universität
wien

DISSERTATION

Titel der Dissertation

„Interaction between ribosomal proteins S1 and S2: A novel target for antimicrobials semi-selective against Gram-negative bacteria“

Verfasser

Dipl. Chem. Konstantin Byrgazov

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, 2011

Studienkennzahl lt. Studienblatt:	A 091 490
Dissertationsgebiet lt. Studienblatt:	Dr.-Studium der Naturwissenschaften Molekulare Biologie UniStG
Betreuerin / Betreuer:	Ass. Prof. Dr. Isabella Moll

Table of Contents

1. Summary 4

2. Zusammenfassung 6

3. Introduction 8

3.1. Translation initiation in prokaryotes 11

3.2. Physical and structural properties of protein S1. 14

3.2.1. Physical characteristics of S1 14

3.2.2. Multidomain structure of protein S1 17

3.3. The potential role of protein S1 in translation initiation 26

3.4. References 33

4. Aims of the study 44

5. Manuscripts 45

Direct interaction between the N-terminal domain of protein S1 and protein S2 on the *Escherichia coli* ribosome 46

Structural insights into the assembly of protein S1 to the ribosome in *Escherichia coli* 78

6. Acknowledgements 112

7. Curriculum Vitae 113

1. Summary

The structure of the *E. coli* ribosome is solved at atomic resolution. In contrast, hitherto the position and structure of the essential ribosomal protein S1 has not been determined due to its intrinsic flexibility. Since protein S1 is pivotal for translation initiation in all Gram-negative bacteria studied so far, the aim of this project was the structural characterization of protein S1 with a special focus on the site of interaction with the ribosome. Previously, we have obtained evidence that protein S1 requires protein S2 for binding to the 30S ribosomal subunit. In this study, I was able to show that assembly of protein S1 to the ribosome is mediated by its N-terminal domain D1 that directly interacts with the coiled-coil domain of protein S2. Using an NMR-based approach, I determined that the N-terminal domain D1 consists of a folded core of four β-strands that are flanked by flexible N- and C-terminal regions. Surprisingly, the flexible N-terminal region of domain D1 of protein S1 comprising eighteen amino acids (referred to as $S1_{18}$) is indispensable for binding of protein S1 and its truncated variants to the ribosome. Moreover, I showed that peptide $S1_{18}$ binds to the ribosome and competes with native protein S1 for its binding pocket on the 30S ribosomal subunit.

In addition, my results indicate that the coiled-coil domain of protein S2 is required and sufficient to allow binding of protein S1 to the ribosome. Noteworthy, changing residue Asn145 of protein S2, the side chain of which is oriented towards the cleft of the head, body, and neck of the 30S ribosomal subunit where protein S1 has been proposed to bind, abrogates the interaction between the coiled-coil domain of protein S2 and the N-terminal region of protein S1. Likewise, two glycine residues of the coiled-coil domain located close to the globular

domain of S2 are required for the interaction with the N-terminus of protein S1, since glycine to alanine mutations at these positions (Gly148 and Gly149) also abolish the interaction between proteins S1 and S2.

Taken together, my data support the notion that the flexible region of eighteen amino acids in length located at the N-terminus of protein S1 could serve as a primary interaction site for S1 on the 30S subunit. Due to its intrinsic flexibility the $S1_{18}$ region could act as an anchoring domain, which interacts specifically with residues at the boundary between the coiled-coil and globular domain of protein S2 *via* an induced fit mechanism. Thus, during the course of these studies I was able to narrow down the site of interaction between proteins S1 and S2. Moreover, I identified several residues which might be directly involved in this interaction. Since assembly of protein S1 to the ribosome is essential for the viability of Gram-negative bacteria, this interaction surface might serve as potential target for the design of novel antimicrobial compounds that act semi-selective against Gram-negative pathogens without affecting the Gram-positive flora, which do not harbor functional homologues of protein S1.

2. Zusammenfassung

Zu Beginn des neuen Milleniums ist es gelungen die molekulare Struktur der ribosomalen Untereinheiten und des gesamten Ribosoms von *Escherichia coli* aufzuklären. Im Gegensatz dazu konnte die Struktur des essentiellen ribosomalen Proteins S1 und seine Bindestelle an der kleinen Untereinheit des bakteriellen Ribosomes aufgrund seiner hohen Flexibilität noch nicht bestimmt werden. Da Protein S1 für die Initiation der Translation in allen bisher bekannten Gram-negativen Bakterien notwendig ist, war das Ziel dieser Studie die strukturelle Charakterisierung dieses Proteines, mit einem speziellen Schwerpunkt auf die Interaktion mit dem Ribosom. In vorangegangenen Studien wurde gezeigt, dass für die Bindung von S1 an das Ribosom das Protein S2 benötigt wird. In der vorliegenden Arbeit konnte ich zeigen, dass diese Interaktion durch die N-terminale Domäne von S1 vermittelt wird, wobei der N-Terminus direkt mit der sogenannten „coiled-coil“ Domäne des ribosomalen Proteins S2 interagiert.

Mithilfe von NMR-Studien konnte ich zeigen, dass die Kernstruktur der N-terminalen Domäne von S1 aus vier ß-Strängen aufgebaut ist, die von flexiblen Regionen am N- und am C-Terminus flankiert sind. Überaschenderweise deuten die Ergebnisse meiner Untersuchungen darauf hin, dass der flexible N-Terminus bestehend aus 18 Aminosäuren (hier als $S1_{18}$ bezeichnet) essentiell für die Bindung von nativem Protein S1 und verkürzten S1-Varianten an das Ribosom ist. Weitere Studien zeigen, dass dieses $S1_{18}$ Peptid an das Ribosom bindet und mit dem nativen Protein S1 um die Bindestelle kompetitiert.

Zusätzlich weisen meine Ergebnisse darauf hin, dass die „coiled-coil“-Domäne des ribosomalen Proteins S2 notwendig, aber auch ausreichend für die

Assemblierung von S1 an das Ribosom ist. Interessanterweise zeigen Mutationsanalysen an Protein S2, dass die Aminosäuren Asparagin an Position 145 und Glycin an Positionen 148 und 149 in der „coiled-coil“ Domäne in sterischer Nähe zur globulären Domäne von Protein S2 wichtig für diese Bindung ist, da Mutationen an diesen Stellen die Interaktion mit S1 verhindert.

Zusammengefasst, konnte ich in meinen Studien die Bindung zwischen den ribosomalen Proteinen S1 und S2 näher charakterisieren und die Interaktionsdomänen eingrenzen. Die Ergebnisse meiner Untersuchungen weisen darauf hin, dass der flexible Bereich am N-Terminus von S1 eine primäre Interaktionsdomäne mit der kleinen Untereinheit des Ribosomes darstellt. Es ist denkbar, dass dieser Bereich durch seine hohe intrinsische Flexibilität über einen „induced-fit“ Mechanismus mit der Region zwischen der „coiled-coil“-Domäne und der globulären Domäne des ribosomalen Proteins S2 interagiert. Da diese Bindung essentiell für das Überleben von Gram-negativen Bakterien ist, ist es vorstellbar, dass dieser Bereich ein potentielles Angriffsziel für die Entwicklung von neuen antimikrobiellen Wirkstoffen darstellen könnten, die semi-selektiv gegen Gram-negative pathogene Bakterien wirken. Weiters könnten diese Wirkstoffe die nützliche Gram-positive Flora nicht zerstören, da diese Bakterien kein homologes Protein S1 besitzen.

3. Introduction

In living organisms information deposited in the genomic DNA has to be converted into functional biopolymers. First, during **transcription** coding or non-coding RNAs are produced from DNA where genomic information is stored. Then, a coding or **messenger RNA** (**mRNA**) is converted into a polypeptide chain in the process of **translation** (Crick, 1958) whereas a non-coding RNA (**ncRNA**) is usually employed for control of transcription or translation (Kaberdin and Bläsi, 2006).

During translation mRNA is decoded into proteins by a large ribonucleo-protein particle called the **ribosome**. The prokaryotic ribosome is a particle of 2.5 MDa in size (**70S** ribosome) which consists of two asymmetric subunits. Approximately one third of the ribosomal mass accounts for the **30S** small ribosomal subunit comprising the **16S ribosomal RNA** (**rRNA**) and proteins **S1-S21** and the other two thirds belong to the large **50S** ribosomal subunit composed of **23S** and **5S rRNAs** and proteins **L1-L34** (Berk and Cate, 2007). The 30S subunit mediates the interaction between mRNA and transfer RNA (**tRNA**), the adapter molecules carrying amino-acylated residues which participate in peptide bond formation. There are three tRNA binding sites on the ribosome: the **A-site**, where incoming aminoacyl-tRNAs bind, the **P-site**, where peptidyl-tRNAs reside, and the E-site (named after “exit”), from which deacylated tRNAs leave the ribosome. In the 50S subunit the peptidyl transferase center (**PTC**) is located where peptide bond formation is catalyzed by atoms belonging to the backbone of 23S rRNA (Rodnina, 2007).

The process of translation can be divided into four steps: **initiation**, **elongation**, **termination**, and **recycling**. Each step is assisted by the corresponding

translation factors of initiation (IFs), elongation (EFs), termination (RF), and recycling (RRF). At the first step of **initiation**, a ternary complex between the 30S ribosomal subunit, the mRNA and the formyl-methionine charged initiator $tRNA_i$ bound to the P-site is formed with aid of IFs. In next step, the 50S subunit is docked on the ternary complex by resulting in formation of the 70S initiation complex (**70SIC**) (Antoun, 2006). Then, upon hydrolysis of initiator factor 2 (IF2) bound GTP and subsequent release of IF2, the ribosome can enter the elongation phase. During the first step of **elongation**, the aminoacyl-tRNA enters the A-site with assistance of EF-Tu, a small GTPase. If the cognate codon-anticodon interaction occurs, the EF-Tu-bound GTP is hydrolyzed enabling the release of EF-Tu and accommodation of aminoacyl-tRNA in the A-site followed by peptide bond formation at in the PTC (Rodnina, 2007). Thus, the nascent peptide chain elongated by one amino acid resides on the peptidyl-tRNA in the A-site whereas deacylated tRNA occupies the P-site (Moazed and Noller, 1989). Next, EF-G inducible translocation takes place: EF-G binds to the A-site-bound tRNA and triggers its shift to the P-site concurrently with movement of the P-site-bound tRNA to the E-site (Frank, 2007). Consequently, the ribosome moves along the mRNA until its A-site gets occupied by one of the three stop-codons. This signal is recognized by release factors RF1 or RF2 which initiates the **termination** phase of translation. These two factors induce the hydrolytic release of the polypeptide chain from the peptidyl-tRNA located in the P-site (Kisselev and Buckingham, 2000). Then, RF1 and RF2 are released from the terminating ribosome by RF3 in a GTP-dependent manner (Zavialov, 2002). Subsequently, ribosome recycling factor RRF with aid of EF-G and IF3 promotes dissociation of the ribosome into its subunits (Petrelli, 2001).

Among these stages of translation, initiation is the rate-limiting step (Gualerzi and Pon, 1990). Hence, the assembly of translation initiation complex on the mRNA is a pivotal step in regulation of protein expression. According to the classical model, the ribosome binding site on the mRNA is recognized by the 30S ribosomal subunit *via* a direct interaction between the **Shine** and **Dalgarno** (**SD**) sequence located upstream of the start codon and the anti-SD sequence at the very 3'-end of the 16S rRNA (Shine and Dalgarno, 1974). In the last decade, availability of crystal structures of the prokaryotic ribosome and its subunits revolutionized ribosome research (Schluenzen, 2000) (Wimberly, 2000) (Ban, 2000) (Yusupova, 2001). In 2009, the Nobel Prize for chemistry has been awarded for studies of the structure and function of the ribosome. In addition, several stages of translation initiation have been studied at molecular level (Simonetti, 2009), which will be discussed in the following chapter. Intriguingly, there is one essential component of the ribosome, the function and structure of which are still not fully understood: protein S1. All the known ribosome structures solved at high resolution lack protein S1 because this flexible protein was intentionally removed before crystallization (Wimberly, 2006). Protein S1 mediates initial binding of the mRNA to the 30S subunit by binding to pyrimidine-rich regions upstream of the SD sequence, thereby increasing the concentration of the translational start site in the vicinity of the decoding site on the ribosome (Boni, 1991). Thus, protein S1 is essential for translation initiation in *E. coli* and most likely all Gram-negative bacteria (Sorensen, 1998). Therefore, this study was performed to characterize the interaction of protein S1 with the ribosome and to shed light on the mode of action of this ribosomal protein.

3.1. Translation initiation in prokaryotes

In a prokaryotic cell translation can already be initiated on a nascent mRNA when its Translation Initiation Region (**TIR**) becomes accessible upon transcription. Therefore, coupling of transcription and translation allows the ribosome to mask sensitive sites on the mRNA from cleavage by **endoribonucleases**, which trigger mRNA decay by creating entry sites for 3'-5' exonucleases (Regnier and Arraiano, 2000).

During initiation the ribosome along with the $tRNA_i$ and the mRNA form an active initiation complex. This step is facilitated by the activity of three initiation factors, IF1, IF2, and IF3 (Boelens and Gualerzi, 2002). IF3 promotes dissociation of the ribosome into its subunits (Karimi, 1999) and acts as a fidelity factor discriminating against non-AUG start codons (Hartz, 1990); IF1 promotes more efficient binding of IF2 and IF3 to the 30S subunit (Wintermeyer and Gualerzi, 1983) (Pon and Gualerzi, 1984) and prevents the $tRNA_i$ binding to the A-site (Milon, 2008); the multidomain GTPase IF2 remains associated with the ribosome throughout the entire translation initiation phase and promotes the interaction of the initiator complex assembled on the 30S subunit with the 50S subunit (Antoun, 2003). Despite extensive studies on translation initiation, the exact chronology of events throughout the translation initiation is still controversial (Laursen, 2005) (Simonetti, 2009). It is suggested to occur in highly cooperative manner (Antoun, 2006) (Wintermeyer and Gualerzi, 1983) but in a random order (Gualerzi and Pon, 1990).

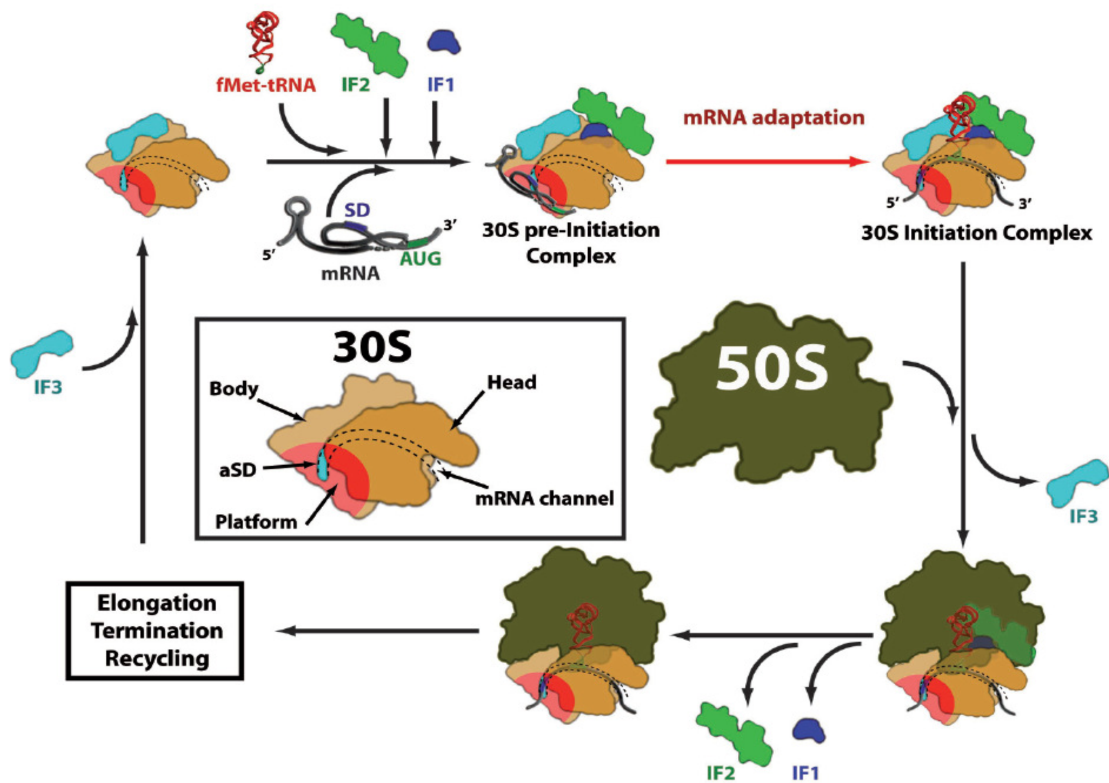


Figure 1. Schematic depiction of the initiation process. Formation of the 30S (30SIC) and 70S (70SIC) translation initiation complexes, containing ribosomes (30S subunit in orange, 50S in brown), initiator fMet-tRNAfMet, mRNA, initiation factors IF1 (in blue), IF2 (in green) and IF3 (in light blue). View of the 30S ribosomal subunit and the ribosome from the top. The platform of the 30S is in red with the anti-Shine-Dalgarno (aSD) sequence in cyan. Structured mRNA binds to the 30S in two distinct steps: the docking of the mRNA on the platform of the 30S subunit forms the pre-initiation complex that is followed by the accommodation of the mRNA into the normal path to promote the codon-anticodon interaction in the P site (Marzi et al., 2007). The resulting 30SIC engages the 50S subunit to form the 70SIC from which the initiation factors are expelled and the synthesis of the encoded protein can proceed through the elongation, termination and ribosome recycling phases (Marzi, 2008) (taken from Simonetti, 2009).

The efficiency of the formation of the translation initiation complex can be affected by secondary and tertiary structures within the ribosome binding site (**RBS**) located on the mRNA. These folded mRNAs are abundant in Gram-negative bacteria possessing a high GC content within their genomes. However,

this class of bacteria can overcome this obstacle with the help of ribosomal protein S1 which is essential for translation initiation in Gram-negative bacteria (Sorensen, 1998). Support for the essentiality of protein S1 in translation initiation has been provided already 30 years ago (van Dieijen, 1976). These results indicated that addition of antibodies specific for protein S1 resulted in inhibition of the ternary complex formation *in vitro*. Moreover, the inhibition can be reversed by neutralizing the antibody with purified S1. Interestingly, after initial mRNA binding the addition of anti-S1 antibodies did not cause inhibition of subsequent translation (van Dieijen, 1978).

One of the crucial roles of S1 in translation initiation is disrupting RNA secondary structures (Szer, 1976). The mono N-ethylmaleimide derivative of S1 (MalN-S1), which contains a modified –SH group on Cys349, binds to the ribosome with the same binding constant as unmodified S1 (Laughrea and Moore, 1978). Nevertheless, 30S subunits containing MalN-S1 are not able to bind the $tRNA_i$ when programmed with structured MS2 RNA (Kolb, 1977). Taken together, these studies support the notion that unwinding RNA secondary structures is an essential function of S1 in translation initiation.

In strong contrast to this pivotal function in protein synthesis, the structure and function of protein S1 are yet to be determined in details. Therefore, the elucidation of the physical properties and complex structure of protein S1 are of great importance to shed the light on the mechanism of protein synthesis.

3.2. Physical and structural properties of protein S1

3.2.1. Physical characteristics of S1

Protein S1 is a peculiar ribosomal protein which exhibits several features distinct from the other ribosomal proteins: *E. coli* S1 has a molecular weight of 61 kDa, which is about two times larger than the majority of ribosomal proteins (Subramanian, 1983); it has a highly elongated structure in solution which spans 210-280 Å, comparable to the largest diameter of the 30S subunit (Laughrea and Moore, 1977) (Giri and Subramanian, 1977); it plays no apparent role in the assembly of the ribosome (Held, 1974) (Culver, 2006) and binds to the ribosome by means of protein-protein interactions (Boni, 1982) with a binding constant of $2x10^8$ M^{-1} (Laughrea and Moore, 1978) (Draper and von Hippel, 1979). Moreover, it can be removed from the ribosome by an excess of mRNA (Suryanarayama and Subramanian, 1983). In addition, protein S1 is exchangeable between ribosome-bound and unbound states (Ulbrich and Nierhaus, 1975) (Pulk, 2010). Nevertheless, more than one binding site on the 30S subunit have been proposed for protein S1: first, ribosomal protein S2 has been shown to be required for binding of S1 to the ribosome, since ribosomes depleted of S2 concomitantly lack protein S1 (Moll, 2002). Second, when 30S subunits are inactivated by decreasing of Mg^{2+} concentration and/or addition of NaCl, a second molecule of S1 can bind to the 30S subunit, however with much lower affinity (Laughrea and Moore, 1977). Intriguingly, the interaction of the second binding site is absent when either the 16S rRNA lacks the 3'-terminal 49-mer containing the aSD sequence (Laughrea and Moore, 1978) or the last C-terminal domain of S1 is absent (Thomas, 1979). Interestingly, protein S1 has been reported to bind to the

aSD sequence at the 3'-end of the 16S rRNA (Dahlberg, 1975). Most likely, both terminal domains of protein S1 play distinct roles in such bimodal binding: the N-terminus of protein S1 is shown to interact with protein S2 (results of this work) whereas the C-terminus is likely involved in a weaker interaction with the aSD sequence.

Immunoelectron microscopic studies revealed that the extended C-terminus of protein S1 expands from the interface of the main morphological domains of the 30S subunit: the head, platform and body (Fig.2A) (Walleczek, 1990). Nonetheless, supporting the existence of the second binding site, Sengupta and co-workers have obtained another visualization for protein S1 on the 30S ribosomal subunit (Fig. 2B) analyzing the 11.5 Å resolution Cryo-EM map of the *E. coli* 70S ribosome containing the $tRNA_i$ and a 46nt long mRNA fragment (Gabashvili, 2000). Their result indicates that both termini of S1 are bound to the ribosome and do not expand to the solvent. The difference between the two models can be explained by the nature and preparation of the particles, as for immunoelectron microscopy high salt washed 30S ribosomal subunits that do not contain additional factors have been used (Fig. 2A), whereas the initiation-like 70S-$tRNA_i$-mRNA complex has been used for Cryo-EM studies (Fig. 2B). Thus, the two models might present distinct snapshots from different steps of translation initiation. Therefore, the comparison of the two models might imply that, in contrast to the first step in initiation, the C-terminus of protein S1 interacts with aSD sequence shortly before elongation. Alternatively, the difference could also be attributed to the different buffer conditions used for samples preparation since spermine and spermidine were added for preparation of the Cryo-EM samples and not for immunoelectron microscopy.

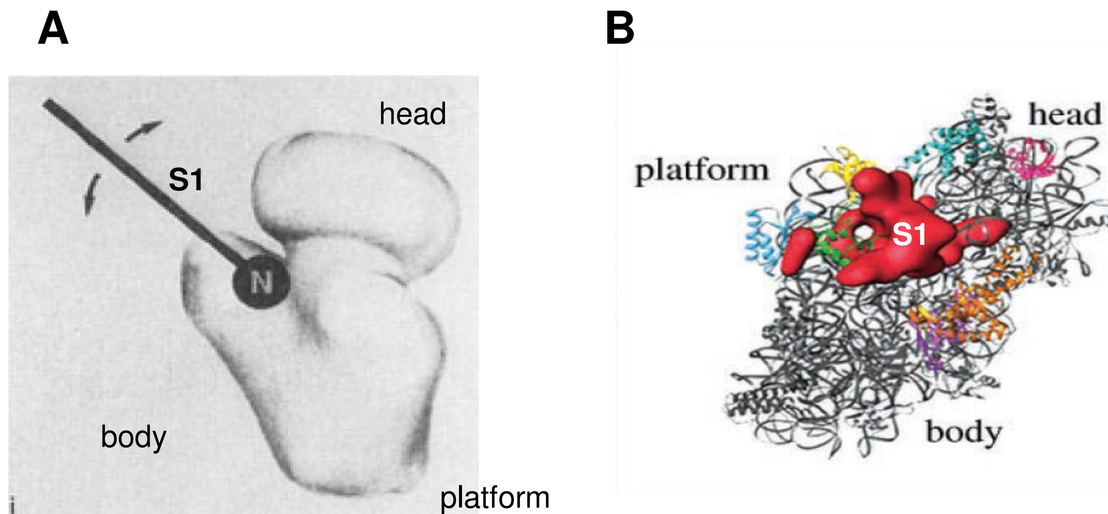


Figure 2. Different visualizations of protein S1 on the 30S ribosomal subunit. (A) Model of S1 bound to the ribosome proposed from immunoelectron microscopy studies (Walleczek, 1990). S1 is visualized as an "mRNA-holding arm" of the ribosome: the N-terminus is represented by a black ball bound to the cleft of the body, head, and platform, whereas the C-terminus, shown as a stick possessing a certain degree of flexibility, is exposed to the solvent. (B) Hybrid Cryo-EM-X-ray map of the complete 30S subunit (Sengupta, 2001). Protein S1 (indicated in red) is docked into the structure of the 30S subunit without prominent exposure of its elongated parts to the solvent.

Due to its elongated structure protein S1 has a large radius of gyration (R_G) of an average value of 70 Å, which is almost three times larger than for globular proteins of the same molecular weight (Oesterberg, 1978). The *in situ* R_G of protein S1 on the 30S subunit has been determined by neutron scattering to be approximately 65 Å (Sillers and Moore, 1981). This fact along with immunoelectron studies might indicate that the extended structure of protein S1 does not significantly change upon ribosome binding. However, there are several lines of evidence indicating that the conformation of the 30S subunit changes upon binding of S1: first, 30S subunits lacking S1 have a higher sedimentation coefficient than its S1-containing counterpart (Dahlberg, 1974); second, hot tritium bombardment

has indicated differences in exposure of some ribosomal proteins upon binding of *T. thermophilus* protein S1 to *E. coli* 30S subunit lacking S1 (Agalarov, 2006). In the presence of S1 proteins S7, S11 and S21 are less labeled, whereas incorporation of tritium in proteins S2 and S6 was increased. The only reasonable explanation why certain proteins become more labelled upon addition of protein S1 is that the binding of the protein causes a conformational change in the 30S subunit resulting in increased exposure of the two ribosomal proteins S2 and S6. Intriguingly, protein S2 has been implicated in stabilization of SD-aSD-duplex during translation initiation (Yusupova, 2006). Thus, it is tempting to speculate that binding of protein S1 to the 30S subunit might facilitate the interaction between the SD-aSD-duplex and protein S2.

