IDENTIFICATION AND CHARACTERIZATION OF NOVEL RIBOSOMAL PROTEIN-BINDING RNA MOTIFS IN BACTERIA

a thesis

by

YANG FU

submitted in partial fulfillment of the requirements

for the degree of

Master of Science

May 2014
As the factory responsible for producing proteins, ribosomes are of great importance. In bacteria, ribosomes are composed of three ribosomal RNAs (rRNA) of different sizes, and around 50 ribosomal proteins (r-protein). During ribosome biogenesis in bacteria, synthesis of rRNAs and r-proteins are both tightly regulated and coordinated to ensure robust growth. In particular, a group of cis-regulatory RNA elements located in the 5′ untranslated regions or the intergenic regions in r-protein operons are responsible for the regulation of r-protein biosynthesis. Based on the fact that RNA-regulated r-protein biosynthesis is essential and universal in bacteria, such unique and varied regulatory RNAs could provide new targets for antibacterial purpose. In this thesis, we report and experimentally verify a novel r-protein L1 regulation model that contains dual L1-binding RNA motif, and for the first time, a S6:S18 dimer-binding RNA structure in the S6 operon. We also describe *Escherichia coli*-based and *Schizosaccharomyces pombe*-based reporter systems for in vivo characterization of RNA-protein interactions. So far, both in vivo systems failed to report RNA-protein interactions, and thus need further tuning. In addition, we performed phage-display to select for regulatory RNA-binding small peptides and examined their effects on bacteria viability. One selected peptide, TVNFKLY-C, caused defective growth when overexpressed in *E. coli*. Yet, further studies must be conducted to verify the possibility that bacteria were killed by direct RNA-peptide interaction that disrupted the native r-protein regulation.
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>ii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>Chapter 1</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>7</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>12</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>20</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>26</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>32</td>
</tr>
<tr>
<td>References</td>
<td>63</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank Dr. Meyer for being a great mentor to me and bearing with me during my down times. I would like to thank the members of the Meyer lab, Betty, Ari, and Shermin, for providing me help of all kinds. I honestly feel so blessed to have you guys around to work with and to hang out with through all these year. I also want to thank Dr. Hoffman for teaching me all the yeast knowledge. Lastly, thank you to Dr. Johnson and Dr. Annunziato, for being my committee members and for their advice and suggestions.
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>An overview of ribosomal protein regulation</td>
<td>41</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Phylogenetic distribution of <em>E. coli</em> r-protein regulatory RNAs</td>
<td>42</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Regulation models of L1 biosynthesis in <em>E. coli</em> and Archaea</td>
<td>43</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Consensus sequence and secondary structure for r-protein L1 regulatory RNA</td>
<td>44</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Multiple regulation model of ribosomal protein L1</td>
<td>45</td>
</tr>
<tr>
<td>Figure 6</td>
<td>L1-binding RNA structures from <em>G. kaustophilus</em></td>
<td>46</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Known transcriptional organizations of the S6 operon</td>
<td>47</td>
</tr>
<tr>
<td>Figure 8</td>
<td>S6 operon-associated RNA secondary structures</td>
<td>48</td>
</tr>
<tr>
<td>Figure 9</td>
<td>rpsF_5-69 titrations</td>
<td>49</td>
</tr>
<tr>
<td>Figure 10</td>
<td>rpsF_5-69 titrations with one of S6:S18 at constant concentration</td>
<td>50</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Mutant RNA titrations</td>
<td>51</td>
</tr>
<tr>
<td>Figure 12</td>
<td>S6:S18 preferentially interacts at H2 of predicted RNA structure</td>
<td>52</td>
</tr>
<tr>
<td>Figure 13</td>
<td>SHAPE experiment for rpsF_1-69</td>
<td>53</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Varied reporter activities seen with <em>E. coli</em> S15-RNA interaction in <em>S. pombe</em></td>
<td>54</td>
</tr>
<tr>
<td>Figure 15</td>
<td>A schematic of the dual-luciferase reporter plasmid</td>
<td>55</td>
</tr>
</tbody>
</table>
Figure 16. Varied reporter activities seen with \textit{G. kaustophilus} L1-RNA interaction in \textit{E. coli}.

Figure 17. \textit{G. kaustophilus} RNAs used in the phage display selections.

Figure 18. \textit{E. coli} BL21(DE3) growth curves for peptide overexpression.

Figure 19. \textit{E. coli} BL21(DE3) growth curves for r-protein overexpression.
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Primers for RACE</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>Phage display selection setups</td>
<td>61</td>
</tr>
<tr>
<td>3</td>
<td>Selected peptides</td>
<td>62</td>
</tr>
</tbody>
</table>
Chapter 1:

Introduction
Ribosomal protein feed-back regulation in bacteria

Prokaryotic and eukaryotic ribosomes differ in their size, composition and the ratio of ribosomal protein to ribosomal RNA (Ramakrishnan et al., 2002; Shajani et al., 2011). They employ different mechanisms to regulate and coordinate the biosynthesis of ribosomal RNAs and ribosomal proteins (Nomura, 1999). Unlike the separate and direct regulation of both rRNA and ribosomal protein syntheses in eukaryotes, in bacteria, the synthesis of ribosomal protein is feed-back regulated in response to rRNA synthesis (Nomura et al., 1984; Nomura, 1999). Genes coding for ribosomal proteins in bacteria are typically co-located in the genome and expressed as large transcriptional units (operons) with strong translational coupling (Jaskunas et al., 1977; Lindahl et al., 1977). The expression of ribosomal proteins is often negatively regulated at the transcriptional or translational level by one or more of the operon products that binds to a structured RNA at the 5'-UTR or intergenic region of the mRNA associated with the operon (Zengel et al., 1994). These non-coding RNAs that bind to excess ribosomal proteins are termed cis-regulatory elements, and are responsible for keeping the ribosomal RNA and protein levels in balance (Nomura et al., 1999). Structural studies have revealed that some of the regulatory elements possess secondary structures that are mimicries of ribosomal protein binding sites on rRNAs in assembled ribosomes, but have lower binding affinity to the same ribosomal protein (Choonee et al., 2007; Guillier et al., 2002; Köhrer et al., 1998; Serganov et al., 2003). A free ribosomal protein preferentially binds to the primary target, ribosomal RNA. However, under conditions where rRNA sites are completely saturated, it binds to the regulatory site on its own mRNA, blocking further ribosomal protein synthesis, and normalizing the disturbed balance between ribosomal proteins and rRNAs (Figure 1).
Ribosomal protein synthesis in bacteria is phylum-specific

As one of the most studied bacterial species, *Escherichia coli* has 12 distinct RNA regulatory elements that control the synthesis of more than half of its ribosomal proteins (Aseev et al., 2008; Choonee et al., 2007; Lindahl et al., 1977; Zengel et al., 1994). Each of them is a structure on the mRNA that is capable of binding to a ribosomal protein to down-regulate gene expression. This feedback regulation has been found to be accomplished via a variety of mechanisms. For example, *E. coli* S4 mediates expression repression of the α operon by ribosomal entrapment (Schlax et al., 2001); *E. coli* L4 represses synthesis of the 11-gene S10 operon by premature transcription termination (Stelzl et al., 2003); *E. coli* L20 down-regulates expression of itself and L35 through ribosome-binding site occlusion (Guillier et al., 2002).

Despite the fact that the regulation of ribosomal protein synthesis by cis-regulatory RNA elements is universal in bacteria, the organization of operons that encode ribosomal proteins and the distribution of RNA elements in each operon, as well as structures of those elements, are not well conserved across different bacteria phyla. Instead, they are likely to be both distinct and diverse. Comparative genomic studies have revealed that only three of the RNA regulatory elements in *E. coli* are widespread across the bacterial world (Figure 2) (Fu et al., 2013). These three widely distributed RNA elements bind ribosomal proteins L1, L10, and S2, and most bacteria have at least one of them. Moreover, these wide-spread RNAs are also found to be both structurally and functionally conserved. For example, S2 proteins from *E. coli* can interact with S2 regulatory RNAs from *Yersinia pestis*, *Haemophilus influenza* and *Pseudomonas aeruginosa*, all of which are also under Proteobacteria phylum (Aseev et al., 2008). As for the narrowly distributed *cis*-regulatory RNA elements category, under Proteobacteria phylum, most of them can be found in only 4 orders of bacteria (Figure 2) (Fu et al., 2013). In addition to *E. coli*,
in *Bacillus subtilis*, a representative bacteria within the Firmicutes phylum, RNA elements binding to L20, S15 and S4 proteins have been experimentally confirmed to be structurally and functionally Firmicutes-specific (Choonee et al., 2007; Grundy et al., 1991; Scott et al., 2005). For example, *cis*-regulatory RNA elements of ribosomal protein S15 from *E. coli* and *Bacillus stearothermophilus* are distinct in both their primary sequences and secondary structures, and they both down-regulate S15 expression level in response to excess S15 in their respective organisms (Philippe et al., 1990; Scott et al., 2001). However, *E. coli* S15 cannot bind to regulatory RNA from *B. stearothermophilus* both *in vitro* and *in vivo*, and vice versa (Scott et al., 2005). More recent comparative genomic studies have also discovered more putative RNA elements that are only found within the Firmicutes phylum (Weinberg et al., 2007, 2010; Yao et al., 2007). Since RNA-based regulation of ribosomal proteins remain vastly unexplored in most Eubacteria, more similarities and diversities among different phyla will be exposed in the future, providing more insights into bacterial ribosomal protein biosynthesis and bacterial evolution.

