experiments with orthologue L1 proteins:

We were working with six L1 protein orthologues, namely with the L1 proteins from *Escherichia coli, Saccharomyces cerevisiae, Caenorhabditis elegans, Homo sapiens, Methanococcus jannaschii and Sulfolobus solfataricus*. The genes were previously cloned into pSU20 vectors. As the positive control we used the known RNA chaperone StpA. In contrast to the experiment with L19, this assay was performed with a radiolabelled NBS-2 primer.

All the vectors were transformed into *E.coli* C600 cells. A bunch of colonies were picked and incubated in TBY-E media containing ampicillin, chloramphenicol and thymine. We also tried to incubate the strain containing the wildtype *td* gene into media without any thymine. This was quite effective as the results of the poisoned primer assay were much better. We observed three times more splicing of the *td* gene in this case.

The culture expressing the *Mja* L1 protein was growing rather slow. In general, this culture needed more than 2 hours longer to reach the optimal OD_{600} value than the other cultures. This was indicating that the overexpression of the *Mja* L1 protein was toxic for the *E.coli* cells. Usually, the cultures expressing StpA, the wildtype *td* gene and the mutated *td* gene were growing fastest.

However, the *in vivo* assay did not work properly as the total splice values were rather low. The splicing ratio between the negative control, the positive control and the putative RNA chaperones were often realistic but the total values were far too low. So far it seemed that the eukaryotic L1 proteins from *Sce* L1, *Cel* L1 and *Hsa* L1 and the bacterial *Eco* L1 showed RNA chaperone activity, while *Mja* L1 did not show any RNA chaperone activity. *Sso* L1 war unclear as it showed a weak RNA chaperone activity sometimes but occasionally no splicing occurred. Still, it is not possible to make any clear statements as the splicing ratio was rather low as one can see in figure 15.

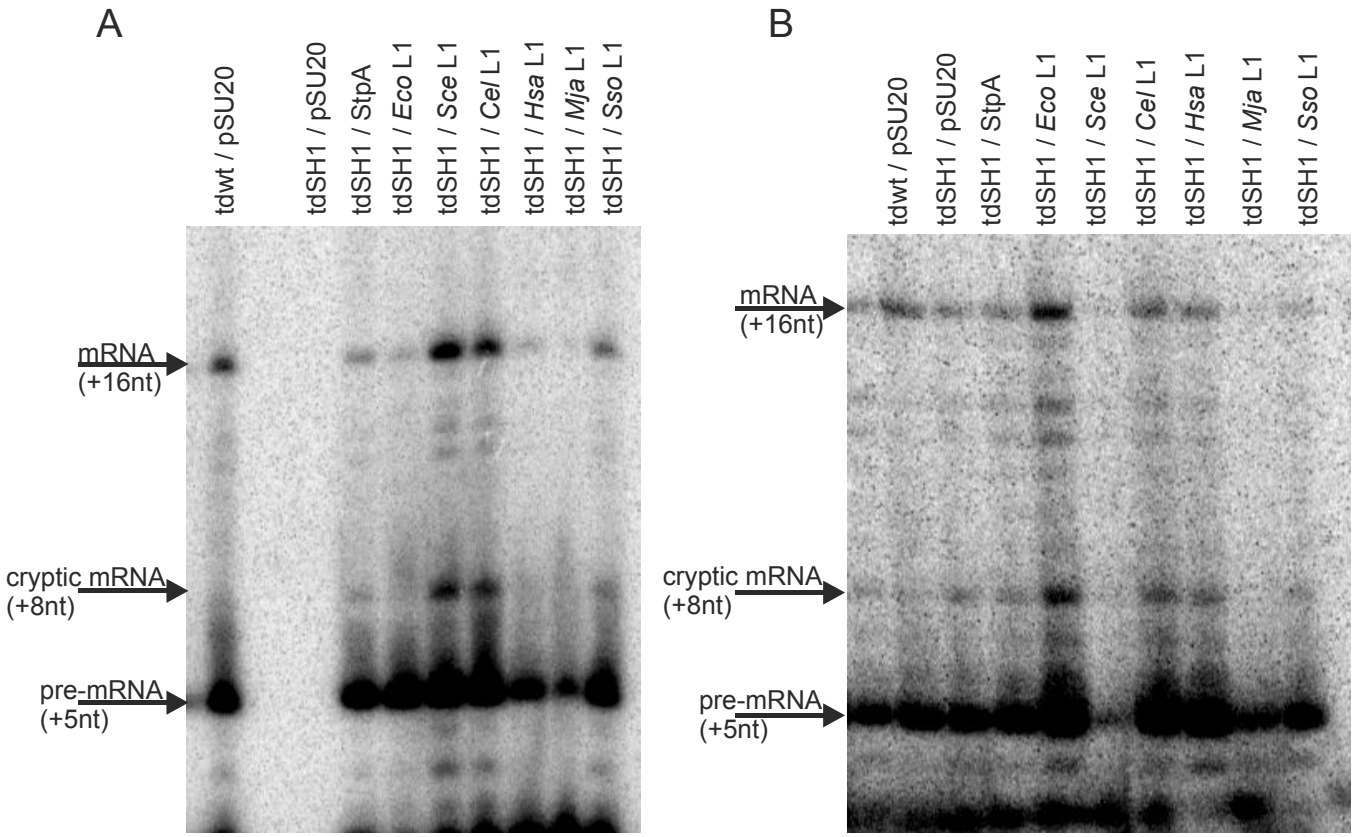


Figure 15: This figure shows two different *in vivo* assay results. The result in **a)** would be quite good if the negative control would not be missing. Still, in these assay *Sce* L1 and *Cel* L1 are promoting the splicing of the *td* gene *in vivo* and are stronger than the wildtype and the positive control. In **b)** another result is shown which should illustrate the contradictory results of this assay. This time *Eco* L1 has the strongest mRNA band but also most of the pre-mRNA. Therefore, none of the proteins tested show any real RNA chaperone activity, not even the positive control StpA.