3.2.2. The multi-domain structure of protein S1

The actual shape of a protein can be determined only from its crystal structure, which is still not solved for protein S1. However, the physical measurements on S1 have given some indications of its shape in solution. The physical constants of protein S1 are consistent with a highly asymmetric shape (Subramanian, 1983). Small angle X-ray scattering (SAXS) has revealed that the actual shape of S1 does not represent a simple triaxial body (Oesterberg, 1978). Several complex shapes have been proposed for protein S1 such as two V-shaped triaxial bodies or a long cylinder with an attached short cylinder at one end (Oesterberg, 1978) (Labischinsky and Subramanian, 1979).

In Gram-negative bacteria S1 is composed of six domains of approximately 70 amino acids each (Fig.3). Two N-terminal repeats, which have been shown to

be essential for binding to the ribosome (Giorginis and Subramanian, 1980), are distinct from domains D3 to D6 (Gribskov, 1992). The C-terminal domains D3-D6 are folded into five-stranded antiparallel β-barrels similar to **oligonucleotide-oligosaccharide-binding** folds (**OBfold**) (Aliprandi, 2008) (Salah, 2009). This β-barrel fold was accordingly termed **S1 motif** and has been found in a number of proteins associated with RNA metabolism, such as the polynucleotide phosphorylase (PNPase), a component of the *E. coli* RNA degradosome (Bycroft, 1997), or the aspartyl- and lysyl-tRNA-synthases (Commans, 1995) (Eiler, 1999). Interestingly, *E. coli* IF1 represents a single S1 domain (Carter, 2001). In addition to PNPase, the RNA degradosome also contains the endoribonuclease RNase E (reviewed in Carpousis, 2007) which harbors a putative S1 domain (Bycroft, 1997). However, the abundance of S1 motives is not restricted to the prokaryotic kingdom of life. Employing primary sequence analysis, the S1 motif was found in several proteins implicated in RNA binding and processing in eukaryotes, for example the yeast RNA helicase-like protein PRP22 which is involved in pre-mRNA splicing (Schwer and Gross, 1998). Another yeast protein, Rrp5p, which is implicated in pre-rRNA processing (Venema and Tollervey, 1996) and present in the 90S pre-ribosome (Grandi, 2002), contains four putative S1 motives. These examples indicate how the implication of the S1 motif in different functions related to RNA binding and metabolism.

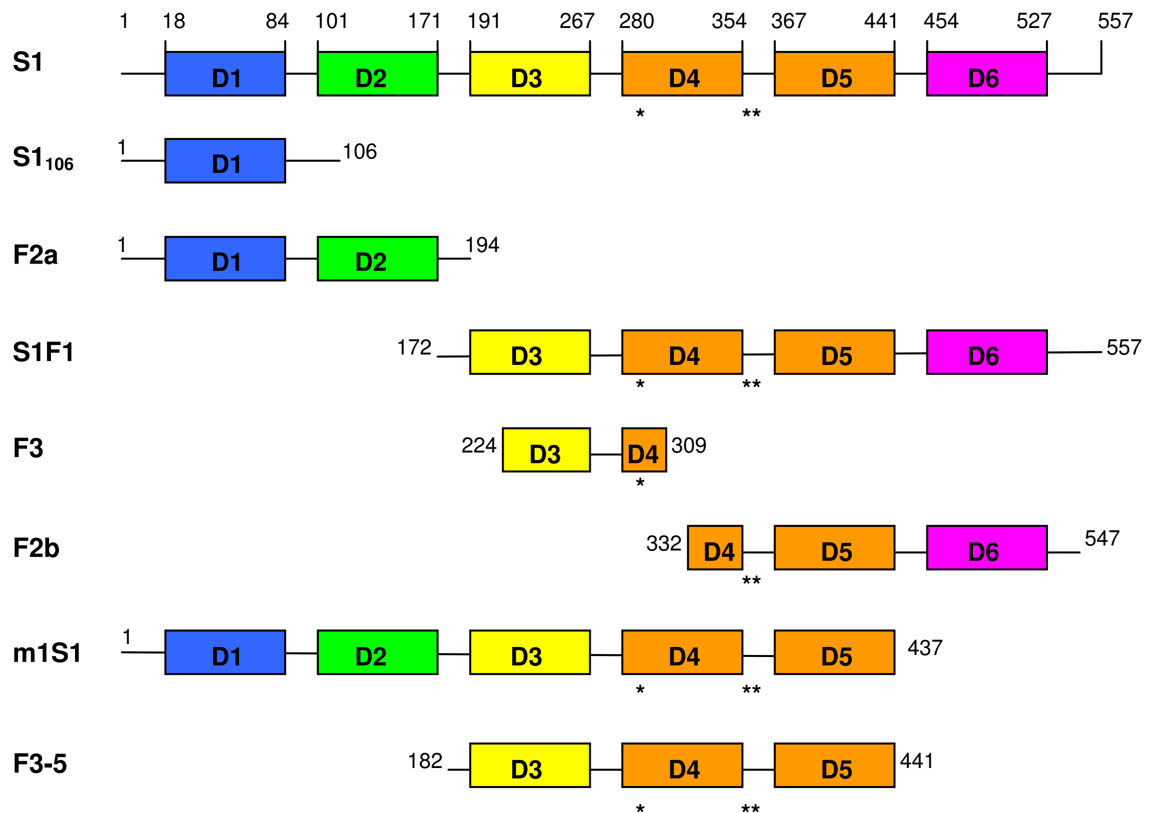


Figure 3. Linear representation of S1, its functional domains and studied fragments. Domains of protein S1 are colored according to the classification from (Salah, 2009). The beginning and ending position of every domain are given according to Pfam. Fragments S1D1 and F2a contains the ribosome binding site of S1, while fragment S1F1 contains the nucleic acid binding domain. The small fragment F3 can bind to nucleic acids strongly whereas F2b does not. Protein M1S1 is 75% as active as S1 in protein synthesis. Two cystein residues in the domain D4, Cys292 and Cys 349, are marked by (*) and (**) correspondently. Modified from (Subramanian, 1983)

Due to the folding behavior of the distinct domains, the structure of several individual domains of protein S1 was determined. The structures of domains D3 and D5 have been obtained by homology modeling (Aliprandi, 2008), whereas the structures of domains D4 and D6 have been solved experimentally (Salah, 2009). Domains D4 and D6 reveal a similar geometry and represent typical β-barrel characteristic of the S1 domain structures (Fig.4) (Bycroft, 1997). The long

loop between strands B3 and B4 is mainly disorganized with a propensity to form a helix turn at each of its extremities. The β-barrels are stabilized by a set of similar hydrophobic interactions. In both barrels, three hydrophobic residues of the strands B1 (L/V-x-G-x-V), B2 (C/A-x-V-x-I/L), B3 (V-x-G-x-V/L) and B5 (I-x-L-x-L/V) participate in the core stabilization. Four hydrophobic residues and aspartate are found in the strand B4 (D-x-V-x-V/A-x-V/F-xx-I/V). In addition, a set of conserved glycines is found at or near the extremities of the strands B1, B2, B3 and B4 which do not participate in the packing of the barrels, but seem important for the connections between the strands (Salah, 2009).

The solution structure of domain D6 is changed similarly upon addition of either poly(A) or poly(U) at 20:1 nucleotide:protein ratio (Salah, 2009). Up to now, no other RNA molecule has been reported to bind to D6. The area affected by RNA binding encompasses residues at the surface of the β-barrel also found in the case of the domains D3, D4 and D5 (Salah, 2009). The fragment F3-5 comprising domains D3, D4 and D5 (Fig.3) has been shown to bind to three different RNA oligonucleotides. Interestingly, the three RNAs are similarly recognized by the F3-5 fragment and the interaction surface is formed on the same region of the three domains as the binding surface of each domain corresponds to the same β-barrel side of the OB fold (Fig. 5). This result confirms the hypothesis of a common interaction surface of all S1-like motives involved in RNA binding (Draper and Reynaldo, 1999).

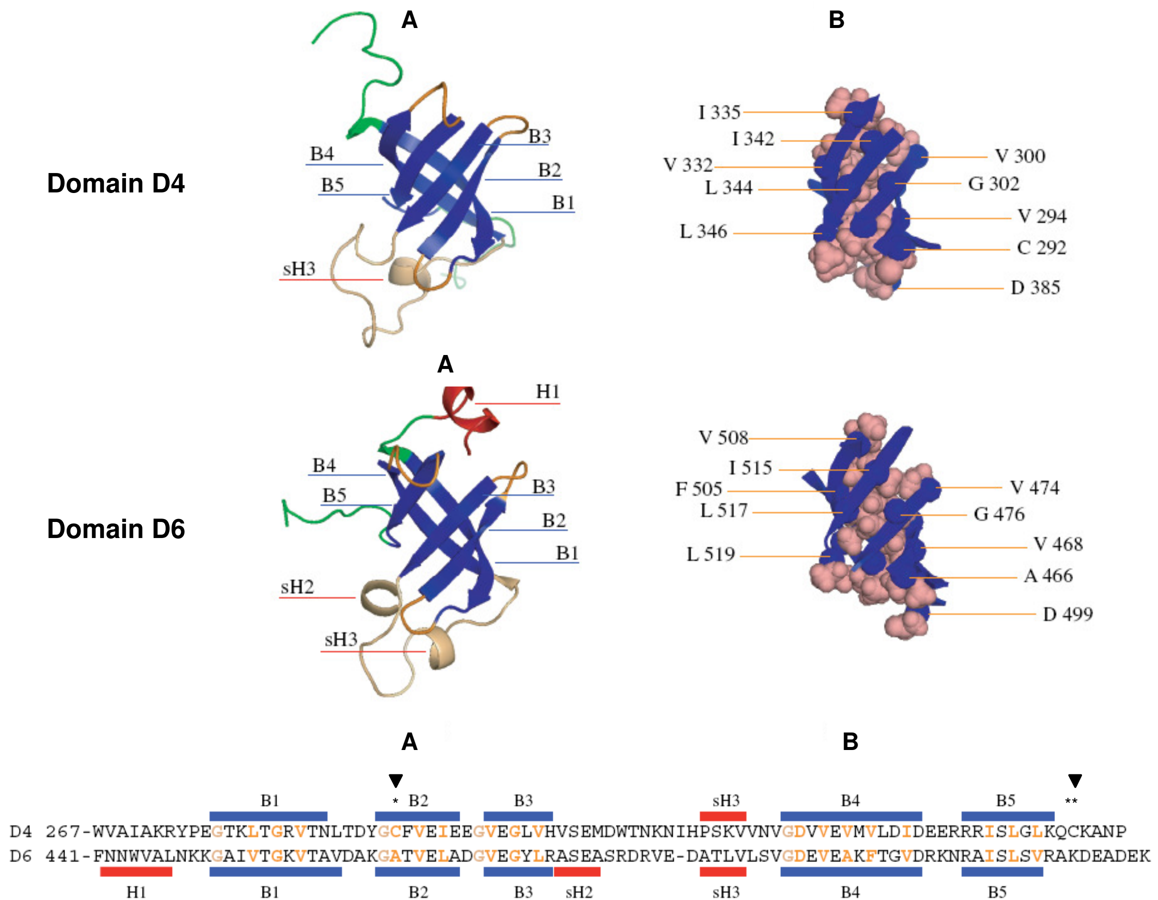


Figure 4. Comparasion of the domains D4 and D6 structures. In (A) is represented a ribbon view of one model; b-strands are in blue, loops in orange, ends in green. The domain D6 has a short α-helix (in red) at its N-terminus. In (B) are represented the residues involved in packing of β-strands forming the β-barrel. They are indicated in orange in the aligned sequences of the domains. Modified from (Salah, 2009).

The RNA binding area in F3-5 spans over all three domains and binds the RNA that is small in size (10-20nt). Topologically, domains D4 and D5 are associated and represent a continuous RNA binding surface while the domain D3 is spatially separated from D4 (Aliprandi, 2008). Therefore, it has been suggested by Aliprandi *et al.*, 2008 that the equilibrium between two forms may occur, namely an “open” (non-interacting D3 and D4) state and a “closed” (loosely interacting D3 and D4) state. Therefore, RNA binding is associated with structural re-

arrangements of the fragment F3-5. It seems that the preformed surface provided by the D4 and D5 is sufficient to bind the RNAs and that D3 can adjust to provide the additional interactions (Aliprandi, 2008).

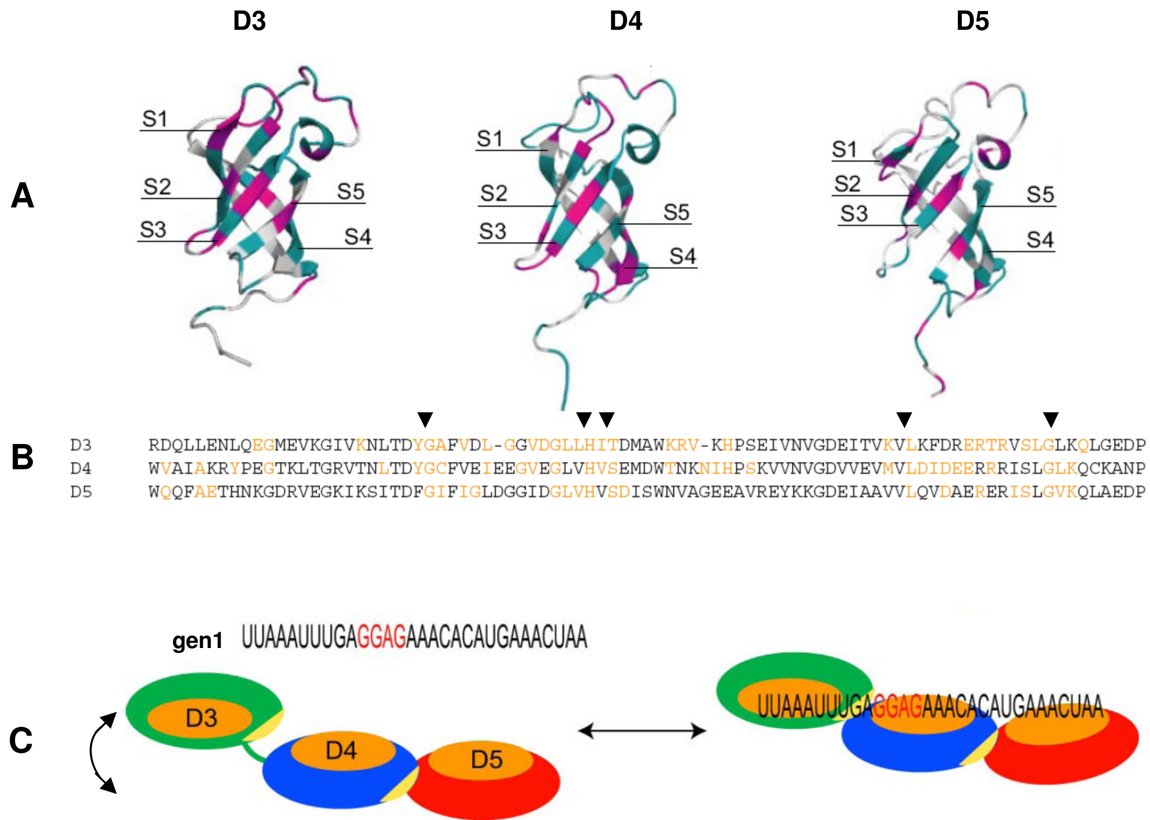


Figure 5. The RNA binding surface of the domains included in the fragment F3-5 and the model of the F3-5 fragment organization and function. In (A) are the model structures of the domains D3, D4 and D5 where the residues affected by RNA binding to F3-5 are colored: in magenta are systematically (by all RNAs) affected residues, in cyan are specifically (by a specific RNA) affected residues. In (B) is alignment of domains D3, D4 and D5 where residues involved in RNA binding are in orange. The highly conserved amino acids systematically affected by RNA binding in all three domains are marked by ▼. They include Gly, His, Thr/Ser, Leu and Gly. In (C) is a likely model of the F3-5 fragment organization. The positions affected by interdomain interactions are in yellow. The domains D4 and D5 are associated and present a continuous interface for RNA interaction (in orange). Taken from (Aliprandi, 2008) and (Salah, 2009) with modifications.

```
D1 ESLKEIETRPGSIVRGVVVAIDKDVVLVDA--GLKSESAIPAEQFKNAQ--GELEIQVGDEVDVALDAVEDGFGETLLSRE---KAKRHEA
D2 WITLEKAYEDAETVTGVINGKVKGGFTVELN-GIRAFLPGSLVDVRPVR--DTLHLE-GKELEFKVIKLDQKRNNVVVSRRAVIESENSAE
D3 RDQLLENLQEGMEVKGIVKNLTDYGAFVDLG-GVDGLLHITDMAWKRVK-HPSEIVNVGDEITVKVLKFDRERTRVSLGL----KQLGEDP
D4 WVAIAKRYPEGTKLTGRVTNLTDYGCFVEIEEGVEGLVHVSEMDWTNKNIHPSKVVNVGDVVEVMVLDIDEERRRISLGL----KQCKANP
D5 WQQFAETHNKGDRVEGKIKSITDFGIFIGLDGGIDGLVHLSDISWNVAGEEAVREYKKGDEIAAVVLQVDAERERISLGV----KQLAEDP
D6 FNNWVALNKKGAIVTGKVTAVDAKGATVELADGVEGYLRASEASRDRVE-DATLVLSVGDEVEAKFTGVDRKNRAISLSVRA--KDEADEK
```

B1 B2 B3 B4 B5

Figure 6. The comparison of the hydrophobic residues (marked in purple) proposed to be involved in the stabilization of β-barrels of the domains of S1. The consensus sequences characteristic of the strands B1, B2 and B4 are well conserved throughout the sequences of the six domains. In the case of the strands B3 and B5 the consensus sequences are well conserved for domains D3, D4, D5 and D6 but are more degenerated in the domains D1 and D2. Positions, where hydrophobic residues substituted by hydrophilic amino acids in the domains D1 and D2, are marked by ▼.

The complex shape of protein S1 indicates that distinct parts of this multi-domain protein are responsible for different roles in protein biosynthesis. Before the availability of detailed bioinformatic sequence profile analysis, several studies on the isolation and functional characterization of discrete fragments of protein S1 had been performed which contributed to dissecting the roles of different parts of protein S1 in its function. All of these fragments (Fig. 3), except m1S1, have been obtained by biochemical approaches probing the structural organization of the S1 molecule.

The protein variant **m1S1**, which harbors a stop codon at position 437, thus lacking the last domain D6, has been serendipitously obtained from an *E. coli* mutant strain which was selected for its resistance to lividomycin (Fig. 3). Protein m1S1 was shown to bind to the 30S subunit lacking native S1 and restores about 75% of the translation activity on synthetic or natural mRNAs (Subramanian and

Mizushima, 1979). This dispensability of domain D6 for essential functions of protein S1 during translation has been confirmed *in vivo* (Schnier, 1986). Interestingly, the unreactive –SH group of Cys292 in full-length S1 becomes reactive in m1S1 (Subramanian, 1981). In addition, whereas a second molecule of S1 can bind to "inactive" 30S subunits, such effect has not been observed for m1S1 indicating that the domain D6 might be required for interaction of protein S1 with the aSD-sequence at the 3'-end of the 16S rRNA (Thomas, 1979). According to SAXS data, protein m1S1 has a complex shape, which is very similar to that of full-length protein S1 (Subramanian, 1983).

Limited proteolysis of protein S1 results in a large fragment **S1F1**, which is totally inactive in restoring its function on ribosomes lacking protein S1 (Suryanarayana and Subramanian, 1981) (Fig. 3). It comprises amino acids from positions 172 to 557 thus lacking the N-terminal residues of native protein S1 (Kimura, 1982). Although preliminary experiments had shown interaction between fragment S1F1 and 30S subunits (Suryanarayana and Subramanian, 1979) later experiments revealed that protein S1F1 cannot bind to the 30S subunits (Giorginis and Subramanian, 1980). Although, fragment S1F1 contains 50 trypsin-sensitive peptide bonds, it is very resistant to further cleavage. In contrast, the N-terminal part of protein S1 is rapidly degraded upon trypsin limited proteolysis even at temperatures of 0°C. Protein S1F1 represents an elongated molecule of maximum length of 210 Å; but, in contrast to S1 and M1S1, it is a simple triaxial body which can be interpreted as a cylinder of 200 Å in length (Subramanian, 1983). Noteworthy, fragment S1F1 is able to bind to nucleic acids (Thomas, 1979) but it cannot unwind RNA secondary structures (Suryanarayama and Subramanian, 1981) despite the presence of the reactive –SH group of Cys349 which is impli-

cated in RNA unwinding function. Taken together, these results reflect the complex nature of unwinding properties of protein S1 which still have to be elucidated.

With the help of BrCN-mediated chemical cleavage at the methionine residues in the S1 chain, three large fragments are produced: **F2a**, **F3** and **F2b** (Subramanian, 1981) (Giorginis and Subramanian, 1980) (Fig.3). Fragment **F2a** contains the sequence from positions 1 to 193, thus including domains D1 and D2 of protein S1 (Fig. 3). It binds to S1-lacking 30S subunits with the same affinity as the full-length S1 (Giorginis and Subramanian, 1980). Over expression of the gene encoding the protein sequence of the F2a variant hinders *E. coli* growth due to displacing the native protein S1 from the ribosome (McGiness and Sauer, 2004). Later, the same effect was observed for protein S1D1 comprising the first 106 amino acids of S1 including the N-terminal domain D1 (Byrgazov, manuscript in prep). Moreover, protein F2a has no detectable affinity either to polyU (Subramanian, 1981) or to natural mRNA (McGiness and Sauer, 2004).

Although no binding affinity to poly(U) had been previously reported for C-terminal fragment **F2b** (Subramanian, 1981) (Fig.3), it contains D6 which has been reported to bind to poly(U) and poly(A) (Salah, 2009).

The fragment **F3** from the central region has been proposed to represent the core of the S1 nucleic acid binding domain. Although its sequence does not contain the intact S1 motif, it binds to poly(U); moreover, the binding is abolished by low concentrations of aurintricarboxylate, which also inhibits the binding between poly(U) and S1. Interestingly, F3 migrates relatively slow on SDS-PAGE, indicating a high apparent molecular weight of 16 kDa when analyzed by this procedure (Subramanian, 1981).

3.3. The potential role of protein S1 in translation initiation

In all organisms gene expression can be regulated post-transcriptionally at the step of translation initiation in response to various external signals. Studying the molecular mechanisms involved in translation initiation can lead to a better understanding how the synthesis level of particular proteins is subjected to regulation (Boni, 2006). According to Shine and Dalgarno, the general mechanism of translation initiation in prokaryotes is based on the interaction between the SD sequence located upstream of the start codon on the mRNA and the complementary anti-SD sequence within the 3'-end of the 16S rRNA (Shine and Dalgarno, 1974). Although a high degree of conservation of the aSD sequence has been found in a huge variety of prokaryotic species (Nakagawa, 2009), extensive research in the field of regulation of protein synthesis has shown that this model is not always applicable. Noteworthy, the translation machinery of *E. coli*, which has been studied to gain support for the classical model, provides diverse examples for alternative translation initiation pathways, which are based on differences in the translation initiation region located on the mRNA. In this chapter the different scenarios of possible pathways leading to translation initiation are presented and the potential function of protein S1 is discussed in detail. Beside the ***type A*** mRNA (Fig.7) which is a perfect example for the "*classic*" RBS, the *E. coli* translation machinery can initiate on mRNAs where the SD sequence is masked by secondary structures (***type C***) and, moreover, on mRNAs lacking the SD sequence (***type B*** and ***type D***).

The perfect model mRNA where translation initiation is accomplished according to the pathway proposed by Shine and Dalgarno is depicted in Fig. 7A. Lacking secondary structures this class of TIRs contains an extended SD sequence to ensure interaction with the anti-SD sequence at the 3' terminus of the 16S rRNA. The SD element plays a key role in selection of the correct start codon. The duplex between SD and anti-SD (SD-duplex) is located on the platform of the 30S subunit, as shown by crystallization of a complex between the *T. thermophilus* ribosome and a short RNA fragment containing the SD sequence (Yusupova, 2006) (Kaminishi, 2007).

In *B. subtilis*, where the majority of genes contain SD elements, the stability of the SD duplex correlates with the efficiency of translation (Ma, 2002). Interestingly, in *E. coli*, formation of a stable and extended SD duplex can decrease the level of protein synthesis (Komarova, 2002). This phenomenon might be explained by a kinetic arrest of the initiation complex. The strong SD-aSD interaction might inhibit the transition from the initiation to the elongation phase. Here, one potential function of protein S1 might be to help the ribosome to proceed to the elongation phase. This hypothesis is supported by the observation that protein S1 has a **strand displacement** activity (Rajkowitsch and Schroeder, 2007), which might accelerate the reaction SD + aSD $\leftrightarrow$ SD-duplex in both directions thus making the SD-duplex kinetically labile. However, this notion is difficult to reconcile with the extended SD-duplex present in *B. subtilis* where protein S1 was shown not to be associated with the 30S subunit (Roberts and Rabinowitz, 1989). Nevertheless, *B. subtilis* still contains domain D6 (Salah, 2009) which might weakly interact with the aSD sequence like it has been shown for domain D6 in *E. coli* (Laughrea and Moore, 1978) (Dahlberg, 1975).

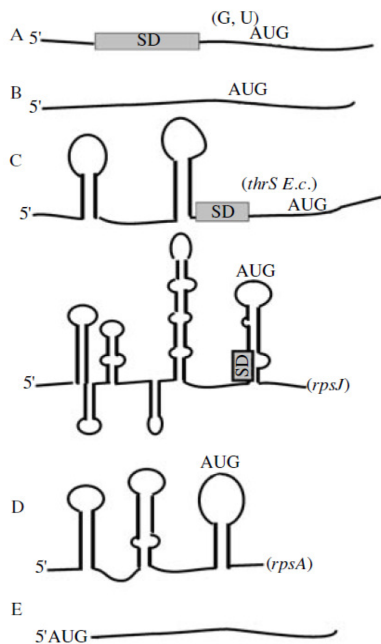


Figure 7. Different types of TIRs present in mRNAs that can be translated by the *E. coli* translation machinery. Type A represents a "*classic*" model with SD sequence and start AUG codon located on unstructured mRNA, which are easily accessible for translation initiation complex formation. Type B is lacking a SD sequence but has no secondary RNA structures around the start codon; in order to bind the 30S subunit on the correct AUG start codon it has to contain AU-rich sequences upstream of the first codon. Type C shows two examples when the SD and/or AUG start codon are hidden by RNA secondary structures. Type D reveals a curious case of TIR present on the *rpsA* mRNA encoding protein S1; despite the presence of extensive and stable secondary structures this TIR is translated with high efficiency in *E. coli* (Boni, 2001). Type E displays an example of a leaderless mRNA (lmRNA) which starts directly with a 5'-terminal start codon. Adopted from (Boni, 2006).