**Control of ribosomal protein biosynthesis as a viable antibacterial target**

Considering the vital role of the ribosome in protein biosynthesis, it is an important antibacterial target for a wide variety of antibiotics (Brodersen et al., 2000). Currently, it is targeted by seven different classes of antibiotics used clinically, most of which function at the large ribosome subunit and rRNAs (Knowles et al., 2002; Yonath, 2005). More specifically, these antibiotics target the ribosome at regions of functional importance, including the decoding site, the peptidyl-transferase center (PTC) and the protein exit tunnel. They are designed to interfere with tRNA at the decoding site (Brodersen et al., 2000), as well as at the PTC (Carter et al., 2000), and block the protein exit tunnel (Schlünzen et al., 2001). However, with the increased use of antibiotics, pathogenic bacteria acquiring antibiotic resistance has become a serious problem.
For example, tetracycline, which targets at the A site of 16S rRNA on the small subunit, was of great clinical importance when it was first discovered, but unfortunately strains of bacteria with resistance to these drugs have become commonplace, limiting its effectiveness (Goldmann, 1992; Roberts, 2005). At the same time, many of the currently used antibiotics have insufficient specificity toward a narrow group of bacteria. These issues call for the identification of new drug targets.

Aside from targeting bacterial ribosomes at the rRNAs and the large subunit, there are currently no antibiotics acting on the regulatory RNAs involved in the biosynthesis of ribosomal proteins. As mentioned previously, cis-regulatory RNA elements regulating ribosomal protein biosynthesis is universal across the bacterial world. Tight and accurate regulation is essential for correct ribosome assembly, thus resulting in robust bacteria growth (Nomura, Yates, & Dean, 1980). In addition to the fact that growth defects associated with deregulation of ribosomal proteins led to the discovery of cis-regulatory RNA elements in E. coli (Dean et al., 1980), we have also observed bacterial growth defects caused by regulated over-expression of some ribosomal proteins in our own studies. Thus, antibiotics that can simulate the protein-RNA-binding feedback regulation of ribosomal protein synthesis can lead to repressed ribosomal protein gene expression, and consequently inhibit bacterial growth. Moreover, unlike the widely distributed bacteria riboswitches, the cis-regulatory RNA elements involved in ribosomal protein regulation are both widely and narrowly distributed, and thus, can be generally or specifically targeted for different purposes. For example, the RNA element for L1 synthesis regulation could be a universal target for anti-bacterial pharmaceuticals, while the RNA element regulating S15 synthesis could be used to specifically target bacteria species within a certain phylum.
Objectives of this thesis

The first objective of this thesis was to experimentally confirm the predicted Firmicute-specific regulation model that includes dual L1 cis-regulatory RNA element. Rapid Amplification of cDNA Ends (RACE) would help examine the transcriptional organization of L1 operon in Firmicutes. The RNA-L1 ribosomal protein interactions were examined in vitro with filter-binding assays.

The second objective was to experimentally validate the predicted S6-operon related RNA structure. Again, RACE was used to examine the transcriptional organization of this operon. Then in vitro methods were used to study the interaction between the predicted RNA structure and its potential protein partners. In addition, nuclease protection assays and SHAPE (2’-hydroxyl acylation analyzed by primer extension) were performed to investigate the RNA-protein interaction dynamics and to identify potential interaction sites.

The third objective was to construct E. coli-based and Schizosaccharomyces pombe-based reporter systems to study RNA-protein interactions in vivo and to test their feasibility as potential high throughput screening platforms for RNA-binding compounds.

The last objective of this thesis was to use M13 phage display technique to select for small peptides that specifically bind to cis-regulatory RNA elements and explore the use of such RNA elements as potential antibacterial targets. Selected peptides were tested in E. coli for their effects on bacteria viability.
Chapter 2:

Ribosomal protein L1 operon in Firmicutes bacteria carries dual L1-binding sites
**Ribosomal protein L1 regulatory models**

L1 is one of the primary RNA-binding ribosomal proteins in ribosome assembly. It specifically associates with 23S rRNA with high affinity, and is one of the first ribosomal proteins recruited to the 50S ribosomal subunit during ribosome assembly (Köhrer et al., 1998). L1 protein is strikingly conserved among most Eubacteria and Archaea in both structure and function (Gourse et al., 1981; Nikulin et al., 2003). In addition to the interaction with the rRNA, L1 also binds to its own mRNA through a conserved RNA structure, acting as a translational repressor. L1-binding RNAs have been thoroughly studied in both *E. coli* (Baughman et al., 1984; Zengel et al., 1994) and some Archaea species (Ameres et al., 2007; Köhrer et al., 1998). Regulation models in *E. coli* and Methanococcus containing a single L1-binding site have been generated from previous studies. The L1-binding site in these models exhibits high similarity in both primary sequence and secondary structure to the L1 binding site on the 23S rRNA (Köhrer et al., 1998; Kraft et al., 1999; Zengel et al., 1994). However, the genomic location of this site varies (Figure 3). In *E. coli*, the gene coding for L1, *rplA*, is preceded by that which encodes L11, *rplK*. The two genes are co-regulated and the L1-binding site is in the 5’-UTR preceding *rplK*. As a result, in response to excess L1, this RNA element interacts with L1 and inhibits translation of both L11 and L1. In Archaea, *rplA* and *rplK* are not co-located in the genome. Instead, *rplA* is followed by *rplJ* and *rplL* in the same operon, and the L1-binding site directly precedes *rplA*. As mentioned before, L1 *cis*-regulatory RNA is one of the three RNA elements that are widespread across the bacterial world. L1 RNA motifs in several organisms have been experimentally proven to be structurally and functionally equivalent to each other. For example, *E. coli* L1 is able to inhibit translation of L1 mRNA from *M. vannielii*. Conversely, *M. vannielii* L1 can inhibit synthesis of both L11 and L1 proteins of *E. coli* (Köhrer et al., 1998).
**L1 model validation**

**Overview of the consensus L1 RNA structure**

Computational studies done in our lab revealed a L1-binding consensus RNA structure across many bacteria phyla that contains a stem (6-12 bp) -loop and an internal bulge with a G-A pair and an unpaired adenosine (Figure 4) (Fu et al., 2013). These nucleotides are known to be involved in L1-binding in both *E. coli* and *Methanococcus Vannielli* (Zengel, 1994; Kraft et al., 1999; Nevskaya et al., 2006). In general, the structure of this consensus RNA is in good agreement with previous studies (Köhrer et al., 1998; Kraft et al., 1999). In addition, our data also show interesting distribution patterns of such structured RNA (Figure 5). In accordance with previous studies, one copy of the RNA structure is found in Proteobacteria, represented by *E. coli* (Baughman et al., 1984), and it precedes *rplK*, whereas the one copy found in Actinobacteria is located in between *rplK* and *rplA*, similar to that found in Archaea (Hanner et al., 1994). Surprisingly, two copies of the same structure are found in approximately 40% of Firmicutes, represented by *Geobacillus kaustophilus* and *Bacillus subtilis*, suggesting a potential novel bacterial L1 regulation model.

**Transcriptional organization**

Since our predicted novel L1 regulation model has two copies of the structured RNA, one preceding *rplK* and the other locating in the intergenic region between *rplK* and *rplA*, we conducted 5’ RACE on total RNA extracted from *B. subtilis* to investigate the transcriptional organization of this model. Our results indicate that *rplK* and *rplA* are co-transcribed from the same promoter. In addition, almost half of the reverse transcripts include both predicted RNA structures, suggesting that both sites are involved in the regulation of L1 and L11 biosynthesis in Firmicutes.
Both RNAs interact with ribosomal protein L1 from *G. kaustophilus* in vitro

To further ensure that the dual RNA structures have the L1-binding ability, we performed filter-binding assays on ribosomal protein L1 and the two RNAs derived from *G. kaustophilus*. Our results show that L1 ribosomal protein binds to both RNA structures in a concentration-dependent manner ($K_D$ 25-50 nM) (Figure 6) (Fu et al., 2013).

To assess the specificity of the RNA-protein interaction, we also made mutant copies of both RNAs. Previous studies on L1 regulation show that the L1-RNA interaction can be destroyed by replacing two guanines with adenines in the big bulge region (Köhrer et al., 1998; Kraft et al., 1999). Our data support the same specificity as the RNA-protein interaction was abolished by substituting G5 and G7 with As on both predicted RNA structures (Figure 6) (Fu et al., 2013). In general, all the results generated from filter-binding assays further verified that the two predicted RNA structures are biologically relevant and both are involved in L1 regulation in Firmicutes.

**Discussion**

In this study, we experimentally validated a novel Firmicute-specific ribosomal protein L1 regulation model predicted previously during our computational study. Transcriptional organization of this model was confirmed by 5’ RACE and dual RNA-L1 protein interaction was confirmed by filter-binding assays.