4.2. DISCUSSION

The ribosomal protein L1 is known for its huge RNA chaperone activity. However, L1 proteins from thermophilic archaea do not show any RNA chaperone activity. In contrast, *Mja* L1 inhibits group I intron splicing (Ameres et al. 2007).

Here, we started to test whether the two domains of the L1 protein have specific roles for the RNA chaperone activity or the inhibition of trans-splicing, respectively. Therefore, we started trans-splicing assays with L1 fusion proteins containing one domain from the RNA chaperone *Mva* L1 and the other domain from the inhibitor *Mja* L1. So far, both fusion proteins, *Mjavaja* L1 and *Mvajava* L1, show hardly any RNA chaperone activity in the trans-splicing assay.

Moreover, we performed *in vivo* experiments with L1 protein orthologues. However, the results are contradictory. On the one hand, the total splicing values are very low. In general, the splicing values are around 8% which is too low to make any clear statement. Only the cultures expressing the *td* wildtype gene spliced sometimes around 30%. On the other hand, the ratio between the samples seems to be quite realistic in most of the experiments. However, there are also exceptions like the *Sso* L1 protein which inhibited the trans-splicing assay but seemed to stimulate the *in vivo* assay occasionally. Moreover, some eukaryotic L1 proteins are showing a stronger RNA chaperone activity than the wildtype in one assay and in the next they hardly have any RNA chaperone activity. Nevertheless, if the splicing activity is that low it is impossible to discover whether a protein shows RNA chaperone activity or not. Further, the *in vivo* assay has strong limitations like reaching the optimal protein concentration. This could explain too why some results of the L1 *in vivo* assays are changing.

5. Material and Methods:

5.1. Preparation of competent *E.coli* cells:

An overnight *E.coli* culture was diluted 1:100 in 100ml TBY-E medium and then grown to an OD_{600} of 0.5 to 0.7. The cells were placed on ice for 15 minutes and afterwards pelleted (5min 4000g). Next, the pellet was resuspended in half a volume 100mM $CaCl_2$ solution and incubated on ice for 5 more minutes. After another harvesting step the cells were resuspended in 1/10x volume of cold 100mM $CaCl_2$ and incubated on ice for 90 minutes. The cells were harvested again, resuspended in 1/10 volume of cold 100mM PIPES, placed on ice for 2 more minutes, pelleted again and finally resuspended in 1/12.5 volume of 100mM $CaCl_2$/ 15% glycerol. The cells were aliquoted, frozen with liquid N_2 and stored at -80°C.

5.2. *E.coli* Transformation:

An adequate amount of plasmid DNA was incubated with 100µl up to 150µl competent *E.coli* cells. The cells were placed on ice for 20 up to 60 minutes, depending on the transformed DNA. A heat shock was performed at 42°C for at maximum 30 seconds and the cells were placed on ice again. 1ml of TBY-E was added and the cells were recovered at 37°C for at least 1 hour. The cells were harvested for 5 minutes at 1430g and then plated on adequate media.

5.3. *In vivo* chaperone assay:

The *in vivo* assay is a method to measure the RNA chaperone activity of a protein *in vivo*. For this assay *E.coli* cultures had to be prepared, the RNA had to be extracted and a poisoned primer assay had to be performed with the isolated RNA. The assay was based on the group I intron splicing of the *thymidylate synthase* (*td*) gene. As a positive control a plasmid containing the wild type *td* gene had to be transformed into *E.coli.* Moreover, a plasmid containing a mutated *td* (*td*SH1) gene (insertion of a stop codon in exon 1 and intronic mutation C865U), which nearly shows no splicing, was transformed as a negative control. An additional vector (pSU20) coding for the putative RNA chaperone and which has an origin of replication compatible to the vector containing the *td* gene was co-transformed. Our control strains contained either the *td* wildtype vector and the empty pSU20 vector, the *td*SH1 vector and the empty pSU20 vector, or (as a positive control for an RNA chaperone) the *td*SH1 vector and the pSU20 vector containing the RNA chaperone StpA. The three individual steps are described next (For more details see (Prenninger et al. 2006)).

RNA cultures for RNA preparation:

An overnight *E.coli* culture was diluted to an OD_{600} of 0,05 to 0,1 in 80ml TBY-E +Chloramphenicol (CAM; 30µg/ml end concentration)+Ampicillin (AMP; 100µg/ml end

concentration)+Thymin (Thy; 36µg/ml), induced with IPTG (1mM end concentration) and grown to an OD_{600} of ~0.3. 50ml of the culture were transferred to 500µl CAM and pelleted for 5 minutes with 3500g. The pellet was resuspended in 1ml TM buffer, transferred to a tube containing 20µl CAM and again harvested for 1 minute at 15115g. The supernatant was removed and the pellet was stored at -80°C.

RNA extraction:

The cell pellets were resuspended in 157µl Solution A (150µl TE; 9.5mM DTT, 0.75µl RNAsin (40U/µl), 4µl Lysoyzm (10mg/ml)) and four times frozen with liquid N_2 and thawed in a water bath. Then solution B was added (200mM $MgAc_2$, 3.5µl Dnase I, 0.1µl RNAsin (40U/µl)) and the samples were incubated on ice for 45 minutes up to one hour. Next, solution C (100mM acetic acid, 5%SDS) was added, the samples were incubated at room temperature for 5 minutes and a phenol extraction was performed with a following RNA precipitation (Addition of 200µl phenol; centrifugation for 5 minutes with 15115g; supernatant added to equal volume of phenol/chloroform; centrifuged; supernatant put to equal volume of chloroform, centrifuged, supernatant precipitated with a tenth volume of 3M NaAc (pH 5.2) and 2.5 volumes of 100% ethanol). The precipitation was performed at -20°C over night. Subsequently, the samples were centrifuged for at least 30 minutes with 15115g at 4°C. The supernatant was removed and the pellets were resuspended in dH_2O.

