2009

Selection and characterization of RNA aptamers that detect a quaternary structure for ribosomal protein S7

Allison Layne Pappas
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Biochemistry, Biophysics, and Structural Biology Commons

Recommended Citation
Pappas, Allison Layne, "Selection and characterization of RNA aptamers that detect a quaternary structure for ribosomal protein S7" (2009). Graduate Theses and Dissertations. 10828.
https://lib.dr.iastate.edu/etd/10828

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Selection and characterization of RNA aptamers that detect a quaternary structure for ribosomal protein S7

by

Allison L. Pappas

A dissertation submitted to the graduate faculty in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Program of Study Committee:
Marit Nilsen-Hamilton, Major Professor
Gloria M. Culver
W. Allen Miller
F. Chris Minion
Michael J. Wannemuehler

Iowa State University
Ames, Iowa
2009

Copyright © Allison L. Pappas, 2009. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>iii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER 1. GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Literature Review</td>
<td>1</td>
</tr>
<tr>
<td>Current Study</td>
<td>29</td>
</tr>
<tr>
<td>Thesis Organization</td>
<td>33</td>
</tr>
<tr>
<td>References</td>
<td>33</td>
</tr>
<tr>
<td>CHAPTER 2. SELECTION OF RNA APTAMERS THAT BIND WITH HIGH AFFINITY AND SPECIFICITY TO RIBOSOMAL PROTEIN S7 AND THAT REVEAL A QUATERNARY STRUCTURE FOR S7</td>
<td>48</td>
</tr>
<tr>
<td>Abstract</td>
<td>48</td>
</tr>
<tr>
<td>Introduction</td>
<td>49</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>51</td>
</tr>
<tr>
<td>Results</td>
<td>59</td>
</tr>
<tr>
<td>Discussion</td>
<td>65</td>
</tr>
<tr>
<td>References</td>
<td>74</td>
</tr>
<tr>
<td>CHAPTER 3. THE BASIS OF A QUATERNARY STRUCTURE IN RIBOSOMAL PROTEIN S7 AS DETECTED BY APTAMER BINDING</td>
<td>88</td>
</tr>
<tr>
<td>Abstract</td>
<td>88</td>
</tr>
<tr>
<td>Introduction</td>
<td>89</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>91</td>
</tr>
<tr>
<td>Results</td>
<td>95</td>
</tr>
<tr>
<td>Discussion</td>
<td>99</td>
</tr>
<tr>
<td>References</td>
<td>110</td>
</tr>
<tr>
<td>CHAPTER 4. GENERAL CONCLUSIONS AND FUTURE DIRECTIONS</td>
<td>120</td>
</tr>
<tr>
<td>General Conclusions</td>
<td>120</td>
</tr>
<tr>
<td>Future Directions</td>
<td>122</td>
</tr>
</tbody>
</table>
LIST OF TABLES

CHAPTER 2. SELECTION OF RNA APTAMERS THAT BIND WITH HIGH AFFINITY AND SPECIFICITY AND THAT DETECT A QUATERNARY STRUCTURE FOR RIBOSOMAL PROTEIN S7

Table 1. Binding affinities, stoichiometry, and $\Delta H$ for aptamer binding to S7 and S4 obtained by ITC 77
Table 2. Affinities and Hill coefficients for S7 aptamers 78
Table 3. Affinities, maximum fraction bound, and Hill coefficients for aptamer A06-TR3 chimeras 79

CHAPTER 3. THE BASIS OF A QUATERNARY STRUCTURE IN RIBOSOMAL PROTEIN S7 AS DETECTED BY APTAMER BINDING

Table 1. Comparison of S7 mutations and their effect on RNA binding 113
Table 2. Affinities and Hill coefficients for S7 aptamers and select S7 mutants 114
LIST OF FIGURES

CHAPTER 1. GENERAL INTRODUCTION

Figure 1. 30S ribosomal subunit assembly

Figure 2. Alignment of the amino acid sequences of ribosomal protein S7 from various bacteria

Figure 3. The structure of S7 bound to 16S rRNA from the crystal structure of the 30S subunit from Thermus thermophilus

Figure 4. A view of S7 showing the concavity that makes intimate contact with 16S rRNA.

Figure 5. Organization of the str operon from E. coli

Figure 6. (a) The secondary structure of 16S rRNA in the region of S7 binding

(b) The secondary structure of the intercistronic region of str mRNA to which S7 binds.

CHAPTER 2. SELECTION OF RNA APTAMERS THAT BIND WITH HIGH AFFINITY AND SPECIFICITY AND THAT DETECT A QUATERNARY STRUCTURE FOR RIBOSOMAL PROTEIN S7

Figure 1. RNA sequences selected after round 14 of SELEX against S7

Figure 2. Representative ITC data obtained from the titration of protein into randomized SELEX pool or aptamer RNA

Figure 3. ITC data obtained from the titration of ribosomal protein S7 into truncated aptamers at 25°C

Figure 4. Binding curves obtained from the titration of ribosomal protein S7 into truncated aptamers and str mRNA at ~24°C using a nitrocellulose filter binding assay

Figure 5. Ability of aptamer truncations and point mutations to bind S7

Figure 6. The secondary structure of the S7 binding site of str mRNA and the predicted secondary structures of the truncated aptamers and binding curves for the aptamer chimeras

Supplemental Figure 1. ITC data obtained from the titration of S7 into binding buffer, the lower half of the 3' domain of 16S rRNA, and aptamers A09, A06, C04, B02, B04, and B11

Supplemental Figure 2. ITC data obtained from the titration of S4 into the randomized RNA SELEX pool and aptamers A06, C04, B04, and B11

CHAPTER 3. THE BASIS OF A QUATERNARY STRUCTURE IN RIBOSOMAL PROTEIN S7 AS DETECTED BY APTAMER BINDING

Figure 1. Crystal structure of Bacillus stearothermophilus adapted from Hosaka et al. showing S7 amino acids mutated in this work
Figure 2. Ability of S7 aptamers A06-TR3, C04-TR2, and D02-TR1 to bind mutant S7 derivatives 116

Figure 3. Representative results from size exclusion chromatography of S7, aptamer A06, and the S7/A06 complex 117

Figure 4. Binding curves for selected mutant S7 derivatives and the aptamers A06-TR3, C04-TR2, and D02-TR1 118

Figure 5. Ability of 16S rRNA, str mRNA, and aptamer RNA to bind various S7 mutants 119

CHAPTER 4. GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

Figure 1. 30S ribosomal subunit assembly in the presence of S7 aptamer A06-TR3 127

Figure 2. Dissociation of the A06-TR3/S7 complex over time 128
Here we report on the selection and characterization of RNA aptamers that recognize *E. coli* ribosomal protein S7. Ribosomal protein S7 plays two important roles in ribosome biogenesis: (1) as an assembly initiator, S7 nucleates the folding of the 3' major domain of 16S rRNA, and (2) it binds to the *str* operon and represses the translation of S12, S7, and EF-G. The primary and secondary structures of the S7 binding sites of rRNA and mRNA share limited sequence and structural homology and the required elements for high affinity binding have not been entirely elucidated. We have selected RNA aptamers that share very little primary sequence homology to either the S7 binding site of 16S rRNA or to the intercistronic region of *str* mRNA. Many of the aptamers are expected to fold into three-helix junctions, a structure particularly reminiscent of the mRNA. Interestingly, the aptamers exhibit cooperative binding with Hill coefficients of ~3 indicating that they are detecting a quaternary structure of S7.

We have found that the S7 aptamers use the same amino acids and structural elements to bind S7 as the rRNA and mRNA indicating that the same binding site is used for all three RNAs. With gel filtration, we were only able to isolate the aptamer/S7 complex at a 1:1 stoichiometry, indicating that the proposed quaternary structure of S7 is weak. However, deletion of the β-ribbon nearly eliminates cooperative aptamer binding suggesting that this structural element may be involved in protein-protein interaction. Furthermore, pre-treatment of native S7
with the N-terminal extension also results in a significant reduction in cooperative aptamer binding.

The results presented here suggest that S7 itself may undergo conformational rearrangement subsequent to 16S rRNA binding, and may help explain the strong temperature-dependent rearrangements at the binding site of S7 within the 16S rRNA. Furthermore, we propose that the weak, multimeric interaction of S7 may have a role in the retroregulation of S12. S7 may bind to the mRNA in a pre-multimerized form or multimerize subsequent to binding, resulting in ribosome stalling due to the multimeric obstacle. If the S7/S7 interaction is weak however, then it may be easily disrupted by repeated ribosome bombardment, causing eventual decay of the multimer and relieving some of the translational repression. Translational repression of the genes encoding S7 and EF-G would remain constant over time however, because the monomeric S7 bound more tightly to the intercistronic region would continue to prevent translational coupling with the upstream gene encoding S12.
CHAPTER 1. GENERAL INTRODUCTION

Literature Review

Ribosomes: providing an essential cellular function

In 1676, Antwon van Leeuwenhoek submitted a letter to the English Royal Society in which he described, through the use of a homemade microscope, the observation of living, single celled entities that he called “animicules” [1]. It was nearly five years later when the Royal Society finally accepted Leeuwenhoek’s discovery of microorganisms, which at the time was both astonishing and crucial to the advancement of science and medicine. Although it is thought that Leeuwenhoek possessed microscopes capable of magnifying an object five hundred times, few scientists would have believed that there would be much to see under such an objective. Indeed, at the time it was unimaginable that the survival and propagation of all living things depends on a microscopic network of molecules so complex that every discovery at the sub-cellular level continues to challenge our understanding of biology to this very day.

Within each living cell is a DNA genome. The genome, often considered a “blueprint” for life, gives rise to a vast assortment of RNAs and proteins which are essential players in the microscopic network of life. Although life, with all of its complexity, currently requires DNA, RNA, and proteins, the workings of the original cell must have been vastly different. The origin of life necessarily began with a molecule both capable of storing hereditary information and of self-replication.
Propagation would require that this molecule serve as a template for the synthesis of identical (or near identical) molecules for passage to its progeny. Moreover, evolution in response to environmental changes would require self-replication, giving rise to the short generation times currently observed (e.g. *E. coli* can replicate its genome and divide within twenty minutes). It is now widely believed that this versatile, primordial molecule capable of both information storage and self-replication may have been RNA [2, 3].

Fortunately, one does not need to travel back billions of years through time in order to observe the ability of RNA to serve both as the molecule of heritable genetic information and as a catalyst of biological processes; current life abounds with examples of RNA’s versatility. Today, RNA continues to store information in the form of messenger RNA (mRNA) which is a necessary intermediate in the process of making proteins from a DNA sequence. Furthermore, for RNA viruses, RNA still serves as the hereditary genomic material contained within the viral capsid. Examples of the ability of RNA to catalyze biological events are also plentiful. RNA enzymes, termed ribozymes, were first discovered in the 1980’s as being capable of RNA hydrolysis (e.g. RNase P [4, 5], Group I and II introns [6-8], and the spliceosome [9]). However, perhaps the best studied and most heralded ribozyme is the ribosome.

Ribosomes provide an essential service to all living cells: they convert information encoded in mRNA into proteins through the process of translation. It is through translation that the majority of a cell’s genes are expressed and normal cell
growth is allowed to occur. Impaired translation can lead to reduction in growth, and can render a cell incapable of responding to its environment. The enzymatic component of the ribosome is its peptidyl transferase activity that creates a peptide bond between two amino acids thus allowing for the formation of polypeptides. Less than three hundred years after Leeuwenhoek’s description of animicules, advancements in microscopy allowed for the direct observation of components within a single cell. Electron micrographs have revealed an abundance of ribosomes within both prokaryotic and eukaryotic cells. Furthermore, microscopy has revealed the overall structure of ribosomes and their in vivo association with other organelles. Leeuwenhoek could not have guessed at the complexity of interplay between subcellular structures yet to be discovered. Even more unimaginable is that a study of these organelles, such as the ribosome, would reveal clues to the very origin of our existence and the persistent success of life on Earth.

The ribosome: a bridge from RNA-based life to modern cells

Speculation that life began with an RNA molecule, and not a DNA molecule, is not a haphazard choice. DNA and RNA are nearly identical in structure and are both capable of storing the same amount of genetic information. The sugar of a ribonucleotide, however, possesses a 2’ hydroxyl group whereas a deoxyribonucleotide does not. It is this 2’ hydroxyl group that allows for the enzymatic activity of RNAs by providing a functional group for catalysis. For a primitive organism this dual functionality of RNA would be a necessity, but the 2’ hydroxyl group could also be a “toxic asset” because it makes the RNA significantly
more susceptible to degradation than a DNA molecule. Thus, evolution towards the more stable hereditable DNA molecule would be favored, but would not render RNA obsolete.

Whereas DNA has replaced RNA as the genomic material in living cells, so too have proteins taken over the vast majority of catalytic processes within cells. Like DNA, proteins are significantly more stable than RNA. Furthermore, their ability to adopt complex tertiary and quaternary structures and their more diverse chemistry allows for a greater variety of more efficient catalytic functions. Nevertheless, RNA continues to play a plethora of roles within the cell, serving as intermediates between DNA and protein (i.e. mRNA), as catalysts (e.g. in the ribosome), and as regulators of RNA degradation (e.g. siRNA), just to name a few.

Although the first life may have been RNA-based with no need for a ribosome, there is currently no organism that can live without it. The ribosome is a large ribonucleoprotein complex with two subunits, each composed of ribosomal RNA (rRNA) and protein. Before it was discovered that RNAs were capable of catalysis, it was believed that the rRNA served as a sort of scaffold onto which the proteins bound, and that the proteins provided the peptidyl transferase activity. We now know, however, that the proteins have a more supportive role; association of the ribosomal proteins with rRNA promotes the correct folding of the rRNA, a necessary step for proper ribosome function. From this collection of rRNAs and proteins arises a complex capable of deciphering the genetic code of mRNA and translating the information contained within. The amazing speed and accuracy of the
The ribosome serves life well, for it allows a single organism to respond to its environment and make adaptations necessary for its survival.

The ribosome is a ubiquitous organelle with a startling conservation across all kingdoms of life. Conservation exists both at the morphological and the primary sequence levels [10]. This conservation suggests that the role of the ribosome is so essential for the continuation of life that a high degree of selective pressure is exerted to maintain its structure and function. So few mutations have been introduced into the rRNA sequences of organisms that phylogenetic studies have been able to use this sequence data to elucidate the evolutionary tree of life. The ribosome is evolution’s artifact— an indispensible biological machine that has survived billions of years of evolution. Studying the ribosome allows us to peak back into time and speculate on how modern life evolved from what we hypothesize was an RNA World. The ribosome provides us with more than just clues as to the origins of life; however, studying the interaction of proteins and rRNA within the ribosome and the assembly of the entire ribonucleoprotein complex provides us with a deeper understanding of the basic molecular principles underlying cellular biology.

Moreover, an intimate knowledge of the ribosome and its assembly will make us better equipped to address a number of threatening medical issues, including various cancers [11], Diamond-Blackfan anemia [12], autism [13], and multi-drug resistant bacterial strains.
The bacterial ribosome: A good model

The ability to make proteins is so crucial to a cell’s viability that the ribosome, the organelle responsible for translating the genetic code of mRNA and synthesizing protein, can be found in every cell. While in bacterial cells ribosomes are found in the cytoplasm, the ribosomes of eukaryotic cells may also be associated with the endoplasmic reticulum, an organelle involved in protein folding and transport. The composition and structures of prokaryotic and eukaryotic ribosomes are very similar. Whereas the ratio of RNA to protein in the bacterial ribosome is 2:1, the eukaryotic ribosome contains twice as much protein as RNA. This ratio is influenced by the numbers of proteins within the ribosome: bacterial ribosomes contain approximately 53 r-proteins, whereas eukaryotic ribosomes have about 80. Despite this difference, electron micrographs reveal that the overall structures of the prokaryotic and eukaryotic ribosomes are the same. Each ribosome is composed of one small and one large ribonucleoprotein subunit that must associate with one another before translation occurs.

Although the prokaryotic and eukaryotic ribosomes share a common structure and function, there are a great many differences in how these two ribosomes are assembled. The eukaryotic cell poses certain challenges to the assembly process because of the sequestration of the genome within a nucleus. For example, the genes encoding ribosomal proteins are transcribed within the nucleus and then exported to the cytoplasm for translation, after which the ribosomal proteins must then be imported back into the nucleus for ribosome assembly. Prokaryotes, with
their lack of a nuclear membrane, avoid these trafficking issues. Moreover, eukaryotic cells make use of over 500 rRNAs, ribosomal proteins, snoRNAs, and other trans-acting factors during ribosome biogenesis, vastly outnumbering those used by prokaryotes [14]. Because of its relatively simple assembly, the bacterial ribosome has been widely studied by the scientific community. Specifically, the *E. coli* ribosome is the model from which we have derived most of our information about both ribosome assembly and the process of translation, and more recently has provided a wealth of structural information as well.

**An overview of the *E. coli* ribosome and translation**

The *E. coli* ribosome is a 2.5MDa complex referred to as the 70S ribosome (70S being a reference to its sedimentation coefficient). The 70S ribosome is composed of a small subunit (i.e. 30S) and a large subunit (i.e. 50S). Each subunit is itself composed of both RNA and proteins. The 50S subunit has two RNAs, the 23S (2,904 nt) and 5S (120 nt) rRNAs, and 33 ribosomal proteins. The 50S subunit contains the peptidyl transferase activity of the ribosome. The 30S subunit has only one RNA, the 16S (1,542nt) rRNA, and 21 ribosomal proteins. The 30S subunit, although not involved in catalysis, is the site of codon-anticodon interaction and thus plays an important role in maintaining the fidelity of the growing polypeptide.

The interface of the two subunits forms the A-site, P-site, and E-site. During translation, an incoming tRNA binds to the A-site where an mRNA codon is exposed. If the anticodon loop of the aminoacyl tRNA is complementary to the codon, then the nascent polypeptide chain is transferred from the tRNA at the P-site and is added to
the aminoacyl moiety at the A-site through the creation of a peptide bond. In order to free the A-site for the decoding of the next mRNA codon, the deacylated tRNA and peptidyl tRNAs are translocated to the E-site and P-site, respectively.

It is now known that the two ribosomal subunits have a similar organization. For each, the r-proteins associate with specific primary rRNA sequences and promote the correct folding of the rRNAs. In the assembled ribosomal subunits, different regions of rRNAs are recognizable as distinct structural domains. Our understanding of protein-RNA interactions and the assembly of the ribosome has been greatly advanced by the study of the 30S subunit.

**In vitro Assembly of the 30S subunit**

**Organization of 16S rRNA**

In 1960, a new cell-free *in vitro* translation system derived from *E. coli* allowed for rapid advancements in the field of translation research [15]. Besides leading to the elucidation of the genetic code [16, 17], it was observed that the ribosome could be split into the 50S and 30S subunits [18], and that dissociation and re-association of the subunits played a role in the process of translation [19]. In 1968, the 30S subunit was reconstituted *in vitro* from its RNA and protein components [20], allowing for rapid advancements in understanding how ribosomal constituents can spontaneously assemble and give rise to a ribosomal subunit.

The first component of either of the two ribosomal subunits to be characterized was the RNA. Sucrose gradient sedimentation allowed for
identification of the 16S and 23S rRNAs as distinct components of the ribosomal subunits [21]. In the time since the first partial 16S rRNA sequences from *E. coli* were obtained four decades ago [22, 23], sequences from numerous other prokaryotic organisms have been obtained. The high degree of selective pressure maintaining the structure and function of ribosomal components allowed for comparison of these sequences which led to predictions about the secondary structure of 16S rRNA [24, 25]. Later, the crystal structure of the ribosome confirmed that most of these base pair predictions were indeed correct [26].