Recently, bioinformatics analysis of the SD-duplex stability in different prokaryotic species has been performed (Nakagawa, 2009). This study revealed a phylum dependence of the SD-duplex stability in a species. Moreover, species, where the protein encoded by the *rpsA* gene associates with the ribosome (based on classification from (Salah, 2009)), have been shown to have a tenden-

cy for a low SD-duplex stability. Thus, there is a reverse correlation between the presence of ribosome-bound protein S1 and the stability of the SD-duplex, as for example, Firmicutes possess a high SD-duplex stability. However, when they harbor S1-like proteins, these lack N-terminal domain D1 responsible for binding to the 30S ribosomal subunit (Salah, 2009).

It has been shown that only 57% of all *E. coli* mRNAs contain a SD sequence, whereas this number rises to 90% in *B. subtilis* (Ma, 2002). Toeprinting analysis revealed that for the formation of a translation initiation complex on the mRNAs lacking a SD element (Fig.7B), proteins S1 and IF3 are indispensable (Tzareva, 1994). While IF3 acts as fidelity factor, which ensures the correct facilities codon-anticodon interactions between the start codon and the initiator $tRNA_i$ in the ribosomal P site (Hartz, 1988), protein S1 was suggested to have a recognition function (Tzareva, 1994). Protein S1 was shown to bind AU-rich elements upstream of the start AUG codon in ***type B*** TIRs. Binding of protein S1 to these recognition elements has been verified by <u>sel</u>ective <u>ex</u>ponential enrichment (SELEX) experiments employing protein S1 and *E. coli* 30S ribosomes (Ringquist, 1995). These experiments revealed that depletion of *E. coli* 30S ribosomes for protein S1 leads to selection of SD-like sequences (Boni, 1991). In contrast, using 30S subunits equipped with protein S1, SELEX procedure resulted in the enrichment of AU-rich RNAs. The same RNAs were selected when free protein S1 was used in the SELEX studies (Ringquist, 1995). The function of these AU-rich elements within the TIR can be further accomplished by the fact that insertion of AU-rich elements upstream of SD element can increase the level of protein synthesis (Komarova, 2002). Thus, AU-rich sequences present in the TIR were termed **enhancers** of translation initiation (Gallie and Kado, 1989).

In Gram-negative bacteria with a high GC content the presence of RNA secondary structures within the TIR of mRNAs is very likely (Fig.7C). Moreover, in the high GC group of Gram-negative bacteria protein S1 has been shown to be essential (Sorensen, 1998). This hypothesis is supported by the finding that folded mRNAs can bind to the platform of the 30S subunit in the absence of S1 (Marzi, 2007), but they do not form a ternary complex as it has been shown by toeprinting experiments (Tedin, 1997). As the formation of secondary structure elements hinders the association of a translation initiation complex (de Smit and van Duin, 1990) (de Smit and van Duin, 1994), it is tempting to speculate that the *E. coli* protein synthesis machinery may employ protein S1 to unwind RNA secondary structures (Kolb, 1977) after the structured mRNA had been bound to the platform of the 30S subunit (de Smit and van Duin, 2003). One example for this scenario is *ompA* mRNA, which contains secondary structures within its TIR. Here, the translation initiation complex only forms in the presence of protein S1 (Tedin, 1997). Interestingly, *E. coli* protein S1 can bind to the *B. subtilis* 30S ribosome and hereby makes it tolerant to secondary structures present within TIRs (Roberts and Rabinowitz, 1989). Thus, unwinding of RNA secondary structures seems to be a universal feature of protein S1 in the process of recognition of the correct AUG start codon.

Type D represents the interesting case of TIR present in the *rpsA* mRNA encoding protein S1. Despite the lack of a SD element and the presence of secondary strucutre, this TIR is highly efficient in protein synthesis. It seems that secondary structures within the TIR are required for high efficiency in translation initiation since the destabilization of the structure elements leads to a decrease in protein synthesis efficiency of *rpsA* mRNA TIR (Boni, 2001). In addition, these

structure elements are highly conserved on the level of the secondary but not the primary structure in γ-Proteobacteria (Tchufistova, 2003). The hypothesis of an existing discontinuous SD sequence in the loops of two stems in *rpsA* TIR has been disproved employing specialized ribosomes (Skorski, 2006). However, protein S1 strongly regulates its own synthesis; the excess of protein S1 protects the AUG start codon as it was shown in DEPC modification assay (Boni, 2001). This autoregulation is physiologically, since free protein S1 recognizes the same AU-rich elements as the 30S subunit (Boni, 1991) thus competing with the ribosome for binding to TIR of mRNAs.

It is noteworthy that TIRs of ***types B-D*** cannot form translation initiation complex with the *B. subtilis* (or closely related *B. stearothermophilus*) 30S ribosomes which do not contain S1-like proteins (Roberts and Rabinowitz, 1989). It has been shown that in the absence of protein S1 prokaryotic ribosomes are intolerant of either lacking SD sequence (Vellanoweth and Rabinowitz, 1992) or the presence of RNA secondary structures within TIR (Tedin, 1997).

Type E mRNA represents leaderless mRNAs (**lmRNA**) which begin directly with the AUG start codon. In contrast to canonical leadered mRNAs, leaderless mRNAs do not require protein S1 for their translation (Tedin, 1997) (Moll, 2002). One interesting example is the CI repressor of bacteriophage λ, which is encoded by a leaderless mRNA. It has been extensively used to study the mechanism of translation initiation on lmRNAs (Grill, 2001) (Moll, 2002). First, it has been discovered that protein S1 is dispensable for translation initiation on λ*cI* lmRNA (Tedin, 1997) and, moreover, the lack of native S1 from the ribosome increases the translation efficiency of cI lmRNAs (Moll, 2002) (Komarova, 2005). Second, the λ*cI* lmRNA can be translated in the presence of the antibiotic Kasugamycin *in*

vivo, which inhibits translation initiation on canonical mRNAs (Moll, 2002) (Kaberdina, 2009). Kasugamycin can trigger the formation of protein deficient 61S ribosomes lacking several r-proteins from the 30S ribosomal subunit, especially protein S1 among the others (Kaberdina, 2009). Taken together with the essential role of S1 in translation of bulk mRNAs, the absence of S1 from the ribosome confers the selective translation of lmRNAs.

3.4. References

Agalarov, S.Ch., Kalinichenko, A.A., Kommer, A.A., and Spirin, A.S. (2006). Ribosomal protein S1 induces a conformational change of the 30S ribosomal subunit. FEBS Lett. *580*, 6797-6799.

Aliprandi, P., Sizun, C., Perez, J., Mareuil, F., Caputo, S., Leroy, J.L., Odaert, B., Laalami, S., Uzan, M., and Bontems, F. (2008). S1 ribosomal protein functions in translation initiation and ribonuclease RegB activation are mediated by similar RNA-protein interactions: an NMR and SAXS analysis. J Biol Chem. *283*, 13289-13301.

Antoun, A., Pavlov, M.Y., Andersson, K., Tenson, T., and Ehrenberg, M. (2003). The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. The EMBO journal *22*, 5593-5601.

Antoun, A., Pavlov, M.Y., Lovmar, M., and Ehrenberg, M. (2006). How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis. Molecular Cell *23*, 183-193.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science. *289*, 905-920.

Berk, V., and Cate, J.H. (2007). Insights into protein biosynthesis from structures of bacterial ribosomes. Curr Opin Struct Biol. *17*, 302-309.

Bisaglia, M., Laalami, S., Uzan, M., and Bontems, F. (2003). Activation of the RegB endoribonuclease by the S1 ribosomal protein is due to cooperation between the S1 four C-terminal modules in a substrate-dependant manner. J Biol Chem. *278*, 15261-15271.

Blumenthal, T., and Carmichael, G.G. (1979). RNA replication: function and structure of Qbeta-replicase. Annu Rev Biochem. *48*, 525-548.

Boelens, R., and Gualerzi, C.O. (2002). Structure and function of bacterial initiation factors. Curr Protein Pept Sci. *3*, 107-119.

Boni, I.V., Artamonova, V.S., Tzareva, N.V., and Dreyfus, M. (2001). Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. EMBO J. *20*, 4222-4232.

Boni, I.V., Isaeva, D.M., Musychenko, M.L., and Tzareva, N.V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res. *19*, 155-162.

Boni, I.V., Zlatkin, I.V., and Budowsky, E.I. (1981) Ribosomal protein S1 associates with Escherichia coli ribosomal 30-S subunit by means of protein-protein interactions. Eur J Biochem. *121*, 371-376.

Boni, I.V. (2006). Diverse molecular mechanisms for translation initiation in prokaryotes. Mol Biol (Mosk). *40*, 658-668.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M., and Murzin, A.G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell. *88*, 235-242.

Carpousis, A.J. (2007). The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. Annu Rev Microbiol. *61*, 71-87.

Carter, A.P., Clemons, W.M. Jr., Brodersen, D.E., Morgan-Warren, R.J., Hartsch, T., Wimberly, B.T., and Ramakrishnan, V. (2001). Crystal structure of an initiation factor bound to the 30S ribosomal subunit. Science. *291*, 498-501.

Commans, S., Plateau, P., Blanquet, S., and Dardel, F. (1995). Solution structure of the anticodon-binding domain of Escherichia coli lysyl-tRNA synthetase and studies of its interaction with tRNA(Lys). J Mol Biol. *253*, 100-113.

Company, M., Arenas, J., and Abelson, J. (1991). Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. Nature. *349*, 487-493.

Crick, F. H. (1958). On protein synthesis. Symp Soc Exp Biol *12*, 138-163.

Dahlberg, A.E. (1974). Two forms of the 30S ribosomal subunit of Escherichia coli. J Biol Chem. *249*, 7673-7678.

Dahlberg, A.E., and Dahlberg, J.E. (1975). Binding of ribosomal protein S1 of *Escherichia coli* to the 3'end of 16S rRNA. (1975). Proc Natl Acad Sci U S A. *72*, 2940-2944.

de Smit, M.H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc Natl Acad Sci U S A. *87*, 7668-7672.

de Smit, M.H., and van Duin, J. (1994). Control of translation by mRNA secondary structure in Escherichia coli. A quantitative analysis of literature data. J Mol Biol. *244*, 144-150.

de Smit, M.H., and van Duin, J. (2003). Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. J Mol Biol. *331*, 737-743.

Draper, D.E., and Reynaldo, L.P. (1999). RNA binding strategies of ribosomal proteins. Nucleic Acids Res. *27*, 381-388.

Draper, D.E., and von Hippel, P.H. (1979). Interaction of Escherichia coli ribosomal protein S1 with ribosomes. Proc Natl Acad Sci U S A. *76*, 1040-1044.

Eiler, S., Dock-Bregeon, A., Moulinier, L., Thierry, J.C., and Moras, D. (1999). Synthesis of aspartyl-tRNA(Asp) in Escherichia coli--a snapshot of the second step. EMBO J. 18, 6532-6541.

Frank, J., Gao, H., Sengupta, J., Gao, N., and Taylor, D.J. (2007). The process of mRNA tRNA translocation. Proceedings of the National Academy of Sciences of the United States of America *104*, 19671-19678.

Gabashvili, I.S., Agrawal, R.K., Spahn, C.M., Grassucci, R.A., Svergun, D.I., Frank, J., and Penczek, P. (2000). Solution structure of the E. coli 70S ribosome at 11.5 A resolution. Cell. *100*, 537-549.

Gallie, D.R., and Kado, C.I. (1989). A translational enhancer derived from tobacco mosaic virus is functionally equivalent to a Shine-Dalgarno sequence. Proc Natl Acad Sci U S A. *86*, 129-132.

Giorginis, S., and Subramanian, A.R. (1980). The major ribosome binding site of Escherichia coli ribosomal protein S1 is located in its N-terminal segment. J Mol Biol. *141*, 393-408.

Giri, L., and Subramanian, A.R. (1977) Hydrodynamic properties of protein S1 from Escherichia coli ribosome. FEBS Lett. *81*, 199-203.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schäfer, T., Kuster, B., Tschochner, H., Tollervey, D., Gavin, A.C., and Hurt, E. (2002). 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Mol Cell. *10*, 105-115.

Gribskov, M. (1992). Translational initiation factors IF-1 and eIF-2 alpha share an RNA-binding motif with prokaryotic ribosomal protein S1 and polynucleotide phosphorylase. Gene. *119*, 107-111.

Grill, S., Moll, I., Hasenöhrl, D., Gualerzi, C.O., and Bläsi, U. (2001). Modulation of ribosomal recruitment to 5'-terminal start codons by translation initiation factors IF2 and IF3. FEBS Lett. *495*, 167-171.

Gualerzi, C.O., and Pon, C.L. (1990). Initiation of mRNA translation in prokaryotes. Biochemistry. *29*, 5881-5889.

Hartz, D., McPheeters, D.S., Traut, R., and Gold, L. (1988). Extension inhibition analysis of translation initiation complexes. Methods Enzymol. *164*, 419-425.

Hartz, D., McPheeters, D.S., and Gold, L. (1990). Selection of the initiator tRNA by Escherichia coli initiation factors. Genes & development *3*, 1899-1912.

Held, W.A., Ballou, B., Mizushima, S., and Nomura, M. (1974). Assembly mapping of 30 S ribosomal proteins from Escherichia coli. Further studies. J Biol Chem. *249*, 3103-3111.

Kaberdin, V.R., and Bläsi, U. (2006). Translation initiation and the fate of bacterial mRNAs. FEMS Microbiol Rev. *30*, 967-979.

Kaberdina, A.C., Szaflarski, W., Nierhaus, K.H., and Moll, I. (2009). An unexpected type of ribosomes induced by kasugamycin: a look into ancestral times of protein synthesis? Mol Cell. *33*, 227-236.

Kaminishi, T., Wilson, D.N., Takemoto, C., Harms, J.M., Kawazoe, M., Schluenzen, F., Hanawa-Suetsugu, K., Shirouzu, M., Fucini, P., and Yokoyama, S. (2007). A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction. Structure. *15*, 289-297.

Karimi, R., Pavlov, M.Y., Buckingham, R.H., and Ehrenberg, M. (1999). Novel roles for classical factors at the interface between translation termination and initiation. Molecular cell *3*, 601-609.

Kisselev, L.L., and Buckingham, R.H. (2000). Translational termination comes of age. Trends in biochemical sciences *25*, 561-566.

Kolb, A., Hermoso, J.M., Thomas, J.O., and Szer, W. (1977). Nucleic acid helix-unwinding properties of ribosomal protein S1 and the role of S1 in mRNA binding to ribosomes. Proc Natl Acad Sci U S A. *74*, 2379-2383.

Komarova, A.V., Tchufistova, L.S., Aseev, L.V., and Boni, I.V. (2005). An Escherichia coli strain producing a leaderless mRNA from the chromosomal lac promoter. Bioorg Khim. *31*, 557-560.

Komarova, A.V., Tchufistova, L.S., Dreyfus, M., and Boni, I.V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in Escherichia coli. J Bacteriol. *187*, 1344-1349.

Komarova, A.V., Tchufistova, L.S., Supina, E.V., and Boni, I.V. (2002). Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. RNA. *8*, 1137-1147.

Labischinski, H., and Subramanian, A.R. (1979). Protein S1 from Escherichia coli ribosomes: an improved isolation procedure and shape determination by small-angle X-ray scattering. Eur J Biochem. *95*, 359-366.

Laughrea, M., and Moore, P.B. (1977). Physical properties of ribosomal protein S1 and its interaction with the 30 S ribosomal subunit of Escherichia coli. J Mol Biol. *112*, 399-421.

Laughrea, M., and Moore, P.B. (1978). On the relationship between the binding of ribosomal protein S1 to the 30 S subunit of Escherichia coli and 3' terminus of 16 S RNA. J Mol Biol. *121*, 411-430.

Laursen, B.S., Sorensen, H.P., Mortensen, K.K., and Sperling-Petersen, H.U. (2005). Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev *69*, 101-123.

Ma, J., Campbell, A., and Karlin, S. (2002). Correlations between Shine-Dalgarno sequences and gene features such as predicted expression levels and operon structures. J Bacteriol. *184*, 5733-5745.

Marzi, S., Fechter, P., Chevalier, C., Romby, P., and Geissmann, T. (2008) RNA switches regulate initiation of translation in bacteria. Biol Chem. *389*, 585-598.

Marzi, S., Myasnikov, A.G., Serganov, A., Ehresmann, C., Romby, P., Yusupov, M., and Klaholz, B.P. (2007). Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. Cell. *130*, 1019-1031.

McGinness, K.E., and Sauer, R.T. (2004). Ribosomal protein S1 binds mRNA and tmRNA similarly but plays distinct roles in translation of these molecules. Proc Natl Acad Sci U S A. *101*, 13454-13459.

Milon, P., Konevega, A.L., Gualerzi, C.O., and Rodnina, M.V. (2008). Kinetic checkpoint at a late step in translation initiation. Molecular cell *30*, 712-720.

Moll, I., Grill, S., Gründling, A., and Bläsi, U. (2002). Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. Mol Microbiol. *44*, 1387-1396.

Moll, I., Hirokawa, G., Kiel, M.C., Kaji, A., and Bläsi, U. (2004). Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. Nucleic Acids Res. *32*, 3354-3363.

Nakagawa, S., Niimura, Y., Miura, K., and Gojobori, T. (2010). Dynamic evolution of translation initiation mechanisms in prokaryotes. Proc Natl Acad Sci U S A. *107*, 6382-6387.

Osterberg, R., and Sjöberg, B. (1978). Small-angle X-ray scattering study of the proteins S1, S8, S15, S16, S20 from Escherichia coli ribosomes. FEBS Lett. *93*, 115-119.

Petrelli, D., LaTeana, A., Garofalo, C., Spurio, R., Pon, C.L., and Gualerzi, C.O. (2001). Translation initiation factor IF3: two domains, five functions, one mechanism? The EMBO Journal *20*, 4560-4569.

Pulk, A., Liiv, A., Peil, L., Maiväli, U., Nierhaus, K., and Remme, J. (2010). Ribosome reactivation by replacement of damaged proteins. Mol Microbiol. *75*, 801-814.

Rajkowitsch, L., and Schroeder, R. (2007). Dissecting RNA chaperone activity. RNA. *13*, 2053-2060.

Régnier, P., and Arraiano, C.M. (2000). Degradation of mRNA in bacteria: emergence of ubiquitous features. Bioessays. *22*, 235-244.

Ringquist, S., Jones, T., Snyder, E.E., Gibson, T., Boni, I., and Gold, L. (1995). High-affinity RNA ligands to Escherichia coli ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. Biochemistry. *34*, 3640-3648.

Roberts, M.W., and Rabinowitz, J.C. (1989). The effect of Escherichia coli ribosomal protein S1 on the translational specificity of bacterial ribosomes. J Biol Chem. *264*, 2228-2235.

Rodnina, M.V., Beringer, M., and Wintermeyer, W. (2007). How ribosomes make peptide bonds. Trends in biochemical sciences *32*, 20-26.

Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., and Bontems, F. (2009). Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. Nucleic Acids Res. *37*, 5578-88.

Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. Cell. *102*, 615-623.

Schwer, B., and Gross, C.H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. EMBO J. *17*, 2086-2094.

Sengupta, J., Agrawal, R.K., and Frank, J. (2001). Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. Proc Natl Acad Sci U S A. *98*, 11991-11996.

Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci U S A. *71*,1342-1346.

Sillers, I.Y., and Moore, P.B. (1981). Position of protein S1 in the 30S ribosomal subunit of Escherichia coli. J Mol Biol. *153*, 761-780.

Simonetti, A., Marzi, S., Jenner, L., Myasnikov, A., Romby, P., Yusupova, G., Klaholz, B.P., and Yusupov, M. (2009). A structural view of translation initiation in bacteria. Cell Mol Life Sci. 2009. *66*, 423-436.

Skorski, P., Leroy, P., Fayet, O., Dreyfus, M., and Hermann-Le Denmat, S. (2006). The highly efficient translation initiation region from the Escherichia coli rpsA gene lacks a shine-dalgarno element. J Bacteriol. *188*, 6277-6285.

Sørensen, M.A., Fricke, J., and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. J Mol Biol. *280*, 561-569.

Studer, S.M., and Joseph, S. (2006). Unfolding of mRNA secondary structure by the bacterial translation initiation complex. Mol Cell. *22*, 105-115.

Subramanian, A.R., and Mizushima, S. (1979). Characterization of a mutant form of ribosomal protein S1 from Escherichia coli. J Biol Chem. *254*, 4309-4312.

Subramanian, A.R., Rienhardt, P., Kimura, M., and Suryanarayana, T. (1981). Fragments of ribosomal protein S1 and its mutant form m1-S1. Localization of nucleic-acid-binding domain in the middle region of S1. Eur J Biochem. *119*, 245-249.

Subramanian, A.R. (1983). Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol. *28*, 101-42.

Suryanarayana, T., and Subramanian, A.R. (1979). Functional domains of Escherichia coli Ribosomal Protein S1. Formation and characterization of a fragment with ribosome-binding properties. J Mol Biol. *127*, 41-54.

Suryanarayana, T., and Subramanian, A.R. (1983). An essential function of ribosomal protein S1 in messenger ribonucleic acid translation. Biochemistry. *22*, 2715-2719.

Szer, W., Hermoso, J.M., and Boublik, M. (1976). Destabilization of the secondary structure of RNA by ribosomal protein S1 from Escherichia coli. Biochem Biophys Res Commun. *70*, 957-964.

Tchufistova, L.S., Komarova, A.V., and Boni, I.V. (2004). A key role for the mRNA leader structure in translational control of ribosomal protein S1 synthesis in gamma-proteobacteria. Nucleic Acids Res. *31*, 6996-7002.

Tedin, K., Moll, I., Grill, S., Resch, A., Graschopf, A., Gualerzi, C.O., and Bläsi, U. (1999). Translation initiation factor 3 antagonizes authentic start codon selection on leaderless mRNAs. Mol Microbiol. *31*, 67-77.

Tedin, K., Resch, A., and Bläsi, U. (1997). Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5' leader sequence. Mol Microbiol. *25*, 189-199.

Thomas, J.O., Boublik, M., Szer, W., and Subramanian, A.R. (1979). Nucleic acid binding and unfolding properties of ribosomal protein S1 and the derivatives S1-F1 and m1-S1. Eur J Biochem. *102*, 309-314.

Thomas, J.O., and Szer, W. (1982). RNA-helix-destabilizing proteins. Prog Nucleic Acid Res Mol Biol. *27*, 157-187.

Tzareva, N.V., Makhno, V.I., and Boni, I.V. (1994). Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. FEBS Lett. *337*, 189-194.

Ulbrich, B., and Nierhaus, K.H. (1975). Pools of ribosomal proteins in Escherichia coli. Studies on the exchange of proteins between pools and ribosomes. Eur J Biochem. *57*, 49-54.

van Dieijen, G., van Knippenberg, P.H., and van Duin, J. (1976) The specific role of ribosomal protein S1 in the recognition of native phage RNA. Eur J Biochem. *64*, 511-518.

van Dieijen, G., Zipori, P., van Prooijen, W., and van Duin, J. (1978). Involvement of ribosomal protein S1 in the assembly of the initiation complex. Eur J Biochem. *90*, 571-580.

Vellanoweth, R.L., and Rabinowitz, J.C. (1992). The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. Mol Microbiol. *6*,1105-1114.

Venema, J., and Tollervey, D. (1996). RRP5 is required for formation of both 18S and 5.8S rRNA in yeast. EMBO J. *15*, 5701-14.

Walleczek, J., Albrecht-Ehrlich, R., Stöffler, G., and Stöffler-Meilicke, M. (1990). Three-dimensional localization of the NH2- and carboxyl-terminal domain of ribosomal protein S1 on the surface of the 30 S subunit from Escherichia coli. J Biol Chem. *265*, 11338-11344.

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

Wintermeyer, W., and Gualerzi, C. (1983). Effect of Escherichia coli initiation factors on the kinetics of N-Acphe-tRNAPhe binding to 30S ribosomal subunits. A fluorescence stopped flow study. Biochemistry *22*, 690-694.

Yusupova, G., Jenner, L., Rees, B., Moras, D., and Yusupov, M. (2006). Structural basis for messenger RNA movement on the ribosome. Nature. *444*, 391-394.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H., and Noller, H.F. (2001). The path of messenger RNA through the ribosome. Cell. *106*, 233-241.

Zavialov, A.V., Mora, L., Buckingham, R.H., and Ehrenberg, M. (2002). Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Molecular cell *10*, 789-798.

4. Aims of the study

The assembly of ribosomal protein S1 to the 30S ribosomal subunit is mediated by protein-protein interaction with protein S2. Since protein S1 is essential for translation initiation in Gram-negative bacteria, the interaction surface between the two ribosomal proteins S1 and S2 may represent a potential target for antimicrobials that could be semi-selective against opportunistic pathogens such as *P. aeruginosa*, *S. flexneri*, etc. without affecting Gram-positive bacteria with low GC content, as there protein S1 is not required for viability. Moreover, mitochondria and eukarya do not employ a homologue of protein S1 for translation initiation. Therefore, the aim of this study is to characterize the interaction between proteins S1 and S2 and to determine the important structural elements participating in this interaction. In addition, this study is anticipated to shed light on the structure of the ribosomal protein S1 and to contribute to the elucidation of the function of the essential protein S1 in translation initiation in *Escherichia coli*.