We show a novel regulation model of ribosomal protein L1 that carries two conserved L1-binding RNA motifs, both of which interact with L1 specifically *in vitro*. This model is seen in more than 40% Firmicutes with complete genomes and may present a new mechanism in which
ribosomal protein L1 and L11 ORFs are in the operon but are regulated separately. Despite its widespread distribution and universal regulation, the genomic locus of L1-binding RNA is not conserved within Eubacteria, suggesting that this RNA-mediated regulation might have evolved from multiple inventions. This finding broadens our knowledge of the currently known models of L1 regulation in Eubacteria and Archaea (Fu et al., 2013). The previous neglect of such novel model may be due to the small size and minimal sequence conservation of the L1-interacting motif, it is likely that there are other undiscovered L1-regulation models associated with certain bacterial phyla still to be uncovered by more updated databases and computational tools.
Chapter 3:

Bacterial RNA motif in ribosomal protein S6 operon interacts with an S6:S18 dimer
Ribosomal proteins S6 and S18 regulation

Ribosomal proteins S6 and S18 are building blocks of the 30S small ribosomal subunit, but unlike L1, they are not primary ribosomal proteins. They are recruited to the 16S rRNA following ribosomal protein S15, and as a result, interact with 16S rRNA as well as S15 (Recht et al., 2001). Previous studies have also shown that S6:S18 bind to the rRNA as a protein dimer (Recht et al., 2001). However, the regulatory mechanism of S6 and S18 biosynthesis is still unclear. In bacterial genomes, rpsF and rpsR, genes coding for ribosomal proteins S6 and S18, are usually located in the same operon. For example, in E. coli and Thermus thermophilus, rpsF appears as the first gene in such a operon, followed by priB, rpsR, and rplL, the gene coding for ribosomal protein L9 (Figure 7) (Isono et al., 1978). In bacteria under Firmicute phyla, such as Bacillus subtilis and Geobacillus kaustophilus, the genomic localization of rpsF and rpsR is very similar whereas rplL is located somewhere else (Figure 7) (Akanuma et al., 2012). Previously, comparative genomic studies have shown that a conserved RNA motif can be found in the 5’-UTR region of the S6-related operon in Firmicutes (Yao et al., 2007). Thus, it is tempting to speculate that such RNA motif is another cis-regulatory RNA and responsible for S6 and S18 regulation.

S6 model validation

Overview of the predicted S6 consensus RNA structure

Computational studies done in our lab reveal a consensus RNA structure at the 5’-UTR of S6 operon across many bacteria phyla (Figure 8) (Fu et al., 2014). The consensus structure has a stem-loop with an internal bulge containing a conserved pair of cytosines. The H1 stem sequesters the Shine-Dalgarno sequence and the H2 stem (5-124 nt) and the internal bulge (7-22 nt) are variable in length. The proximity of the predicted RNA structure to the translational
start site of S6 ribosomal protein and its blocking of ribosome binding site suggest that it could be another widely distributed cis-regulatory RNA element.

**Transcriptional organization**

By performing 5′ RACE on total RNA extracted from *B. subtilis* using primers in the S6 and S18 coding regions, we identified a transcription start site 20-nt upstream of the predicted RNA structure (Figure 8) (Fu et al., 2013); we also verified the co-transcription of *rpsF, ssbA*, and *rpsR*. In addition, 3′ RACE revealed the 3′ end of the transcription, which lies 22-nt downstream of the stop codon for S18.

**Predicted RNA structure only specifically interact with S6:S18 dimer in vitro**

According to previous examples of RNA-regulated ribosomal protein biosynthesis and the transcriptional organization of this operon, we speculate that either S6 or S18 is a possible protein partner of the RNA structure. Also, previous studies show that S6 and S18 ribosomal proteins form a dimer before binding to the 16S rRNA-S15 complex during ribosome assembly (Recht et al., 2001), suggesting that S6:S18 dimer could also be a potential binding partner. The interaction between *B. subtilis* RNA and *G. kaustophilus* S6, S18, and S6:S18 dimer was examined by Electrophoretic Mobility Shift Assays (EMSA). The RNA structure transcribed from *B. subtilis* genome was truncated at the 5′ end (*rpsF* _5-69). However, we observed similar binding results by using RNA (*rpsF_1-69) transcribed from the native start site. During protein purification, we found it difficult to isolate *B. subtilis* S18 due to its insolubility and small size. Since S6 and S18 from *B. subtilis* and *G. kaustophilus* share 75% and 89% sequence identity, we think it is reasonable to substitute *B. subtilis* proteins with *G. kaustophilus* proteins in this experiment. In
addition, the RNA motifs from *G. kaustophilus* and *B. subtilis* are identical with the exception of minor differences in the sequence and length of the variable stem H2.

Our binding results show that the RNA structure specifically interacts with the S6:S18 dimer with a $K_D$ of 155 nM (Figure 9). The RNA structure interacts with S18 alone only at a high protein concentration ($K_D > 500$ nM) while it does not bind S6 alone at all. The S6:S18 binding curve exhibits a high Hill coefficient (4.1) possibly due to the insolubility of S18, which would lead to aggregation of the dimer protein. To determine if this is the cause, we added the detergent IGEPAL CA630 (0.02%) (Recht et al., 2001) to the binding buffer, but observed similar $K_D$ and Hill coefficient. We also looked into the RNA-dimer protein interaction by titrating S6 and S18 individually while keeping the other at a constant concentration (500 nM), which allows most of the S6 and S18 to dimerize before interacting with the RNA structure. As a result, we observed two curves with much smaller Hill coefficient (S6: 1.45, S18: 1.5) compared to that of the previous dimer curve (4.1) (Figure 10). Also, the $K_D$ of each binding curve is lower than that of the dimer curve. Such results suggest that the previous high Hill coefficient is more likely caused by the proportion of incorrectly folded S6:S18 dimer in the binding reactions.

To further investigate the specificity of the RNA-S6:S18 interaction, we constructed six mutant RNAs that have varied secondary structures (M1-M6) (Figure 11). M2 completely abolished the RNA-protein interaction, whereas M1, M3, and M4 significantly impaired the interaction ($K_D > 500$ nM). Such results were expected, since M1/2 and M3 were designed to compromise the H1 stem and H2 stem respectively, whereas M4 loses the conserved cytosine pair. The compensatory mutant M6 rescues the loss of binding caused by M1 and M2, suggesting that the H1 does exist as predicted. Interestingly, the other compensatory mutant M5 fails to restore the
loss of binding caused by M3, indicating that the primary sequence of the H2 area maybe critical for protein-binding.

Nuclease protection assays results

As the next step, we performed RNase V1 and A protection assays to investigate the RNA secondary structure and identify the RNA-protein interaction general regions (Figure 12). RNase A specifically cleaves single-stranded cytosines and uracils and RNase V1 only cleaves double-stranded regions. According to our results, when the RNA (rpsF_5-69) alone is exposed to RNase A its U32, C39, U43, and C48 are cleaved, which is similar to our predictions. However, different from our predictions, no cleavage is observed at C41, U46, and C47, suggesting that there is likely tertiary structure involved in the big bulge region. When RNA (rpsF_5-69) alone is exposed to RNase V1, cleavage is seen at C28-C29 and G35-C38, which is the predicted H2 stem, as well as U24-U25 and A52-G56, the predicted H1 stem. In addition, the RNase probing assays with the RNA mutant M4 alone result in very similar cleavage patterns compared with those of the wildtype RNA, suggesting that mutating the conserved double cytosine does not significantly alter the structure.

The RNase probing assays can also provide information on general RNA-protein interaction regions, since the S6:S18 dimer may protect the RNA structure from RNase attack at certain regions. Our results show that with the dimer protein present, protection from RNase V1 is observed at C28-C29 and G35-C38, suggesting that the RNA structure might have direct contact with the dimer at the predicted H2 region upon protein binding. Protection is also seen at A52-G53, which is the H1 stem. These results indicate that the loss of binding with M1/2 and M3 we observed in EMSA are likely due to loss of the stems, and that of M5 is probably due to loss of
sequence-specific contacts between the mutant RNA and the dimer protein. In accordance with the EMSA results, there is no sign of M4-S6:S18 dimer interaction in RNase probing assays, suggesting that the conserved cytosine pair is critical in dimer binding as predicted. We also tested the RNA-S18 interaction without S6 with RNase probing assays. Our results show no cleavage change even at high S18 concentration (600 nM), suggesting that the RNA:S18 interaction we observed with EMSA might be too weak for probing assays to detect or it could be nonspecific interaction due to S18 aggregation.

Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) results

To study the secondary structure of the predicted RNA, we also conducted SHAPE on protein-bound and unbound full length RNAs (rpsF_1-69) (Figure 12; Figure 13). In a SHAPE experiment, RNA is treated with reagents such as N-methylisotoic anhydride (NMIA), which selectively reacts with the 2'-hydroxyl positions of the RNA backbone at conformationally flexible nucleotides to form 2'-O'-adducts (Wilkinson et al., 2006). These adducts are then detected by primer extension and the results can report the extent to which a nucleotide is constrained by base pairing or other interactions (Deigan et al., 2008). Our data suggest similar flexibility at the 3' of the RNA structure in both states. Reactive nucleotides are seen at the predicted H1/2 loop regions, the big bulge region, and the 3' end, suggesting that these regions indeed are not structured. We were unable to quantify the first six nucleotides of the RNA due to the poor resolution at the 5' end on the gel. In addition, the loss of nucleotides 35-39 quantification data is due to a strong pause site at the base of the H2 stem. In general, the structures of the S6:S18-bound and the unbound state are very similar, indicating that the RNA structure is not formed as a response to protein binding.
Discussion

In this chapter, we experimentally investigated a RNA motif predicted to be associated with bacterial S6 operon. The transcriptional organization of *B. subtilis* S6 operon was studied by 3’ and 5’ RACE. The binding partner of the RNA motif and the binding kinetics were investigated by EMSA. The secondary structure of the predicted RNA motif was confirmed by SHAPE and detailed RNA-protein interaction was profiled by nuclease protection assays.