The 1542 nucleotides of 16S rRNA form a number of helices that can be organized into four distinct domains (Figure 1a). Radiating from a central pseudoknot, these domains fold into distinct structural components of the 30S subunit. The 5’ domain contains ~560 nucleotides and forms the body of the subunit. The central domain has ~360 nucleotides and forms the platform, a structural feature also within the body. The 3’ major domain contains ~480 nucleotides and becomes the head. The 3’ minor domain is significantly smaller with only ~145 nucleotides, and contains the anti-Shine Delgarno sequence [27].

**Cooperative assembly of the 30S subunit**

It was not until the development of two-dimensional gel electrophoresis that the full panoply of ribosomal proteins was appreciated [28]. In the lab of Masayasu Nomura, the binding of single, purified 30S r-proteins to 16S rRNA was assessed by sucrose gradient sedimentation and gel electrophoresis [29]. These studies confirmed previous observations that assembly is a highly cooperative process
wherein the affinity of most of the r-proteins is enhanced by the prior binding of one or more other r-proteins [30]. These results led to the very first assembly map for the 30S subunit, which embodies the highly ordered and cooperative nature of its assembly (Figure 1b).

Refinements of the 30S assembly map were made possible by continued research on \textit{in vitro} reconstitution as well as RNA footprinting results, which mapped the contacts between the r-proteins and 16S rRNA [31, 32]. Although minor changes to the assembly map continue to be made, it is clear that ribosomal proteins are associated with particular domains of the 16S rRNA and that they promote 30S assembly in a cooperative manner. The assembly map reveals that r-proteins can be categorized as either primary, secondary, or tertiary binding proteins based on their ability to associate with 16S rRNA. Primary binding proteins are capable of binding directly to the rRNA in the absence of any other r-proteins and nucleate the folding of the 16S rRNA. This group includes the r-proteins S17, S4, and S20 that bind to the 5’ domain of 16S rRNA; S8 and S15 that bind to the central domain; and S7 that binds to the 3’ major domain. The binding of the five primary binding proteins S4, S7, S8, S15 and S17 and the secondary binding protein S16 is necessary and sufficient for the 16S rRNA to achieve a compactness similar to that seen in the 30S subunit [33].

\textit{In vivo} studies showed that even under sub-optimal growth conditions where rRNA was produced in a three-fold excess, a significant number of active 70S ribosomes were present [34]. This observation indicates that assembly must be
initiated by considerably fewer proteins than actually exist within the ribosomal subunits. The assembly initiator proteins, by definition, do not rely on the presence of other r-proteins in order to associate with 16S rRNA and thus do not exhibit cooperative binding. Furthermore, they are essential for formation of active ribosomal subunits. Pulse-chase experiments using a 30S in vitro reconstitution assay identified the primary binding proteins S4 and S7 as being assembly initiators [35]. The remaining primary binding proteins bind cooperatively to the 16S rRNA, with their affinity being enhanced by the prior binding of S4.

The secondary binding proteins are those that bind to the 16S rRNA only after one or more of the primary binding proteins have bound. This group includes ribosomal protein S16 of the 5' domain; S18 and S6 of the central domain; and S9 and S19 of the 3' major domain. Finally, the tertiary binding proteins are those that require the prior binding of one or more of the primary and secondary proteins in order to associate with the small subunit. This group includes S5 and S12 of the 5' domain; S11 and S21 of the central domain; and S2, S3, S10, S13, and S14 of the 3' major domain. The secondary and tertiary binding proteins serve to promote and stabilize the proper formation of the 30S structural features. For example, S16 has been shown to be a non-essential component of functional ribosomes [36]. Nevertheless, S16 plays an important role by increasing the rate of ribosome assembly through the suppression of an intermediate non-native structure of 16S rRNA that forms following the binding S17, S4, and S20 to the 5' domain [37].
Kinetics and *in vivo* assembly of the 30S subunit

While *in vitro* reconstitution assays of the 30S subunit have allowed us to elucidate the order of protein binding, the derived assembly map does not tell us anything about the kinetics of assembly. In fact, the 30S assembly map can be deceiving; while it does suggest an optimal order for protein association with 16S rRNA, subunit assembly is extremely dynamic. A number of studies have been undertaken to reveal both the kinetic and cooperative nature of assembly. When 30S subunits are reconstituted at a lowered temperature (i.e. 0°C) or the standard temperature (i.e. 42°C), the bases of the 16S rRNA can be analyzed for their reactivity. A comparison of the results allows the r-proteins to be assigned to different kinetic classes [38]. At 0°C, a group of r-proteins identified with similar reactivities as those in native 30S subunits were classified as early binders. Interestingly, most of these proteins were associated with the 5' domain of the 16S rRNA whereas the 3' domain contains mostly mid and late binding proteins. This result underscores the polar nature of subunit assembly, with the 5' domain folding first to form the body and the 3' domain last [39]. Although in this *in vitro* experiment polarity is independent of transcription, *in vivo*, the polarity of both transcription and subunit assembly are linked [40].

The kinetic experiment described above also revealed that 30S assembly is not strictly sequential, as many nucleotides could be placed into more than one kinetic class. Recently, time-resolved synchrotron X-ray footprinting was used to study the kinetics of ribosome assembly [41]. This method provides an advantage
over previous ones because it does not rely on the temperature-dependent assembly of the 30S subunit and allows for the study of base reactivity within ~10ms of initiating reconstitution. These results revealed multiphasic folding kinetics whereby many nucleotides were protected within the first 50-100ms, but whose saturation was achieved over an extended period of time. These results confirm what Talkington et al. refer to as the “assembly landscape” [42]. Assembly can proceed through any number of routes, some of which are more energetically favorable than others, but all of which can drive the formation of the 30S subunit. The ability of r-proteins to bind independently of each other prevents the formation of a bottleneck in the assembly process, which might otherwise occur under sub-optimal growth conditions.

**Ribosome Biogenesis**

**Growth-dependent biogenesis**

The formation of new, functional ribosomes represents a crucial task for actively growing cells. Without ribosomes, a cell is neither able to grow nor divide. In fact, ribosomes are so key to the success of proliferating organisms that they can make up 50% [43] of a bacterial cell’s dry weight. When a bacterium is presented with optimal growth conditions, it may have upwards of 10,000 ribosomes engaged in active translation, and as much as 60% of a bacterium’s energy may be devoted to making more ribosomes [44]. The 70S ribosome is a mega complex of protein and RNA and for all of its complexity- both structural and functional- its assembly is, surprisingly, spontaneous. That is, all of the information required for assembling a
functional ribosome is contained within the primary sequences of the rRNAs and r-proteins. This is not to say that there are not other proteins *in vivo* that aid in the process of assembly by, for example, lowering the activation barrier at certain steps. The ribosome is, however, a fully self-assembling macromolecular machine.

As has been discussed, protein synthesis is a key cellular function and without it there would be no life. In times of abundant nutrients and favorable environmental conditions, cells are able to “cash-in” on this wealth by growing and dividing. Population survival and expansion are the goals of all evolving organisms. For an organism to grow and divide however, it must synthesize a wealth of metabolic enzymes in addition to proteins involved in other processes, such as DNA replication, transcription, and translation. Efficient exploitation of the organism’s environment requires that it sequester and/or utilize the nutrients quickly, before its competitors squander the riches. In such times, it is of utmost importance that the cell produces as many ribosomes as possible, for it is only through the ribosome that environmental exploitation can occur. Furthermore, translational regulation is important because only those mRNAs encoding proteins valuable for survival in the current environmental conditions should be translated.

It is often the case however, that environmental conditions do not favor growth and division of the bacterial cell. Many factors contribute to the ability of a bacterium to grow and divide, including, but not limited to, temperature, available nutrients, and light. Under suboptimal growing conditions, energy spent on making more ribosomes is wasted energy, for a bacterium with no nutrient source neither
needs the enzymes required to metabolize the nutrient nor the ribosomes to produce the metabolic enzymes. Accordingly, bacteria exhibit growth rate dependent control of ribosome biogenesis; that is, as the growth rate increases, the number of ribosomes per unit cell mass increases linearly [45]. Furthermore, the production of rRNAs and r-proteins should be monitored to ensure the correct stoichiometry. In *E. coli*, two negative feedback mechanisms exist to ensure appropriate ribosome formation in conjunction with cell growth: (1) transcriptional repression of rRNA synthesis and (2) translational repression of r-protein synthesis.

**Regulation of rRNA production**

The *E. coli* genome contains seven copies of the *rrn* operon that encodes for the three rRNAs (i.e. 16S, 23S, and 5S): *rrnA*, *rrnB*, *rrnC*, *rrnD*, *rrnE*, *rrnG*, and *rrnH* [46]. In addition to the rRNA genes, there are a number of tRNA genes that are encoded between the genes for 16S and 23S rRNAs. Cotranscription of the rRNA genes allows for stoichiometric amounts of each rRNA to be produced. The regulation of rRNA synthesis is accomplished through four mechanisms: stringent control, growth rate regulation, upstream activation, and anti-termination [47]. While a number of effectors for regulation of rRNA production have been identified, guanosine tetraphosphate, ppGpp, and guanosine pentaphosphate, pppGpp (i.e. (p)ppGpp) are of particular interest.

In the stringent response, amino acid starvation triggers a prokaryotic cell to convert nearly all of its GTP and GDP to (p)ppGpp [48]. (p)ppGpp effectively shuts down transcription from the *rrn* operons, causing an almost immediate halt in rRNA
synthesis [49]. (p)ppGpp accomplishes this regulatory feat by binding to RNA polymerase and weakening the strength of the rrn operons’ P1 promoter [50, 51]. Furthermore, (p)ppGpp inhibits production of the Fis protein, a regulator of the rrn promoter. By preventing the synthesis of additional ribosomes, the existing ribosomes are able to function at a higher rate [52].

More recently, it has been shown that some ribosomal proteins play a role in regulating rRNA synthesis. For example, the production of (p)ppGpp entirely depends on the presence of an uncharged tRNA at the ribosome’s A-site and the r-protein L11 [53]. Ribosomal protein S4 is involved in another mode of rRNA synthesis regulation, antitermination of transcription. *In vivo*, S4 directly binds RNA polymerase and in *in vitro* assays can cause an ~11-fold increase in terminator read-through in conjunction with r-proteins L3, L4 and L13 [54]. These results reinforce the intimate link between rRNA and r-protein production in biosynthesis of the ribosome.

**Regulation of r-protein production**

While the genes encoding the rRNAs exist in seven copies within the *E. coli* genome, the genes encoding the r-proteins are present in only one copy. Many r-proteins lie within the same operon with about half in the *spc*, S10, *str*, and *α* operons. Expression of the r-proteins is tightly linked to rRNA production through feedback inhibition mechanisms. Within each operon exists a region to which one of the ribosomal proteins encoded within the operon can bind. In times of amino acid starvation, the levels of rRNA within the cell decrease. A reduction in rRNA
concentration results in an increased pool of free r-proteins. Some of the free r-proteins are then able to bind the polycistronic mRNAs and down regulate translation. This type of regulation is termed autogenous control and exists for the following operons, where the regulatory protein appears in **bold**: the α-operon, encoding r-proteins S13, S11, S4, and L17 [55]; the str operon, encoding r-proteins S12 and S7 [56]; the spc operon, encoding r-proteins L14, L24, L5, S14, S8, L6, L18, S5, L30, and L15 [57]; the L11 operon, encoding r-proteins L11 and L1 [58]; the S10 operon, encoding r-proteins S10, L3, L4, L23, L2, S19, L22, S3, L16, L29, and S17 [59, 60]; the L10 operon, encoding r-proteins L10 and L12 [61]; the L35 operon, encoding L35 and L20 [62]; and the S15 and S20 operons, encoding r-proteins S15 and S20, respectively [63, 64].

**Mechanisms of translational repression**

Cell survival depends not only on the tightly coordinated synthesis of the rRNA and r-proteins, but also on general transcriptional and translational control mechanisms. A bacterium’s response to, and its survival within, its environment requires transcription of the appropriate genes. One way in which this response is regulated is through the production of different sigma factors. Sigma is a subunit of RNA polymerase that recognizes sequences within the genome and recruits the polymerase to transcribe particular genes. Different sigma factors, such as heat-shock sigma factors (e.g. σ32 in *E. coli* [65] and σ8 in *B. subtilis* [66]), are produced under different environmental conditions, which allows for a coordinated increase in mRNA transcripts coding for genes that will allow the cell to survive and/or flourish in
its surroundings. However, for the rapid adaptation of bacteria to new environmental conditions, translational regulatory mechanisms are essential.

**Riboswitches: RNA-mediated translational repression**

Riboswitches are RNA elements located within 5' untranslated regions (UTRs) that control gene expression in response to a cellular metabolite. Riboswitches have two domains: (1) an aptamer domain that is able to bind a metabolite ligand, and (2) an expression platform that is capable of affecting the level of gene expression. Aptamers are nucleic acid molecules that bind a target molecule with both high affinity and specificity. The ligand binding property of aptamers is imparted by the secondary structure of the nucleic acid, and not exclusively by primary sequence. It is through ligand binding that the riboswitch either promotes or represses translation. Ironically, the *in vitro* selection of aptamers was accomplished before it was recognized that nature had already exploited the ligand-binding capability of nucleic acids. Nevertheless, riboswitches are omnipresent, having been located in all domains of life. In bacteria, riboswitches are extremely well-represented, with more than 2% of all genes being controlled by these metabolite-sensing RNAs [67].

Translational repression of a gene can be effectively accomplished by riboswitches through an allosteric conformational rearrangement of the RNA in response to metabolite binding. Regulation of gene expression requires the aptamer domain to be highly specific, and riboswitches that bind to a variety of metabolites and coenzymes have been reported (e.g. S-adenosylmethionine (SAM) [68],...
adenine and guanine [69], thiamine pyrophosphate (TPP) [70, 71], and Mg$^{2+}$ [72]). Binding of the metabolite to the aptamer domain causes a conformational change within the expression platform that can prevent ribosome binding through sequestration of the Shine-Delgarno (SD) sequence within the secondary structure of the 5'UTR. SD sequestration is observed for the TPP riboswitch of *E. coli* which is upstream of genes involved in importing or synthesizing thiamine and its phosphorylated derivatives [71]. Alternatively, binding of the metabolite may cause a conformational change within in the riboswitch that leads to the formation of a ribozyme that cleavages the mRNA, as is observed for regulation of the *glmS* gene, which encodes for glucosamine-6-phosphate (GlcN6P) synthetase, in Gram-positive bacteria [73].

**Protein-mediated translational repression mechanisms**

With the discovery of riboswitches, the roles of small noncoding RNAs (sRNAs) and cis-acting RNA elements in the control of gene expression have gained a lot of recent interest [74, 75]. Nevertheless, the effect on mRNA translation by cis-acting RNA elements when bound by a regulatory protein has been well documented for decades and the study of the r-protein operons has been particularly fruitful. When an r-protein binds to its mRNA regulatory region, translation is down-regulated either through premature transcription termination, such as for the S10 operon [76-78]; destabilization of the mRNA transcript, as for the *spc* operon[79]; or by prevention of 30S subunit association or entrapment of the 30S subunit in a non-active state, as for the $\alpha$-operon [80]. Translational repression of other operons
within the *E. coli* genome has also been described, although the mechanisms used are often the same as those for the r-protein operons [81, 82].

mRNA stability and degradation play an important role in regulating the proteomes of bacteria [83]. Translation of an mRNA is directly linked to its stability because the actively translating ribosomes provide a protective barrier around the transcript, shielding it from degradation by RNases [84]. In prokaryotes, polycistronic mRNAs have multiple ribosome binding sites, allowing the mRNA to be well protected. Furthermore, not all cistrons within an mRNA will degrade at the same rate since degradation is initiated by the endonuclease RNase E [85, 86]. The binding of a regulatory protein in the 5' UTR of a gene often competes, directly or indirectly, with ribosome binding. Often, the SD sequence can be sequestered within a hairpin structure that is stabilized upon protein binding. Sequestration prevents association of the ribosome with the mRNA and can result in reduced translation of an entire operon via translational coupling. Alternatively, the binding of a regulatory protein can directly interfere with the ribosome’s association with the SD sequence. These types of regulation are both seen in *B. subtilis* for the regulation of *trpE*, *trpD*, *trpG* (tryptophan synthesis), *trpP* (tryptophan transport), and *ycbK* (putative efflux protein) by the tryptophan-activated RNA binding attenuation protein (i.e. TRAP) [87]. Whereas the translation of *trpE* (and *trpD* through translational coupling [87]) is regulated by TRAP binding and sequestration of the SD sequence into a hairpin, translational repression of *trpG*, *trpP*, and *ycbK* are achieved when TRAP binds to
the 5' UTR, overlapping the SD sequence and/or the translation initiation sequence [82, 88, 89].

Protein-mediated translational repression can also be accomplished by entrapment of the ribosome on the mRNA. This type of regulation is observed for the α-operon that encodes r-proteins S13, S11, S4, and L17 and the α-subunit of RNA polymerase. Upstream of the gene encoding S13 is a nested pseudoknot structure that contains the SD and translation initiation sequences. This pseudoknot can exist in two different states, an active and an inactive conformation. While the 30S subunit is capable of binding either conformation, only one structure (i.e. the active) allows for the formation of a ternary initiation complex [90]. The regulatory r-protein S4 causes translational repression of the α-operon by binding and stabilizing the inactive form of the pseudoknot, resulting in entrapment of the 30S subunit [80]. A similar mechanism is observed for the translational regulation of the S15 operon [91, 92].

**Ribosomal protein S7**

**Structural and biophysical characteristics of S7**

The S7 ribosomal protein of *E. coli* strain K is a 19.9 kDa protein with 178 amino acid residues. The protein has a pI of ~10.3 and its highly basic nature is no doubt related to its function as a primary binding protein. A number of basic and hydrophobic amino acids of rpS7 are conserved across bacterial species, as revealed by sequence alignment (Figure 2). Despite the high degree of
conservation, two major forms of S7 exist among *E. coli* cells: whereas *E. coli* strain K harbors the 178 amino acid version, strain B harbors a version of S7 which lacks 24 amino acids at its C-terminus [93, 94]. Among the prokaryotes, the 154 amino acid version of S7 is far more common [95-97], and the C-terminal extension has been shown to be entirely dispensable for 16S rRNA binding [98-100].

Early studies of S7 reported that the protein had a large radius of gyration in solution (27 Å) as measured by small-angle X-ray scattering [101]. Moreover, proton NMR failed to identify any tertiary structure in solution [102]. These results seemed to indicate that the protein existed in an elongated, non-globular state when alone in solution and possibly also when associated with the ribosome [103]. The proteins used for these studies however, were prepared under denaturing conditions and refolded in the presence of 10 mM potassium phosphate. It appears however, that the low salt concentrations used in these studies prevented the refolding of the protein and precluded any analysis of secondary and tertiary structure.

In studies where denatured S7 was refolded in the presence of 100 mM NaCl, it became evident that the protein indeed has a rich secondary and tertiary structure. In contrast with the above results, neutron scattering showed that S7 has a globular shape with a radius of gyration of 14.5 Å [104]. Circular dichroism revealed that the secondary structure of S7 is composed of 53% α-helices and 41-64% of β-sheets [105]. Proton NMR provided evidence of a tertiary structure, demonstrating the spatial proximity of many apolar aliphatic residues to aromatic residues. However, early investigators noted the relative instability of the protein which was nearly
unfolded in the presence of only 2M urea. Furthermore, scanning calorimetry indicated that S7 has a melting temperature of only 43°C and that the unfolding was irreversible [106]. At temperatures above the melting temperature the protein was observed to aggregate, presumably as a result of the high salt concentration (350 mM KCl). Nevertheless, the presence of a secondary and tertiary structure for S7 had been established.

