High-Affinity RNA Ligands to *Escherichia coli* Ribosomes and Ribosomal Protein S1: Comparison of Natural and Unnatural Binding Sites†

Steven Ringquist,‡ Thomas Jones,‡ Eric E. Snyder,‡ Terri Gibson,‡ Irina Boni,§ and Larry Gold*,‡

Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309, and Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, U. I. Miklukho-Maklaya, 16/10, 117871 GSP Moscow, Russia

Received September 6, 1994; Revised Manuscript Received December 29, 1994⊗

ABSTRACT: High-affinity RNA ligands were generated against intact 30S ribosomes, S1-depleted 30S ribosomes, and purified ribosomal protein S1. Sequence analysis indicated two classes of ligand: unstructured RNAs containing a Shine−Dalgarno sequence and structured RNAs containing a pseudoknot. The Shine−Dalgarno-containing ligands were generated against S1-depleted 30S ribosomes but, surprisingly, not against intact 30S ribosomes or ribosomal protein S1. In contrast, pseudoknot-containing ligands were generated against intact ribosomes as well as purified S1 protein. The two classes of ligand exhibited specificity for their respective targets, as well as conserved sequence and secondary structure reminiscent of naturally occurring, cis-acting mRNA elements.

Formation of the translational initiation complex is considered the rate-limiting step in setting the level of translational yield for most messages [reviewed by Gold et al. (1981); Gold, 1988; Gualerzi & Pon, 1990]. Translational yield from open reading frames from within a single polycistronic message can vary over 3 orders of magnitude (Kozak & Nathans, 1972; Ray & Pearson, 1975). The best-studied determinants of initiation rate are found within the ribosome-binding site (RBS)[1] and include the choice of start codon (Munson et al., 1984; Shinedling et al., 1987; Hartz et al., 1991; Ringquist et al., 1992, 1993a), as well as the extent of complementarity between the 3′-end of the 16S rRNA and a region upstream from the start codon, the Shine−Dalgarno (SD) sequence (Shine & Dalgarno, 1974; Steitz & Jakes, 1975; Hui & deBoer, 1987; Jacob et al., 1987; Ringquist et al., 1992). The effect of spacing between the SD and the start codon (Hartz et al., 1991; Ringquist et al., 1992) as well as the degree of secondary structure at the RBS (McPheeters et al., 1986; Looman et al., 1987; Blasi et al., 1989; deSmit & van Duin, 1990a,b, 1994; Ringquist et al., 1992, 1993b, 1994) and autogenous regulation by protein gene products have also been studied extensively (Gold et al., 1984; Winter et al., 1987; McPheeters et al., 1988; Moine et al., 1988; Wulczyn et al., 1989; Tuerk et al., 1990).

Messenger RNA elements other than those within the RBS also influence translational yield (McCarthy & Gualerzi, 1990). Known as translational enhancers, these sequences elevate the rate of formation of the translational initiation complex and are thought to function by mRNA−rRNA pairing (van Knippenberg, 1975; Petersen et al., 1988; Olina & Rangwala, 1989; Thanaraj & Pandit, 1989; Shean & Gottesman, 1992), as well as by specific mRNA−ribosomal protein interactions (Boni et al., 1991; Zhang & Deutscher, 1992). Recently, Loechel et al. (1992) described a novel mRNA element that enhances translational initiation independent of the SD sequence, and Melancon et al. (1990) showed that ribosomes lacking the 3′-end of 16S rRNA correctly select the translational start site in some natural mRNAs, suggesting that mRNA−ribosomal protein interactions may be important determinants of the rate of formation of the translational initiation complex. One likely candidate for a ribosomal protein that acts to enhance translational initiation is S1. Cross-linking studies have aligned the nucleic acid-binding domain of S1 with a region of the mRNA 5′ to the RBS (Boni et al., 1991), suggesting that S1 may interact with mRNA sequences upstream from natural *Escherichia coli* RBSs (Boni et al., 1991; Zhang & Deutscher, 1992).

The present work investigates the role of ribosomal protein S1 during recognition of the translational initiation region (TIR). High-affinity RNA ligands to 30S ribosomes, S1-depleted 30S ribosomes, and purified ribosomal protein S1 were generated by SELEX (systematic evolution of ligands by exponential enrichment) (Tuerk & Gold, 1990). RNA ligands were analyzed for target affinity as well as specificity, and nucleotide residues essential for RNA−target binding were identified. Comparison of RNA ligands to natural *E. coli* and bacteriophage RNAs indicates that ribosomal protein S1 may play an important regulatory role during gene expression.

EXPERIMENTAL PROCEDURES

Materials. AMV reverse transcriptase and T7 RNA polymerase were from U.S. Biochemicals Corporation (Cleveland, OH). Taq polymerase was from Promega Corporation (Madison, WI). T4 polynucleotide kinase, alkaline phosphatase, RNAse T1, DNAse I, and RNA ligase were purchased from New England Biolabs, Inc. (Beverly, MA). $tRNA_f^{Met}$ was purchased from Subriden RNA (Rolling Bay, WA). 30S ribosomes were a gift from V. Makhno

† This work was supported by grants to L.G. from the National Institutes of Health (GM-28685 and GM-19963).

* Corresponding author: telephone, (303) 492-7864; Fax, (303) 492-7744.

‡ University of Colorado.

§ Russian Academy of Sciences.

⊗ Abstract published in *Advance ACS Abstracts,* February 15, 1995.

[1] Abbreviations: RBS, ribosome-binding site; SD, Shine−Dalgarno; SELEX, systematic evolution of ligands by exponential enrichment; TIR, translation initiation region.

0006-2960/95/0434-3640$09.00/0 © 1995 American Chemical Society

Table 1: DNA Oligonucleotides Used in This Study

oligonucleotide[a]	sequence[b]
30N template	5′-GCCGGATCCGGGCCTCATGTCGAA[30N]TTGAGCGTTTATTCTGAGCTCCC-3′
5′-primer	5′-CCGAAGCTTAATACGACTCACTATAGGGAGCTCAGAATAAACGCTCAA-3′
3′-primer	5′-GCCGGATCCGGGCCTCATGTCGAA-3′

[a] DNA oligonucleotides were used to generate the double-stranded DNA template for *in vitro* transcription. The 3′-primer was also used to generate cDNA from the RNA repertoire and was used as a sequencing primer. [b] The symbol 30N indicates a 30-nucleotide randomized region within the template oligonucleotide.

(Petersburg Nuclear Physics Institute, Russia). Other reagents used in this work were of the highest quality obtainable.