5. Manuscripts

Direct interaction between the N-terminal domain of protein S1 and protein S2 on the *Escherichia coli* ribosome

Konstantin Byrgazov[1], Salim Manoharadas[1], Anna Chao Kaberdina[1],

and Isabella Moll[1‡]

[1]Max F. Perutz Laboratories, Center for Molecular Biology,
Department of Microbiology, Immunobiology and Genetics, University of Vienna,
Dr. Bohrgasse 9/4, 1030 Vienna, Austria

[‡] Corresponding author:

Isabella Moll
Phone: +43-4277-54606
Fax: +43-1-4277-9546
e-mail: Isabella.Moll@univie.ac.at

Summary

The structure of the *E. coli* ribosome is solved at atomic resolution. In contrast, hitherto the position and structure of the essential ribosomal protein S1 has not been determined due to its intrinsic flexibility. Previously, it has been shown that protein S1 utilizes its N-terminal domains to bind to the ribosome *via* protein-protein interaction. Moreover, protein S2 has been shown to be required for binding of protein S1 to the ribosome. Here, we show that the N-terminal domain of S1 (amino acids 1-106; $S1_{106}$) is essential for its interaction with protein S2 as well as with the ribosome. Moreover, over production of protein $S1_{106}$ hinders *E. coli* growth by displacing native protein S1 from its binding pocket on the ribosome. In addition, we verify that the coiled-coil domain of protein S2 ($S2\alpha_2$) is sufficient to allow protein S1 to bind to the ribosome.

Taken together, these data reveal the crucial elements required for the S1/S2 interaction, which is pivotal for translation initiation in Gram-negative bacteria. Thus, the interaction surface between proteins S1 and S2 could be considered as a potential target for antimicrobial compounds acting semi-selectively against Gram-negative bacteria.

Introduction

A pivotal step in regulation of gene expression is the initiation of translation, more precisely, the initial interaction between the ribosome and the mRNA (Gualerzi and Pon, 1990). In *Escherichia coli* and most Gram-negative bacteria ribosomal protein S1 is a key player that mediates the primary binding of the 30S ribosomal subunit to the ribosome binding site (rbs) on the mRNA (Sorensen *et al.*, 1998). S1 represents the largest ribosomal protein with a molecular weight of 61.1 kDa. In particular, it is implicated in translation initiation complex formation on mRNAs comprising highly structured 5´-untranslated regions (UTR; Szer *et al.*, 1975; van Dieijen *et al.*., 1976). The protein interacts with a pyrimidine-rich region upstream of the Shine and Dalgarno (SD)-sequence and was suggested to unwind RNA secondary structures (Bear *et al.*., 1976; Thomas *et al.*., 1978; Rajkowitsch and Schroeder, 2007), thereby facilitating positioning of the 30S subunit in close proximity to the translational start site (de Smit and van Duin, 1994). In contrast, S1 is dispensable for translation of leaderless mRNAs (lmRNAs) that start directly with the AUG codon thus lacking a 5´-UTR (Tedin *et al.*., 1997; Moll *et al.*., 2002a).

S1 is composed of six contiguous OB (oligonucleotide–oligosaccharide-binding) folds, the 'so-called' S1 domains, which are approximately 70 amino acids in size (Figure 1; Bycroft *et al.*, 1997). Although structurally related, these domains exhibit distinct functions: the two N-terminal domains (D1-D2) are suggested to be involved in ribosome binding and interaction with the Qβ replicase (McGinness and Sauer, 2004; Subramanian *et. al.*, 1981; Guerrier-Takada *et al.*, 1983), whereas domains D3-D5 can bind ssRNA and RNA pseudoknots (Figure 1; Boni *et al.*, 1991; Subramanian, 1984; Aliprandi *et al.*, 2008; Salah *et al.*,

2009). The most distal domain (R4; Figure 1) is involved in autogenous regulation of *rpsA* (Boni *et al.*, 2001). Recently, the functional specialization of the different domains has been supported by phylogenic trees built from the alignment of domain sequences of the Gram-negative S1 proteins (Salah *et al.*, 2009).

Despite its essentiality in the process of translation, to date there is no structure of the native protein S1, and moreover the protein is missing in the high resolution structures available for the *E. coli* ribosome. However, a tentative position has been determined by comparing cryo-electron data of the 30S ribosomal subunit of *E. coli* with an X-ray crystallographic structure of a 30S subunit of *T. thermophilus* lacking S1 (Sengupta *et al.*, 2001). The data underline the results obtained by cross linking and immuno-precipitation studies, which suggest a direct interaction between proteins S1 and S2 (Laughrea and Moore, 1978; Bollen *et al.*., 1979; Aseev *et al.*., 2008). Moreover, our observation that *E. coli* ribosomes lacking protein S2 are likewise devoid of protein S1 (Moll *et al.,* 2002b) indicates that protein S2 is essential for binding of S1 to the 30S ribosomal subunit. In addition, the formation of a stoichiometric complex of proteins S1 and S2 was reported (Sukhodolets and Garges, 2004), which is implicated in the regulation of the expression of the *rpsB-tsf* operon encoding ribosomal protein S2 and translation elongation factor EF-Ts (Aseev *et al.*., 2008).

This study was performed to gain insights into the binding mode of protein S1 to the ribosome. With the objective to determine structural features required for assembly of the protein on the ribosome, we determined the binding capacity of different truncated protein S1 variants. Our results indicate that solely the N-terminal domain D1 is responsible and required for interaction of S1 with the ribosome. Furthermore, our data indicate that overexpression of the $S1_{106}$ protein,

representing the N-terminal S1 domain, inhibits translation of bulk mRNA, whereas translation of lmRNAs continues. Moreover, we verify the direct interaction between domain D1 and ribosomal protein S2, which is pivotal for binding of protein S1 to the ribosome. As preventing of this binding causes severe affects *E. coli* viability, the interaction surface between proteins S1 and S2 may represent a novel target for antimicrobials which are semi-selective against Gram-negative bacteria.

Results

The N-terminal domain D1 of protein S1 is required for binding to the ribosome *in vivo*

Previous studies indicate that the N-terminal fragment of protein S1 comprising domains D1 and D2 (protein $S1_{194}$, Figure 1) is pivotal for ribosome binding (McGuiness and Sauer, 2004; Subramanian, 1984, Sillers and Moore 1981). However, based on the information of a phylogenic tree built on alignments of S1 protein sequences form Gram-negative bacteria, domains D1 and D2 are suggested to have different roles in ribosome binding (Salah *et al.*, 2009). Therefore, the first aim of this study was to narrow down the point of interaction between S1 and the ribosome. To distinguish, whether domain(s) D1 and/or D2 are required for ribosome binding, FLAG-tagged S1 variants comprising either domain D1 ($S1_{106}$), domain D2 ($S1_{87-194}$), or both domains D1-D2 ($S1_{194}$) were over-expressed *in vivo*. *E. coli* strain JE28 (Ederth *et al.*, 2008) harbouring plasmids pProS1D1F, pProS1D2F, or pProS1D12F (Table 1) coding for the respective S1 fragments under control of the *trc* promoter was grown in LB broth at 37°C. At OD_{600} of 0.2 synthesis of S1 variants was induced by addition of 50 µM IPTG. As shown in Figure 2A, over-expression of protein $S1_{194}$ severely effects growth. This effect can be attributed to the inhibition of protein synthesis as the binding of native S1 is prevented by the ribosome bound $S1_{194}$ variant (McGuiness and Sauer, 2004; Subramanian, 1984). This effect is mirrored by over-expression of protein $S1_{106}$ (comprising domain D1) as cell growth is inhibited in a comparable manner. In contrast, over-expression of domain D2 ($S1_{87-194}$) did not affect growth, already indicating that domain D2 is not involved in ribosome binding.

To support the assumption that domain D1 is essential for protein S1 to interact with the ribosome, the cells were harvested 60 min upon induction and ribosomes were prepared. As *E. coli* strain JE28 harbours a modified *rpsL* gene encoding a His-tagged protein L7/L12 (Ederth *et al.*, 2008), 70S ribosomes were purified employing Ni-NTA agarose as specified in Materials and Methods. Upon separation of ribosomal proteins on SDS-PAGE, western blot analysis employing anti-FLAG antibodies revealed the presence of proteins $S1_{194}$ (Figure 2B, panel b, lane 8) and $S1_{106}$ (Figure 2B, panel b, lane 6) on the ribosome *in vivo*. As expected, this binding severely reduces the amount of native protein S1 present on the ribosome (Figure 2B, panel a, lanes 6 and 8). In contrast, protein $S1_{87\text{-}194}$ comprising domain D2 cannot be detected in the 70S fraction (Figure 2B, panel b, lane 4).

Protein S1 lacking the N-terminal domain D1 does not bind to the ribosome *in vivo*

To verify that only domain D1 is involved in interaction with the ribosome, the affinity of a truncated variant of S1 lacking the N-terminal D1 domain ($S1_{87\text{-}557}$, Figure 1) was tested *in vivo*. Upon over-expression of the C-terminally FLAG-tagged native S1 protein or the $S1_{87\text{-}557}$ variant in *E. coli* strain JE28, ribosomes were separated from the S100 extract. The presence of native S1 and its protein variant on 70S ribosomes was determined by western blot analysis. The result shown in Figure 3 reveals that in contrast to the native S1 (Figure 2C, panel a, lanes 1 and 2), protein $S1_{87\text{-}557}$ does not interact with the ribosome, as it can be detected solely in the ribosome free S100 fraction (Figure 2C, panel a, lanes 3 and 4). This result supports the notion that interaction with the ribosome occurs within the first 86 amino acid residues of protein S1.

Domain D1 inhibits translation of canonical mRNA but does not affect lmRNA translation

Taken together these results indicate that domain D1 interacts with the ribosome and subsequently prevents binding of native protein S1. As S1 is essential for translation initiation on canonical mRNAs (Sorensen *et al.*, 1998) we rationalized that overexpression of Domain 1 might inhibit translation of canonical mRNAs. In contrast, translation of lmRNA is accomplished in the absence of protein S1 (Tedin, *et al.*, 1997; Moll *et al.*, 2002). We thus asked whether over-expression of protein $S1_{106}$ could render the translational apparatus selective for lmRNAs. To answer this question, translation upon over-expression of proteins $S1_{106}$, $S1_{87-194}$ and $S1_{194}$ was monitored *in vivo* by pulse labelling. Briefly, *E. coli* strain JE28 harbouring plasmid pKT35-cI-lacZ (encoding the leaderless *cI-lacZ* fusion gene; Tedin *et al.*, 1997) and either plasmid pProF-S1D12F, pProF-S1D1F or pProF-S1D2F (encoding proteins $S1_{194}$, $S1_{106}$, and $S1_{87-194}$; Table 1), respectively, was grown in M9 minimal medium and pulse labelling was performed before and 15, 30, and 60 minutes upon addition of IPTG as specified in Materials and Methods. As shown in Figure 4, upon over-expression of protein $S1_{87-194}$ lacking domain 1, translation of bulk mRNA was not affected (lanes 5-8). However, upon synthesis of proteins $S1_{106}$ and $S1_{194}$ translation of bulk mRNA ceased, whereas translation of the leaderless *c*I-*lacZ* mRNA continued (lanes 2-4 and lanes 10-11). To ensure overexpression of proteins $S1_{106}$ and $S1_{194}$ under these conditions the respective mRNAs only contain a short leader of 7 nucleotides in length, translation of which likewise does not require protein S1 (Tedin *et al.*, 1997; Figure 4, indicated by stars).

Protein $S1_{106}$ interacts with protein S2

More than 30 years ago, it has been suggested that protein S1 associates with the 30S ribosomal subunit by means of protein-protein interaction (Boni *et al.*, 1981). This notion has been supported by biochemical studies and cross-linking experiments indicating that protein S1 is located in spatial proximity to proteins S2, S10, and S18 (Laughrea and Moore, 1978; Boileau *et al.*, 1981). These results are in good agreement with our observation that *E. coli* ribosomes, which lack ribosomal protein S2 are likewise devoid of protein S1 (Moll *et al.,* 2002). To scrutinize, whether the proximal domain D1 of protein S1 directly interacts with protein S2 we performed a pull down assay employing tagged protein S1 variants. Briefly, *E. coli* strain Tuner harbouring plasmid pProEx-D12F, pProEx-D1F or pProEx-D2F (encoding proteins $S1_{194}$, $S1_{106}$, and $S1_{87\text{-}194}$ containing an N-terminal His-Tag and and C-terminal Flag-tag; Table 1), respectively, was grown in LB medium. Upon overexpression of the protein S1 variants, S30 extracts were prepared and loaded onto a Ni-NTA-agarose column to allow binding of the tagged proteins $S1_{106}$, $S1_{87\text{-}194}$ and $S1_{194}$. After vigorous washing, the proteins bound to the column were eluted and tested for co-purification of protein S2 by Western blot analysis. As shown in Figure 5, concomitantly with the elution of proteins $S1_{106}$ and $S1_{194}$ (panel c, lanes 4 and 8) we obtained a significant amount of endogenous protein S2 (panel c, lanes 4 and 8). In contrast, we did not observe co-purification of protein S2 when protein $S1_{87\text{-}194}$ was bound to the Ni-NTA-matrix (Figure 5, panel b, lane 6), which lacks the N-terminal D1 domain. Taken together, these data suggest that solely domain D1 is involved in direct interaction with protein S2. Moreover, these results were supported employing a

yeast two hybrid approach, which likewise indicates the interaction between protein S1 or its variants, $S1_{106}$ and $S1_{194}$, and S2 (Figure S1, a-e).

The coiled-coil domain of protein S2 is sufficient to allow assembly of protein S1 to the ribosome

During the analysis of the crystal structure of the 30S ribosomal subunit the structure of ribosomal protein S2 has been determined (Brodersen *et al.*, 2002). The protein is located at the solvent side of the 30S subunit at the hinge region between the head and the body of the particle (Brodersen et al., 2002). As shown in Figure 6, the protein consists of a large globular domain (indicated in green) and a protruding coiled-coil domain spanning amino acids 110-150 ($S2\alpha_2$; indicated in red), which are connected by an unstructured neck region. The globular domain of the protein is functionally implicated in the accommodation and stabilization of the SD-aSD duplex in the post-initiation complex (Yusupova *et al.*, 2006), whereas the side of the coiled-coil protrusion $S2\alpha_2$ mediates the interaction with helices 35 and 37 of the 16S rRNA (Brodersen *et al.*, 2002).

Considering the proposed localization of protein S1 on the 30S ribosomal subunit by Sengupta et al. (2001), which indicates that the long arm of protein S1 (LA), representing the N-terminal domain, lies in close proximity to the $S2\alpha_2$ domain, we next tested the direct interaction between these domains as specified in Materials and Methods. Briefly, S30 extracts prepared from *E. coli* Tuner cells overproducing either the his-tagged $S2\alpha_2$ domain, the flag-tagged protein S1 or the flag-tagged protein $S1_{87\text{-}557}$ were mixed and incubated with Ni-NTA-agarose to allow binding of the $S2\alpha_2$ domain. After several washing steps to remove unspecifically bound proteins, protein $S2\alpha_2$ and its potential binding partners were

eluted by addition of imidazol. Western blot analysis of the elution fractions revealed that only full length protein S1 co-purifies with protein S2α_2 (Figure 6B, panel a, lanes 2 and 4). In contrast, we did not detect protein S1$_{87-557,}$ lacking the N-terminal domain D1 upon elution of S2α_2 (Figure 6B, panel a, lanes 3 and 5). Taken together, this analysis strongly supports the notion that the D1 of protein S1 is required for direct interaction with the coiled coil domain of protein S2. In addition, the yeast-two hybrid system mentioned above likewise revealed the direct interaction between proteins S1$_{106}$ and S2α_2 (Figure S1, f and g).

Considering that most interactions between S2 and the 16S rRNA are formed *via* the coiled-coil domain (Brodersen *et al.*, 2002; Yusupova *et al.*, 2006), we anticipated that overexpression of the S2α_2 domain could outcompete native protein S2 from the ribosome. However, taking the interaction between S1 and the S2α_2 domain into account, it seemed conceivable that binding of S2α_2 would not interfere with assembly of protein S1 to the 30S subunit, as it could provide the platform for S1 binding. In order to test for this hypothesis, *E. coli* strain Tuner harbouring plasmid pET-ccS2, encoding the S2α_2 domain (Table 1) was grown in LB broth. Ribosomes were purified by sucrose density gradient centrifugation as specified in Materials and Methods, before (time point 0) and 30, 60, and 90 minutes upon addition of IPTG to induce synthesis of the coiled-coil domain of protein S2. The presence of native S1 and S2 proteins as well as of the S2α_2 domain on crude ribosomes was determined employing antibodies directed against proteins S1 and S2. As shown in Figure 6C, upon induction of S2α_2 synthesis, we were able to verify binding of the S2α_2 domain to the ribosome (panel c, lanes 2-4). Concomitantly, the amount of native protein S2 bound to the ribosome is severely reduced (panel b, lanes 2-4). However, as expected the amount of protein

S1 did not alter during the course of the experiment (panel a, lanes 1-4). Protein S10 served as an internal control to verify that the same amount of ribosomes was applied.

Discussion

In Gram-negative bacteria protein S1 is an essential mediator in translation initiation (Sorensen *et al.*, 1998). Its role is thought to unwind secondary structures within translation initiation regions (TIRs) of mRNAs in order to facilitate recognition of the correct start codon (Bear *et al.*, 1976; Thomas and Szer, 1982). Although the structure of the *E. coli* ribosome has been already solved at atomic resolution (Schuwirth *et al.*, 2005), the precise position of protein S1 on the ribosome still has to be elucidated. Previously, the ribosome binding site of protein S1 has been shown to be located within its N-terminal domains (McGiness and Sauer, 2004; Subramanian, 1984). Here, we showed that the N-terminal domain D1 (also referred as protein $S1_{106}$) plays an important role in binding of protein S1 to the 30S ribosomal subunit. The deletion of domain D1 prevents interaction of protein S1 with the ribosome since protein $S1_{87\text{-}557}$ does not interact with the ribosome *in vivo* (Figure 2). Being over produced, protein $S1_{106}$ is toxic for *E. coli* comparable to over expression of the gene encoding the two N-terminal domains ($S1_{194}$; McGinness and Sauer, 2004). Moreover, we were able to verify binding of protein $S1_{106}$ to crude ribosomes upon over expression. Consequently, this binding inhibits assembly of endogenous protein S1 to the ribosome (Figure 4). Therefore, it could be assumed that the toxicity of synthesis of protein $S1_{106}$ could be attributed to the accommodation of the truncated protein $S1_{106}$ on the binding pocket on the 30S subunit thereby inhibiting assembly of native protein S1. Taken together, our data suggest that the N-terminal domain of S1 ($S1_{106}$) is sufficient for the interaction with the ribosome. This observation is in agreement with the fact that this domain is absent from protein S1 sequences of Gram-positive bacteria with low GC content where protein S1 does not bind to

the 30S subunit and thus is not a true component of the ribosome (Salah *et al.*, 2009) (Vellanoweth and Rabinowitz, 1992).

Previously the possible role for protein S1 in translation elongation was proposed (Potapov and Subramamanian, 1992). However, the depletion of the ribosome by anti-S1-serum does not affect translation elongation (van Dieijen *et al.*, 1978). In addition, leaderless mRNAs do not require protein S1 for their translation indicating that protein S1 is dispensable for translation elongation (Moll *et al.*, 2002). This result is supported by current work, as we were able to show that leaderless mRNA is still translated upon induction of protein $S1_{106}$ synthesis. It implies that the elongation process is not affected by replacement native protein S1 by its truncated variant.

Protein S1 is associated with the ribosome by means of protein-protein interactions (Boni *et al.*, 1981) and protein S2 is required for its binding to the ribosome (Moll *et al.*, 2002). The interaction between proteins S1 and S2 on the ribosome was proposed from cross-linking studies (Laughrea and Moore, 1979). Recently, complex between proteins S1 and S2 was proposed to participate in autogenous control of the *rpsB-tsf* operon encoding protein S2 and elongation factor EF-Ts (Aseev *et al.*, 2008). As shown in this work, protein S1 requires its first domain to bind to the ribosome. To test for the protein-protein interaction between protein $S1_{106}$, representing the first S1 domain, and protein S2. We employed pull down assays where we over expressed his-tagged proteins $S1_{106}$ and $S1_{87\text{-}194}$ followed by precipitation on Ni-NTA agarose. Western blot analysis revealed that protein S2 eluted along with both proteins but in significantly different amount. While fractions of protein $S1_{87\text{-}194}$ do not contain detectable amount of protein S2, protein $S1_{106}$ binds significant amount of endogenous

protein S2. This indicates a strong interaction between these two proteins which was recapitulated employing a yeast-two hybrid system.

Protein S2 is organized in bi-domain form consisting of the globular and coiled-coil domains. We verified that the coiled-coil domain binds to the ribosome and, moreover, is sufficient to allow protein S1 to bind to the ribosome. Furthermore, we showed the direct interaction between the coiled-coil domain of protein S2 ($S2\alpha_2$) and protein S1, which was supported employing yeast two hybrid system. In addition, we were able to show that protein S1 requires its N-terminal domain D1 for interaction with protein $S2\alpha_2$.

The interaction surface will be further characterized in order to determine the residues involved in the interaction between two proteins which give a powerful tool against bacterial translation initiation on mRNAs bearing secondary structures within TIRs. Hence, further characterization of proteins S1/S2 interaction surface will aim to design the novel antimicrobial compounds acting against Gram-negative bacteria (such as opportunistic pathogens *P. aeruginosa*, *S. enterica*, etc.) where assembly of protein S1 on the ribosome is essential for bacterial viability.

Matherials and Methods

Table 1. Bacterial strains and plasmids used in this study.

	Relevant features	Source or reference
E. coli strains:		
JE28	MG1655::*rplL-his*	Ederth et al., 2009
Tuner	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1*	Novagen
Tuner(DE3)	F^- *ompT hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1(DE3)*	Novagen
Plasmids:		
pKT35-cILacZ	cI-lacZ fusion under Plac protmoter	Tedin et al., 1996
pProEX-HTb	vector for Trc driven gene expression	Invitrogen
pProEX-S1F	pProEX derivative encoding flagged S1	this study
pProEX-S1ΔD1F	pProEX derivative encoding flagged $S1_{87\text{-}557}$	this study
pProEX-S1D12F	pProEX derivative encoding flagged $S1_{194}$	this study
pProEX-S1D1F	pProEX derivative encoding flagged $S1_{106}$	this study
pProEX-S1D2F	pProEX derivative encoding flagged $S1_{87\text{-}194}$	this study
pPro-S1F	pProEX-S1F without his-tag	this study
pPro-S1ΔD1F	pProEX-S1ΔD1F without his-tag	this study
pPro-S1D12F	pProEX-S1D12F without his-tag	this study
pPro-S1D1F	pProEX-S1D1F without his-tag	this study
pPro-S1D2F	pProEX-S1D2F without his-tag	this study
pET28b	vector for T7 driven over expression	Novagen
pET-S1D1D2	pET derivative encoding for his-tagged $S1_{194}$	this study
pET-S1D1	pET derivative encoding for his-tagged $S1_{106}$	this study
pET-S1D2	pET derivative encoding for his-tagged $S1_{87\text{-}194}$	this study

Table 2. Synthetic oligonucleotides used in this study

Oligonucleotide (Purpose)

J3 TATA*CCATGG*AATCTTTTGCTCAACTC (NcoI, rpsA from the 1st codon)
K3 TATA*CCATGGA*GAAAGCTAAACGTC (NcoI, rpsA from the 87st codon)
L3 TATA*GAATTC*AGCATCTTCGTAAGC (EcoRI, rpsA until the 106st codon)
M3 TATA*GAATTC*CATGCCTTCCTGCAGG (EcoRI, rpsA until the 194st codon)
B5 TATA*GGCGCCGAATTC*ATGACTGAATCTTTTGCTC (NarI, EcoRI, rpsA from the 1st codon)
D5 TATA*GGCGCCGAATTC*ATGAAAGCTAAACGTCAC (NarI, EcoRI, rpsA from the 87th codon)
G5 TATA*CTCGAG***TTA**TTTTTCATCGTCATCCTTATAGTCAGCATCTTCGTAAGC (XhoI, flag-tag, rpsA until 106th codon)
H5 TATA*CTCGAG***TTA**TTTTTCATCGTCATCCTTATAGTCCATGCCTTCCTGCAGG (XhoI, flag-tag, rpsA until 194th codon)
I5 TATA*CTCGAG***TTA**TTTTTCATCGTCATCCTTATAGTCGCCTTTAGCTGCTTTG (XhoI, flag-tag, rpsA until 557th codon)
J5 TATA*GAATTCCTCGAG*GGTCTGTTTCCTGTG (primer for site-directed mutagenesis)

Bacterial strains and plasmids

E. coli strains, plasmids and oligonucleotides used in this study are listed in tables 1 and 2. Unless otherwise indicate, bacterial cultures were routinely grown in LB medium (Miller, 1972) supplemented with the antibiotics ampicillin (100µg/ml) or kanamycin (40µg/ml). Growth was monitored by measuring the optical density at 600 nm (OD_{600}).

Construction of plasmids

Coding sequences of protein S1 and its variants were amplified and cloned under control of T7 RNA polymerase promoter between *NcoI* and *EcoRI* sites of pET28b or under control of Trc promoter between *NarI* and *XhoI* sites of pProEX-HTb (Invitrogen). The pProEX-HTb derivatives were then amplified by using primer J5 and corresponding forward primer, the resulting PCR products were digested with EcoRI and DpnI, ligated by T4 DNA ligase. This procedure resulted in pProEX-HTb derivatives lacking the sequence encoding for N-terminal His-tag followed by TEV-cleavage site (called pProF backbone). All the sequences were verified by sequencing (AGOWA).

The ribosome purification

JE28 cells transformed with pProF plasmids were grown overnight in LB broth plus 100µg/ml of ampicillin and 20µg/ml of kanamycin, diluted 1:100 into fresh medium, grown to OD_{600} 0.30-0.35 and induced with 100µM of IPTG. After 1h upon induction cells were harvested by centrifugation and lysed by freezing-thawing method in lysis buffer containing 20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 100mM KCl, 10mM Imidazole, 1u/mL RNase-free Dnase I (Roche). After centrifugation, the resultant S30 extracts were applied to Ni-NTA agarose, washed by 10 column volumes of washing buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 20mM Imidazole) followed by elution with elution buffer containing 20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 150mM Imidazole. The protein composition of the ribosomes was estimated by running the same number of A_{260} units on SDS-PAGE followed by Western blot analysis using anti-ECS (Abcam) and anti-L2 antibodies.

