

Structure and Function of the Conserved 690 Hairpin in *Escherichia coli* 16 S Ribosomal RNA: Analysis of the Stem Nucleotides

Svetlana V. Morosyuk¹, KangSeok Lee³, John SantaLucia Jr^{1*} and Philip R. Cunningham^{2*}

¹Department of Chemistry Sciences, Wayne State University, Detroit, MI 48202, USA

²Department of Biological Sciences, Wayne State University, Detroit, MI 48202, USA

³Stanford University School of Medicine, Department of Genetics, Stanford CA 94305, USA

Nucleotides 680 to 710 of *Escherichia coli* 16 S rRNA form a distinct structural domain required for ribosome function. The goal of this study was to determine the functional significance of pairing interactions in the 690 region. Two different secondary structures were proposed for this hairpin, based on phylogenetic and chemical modification studies. To study the effect of pairing interactions in the 690 hairpin on ribosome function and to determine which of the proposed secondary structures is biologically significant, we performed an instant-evolution experiment in which the nine nucleotides that form the proposed base-pairs and dangling ends of the 690 stem were randomly mutated, and functional mutant combinations were selected. A total of 96 unique functional mutants were isolated, assayed *in vivo*, and sequenced. Analysis of these data revealed extensive base-pairing and stacking interactions among the mutated nucleotides. Formation of either a Watson-Crick base-pair or G·U pair between positions 688 and 699 is absolutely required for ribosome function. We also performed NMR studies of a 31-nucleotide RNA which indicate the formation of a functionally important base-pair between nucleotides 688 and 699. Formation of a second base-pair between positions 689 and 698, however, is not essential for ribosome function, but the level of ribosome function correlates with the predicted thermodynamic stability of the nucleotide pairs in these positions. The universally conserved positions G690 and U697 are generally portrayed as forming a G·U mismatch. Our data show co-variation between these positions, but do not support the hypothesis that the G690:U697 pair forms a wobble structure. NMR studies of model 14-nt and 31-nt RNAs support these findings and show that G690 and U697 are involved in unusual stacking interactions but do not form a wobble pair. Preliminary NMR structural analysis reveals that the loop portion of the 690 hairpin folds into a highly structured and novel conformation.

© 2000 Academic Press

Keywords: 690 hairpin; mutation; ribosomal RNA folding; NMR; instant evolution

*Corresponding authors

Introduction

Ribosomal RNA provides the structural core for ribosome assembly and is directly involved in the

Abbreviations used: CAT, chloramphenicol acetyltransferase; MBS, mRNA binding sequence; RBS, ribosome binding sequence; MIC, minimal inhibitory concentration; NOESY, nuclear Overhauser enhancement spectroscopy.

E-mail addresses of the corresponding authors: jsl@chem.wayne.edu; Philip.Cunningham@wayne.edu

catalytic process (Dahlberg, 1989; Noller *et al.*, 1992; Yoshizawa *et al.*, 1999). Several ribosomal RNA sequences within the small subunit are highly conserved evolutionarily, but the functional roles of only a few of these nucleotides have been elucidated (Cunningham *et al.*, 1993; Noller, 1991). The nucleotide sequence corresponding to the 690 region found in *Escherichia coli* 16 S rRNA between positions 680 and 710 is highly conserved. Figure 1 shows the 690 hairpin secondary structure (cf. Table 1 for evolutionary conservation). Recent crys-

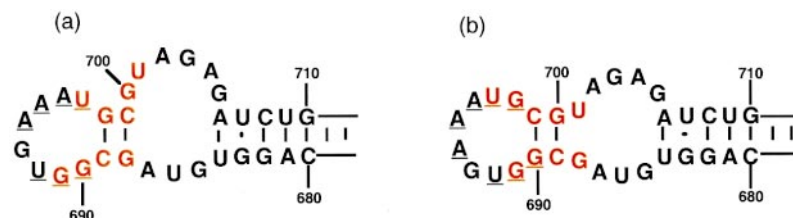


Figure 1. Two potential secondary structures of the *E. coli* 690 region. Nucleotides mutated in this study are in red. Nucleotides underlined are conserved in >95% of all bacteria (Van de Peer *et al.*, 1999). (a) Phylogenetically-derived secondary structure (Gutell, 1994). (b) Secondary structure suggested by chemical modification studies (Kean & Draper, 1985; Serganov *et al.*, 1996).

tal structures of the 30 S and 70 S ribosomes revealed that helix 23 is located at the top of the platform of the small subunit and forms a bridge with the 50 S subunit (Cate *et al.*, 1999; Clemons *et al.*, 1999). Recently, the X-ray crystal structure of the central domain of the small-subunit was resolved to 2.6 Å resolution, although the 690 hairpin was not resolved (Agalarov *et al.*, 2000). Chemical probing studies also support the hypothesis that the 690 loop plays an important role in subunit association (Merryman *et al.*, 1999). Chemical probing and cross-linking studies suggest that the 690 region is a component of the P-site (Joseph *et al.*, 1997). It was shown that P-site binding of tRNA^{Phe}, or a 15 bp anticodon stem-loop fragment of tRNA^{Phe}, strongly protects G693 from chemical modification with DMS, kethoxal and CMCT (Moazed & Noller, 1986, 1990). The 690 loop-stem and the anticodon loop-stem of P-site bound tRNA are proposed to be in parallel, based on the cross-linking pattern of modified tRNAs to the 16 S ribosomal RNA (Doring *et al.*, 1994; Osswald *et al.*, 1995; Rinke-Appel *et al.*, 1995). Direct contact of initiation factor three (IF3) with positions G700, U701 and G703 was also proposed based on strong protection of these positions from chemical modification upon binding of IF3 to the 30 S subunit (Moazed *et al.*, 1995; Muralikrishna & Wickstrom, 1989). Hairpin loop nucleotides of the 690 region were also shown to effect binding of the antibiotics

pactamycin and edeine, which inhibit initiation of protein synthesis (Egebjerg & Garrett, 1991; Mankin, 1997; Oehler *et al.*, 1997; Woodcock *et al.*, 1991). Strong pactamycin protection of G693 from chemical modification and the A694G mutation that confers pactomycin resistance suggest that the 690 hairpin loop nucleotides are directly involved in pactamycin binding (Egebjerg & Garrett, 1991; Mankin, 1997; Woodcock *et al.*, 1991). Tetracycline derivatives with tethered photo-activated moieties crosslink to positions G693, G1300 and G1338 of 16 S rRNA. It was proposed that these residues form a tetracycline-binding site on the 30 S subunit (Oehler *et al.*, 1997). The 690 region was also identified as a primary candidate for structural rearrangements in rRNA during translocation in a number of studies (Joseph *et al.*, 1997; Moazed & Noller, 1989; Mueller *et al.*, 1997; Stark *et al.*, 1997). In addition, two different secondary structures of the 690 region have been proposed (Figure 1). One is based on phylogenetic studies (Gutell, 1994) and the other is based on chemical modification and nuclease digestion *in vitro* (Kean & Draper, 1985; Serganov *et al.*, 1996). To identify base-pairing interactions in the 690 region which are important for ribosome function, we randomly mutated two base-pairs of the stem and five neighboring nucleotides of the 690 region. This instant-evolution study (Lee *et al.*, 1997) was used to identify functional mutants and determine the sequence and