In more recent years, X-ray crystallography has revealed a wealth of information on the secondary and tertiary structure of S7. In 1997, the crystal structures of S7 from Bacillus stearothermophilus and Thermus thermophilus were solved at 2.5 Å and 1.9 Å, respectively [107, 108]. Both structures revealed that S7 consists of six alpha helices and a single anti-parallel β-sheet with the connectivity scheme α1-α2-α3-β1-β2-α4-α5-α6 [109]. Helices one through five comprise the hydrophobic core of the protein with the β-ribbon arm intervening between helices three and four and extending out into the environment. It was noted that a helix-turn-helix motif is adopted by the hydrophobic core, a structure reminiscent of that observed for the DNA architectural factor, histone-like protein HU that binds double-stranded DNA [110, 111]. Helices one, four, and six along with the β-ribbon form a concavity lined with a number of conserved basic and hydrophobic residues (Figure 4). It was proposed that this concavity forms a binding site for double-stranded RNA. The β-ribbon was also identified as a structure common to nucleic acid binding proteins, including r-protein L14 [112].
Interaction of S7 with 16S rRNA

Identifying the location of S7 binding on 16S rRNA has been greatly facilitated by RNA footprinting assays. Limited ribonuclease hydrolysis initially revealed that S7 probably binds to a fragment of 16S rRNA generated from the 3' end [32]. Chemical and enzymatic probing showed that S7 could protect large portions of the 3’ major domain, but not the 3’ minor domain, from modification [113, 114]. Moreover, upon S7 binding, a number of bases experienced enhanced modification suggestive of S7’s role in the structural rearrangement of 16S and cooperative 30S assembly. Additional information was provided through a series of cross-linking studies which also revealed the proximity of S7 with a number of bases within the lower half of the 3' major and 3' minor domains [115-118], including a crosslink to U1240 in the loop between helices 30 and 41 [119]. Further studies narrowed the possible binding site of S7 to bases within helices 29 to 32 and 41 to 43 [120, 121].

The independent assembly of the 16S domains into the 30S structural components having been established [35, 39], a fragment encompassing the presumed S7 binding site was synthesized by in vitro transcription [122]. This fragment, encoding the 3' major domain (nucleotides 926-1393), was bound by S7 with a Kd of ~0.63 µM. A fragment corresponding only to the lower half of the 3' major domain (nucleotides 926-986 and 1219-1393) was able to bind S7 with the same Kd. A series of mutations deleting one or more of the helices identified a minimal fragment required for S7 binding. This fragment is only 108 nucleotides in length and contains part of helices 28 and 42 and all of helices 29 and 43. The loops
connecting the helices also appear to be involved in S7 binding, with helices 28, 29 and 43 forming a three-helix junction [123].

A series of deletion and point mutations made in S7 and tested both in vitro and in vivo identified the structural features and amino acid residues important for 16S rRNA binding [99, 100]. Deletion of the C-terminal extension present in E. coli K was inconsequential, and a number of point mutations located throughout the S7 structure resulted in only a slight decrease in 16S binding. Surprisingly, deletion of the highly conserved β-ribbon resulted in a 0.20-0.34 relative affinity in vitro but was found to incorporate into functional ribosomes in vivo with 45% efficiency. Most mutations that abolished S7 binding were localized to the unstructured N-terminal amino acids. Deletion of the 17 N-terminal amino acids only bound 16S in vitro with 0.07 relative affinity, and in vivo was able to incorporate into functional ribosomes with 1.8-3.1% relative efficiency. Various point mutations within the N-terminal extension and loops between the α1 and α2 helices and the α2 and α3 helices yielded similar results. Localized hydroxyl radical probing using iron tethered at position four of the B. stearothermophilus S7 protein identified helix 43 of 16S rRNA as being in close proximity to the N-terminal extension [124]. It was thus proposed that the N-terminus plays an important role in 16S binding by clamping the rRNA into the electropositive concavity of S7.

The crystal structure of the 30S subunit has confirmed much of the biochemical and mutational information used to characterize the interaction between S7 and 16S rRNA. The 30S crystal structure from Thermus thermophilus shows S7
in intimate contact with 16S helices 28, 29, 41, and 43 (Figure 3) [125, 126]. Both S7 and S4 are bound at multi-helix junctions and presumably stabilize various helices of 16S, in accordance with their roles as assembly initiators. As predicted, the electropositive concavity of S7 is bound to 16S while the opposite, relatively neutral, surface is exposed at the 30S surface. In general, the results of the crystal structure were in good agreement with previous three-dimensional maps of the 30S subunit derived from neutron scattering and tritium bombardment [127, 128].

**S7 regulation of the str operon**

The tight regulation of ribosome biogenesis is, in part, accomplished through translational repression of ribosomal protein expression. A number of ribosomal proteins are able to repress translation of their own operons, including S7. The streptomycin (str) operon, or the str operon, contains four genes: *rpsL* (encoding the S12 r-protein), *rpsG* (encoding the S7 r-protein), *fusA* (encoding the translational elongation factor EF-G), and *tufA* (encoding the translational elongation factor EF-Tu) (Figure 5). In *E. coli*, between the *rpsL* and *rpsG* genes is a 96 nucleotide mRNA sequence to which the S7 r-protein can bind (Figure 5). Upon binding of this sequence, translation of the str operon is prevented. Because a low 16S rRNA concentration leads to a higher concentration of free S7 r-protein, S7 is available to bind to the intercistronic region of the str operon (*in vitro* $K_d = 0.15 \mu M$) and repress both the expression ribosomal proteins and the expression of translational elongation factors. Thus, both ribosomal protein and translational elongation factor
production can be down regulated when the intracellular concentration of 16S rRNA is low.

The translational repression of the genes within the *str* operon is accomplished only through translational coupling. For the coupling of S12 and S7, it appears that following synthesis of S12 the ribosome either remains attached to the mRNA and scans for the start codon of *rpsG* or jumps from the stop codon of *rpsL* to the start codon. Using a β-galactosidase reporter assay, it was determined that S7 is able to repress translation of S12, itself, and EF-G only through translational coupling and has no affect on independent initiation (EF-Tu is only slightly downregulated) [129]. With 10-20% of S7 being produced from independent initiation, it seems probable that there is a relatively large intracellular pool of S7, even when the relative abundance of 16S rRNA is low. Furthermore, the retro-regulation of S12 is not particularly tight and stoichiometric amounts of r-proteins are probably not synthesized from this operon. The benefit of this operon’s organization and its mechanism of translational regulation, if any, is unclear.

A comparison of the *str* intercistronic region and the S7 binding site of 16S rRNA reveals that there is limited sequence and structural homology between the two (Figure 6). A sequence alignment reveals four homologous stretches of five- to six-nucleotides in length that are shared between the S7 binding sites of the rRNA and mRNA [130]. It also appears that both sequences fold into a three-helix junction. The secondary structure of the mRNA is predicted to fold in such a way as to sequester the Shine-Delgarno sequence directly upstream of the *rpsG* gene and part
of the start codon. This structure is thought to facilitate the translational coupling of S12 and S7 by bringing the S12 stop codon and S7 start codon into very close proximity despite the 96 nucleotides between them.

*In vitro* binding assays have demonstrated that *str* mRNA uses the same amino acids and structural features to bind S7 as does the 16S rRNA. Notably, deletion of the N-terminal 17 amino acids was capable of reducing affinity of S7 for *str* mRNA by ~91% [131]. Deletion of the β-ribbon resulted in an ~70% reduction in affinity, whereas deletion of the C-terminal extension characteristic of the *E. coli* K12 S7 had no affect on binding. A point mutation in M115, an amino acid in the loop between the α4 and α5 helices that makes contact with 16S rRNA, reduces affinity of the protein for *str* mRNA by ~90%. The ability of S7 derivatives to bind *str* mRNA *in vivo* can be determined by monitoring the doubling time of bacteria expressing the mutant protein from a plasmid. If the mutant is able to bind the mRNA, then translational repression of the endogenous *str* operon will occur and the population’s doubling time will increase. Results from *in vivo* experiments have generally confirmed the *in vitro* results described above [100, 131].

**S7 within the 70S ribosome**

S7, like most r-proteins, is located at the periphery of the 30S subunit. The crystal structure of the 70S ribosome reveals that S7 and S11 lie at the interface of the 30S and 50S subunits, at the location of the E-site [132, 133]. Additionally, S7 has been shown to be crosslinked to mRNA [134] and tRNA at the A, P, and E-sites [135, 136]. These data suggest that S7 may have a role during translation, apart
from ribosome biogenesis. In 2003, Robert et al. showed that S7 and S11, which connect the head of the 30S subunit to the platform, have a functional interaction [137]. Mutation or deletion of the C-terminal amino acids 148-155, which were proposed to interact with S11, did not affect S7’s incorporation into 30S subunits but prevented the association of the 30S and 50S subunits. Moreover, this C-terminal mutation was shown to cause a marked increase in translational frameshifting, stop codon readthrough, and codon misreading. Deletion of S7’s anti-parallel β-sheets has also been shown to result in an increase of -1 and +1 frameshifting, presumably through destabilization of the E-site tRNA [138]. The integrity of S7 thus appears important for translational fidelity.

**Current Study**

**Aptamer selection**

Aptamers are short sequences of DNA or RNA that bind a target molecule with high affinity and specificity and can be considered functionally similar to antibodies. Unlike antibodies however, aptamers can be developed *in vitro* through iterative rounds of selection. In 1990, the first reports of aptamer selection came out of the labs of Jack Szostak [139] and Larry Gold [140] and since this time the literature has become riddled with reports of aptamer selection. The development of aptamers, termed SELEX (for Systematic Evolution of Ligands by Exponential enrichment), begins with a randomized pool of nucleic acid flanked by fixed sequences at its 5’ and 3’ ends. Randomized DNA pools are initially synthesized for the first round of selection and pools for subsequent rounds are amplified by PCR
using primers complementary to the 5' and 3' fixed sequences. The selection of RNA aptamers requires the inclusion of an RNA polymerase promoter sequence (e.g. T7) located within either the 5' or 3' fixed sequences or within one of the primers that anneals at either end. Following amplification, the randomized pool is mixed with the target molecule and allowed to incubate. After some time, the nucleic acid molecules incapable of binding the target molecule are washed away and the binders are retrieved. Following elution, selected RNA aptamers (but not DNA aptamers) must be reverse transcribed to generate a template for PCR, and the selected pool is again amplified using an error-prone DNA polymerase (e.g. Taq). RNA aptamers are then generated by transcribing the amplified pool, exploiting the inclusion of the promoter sequence within either the 5' or 3' end. The amplified nucleic acid and target molecule are then combined as before. The evolution of the pool is affected by various parameters, including the ratio of nucleic acid to target molecule, which can be differed at various rounds. After multiple rounds of selection, the pool is sequenced and the aptamers are analyzed for similarities in primary sequence and structural motifs.

The length of the randomized region is able to influence the success of selection. When random regions of 16, 22, 26, 50, 70, and 90 nucleotides in length were tested for their ability to evolve a leucine aptamer, both the shortest and longest segments tested did not successfully recover the expected motif [141]. Random sequences of 50-70 nucleotides in length were found to be optimal. A number of aptamers have been selected to small molecules such as
aminoglycosides [142], arginine [143], and ATP [144]. The selection of aptamers that are both high in affinity and specificity have also been reported for much larger biomolecules such as proteins [145-148]. The *in vitro* selection of nucleic acid molecules that adopt secondary and tertiary structures and recognize a target molecule has lent great support to the RNA World hypothesis.

**The use of aptamers for studying ribosomal proteins**

The process of nucleic acid selection has been previously used to study a collection of ribosomal proteins. For example, SELEX was used to develop an RNA aptamer that could bind eukaryotic ribosomal protein L22 [149]. This selection helped to identify the criteria for optimal binding of RNA to L22, and identified elements found within its natural binding partners, rRNA and Epstein-Barr virus expressed RNA 1 (EBER1). Selection has also been used to study the binding site of the *E. coli* primary binding protein S8. Aptamers were selected from a doped sequence corresponding to the S8 rRNA binding site. Selected sequences differed from the rRNA binding site and revealed that a core structure for S8 included three interdependent bases (nucleotides 597, 641, and 643) with an essential intervening adenine nucleotide at position 642 [150].

More recently, a similar strategy to that of SELEX was employed in the study of the binding of S7 to *str* mRNA [151]. S7 was incubated with fragments of *str* mRNA generated from DNase I digestion of the DNA region corresponding to the genes encoding S12, S7 and part of EF-G. After nine rounds of selection, a fragment of 109 nucleotides in length that corresponds to the S12-S7 intercistronic region was
identified. The selected RNA had twice the affinity for S7, presumably as a result of mutations in stem IV that stabilized the overall secondary structure. Surdina *et al.* [151] claim that this strategy of genomic hydrolysis followed by rounds of selection can be used to search for alternative sites of S7 regulation within the genomes of other organisms.

**The use of aptamers to study S7**

In the current study, SELEX was used to develop RNA aptamers that bind S7 with high affinity and specificity. With S7 playing such a critical role in ribosome biogenesis, it was our initial goal to use these aptamers to inhibit ribosome assembly. The selected aptamers did not share any gross sequence homology with either the S7 binding site of 16S rRNA or *str* mRNA. These results suggested that S7 could bind diverse sequences, and that secondary structure may be the most important factor for high affinity binding. During the characterization of the aptamers we discovered a peculiarity: all aptamers with high affinity bound S7 in a cooperative manner. Results from the studies of other aptamers have suggested that cooperative aptamer binding can occur, as for the aptamers of the nucleocapsid protein from human immunodeficiency virus-1 [152] and IgE [153]. The basis of cooperative binding of glycine by the bacterial glycine riboswitch (a tandem aptamer structure) has recently been reported [154]. The strong positive cooperativity among the S7 aptamers is notable however; whereas the glycine riboswitch of *Bacillus subtilis* has a Hill coefficient of ~1.5 [155], the S7 aptamers have reproducible Hill coefficients of ~3.
**Thesis Organization**

In the following work we have characterized the cooperative binding of the S7 aptamers both from the RNA and protein perspectives. As reported in chapter 2, mutational studies of the aptamers have identified primary and (presumed) secondary structures important for S7 binding. We have also used chimeric aptamer constructs to show that cooperative binding can be greatly reduced through shortening the helix lengths of the aptamers. In chapter 3, we show that the aptamers contact many of the same amino acid and secondary structural features on S7 as the 16S rRNA and str mRNA. Binding assays reveal amino acids that may be involved in the formation of an, as yet unreported, quaternary structure of S7.

Chapter 1 includes a literature review and a presentation of the current study. Chapters 2 and 3 contain work as described in the preceding paragraph and are papers to be submitted to journals for publication. The contribution of each author to these papers is as follows: Allison Pappas designed, performed, and analyzed all experiments and wrote the papers; Marit Nilsen-Hamilton and Gloria M. Culver provided suggestions for experimental design and collaborated in data interpretation; Marit Nilsen-Hamilton also edited each paper and aided in their revision. Chapter 4 includes general conclusions and proposed directions for future work.

**References**


148. Takemura, K., et al., *DNA aptamers that bind to PrP(C) and not PrP(Sc) show sequence and structure specificity*. Exp Biol Med (Maywood), 2006. 231(2): 204-14.


Figure 1. 30S ribosomal subunit assembly. (a) The secondary structure of 16S rRNA with its domains identified. (b) Assembly map of the 30S subunit (adapted from Holmes and Culver (J. Mol. Biol. (2005) 354, 340-357). Column coloring corresponds to the 16S domains as identified in (a). The lengths of the arrows indicate the relative time point in which a particular r-protein associates with the assembling 30S subunit. The intensity of the arrows indicates the degree to which an r-protein's association is positively influenced by a previously associated r-protein (i.e. the degree of cooperativity). Figures were created with Adobe Illustrator.
**Figure 2.** Alignment of the amino acid sequences of ribosomal protein S7 from various bacteria. The locations of $\alpha$-helices and $\beta$-sheets are identified above the sequence. Dashes represent the absence of an amino acid. Abbreviations of bacterial names represent the following species: *E. coli*, *Escherichia coli*; *G. stearothermophilus*, *Geobacillus stearothermophilus*; *L. salivarius*, *Lactobacillus salivarius*; *S. pyogenes*, *Streptococcus pyogenes*; *T. thermophilus*, *Thermus thermophilus*. Figure was created using Invitrogen’s alignment software AlignX and Adobe Illustrator.
Figure 3. The structure of S7 bound to 16S rRNA from the crystal structure of the 30S subunit from *Thermus thermophilus* (adapted from Wimberly et al., 2000). S7 is shown within the context of the entire 30S subunit (a) and close-up (b). 16S rRNA bases in helix 28 are colored in red; helix 29, blue; helix 41, orange; helix 42, green; and helix 43, yellow. Figures were created using Pymol.
Figure 4. A view of S7 showing the concavity that makes intimate contact with 16S rRNA. The basic residues are highlighted in blue (a) and the hydrophobic residues in pink (b).
Figure 5. Organization of the str operon from *E. coli* (adapted from Saito and Nomura, 1994). The presumed secondary structure of the intercistronic region between the *rpsL* and *rpsG* genes is numbered with reference to the start codon of *rpsG*. Figure was created using Adobe Illustrator.
Figure 6. (a) The secondary structure of 16S rRNA in the region of S7 binding. Helices that have been demonstrated to be in close contact with S7 are identified. Base-pairing is consistent with the crystal structure of the *Escherichia coli* ribosome. (b) The secondary structure of the intercistronic region of *str* mRNA to which S7 binds. Base numbering is in reference to the start codon of *rpsG*. The stop codon of *rpsL*, the Shine-Delgarno sequence, and the start codon of *rpsG* are boxed in grey. Primary sequences that are shared between the two RNAs are boxed in the same color. Base-pairing is based on RNA footprinting data. Figures were created using Adobe Illustrator.
CHAPTER 2: SELECTION OF RNA APTAMERS THAT BIND WITH HIGH AFFINITY AND SPECIFICITY TO RIBOSOMAL PROTEIN S7 AND THAT REVEAL A QUATERNARY STRUCTURE FOR S7

A paper to be submitted to the journal RNA

Allison L. Pappas, Gloria M. Culver, and Marit Nilsen-Hamilton

Abstract

Ribosome biogenesis is a tightly regulated process in all prokaryotes. The stringent response ensures that at times of amino acid starvation, 16S rRNA production is quickly stopped. Ribosomal protein production is also repressed through the autogenous regulation of many of the operons encoding ribosomal proteins. Ribosomal protein S7 plays two important roles in ribosome biogenesis: (1) as an assembly initiator, S7 nucleates the folding of the 3' major domain of 16S rRNA, and (2) it binds to the str operon and represses the translation of S12, S7, and EF-G. The primary and secondary structures of the S7 binding sites of rRNA and mRNA share limited sequence and structural homology and the required elements for high affinity binding have not been entirely elucidated. Here we report on the selection of RNA aptamers that share very little primary sequence homology to either the S7 binding site of 16S rRNA or str mRNA. Many of the aptamers are expected to fold into three-helix junctions, a structure particularly reminiscent of the mRNA. Interestingly, the aptamers exhibit cooperative binding with Hill coefficients of ~3
indicating that they are detecting a quaternary structure of S7. These aptamers show that S7 is able to bind a wide range of target RNAs and indicate that there may be a number of alternative binding sites for S7 within the bacterial transcriptome.

**Introduction**

Ribosomal protein S7 is a required component of the 30S subunit and its viability is essential for proper ribosome function. As one of only two assembly initiators, S7 plays a critical role in the assembly of the 30S subunit (Nowotny and Nierhaus 1988). S7, a primary binding protein, associates with the 3' major domain of 16S rRNA and nucleates folding of the head. Ribosomes that are assembled *in vitro* in the absence of S7 have gross structural perturbations and are incapable of translation. Chemical and enzymatic modification assays and cross-linking studies revealed that the S7 binding site is localized to helices 29 to 32 and 41 to 43 of 16S rRNA (Zimmermann et al. 1972; Moller and Brimacombe 1975; Wower and Brimacombe 1983; Greuer et al. 1987; Osswald et al. 1987; Powers et al. 1988a; Powers et al. 1988b; Urlaub et al. 1997b). A 108 nucleotide fragment corresponding to the minimal S7 binding site was later identified by Robert *et al*. This fragment contains a three-helix junction connecting part of helix 28 and helices 29 and 43 (Dragon and Brakier-Gingras 1993).