Nucleic Acids. Briefly, RNAs were synthesized by *in vitro* transcription using PCR-amplified DNA oligonucleotides and T7 RNA polymerase (Ringquist et al., 1993b). DNA oligonucleotides (Table 1) were synthesized on an Applied Biosystems Model 394 DNA/RNA synthesizer and purified by 6% preparative polyacrylamide gel electrophoresis. PCR amplification was performed in 50 mM KCl, 10 mM Tris-HCl (pH 8.6), 2.5 mM $MgCl_2$, 170 mg/mL BSA, and dNTPs (present at 1 mM each). Taq DNA polymerase was used at 100 units per 0.1 mL reaction, and the 5′- and 3′-primers were present at 1 mM. Transcription was performed in 8% PEG 8000, 5 mM DTT, 40 mM Tris-HCl (pH 8.0), 12 mM $MgCl_2$, 1 mM spermidine, 0.002% Triton X-100, and 2 mM NTP. T7 RNA polymerase was present at 1 unit/μL. The reaction was incubated at 37 °C for 2 h and then treated with 20 units of DNAse I for an additional 5 min at 37 °C. The reaction was stopped by the addition of an equal volume of loading buffer (72% formamide, 120 mM EDTA, pH 8.0) and heated to 95 °C for 3 min prior to electrophoresis on a 6% polyacrylamide/8 M urea denaturing gel. The RNA transcript was visualized by UV shadowing and was excised from the gel and eluted into distilled H_2O. The RNA transcript was ethanol precipitated, dried under vacuum, and redissolved in distilled H_2O. Transcription of ^{32}P-labeled RNA was performed as described earlier, except that 50 μCi of [α-^{32}P]ATP (800 Ci/mmol) was present in the transcription reaction. The concentration of RNA was quantitated by measuring the A_{260} and assuming that 1 A_{260} unit equaled 40 μg/mL.

Purification of Ribosomal Protein S1 and S1-Depleted 30S Ribosomes. Ribosomal protein S1 and S1-depleted 30S ribosomes were prepared from intact 30S ribosomes by affinity chromatography on poly(U)−Sepharose (Pharmacia, Piscataway, NJ) as described by Subramanian (1983). The purified S1 protein gave a single band when analyzed by SDS−PAGE electrophoresis and was active in binding to S1-depleted 30S ribosomes when assayed by toeprinting (Ringquist et al., 1993a). The concentration of purified S1 was determined as described by Lowrey et al. (1951). The concentration of S1-depleted ribosomes was determined by assuming that 1 A_{260} unit equaled 67 pmol/mL; S1-depleted ribosomes lacked S1 when assayed by SDS−PAGE electrophoresis.

Evolution of High-Affinity RNA Ligands. SELEX was performed as described by Tuerk and Gold (1990) using the oligonucleotides illustrated in Table 1. The DNA template contained a 30-nucleotide (30N) variable sequence generated by mixed-nucleotide DNA synthesis, as well as 5′- and 3′-fixed sequences. The fixed sequences were necessary for PCR amplification of the template, and the 5′-fixed sequence contained a T7 RNA polymerase promoter. The first round of SELEX was performed by PCR amplification of the 30N template to generate a double-stranded DNA template repertoire for RNA transcription. The RNA product was split into five pools of equal concentration that were used to generate ligands to the targets. SELEX consisted of 11−13 rounds of RNA synthesis, binding to target, partitioning of bound and unbound RNA by nitrocellulose filtration, cDNA synthesis, and PCR amplification to regenerate the double-stranded DNA template. Binding of the target to RNA pool was performed in 0.1 mL of binding buffer (100 mM NH_4Cl, 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 1 mM DTT) at 37 °C for 5 min prior to filtration. The concentrations of RNA and target were 1000 and 50 nM, respectively.

Cloning and Sequencing of RNA Ligands. Cloning of the nucleic acid repertoire was performed as described by Tuerk and Gold (1990) using double-stranded DNA that was generated from the RNA repertoire by PCR amplification. PCR-amplified DNA was digested with the restriction enzymes *Hin*dIII and *Bam*HI and ligated into compatible sites within pUC18. Ligated plasmids were transformed into *E. coli* strain JM103 and plated onto LB agar containing 5-bromo-4-chloro-3-indolyl β-D-galactoside, isopropyl thiogalactoside, and 100 μg/mL ampicillin. Colonies not expressing β-galactosidase were analyzed by hybridization to the 3′ PCR primer, and positive colonies were chosen for further analysis by direct sequencing of the cloned insert. Sequencing of DNA was performed as described by Tuerk and Gold (1990) using the dideoxynucleotide procedure of Sanger et al. (1977). Plasmids were isolated from *E. coli* JM103 by the alkaline lysis miniprep procedure (Manitatis et al., 1982). DNA was incubated in 50 mM Tris-HCl (pH 8.3), 60 mM NaCl, 6 mM magnesium acetate, and 1 mM DTT with 0.4 mM dNTP and 0.2 mM dideoxy-NTP for 20 min at 48 °C. AMV reverse transcriptase was present at 4 units per reaction. The reactions were stopped by the addition of 10 μL of loading buffer and heated to 95 °C for 3 min prior to gel electrophoresis on a 6% polyacrylamide/8 M urea denaturing gel.

Boundary Determination. Boundary experiments were performed essentially as described by Ringquist et al. (1994) and followed the procedures of Carey et al. (1983) and Tuerk et al. (1990). Briefly, the RNA transcript was 5′-end-labeled by incubation with alkaline phosphatase (1 unit) at 37 °C for 30 min and with polynucleotide kinase (2 units) in the presence of [γ-^{32}P]ATP for 30 min at 37 °C. Labeling of the 3′-end with [5′-^{32}P]pCp was performed with RNA ligase for 30 min at 37 °C. The 5′- as well as the 3′-labeled RNAs were band purified on a 6% polyacrylamide/8 M urea gel. End-labeled RNA (75 pmol) was hydrolyzed in 50 mM Na_2CO_3 (pH 9.0) and 1 mM EDTA in a volume of 100 μL for 10 min at 90 °C. The hydrolysis reaction was stopped by the addition of 11 μL of 3 M sodium acetate (pH 5.2).

ppp-gggagcucagaauaaacucgcaa[30N]uucgacaugaggcccggauccggc-OH

FIGURE 1: Nucleotide sequence of the starting RNA repertoire. The fixed sequences are shown in lower case letters, and the position of the 30-nucleotide randomized sequence is indicated (30N). The position of the AUG trinucleotide within the 3′-fixed region is underlined.

Table 2: SELEX Summary

experiment	target	apparent K_d (nM)[a] round 0	apparent K_d (nM)[a] round 11	K_d^0/K_d^{11} [b]	round cloned	principal RNA ligand
A	S1-depleted 30S ribosomes	110	20	6	12	SD element
B	S1-depleted 30S ribosomes plus $tRNA_f^{Met}$				12	SD element
C	30S ribosomes	150	4	38	13	pseudoknot
D	30S ribosomes plus $tRNA_f^{Met}$	70	8	9	11	pseudoknot
E	ribosomal protein S1	120	40	3	13	pseudoknot

[a] The apparent K_d was measured by nitrocellulose filtration as described in the Experimental Procedures. [b] The ratio of the apparent K_d of the RNA repertoire measured at rounds 0 and 11.