Table 1. Sequence variation in the 690 hairpin

Nucleotide	688	689	690	691	692-696	697	698	699	700	701
A. Nucleotide distribution of functional mutants^a										
A	13(6) ^b	22(12)	18(12)	10(6)		31(13)	5(2)	7(3)	22(13)	22(13)
C	37(23)	<u>37(21)</u>	35(16)	3(3)		17(10)	39(18)	<u>25(15)</u>	6(2)	35(17)
G	<u>38(19)</u>	27(13)	<u>23(11)</u>	<u>66(35)</u>		1(0)	<u>35(23)</u>	37(23)	<u>51(31)</u>	7(4)
U	7(3)	9(5)	19(12)	16(7)		<u>46(28)</u>	16(8)	26(10)	16(5)	<u>31(17)</u>
B. Nucleotide distribution in all known bacteria^c										
A	82	597	297	2		0	31	0	120	20
C	6	<u>2711</u>	2	2		0	394	<u>3604</u>	13	649
G	<u>3656</u>	392	<u>3437</u>	<u>3737</u>		0	<u>2727</u>	0	<u>3384</u>	107
U	6	47	3	2		<u>3751</u>	584	132	235	<u>2979</u>

Underlined numbers indicate the wild-type *E. coli* sequence.

^a Nucleotide distribution of the 96 functional 690 hairpin mutants.

^b Parenthesis indicate nucleotide distribution of the 51 functional 690 hairpin mutants with MIC \geq 300 μ g/ml chloramphenicol.

^c Phylogenetic variation of the 690 stem nucleotides in bacterial ribosomes (Van de Peer *et al.*, 1999).

structural elements of the 690 loop that are required for function.

Results and Discussion

Random mutagenesis and selection of functional mutants

We developed a novel genetic system to study ribosomal RNA function *in vivo* (Lee *et al.*, 1996, 1997). In this system, the chloramphenicol acetyltransferase (CAT) reporter message (Burns & Crowl, 1987) is translated exclusively by plasmid-encoded ribosomes that cannot translate normal cellular messages (Lee *et al.*, 1996, 1997). This was achieved by mutations made in the mRNA binding sequence (MBS) of a plasmid-encoded 16 S rRNA, which binds only to the complementary mutated ribosome binding sequence (RBS) of the CAT message (Burns & Crowl, 1987; Lee *et al.*, 1996). Cells containing pRNA122 are chloramphenicol resistant when transcription of the rRNA operon is induced with IPTG. The level of this resistance depends directly upon the amount of functional CAT protein produced by the plasmid-derived ribosomes. Thus, the effect of rRNA mutations on ribosome function may be assayed *in vivo* by determining the minimal inhibitory concentration (MIC) of chloramphenicol for each mutant strain. In one application of the system, functional mutants are selected from a pool of cells that contain randomized regions of plasmid-encoded ribosomal RNA by simply growing the cells in the presence of chloramphenicol. Sequence and functional analysis of the selected mutants then reveals nucleotides that play a direct role in function (invariant) and those that play an indirect, structural role. This approach to differentiate functional from structural nucleotides is termed instant evolution (Lee *et al.*, 1997).

Sequence analysis of functional mutants

We showed that by random mutation of each of the loop nucleotides in the highly conserved 790 loop of 16 S RNA, alternative functional sequences could be selected, and that analysis of these alternative sequences revealed functionally important nucleotide interactions within the loop (Lee *et al.*, 1997). Here, we performed an instant-evolution experiment on the 690 loop in which we randomized the stem and neighboring nucleotides to maximize the number of alternative, functional sequences/structures. Nine positions (688 to 691 and 697 to 701) were randomly mutated in the 690 stem region (Figure 1). Equal distribution of the four nucleotides at each randomized position was confirmed by sequence analysis of the plasmid library prior to selection. Functional mutants were selected in *E. coli* DH5 cells, sequenced, and assayed for ribosome function. A total of 2.3 million transformants were obtained to ensure that all of the 262,144 mutant sequences were

represented in the original pool with >99.9% confidence (Clarke & Carbon, 1976). Approximately one out of 1000 plated cells survived on 100 µg/ml chloramphenicol selection media. A total of 100 chloramphenicol-resistant functional mutants were randomly selected and sequenced. Of these, 96 unique mutant sequences were obtained and analyzed (Table 2). Four duplicate sequences were excluded. Of the 96 functional mutants, 82 contain four to seven substitutions out of the nine mutated nucleotides, and 24 of these mutants retain 50% or greater the wild-type activity (MIC assay). The MIC of chloramphenicol for cells expressing wild-type rRNA from pRNA122 plasmid is 700 µg/ml. The MICs of the mutants range from 150 to 650 µg/ml. The mean and median MIC are both 300 µg/ml.

Survivors of the instant-evolution experiment show non-random nucleotide distribution at all mutated positions, except position 690, for which no nucleotide preference was observed (Table 1). The strongest nucleotide preferences are found at positions 691 (χ^2 , $P = 2.4 \times 10^{-22}$), 697 (χ^2 , $P = 3.9 \times 10^{-10}$), and 700 (χ^2 , $P = 3.2 \times 10^{-10}$). At position 697, G was strongly selected against.

The mean activities (MICs) of all mutants at each position were compared by single-factor analysis of variance (ANOVA) to determine if nucleotide identity at a given position affected ribosome function. Although specific nucleotide preferences were observed at several of the mutated positions, no significant functional differences were observed among the mutants at each position, suggesting that nucleotide identity is not required for function. These data are consistent with the current paradigm that the nucleotides mutated in this experiment serve primarily a structural role.

A comparison of the nucleotide distribution of the instant-evolution mutants with the phylogenetic data is shown in Table 1. The selected mutant sequences show significantly more variability than those found in nature. This may be because all of the positions in the stem region were simultaneously mutated, allowing normally deleterious mutations in one position to be compensated for by other mutations at different positions, a process which is unlikely to occur in nature (Lee *et al.*, 1997). In addition, wild-type sequences have been selected for optimal ribosome activity, while all of the isolated functional mutants had lower level than wild-type ribosome activity. Nucleotide variability at a given position suggests that the wild-type nucleotide at this position is not directly involved in ribosome function but, in conjunction with co-variation analysis, provides important information about potential pairing interactions.