We identify the conserved RNA motif found in the 5’-UTR region of S6 operon specifically interacts with an S6:S18 dimer protein *in vitro*. Ribosomal protein S18 is insoluble and only partially folded without binding to S6 (Held et al., 1974). Although we observed weak interaction between the predicted RNA structure and S18, the dissociation constant is high (> 500 nM) compared to that of the RNA-S6:S18 binding curve (150 nM). In addition, we find that in nuclease protection assays, S18 alone cannot protect the RNA structure from nuclease cleavage, proving that this interaction is not specific and is probably the result of insoluble S18 aggregation. The dissociation constant we observed for the rpsF_5-69 and S6:S18 dimer binding is significantly higher than that reported for the 16S rRNA-S15 complex and S6:S18 binding (6 nM) (Recht et al., 2001). The difference is consistent with many previously observed variations between ribosomal protein-rRNA binding and ribosomal protein-mRNA binding. It also fits the theory that when ribosomal proteins are expressed they preferably bind to the rRNA for ribosome assembly (Zengel et al., 1994).

During RNA-S6:18 dimer binding titration, we do observe a steep transition (Hill coefficient = 4.1) between the unbound and bound states. A similar steep transition is also seen in a previous study (Recht et al., 2001) with dissociation constants of S6:S18 dimerization and S6:S18-rRNA
interaction close to each other (8 nM and 6 nM). We conclude that the steep transition in our study is due to a similar reason, which is the inefficient formation of S6:S18 dimer at low protein concentration in binding buffer. This conclusion is supported by the lower dissociation constant numbers observed when titrating S6 and S18 individually while holding the other component at a constant high concentration (500 nM).

Previous studies as well as our computational data suggest that a pair of cytosines (C47 and C48) are very conserved in the internal bulge region of the consensus RNA structure associated with S6 operon and may be directly involved in protein-binding (Yao et al., 2007). Our experimental data show that mutating the cytosine pair (M4) disrupts RNA-S6:S18 interaction and that at least one cytosine (C48) is protected by the dimer protein from nuclease cleavage. In addition, a similar cytosine pair is also conserved in >90% of bacterial 16S rRNA sequences (Figure 8) and is involved in S18-binding during ribosome assembly (Recht et al., 2001). Considering the weak but consistent interaction between S18 and rpsF_6-59 observed in our binding assays and the similarities between the rRNA and mRNA primary sequences, we speculate the mRNA version of the protein-binding structure is a mimic of the rRNA version, which is also a common characteristic seen in other verified cis-regulatory RNA elements for ribosomal proteins (Zengel et al., 1994).

The RNA we have characterized from B. subtilis joins the collection of RNA structures that is widely distributed and universally interact with ribosomal proteins (L1, L10, and S2) (Fu et al., 2013). If its regulatory roles are confirmed by future study, this RNA would also fill a current gap in the regulation of ribosomal protein S6 operon.
Chapter 4:

Investigating RNA-protein interaction using reporter systems in *E. coli* and *S. pombe*
In this chapter, we constructed a lacZ reporter system in *S. pombe* and a dual-luciferase reporter system in *E. coli*. Both systems were designed to carry out *in vivo* characterization of RNA-protein interactions and serve as potential platforms for high throughput screening of RNA-binding small molecules.

**lacZ reporter system in Schizosaccharomyces pombe**

**System design**

An *S. pombe*-based reporter system was designed to detect biological compounds or small molecules that selectively interact with bacterial ribosomal protein *cis*-regulatory RNA elements. The system consists of two plasmids: pNMT41 (a gift from Dr. Hoffman) carries the genes coding for bacterial ribosomal proteins following a thiamine repressible promoter *nmt1*; pCHY24 (a gift from Dr. Hoffman) carries the *cis*-regulatory RNA-lacZ translational fusion that is controlled by the glucose repressible *fbp1* promoter (Hoffman et al., 1989). By expressing ribosomal protein and the RNA-lacZ fusion simultaneously, the expression of β-galactosidase, encoded by lacZ gene, could reflect the RNA-protein interaction. When protein and RNA bind, low Miller units would be observed in Miller assays, whereas high Miller units indicate a lack of RNA-protein interaction.

**β-galactosidase assay results**

We tested three ribosomal protein-RNA pairs, *E. coli* S15, *G. kaustophilus* S15, and *B. subtilis* S4. For *E. coli* S15, lacZ expression was inducible, however no consistent difference was observed in β-galactosidase level when S15 protein overexpression was induced (Figure 14), suggesting that either the protein overexpression was not successful in *S. pombe* or there was not enough protein expression to saturate the RNA motif and change lacZ expression. For the latter two, we
could not induce lacZ expression at all. Considering the fact that the *G. kaustophilus* S15 RNA-lacZ fusion is functional in *E. coli* when expressing from a bacterial plasmid with an L-arabinose-inducible promoter (Betty Dixon, unpublished data), and that the S4 RNA-lacZ fusion was functional in *B. subtilis* with various promoters (Grundy et al., 1991), we speculate that these two bacterial RNA structures may not be compatible with the transcription/translation system in *S. pombe*. Thus, the same motif that allows transcription/translation in the absence of its protein partner in bacteria may block transcription/translation in yeast.

**Discussion**

In this work, we incorporate three ribosomal proteins and their known partner RNA motifs into the lacZ reporter system designed in *S. pombe*, as the proof-of-concept test. We can only collect readable data with the *E. coli* S15-RNA pair, however, our inconsistent results suggest no specific RNA-protein interaction. In previous studies, the RNA-mediated *E. coli* S15 regulation has been extensively studied and validated both *in vitro* and *in vivo* using *E. coli*-based reporter system (Philippe et al., 1990; Scott et al., 2001). One explanation for the inconsistent results we observe is that our *S. pombe*-based system may need further tuning. It is also highly likely that these bacterial RNA motifs are not compatible with eukaryotic environment, considering the fact that we cannot turn on the expression of lacZ gene with *G. kaustophilus* S15-binding RNA motif or *B. subtilis* S4-binding RNA motif attached to its 5’ end in the *S. pombe* system, although both motifs have been successfully validated in bacterial reporter systems (Betty Dixon, unpublished data) (Grundy et al., 1991). In general, we speculate that the *S. pombe* transcription/translation system is acting on the bacterial RNA motifs through some unexpected mechanisms. More literature search as well as experimental research need to be done before we consider this yeast-based reporter system feasible.
Dual luciferase reporter system in *E. coli*

**System design**

To assess RNA-protein interaction in vivo, we designed a dual-luciferase reporter system in *E. coli*. This system consists of two plasmids: a petHT plasmid for protein expression and a pBAD33 derivative plasmid (Figure 15) carrying two reporter genes, coding for firefly and renilla luciferase. The wildtype RNA sequence to be tested, including the native Shine-Dalgarno sequence could be fused to one luciferase gene through translational fusion, and the mutant RNA could be fused to the other luciferase gene. Each RNA-luciferase translational fusion has its own L-arabinose-inducible promoter and a terminator. Since Firefly and Renilla luciferases react with specific substrates (firefly luciferin and renilla luciferin, respectively), such design would allow us to measure two different RNA-protein/peptide interactions in the same system simultaneously (Grentzmann et al., 1998; Harger et al., 2003; McNabb et al., 2005). When a protein interacts specifically with the wildtype RNA target, a change in the expression level of the corresponding fusion luciferase is expected, whereas the expression level of the mutant RNA-fused luciferase is expected to stay the same. In addition, both *E. coli* culture and luciferase color assays can be performed using multi-well plates, providing a potential platform for high-throughput screening for RNA-binding biological compounds and small molecules.

**Preliminary results suggest no L1 protein-RNA interaction in vivo**

We tested two RNA structures, the structured RNA preceding rplA from *G. kaustophilus* (64 nt) and the one preceding rplA from *E. coli* (93 nt, including the native transcription initiation site). Both RNAs are known to interact with L1 ribosomal protein in vitro (Fu et al., 2013). However, our in vivo experiment results do not support the RNA-protein interaction and the RNA-mediated regulation (Figure 16). In both constructs we fused the wildtype RNA to renilla
luciferase gene. Considering that renilla luciferase color assays generally yielded higher readouts than those of firefly luciferase in our pilot experiments, such design would be helpful to reflect even small changes associated with the wildtype RNA-protein interaction. In general, we observed relatively stable firefly luciferase activity after 30-60 min’s overexpression of G. kaustophilus L1 protein in E. coli, and varied and unexpected changes of renilla luciferase level under the same condition. One possibility is that not enough L1 protein were expressed in the system to interact with the RNA and induce changes in luciferase expression. However, as we demonstrated in other experiments that unbalanced ribosomal protein level and ribosomal RNA level lead to ill growth of E. coli (Figure 19), we also observed significant cell density (OD) drop and firefly luciferase expression drop after overexpressing L1 protein in E. coli over 1 hour. Considering our inconclusive assay data, the culturing condition may need further optimization. The difficulty of engineering L1 overpression strain might also be the reason why RNA-mediated L1 regulation has only been proved using in vitro cell-free transcription-translation system (Hanner et al., 1994).