Poisoned primer assay:

Poisoned primer assays were performed with either fluorescence labeled oligo primer NBS-2 or with radioactive labeled oligo primer NBS-2. For the L19 experiment we used the fluorescence labeled primer and for the L1 experiment the radioactive labeled one. For the poisoned primer assay a 2.5µl reaction mixture containing 10µg RNA was prepared. 1µl hybridization buffer was added as well as 1µM of the Cy-5 labelled NBS-2 oligo (or 1µl ^{32}P-labelled NBS-2). The sample was boiled at 95°C for 1 minute and then cooled down to 42°C. 2.3µl of an extension mix were added (see buffers and media) and the sample was incubated for 1 hour at 42°C. 7µl stop solution and 60µl ethanol were added to perform a precipitation (-20°C for at least one hour, centrifugation at 15115g for at least 30 minutes). The pellet was resuspended in 5 to 10µl loading dye and the samples were loaded onto a 10% polyacrylamid gel. The results were analysed using Image Quant.

5.4. Kinasation of NBS-2:

20pMol NBS-2 were used in a 14µl assay (2µl T4 PNK, 6µl γ-^{32}P-ATP). The assay was incubated for 30 minutes at 37°C and then boiled for 2 minutes with 95°C. To eliminate the remained γ-^{32}P-ATP molecules the sample was loaded onto a G25 spin column. Afterwards, the sample was

loaded onto a 20% polyacrylamid gel and the radioactive NBS-2 band was cut out by using the phosphorimager (a photo was taken with the phosphorimager, then printed and used as a template to cut out the band). The DNA was eluted over night at room temperature in elutionbuffer. Next, the DNA was precipitated and the pellet resuspended in an adequate amount of H_2O.

5.5. Purification of Cy-5 labeled NBS-2:

100µM of the Cy-5 labeled NBS-2 primer were loaded onto a 20% polyacrylamid gel. The band was cut out of the gel and placed in elution buffer. The elution occurred over night at 20°C. Afterwards, the DNA was precipitated and the pellet was resuspended in an adequate amount of H_2O.

5.6. Protein purification:

For the protein purification we used the Impact^{TM} TWIN System by NEB:

500ml up to 1l TBY-E media were inoculated with an *E.coli* overnight culture. The culture was grown to an OD_{600} of 0.4 to 0.5 and then induced with IPTG (end concentration 1mM). The induction occurred at 18 to 24°C for six hours or over night depending on the induced protein (In case of the *E.coli*, *A.thaliana* and *C.elegans* induction occurred for six hours at 22°C; *H.sapiens* and *S.cerevisiae* were induced over night at 24°C, *M.jannaschii* was induced over night at 20°C). The culture was pelleted (3400g 15 min. 4°C) and stored at -80°C. All the following steps occurred at 4°C.

The pellet was resuspended in 25-45ml of either buffer B1 or buffer B2 (depending on the used intein) and then six times sonicated for 10 seconds at full power. The sample was spinned (15 min. 10400g) and the lysate was put onto 6-10ml chitin beads. The proteins were loaded with a flowthrough of 0.4 to 0.5ml/min. Then, the column was washed with ~300ml of the corresponding buffer (flowthrough ~1ml/min). Afterwards, the beads were washed with 9ml of the eluting buffer (B2 or B3). Next, 25ml of the eluting buffer were loaded onto the column and incubated over night. In case of the N-terminal fusion the incubation occurred at room temperature (*Eco* S12 and *Eco* truncated S12). The elution was performed the next day with a flowthrough of 0.4 to 0.5ml/min. Aliquots containing the protein were dialysed into Nierhaus buffer and concentrated into a volume of 0.5 to 1.5ml.

Aliquots were taken before and after the induction, from the flowthrough of the cell lysate (P1), from the flowthrough of the column wash (P2) and from the flowthrough of the eluting buffer (P3). Additionally, samples were taken from the pellet and the supernatant after the sonication. At the beginning all purifications were analysed with coomassie staining. Later, all purifications were analysed with western blots.

5.7. Measuring of the protein concentration:

The protein concentration was measured by Bradford analysis.

5.8. Transkription of hammerhead ribozyme:

Equimolar volumes of the T7-Top oligo nucleotide and the HaHe16 ribozyme DNA were pipetted into an eppi. The DNA was boiled with 95°C for one minute and then cooled down to 37°C. Afterwards, the transcription mix was added (see buffers and media) and the sample was incubated for 3 hours up to over night at 37°C. To eliminate the DNA 6µl of Dnase I(Rnase free) were added and the sample was incubated for additional 30 minutes at 37°C. The RNA was loaded on a 10% polyacrylamid gel and the hammerhead transcript (38nt) was cut out (UV-shadowing) and placed in elutionbuffer. The elution was done over night at 20°C.

Analogue to the hammerhead ribozyme transcription the HaHe16 substrate transcription was performed using the T7-Top oligo nucleotide and the HaHe16 substrate oligo nucleotide.

5.9. Hammerhead ribozyme cleavage:

Single turnover assay:

100pMol/µl of the hammerhead ribozyme were added to 1 up to 1.5pMol/µl Cy-5 labeled hammerhead substrate in a volume of 6µl. The sample was boiled at 95°C for 1 minute and afterwards cooled down to a temperature of 20°C. At the desired temperature the 10x hammerhead buffer was added as well a 1µM of the protein. To the negative control the same amount of the protein storage buffer (Nierhaus buffer) was added. 2µl aliquots of the samples were taken after various time points (usually 1´, 2´, 3´, 5´, 10´, 15´, and 30´) and mixed with loading buffer. The samples were loaded onto a 10% polyacrylamid gel. The cleavage was analysed by using Image Quant. The uncleaved substrate had a size of 18nt the cleaved product was 9nt in size.