In genetic complementation experiments, it is often observed that mutation of interacting ligands results in loss of function, even in pairs that retain some ability to interact. To identify functionally significant nucleotide co-variations within the mutated stem region, we first determined a chloramphenicol concentration that would eliminate

Table 2. Sequence and MIC of functional mutants

MIC ^a µg/ ml	Nucleotide sequence ^b										Number of mutations
	688	689	690	691	692-696	697	698	699	700	701	
700 ^c	G	C	G	G	U G A A A	U	G	C	G	U	WT
650	C	A	A	G		U	U	G	G	U	5
550	C	A	A	G		U	U	G	G	U	5
500	G	C	U	U		A	C	C	C	A	3
500	G	C	U	U		A	C	C	C	A	4
450	C	C	A	G		C	C	G	G	C	7
450	C	C	A	G		C	C	G	G	C	5
450	G	A	C	G		A	C	C	G	C	4
450	C	C	C	G		A	C	C	G	C	7
400	G	C	A	G		U	A	C	C	U	4
400	A	A	A	G		A	C	C	A	U	5
400	C	A	A	G		A	C	C	G	U	6
400	C	A	A	G		A	C	C	G	A	6
350	C	C	C	G		U	G	G	A	U	4
350	C	C	C	G		U	G	G	A	U	4
350	G	C	C	G		U	G	G	A	U	3
350	G	C	C	G		A	U	C	C	A	4
350	G	C	C	G		U	G	C	C	A	4
350	C	C	C	G		U	G	C	C	A	5
350	C	C	C	G		U	G	C	C	A	2
350	A	C	C	A		U	G	C	C	U	5
350	A	C	C	A		U	G	C	C	U	4
350	A	C	C	A		U	G	C	C	U	7
350	C	A	C	A		U	G	C	C	U	5
350	C	A	C	A		U	G	C	C	U	6
350	C	A	C	A		U	G	C	C	U	5
350	A	C	C	A		U	G	C	C	U	7
350	A	C	C	A		U	G	C	C	U	8
350	C	C	C	G		U	G	C	C	A	3
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	3
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	8
300	A	A	A	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	3
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	3
300	C	C	C	G		U	G	C	C	A	6
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5
300	C	C	C	G		U	G	C	C	A	7
300	C	C	C	G		U	G	C	C	A	4
300	C	C	C	G		U	G	C	C	A	5

250	G	U	A	G	C	C	C	U	U	5
250	<u>G</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>C</u>	<u>U</u>	<u>C</u>	4
250	<u>G</u>	<u>A</u>	<u>C</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>U</u>	5
250	<u>C</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>A</u>	7
250	<u>C</u>	<u>A</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>A</u>	<u>U</u>	<u>G</u>	<u>C</u>	6
250	<u>G</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>U</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>C</u>	5
250	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>U</u>	<u>A</u>	<u>G</u>	<u>G</u>	<u>A</u>	4
250	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>A</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>A</u>	5
250	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	<u>U</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>A</u>	7
250	<u>U</u>	<u>U</u>	<u>C</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>A</u>	<u>U</u>	<u>A</u>	6
250	<u>A</u>	<u>G</u>	<u>A</u>	<u>U</u>	<u>A</u>	<u>U</u>	<u>U</u>	<u>G</u>	<u>A</u>	8
250	<u>A</u>	<u>G</u>	<u>A</u>	<u>U</u>	<u>A</u>	<u>U</u>	<u>U</u>	<u>A</u>	<u>A</u>	6
200	<u>C</u>	<u>C</u>	<u>C</u>	<u>U</u>	<u>A</u>	<u>U</u>	<u>G</u>	<u>U</u>	<u>A</u>	4
200	<u>C</u>	<u>A</u>	<u>C</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>G</u>	7
200	<u>U</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>A</u>	<u>A</u>	<u>U</u>	<u>C</u>	7
200	<u>C</u>	<u>C</u>	<u>G</u>	<u>G</u>	<u>U</u>	<u>U</u>	<u>G</u>	<u>U</u>	<u>C</u>	5
200	<u>C</u>	<u>G</u>	<u>G</u>	<u>U</u>	<u>U</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>C</u>	7
200	<u>U</u>	<u>U</u>	<u>U</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>A</u>	<u>C</u>	<u>C</u>	8
200	<u>C</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>C</u>	<u>U</u>	<u>G</u>	<u>A</u>	<u>C</u>	8
200	<u>C</u>	<u>A</u>	<u>C</u>	<u>G</u>	<u>U</u>	<u>U</u>	<u>G</u>	<u>A</u>	<u>C</u>	7
200	<u>A</u>	<u>C</u>	<u>A</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>A</u>	<u>C</u>	6
150	<u>U</u>	<u>C</u>	<u>U</u>	<u>G</u>	<u>A</u>	<u>U</u>	<u>A</u>	<u>C</u>	<u>U</u>	5
150	<u>G</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>U</u>	3
150	<u>G</u>	<u>G</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>C</u>	<u>C</u>	7
150	<u>G</u>	<u>G</u>	<u>U</u>	<u>A</u>	<u>U</u>	<u>C</u>	<u>U</u>	<u>C</u>	<u>U</u>	6
150	<u>A</u>	<u>A</u>	<u>A</u>	<u>G</u>	<u>U</u>	<u>C</u>	<u>U</u>	<u>A</u>	<u>U</u>	6
150	<u>C</u>	<u>A</u>	<u>C</u>	<u>G</u>	<u>C</u>	<u>C</u>	<u>G</u>	<u>A</u>	<u>U</u>	7

^a Sequences are ranked by MIC of chloramphenicol.

^b Functional mutants selected at 100 µg/ml chloramphenicol as described in the text. Mutations are underlined.

^c MIC and sequence of the unmutated control, pRNA122, (WT, wild-type).

99.9% of the random pool (not shown) and then examined the distribution between all pairs of nucleotides for the surviving mutants at two levels of function. One set of analyses was performed on all 96 of the mutants that survived selection at 100 µg/ml chloramphenicol. The second set of analyses was performed on a 51-member subset of this pool composed of mutants with an MIC \geq 300 µg/ml chloramphenicol. The co-variation data are summarized in Figure 2. The two analyses produced similar results, although probability values for the more functional subset (Figure 2, red arrows) were reduced because of the smaller sample size. In addition, the marginally significant co-variation observed between 689 and 700, and between 697 and 700 in the total selected pool were absent in the analysis of the more functional mutants. It is possible that interactions between these loci are only functionally important if other interactions within the randomized sequence are disrupted. Among the entire pool of selected mutants, the most significant covariations (Figure 2) were observed between nucleotides at positions 688:699 (χ^2 , $P = 7 \times 10^{-42}$) and 689:698 (χ^2 , $P = 4 \times 10^{-9}$), indicating the functional requirement for a two-base-pair stem as proposed in the phylogenetically derived secondary structure (Figure 1(a)). Examination of the phylogenetic data reveals that 99.7% of known prokaryotic sequences have a stem with two canonical base-pairs or, in rare occasions, one canonical and one G·U wobble base-pair (Van de Peer *et al.*, 1999). In our studies, we found that formation of only one base-pair between positions 688 and 699 is required for ribo-