De novo cI-*lacZ* synthesis upon protein S1 variants over expression

JE28 strain cells containing pKT35-cIlacZ plasmid along with pProEX-HTb, pProF-S1D12F, pProF-S1D1F and pProF-S1D2F vectors were grown in M9 minimal media. At an OD_{600} of 0.3 protein over expression was induced by 100µM of IPTG. A total of 15, 30, 45 and 60 min after IPTG addition, 1 ml aliquots were withdrawn from both cultures. At each time time pulse, labelling was carried out by addition of 1µL [^{35}S]-methionine (10µCi/mL), and by further incubation for 5 min at 37°C. The labelling was stopped by addition of 1mg/ml of "cold" methionine and further incubation for 1 min at 37°C. The reactions were stopped by addition of cold 5% TCA, followed by incubation in ice for 15 min and subsequent centrifugation for 15 min at 15000 g at 4°C. The cell pellets were washed once with 90% acetone dried under vacuum for 5 min, resuspended in SDS-protein sample buffer, and boiled for 5 min prior to loading onto 12% SDS-polyacrylamide gel. For the different OD_{600} values, the same amounts of total cellular protein were subjected to electrophoresis. The gels were dried and exposed to a Molecular Dynamics PhosphoImager for visualization and quantification.

Co-immunoprecipitation

Tuner cells containing plasmids pProEX-HTb, pProEX-S1D12F, pProEX-S1D1F and pProEX-S1D2F were grown overnight in LB broth containing 40µg/ml of kanamycin, diluted 1:100 into fresh medium, grown to OD_{600} 0.30-0.35 and induced with 100µM of IPTG. After 1h upon induction cells were harvested by centrifugation and lysed by freezing-thawing method in lysis buffer containing 50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 10mM Imidazole, 0.1% Tween-20, 0.5 mg/mL Dnase I (Roche), 20 µg/ml RNase A. After centrifugation, the resultant S30 extracts were applied to Ni-NTA agarose, washed by 10 column volumes of washing buffer (50mM Na_2HPO_4, pH 8.0, 500mM NaCl, 20mM Imidazole) followed by elution with elution buffer (50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 250mM Imidazole). Protein concentration was estimated by Bradford assay. The protein composition of the eluted fractions was estimated by running the same number of total protein amount on SDS-PAGE followed by Western blot analysis using anti-ECS (Bethyl) and anti-S2 antibodies

Protein purification

Tuner (DE3) cells containing plasmids pETS1D1 and pETS1D2 were grown overnight in LB broth containing 40µg/ml of kanamycin, diluted 1:100 into fresh medium, grown to OD_{600} 0.70-0.85 and induced with 1 mM of IPTG. After 1h upon induction cells were harvested by centrifugation and lysed by freezing-thawing method in lysis buffer containing 50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 10mM Imidazole, 0.1% Tween-20, 0.5 mg/mL Dnase I (Roche), 20 µg/ml RNase A. After centrifugation, the resultant S30 extracts were applied to Ni-NTA agarose, washed by 10 column volumes of washing buffer (50mM Na_2HPO_4, pH 8.0, 500mM NaCl, 20mM Imidazole) followed by elution with elution buffer (50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 250mM Imidazole). Protein concentration was estimated by Bradford assay. The protein composition of the eluted fractions was estimated by running the same number of total protein amount on SDS-PAGE. The eluted fractions were dialyzed into running buffer (50mM Na_2HPO_4, pH 8.0, 200mM NaCl) and loaded on HiLoad Sephadex 75 16/60 column (GE Healthcare). The size exclusion FPLC was done in running buffer at 4°C. The peak corresponding to pure proteins $S1_{106}$ and $S1_{87\text{-}194}$ were collected, concentrated and subjected to Bradford analysis for estimation of protein concentration.

Figure Legends

Figure 1. Schematic depiction of variants of protein S1 used in this study.
Protein S1 and its variants used in this study are C-terminally tagged to avoid interference with ribosome binding occurring *via* N-terminus of the protein. The deletion of the first S1 domain from proteins S1 and $S1_{194}$ result in proteins $S1_{87-557}$ and $S1_{87-194}$, respectively, which are tested in this studies for their binding to the ribosome. Protein $S1_{106}$ represents the N-terminal domain of protein S1.

Figure 2. Induced synthesis of protein $S1_{106}$ affects *E. coli* growth by displacing native protein S1 from the ribosome.
(A) Effect of over expression of N-terminal variants of protein S1 on bacterial growth. *E.coli* JE28 strain harboring plasmids pProEX-HTb (-♦-), pProF-S1D1(-▲-), pProF-S1D2 (-■-) and pProF-S1D12 (-●-) were grown in LB medium containing ampicilin (100µg/ml) and kanamycin (20µg/ml). At OD_{600} 0.2 (marked by ▼) 100µM IPTG was added to the cultures. 50ml from each culture was harvested for the 70S-His ribosome preparation after 1 hour upon induction (marked by Δ) and 50 ml continued to grow to make the growth curve. (B) Purification of his-tagged ribosomes from *E.coli* strain JE28 over expressing FLAG-tagged proteins $S1_{106}$, $S1_{87-194}$ and $S1_{194}$. The 70S-His ribosomes were purified by applying S30 extracts (lanes 1, 3, 5 and 7) to Ni-NTA-agarose beads followed by washing and elution as described in Materials and Methods. The protein composition of purified 70S-His ribosomes was checked by Western blot analysis using anti-S1 antibodies (upper panel), anti-FLAG antibodies (middle panel) and anti-S2 (lower panel). The latter were used as loading control. (C) Purification of his-tagged ribosomes from *E.coli* strain JE28 over expressing flag-tagged proteins S1 and $S1_{87-557}$. The S30 extracts were incubated with Ni-NTA in lysis buffer as described in Matherials and Methods. S100 extracts (lanes 2 and 4) contain the proteins unbound to the ribosome. 70S-His were eluted by increasing concentration of imidazole. The same amount of A_{260} units was loaded on SDS-PAGE followed by Western blot analysis. Protein L2 is used as a loading control for the ribosome.

Figure 3. **Over production of ribosome-binding variants of protein S1 blocks translation of bulk mRNAs**.

The autoradiogram of SDS-PAGE shows the effect of over expression of proteins $S1_{106}$, $S1_{87\text{-}194}$ and $S1_{194}$ (marked by *) on on-going protein synthesis. Pulse labeling experiment reveals that over production of proteins $S1_{106}$ and $S1_{194}$ but not $S1_{87\text{-}194}$ results in leaderless mRNA translation and overall protein synthesis inhibition. It can be explain by malfunction of the ribosomes with protein S1 displaced by its short N-terminal variants which leads to inhibition of translation of canonical mRNA. However, protein S1 is dispensable for translation of lmRNA. Thus the ribosomes become selective for translation of leaderless mRNA as it is seen from the increasing of the leaderless *cI-lacZ* mRNA translation (marked by **).

Figure 4. **Ribosome-binding variants of protein S1 directly interact with protein S2**.

S30 extracts form *E. coli* strain Tuner transformed with empty vector pProEX-HTb and its derivatives encoding for N-terminally his-tagged and C-terminally flag-tagged proteins $S1_{106}$, $S1_{87\text{-}194}$ and $S1_{194}$ were loaded onto Ni-NTA-agarose. After the washing step protein bound to the matrix were eluted. Both input (S30 extract) and elution fractions were assayed for the presence of protein S2 by SDS-PAGE followed by Western blot analysis.

Figure 5. **Interaction between the coiled-coil domain of protein S2 and protein S1**.

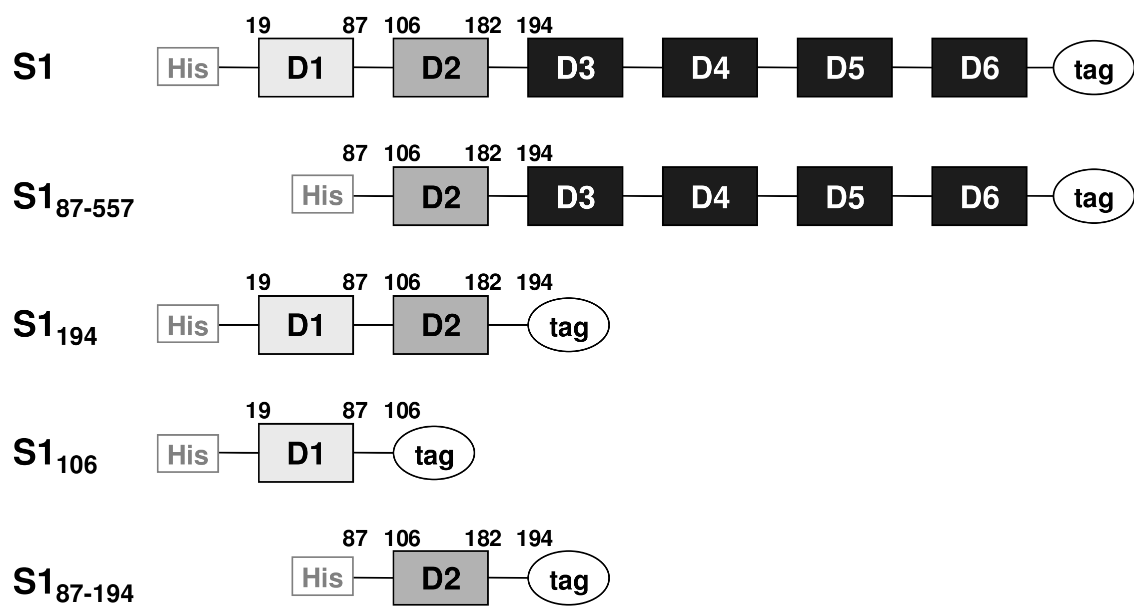
(A) Position of protein S2 and its domains on the 30S ribosomal subunit. The coiled-coil domain (in red) interacts with helices H35-37 (in blue) in the head of the 30S ribosomal subunit, whereas the globular domain (in green) makes contacts with H26 (in olive) in the body of the 30S subunit. (B) Western blot analysis of input and elution fractions in pull down assay showing the direct interaction between his-tagged $S2\alpha_2$ and flag-tagged S1. The inability of protein $S1_{87\text{-}557}$ points out the requirement of the first D1 domain for this interaction. (C) Western blot analysis of the ribosome purified from *E. coli* over expressing $S2\alpha_2$. Ribosome were purified from *E. coli* strain Tuner (DE3) where protein $S2\alpha_2$ was over expressed. Western blot analysis reveled the decreasing amount of protein S2 on

the ribosome while the amount of protein S1 is not changed dramatically. It implies that the coiled-coil domain serves as the binding partner for S1 on the ribosome.

Figure S1. **Yeast two hybrid system indicating interaction of protein S2 and its coiled-coil domain with protein S1 and its N-terminal domain of protein S1**.

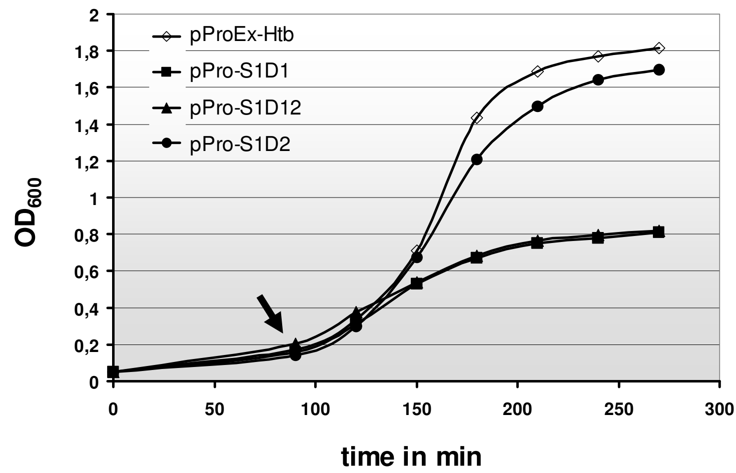
The β-galactosidase activity given in Miller units (MU) was used as reporter for the protein-protein interactions. Lines a and b: controls lacking one interaction partner. Lane c, d and e: MU representing interaction between protein S2 and proteins $S1_{106}$, $S1_{194}$ and native S1, respectively. Lanes f and g: Interaction between the coiled-coil domain of protein S2 and native S1 or its N-terminal domain, respectively.

Byrgazov et al., Fig.1

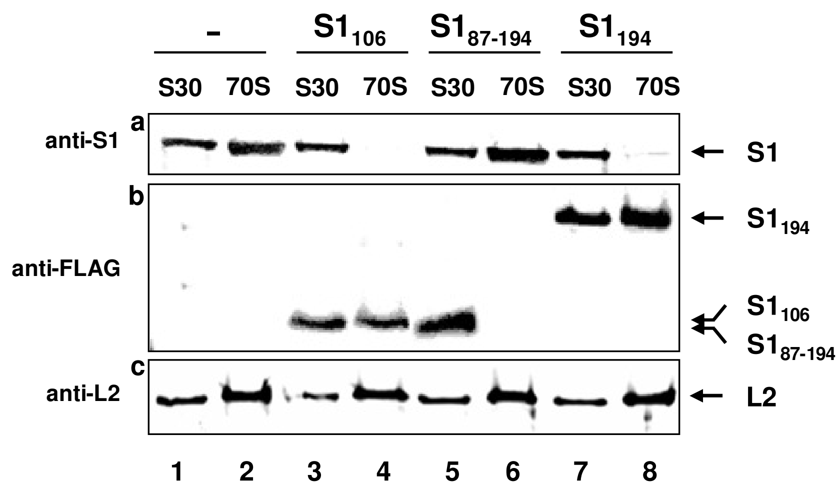


Byrgazov et al., Fig.2

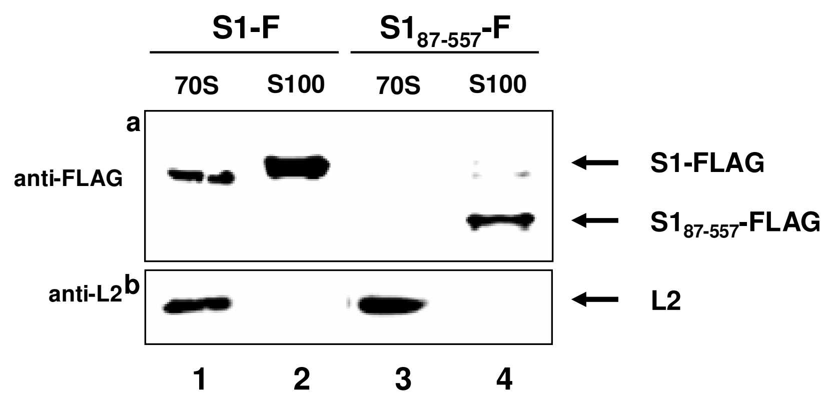
A



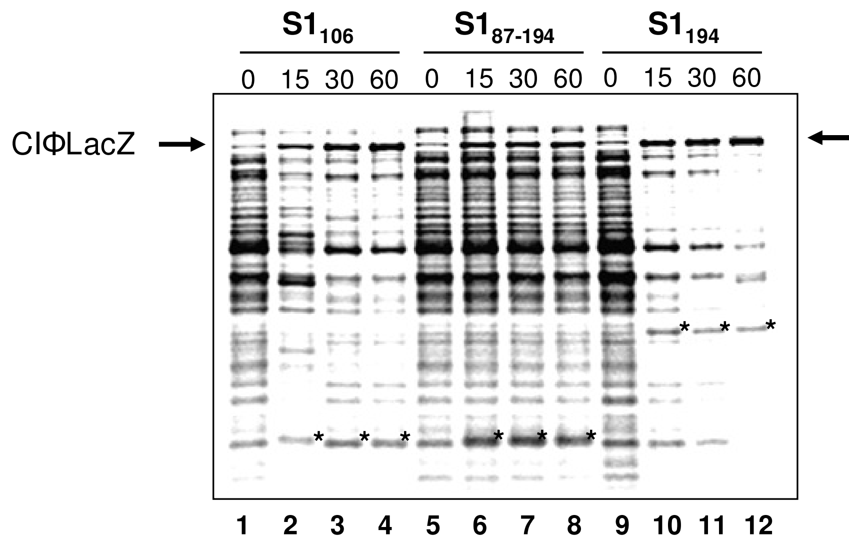
B



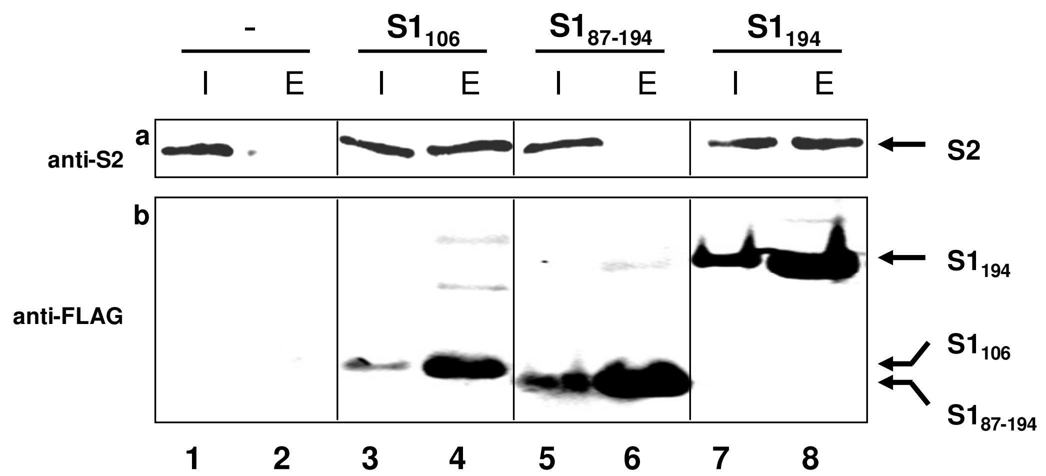
C



Byrgazov et al., Fig.3

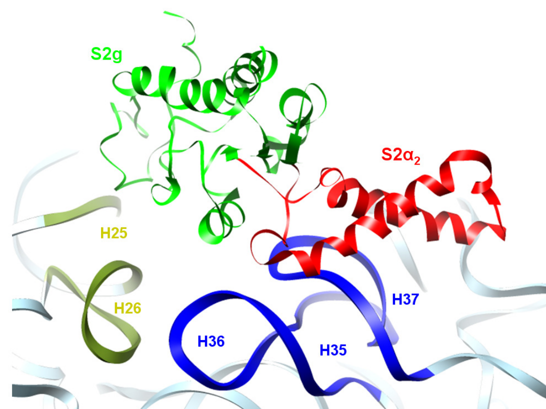


Byrgazov et al., Fig.4

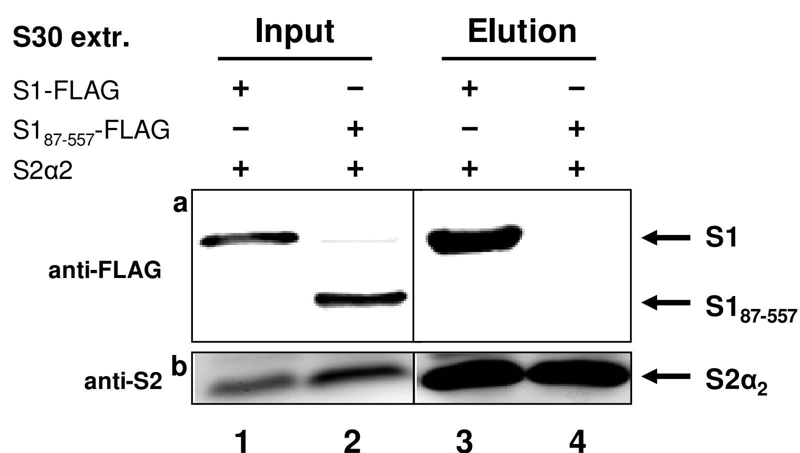


Byrgazov et al., Fig.5

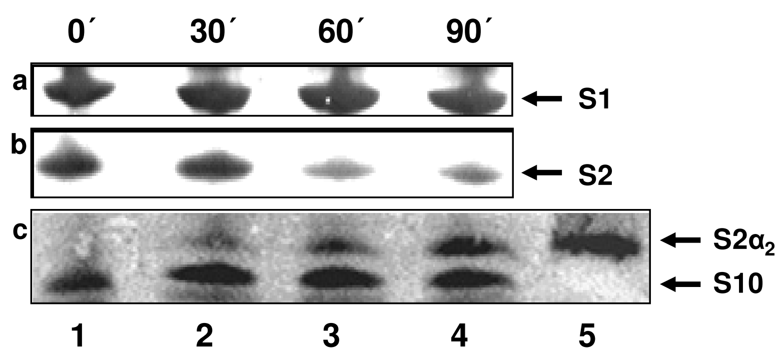
A



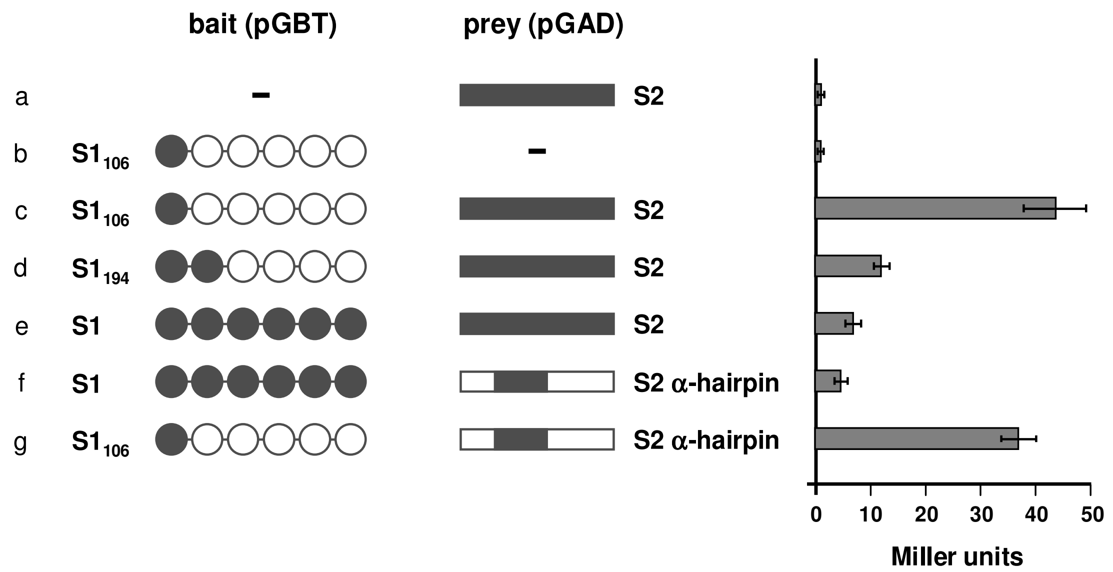
B



C



Byrgazov et al., Fig. S1



References

Aliprandi, P., Sizun, C., Perez, J., Mareuil, F., Caputo, S., Leroy, J.L., Odaert, B., Laalami, S., Uzan, M., and Bontems, F. (2008). S1 ribosomal protein functions in translation initiation and ribonuclease RegB activation are mediated by similar RNA-protein interactions: an NMR and SAXS analysis. J Biol Chem. *283*, 13289-13301.

Aseev, L.V., Levandovskaya, A.A., Tchufistova, L.S., Scaptsova, N.V., and Boni, I.V. (2008). A new regulatory circuit in ribosomal protein operons: S2-mediated control of the rpsB-tsf expression in vivo. RNA. *14*, 1882-1894.

Bear, D.G., Ng, R., Van Derveer, D., Johnson, N.P., Thomas, G., Schleich, T., and Noller, H.F. (1976). Proc Natl Acad Sci U S A. *73*, 1824-1828.

Boileau, G., Sommer, A., and Traut, R.R. (1981). Identification of proteins at the binding site for protein S1 in 70 S ribosomes and 30 S subunits by cross-linking with 2-iminothiolane. J Biol Chem. *256*, 8222-8227.

Bollen, A., Lathe, R., Herzog, A., Denicourt, D., Lecocq, J.P., Desmarez, L., and Lavallé, R. (1979). A conditionally lethal mutation of Escherichia coli affecting the gene coding for ribosomal protein S2 (rpsB). J Mol Biol. *132*, 219-233.

Boni, I.V., Artamonova, V.S., Tzareva, N.V., and Dreyfus, M. (2001). Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. EMBO J. *20*, 4222-4232.

Boni, I.V., Isaeva, D.M., Musychenko, M.L., and Tzareva, N.V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res. *19*, 155-162.

Boni, I.V., Zlatkin, I.V., and Budowsky, E.I. (1981) Ribosomal protein S1 associates with Escherichia coli ribosomal 30-S subunit by means of protein-protein interactions. Eur J Biochem. *121*, 371-376.

Brodersen, D.E., Clemons, W.M. Jr., Carter, A.P., Wimberly, B.T., and Ramakrishnan, V. (2003). Phasing the 30S ribosomal subunit structure. Acta Crystallogr D Biol Crystallogr. *59*, 2044-50.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M., and Murzin, A.G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell. *88*, 235-242.

Cavdar Koc, E., Burkhart, W., Blackburn, K., Moseley, A., and Spremulli, L.L. (2001). The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. J Biol Chem. *276*, 19363-19374.

van Dieijen, G., van Knippenberg, P.H., and van Duin, J. (1976) The specific role of ribosomal protein S1 in the recognition of native phage RNA. Eur J Biochem. *64*, 511-518.

van Dieijen, G., Zipori, P., van Prooijen, W., and van Duin, J. (1978). Involvement of ribosomal protein S1 in the assembly of the initiation complex. Eur J Biochem. *90*, 571-580.

Ederth, J., Mandava, C.S., Dasgupta, S., and Sanyal, S. (2008). A single-step method for purification of active His-tagged ribosomes from a genetically engineered Escherichia coli. Nucleic Acids Res. *37*(2).

Gualerzi, C.O., and Pon, C.L. (1990). Initiation of mRNA translation in prokaryotes. Biochemistry. *29*, 5881-5889.

Guerrier-Takada, C., Subramanian, A.R., and Cole, P.E. (1983). The activity of discrete fragments of ribosomal protein S1 in Q beta replicase function. J Biol Chem. *258*, 13649-13652.

Laughrea, M., and Moore, P.B. (1977). Physical properties of ribosomal protein S1 and its interaction with the 30 S ribosomal subunit of Escherichia coli. J Mol Biol. *112*, 399-421.