Multiple turnover assay:

10pMol/µl of the hammerhead ribozyme were added to a mixture of 10pMol/µl Cy-5 labeled and 45pMol/µl cold hammerhead substrate in a total volume of 5µl. The sample was boiled at 95°C for 1 minute and then cooled down to 20°C. As soon as the desired temperature was reached the 10x hammerhead buffer was added as well as 1µM of the tested protein in a volume of 14µl. The protein storage buffer was added to the negative control. 2µl aliquots were taken after various time points (usually 1´, 5´, 10´, 15´, 20´, 40´, and 60´) and mixed with 6µl loading buffer. The samples were loaded onto a 10% polyacrylamid gel and the cleavage was analysed by using Image Quant.

5.10. Analysis of the hammerhead ribozyme cleavage

The polyacrylamid gels were analysed by using Image Quant. The data graphs were generated comparing the values of the individual samples at every time point. Afterwards, a trendline was set to measure the respective slopes.

To generate the graphs with the k_{rel} of the hammerhead cleavage, the reaction rates in the presence of the proteins of each day were compared with the control reaction rate whereas the control was set to 1. This resulted in the relative reaction rate of the hammerhead cleavage reactions in the presence of the respective protein. Afterwards, the average and the standard deviation of the ratios were calculated. In case of the graphs showing the absolute values, the average and the standard deviation of every single time point was calculated for all the samples. The averages of the single turnover reaction rates were obtained from at least 10 individual experiments. The averages of the multiple turnover experiments were obtained from at least 4 individual experiments.

5.11. *E.coli* colony PCR

A colony was put into 25µl of ddH_2O. The colony was boiled at 90°C for two minutes and afterwards spinned for three minutes with 10000g. 2µl of the supernatant were transferred to 23µl of a PCR mix.

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	25x
20	61	
60	72	
120	72	1

The annealing temperature of the colony PCR was 61°C in case of *Mja* S12, *Hsa* S12 and *Ath* S12, 62°C in case of *Sce* S12, 58°C in case of *Eco* S12 and 65°C in case of *Sso* S12.

5.12. Preparation of cDNA:

To obtain the PCR template DNA for cloning *Hsa* S12, *Sce* S12 total RNA was reverse transcribed into cDNA. Approximately 2µg/µl of the total RNA were mixed with 10mM of an oligo dT in a volume of 10µl. The sample was heated for 5 minutes with 70°C and then put on ice for another 5 minutes. Afterwards, 1mM dNTPs, 40U Rnasin and 10U AMV-RT were added in a total volume of 25µl. The sample was incubated for 1 hour at 42°C.

5.13. Silver staining:

Silver staining was performed with the SilverSNAP Stain Kit II by Pierce according to their protocol.

5.14 MIDI Prep.

MIDI preps were performed with the PureYield Plasmid Mididprep System from Promega according to their protocol.

5.15. PCR product purification and enzymatic reaction clean up:

For the purification of PCR products out of agarose gels and for the elimination of enzymes after restriction digests the Wizard SV Gel and PCR Clean-up system from Promega was used.

5.16. Enzymatic digestion of pTWIN I and pTWIN II for the cloning of S12 genes:

For the cloning of the *Eco* S12 truncated gene into pTWIN I, the vector was first digested with PstI for minimum two hours at 37°C. The cleavage was confirmed with an agarose gel and the vector was precipitated. Next, it was digested with SapI for at least two hours. Afterwards, the vector was purified out of a 1% agarose gel.

For the cloning of the other genes pTWIN I and pTWIN II were digested with the compatible restriction enzymes NdeI and SapI over night at 37°C. The vector was precipitated and then purified out of a 1% agarose gel.

Finally, the purified vectors were dephosphorylated using antarctic phosphatase for 30 minutes at 37°C. The phosphatase was deactivated at 65°C.

5.17. DNA precipitation:

For a DNA precipitation 2.5 volumes of ethanol and a tenth volume of 3M NaAc (pH8) were added to the DNA. The samples were stored for at least 1 hour at -20°C and then spinned with 15115g for at least 25 minutes. The supernatant was completely removed and the DNA resuspended in an adequate amount of H_2O.

5.18. PCR of the S12 genes:

For the PCR a 100µl PCR assay was prepared (0.2mM dNTPs, 100pMol of each primer, 3U Pfu Polymerase, ~50ng DNA). The settings were adapted according to the primers and the DNA length.

Eco S12 and *Eco* Truncated S12:

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	30x
20	58	
240	72	
600	72	1
∞	4	

Sce S12:

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	30x
20	61	
300	72	
600	72	1
∞	4	

Mja S12, *Cel* S12:

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	35x
20	61	
330	72	
420	72	1
∞	4	

Ath S12:

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	30x
20	60	
300	72	
420	72	1
∞	4	

Hsa S12

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	30x
20	57-64	
300	72	
600	72	1
∞	4	

Sso S12:

Time (seconds)	Temperature (°C)	Cycles
300	95	1
20	95	30x
20	64	
300	72	
600	72	1
∞	4	

For the *Eco* S12 PCR genomic MRE 600 DNA was used. For the PCR of *Sce* S12 and *Hsa* S12 RNA was reverse transcribed and then used for the PCR. The *S.cerevisiae* RNA was provided by Andreas Liebeg and the *H.sapiens* RNA was provided by Katarzyna Matylla. The DNA for the PCR of *Cel* S12, *Mja* S12 and *Sso* S12 was previously used by Ameres et al.

5.19. Digestion of the S12 genes:

The S12 gene of *E.coli* truncated was digested with PstI for at least 2 hours at 37°C. Afterwards, the restriction enzyme was removed with the Promega Wizard PCR-cleanup kit and the S12 gene was digested with SapI for a minimum of 2 hours.
All other S12 genes were digested with NdeI and SapI at 37°C.