some function and formation of a second base-pair between positions 689 and 698 is not absolutely required for function. The level of ribosome function, however, is correlated with thermodynamic stability of the 689:698 pair (see below). Of the 96 selected functional mutants, 95 have a canonical or G·U base-pair between positions 688 and 699, but only 67% of the functional mutants have a canonical or G·U base-pair between positions 689 and 698. In the subset of 51 higher functional mutants these numbers were 100% for the 688:699 pair and 74.5% for the 689:698 pair. Thus, our instant-evolution data strongly support the phylogenetically derived secondary structure (Figure 1(a)) as the functionally active conformation of the 690 region. The only evidence for the alternative structure (Figure 1(b)) was a marginally significant covariation between positions 689 and 700 (χ^2 , $P = 0.04$). Even though our genetic data did not support a requirement for the alternative conformation *in vivo*, we did not rule out the possibility that it could form *in vitro*, which would be consistent with the *in vitro* chemical modification results (Kean & Draper, 1985; Serganov *et al.*, 1996). Formation of this alternative secondary structure, however, was not supported by our NMR analysis of a model of the 690 region (see wt31mer, below). Significant covariations observed between mutated loop nucleotides adjacent to the stem, between positions 690 and 691 (χ^2 , $P = 0.0004$), between positions 690 and 697 (χ^2 , $P = 0.0001$), between positions 691 and 697 (χ^2 , $P = 0.008$), and between positions 697 and 700 (χ^2 , $P = 0.025$) are most likely due to important stacking or other unusual

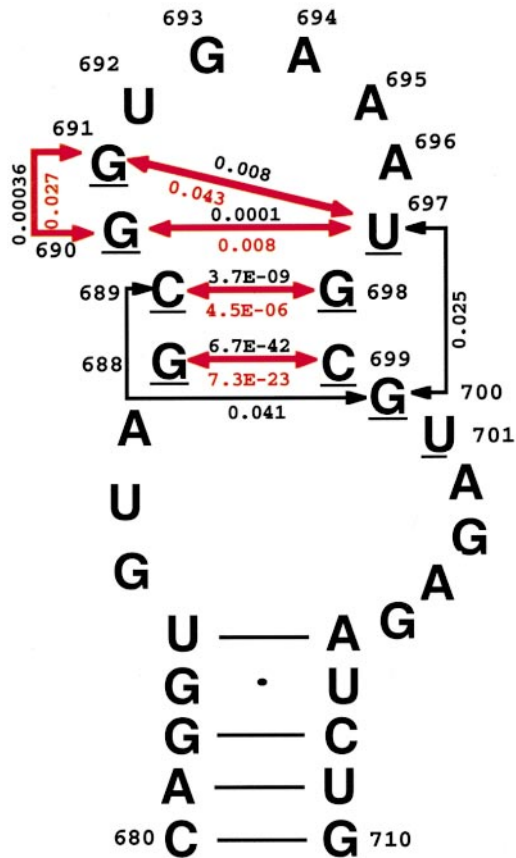


Figure 2. Covariation analysis of the functional mutants. Black χ^2 , P -values are the covariation data derived from analysis of the entire pool of 96 mutants selected at 100 $\mu\text{g}/\text{ml}$ chloramphenicol. Red χ^2 , P -values are derived from analysis of a subset of the selected pool with MICs ≥ 300 $\mu\text{g}/\text{ml}$ chloramphenicol. Black arrows show nucleotide pairs that co-vary only in the selected pool of 96 mutants. Red arrows show nucleotide pairs that co-vary in both the total pool of selected mutants and in the subset of highly functional mutants. Nucleotides mutated in this study are underlined.

base interactions. Position 701 showed no covariation with any of the mutated positions.

The closing nucleotides G690 and U697 are highly conserved and are typically depicted as forming a mismatch pair in the phylogenetically derived secondary structure of *E. coli* 16 S rRNA (Gutell, 1994). Sequence analysis of the instant-evolution mutants revealed that, 78 of the 96 functional mutants have the potential to form a mismatch between positions 690 and 697. Of these 78 sequences, the most frequently found mismatch was the wild-type G·U (22 occurrences). The other 18 sequences all have the potential to form a 690:697 Watson-Crick pair. These sequences included nine A:U pairs, eight U:A pairs, and one C:G pair (MIC 250 $\mu\text{g}/\text{ml}$). No G690:C697 mutant was isolated. These data suggest that strong base-

pairing between these positions inhibits ribosome function. The most frequent substitution for the wild-type G690:U697 pair was the C690:A697 (19 occurrences). This substitution is predicted to cause a reversal of the helical twist of the wobble pair and thus cause a significant change in the local structure (Gautheret *et al.*, 1995). This would imply that G690 and U697 do not form a wobble pair and that the significant covariation observed between these positions (χ^2 , $P = 0.0001$) is a functional requirement for base-pairing other than a wobble conformation. The covariation between G690 and G691 also suggests important stacking interactions. Our NMR experiments on the model 14mer hairpin (see NMR analysis section) also do not support the existence of a G690:U697 wobble pair, but do suggest that stacking plays a critical role in the formation of the functionally active conformation(s). These data, however, do not exclude the possibility of a non-wobble G690:U697 pair.

Thermodynamic stability of the 690 stem

The instant-evolution data indicate that ribosome function is strongly dependent upon formation of a base-pair between positions 688 and 699. Formation of a second base-pair in the stem between positions 689 and 698 is not required for ribosome function even though a significant covariation is observed between positions 689 and 699. One third of the functional mutants have the potential to form a mismatch between position 689 and 698, though a U:U mismatch was never observed. Mutants with the potential to form two base-pairs in the stem, however, are more functional than those with one base-pair and a mismatch. Of 96 functional mutants, 64 can form a two-base-pair stem with an average MIC of 310 $\mu\text{g}/\text{ml}$, and 32 mutants can form a base-pair and a mismatch in the stem with an average MIC of 270 $\mu\text{g}/\text{ml}$. Base-pairs and mismatches are known to make different thermodynamic contributions to RNA helix stability (Kierzek *et al.*, 1999; SantaLucia *et al.*, 1991; Xia *et al.*, 1998). However, the thermodynamic stability of a one or two base-pair stem is strongly affected by the terminal mismatches from the hairpin and internal loops (Serra *et al.*, 1993), and cannot currently be accurately predicted.