Table 3: S1-Depleted 30S Ribosome and S1-Depleted 30S Ribosome Plus $tRNA_f^{Met}$ Ligands[a]

motif	sequence of 30N region
	S1-Depleted 30S Ribosome Ligands
A1	AAG**AGGAGG**UGAUGAAACGAGAAGCCGAGA
A2	**aAGGAGG**UGAGAAGAAGACGAAGUGCCGCGA
A3	**AAAGGAGG**UGAAUGAAGGGACCGGGUAGUA
A4	**AAGGAGG**UGAAAAGGAAUGAUCCGGGUAAA
A5	**aAGGAGG**UGGAAAGAUUGCUGAUCCGGGAGA
A6	**aaGGAGG**UGAGAAGAAGACGAAAGUGCCGCGA
	S1-Depleted 30S Ribosome Plus $tRNA_f^{Met}$ Ligands
B1	AG**AAGGAGG**AAAAAAGAGCAAAGAGGACAA
B2	**aaGGAGG**UGAGAAGAUUGGAUCACGGGCAGAA
B3	**aAGGAGG**UGAAGAGACGAAAGAUCCGGGAAA
consensus	**AAGGAGG**UGA (76% purine-rich sequence)

[a] SD elements are indicated with bold font. Lower case letters indicate fixed sequences, while upper case letters indicate the randominized region. The conserved UGA trinucleotide and the semiconserved GAUCCGGG sequence are underlined.

Table 4: 30S Ribosome and 30S Ribosome Plus $tRNA_f^{Met}$ Ligands[a]

motif	stem 1	stem 2	loop 1	loop 2	fraction[b] of pool
		30S Ribosome Ligands			
C1	GAUG CUAC	CGAACA gcuuGU	ACA	AAGGAAC	0.55
C2	GAUG CUAC	UCGAAG agcuuC	ACA	GAGGAAC	0.14
C3	AGUG UCAC	GAACAG cuuGUC	ACA	AUGGAAC	0.05
		30S Ribosome Plus $tRNA_f^{Met}$ Ligands			
D1	GAUG CUAC	CGAACA gcuuGU	ACA	AAGGAAC	0.20
D2	GGUG CCAC	GAACAC cuuGUG	ACA	AAGGAAU	0.07
consensus	GAUG CUAC	CGAACA gcuuGU	ACA	AAGGAAC	

[a] Upper case letters indicate nucleotides within the 30N randomized region, while lower case letters indicate sequences associated with the 3′-fixed sequence. [b] The frequency of each motif in the ligand-enriched repertoire is calculated from a pool size of 44 sequences.

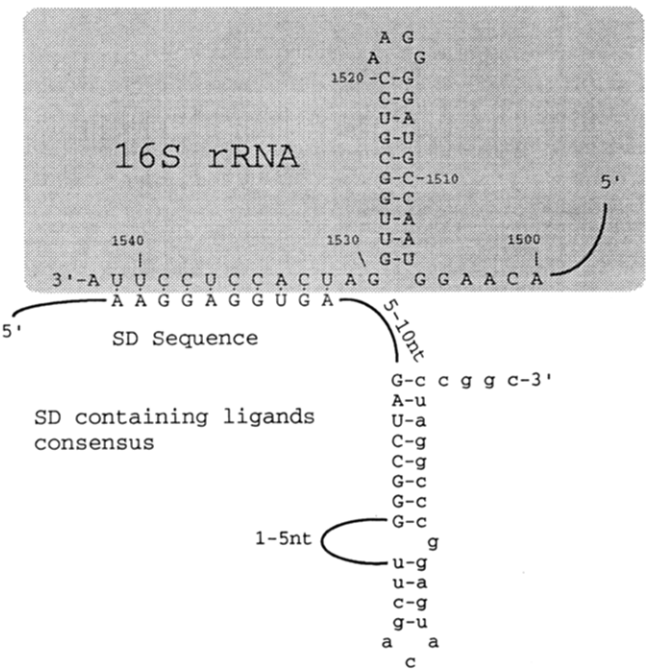

FIGURE 2: Ligands generated against S1-depleted 30S ribosomes can form extra base pairs to the 3′-end of 16S rRNA. The sequence of *E. coli* 16S rRNA from nucleotides 1500 to 1542 are shown (shaded) along with the plausible interaction of the SD-containing ligands.

Binding reactions were performed at the target concentrations described in the Results section and were carried out in binding buffer (10 mM Tris−acetate (pH 7.6), 100 mM NH_4Cl, 10 mM magnesium acetate, and 0.1 mM DTT). The concentration of hydrolyzed RNA was 150 nM. Bound, end-labeled RNA was isolated from unbound RNA by filtration through nitrocellulose membranes (Millipore, Inc., Bedford, MA) and rinsed with 5 mL of wash buffer (10 mM Tris-HCl, pH 7.5). The RNA was eluted from the filters by incubation in an emulsion of 200 μL of 7 M urea and 400 μL of phenol (equilibrated against 1 M Tris-HCl, pH 8.0) for 30 min at room temperature. The urea−phenol emulsion was separated into phases by the addition of 200 μL of chloroform followed by centrifugation. The aqueous phase was removed, and the RNA was precipitated upon the addition of 20 μL of 3 M sodium acetate, 20 mg of carrier tRNA, and 2 vol of ethanol. The pellet was washed once with 70% ethanol, dried, and resuspended in 5 μL of distilled H_2O followed by the addition of 5 μL of formamide loading dye (72% formamide, 120 mM EDTA, pH 8.0) prior to electrophoresis.

A sequence ladder of the RNA ligands was generated by partial RNAse T1 digestion. The RNase T1 digestion was performed in a 10 μL reaction containing 0.5 pmol of end-labeled RNA and 10 units of RNAse T1 in 7 M urea, 20 mM sodium citrate (pH 5.0), and 1 mM EDTA. The RNA was incubated for 10 min at 50 °C without enzyme and then for another 10 min after the addition of enzyme.