Laughrea, M., and Moore, P.B. (1978). On the relationship between the binding of ribosomal protein S1 to the 30 S subunit of Escherichia coli and 3' terminus of 16 S RNA. J Mol Biol. *121*, 411-430.

Marintchev, A., and Wagner, G. (2004). Translation initiation: structures, mechanisms and evolution. Q Rev Biophys. *37*, 197-284.

Moll, I., Grill, S., Gründling, A., and Bläsi, U. (2002). Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. Mol Microbiol. *44*, 1387-1396.

McGinness, K.E., and Sauer, R.T. (2004). Ribosomal protein S1 binds mRNA and tmRNA similarly but plays distinct roles in translation of these molecules. Proc Natl Acad Sci U S A. *101*, 13454-13459.

Potapov, A.P., and Subramanian, A.R. (1992). Effect of E. coli ribosomal protein S1 on the fidelity of the translational elongation step: reading and misreading of poly(U) and poly(dT). Biochem Int. *27*, 745-753.

Rajkowitsch, L., and Schroeder, R. (2007). Dissecting RNA chaperone activity. RNA. *13*, 2053-2060.

Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., and Bontems, F. (2009). Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. Nucleic Acids Res. *37*, 5578-88.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. Science. *310*, 827-834.

Sengupta, J., Agrawal, R.K., and Frank, J. (2001). Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. Proc Natl Acad Sci U S A. *98*, 11991-11996.

Sillers, I.Y., and Moore, P.B. (1981). Position of protein S1 in the 30 S ribosomal subunit of Escherichia coli. J Mol Biol. *153*, 761-780.

Sørensen, M.A., Fricke, J., and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. J Mol Biol. *280*, 561-569.

de Smit, M.H., and van Duin, J. (1990). Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis. Proc Natl Acad Sci U S A. *87*, 7668-7672.

de Smit, M.H., and van Duin, J. (1994). Control of translation by mRNA secondary structure in Escherichia coli. A quantitative analysis of literature data. J Mol Biol. *244*, 144-150.

de Smit, M.H., and van Duin, J. (2003). Translational standby sites: how ribosomes may deal with the rapid folding kinetics of mRNA. J Mol Biol. *331*, 737-743.

Subramanian, A.R., Rienhardt, P., Kimura, M., and Suryanarayana, T. (1981). Fragments of ribosomal protein S1 and its mutant form m1-S1. Localization of nucleic-acid-binding domain in the middle region of S1. Eur J Biochem. *119*, 245-249.

Subramanian, A.R. (1983). Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol. *28*, 101-42.

Sukhodolets, M.V., and Garges, S. (2003). Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. Biochemistry. *42*, 8022-8034.

Sykes, M.T., and Williamson, J.R. (2009). A complex assembly landscape for the 30S ribosomal subunit. Annu Rev Biophys. *38*, 197-215.

Szer, W., Hermoso, J.M., and Boublik, M. (1976). Destabilization of the secondary structure of RNA by ribosomal protein S1 from Escherichia coli. Biochem Biophys Res Commun. *70*, 957-964.

Tedin, K., Resch, A., and Bläsi, U. (1997). Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5' leader sequence. Mol Microbiol. *25*, 189-199.

Thomas, J.O., Boublik, M., Szer, W., and Subramanian, A.R. (1979). Nucleic acid binding and unfolding properties of ribosomal protein S1 and the derivatives S1-F1 and m1-S1. Eur J Biochem. *102*, 309-314.

Thomas, J.O., and Szer, W. (1982). RNA-helix-destabilizing proteins. Prog Nucleic Acid Res Mol Biol. *27*, 157-187.

Ulbrich, B., and Nierhaus, K.H. (1975). Pools of ribosomal proteins in Escherichia coli. Studies on the exchange of proteins between pools and ribosomes. Eur J Biochem. *57*, 49-54.

Vellanoweth, R.L., and Rabinowitz, J.C. (1992). The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. Mol Microbiol. *6*,1105-1114.

Yusupova, G., Jenner, L., Rees, B., Moras, D., and Yusupov, M. (2006). Structural basis for messenger RNA movement on the ribosome. Nature. *444*, 391-394.

Structural insights into the assembly of protein S1 to the ribosome in *Escherichia coli*

Konstantin Byrgazov[1], Nicolas Coudevylle[2], Robert Konrat[2] and Isabella Moll[1‡]

Running title: Interaction between ribosomal proteins S1 and S2

[1]Max F. Perutz Laboratories, Center for Molecular Biology,
Department of Microbiology, Immunobiology and Genetics, University of Vienna,
Dr. Bohrgasse 9/4, 1030 Vienna, Austria

[2]Max F. Perutz Laboratories, Department of Strcutural Biology, University of Vienna, Dr. Bohrgasse, 1030, Vienna, Austria

[‡] Corresponding author:

Isabella Moll
Phone: +43-4277-54631
Fax: +43-1-4277-9546
e-mail: Isabella.Moll@univie.ac.at

Summary

Binding of protein S1 to the ribosome is pivotal for translation in *E. coli* and most Gram-negative bacteria identified so far. Here, we scrutinized the interaction surface between the ribosome and protein S1. During structural analysis we were able to show that the core of the N-terminal domain of protein S1 differs from the general OB-fold of the so-called S1 domain. Moreover, we determined that this core structure is not involved in assembly to the ribosome. In contrast, the flexible (or structurally interchangeable) stretch of eighteen amino acids at the N-terminus of the protein was identified to be essential for the interaction of protein S1 with the ribosome. We show that the 18-mer binds to the ribosome and moreover it competes with native protein S1 for ribosome binding, thus inhibiting translation of canonical mRNAs.

In addition, we identified the residues on protein S2 that are likely involved in the interaction with protein S1. They are located at the C-terminal part of the coiled-coil domain of protein S2, close to the cleft of the head, body, and platform of the 30S subunit.

Introduction

Ribosomal protein S1 represents the largest protein in the *E. coli* ribosome that has a molecular weight of 61 kDa. It is one of the last proteins which associate with the 30S ribosomal subunit during ribosome biogenesis (Sykes and Williamson, 2009). Binding of protein S1 to the 30S subunit is weak and reversible (Subramanian, 1984). In Gram-negative bacteria, protein S1 is a pivotal protein in translation initiation; it interacts with a pyrimidine-rich region within the mRNA upstream of SD-sequence (Boni *et al.*, 1991), thereby increasing the concentration of the translational start site in the vicinity of the decoding site on the ribosome. In addition, protein S1 has been suggested to assist in positioning of the 30S subunit in close proximity to the translational start site by destabilizing secondary structures (de Smit and van Duin, 1994). Hence protein S1 is essential for the translation of canonical mRNAs in *E. coli* and likely all Gram-negative bacteria (Sorensen *et al.*, 1998).

The elongated structure of *E. coli* S1 does not represent a simple triaxial body; small X-ray scattering analysis of S1 in solution proposed a complex shape which might be represented as a long cylinder with an attached short cylinder at N-terminus (Osterberg *et al.*, 1978). Immunoelectron microscopy studies revealed that the extended C-terminal part of protein S1 expands from the interface of the main morphological parts of the 30S subunit: the head, the platform, and the body (Walleczek *et al.*, 1990). Sequence analysis of protein S1 reveals the presence of six repeating motives (Bycroft, 1997). Recently, the solution structures of four C-terminal domains have been determined (Aliprandi *et al.*, 2008) (Salah *et al.*, 2009). Each of these domains possesses an oligosaccharide-oligonucleotide binding (OB) fold and binds to ssRNA. In contrast, phylogenetic

and bioinformatic analysis have shown that the two N-terminal domains are distinct. Moreover, it has been shown that the two N-terminal domains do not bind RNA (McGiness and Sauer, 2004) and that the fragment comprising the two N-terminal domains is implicated in binding to the 30S ribosomal subunit (Subramanian, 1984). This binding of protein S1 to the ribosome occurs *via* protein-protein interactions (Boni *et al.*, 1981) and requires protein S2 (Moll *et al.*, 2002). Recently, we have obtained several lines of evidence that protein S1 employs its N-terminal domain to bind to the coiled-coil domain of protein S2 on the 30S ribosomal subunit (Byrgazov *et.al.*, in prep).

To further elucidate the interaction surface between proteins S1 and S2 we analyzed the structural elements in both proteins required for their interaction. In protein S1 the ribosome-interacting fragment $S1_{106}$ consists of a folded core of four β-strands flanked by flexible N- and C-termini. Here we show that the flexible stretch of eighteen N-terminal residues of protein S1 binds to the ribosome and can interfere with translation of canonical mRNAs.

Studying the coiled-coil domain of protein S2, we were able to identify amino acids which have high propensity to be involved in protein-protein interactions. Point mutations introduced at the C-terminal neck region of protein S2 at positions Asn145 and Gly148 abolish the ability of the coiled-coil domain to interact with ribosome-binding fragment of S1. Since inhibition of the assembly of protein S1 to the ribosome severely affects growth of *E. coli* and most potentially all Gram-negative bacteria, where S1 is essential for translation initiation, interaction surface characterized here more closely could represent a potential target for the design of novel antimicrobial compounds.

Results

The N-terminal domain of S1 is distinct from “S1 domains”

Bioinformatic comparison of the secondary structure of the N-terminal domain D1 of protein S1 (termed protein $S1_{106}$ throughout the manuscript) with the known structure of domain D4 as representative for the RNA-binding domains of protein S1, reveals a slightly different core organization of protein $S1_{106}$ (Figure 1). The hydrophobic residues of the first two strands, b1 and b2 of $S1_{106}$, can be easily aligned to the stabilization elements present in strands B1 and B2 of domain D4 (Figure 1, marked by ▼). However, non-homologous substitutions in $S1_{106}$ within the stabilizing residues of B3 and B5 present in domain D4 (Figure 1, marked by Δ) result in degeneration of these strands: strand b3 of $S1_{106}$ is shortened when compared to B3, and strand B5 is absent in the domain D1 (Figure 1). Strand b4 of $S1_{106}$ is as long as B4 but shifted towards the N-terminus resulting in shortening of the long loop between the b3 and b4. Thus, the core structure of domain D1 is formed by four β-strands instead of five and may not represent OB fold characteristic for RNA-binding domains of protein S1 (Aliprandi *et al.*, 2008) (Salah *et al.*, 2009). As these structural differences observed by computational sequence analyses are in agreement with the observed functional difference between the N-terminal domain D1 and the C-terminal RNA-binding domains of protein S1, we further investigated the structural organization of protein $S1_{106}$ employing NMR (Figure S1). These analyses verified the results of the bioinformatic analysis, revealing that protein $S1_{106}$ consists of a folded core comprising four β-strands, b1, b2, b3, and b4. This core structure is flanked by flexible, structurally interchangeable termini f1 and f2 (Figure 1). Interestingly, these flexible termini

do not affect the folding of the domain core. Comparing the HSQC spectra of ^{15}N-labelled proteins $S1_{106}$ and $S1_{86}$, lacking the flexible stretch f2, or $S1_{19\text{-}86}$, lacking both terminal regions f1 and f2, does not reveal chemical shift perturbations of the core residues (Figure S1).

The first eighteen N-terminal amino acids of protein S1 are required for binding to the ribosome *in vivo*

To dissect the role of the different structural fragments of protein $S1_{106}$ in binding to the ribosome, the affinity of a truncated variant of protein $S1_{106}$ lacking the N-terminal stretch of eighteen amino acids, f1, ($S1_{19\text{-}106}$, Figure 2) was tested *in vivo*. Upon over expression of the C-terminally FLAG-tagged protein $S1_{106}$ or the variant $S1_{19\text{-}106}$ in *E. coli* strain JE28, ribosomes were purified on Ni-NTA-agarose. The presence of the respective protein variants on the ribosome was examined by western blot analysis employing antibodies directed against the FLAG tag. The result shown in Figure 3 reveals that protein $S1_{19\text{-}106}$ does not interact with the ribosome, as it can not be detected in the ribosome containing fraction (Figure 3A, lane 2). In contrast, FLAG-tagged $S1_{106}$ co-precipitates with the ribosome (Figure 3A, lanes 3 and 4). This result indicates that the N-terminal 18 amino acids of S1 are required for binding to the ribosome. To support this observation and to verify that the N-terminal 18 amino acids are required for binding of full-length protein S1 to the ribosome, we deleted these residues from full-length protein S1. As shown in Figure 3B, this deletion abrogates binding of protein S1 to the ribosome as protein $S1_{19\text{-}557}$ can not be detected at the ribosome containing fraction (Figure 3B, lane 4). Thus, these results suggest that the primary interaction site of protein S1 with the ribosome is located within its N-

terminal eighteen amino acid residues. It can be envisioned that f1 can guide the folded core of D1 to the ribosome.

Fusion of the eighteen N-terminal amino acids to S1 lacking the N-terminal domain allows assembly of the fusion protein onto the ribosome *in vivo*

We have shown that the N-terminally truncated variant of protein S1, $S1_{87\text{-}557}$, lacking the first domain D1, cannot bind to the ribosome (Figure 3B, lane 6). To further elucidate the potential role of the N-terminal 18 amino acids in conferring binding activity to the ribosome, we tested whether replacement of f2 with f1 on protein $S1_{87\text{-}557}$ can allow binding to the ribosome (Figure 2). Therefore, the gene encoding this fusion protein was over expressed in *E. coli* strain JE28 followed by co-precipitation with his-tagged ribosomes. The western blot analysis revealed that the fusion protein $S1_{18\Phi106\text{-}557}$ can bind to the ribosome (Figure 3B, lanes 7 and 8) whereas protein $S1_{87\text{-}557}$ does not (Figure 3B, lanes 5 and 6).

Peptide $S1_{18}$ interacts with the 30S ribosomal subunit

To verify the direct interaction of the peptide comprising the N-terminal eighteen amino acids of S1 ($S1_{18}$) with the 30S ribosomal subunit, we employed an ultrafiltration assay using a membrane with a molecular weight cut off (MWCO) of 100 kDa. Therefore, the membrane used in the assay is able to retain the 30S ribosomal subunit (Figure 4, lanes 1 and 2) but neither unbound native S1 (Figure 4, lanes 5 and 6) nor unbound peptide $S1_{18}$ (Figure 4, lanes 3 and 4). To facilitate the detection of the peptide, it was synthesized as FITC-labeled derivative. The 30S ribosomal subunits depleted of S1 (30S(-S1)) were incubated in the presence of peptide followed by ultrafiltration as described in Materials and methods.

As shown in Figure 4, lanes 9 and 10, protein S1 binds to the 30S(-S1) ribosomal subunits as expected. In addition, we were able to determine the presence of FITC-$S1_{18}$ (Figure 4, lanes 7 and 8) in the ribosome fraction. Interestingly, when a mixture of FITC-$S1_{18}$ and full-length protein S1 in a molar ratio 5:1 was used , the affinity of both molecules to the 30S(-S1) ribosome is reduced (Figure 4, lanes 11 and 12). Hence, this result supports the notion that native protein S1 and its N-terminal eighteen amino acids peptide compete for the same binding site on the 30S ribosomal subunit.

The peptide $S1_{18}$ inhibits translation of canonical mRNA *in vitro*

Taken together these results indicate that peptide $S1_{18}$ representing the first 18 residues of S1 binds to the ribosome and subsequently prevents binding of native protein S1. Previously we have shown that over expression of the ribosome-binding protein $S1_{106}$ inhibits translation of bulk mRNA in *E. coli* (Byrgazov *et al.*, in prep.). Since induced synthesis of short peptides *in vivo* is not applicable due to rapid proteolytic degradation of short oligopeptide fragments, we tested the influence of $S1_{18}$ on *in vitro* translation of a model mRNA, which requires S1 for translation initiation, namely *ompA* mRNA (Tedin *et al.*, 1997). Considering the results of the binding studies, we hypothesized that *in vitro* translation of canonical mRNAs might be inhibited in the presence of peptide $S1_{18}$ interfering with binding of native S1 to the ribosome. As a control we used protein $S1_{106}$ which has been shown to hinder translation of bulk mRNAs *in vivo*. As shown in Figure 5, addition of both, peptide $S1_{18}$ and protein $S1_{106}$, inhibits *in vitro* translation of the *ompA* mRNA when present in 100-fold molar excess over the ribosome (Figure 5). Since for the *in vitro* translation assay fully assembled ribosomes compris-

ing protein S1 were used, this result indicates that both protein $S1_{106}$ and peptide $S1_{18}$ displace native protein S1 from the ribosome.

Peptide $S1_{18}$ interacts with the coiled-coil domain of protein S2

Previously we have determined the direct interaction between the coiled-coil domain of protein S2 ($S2\alpha_2$) and domain D1 of protein S1 (Byrgazov *et.al.*, in prep). Since the 18 N-terminal residues of protein S1 are required to interact with the ribosome, we hypothesized that peptide $S1_{18}$ might directly interact with protein $S2\alpha_2$ representing the coiled-coil domain of protein S2. To determine this interaction, we employed a comparative analysis of HSQC spectra of ^{15}N-labelled protein $S2\alpha_2$ upon addition of increasing amount of peptide $S1_{18}$. As can be seen in Figure 6, upon addition of the peptide $S1_{18}$ in a 4:1 molar ratio over $S2\alpha_2$, several significant shifts can be observed. Thus this experiment strongly supports our hypothesis that peptide $S1_{18}$ interacts with the coiled-coil domain of protein S2. Noteworthy, the signals highlighted in the HSQC spectrum of ^{15}N-$S2\alpha_2$ were strongly affected upon addition of peptide $S1_{18}$ (Figure 6). Since generally the signals of N-H bonds of glycine residues are located in this area of a ^{15}N-^{1}H-HSQC-spectrum, this result indicates that the glycine residues present in the coiled-coil domain of S2 are involved in binding of protein S1.

The glycine residues located at positions 148/149 as well as residue Asn145 of protein S2 are required for the interaction between the coiled-coil domain of S2 and ribosome-binding fragment of S1.

The perturbations observed in the HSQC-spectrum of protein $S2\alpha_2$ upon titration of peptide $S1_{18}$ indicated that glycine residues present in the protein may be in-

volved in interaction with peptide $S1_{18}$. In addition, the meta-structure analysis (Konrat, 2009) of protein S2 indicated that residues located at the N- and C-termini of the coiled-coil domain possess a high propensity for protein interactions (Figure S4). Since structural analysis of the ribosome revealed that the N-terminal α-helix of $S2\alpha_2$ is involved in interaction with the 16S rRNA (Brodersen *et al.*, 2002), it was tempting to speculate that the interaction with protein S1 occurs on the C-terminus of the coiled-coil domain (Figure 7A). Interestingly, two duplets of glycine residues are located within the C-terminus of $S2\alpha_2$ (Figure S3B). Therefore, we assayed the importance of the glycine duplets in the interaction with protein S1, employing site-directed mutagenesis. The mutant protein variants of $S2\alpha_2$ were tested for interaction with the ribosome-binding fragment of S1, protein $S1_{194}$ (McGiness and Sauer, 2004). In addition, we assayed the role of asparagine residue Asn145, the side-chain of which is oriented towards the cleft of three main morphological parts of the small ribosomal subunit, the head, body, and platform (Fig. 7A) where the N-terminus of S1 is proposed to bind (Walleczek *et al.*, 1990). The genes encoding the mutant variants of the $S2\alpha_2$ domain, harbouring the G148A and G149A mutations or the N145L mutation were over expressed in *E. coli* strain Tuner as described in Materials and Methods. Obtained mutant and wild-type protein variants of $S2\alpha_2$ were used as baits in pull-down assay whereas FLAG-tagged protein $S1_{194}$ served as a prey. The interaction between proteins was assayed by Western blot analysis employing antibodies directed against FLAG-tag. As shown in Figure 7B, in contrast to wild-type protein $S2\alpha_2$, the mutations GG148AA and N145L abolished interaction with protein $S1_{194}$. Therefore, these results strongly indicate that these residues on protein S2 might be involved in interaction with protein S1.

Discussion

In Gram-negative bacteria protein S1 consists of six domains (Gribskov, 1992). The four C-terminal domains bind to single-stranded RNA (Subramanian, 1984). However, two N-terminal domains have been shown to contain the ribosome-binding site of protein S1 and not to interact with RNA (McGiness and Sauer, 2004). Interestingly, the phylogenetic analyses revealed that these two domains are divergent from RNA-binding domains of S1 and from each other (Salah *et al.*, 2009). Previously, we identified that the first N-terminal domain D1 is responsible for interaction with the ribosome (Byrgazov *et al.*, in prep.). The over expression of protein $S1_{106}$ comprising the first 106 amino acids of protein S1 representing domain D1 hinders growth of *E. coli* and inhibits translation of bulk mRNA. With the objective to narrow down the ribosome-binding site of protein S1, we studied the structural arrangement of domain D1. Employing NMR techniques, we were able to show that $S1_{106}$ consists of a folded core flanked by flexible N- and C-termini. Supporting the results of the phylogenetic analysis, we showed that the core of domain D1 is strikingly different from the RNA binding domains of protein S1 representing the so called "S1 domain". In contrast to the β-barrel structure of an OB-fold comprising five β-strands (Bycroft, 1997), domain D1 harbours only four β-strands, which form the stable core of the N-terminal domain of protein S1. However, we were able to show that the lack of the N-terminal stretch consisting of eighteen amino acids ($S1_{18}$) abolishes ribosome-binding activity of protein S1.

To dissect the role of $S1_{18}$ in binding of S1 to the ribosome, we tested whether a peptide representing the N-terminal flexible stretch of S1 ($S1_{18}$) can bind to the

30S ribosomal subunit. We observed that $S1_{18}$ peptide can bind to the 30S subunits depleted of protein S1. Moreover, addition of the peptide decreases the amount of native protein S1 which is able to potentially bind to the 30S(-S1) ribosome. Hence, we concluded that the peptide efficiently competes with native protein S1 for the same binding pocket on the ribosome.

To test whether this competition has an impact on translation of canonical mRNA, we monitored the impact of addition of the peptide $S1_{18}$ on an *in vitro* translation system programmed with *ompA* mRNA containing secondary structures within its translation initiation region (TIR). When added in high molar excess, peptide $S1_{18}$ is able to decreases the translation yield, comparable to protein $S1_{106}$ representing the N-terminal fragment of S1, over expression of which has been previously shown to inhibit translation of bulk mRNA in *E. coli* (Byrgazov *et.al.*, in prep). The requirement of high excess of peptide $S1_{18}$ and protein $S1_{106}$ can be explained by the presence of native S1 assembled to the ribosomes present in the *in vitro* translation reaction. The inhibitory activity of both protein $S1_{106}$ and peptide $S1_{18}$ on translation supports the idea that both molecules bind to the same site on the ribosome.

Previously, we have observed that the coiled-coil domain of ribosomal protein S2 is sufficient to allow protein S1 to bind to the ribosome (Byrgazov *et.al.*, in prep). In addition, we have shown that the N-terminal fragment of S1 comprising the first domain D1 is required for direct interaction between the proteins. Therefore, we hypothesized that peptide $S1_{18}$, since it is required for protein S1 to bind to the ribosome, could directly interact with the coiled-coil domain of S2 (referred here as $S2\alpha_2$). The significant shifts of several signals in the ^{15}N-^{1}H-HSQC spectrum of $S2\alpha_2$ observed upon addition of $S1_{18}$ indicated conformational changes in the

structure of S2α_2. These perturbations can likely be attributed to a direct interaction between protein S2α_2 and peptide S1$_{18}$. One of the most significant shifts on the ^{15}N-^{1}H-HSQC spectrum of S2α_2 occurred in the area where the signals of N-H bonds of glycine residues are observed. Employing site-directed mutagenesis we investigated if the glycine residues present in the coiled-coil domain of S2 play a significant role in binding to protein S1, particularly, its N-terminal fragment. Using S2α_2 as bait in pull-down assay, we confirmed the direct interaction between the coiled-coil domain of S2 and the N-terminal fragment of S1 (S1$_{194}$) which is involved in the ribosome binding. In addition, the exchange of glycine residues duplet within the C-terminus of the second α-helix of the coiled-coil domain (Gly148 and Gly149 according to numeration within full-length protein S2) inhibits the interaction between S2α_2 and S1$_{194}$. Thus, two glycine residues located on the C-terminal turn of the second helix within coiled-coil domain of S2 are required for interaction with S1. Also, we identified another residue on protein S2 which mutation abrogates S2α_2/S1$_{194}$ interaction. Replacement of Asn145 with leucin on the coiled-coil domain of S2 resulted in non-interacting form of S2α_2. Taken together, our data support the notion that the flexible region of eighteen amino acids in length located at the N-terminus of protein S1 could serve as a primary interaction site for the ribosome. Due to its intrinsic flexibility the S1$_{18}$ region could serve as an anchoring domain, which interacts specifically with residues at the boundary between the coiled-coil and globular domains of protein S2 *via* an induced fit mechanism. This hypothesis is supported by the available structures of the 30S ribosomal subunit showing that the side chain of Asn145 as well as the glycine residues Gly148 and Gly149 are oriented towards the cleft of three main morphological parts of the 30S subunit: the head, body,

and platform. Intriguingly, the binding site for N-terminus of S1 has been proposed to be situated at this cleft (Walleczek *et al.*, 1990). Thus, side chain of Asn145 likely participates in the interaction with protein S1.

In this work we shed a light on composition of the binding pocket of protein S1 on the ribosome. Due to essentiality of protein S1 for translation initiation in Gram-negative bacteria, this binding pocket might represent a potential target for antimicrobials semi-selective against Gram-negative opportunistic pathogens.

Materials and Methods

Table 1. Bacterial strains and plasmids used in this study.