The 688:699 base-pair

To investigate the effect of thermodynamic stability of the two-base-pair stem on ribosome function, we examined the relationship between ribosome function in the mutants and the predicted thermodynamic stability of the stem in both sets of mutants. The dependence of ribosome function (MIC) upon identity of the 688:699 base-pair is shown in Figure 3. The observed trend in activity (G:C > C:G > A:U > G:U \geq U:A = A:C) parallels the predicted thermodynamic stability for each pair (SantaLucia *et al.*, 1991). This suggests that the level of function of the 690 stem is thermo-

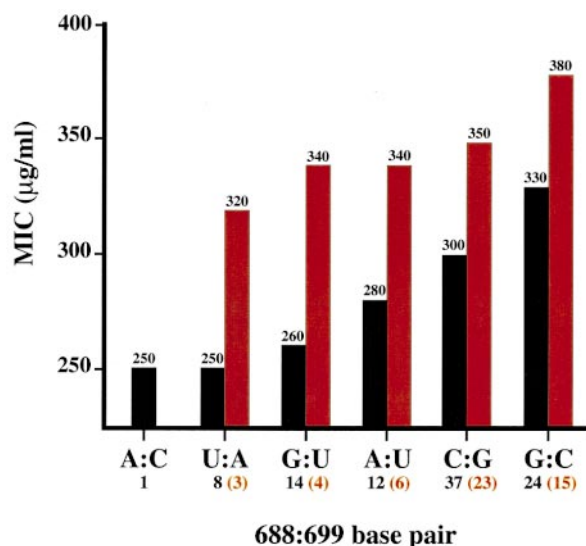


Figure 3. Effect of base-pair stability at positions 688:699 on ribosome function. The 690 mutants were grouped according to the nucleotides at positions 688 and 699 and the mean MIC for each group was determined. Black bars show the mean activity of mutants in the selected pool of 96 mutants. Red bars show the mean activity of mutants in the subset of 51 highly functional mutants as described in the text. The mean MIC in $\mu\text{g/ml}$ is indicated for each group at the top of the bar. The number of mutants in each group are shown below the indicated nucleotide pairs. The number of mutants from the highly-functional subset is in parentheses.

dynamically controlled. A remarkably similar trend was also observed for the closing pair of a 9 nt hairpin loop (Serra *et al.*, 1997). In a similar way in 16 S rRNA the 688:699 base-pair is the closing base-pair of the 9 nt internal loop and its effect on internal loop stability is comparable to the closing base-pair effect reported by Serra *et al.*, (1997) on the 9 nt hairpin loop stability. These data suggest that the role of the 688:699 base-pair separating the internal and hairpin loops is to stabilize the internal loop and to facilitate folding of the conserved hairpin loop.

Our statistical analyses are specifically designed to identify non-random sequence motifs within a pool of other mutations. To determine if base-pair formation at positions 688 to 699 correlated with nucleotide identity at any of the other mutated positions, we performed additional statistical analyses (not shown). These data clearly indicate that base-pair formation at these sites is independent of the presence of nucleotide identity at other positions ($P > 0.15$ in every case).

The 689:698 base-pair

The effect of mutations at the 689:698 pair on ribosome function cannot be entirely explained by the requirement to form a stable base-pair. This is

presumably because the closing base-pair effects both the stability of the stem and the conformation of the hairpin loop. In the instant-evolution survivors, the closing pair must stabilize the stem and compensate for distortions in the loop structure caused by mutations made at the base of the hairpin. Therefore, more variability is observed in the 689:698 pair. The mean MIC of all 96 selected mutants was $300 \mu\text{g/ml}$ and $340 \mu\text{g/ml}$ for the subset of 51 mutants with $\text{MIC} \geq 300$. As shown in Figure 4, mutants with activities below the mean contain nucleotide combinations at positions 689 and 698 that can only form unstable mismatch pairs (Kierzek *et al.*, 1999; Morse & Draper, 1995). Mutants with activities above or at the mean have nucleotides at positions 689 and 698 that are able to form stable pairs. A similar pattern was observed when only the subset of highly functional mutants was analyzed (Figure 4). It is interesting that the normally destabilizing mismatch pairs G:A and A:G are in the group of mutants with activities above the mean. These mismatches, however, are known to stabilize hairpins when present at the base of a loop by forming a stable network of hydrogen bonds (SantaLucia *et al.*, 1992; Serra *et al.*, 1994).

Secondary structure analysis

To examine the potential effect of mutations on folding of the hairpin, the minimum-energy secondary structure of each mutant was predicted using MFOLD (Mathews *et al.*, 1999; Serra & Turner, 1995; Zuker, 1989). Of the 96 unique mutant sequences, 74 were predicted to form minimum-energy secondary structures that violate the requirement for base-pairing between positions 688 and 699 revealed in our genetic analysis. It is possible that for some of the less functional mutants, decreased function could be due to formation of non-functional secondary structures. For instance, all six mutants with the lowest ribosome activity ($\text{MIC } 150 \mu\text{g/ml}$) were predicted to form very stable structures that do not contain the required 688:699 base-pair. On the other hand, formation of a minimum energy structure with a 688:699 base-pair was predicted for all four mutants with $\text{MIC} \geq 500 \mu\text{g/ml}$. Folding of the wild-type sequence revealed that the minimum-energy structure contained the 688:699 base-pair. Although most of the mutants we isolated did not contain a 688:699 base-pair in the minimum energy secondary structure predicted by MFOLD, our genetic data suggest that the 688:699 base-pair must form at some stage of protein synthesis. One possible explanation is that base-pairing between positions 688 and 699 is facilitated by another ligand (i.e. protein or other RNA sequence). An electron density map of the 70 S ribosome showed that the bridge between the two subunits containing the 690 loop is RNA-rich, and it was proposed that the 690 helix makes minor-groove contacts to an RNA element of the 50 S subunit (Cate *et al.*, 1999).