Ribosome biogenesis is a tightly controlled process in prokaryotic cells, and many of the ribosomal proteins also act as translational repressors of their own operons (Fukuda 1980; Zengel *et al*. 1980; Dean *et al*. 1981a; Changchien *et al*. 1986). In times of amino acid starvation, the levels of rRNA severely decrease
allowing the abundant pool of free r-proteins to bind the polycistronic mRNA (Baracchini and Bremer 1988). S7 is involved in the autogenous control of the \textit{str} operon, which contains the genes encoding ribosomal proteins S12 (\textit{rpsL}) and S7 (\textit{rpsG}) and elongation factors EF-G (\textit{tufA}) and EF-Tu (Dean et al. 1981b). S7 binds to the 96 nucleotide intercistronic sequence between \textit{rpsL} and \textit{rpsG} (Saito and Nomura 1994). Overexpression of S7 retroregulates translation of S12 and prevents the translational coupling of S7 and EF-G (Saito et al. 1994). Thus, S7 plays an important role in both the assembly of the 30S subunit and the regulation of ribosome biosynthesis.

A comparison of the rRNA and mRNA binding sites of S7 revealed that the two share limited primary and secondary structural similarities. Both RNAs were found to contain four sequences of five to six nucleotides in length and it was postulated that at least some of these sequences may be directly recognized by S7 (Nomura et al. 1980). This idea was supported by the direct crosslinking of S7 to U1240 of 16S rRNA (Urlaub et al. 1997a), and to U34 of \textit{str} mRNA (Golovin et al. 2006) both of which are found within one of these shared sequences. The secondary structure of the mRNA has a three-helix junction that may mimic the structure of the 16S rRNA site and promote binding of S7. However, the \textit{str} mRNA binding site of S7 is not entirely understood. Chemical and enzymatic probing localized the S7 binding site to a 59 nucleotide hairpin encompassing helices III, IV, and V within the intercistronic region (Saito and Nomura 1994), but the corresponding fragment experienced a six-fold loss in affinity compared with the full-length mRNA.
(Spiridonova et al. 1999). It thus appears that S7 requires some additional secondary structural feature or structural stabilization of the mRNA in order to bind.

To identify key structural elements and sequences needed for RNA binding by S7, we have used the method of SELEX (i.e. Systematic Evolution of Ligands by Exponential enrichment) to select RNA aptamers that bind S7 with a high affinity and specificity. These aptamers share some primary sequence homology with both the 16S rRNA and str mRNA and many appear to contain a three-helix junction. These results demonstrate that S7 can recognize a wide array of primary RNA sequences and that secondary structure may be a more important factor in determining high affinity. Furthermore, these aptamers were found to bind S7 in a highly cooperative manner, a characteristic at least partially dependent on helix length.

Materials and Methods

Expression and purification of S7 and S4

For isothermal titration calorimetry (ITC) and nitrocellulose filter binding assays, S7 and S4 were expressed from pET21 derivatives and purified using a MonoS cation exchange column attached to an AKTA FPLC (GE Healthcare) as previously described (Culver and Noller 1999). For SELEX, the gene encoding S7 was amplified from the pET21 plasmid using the forward primer

```
AAAAAAAAAAAAAAAAAAAAACGGAATTCGACGATGACGATAAGC
```

and the reverse primer

```
AAAAAAAAAAAAAAAAAAAAAAAAAGATCGAAGCTTTCAATTTAAGTAG
```
where the *EcoRI* and *HindIII* restriction enzyme sites are underlined, respectively. The amplicon was digested with *EcoRI* and *HindIII* and subsequently cloned into the pTrcHis vector (Invitrogen), allowing for production of S7 with an N-terminal hexahistidine tag.

XL1-blue cells carrying pTrcHisS7 were grown at 37 °C liquid LB broth supplemented with 100 µg/ml ampicillin until an OD600 of ~0.7 and then induced with 1mM IPTG (i.e. isopropyl β-D-1-thiogalactopyranoside). Cells were allowed to grow for another 3-4 h at 37 °C before harvesting the cell pellet. The cell pellet from 1 L of culture was resuspended in 20 ml 100 mM NaH₂PO₄, 10 mM Tris-HCl, and 8 M urea, pH 8.0 at 25 °C and cells were lysed by sonication. The cell lysate was mixed with Ni-NTA agarose (Qiagen) for 1.5 h at room temperature with constant stirring. The resin was then washed two times with resuspension buffer at pH 6.3 and S7 was subsequently eluted by lowering the pH to 5.9 and then to 4.5. S7 was refolded by dialyzing the protein against four 1 L changes of binding buffer (20 mM K⁺-HEPES, 330 mM KCl, 20 mM MgCl₂, pH 7.6) at 4 °C.

**Selection of S7 aptamers**

SELEX was performed using a 53 nucleotide randomized pool of the sequence GCCTGTTGTGAGCCTCCTGTCGAA-N(53)-TTGAGCGTTTATTCTTGTCTCCC (N=25A:25G:25C:25T). Initial amplification of the pool was done by Klenow extension using the reverse primer TAATACGACTCACTATAGGGAGACAGAGAATAAACGCTCAA, where the T7 promoter is underlined. Four nanomoles of primer were annealed to two nanomoles
of template in 90 µl by heating at 95 °C for five min followed by at least five min on ice in 10 mM Tris-HCl (pH 7.5) and 1 mM MgCl₂. The 90 µl annealing reaction was then added to a 1.2 ml extension reaction containing 0.05 U/µl DNA polymerase I large fragment (Promega), 0.5 mM each dNTP, 50 mM Tris-HCl (pH 7.2), 10 mM MgSO₄, and 0.1 mM DTT. Extension occurred at 37 °C for 30 min and the reaction was stopped by adding EDTA (i.e. ethylenediaminetetraacetic acid; pH 8.3) to 2.0 mM. Subsequent rounds of selection used PCR which also included the forward primer GCCTGTTGTGAGCCTCCTGTCGAA. PCR was carried out in 1 ml reactions containing 1 mM each dNTP, 2 µM forward primer, 2 µM reverse primer, 25 U/ml Taq polymerase (GenScript), 50 mM KCl, 10 mM Tris-HCl (pH 9.0 at 25 °C), and 0.1% Triton X-100. A relatively high concentration of MgCl₂, 7.5 mM, was included in all reactions to decrease the fidelity of the polymerase. PCRs contained variable concentrations of template which were derived from previous rounds of selection using reverse transcription as described below. PCR amplicons were purified by phenol/chloroform/isoamylalcohol (25:24:1) extraction and isopropanol precipitation.

For the first round of selection, a 125 µl transcription reaction was done using the Ampliscribe T7-Flash transcription kit (Epicentre) according to the manufacturer’s instructions. The reactions contained 0.4 nanomoles of Klenow-extended DNA template, 9 mM each NTP, 10 mM DTT, and 0.8 pmol/µl α-³²P ATP (MP Biomedicals) in 1X T7-Flash enzyme solution (components are proprietary) and were incubated at 42 °C for one hour. Subsequent rounds of transcription were
performed using the same conditions with purified PCR amplicons in a total volume of 100 µl. Transcripts were treated with 0.05 U/µl RNase-free DNase I (Epicentre) for 15 min at 37 °C and subsequently purified by phenol/chloroform/isoamylalcohol (25:24:1) extraction and isopropanol precipitation.

Binding reactions for S7 and the aptamer pool were conducted at ~24 °C in 20 mM K⁺-HEPES (pH 7.6), 330 mM KCl, and 20 mM MgCl₂ for 45 min to 2.5 h. Prior to assembling the binding reactions, RNAs were heated at 95 °C for five min and cooled to room temperature on the bench. Two selections were performed in parallel initially using a ratio of S7:RNA of 1:2 and 2:1 and these ratios were adjusted during SELEX to increase competition for S7 binding among the RNA pool. Throughout selection, the binding reaction volume varied between 0.1 and 1 ml, and the concentration of S7 was between 0.50 and 4.0 µM. To avoid selecting aptamers with an affinity for the partitioning matrix, S7-bound RNAs were separated from unbound RNAs either by filtration through a nitrocellulose filter (0.22 µm GSTF, 25 mm, from Millipore) or by passage through a Ni-NTA (Qiagen) column pre-equilibrated with 20 mM K⁺-HEPES (pH 7.6), 330 mM KCl, and 20 mM MgCl₂. RNA and protein was eluted from the filters by incubating with 7 M urea or from the columns by washing with 300 mM imidazole. RNA was separated from protein by phenol/chloroform/isoamylalcohol extraction and alcohol precipitated. Pre-negative selections were conducted during the later rounds of SELEX by passing the RNA pool through either the filter or Ni-NTA column prior to assembling the binding reaction.
For initiation of the next round of selection, eluted RNAs were reverse transcribed in 20 µl reactions using ThermoScript reverse transcriptase (Invitrogen). 20 pmoles of forward primer were annealed to the entire eluted pool from the previous round of selection (10 µl) in the presence of 1.7 mM of each dNTP by heating at 65 °C for five min then placing the reaction on ice for >2 min. Reverse transcription of the above annealing reaction was done with 5 mM DTT, 2 U/µl RNase OUT ribonuclease inhibitor (Invitrogen), and 0.75 U/µl ThermoScript reverse transcriptase in 1X cDNA synthesis buffer (50 mM Tris acetate (pH 8.4), 75 mM potassium acetate, 8 mM magnesium acetate) at 65 °C for 60 min. Reactions were stopped by incubation at 85 °C for five min. Reverse transcribed DNA was added directly to a PCR reaction without prior purification.

After fourteen rounds of selection, the aptamers were cloned into the pCR4 vector using the TOPO TA Cloning kit for Sequencing (Invitrogen) according to the manufacturer’s instructions. Plasmids carrying the selected aptamers were transformed into XL1-Blue cells and individual transformants were inoculated into 1 ml LB-amp. Cell cultures were submitted to the Iowa State University DNA Sequencing Facility for high-throughput plasmid preparation and sequencing. Aptamer sequences were aligned using the program AlignX (Invitrogen) and secondary structural predictions were done using RNAstructure 4.6 (Mathews et al. 2004).
In vitro transcription and preparation of aptamers and other RNAs

pFD3LH is a plasmid which contains the coding sequence for the lower half of the 3’ major domain and the 3’ minor domain of 16S rRNA (nucleotides 926-986/1219-1542), including the binding site for S7. Prior to transcription, 0.05 µg/µl of p16S3’ was linearized with 0.05 U/µl Rsa I at 37 °C for 3 h. This restriction fragment allowed for the subsequent in vitro transcription of the lower half of the 16S 3’ major domain (nucleotides 926-986/1219-1393) using T7 polymerase as described below.

Templates for the amplification of selected aptamers and their derivatives were synthesized by the Iowa State University DNA Facility. A DNA oligo encoding the intercistronic region plus the first 40 nucleotides of the rpsG gene from str mRNA was obtained from Integrated DNA Technologies, Inc. All templates included the T7 promoter for in vitro transcription. PCR was done using 0.2 mM of each dNTP, 0.5 µM forward primer, 0.5 µM backward primer, 0.02 ng/µl DNA template, and 0.025 U/µl Taq DNA polymerase (GenScript) in 50 mM KCl, 10 mM Tris HCl (pH 9.0 at 25 °C), 1.5 mM MgCl2, and 0.1 % Triton X-100. Amplicons were purified using the MinElute PCR purification kit (Qiagen) or the Wizard SV PCR & Gel purification kit (Promega) and quantified using a NanoDrop ND-1000 spectrophotometer at A260.

In vitro transcription of purified amplicons was done using 5 mM each NTP (i.e. ATP, GTP, CTP, and UTP), ~0.7 µCi/µl α-32P ATP, ~0.05 µg/µl template DNA, and 0.77 µM T7 polymerase in 10 mM DTT, 30 mM Tris, 2 mM spermidine, 0.001 % Triton X100, 20 mM MgCl2 at pH 8.5. Transcription reactions were incubated at 37 °C for ≥2 h. Following transcription, 0.05 U/µl RNase-free DNase I was added.
directly to the reactions and they were incubated at 37 °C for ≥15 min. Transcripts were purified by phenol/chloroform extraction and ethanol precipitation. RNA pellets were resuspended in ddH₂O and run through a Bio-spin P-30 Tris column (BioRad) to eliminate free radionucleotides. RNAs were then assessed for the percent incorporation of radionucleotides by TCA precipitation using ascending thin layer chromatography. Initial assays heat denatured the RNAs at 95 °C for ~3 min until it was determined that this step was unnecessary for aptamer binding.

For ITC, in vitro transcribed RNAs were synthesized as described above but were purified by phenol/chloroform extraction followed by ethanol precipitation. RNA pellets were resusupended in binding buffer from the last dialysis step of S7. RNAs were quantified by A260 using a Nanospec and diluted for use in isothermal titration calorimetry (ITC). All RNAs were heat denatured at 95 °C for ~3 min prior to use in ITC.

**Isothermal titration calorimetry**

All ITC was done using a VP-ITC machine from MicroCal. The last buffer change from S7 dialysis was used to thoroughly rinse the cell and syringe prior to titration experiments. Prior to titration, samples were degassed for ~5 min with stirring. For all titrations, 5-10 µM RNA was placed in the cell and the titrant was 31-70 µM S7 or S4. An initial injection of 2 µl (removed before data analysis) was followed by 30 injections of 10 µl each. Data was analyzed using Origin 7.0 with a one site binding model. The titration of S7 or S4 into the randomized SELEX pool was used as the reference data for titrations into aptamer RNA.
Nitrocellulose filter binding assays

For nitrocellulose filter binding assays, 30 µl binding reactions containing body-labeled RNA and S7 were allowed to incubate in binding buffer at ~24 °C for ≥15 min. For S7 binding to str mRNA, the RNA was first heated in buffer at 42 °C for 30 min prior to addition of S7. S7/mRNA binding was done at 30 °C for 30 min and the reaction was placed on ice for ≥10 min before filtration. For all binding reactions, the RNA concentration was held at a constant 1 µM while the protein was titrated. For filter binding assays done using GSTF nitrocellulose filters, filters were washed with 2 ml binding buffer prior to filtering the binding reactions. Filters were then washed with 3 ml binding buffer and counted using a Cerenkov protocol.

For filter binding assays using the BioDot filtration apparatus (BioRad), nitrocellulose (NitroBind, GE Water & Process Technologies) and nylon (MagnaCharge, GE Water & Process Technologies) membranes were soaked in binding buffer for ≥ 10 min prior to filtration and the device was used according to the manufacturer’s instructions. Briefly, 30 µl binding reactions were added to the wells and samples were filtered first through nitrocellulose and then through nylon. Each well was washed three times with 150 µl binding buffer. Following filtration, membranes were exposed over night to a storage phosphor screen that was subsequently imaged using a Typhoon (Molecular Probes). Quantification of bound RNA was done using ImageQuant. For non-linear regression of all binding curves, Origin 7.0 was used.
Results

Selection of RNA aptamers that bind r-protein S7

Selection of RNA aptamers resulted in a robust number of aptamer families that were able to bind S7. SELEX was carried out using an oligonucleotide mixture of 53-nucleotide randomized sequence and the hisS7 r-protein. Following 14 rounds of selection, the aptamer pool was cloned and a total of 84 oligonucleotides were obtained and sequenced. Analysis of the sequences revealed that most appeared more than once, with the largest aptamer family containing a total of ten members (Figure 1). Notably, a number of the oligonucleotides shared a five nucleotide sequence with the S7 binding sites of both 16S rRNA and the intercistronic region of the str operon mRNA.

Nine different aptamer families were assessed for their ability to bind S7 using isothermal titration calorimetry (ITC). The minimal 16S rRNA fragment, which was previously determined to bind S7, was used as a positive control and to establish binding conditions for ITC. The 16S rRNA fragment was found to bind S7 with a Kd of 0.89 µM (Table 1, Supplemental Figure 1b), which is very similar to the apparent dissociation constant of 0.63 µM determined by Dragon et al. (Dragon and Brakier-Gingras 1993) using a filter binding assay and slightly different buffer conditions. Titration of the S7 r-protein into the RNA SELEX pool revealed that S7 has no apparent affinity for the randomized RNA (Figure 2a). These data were used as the reference data and were subtracted from all binding curves generated from the titration of S7 into the RNA aptamers.
Most aptamers yielded a typical sigmoidal binding curve as exemplified by the titration of S7 into aptamer D02 seen in Figure 2b (see Supplemental Figure 1 for additional binding curves). Most ITC titrations resulted in ΔH (i.e. the change in enthalpy) values below -11 kcal and had typical binding curves that could be fit to the data using a single binding site model. Analysis revealed that the aptamer families exhibited a wide range of Kds from ~0.6 µM down to ~17 nM (Table 1). Aptamer families A06, C04, and D02 demonstrated the highest affinities for the S7 r-protein, with Kds around 20 nM. It is noteworthy that the aptamers with the highest affinities were all predicted to fold into three-helix junctions whereas the lower affinity aptamers were predicted to fold into stable extended stem loops (see, e.g., Figure 6b-d).

In an attempt to reduce the length of the aptamers and potentially stabilize the secondary structures, aptamers A06, C04, and D02 were truncated by removing parts of the 5′ and 3′ fixed ends that were not predicted to contribute to the secondary structure of the selected portion. These truncated aptamers were also tested for their ability to bind S7 by ITC (Figure 3). All binding curves were similar to those obtained for the full-length aptamers, including the initially increasing heat releases from the titration of S7 into the truncated A06-TR3 (Figure 3a).

ITC also reports the N-value for a given bimolecular interaction, where the N-value represents the number of ligand binding sites on the macromolecule. Interestingly, the N-values obtained by ITC indicated a stoichiometry of ≥ 3 S7 molecules per aptamer. This result was also obtained for the binding of the 16S
rRNA fragment, and thus probably partially indicates the fractional activity the RNA.
It is notable, however, that the N-values of the truncated aptamers increase by about 2-fold in all cases, indicating that the truncated aptamers make contact with more molecules of S7 than do the full-length aptamers.

**Specificity of the S7 aptamers**

All aptamer families were assessed for their specificity by testing their ability to bind the S4 r-protein. Like S7, S4 is a 30S subunit assembly initiator that binds directly to the 16S rRNA, thus making it a good indicator of the general protein-binding capacity of the S7 aptamers. As for the S7 titrations, S4 was first titrated into the randomized SELEX pool (Supplemental Figure 2a) and these data were used as a reference and subtracted from subsequent titrations during analysis. In the case of the highest affinity aptamer, D02, the $\Delta H$ value was too low to accurately fit a line to the data (Figure 2c). Similar results were obtained for aptamers A06, C04, B04, and B11, with all having extremely low $\Delta H$ values (Table 1, Supplemental Figure 2). These results demonstrated that the S7 aptamers are indeed selective in their ability to bind S7, despite both S7 and S4 being RNA-binding proteins and having pI’s of ~10.

**Aptamer binding is cooperative**

ITC was performed with high-affinity aptamers A06, C04, and D02 at 37 °C (Figure 2d). At this temperature, all titrations resulted in an initially progressive increase in heat release followed by a progressive decrease. This type of curve was
also observed for aptamer A06 at 25 °C. These curves could not be fit with a single binding site model, and generally required analysis with a five binding site model in order to achieve a good fit of the data. Although it is difficult to obtain an accurate estimate with this data, it appears that an approximately 10-fold increase in the Kds occurs at 37 °C (data not shown).

To elucidate the significance of the ITC N-values and the binding curves obtained at 37 °C, filter binding assays were performed for aptamers A06, C04, D02, and the truncated versions of all three. Titration of the S7 protein into each aptamer resulted in a sigmoid binding curve characteristic of a Hill plot (Figure 4). The Kds determined by this method were significantly higher than those obtained by ITC, with a 10-fold or more increase. The addition of 0.02% BSA to the binding reactions had very little affect on the Kds, but did reduce the maximum fraction bound. To determine if the sigmoidal binding curves were an artifact of our S7 preparation, we also titrated the protein into the intercistronic str mRNA fragment to which S7 has been previously shown to bind. This titration resulted in the expected hyperbolic plot with a Kd of ~0.17 µM, similar to the ~0.15 µM previously reported (Robert and Brakier-Gingras 2001) (Figure 4). We thus presumed that there was nothing fundamentally wrong with our preparation of S7.