Discussion

Dual-luciferase reporter system in E. coli

We do not observe consistent results suggesting specific G. kaustophilus L1 RNA-protein interaction using the dual luciferase reporter system constructed in E. coli. It might suggest that the system need to be further fine-tuned to report the RNA-mediated regulation. In addition, previous literatures show that the RNA-mediated L1 regulation, besides the native E. coli L1 regulation system (Thomas & Nomura, 1987), is hardly investigated using in vivo reporter systems in E. coli. For example, the regulation of L1 operon in Methanococcus Vannielli, M. thermolithotrophicus, and M. jannaschii are characterized using cell-free in vitro systems
(Hanner et al., 1994; Kraft et al., 1999). This difficulty might result from the conservation of the L1-binding RNA motif, as well as the L1 protein across bacteria (Fu et al., 2013), and may imply the limitation of applying *E. coli* to L1 related studies. Thus, to test the feasibility of this *E. coli*-based dual luciferase reporter system, we could incorporate other known ribosomal protein and its partner RNA motif into the system, for example, *G. kaustophilus* S15 protein and its specific RNA motif.
Chapter 5:

Selection of small peptides that target bacterial cis-regulatory RNAs
Phage display for selecting RNA-binding small peptides

In this chapter, we aimed to use the phage display technique to select for small peptides that can simulate the native interaction between the regulatory RNA and its ribosomal protein binding partner. Previously, phage display technology has been used extensively in vitro to identify peptides, proteins, and antibodies based on specific interaction or catalytic activity with a given target (Parmley et al., 1988; Sidhu, 2000; Smith, 1985). Phage display describes a selection technique in which a peptide library is expressed together with the capsid protein of a phage virion, while the gene coding for each variant is located on the inside (Rodi & Makowski, 1999). Viruses displaying peptides bind specifically to immobilized targets and can be separated from phage library pool by panning, with the unbound phage being washed away. Phage display has been experimentally proven to be feasible for targeting proteins (Biorn et al., 2004), protein-binding DNAs (Cheng et al., 1996), or protein-binding RNAs (Danner et al., 2001; Li et al., 2009; Llano-Sotelo et al., 2009). The Ph.D.™ Phage Display system from NEB is based on a M13 phage vector, modified for displaying peptide libraries as N-terminal fusions to the minor coat protein pIII (Cwirla et al., 1990; Devlin et al., 1990; Scott et al., 1990). pIII modulates phage infection by binding to the F-pilus of the host cell, and is present as a five copy cluster at one end of M13 virion (Rodi et al., 1999). Fusing a peptide with no more than 50 residues to pIII does not affect its infectivity function. All of the commercially available libraries from NEB have complexities on the order of $10^9$ independent clones, which is sufficient to encode most of the possible 7-mer ($1.28 \times 10^9$) peptide sequences. The large number of clones is also the main advantage of phage display library compared with other libraries with typical $\sim 10^4$ clones (i.e. cDNA libraries expressed in phage lambda). Theoretically, such in vitro system could be used for selecting small peptides that interact with a specific cis-regulatory RNA element.
**RNA target design**

A 62-nt RNA (same as that in filter-binding assays) and a 81-nt RNA (longer at the 5’ end), containing the L1-binding site and RBS that precede gene rplK, was *in vitro* transcribed from *G. kaustophilus* genome. The last 17 nucleotides at the 3’ end was hybridized to a biotin-labeled DNA primer (Figure 17) (Li et al., 2009). Both RNA target sequences cover the minimal requirement for forming the L1 protein-binding structure and can be isolated by magnetic beads through biotin-streptavidin binding. To increase RNA-protein interaction specificity, we also constructed a mutant version of the 62-nt RNA target without the biotin tail by introducing the same two G-to-A mutations as we did for filter-binding assays (Figure 6).

**RNA-binding heptapeptide selection**

The selection for RNA-binding heptapeptides was designed to include 4 rounds. Round 1 through 3 were positive selections where the wildtype RNA is used as the target, whereas in Round 4 a counter selection was carried out in the presence of the mutant target. By using this approach, we aimed to eliminate all the non-specific peptide binders and save peptides that interact with the RNA structure at the same site as L1 protein does. In addition, we also employed two elution methods, specific and low-pH, to isolate RNA-bound phage after each round of selection, as suggested in manufacture’s protocol. The specific elution relies on excessive nonbiotinylated elution RNAs to compete the RNA-bound phage off the beads, whereas the low-pH elution releases the RNA-bound phage from the beads by using acidic buffer to break the biotin-streptavidin bound.

With the 62-nt RNA target, the “winner” peptides was FSGGGNH (10/10, low-pH elution; 8/10, specific elution (Table 2). The other two were NYSYIPP and TVNKLY). With the 81-nt RNA target,
the “winner” peptides were QMLLRLP (QMLLRLP, low-pH elution), NPTRKPK (2/9, low-pH elution), MPTRPNK (3/9, specific elution), and MITTTRK (2/9, specific elution) (Table 2). All of these selected peptides are positively charged which is a characteristic commonly seen in RNA/DNA binding peptides due to the negatively charged phosphate backbone. Normally, the low-pH elution method generates a wider variety of sequences due to the strong cleavage of the biotin-streptavidin bond. However, such results were not seen in our experiments.

**Growth curve assays**

To examine how select peptides affect bacteria growth, we introduced 6 of them into *E. coli* BL21 strain. We overexpressed FSGGGNH, FSGGGNHGGGS, TVNFKLY, TVNFKLYGGGS, NYSYIPP, and NYSYIPPGGGS in *E. coli* cells and monitored *E. coli* growth by measuring the cell density over a period of time (Figure 18). The GGGS sequence acts as the C-terminal spacer between each displayed peptide and the M13 phage virion in the Ph.D. system (NEB). We included the spacer-attaching peptides in this experiment as well since such spacer might be involved in the RNA-peptide interaction during the selection. Our data show that the overexpression of TVNFKLY leads to significant cell death compared to that of the control group (empty pET-HT), whereas the overexpression of the other five peptides does not make notable differences in cell growth under the same growth condition. To exclude the possibility that TVNFKLY overexpression-associated growth defect was due to the accumulation of insoluble small peptides, a peptide with reversed amino acid sequence (YLKFNVT) was also tested using the same overexpression system. To our surprise, normal cell growth was observed, suggesting that the insolubility was less likely a problem.
In addition, to demonstrate that the balance between the level of ribosomal proteins and the level of ribosomal RNAs is important for healthy bacterial growth, we also overexpressed *G. kaustophilus* ribosomal protein L1, S6, S18, and S15 using the same method in *E. coli* BL21 (Figure 19). For example, we speculated that overexpressed L1 would bind to the RNA element on L1 operon and inhibit expression of both native L1 and L11. Thus, the accumulation of *G. kaustophilus* L1 and the deficiency of native *E. coli* L11 would lead to *E. coli* cell death. Interestingly, all of these ribosomal proteins caused defective bacterial growth when overexpressed. L1 overexpression led to significant cell death in a way very similar to that of the peptide TVNFKLY overexpression. S18 and S6 overexpression also resulted in notable bacterial death but not as fatal as that observed for L1. Defective growth seen with the S15 overexpression was not quite expected since *G. kaustophilus* S15 does not interact with the S15 RNA element from *E. coli in vivo* (Betty Dixon, unpublished data), and the gene coding for S15, *rpsO*, is the only gene on the S15 operon in *E. coli*. However, it could also be interacting with the *E. coli* ribosome. In general, the growth curve data reveal the negative effect of one selected peptide on *E. coli*, however, without further *in vivo* and *in vitro* experiments, we could not conclude that the defective growth seen with TVNFKLY overexpression was caused by direct peptide-RNA interaction.

Discussion

In this work, we carry out phage display selections using the L1-binding RNA as the target. We use both specific and low-pH elution methods during selection Round 2-4 and observe variations in output peptides. Theoretically, the low-pH elution method provides more diverse output, since it releases the RNA-bound phage by releasing the RNA target off the beads, whereas the specific elution competes the bound phage off the beads by increasing the concentration of free
RNA targets in solution, where strong binders may still be on-bead (Li et al., 2009). However, we see more diverse output with the specific elution method, which could be due to the small size of sequenced samples (10 for each elution method) or that some RNA-binding phage were killed by low pH buffer before neutralization.

We show with in vivo assays that the overexpression of one particular peptides, TVNKLY, in *E. coli* leads to severe growth defect. Without in vitro data we cannot yet conclude that this observation is directly linked to L1 RNA-peptide interaction. However, we see normal growth when overexpressing the same peptide with reversed amino acid sequence (YLKFNVT), which lowers the possibility that the abnormality in *E. coli* growth is caused by accumulation of insoluble peptides.

One potential problem with phage display selection is that the most strong RNA-binders may kill *E. coli* during amplification step by disrupting the native ribosomal protein regulation system and never get selected out. However, considering our specific aim, which is to select for small peptides that could target at a specific RNA motif and act as potential antibacterial leads, such down-side is acceptable.
Materials and Methods

Chapter 2

5’ RACE

5’ RACE was performed with the Invitrogen GeneRacer kit on the total RNA extracted from 
*Bacillus subtilis* log phase culture to identify the 5’ end transcription initiation site. All primers 
were listed in Table 1. First strands were synthesized using two gene-specific primers derived 
from *rplK* and *rplA* coding regions respectively. RACE PCR was done using the two first strands as 
templates, a linker-specific primer and a second set of gene-specific primers. All PCR products 
were cloned into TOPO vectors (Invitrogen TOPO-cloning kit) and sequenced.