Relevant features		Source or reference
E. coli strains:		
JE28	MG1655::*rplL-his*	Ederth *et al.*, 2009
Tuner	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm lacY1*	Novagen
Tuner(DE3)	F^- *ompT* $hsdS_B$ (r_B^- m_B^-) *gal dcm lacY1(DE3)*	Novagen
Plasmids:		
pProEX-HTb	vector for Trc driven gene expression	Invitrogen
pProEX-S2ccWT	encodes protein his-tagged $S2\alpha_2$ WT	this study
pProEX-S2ccN	encodes protein his-tagged $S2\alpha_2$ N145L	this study
pProEX-S2ccGG	encodes protein his-tagged $S2\alpha_2$ GG148AA	this study
pPro-S1F	encodes flag-tagged S1 without his-tag	Byrgazov *et al.*, in prep
pPro-S1d18F	encodes flag-tagged $S1_{19-557}$ without his-tag	this study
pPro-S1D1F	encodes flag-tagged $S1_{106}$ without his-tag	Byrgazov *et al.*, in prep
pPro-S1D1d18F	encodes flag-tagged $S1_{19-106}$ without his-tag	this study
pPro-S1D12	encodes flag-tagged $S1_{194}$ without his-tag	Byrgazov *et al.*, in prep
pET22b	vector for T7 driven over expression	Novagen
pET-$S1_{106}$	pET derivative encoding for his-tagged $S1_{106}$	this study
pET-$S1_{86}$	pET derivative encoding for his-tagged $S1_{86}$	this study
pET-$S1_{19-86}$	pET derivative encoding for his-tagged $S1_{19-86}$	this study

Table 2. Synthetic oligonucleotides used in this study

Oligonucleotide (Purpose)

P1 TATAGGCGCCGAATTCAACCATCGCTGGCTGG (fwd primer to clone S2cc)
P2 TATACTCGAGTTAGTCCGGCAGACCGC (rev primer to clone S2cc)
P3 P-gcagctatcaaagacatg (fwd primer to perform GG148AA mutagenesis)
P4 P-agctgccaggctgttttc (rev primer to perform GG148AA mutagenesis)
P5 P-ctgagcctgggcgg (fwd primer to perform N145L mutagenesis)
P6 P-cagttccagtttctc (rev primer to perform N145L mutagenesis)
P7 P-CCGGGTTCTATCGTTCG (fwd primer to remove the first 18 codons from *rpsA* sequence)
P8 P-CATGGTCTGTTTCCTGTG (rev primer to remove the first 18 codons from *rpsA* sequence)
P9 TATACATATGACTGAATCTTTTGCTC (fwd primer to amplify *rpsA* sequence from the 1st codon)
P10 TATACATATGACCCGCCCGGGTTC (fwd primer to amplify *rpsA* sequence from 19th codon)
P11 TATACTCGAGTTATTCAGCATCTTCGTAAGC (rev primer to amplify *rpsA* until 106th codon)
P12 TATACTCGAGTTACAGCAGAGTTTCAC (rev primer to amplify *rpsA* until 86th codon)

Bacterial strains and plasmids

E. coli strains, plasmids and oligonucleotides used in this study are listed in tables 1 and 2. Unless otherwise indicate, bacterial cultures were routinely grown in LB medium supplemented with ampicillin (100µg/ml). When appropriate, kanamycin (20µg/ml) was added. Where indicated, isopropyl β-D-1-thiogalactopyranoside (IPTG). Growth was monitored by measuring the optical density at 600 nm (OD_{600}).

Construction of plasmids

The DNA sequence encoding protein $S2\alpha_2$ was amplified by PCR employing primers P1 and P2 and cloned under control of the Trc promoter at restriction sites *NarI* and *XhoI* of plasmid pProEX-HTb (Invitrogen). The obtianed plasmid was used as a template for Phusion site-directed mutagenesis kit (NEB) resulting in plasmids pProEX-S2ccN (employing primers P5 and P6) and pProEX-S2ccGG (employing primers P3 and P4).

The DNA sequences encoding proteins $S1_{106}$, $S1_{86}$ and $S1_{19\text{-}86}$ were amplified by PCR employing pairs of primers P9/P11, P9/P12 and P10/P12 respectively and cloned under control of T7 RNA polymerase promoter in sites *NdeI* and *XhoI* of plasmid pET22b (Novagen).

Plasmids pPro-S1F and pPro-S1D1F were used as templates to construct plasmids pPro-S1d18F and pPro-S1D1d18. Briefly, the coding sequence for the N-terminal 18 amino acids was removed employing Phusion site-directed mutagenesis kit (NEB) following manufacturer protocol employing primers P9 and P10. All the constructs were verified by sequencing (AGOWA).

Purification of 30S and 50S ribosomal subunits

To purify 30S and 50S ribosomal subunits we modified a novel method for the ribosome isolation avoiding ultracentrifugation in sucrose gradients (Ederth *et al.*, 2008), which is based on the presence of the His-tagged protein L7/L12 on the ribosome. This allows purification of the ribosomes employing the Ni-NTA-agarose. *E. coli* strain JE28 was cultivated in 2L of LB medium supplemented with kanamycin (20µg/ml). At OD_{600} 0.7-0.9 the culture was harvested by centrifugation at 5000g for 20 min at $4^{o}C$. The cell pellet was resuspended in lysis buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 100mM KCl, 10mM Imidazole, 1u/mL RNase-free DNase I (Roche), 0.1mM PMSF). The cells were disrupted by the freezing-thawing

method. The lysate was clarified by centrifugation at 15000g for 20min at 4°C. The clear extracts were applied to 10 ml of pre-equilibrated in lysis buffer Ni-NTA-agarose (Invitrogen), washed by 10 column volumes of washing buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 20mM Imidazole). Then, the flow through was stopped and the Ni-NTA-agarose with bound ribosomes was resuspended in 10 column volumes of dissociation buffer (20mM Tris·HCl, pH 7.4, 1mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 20mM Imidazole) and kept for 4-12 hrs at 4°C. After incubation, the flow through fractions were collected and the Ni-NTA-agarose was washed with 5 column volumes of dissociation buffer. The combined flow through and washing fractions contained 30S ribosomal subunits and the $MgCl_2$ concentration in there was adjusted to 10mM. The tetra-His-tagged 50S subunits were eluted by 10 column volumes of elution buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 150mM Imidazole). The ribosome containing fractions were dialyzed against 5000 volume of TICO buffer (20mM HEPES-HCl pH 7.6, 6mM $MgCl_2$, 30mM NH_4Cl and 4mM 2-mercaptoethanol) and concentrated on Amicon filter devices with 100 kDa cut off membrane. This method allowed to obtained 10nmol of each subunit.

Preparation of 30S ribosomal subunits depleted of S1

Protein S1-depleted 30S ribosomes were prepared by affinity chromatography using poly(U)-Sepharose 4B (Pharmacia) (Suryanarayana and Subramanian, 1983).

Co-purification of the tetra-His-tagged ribosome with flag-tagged protein S1 variants

E. coli JE28 cells transformed with plasmids pProS1F, pProS1d18F, pProS1D1F and pProS1d18F were grown overnight in LB broth supplemented with 100µg/ml ampicillin and 20µg/ml kanamycin, diluted 1:100 into fresh medium, grown to OD_{600} 0.30-0.35 and synthesis of FLAG-tagged protein S1 variants was induced by addition of 100µM of IPTG. 1h upon induction cells were harvested by centrifugation and lysed by freezing-thawing method in lysis buffer containing 20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 100mM KCl, 10mM Imidazole, 0.1mM PMSF, 1u/mL RNase-free Dnase I (Roche). After centrifugation at 4°C, 15000g for 30min, the resultant supernatants (S30 extracts) were applied to Ni-NTA agarose, washed by 10 column volumes of washing buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl,

150mM KCl, 20mM Imidazole) followed by elution with elution buffer containing 20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 150mM KCl, 150mM Imidazole. The protein composition of the ribosomes was estimated by running the same number of A_{260} units on SDS-PAGE followed by Western blot analysis using anti-FLAG (Abcam) and anti-L2 antibodies.

In vitro translation assays

Full-length *ompA* mRNA was translated *in vitro* with *E. coli* S30 Extract System for Linear Templates (Promega). The *in vitro* translation reactions containing 1µCi/ml of [^{35}S]-methionine, 0.2 µM of *ompA* mRNA and 0.3 µM of ribosomes (according to the assumption that 2/3 of absorbance at 260 nm is given by the ribosome in S30 extract) were incubated for 60 min at 37°C either in the absence or in the presence of 3µM or 30µM of purified proteins $S1_{106}$, $S1_{87\text{-}194}$ or peptide FITC-$S1_{18}$. Translation was stopped by addition of SDS-protein sample buffer followed by heating for 5min at 95°C. The samples were loaded on a 12% SDS-polyacrylamide gel. Gels were dried and exposed to a Molecular Dynamics PhosphoImager for visualization and quantification.

Purification of ^{15}N-labeled proteins $S1_{106}$, $S1_{86}$, $S1_{19\text{-}86}$ and $S2\alpha_2$

E.coli strain Tuner (DE3) containing plasmids pET-$S1_{106}$, pET-$S1_{86}$, pET-$S1_{19\text{-}86}$ and pProEX-S2ccWT were grown overnight in M9 minimal media supplemented with $^{15}NH_4Cl$ (Sigma, 1 g/l) and 100µg/ml ampicilin, diluted 1:50 into fresh medium, grown to OD_{600} 0.8-0.9 and induced by addition of 1 mM of IPTG. After 2h upon induction cells were harvested by centrifugation and lysed by three times freezing-thawing in lysis buffer containing 50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 10mM Imidazole, 0.1% Tween-20, 0.5 mg/mL DNase I (Roche), 20 µg/ml RNase A. After centrifugation at 4°C, 15000g for 30min, the resultant supernatants (S30 extracts) were applied to Ni-NTA agarose, washed by 10 column volumes of washing buffer (50mM Na_2HPO_4, pH 8.0, 500mM NaCl, 20mM Imidazole) followed by elution with elution buffer (50mM Na_2HPO_4, pH 8.0, 300mM NaCl, 250mM Imidazole). The protein composition of the eluted fractions was determined by running the same number of total protein amount on SDS-PAGE. The eluted fractions were dialyzed into running buffer (1x PBS, Amicon) and loaded on HiLoad Sephadex 75 16/60 column (GE Healthcare). The size exclusion FPLC was performed in running buffer at 4°C. The

purified proteins were concentrated on Amicon devices with 3 kDa cut of membrane. Protein concentration was determined employing Bradford assay.

In vitro binding of FITC-labeled peptide $S1_{18}$ to the 30S(-S1) subunit

The N-terminally FITC-labeled $S1_{18}$ peptide (FITC-MTESFAQLFEESLKEIE-COOH) was *in vitro* synthesized, where FITC denotes Fluorescein IsoThioCyanate. The average molecular mass of the peptide was determined to be 2,254 with an Applied Biosystems Voyager System 1105 mass spectrometer. 40 pmol of 30S(-S1) was incubated either with 80 pmol of native S1 or with 400 pmol of FITC-$S1_{18}$ in 50 µL of TICO buffer at 37°C for 30'. After addition of 50µL of TICO buffer the resultant mixtures were applied to Amicon centrifugation device with 100 kDa cut off membrane and concentrated to volume of 50µl at 5000g. This step was repeated twice. The retained fractions were subjected to SDS-PAGE analysis. After electrophoresis polyacrylamide gel was stained with SYPRO Ruby (Invitrogen) according to manufacturer protocol. The gel was scanned on Typhoon using 488nm laser and two filters: 520nm to detect FITC and 610nm to detect SYPRO Ruby stained proteins.

Pull-down analysis

In pull-down assays wild type and mutant variants of protein $S2\alpha_2$ were used as his-tagged baits. Flag-tagged protein $S1_{194}$ was used as a prey. *E. coli* strain Tuner transformed with either pPro-S1D12 encoding flag-tagged protein $S1_{194}$ or pProEX-S2cc plasmids encoding his-tagged wild type and mutant forms of protein $S2\alpha_2$ were incubated at 37°C. Over expression of the respective genes was induced at OD_{600} 0.6-0.8 by addition of 1mM IPTG. The cells were harvested and disrupted by sonication in lysis buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 100mM KCl, 10mM Imidazole, 0.1mM PMSF, 20µg/mL Rnase A, 10µg/mL Dnase I). The lysates were clarified by centrifugation at 15000g for 30min at 4°C followed by ultracentrifugation at 100000g for 1h at 4°C. Prey- and bait-containing clarified lysates were mixed and incubated at 37°C for 30 min. After that pre-equilibrated Ni-NTA-agarose was added and incubation was continued at 4°C for 4 hours under continuous agitation. Ni-NTA-agarose was subsequently washed with 5 column volumes of washing buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 200mM KCl, 30mM Imidazole). Bound proteins were eluted by elution buffer (20mM Tris·HCl, pH 7.4, 10mM $MgCl_2$, 30mM NH_4Cl, 100mM KCl, 200mM Imidazole) and analysed on

SDS-PAGE followed by Western blot analysis employing anti- $S2\alpha_2$ and anti-ECS (Bethyl) antibodies.

Figure Legends

Figure 1. Secondary structure comparison of domains D1 and D4 of protein S1. The N-terminal domain of protein S1 consists of a folded core formed by four β-strands, flanked by flexible termini f1 and f2. Alignment of D1 and D4 sequences shows the non-homologous substitutions in residues which have been proposed to participate in stabilization of hydrophobic core of so called “S1 domain” (marked by Δ). Flexible flanking regions f1 and f2 are depicted on the sequence of domain D1.

Figure 2. Variants of protein S1 used in this study.
The domains of protein S1 and position of flexible linkers f1 and f2 are depicted. Numbering is according to amino acid position in native protein S1.

Figure 3. Co-purification of protein S1 variants with his-tagged ribosomes.
(A) Western blot analysis of input (S30, lanes 1 and 3) and elution (70S, lanes 2 and 4) fractions obtained from *E. coli* strain JE28 transformed with plasmids pProS1D1F (lanes 3 and 4) and pProS1D1d18F (lanes 1 and 2) encoding for FLAG-tagged N-terminal domain $S1_{106}$ and its truncated version lacking f1, respectively. The amount of proteins S1 (upper panel), $S1_{106}$ and $S1_{19\text{-}106}$ (middle panel) were determined employing anti-S1 and anti-FLAG antibodies, respectively. Protein L2 served as an internal control (lower panel). (B) Western blot analysis of input (S30, lanes 1, 3, 5 and 7) and elution (70S, lanes 2, 4, 6 and 8) fractions prepared from *E. coli* strain JE28 harboring plasmids pProS1F, pProS1d18F, pProS1ΔD1F and pProS1f1ΦF encoding FLAG-tagged form of native protein S1 and its variants $S1_{19\text{-}557}$ lacking f1, $S1_{87\text{-}557}$ lacking D1 except f2 and fusion protein $S1_{18\Phi106\text{-}557}$ which is similar to $S1_{87\text{-}557}$ but containing f1 instead of f2. The amount of proteins was determined using anti-FLAG antibodies (upper panel), whereas protein S3 served as an internal control (lower panel).

Figure 4. Ultrafiltration assay.
30S ribosomes depleted of protein S1 were tested in binding of native protein S1 as well as FITC-labeled peptide $S1_{18}$. SDS-PAGE analysis of input (lanes 1, 3, 5, 7, 9 and 11) and fractions retained on the membrane (lanes 2, 4, 6, 8, 10 and 12) is

depicted on the picture. The amount of ribosomal proteins was estimated as described in Materials and Methods.

Figure 5. Influence of peptide $S1_{18}$ and N-terminal domains D1 and D2 on efficiency of *in vitro* translation system programmed with the *ompA* mRNA.

Increasing amounts of peptide $S1_{18}$ (lanes 6 and 7) as well as proteins $S1_{106}$ (lanes 4 and 5) and $S1_{87\text{-}194}$ (lanes 2 and 3) served as positive and negative controls, respectively were added to *in vitro* translation system programmed with the *ompA* mRNA. SDS-PAGE analysis and its quantification was done as described in Materials and Methods.

Figure 6. HSQC-titration of ^{15}N-labelled coiled-coil domain of protein S2 by peptide $S1_{18}$.

Increasing amount of peptide $S1_{18}$ was added to ^{15}N-labelled protein $S2\alpha_2$. The spectrum of $^{15}N\text{-}S2\alpha_2$ is in blue whereas the spectrum of $^{15}N\text{-}S2\alpha_2$ with added peptide $S1_{18}$ is in red. Several significant signal shifts were observed on the $^{15}N\text{-}^{1}H$-HSQC spectrum of protein $^{15}N\text{-}S2\alpha_2$. The squared area contains the signals of glycine residues present in $S2\alpha_2$.

Fugure 7. Pull-down assay assaying interaction between protein $S1_{194}$ and wild type and mutant variants of protein $S2\alpha_2$.

(A) Position of protein S2 and its domains on the 30S ribosomal subunit. The coiled-coil domain (in red) interacts with helices H35-37 (in blue) in the head of the 30S ribosomal subunit, whereas the globular domain (in green) makes contacts with H26 (in olive) in the body of the 30S subunit. The residues Asn145 and Gly149 are depicted by arrows. The side chain of Asn145 is oriented towards the putative binding site of protein S1 on the ribosome. (B) Pull down assay showing the direct interaction between ribosome-binding variant of S1, protein $S1_{194}$, and protein $S2\alpha_2$; single mutation N145L and double mutation GG148AA abrogate this interaction. Western blot analysis of flow through (FT, lanes 1, 3 and 5) and elution (E, lanes 2, 4 and 6) fractions. The amount of prey protein $S1_{194}$-FLAG was determined employing anti-FLAG antibodies (upper panel) whereas the amount of prey proteins was determined employing anti-$S2\alpha_2$ antibodies (lower panel).

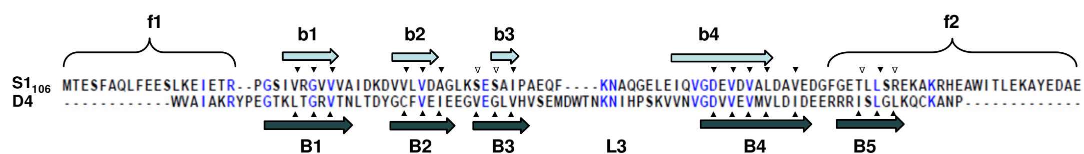
Figure S1. Comparison of folding of three different construct representing the N-terminal domain D1: proteins $S1_{106}$, $S1_{86}$ and $S1_{19\text{-}86}$.

(A) Over lapping of ^{15}N-^{1}H-HSQC spectra of proteins $S1_{106}$ (in blue) and $S1_{86}$ (in red) shows no significant shifts thus it verifies that the folding of the structured core of domain D1 is not affected upon removal of f2 region. (B) Enlarge of crowded area of ^{15}N-^{1}H-HSQC spectra of proteins $S1_{106}$ and $S1_{86}$. (C) Comparison of ^{15}N-^{1}H-HSQC spectra of proteins $S1_{86}$ (in blue) and $S1_{19\text{-}86}$ (in red) shows that both proteins have similar folding since no significant shifts of signals are observed upon removal of f1.

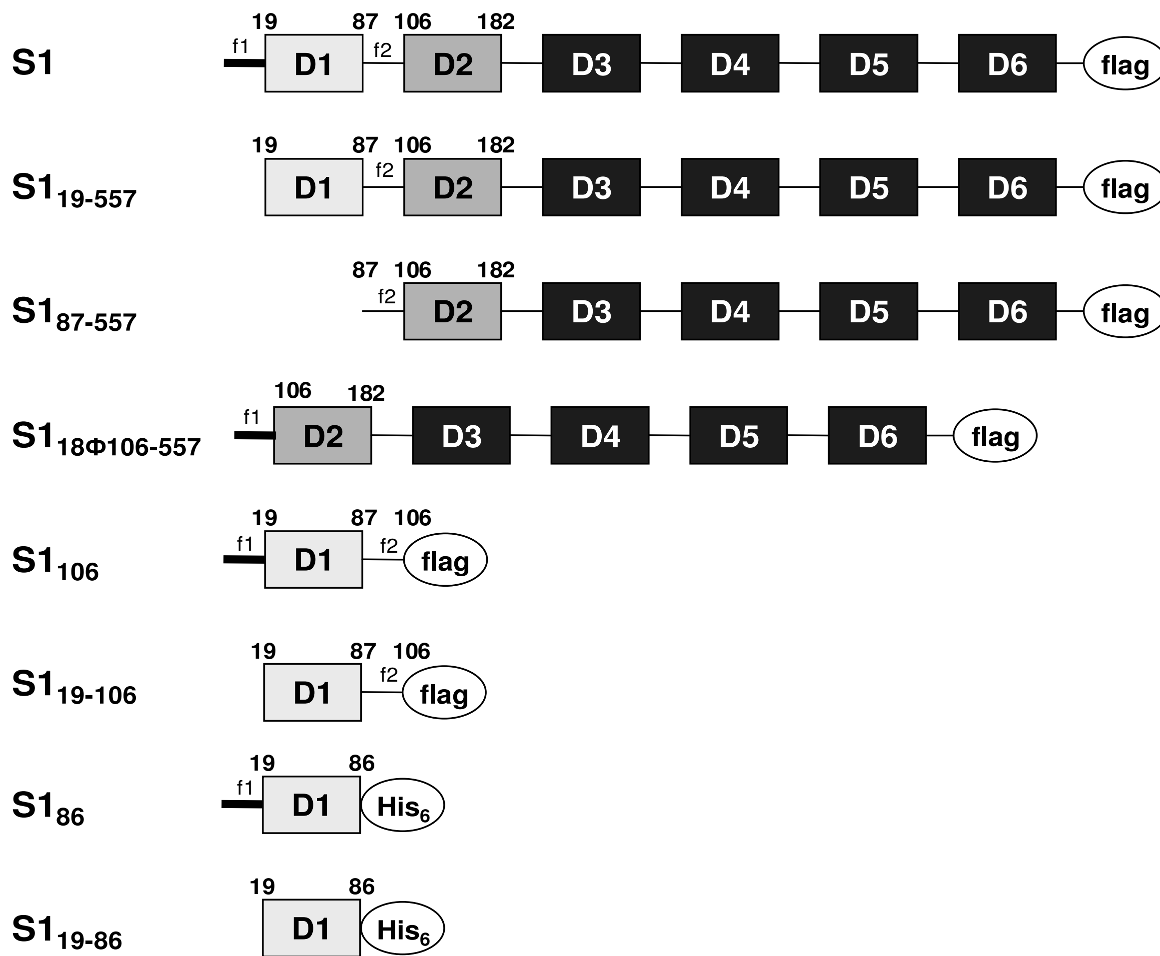
Figure S2. Meta-structure analysis of protein S2.

Meta-structure analysis of full-length protein S2 (A) and its coiled-coil domain (with logarithmic scale along Y-axis) (B). The residues which were mutated for testing interaction with protein $S1_{194}$ are underlined by red lines. Another two pairs of glycine residues within the primary structure are underlined by black and blue lines.