5.20. Ligation:

The vector DNA and the insert DNA were loaded onto an agarose gel to analyse the ratio between them. According to the gel, the amount of vector and insert DNA was added to a ligation assay. In general, we used three times as much insert as vector. Moreover, 10mM dATP, the ligase buffer and the T4 ligase were added to a 10µl assay. The ligation was performed for at least three hours at room temperature.

5.21. Trans-splicing assay:

For this assay the pre-mRNA of the *thymidylate synthase* gene containing a group I intron was used. It was split in two halves, H1 and H2. H1 contained 549 nucleotides of the exon 1 sequence and 131nt of the intron. H2 contained 147 nucleotides of the 3' half of the intron sequence and the first 23 nucleotides of exon2. The trans-splicing reaction was started with the addition of ^{32}P-GTP.

In vitro transcription of H1 and H2:

100-120µg of H1 plasmid were linearized with SalI at 37°C while 100-120µg of H2 plasmid were linearized with BamHI at 37°C. The DNAs were phenol chloroform extracted and afterwards precipitated. Afterwards the *in vitro* transcription mixes were added to ~60µg of H1 or H2. The *in vitro* transcription was performed until the samples became cloudy. Next, a Dnase I digest was performed for 50 minutes at 37°C. Further, the samples were loaded onto G-50 spin columns to remove the excess α-^{35}S-UTP. Afterwards, H1 and H2 were loaded onto a 5% polyacrylamidgel, the bands were cut out and eluted for 3 hours up to over night in elutionbuffer.

Trans-splicing assay:

200nM of H1 and 200nm of H2 were mixed, boiled for 1 minute at 95°C and then cooled down to 55°C (positive control) or 37°C. Now, splicing buffer, γ-^{32}P-GTP, and 2µM of the respective protein were added. In case of the negative controls without any protein the protein storage buffer (Nierhaus buffer) was added. Aliquots were taken after various time points and mixed with stop solution and TE. The samples were phenol- chloroform extracted, precipitated and loaded onto 5% polyacrylamid gels.

6. Buffers and Media:

TBY-E:

10g tryptone
5g yeast extract
5g NaCl
ad 1l H_2O destilled

Loading dye:

7M Urea
0.25% bromphenol blue
0.25% xylene cyanol
in 1x TBE (89mM Tris, 89mM boric acid; 2mM EDTA)

20x TBE buffer

1.78M Tris
40mM EDTA (pH 8.0)
1.78M boric acid

PCR Mix (Colony PCR):

0.2mM dNTP
Taq buffer
2pMol/µl primer 1
2pMol/µl primer 2
Taq Polymerase

Poisoned Primer:

4.5x hybridisation buffer:

225mM K-Hepes (pH7)
450mM KCl

10x extension buffer:

500mM Tris-HCl (pH8)
50mM $MgCl_2$
50mM DTT

2x stop solution:

300mM NaAc
10mM EDTA (pH8)

Poisoned mix:

0.9mM dATP
0.9mM dCTP
0.9mM dGTP
3.7mM ddTTP

Extension mix:

1µl H_2O
0.67µl extension buffer
0.34µl poisoned mix
0.34µl AMV-RT

TM buffer:

10mM Tris (pH 7,5)
10mM $MgAc_2$

TE buffer

10mM Tris (pH 7,5)
1mM EDTA

***In vitro* transkription:**

10x Tra buffer:

400mM Tris (pH 6,9)
260mM $MgCl_2$
30mM spermidine

Transcription mix for hammerhead:

5mM ATP
5mM CTP
5mM GTP
5mM UTP
10mM DTT
1x Tra buffer
10µl T7 RNA polymerase

Transcription mix for H1/H2:

5mM ATP
5mM CTP
5mM GTP
2.5mM UTP
2.5mM ^{35}S-α-UTP
10mM DTT
1x Tra buffer
20µl T7 RNA polymerase

Elution buffer:

10mM Tris (pH7.5)
250mM NaAc
2mM EDTA

Hammerhead assay:

10x hammerhead buffer

50mM Tris (pH 7.5)
10mM $MgCl_2$

Nierhaus buffer:

20mM Tris (pH 7,4)
400mM NH_4Cl
4mM $MgAc_2$
0.2mM EDTA
5mM β-mercaptoethanol

Trans-splicing assay:

10x splicing buffer

40mM Tris (pH 7.6)
30mM $MgCl_2$
4mM spermidine
40mM DTT

Stop solution

40mM EDTA
300µg/ml tRNA

Protein purification

Buffer B1:

20mM Tris (pH 8.5)
500mM KCl
1mM EDTA

Buffer B2:

20mM Tris (pH 7,0)
500mM up to1M KCl
1mM EDTA

Buffer B3:

20mM Tris (pH 8.5)
500mM KCl
1mM EDTA
120mM DTT

Silver Staining:

SilverSNAP Stain Kit II (Pierce)

10x SDS running buffer:

250mM Tris
1.92M glycine
34.6mM SDS

SDS-PAGE loading buffer:

4%SDS
0.1M Tris pH8.9
2mM EDTA
0.1% bromophenol blue
20% glycerol

Coomassie staining solution:

50% methanol
10% acetic acid
40% H_2O
0.05% Coomassie brilliant blue R-250

Coomassie destaining solution

30% methanol
10% acetic acid

Western Blot

10x transferbuffer (Western Blot)

184mM Tris

1.4M glycine

Up to 1l

1x transferbuffer (Western Blot)

100ml 10x transferbuffer

20% methanol

20x TBS (pH 7,6)

201.17mM Tris

1.198M NaCl

TBS-T:

50ml 20x TBS

0.1% Tween 20

ECL Working solution:

100mM Tris (pH 8.7)

2mM 4-IPBA

1.25mM luminol

Enzymes and Markers:

Protein Markers:

ColorPlus Prestained Protein Marker, Broad Range (7-175kDa), P7709V (NEB)

Prestained Protein Marker, Broad Range (7-175kDa); P7708S (NEB)

Spectra Multicolor Broad Range Protein Ladder; SM1841 (Fermentas)

DNA Markers and Loading dye:

O'RangeRuler 200bp DNA Ladder; SM0633 (Fermentas)

GeneRuler express DNA ladder; SM1552 (Fermentas)

GeneRuler 1kb DNA Ladder; SM0311 (Fermentas)

pBR322 DNA MspI digest; N3032 (NEB)

2-Log DNA Ladder (0,1-10kb); N3200 (NEB)

50bp DNA Ladder; N3236 (NEB)

1kb DNA Ladder; N3232 (NEB)

Lambda DNA/EcoRI + HindIII Marker; G1731 (Promega)

Gel loading dye, blue (6x); B7021 (NEB)

Antibodies:

Anti-CBD Monoclonal Antibody; E8034S (NEB)

Anti-Mouse (goat) peroxidase conj.; 115-035-008 (Jackson Immuno Research)

Anti-S12 antibody (provided by Isabella Moll)

Anti-Goat peroxidase conj.

Restriction Enzymes:

SapI; R0569; 10U/µl (NEB)

PstI; R0140; 100U/µl (NEB)

NdeI; R011; 20U/µl (NEB)

HindIII; R0104 ;20U/µl (NEB)

BamHI; R0136; 20U/µl (NEB)

DpnI; R0176; 20U/µl (NEB)

SalI; R0138; 20U/µl (NEB)

Miscellaneous:

Chitin Beads; S6651L (NEB)

Antarctic Phosphatase; M0289L; 5U/µl (NEB)

AMV-RT; M5108, 1U/µl (Promega)

Dnase I; M0303L, 2U/µl (NEB)

RNAsin; N2615, 40u/µl (Promega)

T7 RNA polymerase; M0251, 50U/µl (NEB)

GoTaq DNA Polymerase; M830; 5U/µl (Promega)

Pfu DNA polymerase; M774; 3U/µl (Promega)

T4 DNA ligase; M0202; 400U/µl (NEB)

T4 PNK; M0201; 10U/µl (NEB)

Bradford; A6932 (AppliChem)

Illustra Micro Spin G-50 Micro Columns (GE Healthcare)

Illustra Micro Spin G-25 Columns (GE Healthcare)

***E.coli* Strains:**

XL1 Blue: F-*, thi, lac, supE44, hsdR17, (F`::Tn10, proAB, laclq,))lacZ(M15), recA, endA*

C600: *F-, thi-1, thr-1, leuB6, lacY1, tonA21 supE44, thy-*

BL21: *F-, gal, hsdS BNL, ompT* (DE3)

Rosetta: *F- ompT, hsdS, gal, dcm* (DE3) pLysSRARE2 (Cam^R); (Novagen 71401)

Sequences and Primers:

NBS-2	5`-GACGCAATATTAAACGGT-3`
T7-top	5`-CGCTAATACGACTCACTATA-3`
HaHe16-ribozyme	5´-GGGAACGTTTCGGCCTTTCGGCCTCATCAGGTCATCGCTATAGTGAGTCGTATTAGCG-3`
HaHe 16-substrate	5'- GGCGACGACGACGTTCCCTATAGTGAGTCGTATTAGCG-3'

S12 primers:

NdeI-S12-Ath	5´-GGCGACCATATGGGTAAGACACGTGGTATG-3´
SapI-S12-Ath	5´-GGTGGTTGCTCTTCCGCAAGACCTAGGCTTCTCCTTCTT-3´
NdeI-S12-Hsa	5´-GGCGACCATATGGGGCAAGTGTCGTGGACTT-3´
SapI-S12-Hsa	5´-GGTGGTTGCTCTTCCGCATGATCTTGGTCTTTCCTTCTTG-3´
NdeI-S12-Cel	5´-GGTGGTCATATGGGAAAGCCGAAGGGACTC-3´
SapI-S12-Cel	5´-GGTGGTTGCTCTTCCGCACGAACGTGGTCTCTCCTTCTT-3´
NdeI-S12-Sce	5´-GGTGGTCATATGGGTAAAGGTAAGCCAAGAG-3´
SapI-S12-Sce	5´-GGTGGTTGCTCTTCCGCATGATCTTGGCTTTTCCTTCTTTTC-3´
NdeI-S12-Mja	5´-GGTGGTCATATGAGTGGAAGTAAATCACCAAG-3´
SapI-S12-Mja	5´-GGTGGTTGCTCTTCCGCATCTTTTGATTTTCTCCTGTCTTC-3´
NdeI-S12-Eco	5´-GGTGGTCATATGGCAACAGTTAACCAGCTG-3´
SapI-S12-Eco	5´-GGTGGTTGCTCTTCCGCAAGCCTTAGGACGCTTCACG-3´
SapI-S12-Eco-trunc	5´-GGTGGTTGCTCTTCCAACAACGTGCCTGCGCTGGAAC-3´
PstI-S12-Eco-trunc	5´-GGTGGTCTGCAGTTAAGCCTTAGGACGCTTCAC-3´
NdeI-S12-Sso	5´-GGCGACCATATGGTTAAAAGCAAGTCACCTAAGG-3´
SapI-S12-Sso	5´-GGTGGTTGCTCTTCCGCACCTGACTGGCTTTTGCTTCTTAC-3´
NdeI-S12-Sso-NEU	5´-GGCGACCATATGAGTAAGAGTAAATCATCTAAG
SapI-S12-Sso-NEU	5´-GGTGGTTGCTCTTCCGCATCTAACTGGCTTTTGCTTC

L1 primers:

SapI-*Mja*-DI	5´-GGTGGTTGCTCTTCCAACACAGTTGTTATAAACACAAGAG-3´
Blunt-*Mja*-DI2	5´-TTTTTTCTTAGCCTTCTCTTTC-3´
Blunt-*Mja*-DI1	5´-ATGGACAGAGAAGCACTGTT-3´
PstI-*Mja*-DI	5´-GGTGGTCTGCAGTTATCCATGTGGAAGCACTACTTC-3´
SapI-*Mja*-DII	5´-GGTGGTTGCTCTTCCAACAGAGGGAAAGAAGCTAAAATAG-3´
PstI-*Mja*-DII	5´-GGTGGTCTGCAGTTATTTCTTTAATCTTTCAACTAATGG-3´
SapI-*Mva*-DI	5´-GGTGGTTGCTCTTCCAACACAGTAGCCATAAATACAAGG-3´
Blunt-*Mva*-DI2	5´-TGATATGGGGGCTGCTGC-3´
Blunt-*Mva*-DI1	5´-ATGGACAGTGCACAAATACAA-3´

PstI-*Mva*-DI	5´-GGTGGTCTGCAGTTATCCGCTAGGTAAGACGATC-3´
SapI-*Mva*-DII	5´-GGTGGTTGCTCTTCCAACACAGTAGCCATAAATACAAGG-3´
PstI-*Mva*-DII	5´-GGTGGTCTGCAGTTATTTTTTGAATCTAGCAACCAATG-3´

7. References

Ameres, S. L., D. Shcherbakov, E. Nikonova, W. Piendl, R. Schroeder and K. Semrad (2007). "RNA chaperone activity of L1 ribosomal proteins: phylogenetic conservation and splicing inhibition." Nucleic Acids Res **35**(11): 3752-3763.

Bashan, A. and A. Yonath (2008). "Correlating ribosome function with high-resolution structures." Trends Microbiol **16**(7): 326-335.

Bilgin, N., F. Claesens, H. Pahverk and M. Ehrenberg (1992). "Kinetic properties of Escherichia coli ribosomes with altered forms of S12." J Mol Biol. **224**(4): 1011-1027.

Brodersen, D. E., W. M. Clemons, Jr., A. P. Carter, B. T. Wimberly and V. Ramakrishnan (2002). "Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: structure of the proteins and their interactions with 16 S RNA." J Mol Biol **316**(3): 725-768.

Clodi, E., K. Semrad and R. Schroeder (1999). "Assaying RNA chaperone activity *in vivo* using a novel RNA folding trap." EMBO J. **18**(13): 3776-3782.

Coetzee, T., D. Herschlag and M. Belfort (1994). "*Escherichia coli* proteins, including ribosomal protein S12, facilitate in vitro splicing of phage T4 introns by acting as RNA chaperones." Genes Dev. **8**(13): 1575-1588.

Dunker, A., C. Oldfield, J. Meng, P. Romero, J. Yang, J. Chen, V. Vacic, Z. Obradovic and V. Uversky (2008). "The unfoldomics decade: an update on intrinsically disordered proteins." BMC Genomics **9**(Suppl 2:S1).

Dyson, H. and P. Wright (2005). "Intrinsically unstructured proteins and their functions." Nat Rev Mol Cell Biol **6**(3): 197-208.

Garner, E., P. Romero, A. Dunker, C. Brown and Z. Obradovic (1999). "Predicting Binding Regions within Disordered Proteins." Genome Inform Ser Workshop Genome Inform **10**: 41-50.

Griffiths, A., S. Wessler, R. Lewontin, W. Gelbart, D. Suzuki and M. J. (2005). 8.1. RNA. Introduction to Genetic Analysis. J. Noe, S. Moran, M. Byrd, M. Kuhta and B. Moscatelli, W.H. Freeman 257-259.

Hammann, C. and D. M. Lilley (2002). "Folding and activity of the hammerhead ribozyme." Chembiochem **3**(8): 690-700.

Harms, J., F. Schluenzen, R. Zarivach, A. Bashan, S. Gat, I. Agmon, H. Bartels, F. Franceschi and A. Yonath (2001). "High resolution structure of the large ribosomal subunit from a mesophilic eubacterium." Cell **107**(5): 679-688.

Herschlag, D. (1995). "RNA chaperones and the RNA folding problem." J. Biol. Chem. **270**(36): 20871-20874.

Herschlag, D., M. Khosla, Z. Tsuchihashi and R. L. Karpel (1994). "An RNA chaperone activity of non-specific RNA binding proteins in hammerhead ribozyme catalysis." EMBO J. **13**(12): 2913-2924.

Holberger, L. and C. Hayes (2009). "Ribosomal protein S12 and aminoglycoside antibiotics modulate A-site mRNA cleavage and transfer-messenger RNA activity in Escherichia coli." J Biol Chem **284**(46): 32188-32200.

Iakoucheva, L., C. Brown, J. Lawson, Z. Obradović and A. Dunker (2002). "Intrinsic disorder in cell-signaling and cancer-associated proteins." J Mol Biol. **323**(3): 573-584.

Karbstein, K. (2010). "Chaperoning ribosome assembly." J Cell Biol **189**(1): 11-12.

Kramer, G., D. Boehringer, N. Ban and B. Bukau (2009). "The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins." Nat Struct Mol Biol **16**(6): 589-597.

Maisnier-Patin, S., W. Paulander, A. Pennhag and D. Andersson (2007). "Compensatory evolution reveals functional interactions between ribosomal proteins S12, L14 and L19." J Mol Biol. **366**(1): 207-215.

Mandal, M. and R. R. Breaker (2004). "Gene regulation by riboswitches." Nat Rev Mol Cell Biol **5**(6): 451-463.

Nevskaya, N., S. Tischenko, R. Fedorov, S. Al-Karadaghi, A. Liljas, A. Kraft, W. Piendl, M. Garber and S. Nikonov (2000). "Archaeal ribosomal protein L1: the structure provides new insights into RNA binding of the L1 protein family." Structure Fold. Des. **8**(4): 363-371.