Filter-binding assays

The two predicted RNA structure sequences (preceeding *rplA*: 37 nt, preceding *rplK*: 41 nt, both 
including the T7 promoter sequence) were generated from *Geobacillus kaustophilus* genomic 
DNA by PCR, transcribed *in vitro* with T7 RNA polymerase, and 5’-labeled with \[^{32}\text{P}]\text{ATP}\ in the 
presence of T4 PNK. Mutants of these RNAs (including two G-to-A mutations) were prepared 
with the same methods. *G. kaustophilus* ribosomal protein L1 was over-expressed from pET-HT 
vector in *E. coli* BL21 (Block, Puerta-Fernandez, Wallace, & Breaker, 2011) and purified using 
denaturing ion-exchange chromatography. Filter-binding assays were conducted as previously 
described (Köhrer et al., 1998). A fixed amount of RNA (approx. 1 nM, 500 cpm) was incubated 
with L1 protein in serial dilutions (0 – 500 nM) in a total volume of 50 μL for 15 min at 42 °C (L1 
binding buffer: 50mM Tris–HCl pH 7.6 at 25°C, 20mM MgCl2, 500mM KCl, 1mM beta-
mercaptoethanol, 0.04% BSA). The RNA-protein complexes were captured by vacuum suction 
through a nitrocellulose membrane (Optitran BA S-85 reinforced nitrocellulose, Whatman), 
whereas the unbound RNAs were retained on a nylon membrane (N+ hybond, GE Healthcare).
Both membranes were washed once with 50 μL binding buffer. The radioactivity was quantified using a GE Healthcare STORM 820 phosphoimager and ImageQuant. The protein-binding RNA fraction is calculated as (counts on the nitrocellulose membrane)/(counts on both membranes).

Chapter 3

5’ and 3’ RACE

5’ and 3’ RACE were both performed with the Invitrogen GeneRacer kit on Bacillus subtilis total RNA extracted from a log phase culture to identify the transcription initiation site, the termination site, and the transcriptional organization of the S6 operon. All primers were listed in Table 1. In 5’ RACE, first strands were synthesized with a set of gene-specific primers derived from rpsF and rpsR respectively. RACE PCR was done using the first strands as templates, a linker-specific primer and a second set of gene-specific primers, and products were cloned into TOPO vectors (Invitrogen TOPO-cloning kit) for sequencing. 3’ RACE was conducted by adenylation of the total RNA extract followed by reverse transcription with a dT primer. RACE PCR was done using the reverse transcript as the template, a gene-specific primer, and an Anchor primer, and products were cloned into TOPO vectors (Invitrogen TOPO-cloning kit) for sequencing.

RNA and protein preparation

The sequence preceding rpsF was PCR amplified from Bacillus subtilis genomic DNA with a T7 promoter and was in vitro transcribed with T7 RNA polymerase to generate the rpsF_5-69 RNA and rpsF_1-69 RNA. Purified RNA was 5’-labeled with [γ-32P] ATP in the presence of T4 PNK. To produce RNA mutants (65 nt, M1-M6), DNA templates carrying mutations were generated.
through QuickChange mutagenesis as the first step, followed by the same methods described above.

Genes coding for *Geobacillus subtilis* were cloned in to pET-HT and transformed into *E. coli* BL21. S6 and S18 proteins were overexpressed from the T7 promoter and purified similarly to that previously described (Culver et al., 1999). Cells were lysed in Buffer E (20 mM HEPES [pH 7.6], 20 mM KCl, and 6 mM BME). S6 was purified with Buffer C (20 mM NaOAc [pH 5.0], 20 mM KCl, and 6 mM BME) by nondenaturing FPLC cation exchange. S18 was insoluble and was purified from inclusion bodies using Buffer B (20 mM Tris-HCl [pH 7.6], 20 mM KCl, 6 M urea, and 6 mM BME) by denaturing chromatography. Purified proteins were dialyzed against Buffer D (S18 with 4 M urea) and stored at 4 °C.

**Electrophoretic mobility shift assays (EMSA)**

EMSA assays were performed to study the RNA-protein interaction related to the S6 operon. A fixed amount of RNA (rpsF_5-69, approx. 1 nM, 1000 cpm) was incubated with proteins (S6, S18, and S6:S18 [1:1]) in serial dilutions (0 – 500 or 600 nM) in a total volume of 10 μL at 25 °C (binding buffer: 50 mM Tris-HCl [pH 7.6], 20 mM MgCl₂, 150 mM KCl, 1 mM BME, 0.04% BSA). After 30 min’s incubation, 5 μL 50% glycerol was added before loading onto a native 10% acrylamide gel which had been prerun for 60 min at 160 V under 25 °C. The loaded gel was run at 350 V at 4 °C and dried and imaged using GE Healthcare STORM 820 phosphorimager and radioactivity quantified using ImageQuant. Quantifications were fit to a standard curve generated by \( \frac{F_{\text{max}} \cdot y^n}{x^n + K_D^n} \), where \( F_{\text{max}} \) is the maximum fraction bound, \( y \) is the measured fraction bound, \( n \) is the Hill coefficient, and \( x \) is the protein concentration.
For individual protein titrations, 600 nM S6 was incubated with S18 in serial dilutions (0 nM – 600 nM) in binding buffer for 10 min before a fixed amount of RNA (rpsF_5-69, approx. 1 nM, 1000 cpm) was added. The reactions were further incubated for 20 min. The RNA-protein mixture was examined by EMSA as described above. For the S18-concentration-fixed titration, 500 nM was held as the maximum concentration, since S18 at 600 nM displayed significant aggregation and uninterruptable EMSA results.

Nuclease protection assays

Nuclease protection assays were conducted to investigate the secondary structure of the bound and unbound RNAs. The RNA-protein binding reactions in these assays were identical to those used for EMSA. After incubation, each reaction was treated with RNase A (Ambion, 20 pg) and RNase V1 (Ambion, 0.002 units) for 10 and 5 min at 25 °C, respectively followed by inactivation of the nuclease. Processed RNA fragments were isolated by ethanol precipitation and were resuspended in 10 μL water and 10 μL urea loading solution (10 M urea, 1.5 mM EDTA). Partial alkaline hydrolysis reactions and denaturing RNase T1 reactions were used as ladders. Partial alkaline hydrolysis reactions were prepared by incubating RNA in reaction buffer (50 mM Na₂CO₃ [pH 9.0], 1 mM EDTA) for 10 min at 95 °C. Denaturing T1 reactions were prepared by incubating RNA with RNase T1 (1U, Roche) for 15 min at 55 °C (25 mM Na Citrate [pH 5.0], 5.5 M urea). 10 μL of each reaction and ladder was loaded on a 12% denaturing Acrylamide/Bis gel and run at 35W at 25 °C. The gel was dried and examined using a GE Healthcare STORM 820 phosphorimager and ImageQuant.
SHAPE (selective 2’-hydroxyl acylation analyzed by primer extension) analysis

SHAPE was performed to further study the secondary structure of the predicted RNA. The predicted RNA (rpsF_1-69) with a 3’ end primer binding site was prepared as previously described by in vitro transcription. The RNA-protein dimer reaction and the no-protein control reaction were prepared as described above and incubated with 3 μL NMIA (130 mM) for 15 min at 25 °C to form acylated products. Modified RNAs were purified by phenol-chloroform extraction and ethanol precipitation and resuspended in 10 μL TE buffer (0.5X). Primer extension was conducted by incubated the modified RNA with a 5’-labeled primer for 30 min at 55 °C using Invitrogen Superscript III kit. Reverse transcription was stopped by addition of 1 μL NaOH (4 M) and 29 μL acid stop mix (Wilkinson et al., 2006). cDNAs were loaded on a 10% denaturing Acrylamide/Bis gel and run at 55W at 25 °C. The gel was dried and examined using a GE Healthcare STORM phosphorimager and SAFA footprinting software (Das et al., 2005; Laederach et al., 2008). Data was normalized as previously described (Deigan et al., 2008; Wilkinson et al., 2006).

Chapter 4

S. pombe lacZ reporter system

System setup

The S. pombe strain CHP902 (h' ura4-D18 leu1-32) was a generous gift from Dr. Hoffman. DNA sequences containing the minimal structure of each RNA element were PCR amplified from bacterial genomic DNAs and integrated into pCHY24 (Hoffman et al., 1989) through DMSO transformation as previously described (Bähler et al., 1998). Recombinant plasmids were isolated from yeast cells using “Smash and Grab” method (Hoffman et al., 1987) and transformed into E. coli XL1-Blue for sequencing. Sequences coding for each ribosomal protein
were also PCR amplified, integrated into pNMT41 (a gift from Dr. Hoffman), and sequenced in the same way. Two plasmids were co-transformed into S. pombe.

β-galactosidase assays

Individual colonies containing both plasmids were inoculated in YES-uracil-leucine media overnight at 30 °C before sub-cultured into EMM-uracil-leucine (8% glucose) with and without thiamine (5 μg/mL). Sub-cultured cells were incubated for another 24 hours at 30 °C before sub-cultured again into EMM (0.1% glucose + 3% glycerol) with and without thiamine (5 μg/mL) for 3h at 30 °C to 10^7 cells/mL for β-galactosidase assays. β-galactosidase assays were performed as previously describe (Rose & Botstein, 1983). Protein concentrations were measured using the BCA Protein Assay kit (Pierce). β-galactosidase units = 

\[
\frac{(OD_{420}) \times 378}{time \times volume \ of \ extracts \times protein \ (mg/mL)}
\]

Data were analyzed using Excel.