The filter binding assays resulted in Hill coefficients from approximately two to more than five, indicating strong positive cooperativity in aptamer binding. Truncations C02-TR2 and D02-TR1 both exhibited Hill coefficients approximately double those for the untruncated aptamers (i.e. from ~2 or 3 to ~5.5), but inclusion of
0.02% BSA was able to reduce this number back down to ~3. Reversing the titration such that S7 was held at a constant concentration for which aptamer A06-TR3 binding was demonstrated to occur resulted in a hyperbolic plot (data not shown). This result indicated that cooperative binding is largely, if not exclusively, a result of protein concentration.

**Aptamer mutations**

To determine the primary and secondary features of the aptamers important for S7 binding, we screened a number of deletions and point mutations. For aptamers A06 and C04, we introduced point mutations in the five base sequences that also appear in the S7 binding sites of 16S rRNA and str mRNA. In both cases, we were careful to only make changes that were not predicted to perturb the presumed secondary structure. Changing the A06 sequence from UGAAU to UAAAU or UGGAU, and the C04 sequence from AGUAA to AGUGA completely abolished binding of both aptamers (Figure 5). Furthermore, we also changed the putative S7 binding site of D02 from AGUAC to include the five base sequence AGUAA and this also abolished binding. While it is possible that our point mutations did cause some structural change in the aptamers that prevented S7 from binding, it is also highly possible that these primary sequences are being directly recognized by S7 as in the rRNA and mRNA sequences.

We also tested a number of A06, C04, and D02 aptamer derivatives with whole or partial deletions of helices. All three aptamers and their truncations were predicted to fold into three-helix junctions (Figure 6b-d). Deleting single helices
resulted in the prediction of extended stem loop structures. None of these truncated aptamers bound S7, indicating a possible requirement for a three-helix junction to confer the ability to bind (Figure 5). We decided not to pursue obtaining Kds for these aptamer mutants since none of them appeared to have any affinity for S7 in our initial screen.

**Aptamer chimeras**

Our inability to obtain aptamer mutants that were capable of binding S7 led us to question the degree of flexibility in making changes to the primary and secondary structures of RNAs that bind S7. Having determined that the presence of a three-helix junction is probably a prerequisite for high affinity binding, we decided to construct a number of chimeric aptamers using aptamer A06-TR3 as the reference. Aptamer A06-TR3 has a comparatively short helix I and two other much longer helices that form its three-helix junction (Figure 6b). We initially replaced the 18 nucleotide helix IIIb with the approximately equivalent helix V from *str* mRNA (Figure 6a), helix IIIb from C04-TR2 (Figure 6c), helix III from D02-TR1 (Figure 6d), or a generic helix predicted to maintain the secondary structure of the aptamer. Eliminating A06-TR3 helix IIIb (i.e. A06-TR8) essentially abolished binding of S7, indicating that helix IIIa is not sufficiently long enough or the aptamer is no longer stable enough to allow binding (Figure 6f, Table 3). Alternatively, replacing helix IIIb with any other helix maintained binding and resulted in only slight increases in the Kd (Table 3).
Perhaps most striking is the reduction in the Hill coefficient as helix length is shortened. Replacing the 18 nucleotide A06-TR3 helix IIIb with the 10 nucleotide D02-TR1 helix III causes the Hill coefficient to be reduced from ~3.3 to ~1.7 (Table 3). Most impressive, however, is the 12 nucleotide generic helix IIIb for which cooperative binding is completely eliminated. Furthermore, this chimera has little or no change in its Kd compared to its parent A06-TR3. The other chimeras, with 15 or 16 nucleotides per helix, also experienced a reduction in their Hill coefficients.

A06-TR3 helix IIb chimeras experienced a significant reduction in their abilities to bind S7 (Figure 6e, Table 3). Helices derived from aptamers C04 and D02 and the generic IIb helix exhibit biphasic plots with a maximum fraction bound of ~0.2. Of the chimeras, the str helix IV derivative was best able to bind S7, although the binding curve is not saturated. Surprisingly, deletion of A06-TR3 helix IIb allowed for the greatest fraction of aptamer binding.

Discussion

In this study we have used the method of SELEX to obtain a number of RNA aptamers that bind ribosomal protein S7 with high affinity and specificity. S7 is a ribosomal protein of the 30S subunit and is essential for subunit assembly. As a primary binding protein S7 coordinates folding of the 30S head, binding directly to 16S rRNA and makes contact with the secondary r-protein S11 in the platform (Robert and Brakier-Gingras 2003). In addition to its association with the ribosome, S7 also controls ribosome biogenesis through translational repression of the str operon to which it belongs. In response to lower levels of intracellular 16S rRNA (i.e.
during times of amino acid starvation), S7 binds a region of mRNA between the genes encoding S12 and S7 and prevents the further translation of S12, S7, and elongation factors G and Tu. The S7 binding sites of the 16S rRNA and \textit{str} operon mRNA, although apparently structurally different, do share some primary sequence homology. Namely, each sequence shares four different five nucleotide sequences: AGUAA, UGGAU, UGAAU and ACAAU. Some of these sequences have been shown to make intimate contact with S7 in the mRNA, the first U of the UUGGA sequence, and in the rRNA the U in the ACAAU have been crosslinked to S7 (Urlaub et al. 1997a; Golovin et al. 2006).

After 14 rounds of SELEX, we identified a number of RNA aptamers that were able to bind ribosomal protein S7, but not S4. A total of 84 different sequences were obtained and the largest families were chosen for further study (Figure 1). Sequence analysis revealed that some of the aptamer families contained one of the five-base sequences shared by the S7 binding sites of 16S rRNA and \textit{str} operon mRNA mentioned above. The best represented family selected, B11, contains ten individual clones as well as the common S7 binding motif UGGAU. Somewhat surprisingly, however, was the fact that this family exhibited one of the lowest affinities tested (Table 1). Other smaller families with a native S7 RNA binding motif demonstrated much higher affinities for S7 in the low nanomolar range. Through SELEX we have selected aptamers with limited, but potentially significant, similarity to the 16S rRNA and \textit{str} operon mRNA at the primary sequence level, suggesting that these five nucleotide sequences are generally preferred by S7. As demonstrated by the low
affinity of family B11 though, possession of one of the five nucleotide sequences is not sufficient to confer high affinity. Furthermore, family D02 with a Kd of 17 nM lacks any of these consensus sequences altogether, although it does contain the similar sequence AGUACA.

When the high affinity aptamers A06, C04, D02, and their truncations A06-TR3, C04-TR2, and D02-TR1 were assessed for their abilities to bind S7 by a filter binding assay, we discovered that all aptamers exhibited a high degree of positive cooperativity in their binding (Figure 4, Table 2). These results confirmed our suspicion, based on the isothermal titration calorimetry results obtained at 37 °C, of cooperative binding (Figure 2d). While the N-values obtained by ITC could be questioned because our RNA purification procedures were not particularly effective at removing free nucleotides resulting in high error in the estimated RNA concentration, the Hill coefficients obtained from filter binding are more reliable. In the filter binding assays, the percent nucleotide incorporation of the RNA was determined by TCA precipitation, and RNA was kept at a constant concentration for all protein titrations. At S7 concentrations below ~0.2 µM, the aptamers are incapable of binding the protein (Table 2). These results suggest a quaternary structure for S7 that is detected by neither 16S rRNA nor str mRNA, and help to explain the difference in Kd’s obtained by ITC and filter binding assays. If aptamer binding is dependent on a multimeric form of S7, then at low protein concentrations where S7 is a monomer we expect to observe no binding as in Figure 4. The aptamers therefore have no affinity for monomeric S7, but have a high affinity for S7
in the multimeric form. During ITC, the concentration of S7 (the titrant) was between 31 and 70 µM and was thus in a multimeric form at the moment of injection. Upon injection, the aptamers are immediately able to bind, resulting in low nanomolar Kd’s. Thus, ITC experiments report the affinity of the aptamers for multimeric S7 and curves can be fit using a single binding site model (Figure 2b and 3). Filter binding assays do not result in hyperbolic curves, however, because a protein dilution series that allows for dissociation of multimeric S7 at lower stock concentrations is made prior to titration into the aptamers. The filter binding assays therefore monitor aptamer binding as a function of S7 quaternary structure resulting in a sigmoidal binding curve, a higher Kd value, and Hill coefficients that represent the degree of S7 multimerization.

The detection of a quaternary structure of S7 by the aptamers suggests that S7 may be recognizing a different set of structural features and/or primary sequences in the aptamer RNAs than in the rRNA or mRNA sequences to which it binds. We made a series of point mutations in the five base sequences that are shared by the aptamers and 16S rRNA and str mRNA. These point mutations completely abolished binding of S7, and although none were predicted to perturb the secondary structure of the aptamers we cannot eliminate this possibility (Figure 5). A compensatory mutation in C04 which restored the base pairing disrupted by the initial point mutation was also unable to bind S7, reinforcing the idea that the primary sequence plays an important role in the S7/RNA interaction. We also changed the D02 sequence AGUACA to the consensus sequence AGUAA and anticipated an
increase in binding. Interestingly, this mutation also eliminated S7 binding. These results show that the same five-base sequences that are important for S7 binding to 16S rRNA and str mRNA are also important for aptamer binding, although we cannot rule out the possibility that changes to the primary sequence resulted in a secondary structure that precluded S7 binding.

A secondary structure comparison of the aptamers with the intercistronic region of str mRNA to which S7 binds indicate a striking similarity: the high affinity aptamers A06-TR3, C04-TR2, and D02-TR1 are all predicted to fold into three-helical junctions similar to the one found in the mRNA (Figure 6a-d). All aptamer variants lacking a three-helical junction were incapable of binding S7 (Figure 5), reinforcing the idea that the three-helical junction is important for aptamer/S7 interaction. Thus, it appears that S7 probably recognizes similarities in both primary sequence and secondary structure among the aptamers and the native RNA sequences.

If the S7/aptamer interaction is based on the same RNA elements as the S7 interaction with 16S rRNA and str mRNA, then we wondered why we measure three or more aptamer binding sites per molecule of S7. By using a series of chimeric aptamer constructs derived from A06-TR3, we have shown that shortening (but not deleting) helix IIIb maintains aptamer binding but significantly reduces the Hill coefficient (Figure 6f, Table 3). Among the helix IIIb chimeras that demonstrate cooperative binding there is an apparent linear relationship between helix length and cooperative binding, with longer helices giving rise to larger Hill coefficients. Most
surprisingly, when helix IIIb is replaced by a non-selected, 12 nucleotide generic helix containing a GNRA tetraloop a hyperbolic plot is obtained. The elimination of cooperativity is accompanied by a Kd very similar to that seen for the parental aptamer A06-TR3. These results demonstrate that cooperative binding can be significantly affected by the aptamer sequence, not just helix length since the shorter, selected 10 nucleotide helix III from aptamer D02-TR1 has a Hill coefficient of ~1.7. We have proposed that the Hill coefficients may entirely reflect the dependence of aptamer binding on protein multimerization. Elimination of cooperative binding by changing the identity of helix IIIb could be achieved through a more stable, compact structure of the RNA. A compact aptamer may have better access to S7’s RNA binding site, reducing the chimeras’ dependence on the formation of a S7 quaternary structure. This idea is supported by the low Hill coefficient of the generic helix IIIb substitution that contains a GNRA tetraloop, an RNA motif known to stabilize hairpin structures (Heus and Pardi 1991). Lengthening helix IIIb may allow individual chimeras to form multiple alternative structures, leading to an increase in the Hill coefficients due to increased steric hinderance.

A06-TR3 chimeric aptamers that have had the identity of helix IIb changed experience a significant reduction in the fraction of aptamer that is able to bind S7. Furthermore, an S7 concentration of 5 µM was not enough to confirm saturation of the binding curves. Examination of the binding curves reveals a distinctly biphasic plot for all chimeras (Figure 6e, inset). This second phase in the data probably reflects non-specific binding, such as that observed for the deletion of A06-TR3 helix
IIIb. The first phase observed in these plots is probably the result of specific binding, although the aptamer has a much lower fractional activity. The low fraction bound of each of the helix IIb chimeras is most likely indicative of the aptamers adopting a structure not conducive to S7 binding (e.g. an extended stem-loop lacking a three-helix junction), and may indicate a discrepancy between the predicted and actual secondary structure of aptamer A06-TR3. It seems possible therefore, that S7 uses the same primary sequence and three-helix junction features to recognize both its native RNA targets and the aptamers. The dependency of aptamer binding on a multimeric form of S7 can be at least partially bypassed by changing the identity of A06-TR3 helix IIIb, presumably through stabilization of a more compact aptamer structure.

S7 is known to have a particularly low melting temperature and is prone to aggregation (Khechinashvili et al. 1978; Lamb et al. 2007). The expression, purification, and refolding of our S7 preparation followed a protocol previously published and our binding buffer is the same as that used for reconstitution of functional 30S ribosomal subunits (Culver 2000). If our protein preparation was aggregated due to incomplete folding or subsequent denaturation, then this probably would have resulted in an inability of the 16S rRNA or str mRNAs to bind. In testing the affinity of the rRNA and mRNA we obtained Kd's of 0.89 µM and 0.17 µM, respectively, values nearly identical to those reported within the literature (Dragon and Brakier-Gingras 1993; Robert et al. 2000). Furthermore, size exclusion chromatography and MALDI-TOF analysis revealed an abundance of monomeric S7
in our preparation (data not shown). These results show that, although S7 is prone to aggregation resulting in its precipitation from solution, the protein is largely monomeric while it is soluble.

A sigmoidal binding curve for S7 and an RNA target has been observed previously by Dragon et al. (Dragon and Brakier-Gingras 1993) who published the binding curve for a derivative of the 16S rRNA binding site of S7. This 16S rRNA fragment contained helices 28, 29, 30, 41, and 43, but not helix 42. Elimination of this helix resulted in only an approximately two-fold increase in the Kd, but resulted in a very characteristic Hill plot when S7 was titrated into the RNA over the same range used here, an interesting result which the authors did not address. It may be that the larger rRNA and mRNA sequences are unable to detect the quaternary S7 structure because they make a number of alternative contacts with the protein that result in sequestration of single protein molecules or that open the S7 molecule in a one-on-one interaction to allow binding to its basic cavity.

We are interested in exploring the in vivo significance of a quaternary structure of S7. A BLAST search using the aptamer sequences as the query revealed that there are numerous potential binding sites for S7 within bacterial transcriptomes, including intergenic regions. If these potential binding sites have short lengths, as in the case of the aptamers, they may also exhibit cooperative binding. An inability of S7 at low concentrations (such as those during growth) to bind these sequences may protect the cell against injurious translational repression. Alternatively, an ability of these sequences to bind S7 at higher concentrations (such
as those experienced just after rRNA synthesis has ceased) may facilitate translational repression and degradation of a variety of mRNAs beyond that of the \textit{str} operon. Analysis of the aptamers may reveal some of these alternative S7 binding sites, a method perhaps more robust than the SERF (Selection of Random RNA Fragments) approach proposed by Surdina \textit{et al.} (Surdina \textit{et al.} 2008).

Here we have described the selection of aptamers that bind ribosomal protein S7 with high specificity and affinity. These aptamers appear to recognize a heretofore uncharacterized quaternary structure of S7, revealing yet another potential application of aptamer selection: screening for higher-order structure among proteins. Moreover, the Hill-coefficients of the S7 aptamers are exceptional, being about one or more units larger than those reported for the cooperatively binding IgE aptamer (Gokulrangan \textit{et al.} 2005) or the tandem aptamer glycine riboswitch (Kwon and Strobel 2008).

The selection of RNA aptamers with a variety of sequences that bind S7 has shown that S7 can accept a highly diverse set of primary sequences for binding. While some of the aptamers reported here include an S7 consensus sequence, binding studies with aptamer D02 demonstrate that this is not a prerequisite for high affinity binding. The information gathered from the primary sequences and predicted secondary structures of these aptamers could reveal a robust strategy for predicting and identifying additional sites of translational repression by S7 within the genomes of various prokaryotes.
There has been recent interest in the development of novel antibiotics that target ribosome assembly (Champney 2003; Yassin et al. 2005; Maguire 2009). Aptamers against S7 have the potential of being developed into such antibiotics that could interfere with both 30S subunit assembly and translational regulation of the str operon. If they are effective antibiotics, the ability to obtain different aptamers that bind the same protein might provide a means of combating the anticipated bacterial resistance. If resistance to the aptamer antibiotics evolved, then one of the alternative chimeric aptamers could be substituted to restore efficacy.

References


Greuer, B., Osswald, M., Brimaxcombe, R., and Stoffler, G. 1987. RNA-protein cross-linking in Escherichia coli 30S ribosomal subunits; determination of sites on 16S RNA that are cross-


Table 1. Binding affinities, stoichiometry, and $\Delta H$ for aptamer binding to S7 and S4 obtained by ITC

<table>
<thead>
<tr>
<th>RNA</th>
<th>predicted structure(s)$^a$</th>
<th>Kd (μM)</th>
<th>~N (S7:RNA)</th>
<th>$\Delta H$ (kcal)</th>
<th>Kd (μM)</th>
<th>~N (S4:RNA)</th>
<th>$\Delta H$ (kcal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S 3' fragment</td>
<td>three-helix junction</td>
<td>0.89 ±7.8</td>
<td>3.1</td>
<td>-20 ±0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer A06</td>
<td>three-helix junction, extended stem loop</td>
<td>0.032 ±0.049</td>
<td>3.5</td>
<td>-10 ±0.52</td>
<td>2.0 ±1.7</td>
<td>0.64</td>
<td>-2.2 ±0.41</td>
</tr>
<tr>
<td>Truncation A06-TR3</td>
<td>three-helix junction</td>
<td>0.0007 ±0.031</td>
<td>8.3</td>
<td>-12 ±0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer A09</td>
<td>extended stem loop, three-helix junction</td>
<td>1.1 ±2.7</td>
<td>4.9</td>
<td>-43 ±25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer B02</td>
<td>extended stem loop, three-helix junction</td>
<td>0.57 ±2.6</td>
<td>2.3</td>
<td>-11 ±0.62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer B04</td>
<td>extended stem loop</td>
<td>0.22 ±0.88</td>
<td>2.8</td>
<td>-18 ±0.74</td>
<td>0.71 ±1.4</td>
<td>0.86</td>
<td>-2.9 ±0.21</td>
</tr>
<tr>
<td>Aptamer B11</td>
<td>extended stem loop</td>
<td>0.17 ±1.5</td>
<td>2.2</td>
<td>-19 ±0.27</td>
<td>1.0 ±0.46</td>
<td>36</td>
<td>6.1 ±59</td>
</tr>
<tr>
<td>Aptamer C04</td>
<td>three-helix junction, extended stem loop</td>
<td>0.024 ±0.16</td>
<td>3.4</td>
<td>-22 ±0.24</td>
<td>no detected binding$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncation C04-TR2</td>
<td>three-helix junction</td>
<td>0.0064 ±0.0015</td>
<td>6.3</td>
<td>-23 ±0.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer D02</td>
<td>three- and four-helix junctions</td>
<td>0.017 ±0.052</td>
<td>3.4</td>
<td>-11 ±0.25</td>
<td>no detected binding$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncation D02-TR1</td>
<td>three-helix junction</td>
<td>0.0059 ±0.0010</td>
<td>5</td>
<td>-13 ±0.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer D02b</td>
<td>three- and four-helix junctions, extended stem</td>
<td>0.29 ±0.69</td>
<td>2.8</td>
<td>-35 ±4.4</td>
<td>0.47 ±72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aptamer E05</td>
<td>stem loop</td>
<td>no detected binding$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$Secondary structure predictions were done with the program RNAstructure 4.6. The optimal structures (i.e. those with the lowest $\Delta G$) are listed.