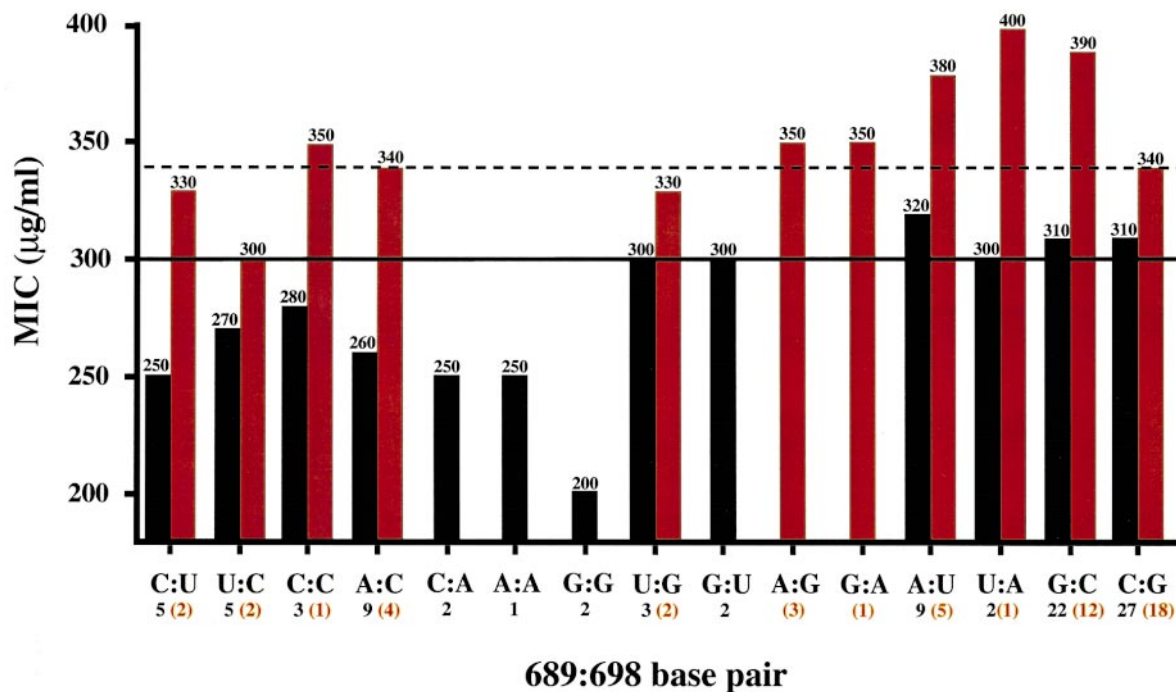


Figure 4. Effect of base-pair stability at positions 689:698 on ribosome function. The 690 mutants were grouped according to the nucleotides at positions 689 and 698 and the mean MIC for each group was determined. Black bars show the mean activity of mutants in the selected pool of 96 mutants. Red bars show the mean activity of mutants in the subset of 51 highly functional mutants as described in the text. The mean MIC in $\mu\text{g/ml}$ is indicated for each group at the top of the bar. The number of mutants in each group are shown below the indicated nucleotide pairs. The number of mutants from the highly functional subset is in parentheses. The sequences are listed from left to right in approximate order of predicted stability (SantaLucia *et al.*, 1991). The horizontal line at MIC 300 $\mu\text{g/ml}$ is the mean MIC value for all mutants and the dashed line at MIC 340 $\mu\text{g/ml}$ is the mean MIC for the subset of highly-functional mutants. The mismatch U:U was never observed.

NMR structural analysis

To investigate base-pairing in the wild-type and functional mutants, imino-proton NMR studies of model RNAs were performed. Figure 5 shows four oligonucleotide RNA models of the 690 region that were designed to characterize the secondary structure of the wild-type sequence and to identify which of the two proposed conformations of the wild-type sequence (Figure 1) exists *in vitro*. A 31-nt molecule, wt31mer, corresponds to nucleotides 680 to 710 of *E. coli* 16 S RNA. The 14 nucleotide hairpin, 14mer, was designed to “lock” the hairpin into the phylogenetic conformation (Gutell, 1994). The 16-nucleotide hairpin, 16mer, was designed to lock the hairpin into the alternative conformation, based on chemical modification studies (Kean & Draper, 1985; Serganov *et al.*, 1996). The 13 nt hairpin, 13mer, is an unaltered sequence of the hairpin-stem portion of the *E. coli* 690 region (nucleotides 688 to 700). The 13mer has the potential to form both phylogenetic, 13merA, and alternative, 13merB, conformations (Figure 5). 1D imino-proton NMR spectra of the wt31mer, 14mer, 13merA/B and 16mer hairpins are shown in Figure 6. The imino protons of the 14mer and

wt31mer were assigned using 1D NOE difference spectroscopy and 2D H₂O NOESY. Figure 7 shows the imino proton region of the 14mer 2D H₂O NOESY spectrum. The wt31mer and 14mer have nearly identical imino proton chemical shifts of the hairpin residues between 10-12 ppm, indicating that both molecules form the same hairpin structure. The spectrum of the 13mer hairpin has chemical shifts characteristic of the hairpin loop structure of the 14mer and wt31mer. Therefore, the 13mer appears predominantly to form the phylogenetically predicted structure, 13merA. The broad resonance at 10.65 ppm of the 13mer is from G700, which is consistent with previous data for a 3' dangling-end guanine residue (SantaLucia *et al.*, 1990). The 13mer resonances were significantly broadened at temperatures higher than 15 °C due to instability of the 13mer hairpin. Thus, the absence of the resonances in the 13mer spectrum in the 10-12 ppm region characteristic of the conformation B (cf. 16mer spectrum in Figure 6) indicates that formation of the alternative conformation, 13merB, is not thermodynamically favorable. In addition, the 1D NOE difference spectra (not shown) and 2D NOESY in H₂O (Figure 7) of the 14mer hairpin revealed that G690 and U697 do not form a wobble

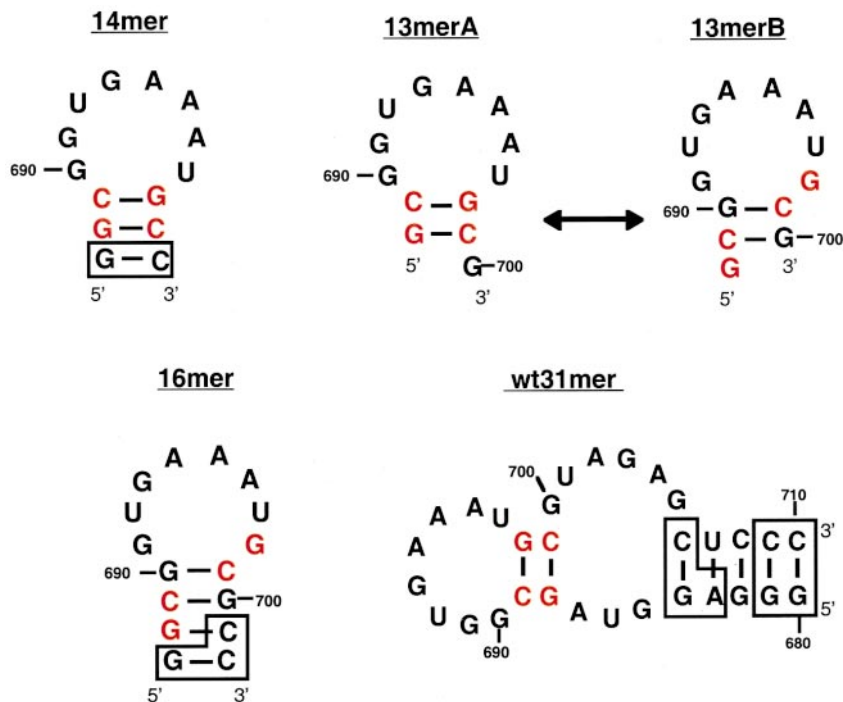


Figure 5. Model RNA hairpins used in NMR studies. Some of the stem nucleotides were modified (boxed) to stabilize NMR constructs. An extra G:C base-pair was added to the stems of the 14mer and the 16mer hairpins to lock hairpin loops in functional and alternative conformations, respectively. Stem nucleotides mutated in this study are in red to show the shift in base-pairing for the alternative conformation, 16mer and 13merB.

pair, which is consistent with the genetic data. However, the NMR data do not preclude the existence of a G690-U697 pair with non-wobble type hydrogen bonding. NMR temperature dependence

studies of the 14mer and wt31mer RNA (data not shown) showed that the resonances between 10.0 and 11.75 ppm, which characterize the loop, gradually disappear with increasing temperature, and