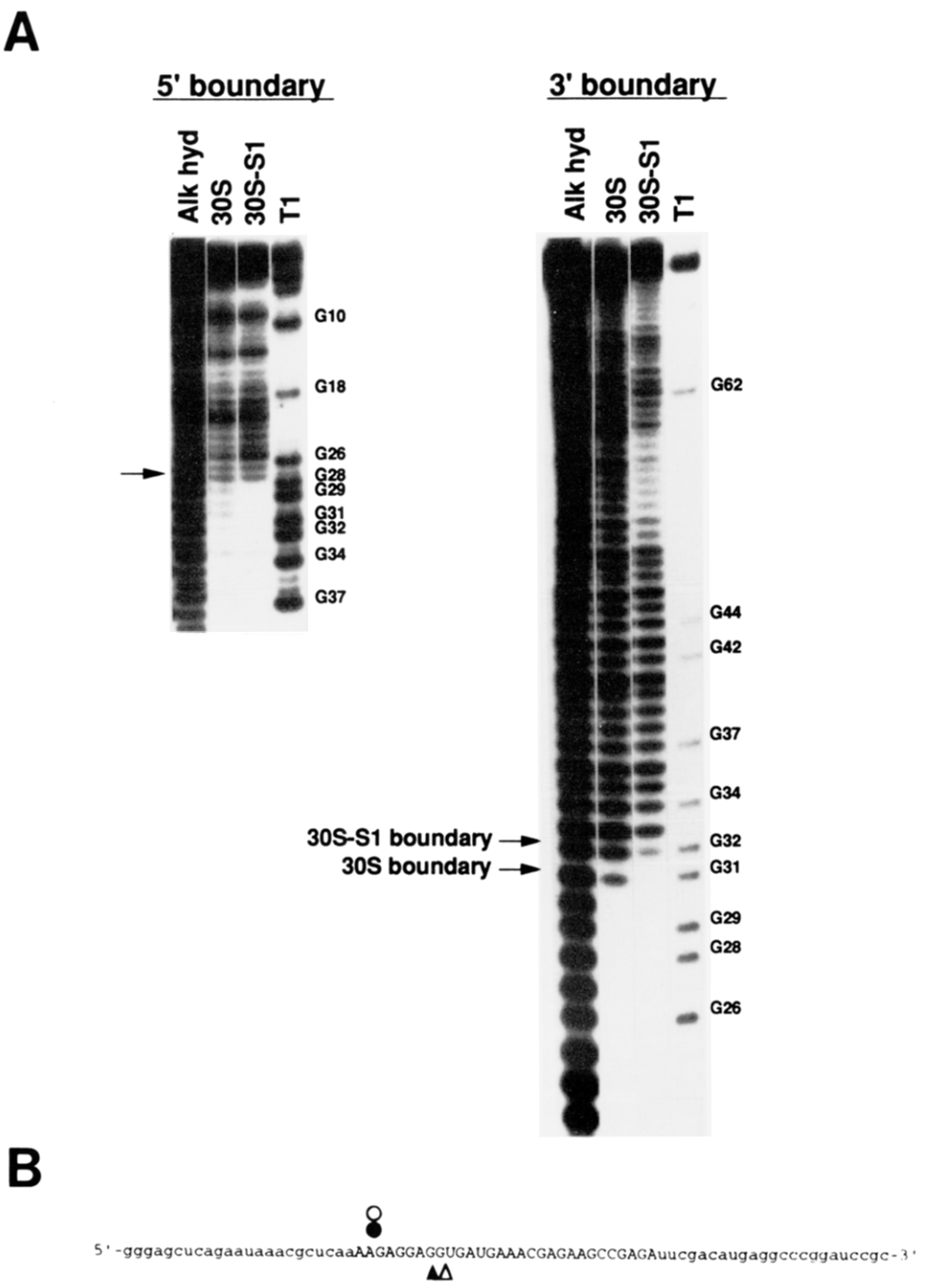

FIGURE 3: Intact 30S ribosomes and S1-depleted 30S ribosomes resulted in overlapping boundaries on ligand A1. (A) Autoradiograms illustrating the boundary data for the binding of ligand A1 to intact and S1-depleted 30S ribosomes. The 5′- and 3′-boundaries, along with the partial RNAse T1 digest and alkaline hydrolysis ladders, are shown. (B) Schematic of diagram ligand A1. Triangles and circles indicate the 5′- and 3′-boundaries for binding to intact ribosomes (closed symbols) and S1-depleted ribosomes (open symbols), respectively. The concentration of intact and S1-depleted 30S ribosomes was 1 nM, while the concentration of RNA ligands was 150 nM. Upper and lower case letters indicate sequences derived from the randomized and fixed regions, respectively.

The reaction was stopped by adding 10 μL of loading buffer.

Equilibrium Binding Analysis. Briefly, equilibrium binding analysis was performed by incubating ^{32}P-labeled RNA with protein in binding buffer. Samples were incubated at 37 °C for 5 min followed by nitrocellulose filtration as described by Tuerk and Gold (1990). Nitrocellulose filters were washed with 10 mM Tris-HCl (pH 8.0) and dried, and the amount of [^{32}P]RNA was determined by scintillation counting in a Beckman LS-133 liquid scintillation counter using ecolume scintillation fluid (ICN Biomedicals, Inc., Costa Mesa, CA). The apparent dissociation constants were determined from a plot of the fraction of bound RNA versus the log concentration of protein (Klotz, 1985).

RESULTS

Selection of High-Affinity RNA Ligands. SELEX was used to generate high-affinity RNA ligands against intact 30S ribosomes, S1-depleted 30S ribosomes, and purified ribosomal protein S1. Ligands were isolated from a single starting repertoire in the presence and absence of $tRNA_f^{Met}$. The starting repertoire consisted of a 30-nucleotide random-

ized region flanked by two fixed sequences (Figure 1) and contained a translational start codon, AUG. SELEX was performed for between 11 and 13 rounds and resulted in a 3−38-fold increase in the apparent affinity of the ligand-enriched RNA repertoire for target (Table 2). The largest increase was observed for intact 30S ribosomes; the round 11 repertoire exhibited an apparent dissociation constant of 4 nM. Each ligand-enriched RNA repertoire was analyzed by cloning and sequencing. Comparison of the nucleotide sequence of individual RNAs indicated the presence of at least two classes of RNA ligand.

Comparison of High-Affinity RNA Ligands. (A) S1-Depleted 30S Ribosomes. RNA ligands generated against S1-depleted 30S ribosomes were analyzed by sequencing individual, cloned ligands. Representing a class of relatively unstructured RNAs, the ligands were purine-rich and contained an SD sequence (AAGGAGG) at the 5′-edge of the randomized region (Table 3). The SD sequence was identical to the consensus, natural RBSs from *E. coli* (Scherer et al., 1980; Stormo et al., 1982; Dreyfus, 1988; Rudd & Schneider, 1992). The SD sequence was generated to S1-depleted ribosomes but, surprisingly, not to intact 30S ribosomes (see Table 4), and the addition of $tRNA_f^{Met}$ to the selection had no discernible effect on the sequence of the ligands generated. Moreover, the ligands shared a conserved UGA trinucleotide that occurred immediately 3′ to the SD sequence and may have resulted from increased base pairing to the 3′-end of 16S rRNA (Figure 2) (Thanaraj & Pandit, 1989). Ligands also shared a weakly conserved octanucleotide sequence (GAUCCGGG) within the 3′ randomized region that could base pair to the 3′-fixed sequence. Overall, the nucleotide sequence of the randomized region was purine-rich (Table 3), consistent with the high G and A content of natural *E. coli* RBSs (Scherer et al., 1980; Stormo et al., 1982; Rudd & Schneider, 1992).