$^b$A Kd could not be obtained from this data, either because a line could not be fit to the binding isotherms or because the $\Delta H$ values were too low to accurately assess binding.
Table 2. Affinity and Hill coefficients for S7 aptamers

<table>
<thead>
<tr>
<th>RNA</th>
<th>±0.02% BSA</th>
<th>Maximum fraction bound</th>
<th>Kd (µM)</th>
<th>Hill coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aptamer A06&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-</td>
<td>0.48 ±0.01</td>
<td>0.33 ±0.0094</td>
<td>3.6 ±0.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncation A06-TR3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
<td>0.59 ±0.00</td>
<td>0.54 ±0.0086</td>
<td>3.3 ±0.070</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0.34 ±0.00</td>
<td>0.71 ±0.014</td>
<td>3.3 ±0.60</td>
</tr>
<tr>
<td>Aptamer C04&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-</td>
<td>0.65 ±0.02</td>
<td>0.75 ±0.058</td>
<td>2.3 ±0.080</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncation C04-TR2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
<td>0.71 ±0.02</td>
<td>0.61 ±0.044</td>
<td>5.5 ±0.48</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0.25 ±0.0036</td>
<td>0.69 ±0.13</td>
<td>2.8 ±0.063</td>
</tr>
<tr>
<td>Aptamer D02&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-</td>
<td>0.39 ±0.04</td>
<td>0.73 ±0.13</td>
<td>3.2 ±0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truncation D02-TR1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
<td>0.65 ±0.00</td>
<td>0.68 ±0.017</td>
<td>5.4 ±0.35</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0.38 ±0.011</td>
<td>0.63 ±0.022</td>
<td>3.2 ±0.074</td>
</tr>
</tbody>
</table>

<sup>a</sup> These samples were filtered using 25mm nitrocellulose filters.

<sup>b</sup> These samples were filtered through nitrocellulose and nylon membranes using the BioDot filtration apparatus.
<table>
<thead>
<tr>
<th>RNA</th>
<th>helix removed</th>
<th>nucleotides/ helix</th>
<th>replacement helix</th>
<th>nucleotides/ helix</th>
<th>Maximum fraction bound</th>
<th>Kd (µM)</th>
<th>Relative affinity</th>
<th>Hill coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>A06-TR3</td>
<td>none</td>
<td>n/a</td>
<td>n/a</td>
<td>0.34 ±0.0</td>
<td>0.71 ±0.01</td>
<td>1.0</td>
<td>3.3 ±0.1</td>
<td></td>
</tr>
<tr>
<td>A06-TR8</td>
<td>llb</td>
<td>18</td>
<td>none</td>
<td>0</td>
<td>no detected binding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A06-str-V</td>
<td>llb</td>
<td>18</td>
<td>str mRNA helix V</td>
<td>0.69 ±0.0</td>
<td>0.97 ±0.05</td>
<td>0.73</td>
<td>2.6 ±0.2</td>
<td></td>
</tr>
<tr>
<td>A06-C04-III</td>
<td>llb</td>
<td>18</td>
<td>C04-TR2 helix III</td>
<td>1.0 ±0.0</td>
<td>1.1 ±0.04</td>
<td>0.65</td>
<td>2.0 ±0.1</td>
<td></td>
</tr>
<tr>
<td>A06-D02-III</td>
<td>llb</td>
<td>18</td>
<td>D02-TR1 helix III</td>
<td>0.80 ±0.0</td>
<td>0.71 ±0.04</td>
<td>1.0</td>
<td>1.7 ±0.1</td>
<td></td>
</tr>
<tr>
<td>A06-gen-IIIb</td>
<td>llb</td>
<td>18</td>
<td>generic helix llb</td>
<td>0.86 ±0.1</td>
<td>0.94 ±0.15</td>
<td>0.76</td>
<td>1.2 ±0.1</td>
<td></td>
</tr>
<tr>
<td>A06-TR9</td>
<td>llb</td>
<td>17</td>
<td>none</td>
<td>0.72 ±0.1</td>
<td>3.5 ±0.4</td>
<td>0.20</td>
<td>1.4 ±0.0</td>
<td></td>
</tr>
<tr>
<td>A06-str-IV</td>
<td>llb</td>
<td>17</td>
<td>str mRNA helix IV</td>
<td>0.25</td>
<td>n/a (biphasic plot)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A06-C04-II</td>
<td>llb</td>
<td>17</td>
<td>C04-TR2 helix II</td>
<td>0.25</td>
<td>n/a (biphasic plot)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A06-D02-IIb</td>
<td>llb</td>
<td>17</td>
<td>D02-TR1 helix llb</td>
<td>0.25</td>
<td>n/a (biphasic plot)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A06-gen-IIb</td>
<td>llb</td>
<td>17</td>
<td>generic helix llb</td>
<td>0.25</td>
<td>n/a (biphasic plot)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All titrations included 0.02% BSA and were filtered using the BioDot filtration apparatus.
<table>
<thead>
<tr>
<th>Family</th>
<th>Sequence</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>D02</td>
<td>AAGACCU-CUUGU-CAGUACGUACACUAGAAUCGAAACGACAUCUCGUUUACGAAACCACUGGAAGAUC</td>
<td>5/84</td>
</tr>
<tr>
<td>C04</td>
<td>CAACCU-CUUGU-CGUAACGGAAGACUUGAUGAGCCGA-UGAGCUUGAGAAGA</td>
<td>4/84</td>
</tr>
<tr>
<td>A09</td>
<td>UCGGACCCCU-CUCGGAGACGUACGUGACGUGAGAGUGUC</td>
<td>2/84</td>
</tr>
<tr>
<td>D02b</td>
<td>UAGGACCUUGUCAACACGUGAAGAAUUGAAUUCGCAACCUCCGAU</td>
<td>1/84</td>
</tr>
<tr>
<td>B11</td>
<td>ACAGAGACUC--UGACG-AGGAACGGAACUUGAUGACGUGAAGGUGA</td>
<td>7/84</td>
</tr>
<tr>
<td>B04</td>
<td>AAGGACUCUUGUCCACGGAACUUGAUGACGUGAAGGUGA</td>
<td>6/84</td>
</tr>
<tr>
<td>A06</td>
<td>CGUGAGGACUCUGCGAGCCUCACGGA-AUGACAGUGAAGGUGA</td>
<td>4/84</td>
</tr>
<tr>
<td>B02</td>
<td>UUCGGACUUGACGACUGGCACGUGAUCAGCCCGUCCCAACGCUUUCUGGCC</td>
<td>10/84</td>
</tr>
<tr>
<td>E05</td>
<td>AAGAAAGACCATCTCCCATCGACG</td>
<td>1/84</td>
</tr>
</tbody>
</table>

**Figure 1.** RNA sequences selected after round 14 of SELEX against S7. The sequences corresponding to the randomized regions for families assessed for S7 binding. The frequency of selection is indicated by the parenthetical numbers after each sequence. Families with 50% or more sequence identity are aligned with identical sequences boxed in grey. Sequences that appear in the S7-binding region of both 16S rRNA and the intercistronic region of *str* operon mRNA are outlined in a black box.
Figure 2. Representative ITC data obtained from the titration of protein into randomized SELEX pool or aptamer RNA. The titration of 31 µM S7 into 10 µM randomized SELEX pool (a) was subtracted from the titration into 10 µM aptamer D02 (b). To test for specificity, 72 µM ribosomal protein S4 was titrated into 10 µM randomized SELEX pool (see Supplemental Figure 2a), and these data were subtracted from the titration of S4 into 10 µM D02 (c). Whereas previous titrations were performed at 25 °C, aptamers demonstrating high affinity binding were also assessed for their ability to bind S7 at 37 °C. Panel (d) shows the titration of 46 µM S7 into 10 µM D02 at this elevated temperature. Dissociation constants and N-values for all aptamers at 25 °C can be found in Table 1. It should be noted that the highly cooperative binding observed at 37 °C precluded an accurate analysis of the data and Table 1 does not include this information.
Figure 3. ITC data obtained from the titration of ribosomal protein S7 into truncated aptamers at 25 °C. 28 µM S7 was titrated into ~10 µM RNA. The titration of S7 into the randomized SELEX pool (see Figure 2a) was subtracted from the titration into aptamer A06-TR3 (a), aptamer C04-TR2 (b), and aptamer D02-TR1 (c). Dissociation constants and N-values can be found in Table 1.
**Figure 4.** Binding curves obtained from the titration of ribosomal protein S7 into truncated aptamers and *str* mRNA at ~24 °C using a nitrocellulose filter binding assay. Binding data were plotted and non-linear curve fitting was done using the Hill equation for the aptamers and a hyperbolic model for *str* mRNA. Dissociation constants, maximum fractions bound, and N-values for the aptamers can be found in Table 2.
Figure 5. Ability of aptamer truncations and point mutations to bind S7. Binding was tested using a filter-binding assay and percent binding is normalized to the results for the unmutated aptamers. Aptamers were screened in triplicate and the standard deviation for all samples was less than 3%. Deletion mutations corresponding to A06-TR3, C04-TR2, and D02-TR1 have deletions of the 5' and 3' ends that were not predicted to perturb the secondary structure. C04 mutation A41G, U50C is a compensatory mutation that restores base pairing disrupted by the AGUAA to AGUGA mutation. Aptamer helices are identified in Figure 6.
Figure 6. The secondary structure of the S7 binding site of *str* mRNA and the predicted secondary structures of the truncated S7 aptamers. (a) The proposed secondary structure of the S12-S7 intercistronic region of *str* mRNA to which S7 binds. The secondary structures of aptamers A06-TR3 (b), C04-TR2 (c), and D02-TR1 (d) were predicted by the program RNAstructure 4.6. For the *str* mRNA, the *rpsL* stop codon and *rpsG* start codon are outlined in black. The five-base sequences shared between the S7 binding sites of 16S rRNA, *str* mRNA, and the aptamers are boxed in grey. Black dots appear next to bases that were mutated in this study. The binding curves for the interaction of S7 with A06-TR3 helix IIb (e) and IIIb (f) chimeras are shown. The inset graph in (e) has the y-axis scale adjusted to show the biphasic plot. Binding constants, maximum fraction bound, and Hill coefficients are given in Table 3.
Supplemental Figure 1. ITC data obtained from the titration of S7 into binding buffer (a), the lower half of the 3' domain of 16S rRNA (b), and aptamers A09 (c), A06 (d), C04 (e), B02 (f), B04 (g), and B11 (h). The titration into the SELEX pool (see Figure 2a) was subtracted from all other titrations prior to analysis. All titrations were performed at 25°C as described in the Materials & Methods section.
Supplemental Figure 2. ITC data obtained from the titration of S4 into the randomized RNA SELEX pool (a) and aptamers A06 (b), C04 (c), B04 (d), and B11 (e). The titration into the SELEX pool was subtracted from all other titrations prior to analysis. All titrations were performed at 25 °C as described in the Materials & Methods.
CHAPTER 3: THE BASIS OF A QUATERNARY STRUCTURE IN RIBOSOMAL PROTEIN S7 AS DETECTED BY APTAMER BINDING

A paper to be submitted to The Journal of Biological Chemistry

Allison L. Pappas, Gloria M. Culver, and Marit Nilsen-Hamilton

Abstract

S7 is one of two assembly initiator ribosomal proteins of the 30S subunit. In addition to binding 16S rRNA, S7 also translationally represses the \( \text{str} \) operon and regulates the expression of S12, S7, and EF-G. The integrity of S7 is important and growth defects can result from its inability to both assemble into ribosomes and associate with \( \text{str} \) mRNA. We previously selected a number of RNA aptamers that bind S7 with low nanomolar affinities. Surprisingly, aptamer binding exhibited strong positive cooperativity with Hill coefficients of \(~3\) suggesting that S7 has a quaternary structure in solution. Here we report that the aptamers bind the same amino acids and structural elements of S7 as the rRNA and mRNA, indicating that the same binding site is used for all three RNAs. With gel filtration, we were only able to isolate the aptamer/S7 complex at a 1:1 stoichiometry, indicating that the proposed quaternary structure of S7 is weak. However, deletion of the \( \beta \)-ribbon nearly eliminates apparent cooperative aptamer binding suggesting that this structural element may be involved in protein-protein interaction. Pre-treatment of native S7 with its own N-terminal fragment also caused a significant reduction in apparent
cooperative binding. We propose that S7 exists in both monomeric and multimeric forms, which may have important implications for \textit{in vivo} ribosome biogenesis.

**Introduction**

S7 is a 19.8 kD protein of the bacterial small ribosomal subunit. Its role in ribosome biogenesis is twofold: first, it is one of only two assembly initiators for the 30S subunit and second, it acts as a translational repressor of the \textit{streptomycin} (\textit{str}) operon, regulating expression of S12, S7, and EF-G (1). The crystal structures of S7 from both \textit{Bacillus stearothermophilus} and \textit{Thermus thermophilus} have been solved down to 2.5 Å and 1.9 Å, respectively (2,3). Both structures revealed that S7 consists of six alpha helices and a single anti-parallel β-sheet with the connectivity scheme α1-α2-α3-β1-β2-α4-α5-α6 (4). Helices one through five comprise the hydrophobic core of the protein with the β-ribbon arm intervening between helices three and four and extending out into the environment. It was noted that a helix-turn-helix motif is adopted by the hydrophobic core, a structure reminiscent of that observed for the DNA architectural factor, histone-like protein HU (5,6). Helices one, four, and six along with the β-ribbon form a concavity lined with a number of conserved basic and hydrophobic amino acid residues. The crystal structure of the 30S subunit reveals that this electropositive concavity of S7 binds 16S rRNA (7).

S7 binds selectively to mRNA and rRNA despite significant differences in their primary and secondary structures (8). Nevertheless, mutational analysis of S7 has revealed that the same amino acids and structural features are used in the recognition of both of these RNAs. For example, although the β-ribbon contributes to
formation of the S7 concavity, mutants without this structure are able to incorporate into functional ribosomes \textit{in vivo} with 45\% efficiency (9). Furthermore, this deletion mutant maintains its ability to repress translation of the \textit{str} operon. Deletion and point mutations of the N-terminal extension of S7, which may be responsible for securing RNA within the binding site of S7 by acting as a clamp (10), are most effective at abolishing binding. \textit{In vitro} assays show that such constructs neither bind 16S rRNA nor \textit{str} mRNA and results from \textit{in vivo} studies corroborate this data (9,11,12).

We previously reported the selection of RNA aptamers that bind S7 with apparent cooperativity and that have Hill coefficients of ~3. Aptamers are small nucleic acids that bind their target molecule with a high degree of affinity and specificity. While there are many reports of aptamer selection, very few reports of cooperatively binding aptamers exist. Only the prokaryotic glycine riboswitch, a naturally occurring tandem arrangement of two aptamers, has a demonstrated cooperative binding mechanism and a Hill coefficient of ~1.6 (13,14). Aptamers selected against the nucleocapsid (NC) protein of the human immunodeficiency virus-1 appear to have two binding sites for NC and the authors propose a multiple and cooperative binding scheme (15). Possible cooperative binding of aptamers to IgE has also been reported (16). The selection of aptamers that demonstrate cooperative binding to NC is not particularly surprising since nucleocapsid proteins are known to associate both \textit{in vitro} and \textit{in vivo}. Among these reports, the S7
aptamers are exceptional because their Hill coefficients far surpass those identified for other aptamers.

Here we characterize aptamer binding with respect to S7. Mutational analysis has revealed that the aptamers use the same amino acid and structural features to bind S7 as do the 16S rRNA and str mRNA. We find that cooperative binding of the aptamers can be nearly eliminated by deleting the β-ribbon of S7. Furthermore, pre-treatment of native S7 with its own N-terminal fragment causes a significant reduction in the Hill coefficients of all aptamers. We propose that the aptamers recognize a weak quaternary structure of S7 and that this structure results from the interaction of the β-ribbon with the N-terminus of adjacent S7 molecules.

**Materials and Methods**

**Creation of S7 mutants**

The pET21 plasmid containing the *rpsG* gene encoding S7 was the template for site-directed mutagenesis. All mutations were created using the QuikChange II kit (Stratagene) according to the manufacturer’s instructions and using primers as previously described (9,12). The identity of all plasmids was confirmed by sequencing using the primers PETFOR (TAATACGACTCACTATAGGG) and/or PETREV (GCTAGTTATTGCTCAGCGG) at the Iowa State University DNA Facility.

**Expression and purification of S7 and mutants**
The pET21 plasmid containing the gene encoding S7 and its derivatives were expressed and purified using a MonoS cation exchange column (GE Healthcare) attached to an AKTA FPLC (GE Healthcare) as previously described (17).

The N-terminal region of the S7 protein, including amino acids 1-19, was synthesized by GenScript and was reported to be 89% pure by HPLC. The lyophilized N-terminal fragment was resuspended in distilled deionized water ddH₂O and diluted into binding buffer (20 mM K⁺-HEPES, 330 mM KCl, 20 mM MgCl₂) for further use. No further purification of the N-terminal fragment was done.

**In vitro transcription and RNA preparation of aptamers**

Templates for the amplification of aptamers A06-TR3, C04-TR2, and D02-TR1 were synthesized by the Iowa State University DNA Facility. All templates included the T7 promoter for *in vitro* transcription. PCR was done using 0.2 mM each dNTP, 0.5 µM forward primer, 0.5 µM backward primer, 0.02 ng/µl DNA template, and 0.025 U/ul Taq DNA polymerase (GenScript) in 50 mM KCl, 10 mM Tris HCl (pH 9.0 at 25°C), 1.5 mM MgCl₂, and 0.1 % Triton X-100. Amplicons were purified using the MinElute PCR purification kit (Qiagen) or the Wizard SV PCR & Gel purification kit (Promega) and quantified using a NanoDrop ND2000 spectrophotometer at A260.

*In vitro* transcription of purified amplicons was done using 5 mM each NTP (i.e. ATP, GTP, CTP, and UTP), ~0.7 uCi/µl α-³²P ATP, ~0.05 µg/µl template DNA, and 0.77 µM T7 polymerase in 10mM DTT, 30mM Tris, 2mM spermidine, 0.001 % Triton X100, and 20 mM MgCl₂ at pH 8.5. Transcription reactions were incubated at
37 °C for ≥ 2 h. Following transcription, the DNA template was digested by adding RNase-free DNase I to 0.05 U/µl directly to the reactions and incubating at 37 °C for ≥15 min. Transcripts were purified by phenol/chloroform extraction and ethanol precipitation. RNA pellets were resuspended in ddH2O and run through a Bio-spin P-30 Tris column (BioRad) to eliminate free radionucleotides. RNAs were then assessed for their percent radionucleotide incorporation by TCA precipitation using ascending thin layer chromatography. No RNAs were heat-denatured prior to use in binding assays as this step was found to be unnecessary.

**Nitrocellulose filter binding assays**

For screening the protein mutants, nitrocellulose filter binding assays were performed with 0.22 µm 25 mm GSTF nitrocellulose filters (Millipore). Thirty µl binding reactions containing 0.5 µM radiolabeled RNA and 0.5 µM S7 were allowed to incubate in binding buffer at ~24°C for ≥ 15 min. Filters were washed with 2 ml binding buffer prior to filtering the reactions. Following filtration, filters were washed with 3 ml binding buffer and counted using a Cerenkov protocol in a scintillation counter.

For determination of affinity, protein was titrated into 1 µM aptamer RNA and 0.02 % Ultrapure BSA (Ambion) in binding buffer. The 30 µl binding reactions were filtered using the BioDot filtration apparatus (BioRad) according to the manufacturer’s instructions. Nitrocellulose (NitroBind, GE Water & Process Technologies) and nylon (MagnaCharge, GE Water & Process Technologies) membranes were soaked in binding buffer for ≥10 min prior to filtration. Samples
were filtered first through nitrocellulose and then through nylon filters. Each well was washed three times with 150µl binding buffer. Following filtration, membranes were exposed over night to a storage phosphor screen that was subsequently imaged using a Typhoon 8600 (Amersham Pharmacia Biotech). Quantification of bound RNA was done using ImageQuant software. For non-linear regression of all binding curves, Origin 7.0 was used. Curves were fit using the Hill equation.