Dual luciferase reporter system

System setup

The dual luciferase reporter plasmid and protein overexpression plasmid were constructed in E. coli XL1-Blue cells. pBAD33 (ATCC) was modified to eliminate the Ncol site through silent mutation (QuickChange Site-Directed Mutagenesis Kit, Stratagene). The firefly luc+NF gene was PCR amplified from pSP-luc+NF (Promega) and inserted into the MCS of the modified pBAD33, followed by an inserted rrnB terminator. A second promoter was inserted after the first luciferase reporter cassette, followed by the renilla luc+NF gene (a gift from Dr. Gubbels), and the native rrnB terminator on pBAD33. A 93-nt sequence and a 69-nt sequence containing the L1-binding structures from E. coli and G. kaustophilus, respectively, were PCR amplified from the genomic DNAs, and the mutant forms were prepared in the same way. Sequence of the wildtype
structure was inserted into the renilla luc-NF cassette, whereas sequence of the mutant structure was inserted into the firefly luc-NF. The protein overexpression plasmid was constructed with rplA (amplified from G. kaustophilus genomic DNA) inserted into pET-HT. Two plasmids were co-transformed into E. coli BL21(DE3) (no plysY, NEB).

Luciferase color assays
To perform dual luciferase, individual colonies that contain both plasmids were cultured in LB media with antibiotics and L-arabinose (15 mM) for 90 min before IPTG induction (30 μM). Cells were collected 1 hour after IPTG induction and quick-freezed at -80 °C. Luciferase assays were conducted according to manufacturer’s protocols (Luciferase Assay System and Renilla Luciferase Assay System, Promega). Data were collected using SpectraMax M3 (Molecular Devices). The ratio of renilla to firefly luciferase activity (Rluc / Fluc) was calculated using the equation \(\frac{(\text{renilla luciferase activity/OD600})}{(\text{firefly luciferase activity/OD600})}\).

Chapter 5
RNA target preparation
The G. kaustophilus RNAs were prepared as described above (Figure 17). The last 17 nt of the wildtype RNA was hybridized to a 5’ biotin-tagged DNA (purchased from Eurofins) by mixing at 1:2.5 molar ratio in 50 μL hybridization buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.2) and incubating for 5 min at 95 °C (Li et al., 2009). Hybridization was cooled down slowly to 25 °C and examined on an 8% nondenaturing acrylamide gel.

Phage display selection
Phage display was conducted to select for heptapeptides that specifically interact with L1 cis-regulatory RNA element from *G. kaustophilus*. The Ph.D. -7 Phage-Display Peptide Library was purchased from NEB and the methods were adapted from NEB protocol and previous studies (Li et al., 2009; Llano-Sotelo et al., 2009). The phage library (10 μL, 2×10^{13} pfu/mL) was preselected against the streptavidin-coated magnetic beads (100 μL, Invitrogen Dynabeads M-280 Streptavidin) overnight at 4 °C in L1 binding buffer (see L1 methods) to remove streptavidin-binding phage. First round selection was conducted using the preselected phage library by incubating with 30 pmol RNA targets in the presence of 150 pmol competitor tRNAs (Ribonucleic acid, transfer, from baker’s yeast. Sigma) in 100 μL L1 binding buffer for 30 min at 25 °C on a rocker. After washing five times with washing buffer (L1 binding buffer with 0.3 % Triton X-100), the RNA-phage complexes were captured by adding 0.1 mg beads into the reaction and incubating for another 30 min at 25 °C with rocking. Collected RNA-phage complexes were eluted by incubating with 300 pmol elution RNAs in 100 μL L1 binding buffer for 30 min at 25 °C. Eluted phage were amplified in *E. coli* ER2738 (NEB) and titered before applied in the next round of selection. The same methods were used throughout Round 2-4 with increasing stringency in the washing step and the addition of a nonspecific elution method (Table 1), in which the RNA-phage complexes were incubated with 100 μL elution buffer (0.2 M glycine-HCl [pH 2.2], 1 mg/mL BSA) for 9 min at 25 °C followed by neutralization with 15 μL of 1 M Tris-HCl, pH 9.1. In addition, in Round 4, 300 pmol mutant RNA targets were added for counter-selection and no amplification was carried out. After the final round, individual phage plaques were picked out and resuspended in water and were PCR amplified using a pair of phage sequencing primers (Li et al., 2009). PCR products were sequenced to identify selected peptide sequences.