Boundary analysis of ligand A1 was performed in order to identify the nucleotide residues essential for the binding of S1-depleted 30S ribosomes (the target), as well as intact 30S ribosomes. The results indicated overlapping binding sites for these ribosome preparations (Figure 3). Comparison of the boundary data showed that the position of the 5′-boundary was independent of the presence of ribosomal protein S1, and binding correlated to the edge of the conserved SD sequence (Figure 3). In contrast, the 3′-boundary shifted by one nucleotide when S1 was present. Moreover, the conserved UGA trinucleotide was not part of the binding sequence for either target, suggesting that factors directing ligand selection during SELEX and affecting binding during boundary analysis may differ. Small differences in the stability of complexes may have a large effect during SELEX on the identity of the winning ligand (Irvine et al., 1991).

(B) 30S Ribosomes and Purified Ribosomal Protein S1. RNA ligands to intact ribosomes and purified ribosomal protein S1 were also examined. These ligands exhibited substantial secondary structure, as well as highly conserved sequence elements (Tables 4 and 5). Table 4 summarizes the different structural elements for ligands generated against intact ribosomes along with their frequency in the ligand-enriched repertoire. Ligands to 30S ribosomes were almost identical to those generated against purified ribosomal protein S1 (Table 5), probably reflecting identical RNA-binding sites as well as the use of a single starting RNA repertoire. Both

Table 5: Ribosomal Protein S1 Ligands[a]

motif	stem 1	stem 2	loop 1	loop 2	fraction[b] of pool
E1	GAUG CUAC	CGAAAG gcuuUC	ACA	AAGGAAC	0.46
E2	GGAG CCUC	GAGUGC uuCACG	ACC	AAGGAAC	0.13
E3	ACAG UGUC	CGAACA gcuuGU	ACG	GUGGAAC	0.13
E4	GCUG CGAC	CGAAGA gcuuCU	ACA	ACGGAAC	0.13
E5	GAUG CUAC	CGAACA gcuuGU	ACA	AAGGAAC	0.08
E6	AGUG UCAC	UCGAGG agcuuC	ACG	AAGGAAC	0.08
consensus	GAUG CUAC	CGAAAG gcuuUC	ACA	AAGGAAC	

[a] Upper case letters indicate nucleotides from the 30N randomized region, while lowercase letters represent the 3′-fixed sequence. [b] The frequency of each motif is calculated from a pool size of 24 sequences.

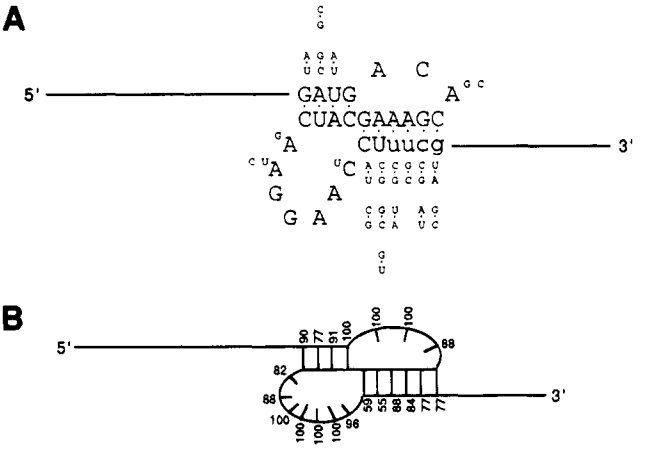

FIGURE 4: RNA ligands generated against intact 30S ribosomes and purified ribosomal protein S1 were identical and formed an RNA pseudoknot. (A) Consensus nucleotide sequence and structure of the ligands generated to intact 30S ribosomes and purified ribosomal protein S1. Large font indicates the consensus sequence. Small font indicates the positions and base pairs that vary from those of the consensus. (B) Schematic representation of the RNA pseudoknot. Numbers indicate the frequency of individual positions compared to the consensus.

targets resulted in ligands that contained a potential pseudoknot and shared identical consensus sequences at loops 1 and 2. Importantly, base changes within the two stems of the pseudoknot consisted entirely of compensatory base changes (Figure 4); in some ligands stem 1 could make an additional base pair, but it would have represented the only position where noncompensatory base changes were observed (data not shown). Another characteristic of the pseudoknot-containing ligands was the selection of AUG or GUG sequences within stem 1 and a Shine−Dalgarno-like element within loop 2. These motifs have been shown to regulate translational initiation (Shine & Dalgarno, 1974; Steitz & Jakes, 1975; Scherer et al., 1980; Stormo et al., 1982; Hui & deBoer, 1987; Jacob et al., 1987; Dreyfus, 1988; Ringquist et al., 1992; Rudd & Schneider, 1992) and may reflect a specific RNA−protein interaction.

The pseudoknot-containing ligand E1 was analyzed by boundary and binding assays using purified ribosomal protein S1 as well as intact and S1-depleted 30S ribosomes (Figure