Nikonov, S., N. Nevskaya, I. Eliseikina, N. Fomenkova, A. Nikulin, N. Ossina, M. Garber, B. H. Jonsson, C. Briand, S. Al-Karadaghi, A. Svensson, A. Aevarsson and A. Liljas (1996). "Crystal structure of the RNA binding ribosomal protein L1 from Thermus thermophilus." EMBO J. **15**(6): 1350-1359.

Patil, A. and H. Nakamura (2006). "Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks." FEBS Lett **580**(8): 2041-2045.

Prenninger, S., R. Schroeder and K. Semrad (2006). "Assaying RNA chaperone activity in vivo in bacteria using a ribozyme folding trap." Nat Protoc **1**(3): 1273-1277.

Radivojac, P., L. Iakoucheva, C. Oldfield, Z. Obradovic, V. Uversky and A. Dunker (2007). "Intrinsic disorder and functional proteomics." Biophys J **92**(5): 1439-1456.

Ramakrishnan, V. and P. Moore (2001). "Atomic structures at last: the ribosome in 2000." Curr Opin Struct Biol **11**(2): 144-154.

Romero, P., S. Zaidi, Y. Fang, V. Uversky, P. Radivojac, C. Oldfield, M. Cortese, M. Sickmeier, T. LeGall, Z. Obradovic and A. Dunker (2006). "Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms." Proc Natl Acad Sci U S A **103**(22): 8390-8395.

Ruusala, T., D. Andersson, M. Ehrenberg and C. Kurland (1984). "Hyper-accurate ribosomes inhibit growth." EMBO J. **3**(11): 2575-2580.

Schmeing, T. and V. Ramakrishnan (2009). "What recent ribosome structures have revealed about the mechanism of translation." Nature **467**(7268): 1234-1242.

Schroeder, R., A. Barta and K. Semrad (2004). "Strategies for RNA folding and assembly." Nat. Rev. Mol. Cell. Biol. **5**(11): 908-919.

Semrad, K., R. Green and R. Schroeder (2004). "RNA chaperone activity of large ribosomal subunit proteins from Escherichia coli." RNA **10**(12): 1855-1860.

Semrad, K. and R. Schroeder (1998). "A ribosomal function is necessary for efficient splicing of the T4 phage *thymidylate synthase* intron *in vivo*." Genes Dev. **12**(9): 1327-1337.

Shcherbakova, I., S. Mitra, A. Laederach and M. Brenowitz (2008). "Energy barriers, pathways, and dynamics during folding of large, multidomain RNAs." Curr Opin Chem Biol **12**(6): 655-666.

Subramanian, A. R. and E. R. Dabbs (1980). "Functional studies on ribosomes lacking protein L1 from mutant Escherichia coli." Eur J Biochem **112**(2): 425-430.

Timsit, Y., Z. Acosta, F. Allemand, C. Chiaruttini and M. Springer (2009). "The role of disordered ribosomal protein extensions in the early steps of eubacterial 50 s ribosomal subunit assembly." Int J Mol Sci **10**(3): 817-834.

Tompa, P. and P. Csermely (2004). "The role of structural disorder in the function of RNA and protein chaperones." Faseb J **18**(11): 1169-1175.

Warner, J. and K. McIntosh (2009). "How common are extraribosomal functions of ribosomal proteins?" Mol Cell **34**(1): 3-11.

Williamson, J. R. (2005). "Assembly of the 30S ribosomal subunit." Q Rev Biophys **38**(4): 397-403.

Wilson, D. and K. Nierhaus (2005). "Ribosomal proteins in the spotlight." Crit Rev Biochem Mol Biol **40**(5): 243-267.

Wimberly, B. T., D. E. Brodersen, W. M. Clemons, Jr., R. J. Morgan-Warren, A. P. Carter, C. Vonrhein, T. Hartsch and V. Ramakrishnan (2000). "Structure of the 30S ribosomal subunit." Nature **407**(6802): 327-339.

Wool, I. G. (1996). "Extraribosomal functions of ribosomal proteins." Trends Biochem Sci **21**(5): 164-165.

Zhang, A., V. Derbyshire, J. L. Salvo and M. Belfort (1995). "Escherichia coli protein StpA stimulates self-splicing by promoting RNA assembly in vitro." RNA **1**(8): 783-793.

Danksagungen

Ich möchte mich recht herzlich für die Unterstüzung bedanken die mir während meiner Diplomarbeit zuteil wurde.

Ich bedanke mich bei meiner Betreuerin Dr. Katharina Semrad, sowie bei den Kollegen von der Arbeitsgruppe Waldsich die mir mit Rat und Tat zur Seite standen wenn ich Hilfe benötigte.

Ein spezieller Dank geht an Wolfgang Reiter, Syam Yelamanchi, Boris Fürtig und Simon Stael deren Tipps für die Proteinaufreinigungen erst zum Erfolg geführt haben.

Weiters bedanke ich mich noch bei der Arbeitsgruppe Schröder für deren freundliche Unterstützung.

Curriculum vitae Martin Alexander López

Date of Birth:	May 14th, 1985
Place of Birth:	Vienna, Austria
Citizenship:	Austria

Education:

1991- 1995:	Elementary school in Vienna
1995- 2003:	Secondary school in Vienna; Graduation with honours
10/2003-9/2004:	Civilian Service at the Red Cross in Vienna
10/2004-10/2010:	Studies of Molecular biology in Vienna
6/2009-6/2010:	Diploma Thesis in the laboratory of Dr. Katharina Semrad at the university of Vienna

Diploma Thesis:

Analysis of RNA chaperone activity of truncated ribosomal protein S12 from *E.coli* and of eukaryal and archaeal S12 orthologues