**Gel filtration**

Gel filtration was performed using a Superdex 75 column attached to an AKTA-FPLC (GE Healthcare) at both 4 °C and 25 °C. The column was pre-equilbrated with ~2 column volumes of binding buffer before the first sample was run, and was then washed with ~0.5 column volumes before applying the next sample. Following centrifugation at 16,000 g at 4 °C for ≥10 min, ~300 µl of ~1 mg/ml blue dextran (to establish V₀), BSA, carbonic anhydrase, aprotinin, or vitamin B12 in binding buffer was injected onto the column. Plotting the Ve/V₀ versus molecular weight allowed for linear regression to generate the standard curve. Samples of 10 or 80 µM S7 and 10 µM aptamer A06 were centrifuged as above and injected onto the column. For filtration of the S7/A06 complex, 2 µM A06 was incubated with 2 µM S7 in binding buffer for 30 min at ~24 °C. The binding reaction was centrifuged as above and injected onto the column. Comparison of the values obtained for Ve/V₀ to the standard curved allowed for an estimation of the molecular masses of S7, A06, and the S7/A06 complex.
Results

S7 residues important for aptamer binding

We previously identified three RNA aptamers that bind with high affinities and specificities to ribosomal protein S7. These aptamers exhibited cooperative binding suggesting that they all recognize a quaternary structure of S7. In vivo, S7 binds to two RNAs, the 16S rRNA and the S12-S7 intercistronic region of the str operon mRNA. We decided to determine if the aptamers were binding to a different set of S7 residues and/or secondary and tertiary structural elements to bind S7 than the two native RNA sequences. Accordingly, we made a number of S7 point mutations and deletion mutations that were identified by Fredrick et al. (9) as resulting in reduced incorporation of the protein into 30S subunits in vivo or by Robert et al. (11,12) as having reduced affinities for 16S rRNA or str mRNA in vitro (Figure 1).

For screening the ability of the aptamers to bind the S7 mutants, we used a nitrocellulose filter binding assay. Our previous results suggested that aptamer binding is dependent on a quaternary structure for S7 and the manifestation of this dependency is a characteristic Hill plot (18). Apparent dissociation constants for the aptamers were determined by non-linear line fitting of the sigmoid binding curves using the Hill equation. The concentrations of aptamer A06-TR3, C04-TR2, and D02-TR1 RNA and protein used in screening the aptamers for their ability to bind S7 mutants were kept at a constant 0.5µM each, a value approximately equal to the Kd’s previously determined by filter binding assay as described above. The aptamers were assessed for their ability to bind two different deletion mutations of
S7: (1) deletion of the N-terminal 17 amino acids and (2) replacement of the \( \beta \)-hairpin (amino acids 73-89) with two glycine residues. Despite deletion of the highly conserved \( \beta \)-hairpin structure, all aptamers bound the S7 mutant with a 0.6-0.8 relative efficiency (Figure 2). By contrast, deletion of the N-terminal 17 amino acids nearly abolishes aptamer binding, indicating that the N-terminus is particularly important for aptamer binding as it is for 16S rRNA and str mRNA binding.

All aptamers exhibited the same trend in their ability to bind various S7 derivatives containing distinct point mutations (Figure 2). With the exception of D14, all residues targeted for mutation are well exposed on the surface of the protein. Of the three aptamers tested, A06-TR3 appears to use the largest number of S7 amino acids in recognizing S7, a characteristic perhaps related to its relatively longer sequence (80nts, versus 65nts for C04-TR2 and 61nts for D02-TR1). Some point mutations had a large impact on aptamer binding while others only caused a 20-30% reduction in binding. For example, mutation K113A experienced the single largest decrease in binding, but methionine 115, which can be cross-linked to the 16S rRNA (19,20), maintained 20-30% binding indicating that this amino acid may be less important for aptamer binding than for 16S rRNA binding. The next largest effect was observed in the double mutant K34A, K35A that targets two conserved lysine residues within the \( \alpha 1-\alpha 2 \) loop and results in a 70-80% decrease in aptamer binding. Additionally, the D14A mutation within the N-terminus also reduced aptamer binding to 20-30%. With the exception of the K113A mutant, it appears that the aptamers
use the same amino acid residues and structural elements as 16S rRNA and *str* mRNA to bind S7.

**The stoichiometry of S7/aptamer binding**

Although apparent cooperative aptamer binding can be nearly eliminated by changing the identity of at least one of the aptamer helices (18), the contribution of S7 in cooperative binding has not been established. Our previous results showed that aptamers A06-TR3, C04-TR2, and D02-TR1 all have Hill coefficients of ~3, indicating that more than three aptamers bind per S7 molecule. Presumably the cooperative binding is facilitated by the multimerization of S7 above concentrations of ~0.2 µM. We attempted to observe this multimeric form of S7 by using size exclusion chromatography.

Gel filtration of S7 resulted in an apparent molecular weight of 18.5 kD, similar to the actual molecular weight of 19.8 kD. Aptamer A06, with an actual molecular weight of 32.4 kD, has an apparent molecular weight of 56.4 kD by gel filtration. The inherently dynamic nature of the RNA probably accounts for this discrepancy. The A06/S7 complex has an apparent molecular weight of 44.2 kD, indicating an RNA:S7 stoichiometry of 1:1. Contradictory to the information obtained from the Hill plots from the S7 titration into the aptamers, these results demonstrate that S7 exists in a monomeric form, both alone and in the aptamer/protein complex.
The role of the N-terminus and β-sheets in S7 multimerization

While gel filtration results indicate that the aptamer/S7 complex has a stoichiometry of 1:1, filter binding assays have shown that the aptamers bind S7 in a highly cooperative manner. These results led us to speculate that at low concentrations of S7, the protein adopts a conformation that the aptamers are unable to bind. Alternatively, at higher concentrations, S7 weakly interacts with itself allowing for exposure of the primary aptamer binding site. We speculated that the flexible N-terminus and β-hairpin may be able to interact with each other intra- and inter-molecularly to sequester or expose the aptamer binding site, respectively.

To test this hypothesis, a series of filter binding assays was performed in which the relevant S7 derivatives were titrated into aptamer A06-TR3, C04-TR2, and D02-TR1 RNA. It can be seen that very little binding occurs between the aptamer and a 19 amino acid fragment corresponding to the N-terminus of S7, which is crucial for aptamer binding to native S7 (Figure 4). The binding curve for this fragment is hyperbolic and saturable, but only ~3% of the total aptamer RNA is able to bind. This low amount of total aptamer binding probably results from only a few of the fragments having amino acids oriented in the proper conformation for binding. The binding curves for the Δβ-hairpin derivatives are also much more hyperbolic than those for the native S7. While there is a 3-6 fold increase in the Kds for this mutant, the maximum fraction of aptamer bound is at least doubled (Table 2). It seems possible therefore that deletion of the β-hairpin allows easier access to the aptamer binding site while at the same time removing residues that contribute to a
higher stability of the S7/aptamer complex. Furthermore, the Hill coefficients for the
Δβ-hairpin derivatives were reduced almost to one, indicating a near elimination of
apparent cooperative aptamer binding.

If our hypothesis is correct, and the flexible N-terminus and β-hairpin of S7
interact with one another at low concentrations to hide the aptamer binding site, then
pre-treatment of native S7 with the N-terminal 19 amino acids should severely
reduce the degree of cooperativity. When a 20-fold molar excess of the N-terminal
fragment was introduced into the S7/aptamer binding reaction, the Hill coefficients
for all aptamers were reduced to ~2 (Figure 4, Table 2). These results indicate that
an excess of N-terminus is able to disrupt the structure of S7 that leads to apparent
cooperative aptamer binding but restores much of the affinity that is lost when the β-
hairpin is deleted.

Discussion

We previously selected aptamers against ribosomal protein S7 that exhibit a
high degree of cooperativity in binding. At concentrations below ~0.2 µM, the
aptamers are incapable of binding to S7 and binding curves exhibit typical sigmoidal
Hill plots. Filter binding assays determined that aptamers A06-TR3, C04-TR2, and
D02-TR2 have Kds in the 0.5 µM range with Hill coefficients of ~3-5. These results
provided evidence for a heretofore uncharacterized quaternary structure of S7. Here,
we provide evidence that at low concentrations, S7 exists in a conformation that
sequesters the aptamer binding site. At higher concentrations of S7 however, we
propose that multimerization of the protein occurs through interaction of the β-sheets
and N-terminal regions of neighboring molecules and that this quaternary structure exposes the electropositive concavity of S7 to allow RNA binding.

The aptamers may interact with the S7 protein by two distinct types of interaction. First is a specific recognition of the electropositive concavity of the protein, which is of high affinity and maintained under conditions such as gel filtration, where shear forces might destroy a weak interaction. Second, there might be a lower affinity (and perhaps less specific) interaction that involves an aptamer binding across a multimer that stabilizes the aptamer-protein interaction. The two interactions may involve entirely different sets of amino acids and structural elements on the protein. All aptamers were predicted to fold into a three-helix junction, and it is conceivable that each helix binds to one or more molecules of S7 giving rise to a stoichiometry of more than three S7 molecules per one aptamer. If this is the case, then we might expect aptamer binding to be affected by a different set of mutations in S7 than the 16S rRNA or str mRNA. The ability of similar S7 mutants to bind in vitro to the minimal S7 binding site of 16S rRNA (11) and the str mRNA (12) has been previously determined. A comparison of our results with this information reveals that the aptamers probably do use the same amino acids for binding, and thus probably use the same electropositive concavity for binding to S7 (Table 1, Figure 5).

One clear exception to the conservation of amino acids important for rRNA, mRNA, and aptamer binding is K113. While a K113Q mutation only causes about a 50% reduction in the affinities of S7 for 16S rRNA and str mRNA (12), the K113A
mutation made in this study abolishes aptamer binding. It is possible that our alanine substitution perturbed the global structure of S7 in a way that the glutamine substitution did not and resulted in loss of aptamer binding. The K113A substitution has already been shown to bind str mRNA \textit{in vivo} as efficiently as native S7 (9), so this explanation is not favored. Another explanation for the observed difference in binding to the K113 mutant is that the aptamer makes a highly stabilizing requisite contact with K113 either before or after binding to the electropositive concavity. K113 is proximal to the concavity of S7 and the adjacent M115 residue has been crosslinked to 16S rRNA (19,20). The crystal structure of S7, both free (from \textit{B. stearothermophilus} (4)) and within the \textit{E. coli} ribosome (21), show that the functional group of K113 is pointed in the opposite direction of the concavity, towards the solvent. This orientation of K113 may indicate that it is involved in the S7 intermolecular interaction we are proposing. If a mutation in K113 abolishes protein-protein interaction, then S7 may be trapped in a monomeric state at higher concentrations than what is observed for native S7. This mutation would therefore have little effect on the affinities of 16S rRNA and str mRNA, both of which bind monomeric S7, but would eliminate aptamer binding, as observed (Figure 2).

Of particular interest is the N-terminus of S7 which is required for binding of the aptamer RNAs, 16S rRNA, and str mRNA (Figure 2, Table 1). \textit{In vivo} work has confirmed that S7 lacking the N-terminal 17 amino acids assembles into 70S ribosomes with only a 1.8-3.1\% efficiency and appears unable to bind str mRNA (9). Here we tested the ability of the aptamers A06-TR1, C04-TR2, and D02-TR1 to
associate with a fragment corresponding to the N-terminal 19 amino acids of S7, but found that only ~3% of the total RNA was capable of binding. The N-terminus of S7 lacks secondary structure, and the first eight amino acids of S7 from *Bacillus stearothermophi us* (22) and the first eleven amino acids of the *Thermus thermophilis* (3) crystal structures are conspicuously absent although they are visible within the crystal structure of the 30S subunit (7). Presumably upon binding to RNA, the N-terminal extension adopts a more static conformation and helps to stabilize the protein/RNA interaction (10). Furthermore, a point mutation within the N-terminal extension, D14A, also severely reduced aptamer binding. D14 was the only point mutation made whose functional group is not exposed on the surface of S7. As noted by Fredrick et al. (9), D14 makes hydrogen bonds to G18, S19, and Y43 and thereby packs the N-terminal extension against the adjacent α-helices. By breaking the tether between these two structural features, the position of the N-terminal extension in the D14A mutant may become entirely arbitrary and functionally equivalent to the N-terminal deletion.

Unlike the N-terminal deletion mutant, deletion of the conserved β-hairpin of S7 assembles into ribosomes with 45% efficiency *in vivo* and appears to associate with *str* mRNA equally as well as unmutated S7 (9). Here too we have found that deletion of the β-hairpin does not severely reduce binding of the aptamers (Figure 2). The β-hairpin thus seems to be generally dispensable for RNA binding, but is important for other tasks such as stabilizing the E-site tRNA and preventing +1 and -1 frameshifting during translation (23).
The crystal structures obtained for S7 have revealed a concavity formed by the helical hydrophobic core domain and the β-hairpin. This concavity is lined with basic residues to which the 16S rRNA binds. We propose that the S7 aptamers also bind to this site, but that at low protein concentrations the flexible N-terminus and β-hairpin of S7 interact to block access to it. At higher concentrations of S7, the binding site is exposed because the β-hairpin of one S7 can interact with the N-terminus (and possibly K113, as discussed above) of an adjacent S7. During our selection of the aptamers, it is reasonable to assume that RNAs were selected for their ability to bind this exposed concavity because we used relatively high concentrations of S7 at about 0.7 µM or above. Here we have reduced the Hill coefficients to ~1.2 by eliminating the β-hairpin from S7 (Table 2), nearly abolishing cooperativity. This result is consistent with the idea that the β-hairpin is at least partially responsible for blocking access to the aptamer binding site. Apparent cooperative binding could also be reduced by pre-treating S7 with the N-terminal 19 amino acids, supporting the idea that these two domains interact.

Although we have detected a quaternary structure for S7, this interaction must necessarily be weak. Neither binding of 16S rRNA nor str mRNA to S7 has been shown to demonstrate cooperativity despite titration of the protein over the same range used here (11,12,24,25). Moreover, our own gel filtration results indicate a 1:1 stoichiometry for the aptamer/S7 complex, indicating that shear force of the FPLC is enough to disrupt the multimeric S7 complex to which the aptamers initially bind (Figure 3). The aptamers, having bound to the common RNA binding site,
remain associated with a single monomer of S7 following disruption of the weak multimeric complex. Thus, formation of the multimeric S7 complex is only necessary for initial aptamer binding and is completely dispensable for maintenance of the S7/aptamer complex. The 16S rRNA (~1500nts) and str mRNA (~100nts) sequences are both larger than the S7 aptamers (61-80nts) and may be capable of destabilizing the inter- and intra-molecular interaction of the N-terminal and β-hairpin domains through making alternative secondary contacts with the S7 protein. The shorter aptamer sequences however, require prior exposure of the concave binding site, which can be accomplished through intermolecular S7 interactions at concentrations above ~0.2 µM. Once the aptamers have bound to the concavity of S7 for which they have a high affinity, the S7/aptamer complex can assume a 1:1 stoichiometry with absolutely no affect on the stability of the complex.

A multimeric structure of S7 probably makes the protein more susceptible to aggregation due to increased exposure of the hydrophobic residues lining the concave surface. We, and other groups (26,27), have observed in vitro aggregation of S7 which appears to be temperature and concentration dependent (results not shown). R-proteins are some of the most abundant proteins within the E. coli cell, and among these, S7’s concentration seems to be one of the highest (28). It is interesting to note that S7 (25-60 µM) has a particularly low melting temperature of only 43°C and that thermal renaturation in vitro does not occur (26). With such a low melting temperature, a fraction of unfolded S7 is expected even at 37°C.
The low melting temperature of S7 and its propensity to aggregate certainly has some interesting in vivo implications. In vivo aggregation of proteins is known to be a common phenomenon in prokaryotic cells that are grown at elevated temperatures. It has been generally accepted that no significant free pool of ribosomal proteins exists within the prokaryotic cytosol (29-32), and ribosomal proteins have not been identified within thermally induced aggregates extracted from bacterial cells (33,34). However, a more recent study by Maisonneuve et al. (35) has shown that protein aggregation occurs in healthy E. coli cells and that these aggregates do in fact contain degradation products of ribosomal proteins. It is interesting to note that of the thirteen 30S subunit r-proteins detected in the aggregates, six belong to the S7 assembly branch: S7 itself; the secondary binding protein S9; and the tertiary binding proteins S2, S3, S10, and S13. Furthermore, r-proteins S2, S3, and S10 all require thermal activation of the reconstitution intermediate in vitro prior to being able to bind the assembling 30S subunit (36,37). The large number of r-proteins from the S7 assembly branch within bacterial aggregates and the in vitro requirement for thermal activation of the reconstitution intermediate for 30S head formation may be indicative of some general assembly inefficiency within the S7 assembly branch.

The thermal instability of S7 may play a role in ribosome biogenesis, necessitating temperature-dependent in vitro assembly of the 30S subunit and chaperone-dependent in vivo assembly. 16S binding to the concavity of S7 is not expected to be impeded by the monomeric structure, as is the case for the S7
aptamers, because 16S makes a number of alternative, probably weak contacts with the protein to stabilize an open conformation of the binding site. Initial binding of S7 to 16S does not, however, preclude subsequent interaction between the flexible N-terminus and $\beta$-hairpin. The thermal activation of the 30S reconstitution intermediate may serve two purposes: (1) promoting the proper folding of the 16S rRNA (38), and (2) promoting a conformational shift in S7 that destabilizes the N-terminal and $\beta$-hairpin interaction that allows S7 to make additional stabilizing contacts with 16S.

Examination of the thermal dependence of 16S rRNA conformational rearrangements by RNA footprinting in the presence of single r-proteins has revealed that the primary binding proteins S4 and S7 yield the largest number of temperature-dependent conformational rearrangements (39,40). Dutca et al. (40) have suggested that binding of these two proteins to 16S rRNA occurs in two sequential phases in which one domain of the protein can bind 16S rRNA at 0 °C and the second domain requires a temperature shift to 42 °C. S4 and S7 are the largest of the primary r-proteins, and it was proposed that their multi-domain structures allow for the biphasic, temperature-dependent rearrangements. It is also interesting to note however, that S4 and S7 probably have the lowest melting temperatures of all the primary r-proteins, with 45 °C and 43 °C, respectively. The results presented here suggest that temperature-dependent rearrangements in 16S rRNA may also be facilitated by a structural rearrangement of S7 at 42 °C. Thus, both protein and RNA conformational changes may contribute to the temperature-dependent structural rearrangements of 16S rRNA and the widely variable rate
constants for the interaction of S7, and possibly of S4, with different 16S residues (41).

The thermal instability of S7 presents a challenge to ribosome biogenesis. Prokaryotic chaperones, which are up-regulated during stress conditions such as elevated temperatures (42-44), are involved in promoting proper protein conformation (45-47) as well as protein refolding (48-50) and disaggregation (51-53). More recently, chaperones have been implicated in the process of ribosome biogenesis. For example, GroEL, which promotes the folding of hydrophobic surfaces (54), is required for the late 45S to 50S transition of the large ribosomal subunit (55). Temperature-dependent ribosome biogenesis in E. coli cells lacking the chaperones DnaK or DnaJ has recently been demonstrated (56), and the ability of DnaK to promote the in vitro assembly of functional 30S subunits at otherwise non-permissive temperatures has also been reported (57,58). High β-sheet content and exposed hydrophobic residues are two major contributing factors leading to protein aggregation (59,60). An intra-molecular interaction between the N-terminus and β-hairpin that sequesters the hydrophobic residues within the concavity of S7 may help to promote S7’s solubilization prior to association with 16S rRNA. Furthermore, at temperatures above 37 °C, where heat-shock induced chaperones are essential for cell division, it is conceivable that S7 partially denatures after binding to 16S or is partially denatured prior to binding 16S. While a DnaK affinity column did not identify S7 as being a binding partner, r-proteins S2, S3, and S19 within the 30S head were found to bind DnaK with medium to high affinity (57). Under both normal growing
and stress conditions, DnaK may accelerate or slow the association of the 30S and 50S subunits based on an indirect assessment of the integrity of S7. Without DnaK, a major bottleneck in 30S maturation occurs wherein all tertiary binding proteins of the S7 pathway are missing (61). Furthermore, the \textit{in vitro} requirement for thermal activation can be bypassed by the inclusion of DnaK in the reconstitution mixture (57,58), perhaps by indirectly forcing a conformational shift in S7. If the N-terminal extension and $\beta$-hairpin of S7 undergo a temperature-dependent shift during \textit{in vitro} reconstitution, then it follows that the thermal activation of 30S subunits assembled with the S7 $\Delta \beta$-hairpin mutant may not exhibit the same degree of thermal dependence as those assembled with native S7.