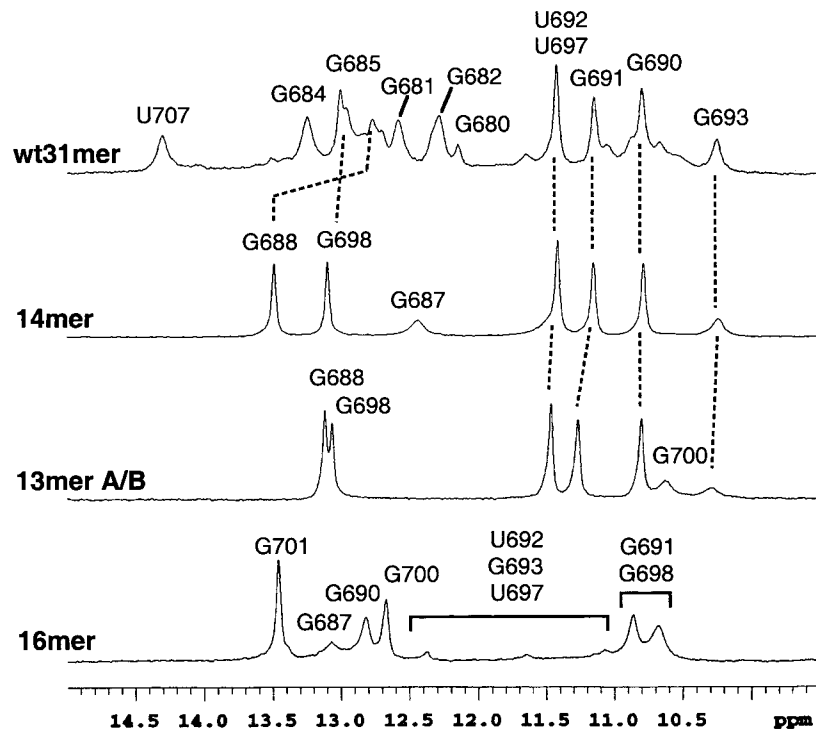


Figure 6. Imino proton region of the 1D NMR spectra of wt31mer, 14mer, 13merA/B and 16mer hairpins. Signature loop imino proton resonances appear between 10-12 ppm. The similarity of the 31mer and 14mer loop chemical shifts indicates that both models have identical hairpin loop structures. Thus, the wt31mer *in vitro* secondary structure is consistent with the genetically derived structure (Figure 1(a)). The 13mer also forms the phylogenetically predicted structure.

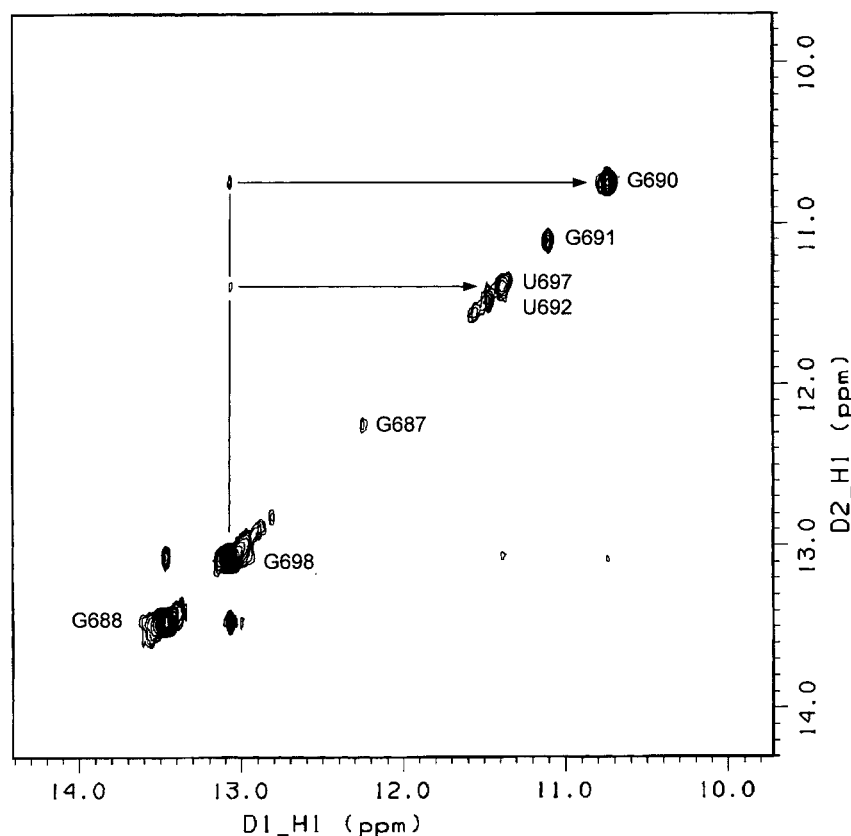


Figure 7. Imino proton region of the 2D H_2O NOESY of the 14mer hairpin. Note the weak NOEs (indicated by lines connecting resonances) from G698 to G690 and U697 and note the absence of an NOE between G690 and U697. These data indicate that G690 and U697 do not form a wobble pair. G693 at 10.3 ppm, which is observed in the 1D NMR spectrum (Figure 6) is not observed here due to solvent exchange. The imino protons of G691 and G693 were assigned based upon key NOEs to non-exchangeable protons that were definitively assigned by 2D H_2O NOESY of the 14mer (data not shown).

exchange with water is nearly complete at 20 °C. These data suggest that pairing between G690 and U697 does not involve their imino protons, but instead that solvent protection is due to stacking with the G689-C698 pair and the other loop nucleotides. Thus, NMR analysis of the wild-type 31mer sequence is in agreement with the phylogenetic covariation, our instant-evolution experiments, and the minimum energy structure predicted by MFOLD.

Next, we investigated the structure formed by one partially functional mutant *in vitro*. A model RNA sequence of one of the partially functional mutants, mut31mer, was synthesized and studied by NMR spectroscopy. Figure 8 shows the mut31mer in both the genetically derived structure (a) and the MFOLD-predicted minimum-energy secondary structure (b). To test the hypothesis that the genetic structure is stabilized by hairpin loop and internal loop tertiary interactions unaccounted for in the MFOLD energy minimization calculations, we selected a mutant sequence that did not have any mutations in the hairpin loop sequence and had only one mutation U701C in the internal loop. Thus, if the mutant sequence forms the functional structure (Figure 8(a)) we would expect to observe the signature imino protons between 9-12 ppm, characteristic of the 14mer and wt31mer loops. Note that the U701C substitution is one of the most frequently found in naturally occurring sequences (Table 1) and thus the U701C mutation is unlikely to have a great effect on function. The

selected sequence has 42% of the wild-type activity, MIC = 300 $\mu\text{g}/\text{ml}$. The imino proton region (a) and 1D NOE difference spectra (b) and (c) of the mutant sequence are shown in Figure 9. The strong NOEs observed from the G690 imino proton (11.2 ppm) to the U698 imino proton (11.5 ppm), and *vice versa* indicate formation of a wobble structure between these two nucleotides. A weak NOE at 13.1 ppm indicates that the G·U wobble pair is located next to the G·C pair in the stem region. These data are consistent with the alternative secondary structure predicted by MFOLD (Figure 8(b)). NMR temperature dependence studies of the mut31mer RNA *versus* wt31mer RNA also showed differences in the behavior of these structures (data not shown). Resonances between 10.0 and 11.75 ppm, which characterize the hairpin loop portion of the wt31mer, gradually disappear with increasing temperature and exchange with water is nearly complete at 20 °C for the wild-type sequence. On the other hand, resonances corresponding to the wobble G·U pair of the mutant sequence retained full intensity at 20 °C (not shown). Thus, the minimum-energy secondary structure predicted by MFOLD for the wild-type and the functional mutant sequences are consistent with our *in vitro* NMR data. The NMR data suggest that the C688-U698-G699-C701 mutant has reduced function *in vivo* because there is a thermodynamic penalty for folding the mutant into a functional secondary structure that contains the 688:699 pair.