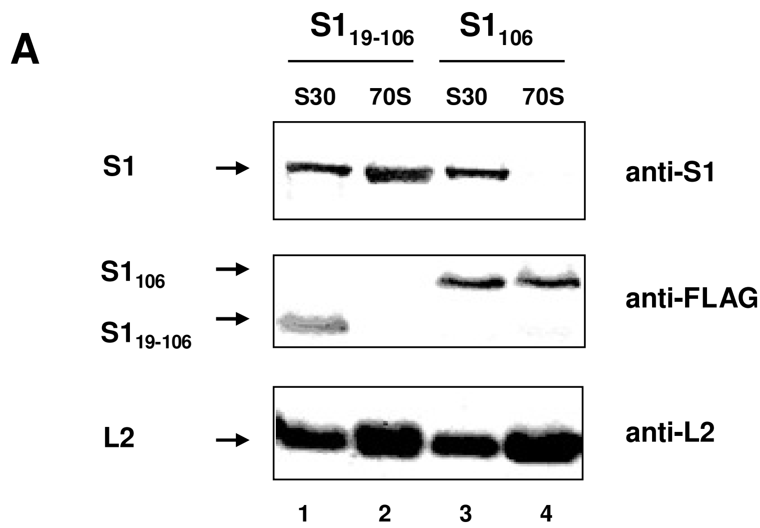
Byrgazov et al., Fig.1

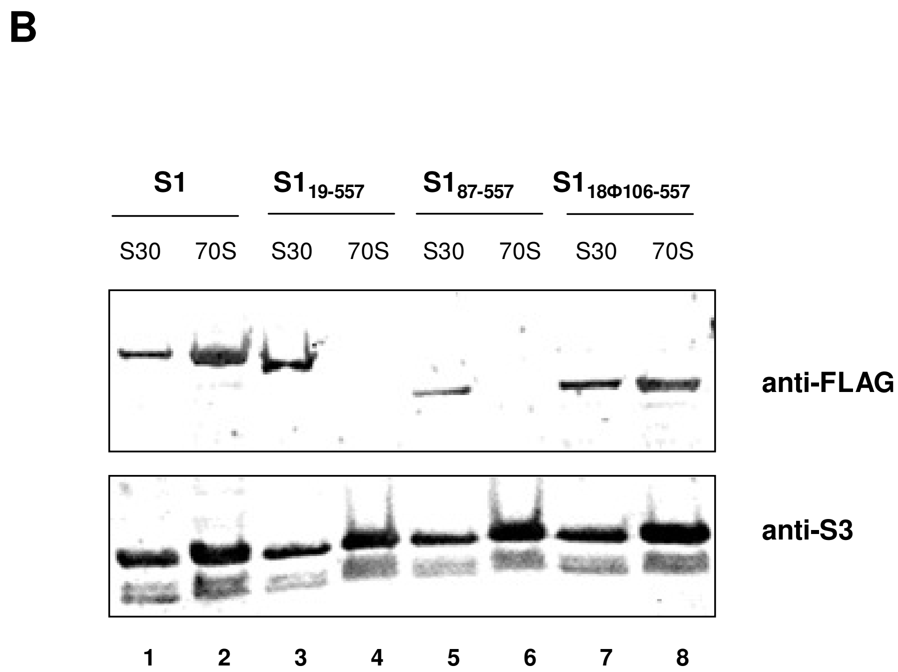


Byrgazov et al., Fig.2

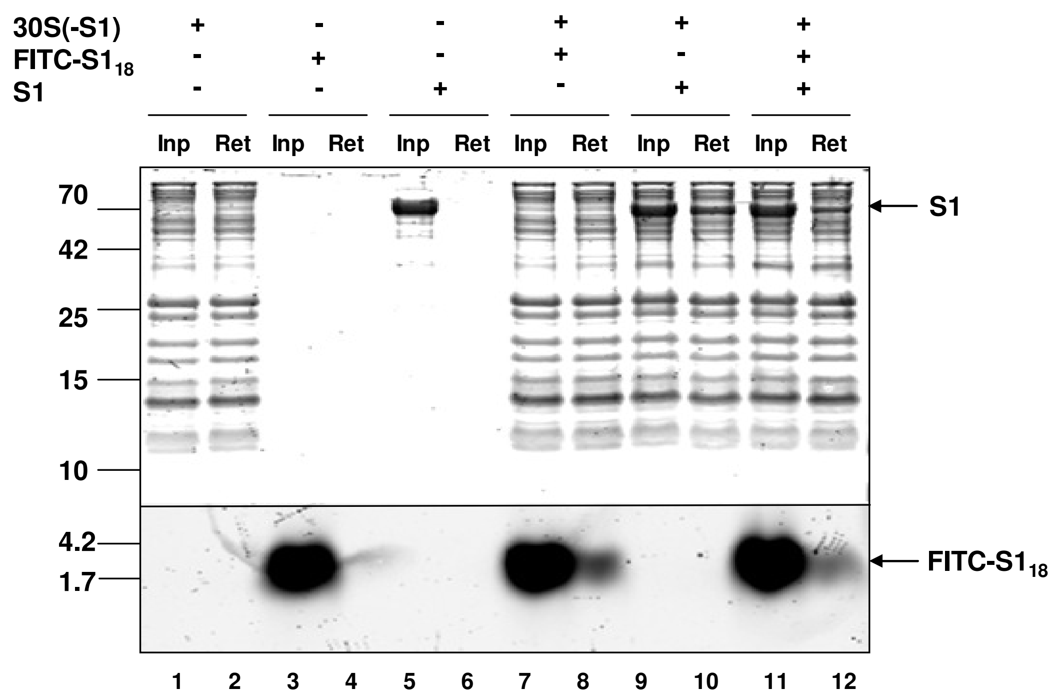


Byrgazov et al., Fig.3

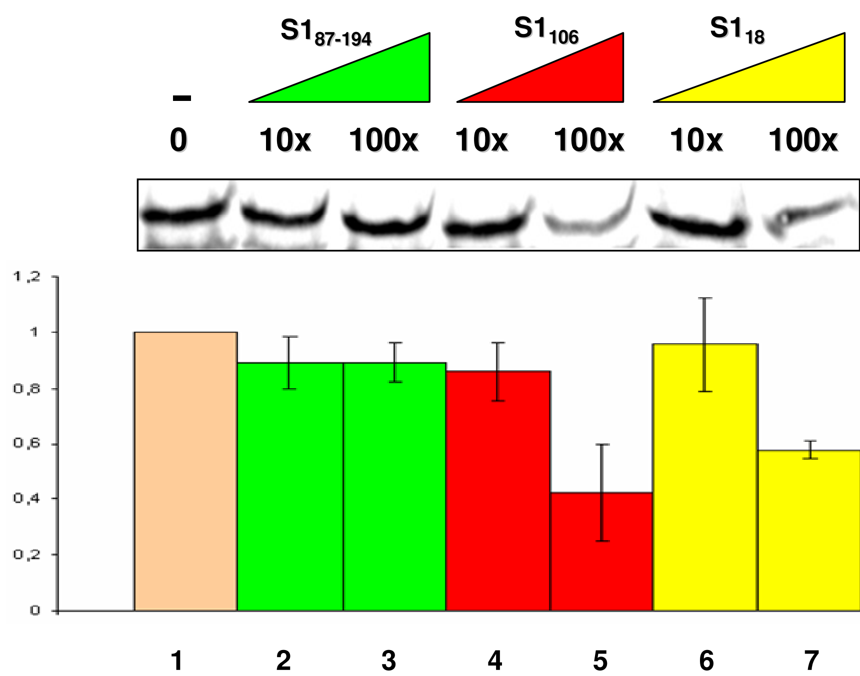




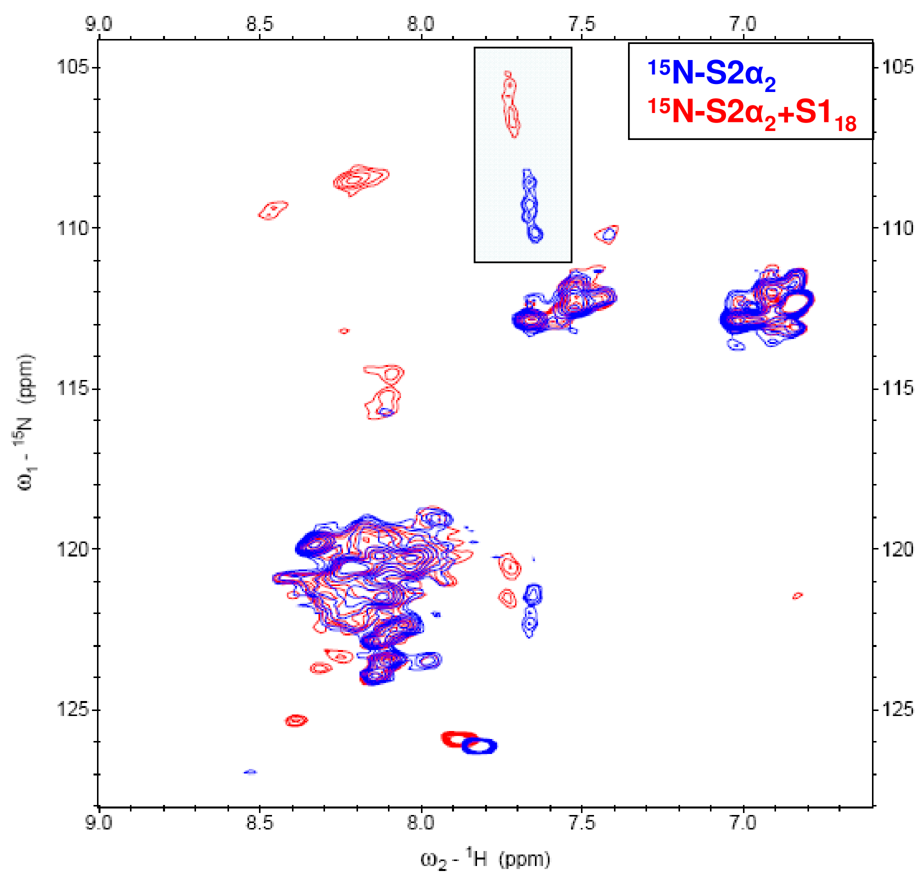
Byrgazov et al., Fig.4



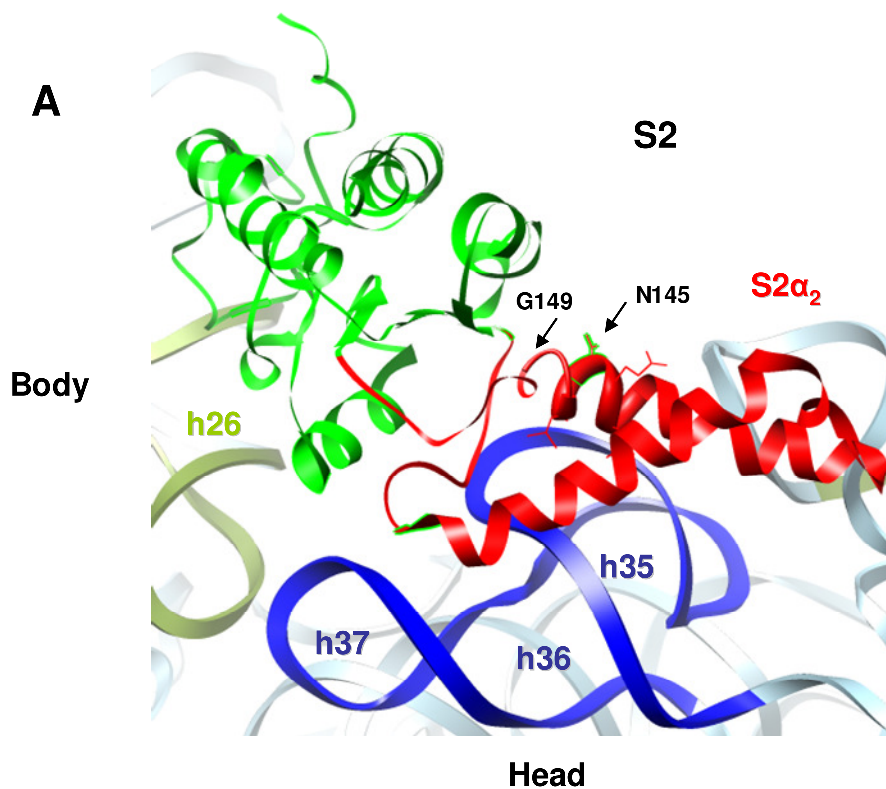
Byrgazov et al., Fig.5

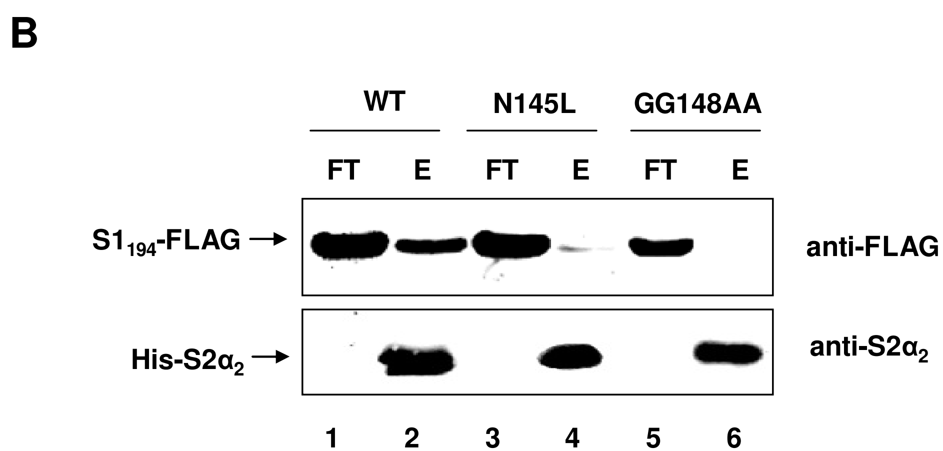


Byrgazov et al., Fig.6

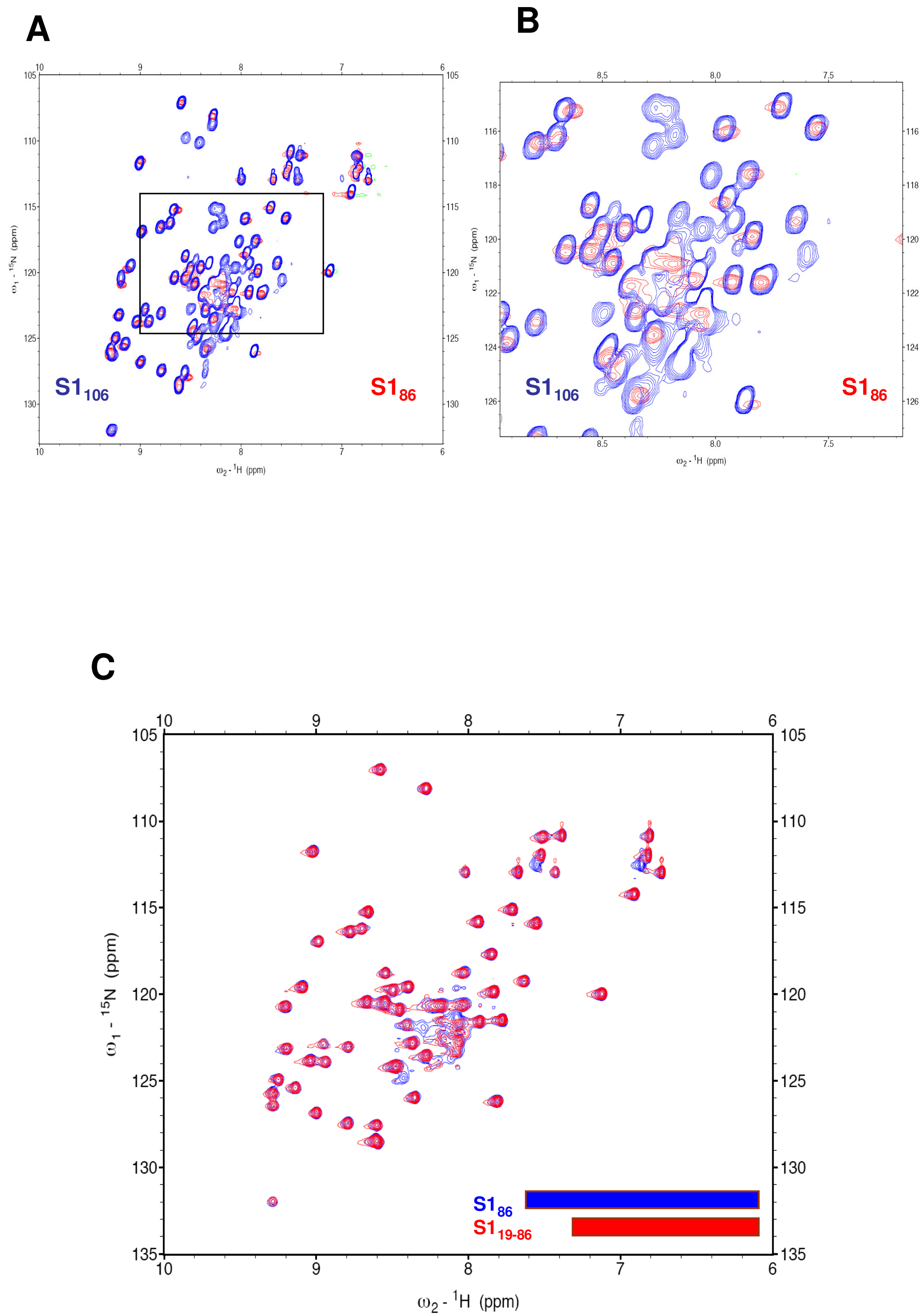


Byrgazov et al., Fig.7

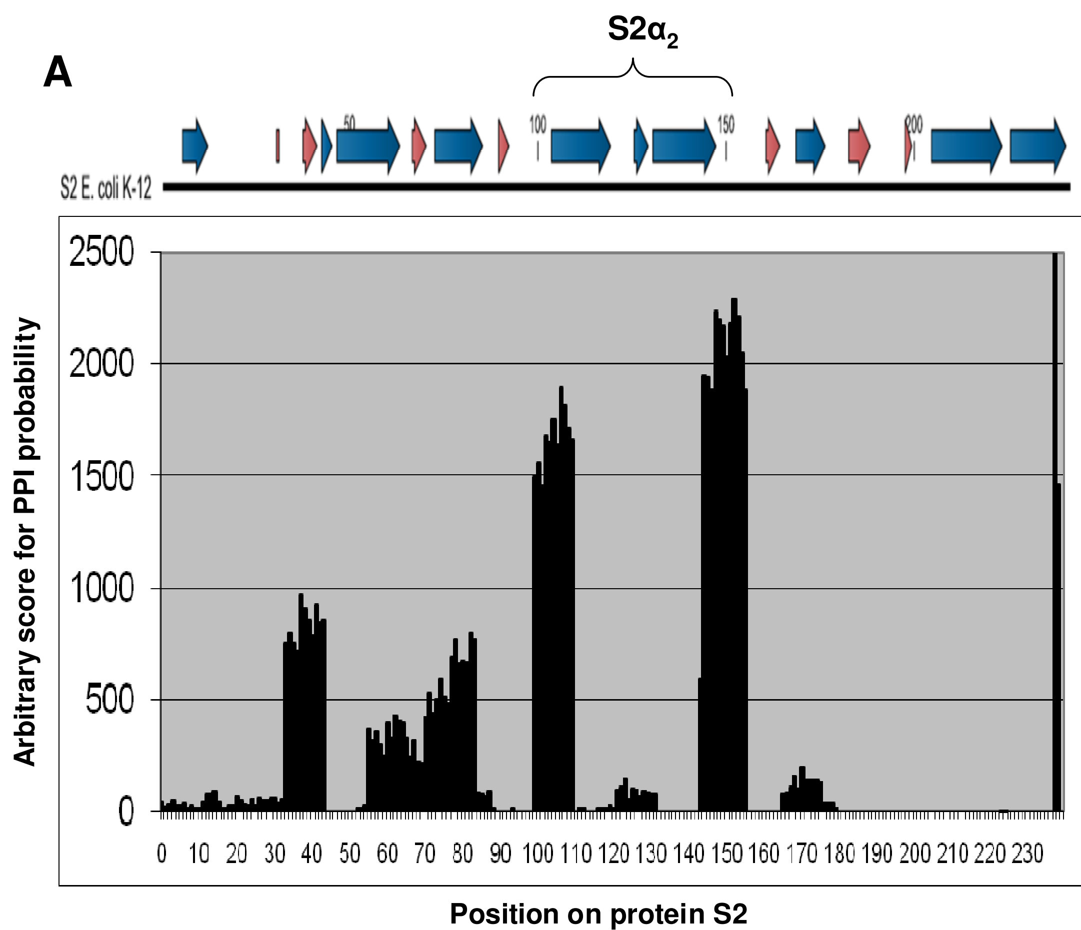


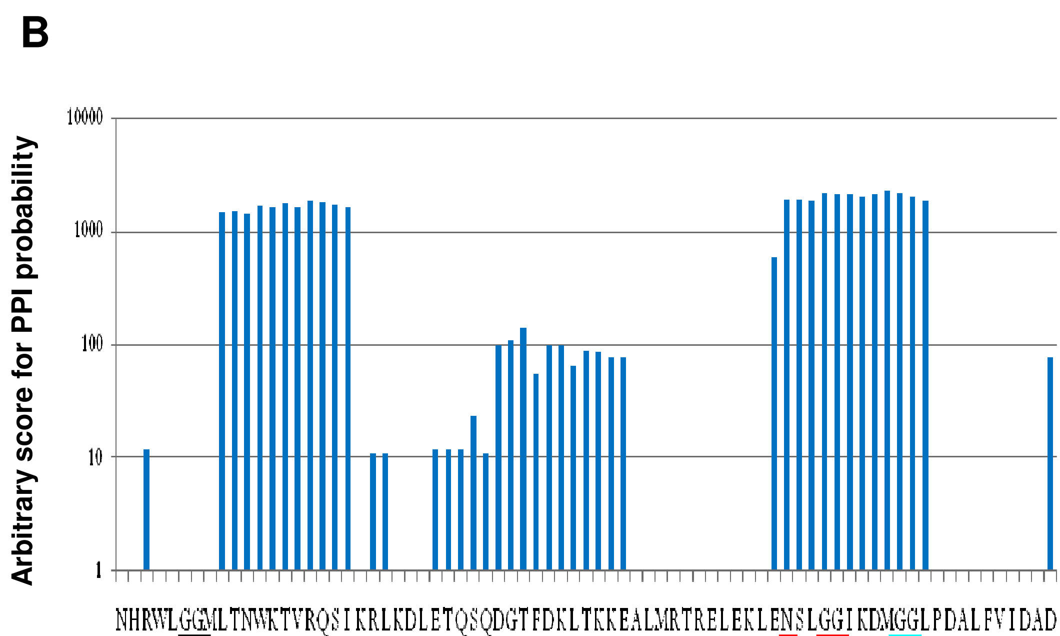


Byrgazov et al., Fig.S1



Byrgazov et al., Fig.S2





References

Aliprandi, P., Sizun, C., Perez, J., Mareuil, F., Caputo, S., Leroy, J.L., Odaert, B., Laalami, S., Uzan, M., and Bontems, F. (2008). S1 ribosomal protein functions in translation initiation and ribonuclease RegB activation are mediated by similar RNA-protein interactions: an NMR and SAXS analysis. J Biol Chem. *283*, 13289-13301.

Boni, I.V., Isaeva, D.M., Musychenko, M.L., and Tzareva, N.V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. Nucleic Acids Res. *19*, 155-162.

Boni, I.V., Zlatkin, I.V., and Budowsky, E.I. (1981) Ribosomal protein S1 associates with Escherichia coli ribosomal 30-S subunit by means of protein-protein interactions. Eur J Biochem. *121*, 371-376.

Brodersen, D.E., Clemons, W.M. Jr., Carter, A.P., Wimberly, B.T., and Ramakrishnan, V. (2003). Phasing the 30S ribosomal subunit structure. Acta Crystallogr D Biol Crystallogr. *59*, 2044-50.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M., and Murzin, A.G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. Cell. *88*, 235-242.

Byrgazov, K., Manoharadas, S., Chao Kaberdina, A., and Moll, I. *In preparation*

Ederth, J., Mandava, C.S., Dasgupta, S., and Sanyal, S. (2008). A single-step method for purification of active His-tagged ribosomes from a genetically engineered Escherichia coli. Nucleic Acids Res. *37*(2).

Gribskov, M. (1992). Translational initiation factors IF-1 and eIF-2 alpha share an RNA-binding motif with prokaryotic ribosomal protein S1 and polynucleotide phosphorylase. Gene. *119*, 107-111.

Konrat, R. (2009). The protein meta-structure: a novel concept for chemical and molecular biology. Cell Mol Life Sci. *66*, 3625-3639.

McGinness, K.E., and Sauer, R.T. (2004). Ribosomal protein S1 binds mRNA and tmRNA similarly but plays distinct roles in translation of these molecules. Proc Natl Acad Sci U S A. *101*, 13454-13459.

Moll, I., Grill, S., Gründling, A., and Bläsi, U. (2002). Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. Mol Microbiol. *44*, 1387-1396.

Osterberg, R., and Sjöberg, B. (1978). Small-angle X-ray scattering study of the proteins S1, S8, S15, S16, S20 from Escherichia coli ribosomes. FEBS Lett. *93*, 115-119.

Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., and Bontems, F. (2009). Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. Nucleic Acids Res. *37*, 5578-88.

de Smit, M.H., and van Duin, J. (1994). Control of translation by mRNA secondary structure in Escherichia coli. A quantitative analysis of literature data. J Mol Biol. *244*, 144-150.

Sørensen, M.A., Fricke, J., and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. J Mol Biol. *280*, 561-569.

Subramanian, A.R. (1983). Structure and functions of ribosomal protein S1. Prog Nucleic Acid Res Mol Biol. *28*, 101-42.

Suryanarayana, T., and Subramanian, A.R. (1983). An essential function of ribosomal protein S1 in messenger ribonucleic acid translation. Biochemistry. *22*, 2715-2719.

Sykes, M.T., and Williamson, J.R. (2009). A complex assembly landscape for the 30S ribosomal subunit. Annu Rev Biophys. *38*, 197-215.

Tedin, K., Resch, A., and Bläsi, U. (1997). Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5' leader sequence. Mol Microbiol. *25*, 189-199.

Walleczek, J., Albrecht-Ehrlich, R., Stöffler, G., and Stöffler-Meilicke, M. (1990). Three-dimensional localization of the NH2- and carboxyl-terminal domain of ribosomal protein S1 on the surface of the 30 S subunit from Escherichia coli. J Biol Chem. *265*, 11338-11344.

7. Acknowledgements

The many thanks to my wife for the support at the difficult minute.

Enormous thanks to my parents and to sister.

Thanks to my scientific leader for valuable discussions.

I am also grateful to all folks, who supported me for all these years.

Большое спасибо моей жене за поддержку в трудную минуту.

Огромное спасибо моим родителям и сестре.

Спасибо моему научному руководителю.

Я также благодарен всем тем, кто поддерживал меня в течение всех этих лет.

8. Curriculum Vitae

KONSTANTIN BYRGAZOV

Schelleingasse 14-16/3/11 • Vienna 1040 • Austria • (+43)664-260-5805 konstantin.byrgazov@univie.ac.at

SUMMARY

Dynamic, bilingual Ph.D. with strong academic background in Molecular (RNA) Biology and Chemistry. Adeptly balances attention to detail and analysis with high-level strategic perspective. Utilizes strong interpersonal skills to build enduring relationships with collaborators and colleagues. Self-motivated team player with strong leadership potential who thrives in intellectually stimulating environments.

EDUCATION

University of Vienna **Vienna, Austria**
Ph.D. Molecular Biology 2007-2011

- **Thesis:** Deciphered breakthrough antibiotic design against Gram-negative bacteria by studying interaction of regulator protein S1 with ribosome.
- **Publications:**
  1. Byrgazov K., Manoharadas S., Kaberdina A., and Moll I. "Only the first N-terminal domain of protein S1 is required for binding to the ribosome", *in preparation*.
  2. Byrgazov K., Coudevylle N., Konrat R., and Moll I. "Structural studies of the N-terminal domain of protein S1: Flexible parts are important for interaction with the ribosome", *in preparation.*
- **Distinctions:** Recipient, RNA Biology Doctorate Programme Fellowship.
- **Relevant Coursework:** "RNA Biology Journal Club", "Presentation in English", "Scientific writing in English", "Project Management".
- **Activities:** Regular participant (with poster/oral presentation), numerous international conferences and workshops devoted to RNA Biology field.Committee Member, Max F. Perutz Laboratories Ph.D. and PostDoc Retreat (organized sponsors and speakers). Vice-captain of MFPL Dragon Boat team (annual team-building event).

Lomonosov Moscow State University **Moscow, Russia**
M.S. in Chemistry. 2002-2007

- **Thesis**: Synthesis and characterization of oligosaccharides, human blood group B antigens.
- **Distinctions:** GPA (3.95/4.0). Top 5 students. Diploma with Honors. GRE Chemistry above 95%. Published Masters thesis
- **Publication:** Korchagina E., Byrgazov K, Ryzhov I, Popova I, and Bovin N. "Block Synthesis of Blood Group Tetrasaccharides B". *Mendeleev Comm.*, 3(19), 2009.
- **Relevant coursework:** Chemistry (Biochemistry, Organic, Inorganic, Analytical, Physical and Medicinal; lectures and practical courses), Physics (lectures and practical courses), Mathematics, Statistics.
- **Activities:** Participant, international conferences devoted to Immunology, and Carbohydrate Chemistry.

EXPERIMENTAL EXPERIENCE

- Molecular Cloning, Western and Nothern Blot Analysis, Primer Extension, Toeprinting Analysis, SHAPE Analysis, RNA Sequencing, Protein Expression and Purification, Pull Down, Immunoprecipitation, In vitro Translation, Pulse Labeling, ITC, DLS, NMR techniques: HSQC, NMR-based structure calculation
- ELISA, Antibody Purification, Chromatography (Affinity, Size-Exclusion, IEX), Organic Synthesis and Purification of Small Molecules, NMR analysis: ^{1}H, NOESY, TOEXY

PERSONAL

- **Languages:** English (Fluent). Russian (Fluent). German (Intermediate).
- **Interests:** RNA Biology. Playing volleyball, basketball, badminton. Reading business magazines, newspapers, novels.

.