We propose that S7 multimerization may also play a role in the translational inhibition of the \textit{str} operon. The $\beta$-hairpin is dispensable for S7 binding to \textit{str} mRNA, and may be completely available for association with adjacent molecules of S7. A deficiency in 16S rRNA allows free S7 to bind to the S12-S7 intercistronic region of \textit{str} mRNA, reducing translation of S12, S7, and EF-G. It has been proposed that upon S7 binding, a structure of the intercistronic mRNA is stabilized that sequesters the stop codon of \textit{rpsL} and the start codon of \textit{rpsG} (1,62,63). The translational repression of the upstream gene encoding S12 is thought to be the result of ribosome stalling due to the sequestered stop codon. However, ribosome stalling may be reinforced by the polymerization of the S7 protein bound at the intercistronic region and low affinity binding of this polymer to the upstream RNA by way of its basic domain. These results may help explain why Saito \textit{et al.} (1) observed a
progressive de-repression of S12 expression over time, but not in the expression of EF-G and EF-Tu, when S7 was over expressed *in trans*. S7 may bind to the mRNA in a pre-multimerized form, resulting in ribosome stalling due to the multimeric obstacle. If the S7/S7 interaction is weak it may be easily disrupted by repeated ribosome bombardment, causing eventual decay of the multimer and relieving some of the translational repression. Translational repression of the genes encoding S7 and EF-G would remain constant over time because the monomeric S7 remains bound with high affinity to the intercistronic region and prevents translational coupling with the upstream gene encoding S12. Alternatively, monomeric or multimeric S7 may bind to the *str* operon with continued protein multimerization. Because the intermolecular interactions of S7 are weak, the multimer may dissociate from the mRNA and the single S7 molecule bound at the intercistronic site. The multimeric S7 will be more prone to aggregation because of the increased exposure of hydrophobic residues in combination with the low melting temperature of the protein. Repeated dissociation and aggregation will eventually lead to a much smaller pool of free S7, primarily in the monomeric form. Monomeric S7 would provide much less of an obstacle for ribosomes translating S12 and would probably result in global derepression of S12 expression over time. This latter explanation would allow for deregulation that does not depend on the half-life of individual molecules of *str* mRNA which is probably no more than four minutes (64).
References

<table>
<thead>
<tr>
<th>Mutation in S7</th>
<th>Aplamer A06-TR3</th>
<th>\textit{in vivo} assembly (30S)\textsuperscript{a}</th>
<th>\textit{in vitro} affinity (16S)\textsuperscript{b}</th>
<th>\textit{in vitro} affinity (str)\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>ΔA1-17</td>
<td>0.04</td>
<td>0.03</td>
<td>n.d.</td>
<td>0.09</td>
</tr>
<tr>
<td>D14A</td>
<td>0.18</td>
<td>0.09</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>K34A, K35A</td>
<td>0.08</td>
<td>0.01</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>G54S</td>
<td>0.35</td>
<td>n.d.</td>
<td>0.40</td>
<td>0.49</td>
</tr>
<tr>
<td>Δβ-heapin</td>
<td>0.63</td>
<td>0.45</td>
<td>0.34</td>
<td>0.28</td>
</tr>
<tr>
<td>R77A, R78A</td>
<td>0.48</td>
<td>0.33</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>K113A or K113Q</td>
<td>0.03</td>
<td>n.d.</td>
<td>0.42</td>
<td>0.59</td>
</tr>
<tr>
<td>M115G</td>
<td>0.25</td>
<td>n.d.</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>K133A</td>
<td>0.35</td>
<td>n.d.</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td>K135A, D139A, R142</td>
<td>0.44</td>
<td>1.6</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

\textsuperscript{a}The relative \textit{in vivo} assembly efficiency into 30S subunits is taken from Fredrick \textit{et al.} (9).

\textsuperscript{b}The relative \textit{in vitro} affinity for the lower half of the 3' major domain of 16S rRNA is taken from Robert \textit{et al.} (11).

\textsuperscript{c}The relative \textit{in vitro} affinity for the intercistronic region of str mRNA is taken from Robert \textit{et al.} (12).
<table>
<thead>
<tr>
<th>Aptamer</th>
<th>S7 derivative</th>
<th>Maximum fraction bound</th>
<th>Kd (µM)</th>
<th>Hill coefficient</th>
<th>Relative Affinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A06-TR3</td>
<td>native</td>
<td>0.34 ± 0.00</td>
<td>0.69 ± 0.02</td>
<td>3.13 ± 0.16</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Δβ-hairpin</td>
<td>1.17 ± 0.19</td>
<td>4.65 ± 1.03</td>
<td>1.41 ± 0.11</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>N-terminal 19 aa</td>
<td></td>
<td>no detected binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S7 pre-treated with N-terminal 19 aa</td>
<td>0.82 ± 0.02</td>
<td>1.03 ± 0.05</td>
<td>2.26 ± 0.17</td>
<td>0.67</td>
</tr>
<tr>
<td>C04-TR2</td>
<td>native</td>
<td>0.25 ± 0.00</td>
<td>0.69 ± 0.01</td>
<td>2.80 ± 0.06</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Δβ-hairpin</td>
<td>0.60 ± 0.11</td>
<td>2.80 ± 0.66</td>
<td>1.30 ± 0.17</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>N-terminal 19 aa</td>
<td></td>
<td>no detected binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S7 pre-treated with N-terminal 19 aa</td>
<td>0.72 ± 0.02</td>
<td>0.90 ± 0.06</td>
<td>1.90 ± 0.18</td>
<td>0.77</td>
</tr>
<tr>
<td>D02-TR1</td>
<td>native</td>
<td>0.38 ± 0.01</td>
<td>0.63 ± 0.02</td>
<td>3.20 ± 0.07</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Δβ-hairpin</td>
<td>0.64 ± 0.02</td>
<td>2.00 ± 0.12</td>
<td>1.20 ± 0.01</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>N-terminal 19 aa</td>
<td></td>
<td>no detected binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S7 pre-treated with N-terminal 19 aa</td>
<td>0.78 ± 0.02</td>
<td>0.87 ± 0.04</td>
<td>2.20 ± 0.18</td>
<td>0.72</td>
</tr>
</tbody>
</table>

*Titrations were done in the presence of 0.02% BSA and filtered using the BioDot filtration unit.*
Figure 1. Crystal structure of r-protein S7 from *Bacillus stearothermophilus* adapted from Hosaka *et al*. Secondary structural features are identified and residues mutated in this study are colored magenta. The β-hairpin is formed by the β1 and β2 sheets which are labeled above. *E. coli* K12 amino acids 1-9 and 148-178 are not included in this model.
Figure 2. Ability of S7 aptamers A06-TR3, C04-TR2, and D02-TR to bind mutant S7 derivatives. Nitrocellulose filter binding assays were performed with concentrations of RNA and protein at the ~Kd for the aptamer/S7 complex. Experiments were performed in duplicates or triplicates and were repeated at least one time. Error bars represent standard deviation. The results from a single representative experiment are shown.
**Figure 3.** Representative results from size exclusion chromatography of S7, aptamer A06, and the S7/A06 complex. All samples were filtered on a Superdex 75 column attached to an AKTA-FPLC and equilibrated with binding buffer. The results shown were obtained at 4°C, and are the same as those obtained at 25°C.
Figure 4. Binding curves for selected mutant S7 derivatives and the aptamers A06-TR3 (a), C04-TR2 (b), and D02-TR1 (c). Maximum fraction bound, affinities, and Hill coefficients can be found in Table 2.
Figure 5. Ability of 16S rRNA (i.e., the lower half of the 3’ major domain), str mRNA, and aptamer RNA to bind various S7 mutants. The relative affinities for the 16S rRNA and str mRNA are from Robert et al. (2000) and Robert et al. (2001), respectively. The relative binding of the aptamers is based on the information in Figure 2. The primary sequence and secondary structure from *E. coli* K12 is shown, and coloring is the same as in Figure 1. Deletion mutants are indicated by a double-headed arrow. The results for double- and triple-point mutations are indicated by a shared symbol.
CHAPTER 4: GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

General conclusions

Here we have reported on the selection of RNA aptamers that bind to prokaryotic ribosomal protein S7. Isothermal titration calorimetry (ITC) has been employed to obtain affinities for the aptamer to both rpS7 and rpS4. The highest affinity aptamers have low nanomolar Kd’s, and the truncated aptamer A06-TR3 was determined to have a Kd of 0.71 nM. None of the aptamers exhibited specific binding to rpS4, and the ΔH for these titrations was too low to obtain reliable thermodynamic data. Our most surprising finding was the cooperative nature of aptamer binding, detected both by ITC and filter binding assays. This cooperativity is indicative of a quaternary structure for S7 that has not been previously characterized. Using aptamer binding as an indicator of the multimeric S7 complex, we have determined that at low concentrations, S7 is a monomer whose primary RNA binding site is blocked by the interaction of the flexible N-terminal extension and anti-parallel β-sheets. The longer 16S rRNA and str mRNA sequences are unable to detect this monomeric form of S7 probably because they are able to make alternative contacts with the protein that stabilize an open conformation for the RNA binding site.

Aptamer development can be a risky business. While in theory it should be possible to obtain an aptamer to nearly any target molecule, in practice SELEX is more of an art rather than a science. A number of variables can contribute to the successful selection of an aptamer, including selection temperature, the inclusion of mono- and divalent cations, the ratio of nucleic acid to target molecule, incubation
time, and the length of the randomized pool among other things. Perhaps the most important factors affecting the successful (and timely) selection of an aptamer however, are the biochemical and biophysical properties of the target molecule itself. The selection of aptamers that bound rpS7 can hardly be considered a coincidence; rpS7 has a large electropositive surface known to bind both rRNA and mRNA. Certainly the development of a panel of aptamers capable of binding to all of the ribosomal proteins is possible and could potentially be done simultaneously (at least initially) using a mixture of the total ribosomal proteins.

It is important to consider the potential knowledge gained and downstream applications of developing an aptamer to any target molecule. Aptamers can be tools with which the researcher is able to test hypotheses concerning nucleic acid-protein interactions either specifically or in general. Subsequent to aptamer selection, it is first desirable to characterize the aptamer/target molecule complex by determining both the affinity and specificity. With this information in hand, the researcher may conduct a more extensive characterization or develop the aptamer technology for a particular application. In the preceding work, our long-term goal was to select a novel class of aptamer bacterial antibiotics that would be capable of interfering with 30S ribosomal subunit assembly. Initial characterization of the aptamer/S7 complexes revealed cooperative binding of the aptamers to S7. This result was both peculiar and interesting and we decided that further study of the S7 aptamers was warranted. While our initial research was goal driven, the cooperative aptamer binding prompted a more hypothesis driven approach to our studies. Cooperative
aptamer binding has facilitated the study of S7/S7 and RNA/S7 interactions with potential \textit{in vivo} significance and has allowed us to provide explanations for anomalies present within the S7 literature.

\textbf{Future Directions}

\textbf{Development of aptamer antibiotics}

The long-term goal of the ribosomal protein aptamer project is the inhibition of 30S or 50S subunit assembly. For S7, in theory, one need only select an aptamer capable of binding the protein with a higher affinity than that of 16S rRNA or the \textit{str} mRNA. It should be noted however that a number of factors can affect the success of the \textit{in vivo} implementation of such aptamer antibiotics derived through SELEX. \textit{In vivo} conditions are significantly more complex than those used during the aptamer selection process and it is difficult, if not impossible, to know if the selected aptamer will have \textit{in vivo} efficacy. Of great consideration is the ability of the aptamers to fold into a structure that S7 can recognize. To test this, an aptamer expression plasmid has been designed that uses \(\beta\)-galactosidase as a reporter of aptamer/S7 interaction. In this plasmid, an aptamer or the \textit{str} mRNA intercistronic region can be cloned upstream of the \textit{lacZ} gene at one of three possible cloning sites. Induction of aptamer/\textit{lacZ} fusion expression in either the presence or absence of S7 overexpression will be done. The ability of the aptamer to inhibit translation of the downstream \textit{lacZ} gene, either alone or bound to S7, can then be assayed by testing for \(\beta\)-galactosidase activity. Prior to \textit{in vivo} experiments, this system can be tested \textit{in vitro} because transcription of the aptamer/\textit{lacZ} fusion is under the control of the T7
promoter. This system, in conjunction with UV crosslinking, may provide a general screening method for testing any aptamer’s ability to fold and bind its target molecule within a bacterial cell.

It is also possible to test the S7 aptamers’ ability to inhibit 30S ribosomal subunit assembly using an \textit{in vitro} reconstitution assay. This experiment has the benefit of being conducted under the same buffer conditions as those used during aptamer selection. Preliminary reconstitution assays in the presence of aptamer A06-TR3 have been conducted by Nathan Napper in Gloria Culver’s laboratory. The results indicate that the aptamer is capable of interfering with 30S assembly, although significantly higher aptamer concentration than that of 16S rRNA or TP30 is necessary (Figure 1). Although we detect low nanomolar dissociation constants for many of the aptamers using ITC, these Kd’s probably represent the affinity of the aptamers for S7 in which the RNA binding site is exposed. Furthermore, cooperative aptamer binding is probably not the best binding model for interfering with S7’s association with 16S rRNA. The affinity of S7 for its 16S rRNA binding site is extremely high and the rRNA will easily out-compete the cooperatively binding aptamers at low concentrations of S7. Moreover, higher intracellular concentrations of free S7 are expected only after rRNA synthesis has shut down and when aptamer binding can no longer affect 30S subunit assembly.

To inhibit 30S subunit assembly, it is desirable to obtain an S7 aptamer that does not bind cooperatively. While we have taken steps towards the development of such an aptamer, the affinities of the aptamers will need to be improved. Another
important factor for successful *in vivo* inhibition of 30S assembly is the off-rate of the aptamer. Results from the A06-TR3 aptamer indicate a half-life of ~9 minutes for the aptamer/S7 complex (Figure 2). While this is a reasonable half-life, it is probably not good enough to sustain inhibition of 30S subunit assembly. The 30S reconstitution assays show an increase in assembly the longer TP30 pre-treated with A06-TR3 is allowed to incubate with 16S rRNA (data now shown). These results provide preliminary evidence that the aptamer dissociates from S7 and is replace by 16S rRNA and that the 16S/S7 complex is more stable than the aptamer/S7 complex.

Cooperative aptamer binding could potentially be eliminated by creating a much larger aptamer. The additional helix length may be able to make alternative contacts with the monomeric form of S7 and stabilize access to the primary RNA-binding site in much the same way that has been observed for the 16S rRNA binding site. As an alternative approach, an aptamer antibiotic may not necessarily need to directly interfere with 30S subunit assembly. If a temporary reduction in 16S rRNA production could be created, then the S7 aptamers may have an opportunity to interfere with repression of the *str* operon. This would lead to overexpression of S7, which is also known to interfere with bacterial growth.

*In vivo* significance of S7 multimerization

The cooperative aptamer binding described above provides strong evidence for a multimeric form of S7. The interactions that contribute to S7 multimer formation are apparently weak and limit the observation of this complex to methods which will not disrupt these interactions. Our studies show that at concentrations above
~250nM, S7 undergoes intermolecular interactions that expose its primary RNA binding site. Many in vitro techniques for studying protein structure (e.g. dynamic light scattering, X-ray crystallography, NMR, small angle X-ray scattering, etc.) require relatively high protein concentrations. However, the propensity of S7 to aggregate even at ~24 °C is expected to interfere with its analysis in vitro using these techniques.

Despite the possible limitations for studying monomeric and multimeric S7 in vitro, the potential role of the multimeric complex in translational repression of the str operon, and particularly the retroregulation of S12 expression, should be further explored. During S7 overexpression, we speculate that gradual translational derepression of S12 expression is a result of the depletion of the multimeric S7 complex either by repeated ribosome bombardment or by eventual precipitation and dissociation of the S7 aggregate from the str mRNA. This hypothesis will be tested both in vitro and in vivo by cloning rpsL, the intercistronic region, and rpsG downstream of the T7 promoter. In vitro analysis of the retroregulation of S12 expression will first be explored by using the plasmid’s corresponding transcript and an in vitro translation system derived from E. coli. We have hypothesized that a multimeric form of S7 is capable of binding to the intercistronic region of str mRNA. Following transcription of the partial str operon, translation in the presence or absence of an excess of S7 will be done. At various time points, chloramphenicol will be added to the in vitro or in vivo reactions in order to “freeze” the ribosomes on the mRNA followed by UV crosslinking. If necessary, isolation of the 70S/mRNA
complex will be achieved by sucrose gradient sedimentation. Primer extension will then be used to determine the position of the ribosomes on the mRNA. If our hypothesis concerning the role of the multimeric S7 complex during retroregulation of S12 is correct, we should see a time-dependent progression of ribosomes on the mRNA in the direction of the intercistronic region. After 30 minutes, we expect to see complete progression of the ribosome through the *rpsL* gene, in accordance with the results of Saito and Nomura. These results could reveal a previously unrecognized mechanism of translational repression and may hint at yet more potential targets for antibiotic development.
Figure 1. 30S ribosomal subunit assembly in the presence of S7 aptamer A06-TR3. (a) The ability of tRNA to bind 30S subunits that were reconstituted following the pre-treatment of TP30 with aptamer A06-TR3. Reconstitution assays contained 0.4 µM 16S rRNA, 0.8 µM total proteins purified from 30S subunits (TP30), and the specified concentration of aptamer A06-TR3 in 80 mM HEPES (pH 7.6), 330 mM KCl, 20 mM MgCl₂, and 0.01% Nikkol. 16S rRNA and TP30 were pre-incubated separately at 42 °C and room temperature, respectively, for 30 minutes and then placed on ice for five minutes. Aptamer RNA was added to TP30 and incubated at 42 °C for 30 minutes. Finally, 16S rRNA was added to the aptamer/TP30 mixture and incubated for 30 minutes at 42 °C, followed by placement on ice. To assess assembly efficiency, poly(U) mRNA and radiolabeled tRNA^{Phe} were mixed with the reconstituted 30S subunits and incubated at 15 °C for 30 minutes. tRNA^{Phe} binding was determined using a nitrocellulose filter binding assay and a scintillation counter. (b) The ability of tRNA to bind 30S subunits that were reconstituted in the presence of aptamer A06-TR3. This experiment was performed as in (a), except the TP30 was not pre-treated with aptamer prior to mixing with 16S rRNA. Rather, following separate incubations of RNA and protein, the TP30, 16S rRNA, and aptamer A06-TR3 RNA were added simultaneously and reconstitution was allowed to occur for 30 minutes at 42 °C. 30S reconstitution assays were done by Nathan Napper in Gloria Culver’s laboratory.
Figure 2. Dissociation of the A06-TR3/S7 complex over time. Radiolabeled aptamer A06-TR3 was competed off of S7 by adding an excess of unlabeled A06-TR3. An aliquot of the reaction was removed at each time point and filtered on nitrocellulose. Bound cpms were determined by Cerenkov.