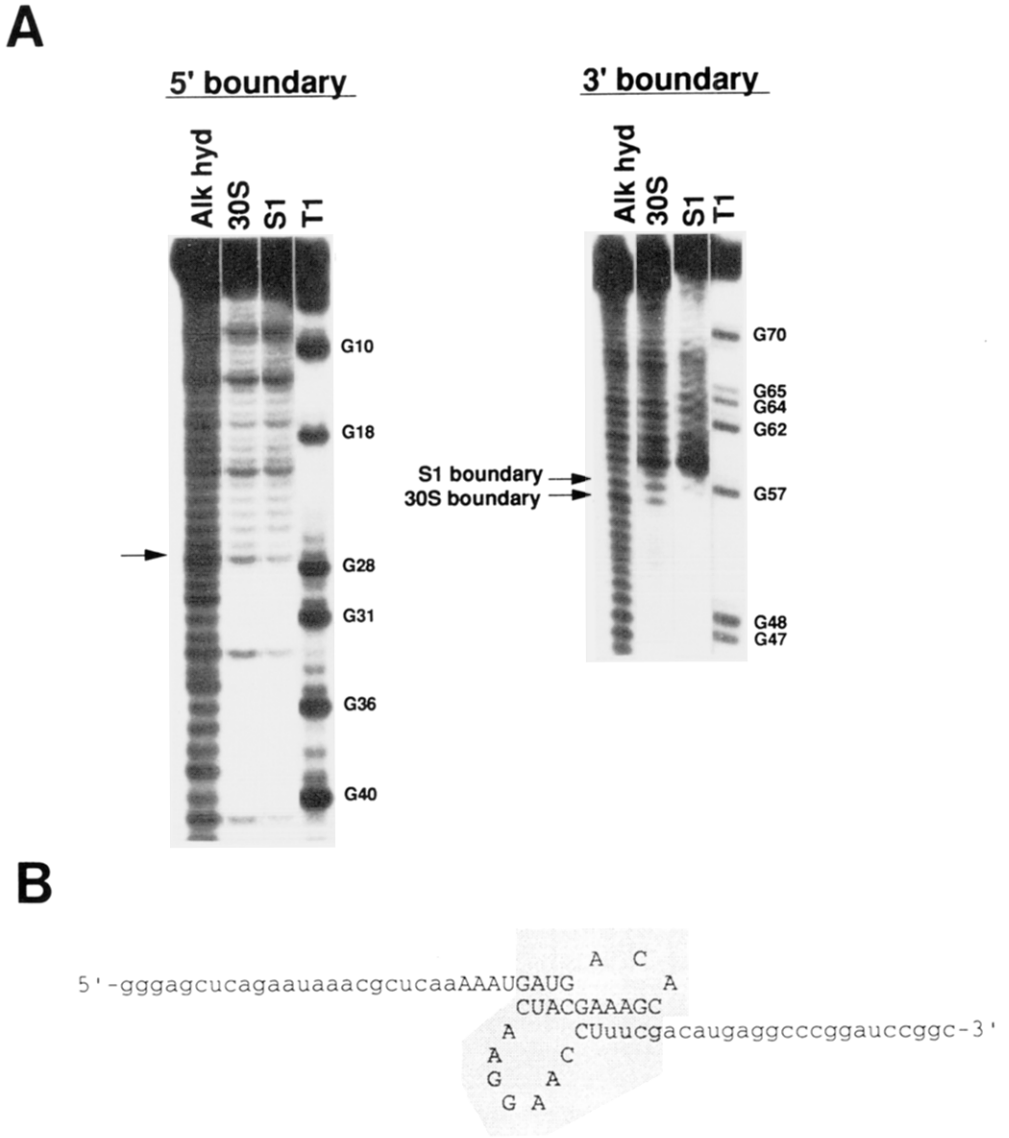

FIGURE 5: Intact 30S ribosomes and purified ribosomal protein S1 gave identical boundaries for binding to ligand E1. (A) Autoradiograms illustrating the boundary data for intact 30S ribosomes and purified S1 protein. The 5′- and 3′-boundaries, along with the partial RNAse T1 digest and alkaline hydrolysis ladders, are shown. (B) Schematic diagram of ligand E1. The shaded region indicates the boundaries for binding to intact 30S ribosomes. Upper and lower case letters indicate residues derived from the 30-nucleotide randomized region and the fixed sequences, respectively. The concentration of 30S ribosome and S1 protein was 4 nM, while RNA was present at 150 nM.

Table 6: Specificity of RNA Ligands

	apparent K_d (nM)	
ligand	ribosomal protein S1	30S ribosome
round 0 repertoire	120	150
SD element (A1)	200	7
pseudoknot (E1)	4	5

5 and Table 6). Consistent with its similarity to ligands generated against intact ribosomes, ligand E1 exhibited overlapping boundaries when incubated with free S1 protein (the target) or intact 30S ribosomes (Figure 5), but failed to bind S1-depleted 30S ribosomes (data not shown). The boundaries for binding S1 protein or intact ribosomes corresponded to the edges of stems 1 and 2. Binding analysis indicated that ligand E1 bound purified S1 and intact ribosomes with identical affinities (Table 6), suggesting that the pseudoknot-containing ligand recognized free and ribosome-bound S1 protein. Moreover, competition experiments between ligands A1 and E1 for binding to intact 30S ribosomes were consistent with the presence of an independent, high-affinity site for each ligand (data not shown). Interestingly, in a separate experiment, a similar pseudoknot-containing ligand was generated against Qβ-replicase (containing S1 protein) (D. Brown, personal communication).

DISCUSSION

High-affinity RNA ligands were generated against intact 30S ribosomes, S1-depleted 30S ribosomes, and purified ribosomal protein S1. Sequence analysis indicated two classes of high-affinity ligand: SD-containing ligands generated against S1-depleted ribosomes and RNA pseudoknot-containing ligands generated against intact ribosomes and ribosomal protein S1. Analysis of each class indicated that binding to S1-depleted ribosomes was governed by base pairing to the 3′-end of 16S rRNA, while binding to intact 30S ribosomes was dominated by ribosome-bound S1 protein. This view was supported by equilibrium binding analysis as well as boundary determinations and competition experiments, suggesting that mRNA elements (in addition

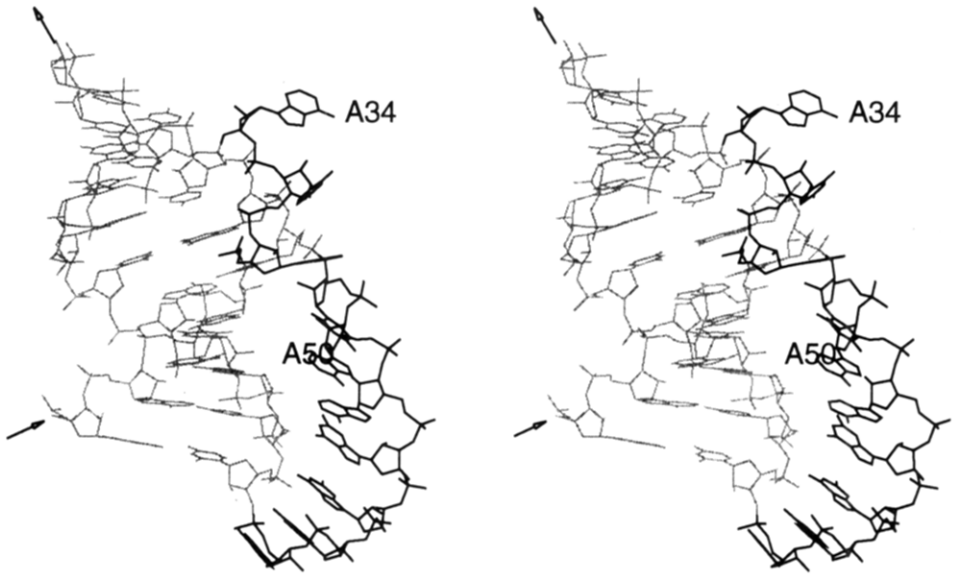

FIGURE 6: Stereodrawing of the plausible three-dimensional structure of the RNA pseudoknot. The loops are indicated (bold) along with residues A34 (loop 1) and A50 (loop 2).

```
Pseudoknot Loops 2 & 1:        AAGGAA  C      ACA
Qβ Binding Site:       AAUAAAUUA  UC     ACAAUUACUCUUACG
Loechel Enhancer:              AAUUAA  CA   ACAUAAUUUAAACAAAUG
S1 RBS:                UGAAGAUUAAACAUGAC  UGAAUC
```

FIGURE 7: Loops from the pseudoknot-containing ligands homologous to natural S1-binding sites in *E. coli* and bacteriophage. (A) Sequence of loops 2 (left) and 1 (right) along with natural RNA elements (Senear & Steitz, 1976; Skouv et al., 1980; Loechel et al., 1992). The translational start codons from the translation enhancer described by Loechel et al. (1992) and the S1 RBS (Skouv et al., 1980) are underlined. Shaded regions indicate the positions of highest homology to loops 2 and 1 of the RNA pseudoknot.

to the well-studied SD sequence) may regulate translational initiation.