**Growth curve assays**
Growth curve assays were performed to investigate the effect of ribosomal proteins and peptides overexpression on *E. coli* growth. Coding sequences of ribosomal proteins were PCR amplified from bacteria genomic DNA and cloned onto pET-HT. Coding sequences for selected peptides with a N-terminal start codon (ATG) were assembled by annealing complementary primer pairs and inserted into pET-HT as well. pET-HT constructs were transformed into *E. coli* BL21(DE3) strain. Individual colonies were cultured in LB media with antibiotics overnight at 37 °C, which were then diluted with fresh media (OD600 < 0.01) with and without IPTG (1mM) the next morning. Culture were incubated at 37 °C for 4-5 hours and cell density was measured with samples collected every 30 min. Growth curves were generated using Excel.
**Figure 1. An overview of ribosomal protein regulation.** To maintain a balance between the level of rRNAs and the level of ribosomal proteins, expression of the ribosomal protein can be autogenously regulated using a negative feedback mechanism. When rRNAs are saturated with bound ribosomal proteins, excess ribosomal proteins can bind to a RNA regulatory element associated with its operon to stop transcription or translation. Because multiple ribosomal proteins are often arranged in the same operon, one regulatory mechanism can control many ribosomal protein concentrations.
Figure 2. Phylogenetic distribution of *E. coli* r-protein regulatory RNAs. (A) Distribution of ribosomal protein autogenous regulatory RNAs in bacterial phyla. (B) Distribution of regulatory RNAs in orders of Gammaproteobacteria. (Fu et al. 2013)
Figure 3. Regulation models of L1 biosynthesis in *E. coli* and Archaea. Transcriptional organization and regulation of L1-related genes. (A) *E. coli*; (B) Archaea. Regulatory RNA elements are shown in grey. (Kraft et al. 1999)
Figure 4. Consensus sequence and secondary structure for r-protein L1 regulatory RNA. This diagram indicates the conserved secondary structure and sequence of the RNA motif. Base pairs supported by covariation are indicated by green shading only if Watson-Crick pairing occurs in >95% of the aligned sequences. Putative L1 binding site is indicated by dashed lines. (Fu et al. 2013)
**Figure 5. Multiple regulation model of ribosomal protein L1.** The L1 binding site is commonly found before the *rplK* gene or before *rplA*. However, in Firmicutes, as represented by *B. subtilis* and *G. kaustophilus*, the L1 binding site is located before both of the *rplK* and *rplA* genes, which presents a unique model. Conserved RNA structures are shown as red arrows.
Figure 6. L1-binding RNA structures from *G. kaustophilus*. (A) *G. kaustophilus* RNA preceding *rplK*. (B) *G. kaustophilus* RNA preceding *rplA*. The guanosine marked by * were mutated to adenosine to create negative control RNAs. Interactions between the RNAs and L1 protein as measured by filter-binding assays. RNAs pictured in parts A and B are represented by filled squares and circles, respectively. Their mutants are represented by open symbols. Error bars represent the standard error over three replicates. (Fu et al. 2013)
Figure 7. Known transcriptional organizations of the S6 operon. (A) Represented by E. coli and T. thermophilus (with the middle gene unannotated). (B) Represented by B. subtilis and G. kaustophilus (with the middle gene unannotated).
Figure 8. S6 operon-associated RNA secondary structures. (A) Consensus secondary structure of the RNA motif preceding rpsF derived from more than 1300 sequences. This RNA motif is widely distributed across many bacterial phyla. (B) Example of the RNA motif from *B. subtilis* 169. The transcription start site, an alternative pairing element, and potential regulatory features are indicated. (C) Consensus secondary structure for the more than 500 RNA sequences primarily identified from Firmicutes and Actinobacteria that contain a potential conserved H3 helix. (D) rRNA binding site for the S6:S18 complex. Sequence >98% across 4214 bacterial 16S rRNA sequences (Cannone et al., 2002). Bases in contact with S18 in *T. thermophilus* structure are boxed (Agalarov et al., 2000). (Fu et al., 2014)
Figure 9. rpsF_5-69 titrations. (A) Gel-shift of rpsF_5-69 with increasing concentration of S18 (0 nM, 19-600 nM). (B) Gel-shift of rpsF_5-69 with increasing concentration of S6 (0 nM, 9.7-600 nM). (C) Gel-shift of rpsF_5-69 with increasing S6 and S18 (mixed 1:1, 0 nM, 9.7-600 nM). Arrows indicate quantified bands. (D) Quantification of the fraction of rpsF_5-69 bound by the protein with S6, S18, and S6:S18. Error bars represent standard deviation on three to six independent experiments. Some error bars fall within the boundaries of the marker. Curve represents a standard binding curve with a $K_D$ of 155 nM, and a Hill coefficient of 4. (Fu et al., 2014)
Figure 10. rpsF_5-69 titrations with one of S6:S18 at constant concentration. (A) Gel-shift of rpsF_5-69 with 500 nM S18 and increasing concentration of S6 (no protein, 0nM S6, 4.6 – 500 nM S6). (B) Gel-shift of rpsF_5-69 with 600 nM S6 and increasing concentration of S18 (0 nM, 18.5 – 600 nM S18, no protein). Arrows on A and B indicate quantified bands. (C) Quantification of the fraction of rpsF_5-69 bound by the S6:S18 complex. Solid thick line corresponds to the titration of S6:S18 mixed in equal amount (K\textsubscript{D} of 37 nM, Hill coefficient 1.45, maximum fraction bound 93.5%). Dashed line is the curve fit of the S6 titration (K\textsubscript{D} = 37 nM, Hill coefficient = 1.45, maximum fraction bound 93.5%). Thin line is the curve fit for the S18 titration (K\textsubscript{D} = 60 nM, Hill coefficient = 1.5, maximum fraction bound is 100%). (Fu et al., 2014)
Figure 11. Mutant RNA titrations. (A) The RNA construct rpsF_5-69 used for in vitro binding assays with mutations indicated. (B) Interactions of S6:S18 with mutants displayed in A. Solid black curve is plotted for M6 (Kd = 115 nM, Hill coefficient = 3). For the remaining mutants RNA binding was not saturated, indicating a Kd > 500 nM and no curves are drawn. (Fu et al., 2014)
Figure 12. S6:S18 preferentially interacts with H2 of predicted RNA structure. (A) Highly reactive (black circles) and moderately reactive (gray circles) nucleotides determined from SHAPE analysis of RNA in isolation mapped to the rpsF_1-69 structure. Positions not resolved are in gray. (B) Reactive nucleotides in SHAPE analysis of RNA bound to the S6:S18 complex. (C) Nuclease probing data mapped to the secondary structure of the rpsF_5-69 RNA structure. Starred bases are cleaved by V1 nuclease and arrows indicate bases cleaved by RNase A. Bases protected from cleavage by the addition of protein are circled. Bases not resolved on the gel are in gray. Numbering starts from the transcription start site (A). (D,E) Nuclease probing gels of rpsF_5-69. (D) and rpsF_5-69 M4 (E) using RNase A (left), which cleaves single-stranded uridine and cytosine, and RNase V1 (right), which cleaves double-stranded RNA. RNA was incubated with increasing concentrations of S6 and S18 mixed at a 1:1 ratio. (F) V1 nuclease probing gel of rpsF_5-69 in the presence of S18 alone. OH− and T1 lanes indicate partial alkaline hydrolysis and RNase T1 digest under denaturing conditions used to map the RNA sequence. On the right, bases cleaved by RNase V1 are indicated. On the left, nucleotides cleaved by RNase T1 (G) and RNase A (C or U, bold) are indicated. (Fu et al., 2014)
Figure 13. SHAPE experiment for rpsF_1-69. The experiment was done as described in Materials and Methods section. T, C, and A are sequencing reactions performed using adenosine, guanine, and thymidine dideoxy terminating nucleotides. The middle panel is shown with increased contrast on the right.
**Figure 14. Varied reporter activities seen with *E. coli* S15-RNA interaction in *S. pombe*.** Each number represents an individual colony. Black columns show β-galactosidase unit measured for yeast cells with thiamine (5 μg/mL) – repressed S15 overexpression; grey columns show β-galactosidase unit measured for yeast cells with de-repressed S15 overexpression. The RNA-lacZ translational fusion expression was de-repressed in all cultures in the presence of 0.1% glucose and 0.3% glycerol as described in the Materials and Methods section.
Figure 15. A schematic of the dual-luciferase reporter plasmid. The reporter plasmid was modified from pBAD33. Arrows represent the L-arabinose-inducible promoters. Black squares represent rrnB terminators. Black and grey open circles represent the wildtype and mutated E. coli L1 regulatory RNA motif, respectively. Each RNA motif and the following luciferase gene is constructed as translational fusion.
Figure 16. Varied reporter activities seen with *G. kaustophilus* L1-RNA interaction in *E. coli*. Each number represents an individual colony. Black columns represent cultures without IPTG induction; grey columns represent cultures with IPTG induction (30 μM) for 1 h. The data are presented as the ratio of renilla to firefly luciferase activity (Rluc / Fluc).
Figure 17. *G. kaustophilus* RNAs used in the phage display selections. (A) r-protein L1-binding RNA target (62-nt) hybridized to a biotin tag at the 3’ end. (B) L1-binding RNA (39-nt) for specific elution. (C) Mutated RNA (39-nt) with abolished L1-binding ability for Round 4. The native Shine-Dalgarno sequence is labeled in grey; the native start codon is labeled in red. The adenosines marked by * were mutated from guanines to create the negative control.
Figure 18. *E. coli* BL21(DE3) growth curves for peptide overexpression. (A) N-FSGGNH-C. (B) N-FSGGNHGGGS-C. (C) N-TVFKLY-C. (D) N-TVFKLYGGGS-C. (E) N-YLFNVT-C. (F) N-NYSYIPP-C. (G) N-NYSYIPPGGGS. Black circles represent negative controls without peptides overexpression; open squares represent cultures with IPTG (1 mM)-induced peptides overexpression.
Figure 19. *E. coli* BL21(DE3) growth curves for r-protein overexpression. (A) Empty pET-HT vector. (B) *G. kaustophilus* L1. (C) *G. kaustophilus* S18. (D) *G. kaustophilus* S6. (E) *G. kaustophilus* S15. Black circles represent negative controls without peptides overexpression; open squares represent cultures with IPTG (1 mM)-induced peptides overexpression.
**Table 1. Primers for RACE.**

<table>
<thead>
<tr>
<th>L1 operon 5’ RACE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene-specific primer <em>rplK</em> 1</td>
<td>5’-AGCAATACTGCAGCAGGTGGAGTT</td>
</tr>
<tr>
<td>Gene-specific primer <em>rplA</em> 1</td>
<td>5’-TGCGAAAAACGAGAACCGCTGAGT</td>
</tr>
<tr>
<td>5’-linker</td>
<td>5’-GACTGGAGCAGAGGACACTGA</td>
</tr>
<tr>
<td>Gene-specific primer <em>rplK</em> 2</td>
<td>5’-CCAACCTGGTGGTGGGTTAGC</td>
</tr>
<tr>
<td>Gene-specific primer <em>rplA</em> 2</td>
<td>5’-CGGTCTACAGCTTAGCAGCTTCA</td>
</tr>
<tr>
<td>S6 operon 5’ RACE</td>
<td></td>
</tr>
<tr>
<td>Gene-specific primer <em>rpsF</em> 1</td>
<td>5’-ACCAAATCTTTGGTCCAGTG</td>
</tr>
<tr>
<td>Gene-specific primer <em>rpsR</em> 1</td>
<td>5’-CATCTACATCTTTGTAGCTAGTGC</td>
</tr>
<tr>
<td>Gene-specific primer <em>rpsF</em> 2</td>
<td>5’-GACTCTGGTCAATTTGGGCGG</td>
</tr>
<tr>
<td>Gene-specific primer <em>rpsR</em> 2</td>
<td>5’-CGCAGACCGCTGCTAGTCCTC</td>
</tr>
<tr>
<td>S6 operon 3’ RACE</td>
<td></td>
</tr>
<tr>
<td>dT primer</td>
<td>5’-GCGGTCAGCTTACTTAGCCCTCACCTGAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT</td>
</tr>
<tr>
<td>Gene-specific primer</td>
<td>5’-GATCAAACGCGCTAAATGG</td>
</tr>
<tr>
<td>Anchor primer</td>
<td>5’-GCGGTCAGCTTACTTAGCCCTACCTGAA</td>
</tr>
</tbody>
</table>
Table 2. Phage display selection setups.

<table>
<thead>
<tr>
<th>Round</th>
<th>Washing</th>
<th>Competitor RNA</th>
<th>Elution method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5 times; 0.3% Triton X-100</td>
<td>1.5 nmol tRNA</td>
<td>Specific</td>
</tr>
<tr>
<td>2</td>
<td>5 times; 0.3% Triton X-100</td>
<td>1.5 nmol tRNA</td>
<td>Low-pH; specific</td>
</tr>
<tr>
<td>3</td>
<td>5 times; 0.5% Triton X-100</td>
<td>1.5 nmol tRNA</td>
<td>Low-pH; specific</td>
</tr>
<tr>
<td>4</td>
<td>10 times; 0.5% Triton X-100</td>
<td>1.5 nmol tRNA; 0.3 nmol mutant</td>
<td>Low-pH; specific</td>
</tr>
</tbody>
</table>
Table 3. Selected peptides.

<table>
<thead>
<tr>
<th>Elution method</th>
<th>Selection 1</th>
<th>Selection 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-pH</td>
<td>(10) N-FSGGGNH-C</td>
<td>(2) N-QMLLRLP-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) N-NPTRKPK-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-IPPIQRR-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-IKRKSHR-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-PNRMPTM-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-RMKRPNN-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-HPRLRKH-C</td>
</tr>
<tr>
<td>Specific</td>
<td>(8) N-FSGGGNH-C</td>
<td>(3) N-MPTRPNK-C</td>
</tr>
<tr>
<td></td>
<td>(1) N-NYSYIPP-C</td>
<td>(2) N-MITTTRK-C</td>
</tr>
<tr>
<td></td>
<td>(1) N-TVNFKLY-C</td>
<td>(1) N-IITLSRK-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-RNIIPKS-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-IQLKINS-C</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1) N-KSKKRTI-C</td>
</tr>
</tbody>
</table>
References