SD-Containing Ligands. Ligands generated against S1-depleted ribosomes contained SD sequences and exhibited extensive complementarity to the 3′-end of 16S rRNA, consistent with annealing to the anti-SD sequence (Shine & Dalgarno, 1974; Steitz & Jakes, 1975; Hui & de Boer, 1987; Jacobs et al., 1987). Moreover, these ligands were reminiscent of strong RBSs from *E. coli* as well as *Bacillus subtilis* and other Gram-positive bacteria lacking an S1 homologue (Roberts & Rabinowitz, 1989). Located at the 5′-edge of the randomized region, the placement of the SD sequence suggested that elements downstream were also important for high-affinity binding. For instance, the SD-containing ligands exhibited a conserved UGA trinucleotide immediately adjacent to the SD sequence; the UGA may have resulted from enhanced base pairing to 16S rRNA (Shine & Dalgarno, 1974; Thanaraj & Pandit, 1989). Naturally occurring, highly expressed mRNAs from *E. coli* also exhibit a similar element (Thanaraj & Pandit, 1989). Addition of UGA to the SD sequence may affect the rate of dissociation of the RNA−ribosome complex: its amplification during SELEX may represent an effect on the rate of dissociation of the translational initiation complex (Irvine et al., 1991; Ringquist et al., 1993a; Spedding et al., 1993).

The unusually long spacing between the SD element and the AUG trinucleotide, located within the 3′-fixed region, suggests that the SELEX protocol may have been insensitive to the presence of initiator tRNA. Successful competition between high-affinity nucleic acids determines the winning ligand (Irvine et al., 1991). Ligands with enhanced on-rates may also have a decided advantage, especially for a target such as the ribosome, which exhibits an extremely slow rate of dissociation from natural RBSs (Calogero et al., 1988; Ringquist et al., 1993a; Spedding et al., 1993). The UGA-containing ligands were probably evolved this way, and the apparent secondary structure near the AUG trinucleotide (see Figure 2) suggests that the structural motif may have resulted in an advantage during SELEX. Moreover, the apparent absence of SD ligands in the enriched repertoire, generated against intact ribosomes (see the following), suggests that high-affinity ligands to different domains of a target may compete during SELEX.

RNA Pseudoknot-Containing Ligands. High-affinity RNA ligands were generated against purified S1 protein and intact 30S ribosomes. Ligands generated against these targets were identical, indicating that free and ribosome-bound S1's bind RNA. Surprisingly, ligands generated against intact ribosomes consisted of an RNA pseudoknot and not a typical SD sequence. The prevalence of the pseudoknot-containing ligands suggests that, at the concentrations of ribosome and RNA used during SELEX, ribosome-bound S1 protein may be the dominant RNA epitope on the ribosome. Moreover, the loops were highly conserved, directly implicating them in binding. How does the pseudoknot affect binding between RNA and S1 protein? One explanation is that the pseudoknot structure holds the loops in a conformation advantageous for binding. Importantly, in its folded conformation, the pseudoknot presents the loops on the same side of the helix but in reversed order; the RNA is folded so that the sequence within loop 2 is placed 5′ to loop 1 (Figure 6). The relatively fixed structure of the loops may act to decrease the loss in entropy upon binding to S1, subsequently enhancing the affinity of the ligand for S1 protein.

Binding of S1 protein to specific elements within the TIR has been proposed to modulate translational initiation (van Dieijen et al., 1976; Suryanarayana & Subramanian, 1984; Roberts & Rabinowitz, 1989; Skouv et al., 1990; Boni et al., 1991; Zhang & Deutscher, 1992). However, while earlier studies presented evidence from homopolymer-binding studies (Draper & von Hippel, 1979; Draper et al., 1977) or cross-linking data (Boni et al., 1991), the current work presents the amplification of specific, high-affinity ligands against ribosomal protein S1. The pseudoknot-containing ligands reported here were homologous to natural TIRs (Figure 7). For instance, the translational initiation site at the autogenously regulated S1 RBS (Christiansen & Pedersen, 1981; Skouv et al., 1990) and the *E. coli* translational enhancer element described by Loechel et al. (1992) exhibit RNA sequences similar to those of the loops of the pseudoknot-containing ligands. These elements, along with a similar sequence within the S1-binding domain from the bacteriophage Qβ genome (Senear & Steitz, 1976; Goelz & Steitz, 1977; Barrera et al., 1993), are illustrated in Figure 7 and suggest a potential regulatory role for S1 protein during gene expression. S1 may recognize specific sequences within these RNAs and through protein−protein contacts link them to the translational machinery. Moreover, Ruckman et al. (1994) showed that the sequence-specific RNA endonuclease from bacteriophage T4, the product of the *regB* gene, requires S1 for efficient cleavage at GGAG sequences. Sequence-specific binding of RNA by S1 protein may provide a general mechanism to regulate gene expression. Association of S1 with different proteins (e.g., the ribosome, Qβ-replicase, and the product of the *regB* gene) may allow nucleic acid binding by S1 to regulate a wide assortment of processes. S1, with its multiple nucleic acid-binding sites (Draper et al., 1977; Draper & von Hippel, 1979) and elongated structure (Subramanian, 1983, 1984), may provide a convenient link between a variety of proteins (e.g., those involved in translation as well as replication) and their RNA substrates.

ACKNOWLEDGMENT

We are grateful to the W. M. Keck Foundation for their generous support of RNA science at the Boulder campus. We also thank Dan Schneider for pointing out the potential pseudoknot in some of the RNA ligands, Christine Brunel for computer modeling of the pseudoknot ligands, and Dave Brown for sharing data prior to publication.

REFERENCES

Barrera, I., Schuppli, D., Sogo, J. M., & Weber, H. (1993) *J. Mol. Biol. 232*, 512−521.

Blasi, U., Nam, K., Hartz, D., Gold, L., & Young, R. (1989) *EMBO J. 8*, 3501−3510.

Boni, I. V., Isaeva, D. M., Musychenko, M. L., & Tzareva, N. V. (1991) *Nucleic Acids Res. 19*, 155−162.

Calogero, R. A., Pon, C. L., Canonaco, M. A., & Gualerzi, C. O. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6427−6431.

Carey, J., Cameron, V., de Haseth, P. L., & Uhlenbeck, O. C. (1983) *Biochemistry 22*, 2601−2610.

Christiansen, L., & Pedersen, S. (1981) *Mol. Gen. Genet. 181*, 548−551.

deSmit, M. H., & van Duin, J. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7668−7672.

deSmit, M. D., & van Duin, J. (1990b) *Prog. Nucleic Acid Res. Mol. Biol. 38*, 1−35.

deSmit, M. D., & van Duin, J. (1994) *J. Mol. Biol. 235*, 173−184.

Draper, D. E., & von Hippel, P. H. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1040−1044.

Draper, D. E., Pratt, C. W., & von Hippel, P. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4786−4790.

Dreyfus, M. (1988) *J. Mol. Biol. 204*, 79−94.

Goelz, S., & Steitz, J. A. (1977) *J. Biol. Chem. 252*, 5177−5179.

Gold, L. (1988) *Annu. Rev. Biochem. 57*, 199−233.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Swebilius-Singer, B., & Stormo, G. (1981) *Annu. Rev. Microbiol. 35*, 365−403.

Gold, L., Stormo, G., & Saunders, R. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7061−7065.

Gualerzi, C. O., & Pon, C. L. (1990) *Biochemistry 29*, 5881−5889.

Hartz, D., McPheeters, D., & Gold, L. (1991) *J. Mol. Biol. 218*, 83−97.

Heus, H. A., & Pardi, A. (1991) *Science 253*, 191−194.

Hui, A., & deBoer, H. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4762−4766.

Irvine, D., Tuerk, C., & Gold, L. (1991) *J. Mol. Biol. 222*, 739−761.

Jacob, W. F., Santer, M., & Dahlberg, A. E. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4757−4761.

Klotz, I. M. (1985) *Q. Rev. Biophys. 18*, 227−259.

Kozak, M., & Nathans, D. (1972) *Bacteriol. Rev. 36*, 109−134.

Loechel, S., Inamine, J. M., & Hu, P.-C. (1992) *Nucleic Acids Res. 19*, 6905−6911.

Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., deBoer, H. A., & van Knippenberg, P. H. (1987) *EMBO J. 6*, 2489−2492.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265−275.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

McCarthy, J. E. G., & Gualerzi, C. O. (1990) *Trends Genet. 6*, 78−85.

McPheeters, D. S., Christensen, A., Young, E., Stormo, G. D., & Gold, L. (1986) *Nucleic Acids Res. 14*, 5813−5826.

McPheeters, D. S., Stormo, G. D., & Gold, L. (1988) *J. Mol. Biol. 201*, 517−535.

Melancon, P., Leclerc, D., Destroismaisons, N., & Brakier-Gingras, L. (1990) *Biochemistry 29*, 3402−3407.

Moine, H., Romby, R., Springer, M., Grunberg-Manago, M., Ebel, J. P., Ehresmann, C., & Ehresmann, B. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7892−7896.

Munson, L. M., Stormo, G. D., Niece, R. L., & Reznikoff, W. S. (1984) *J. Mol. Biol. 177*, 663−683.

Olins, P. O., & Rangwala, S. H. (1989) *J. Biol. Chem. 264*, 16973−16979.

Petersen, G. B., Stockwell, P. A., & Hill, D. F. (1988) *EMBO J. 7*, 3957−3962.

Ray, P. N., & Pearson, M. L. (1975) *Nature 253*, 647−650.

Ringquist, S., Shinedling, S., Barrick, B., Green, L., Binkley, J., Stormo, G. D., & Gold, L. (1992) *Mol. Microbiol. 6*, 1219−1229.

Ringquist, S., Cunningham, P., Weitzmann, C., Formenoy, L., Pleij, C., Ofengand, J., & Gold, L. (1993a) *J. Mol. Biol. 234*, 14−27.

Ringquist, S., MacDonald, M., Gibson, T., & Gold, L. (1993b) *Biochemistry 32*, 10254−10262.

Ringquist, S., Schneider, S., Gibson, T., Baron, C., Böck, A., & Gold, L. (1994) *Genes Dev. 8*, 376−385.

Roberts, M. W. & Rabinowitz, J. C. (1989) *J. Biol. Chem. 264*, 2228−2235.

Ruckman, J., Ringquist, S., Brody, E., & Gold, L. (1994) *J. Biol. Chem. 269*, 26655−26662.

Rudd, K. E., & Schneider, T. D. (1992) in *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria* (Miller, J., Ed.) pp 17.19−17.45, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.

Scherer, G. F. E., Walkinshaw, M. D., Arnott, S., & Morre, D. J. (1980) *Nucleic Acids Res. 8*, 3895−3907.

Senear, A. W., & Steitz, J. A. (1976) *J. Biol. Chem. 251*, 1902−1912.
Shean, C. S., & Gottesman, M. E. (1992) *Cell 70*, 513−522.
Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342−1346.
Shinedling, S., Gayle, M., Pribnow, D., & Gold, L. (1987) *Mol. Gen. Genet. 207*, 224−232.
Skouv, J., Schnier, J., Rasmussen, M. D., & Subramanian, A. R. (1990) *J. Biol. Chem. 265*, 17044−17049.
Spedding, G., Gluick, G., & Draper, D. E. (1993) *J. Mol. Biol. 229*, 609−622.
Steitz, J. A., & Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4734−4738.
Stormo, G., Schneider, T., & Gold, L. (1982) *Nucleic Acids Res. 10*, 2971−2996.
Subramanian, A. R. (1983) *Prog. Nucleic Acids Res. Mol. Biol 28*, 101−142.
Subramanian, A. R. (1984) *Trends Biochem. Sci. 9*, 491−494.
Suryanarayana, T., & Subramanian, A. R. (1984) *Biochemistry 23*, 1047−1051.
Thanaraj, T., & Pandit, M. W. (1989) *Nucleic Acids Res. 17*, 2973−2985.
Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.
Tuerk, C., Eddy, S., Parma, D., & Gold, L. (1990) *J. Mol. Biol. 213*, 749−761.
Van Knippenberg, P. H. (1975) *Nucleic Acids Res. 2*, 79−85.
Winter, R. B., Morrissey, L., Gauss, P., Gold, L., Hsu, T., & Karam, J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7822−7826.
Wulczyn, F. G., Bolker, M., & Kahman, R. (1989) *Cell 57*, 1201−1210.
Zhang, J., & Deutscher, M. P. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2605−2609.

BI942101D