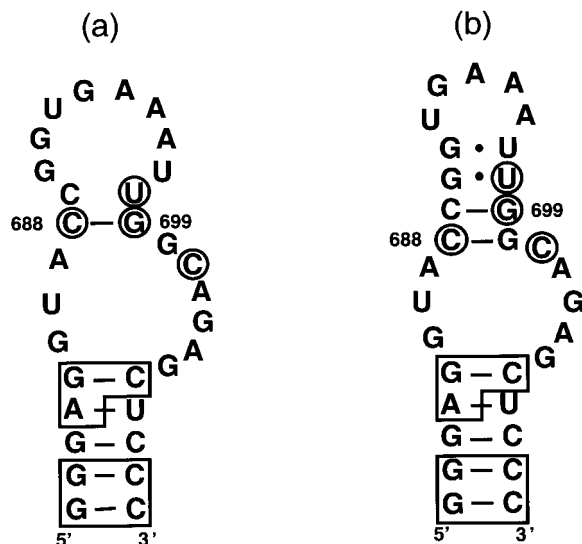


Figure 8. 31mer RNA NMR construct of a partially functional sequence with mutations at C688, U698, G699 and C701. Mutated positions are circled. (a) Genetically predicted structure of the functional mutant with the 688-699 base-pair requirement. Boxed nucleotides were substituted for NMR studies. (b) Minimum-energy predicted secondary structure. NMR spectroscopy data are consistent with the predicted minimum energy structure. The results suggest that the partial loss of function is due to the energy cost required to rearrange the structure to form a 688-699 base-pair.

Conclusions

By random mutagenesis of the stem and adjacent nucleotides and selection of functional mutants, we identified base-pairing interactions in the 690 region that are important for ribosome function *in vivo*. The data unambiguously demonstrate that highly active ribosomes require the formation of a stable 688:699 base-pair (Figure 1(a)). NMR studies were fully consistent with the *in vivo* results and showed that the wild-type sequence forms a 688:699 pair, but a partially functional mutant formed a secondary structure without the 688:699 pair. Since all of the mutants were strongly selected for formation of a 688:699 base-pair, this suggests that formation of the correct secondary structure may be stabilized by specific interaction of the 690 region with another protein or RNA ligand. Since we observed considerable nucleotide variability at all of the mutated positions except 697, it is unlikely that any of the nucleotides at positions 688 to 691 and 698 to 701 have specific interactions with IF3, tRNA, rRNA, or ribosomal proteins. Conserved hairpin loop and internal loop nucleotides are the primary candidates for interaction with the ligand. In the X-ray crystal structure of the 70 S ribosome, the 690 helix was identified to be on the interface with the 50 S subunit and to form one of the bridges with the 50 S subunit (Cate *et al.*, 1999). On the other hand, none of the functional structures, including the wild-type, are exceptionally stable, which

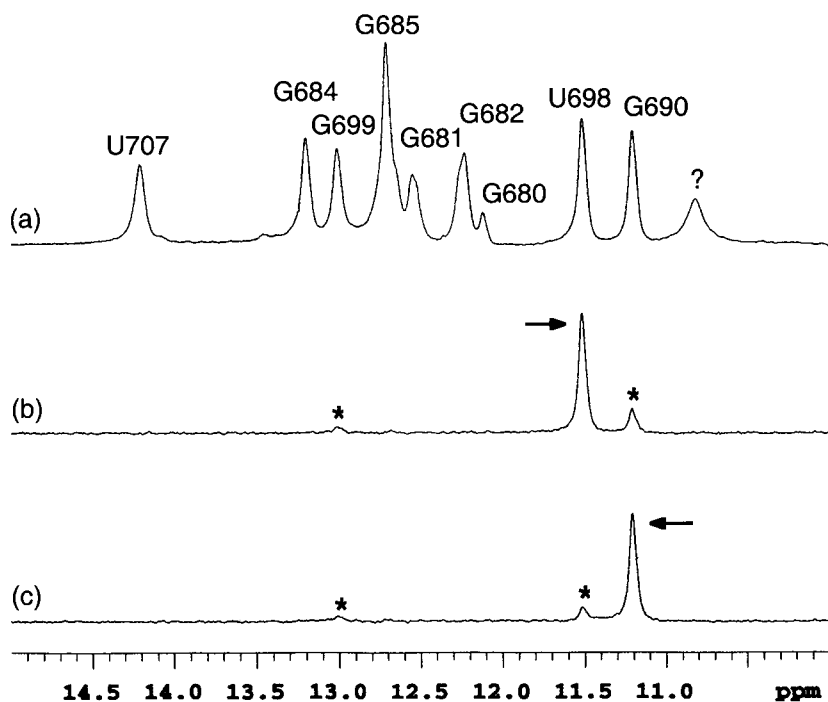


Figure 9. (a) The imino proton region of the 1D NMR spectrum of the mut31mer. (b) and (c) 1D NOE difference spectra of the mutant 31mer shown in Figure 8. The irradiated resonances are indicated by arrows and NOEs by asterisks. Note the intense NOEs, indicative of wobble G·U structure between U698-G690.

suggests that there is a need for conformational flexibility in the 690 region.

Materials and Methods

Reagents

Restriction enzymes, ligase and calf intestine alkaline phosphatase, were from New England Biolabs and from Gibco-BRL. DNA sequencing reactions were performed using the SequiTherm EXCEL[™] II DNA Sequencing Kits from Epicentre and processed with a Li-Cor Global IR² automated DNA sequencer. Oligonucleotides were either purchased from Midland Certified Reagent Company (Midland, TX) or synthesized on-site using a Beckman Oligo 1000 DNA synthesizer. Amplitaq DNA polymerase and PCR reagents were from Perkin-Elmer-Cetus.

Plasmids

Construction of pRNA122 (Lee *et al.*, 1997) and pRNA16 ST will be described elsewhere. The key features of pRNA122 are: (i) it contains a copy of the *rrnB* operon from pKK3535 (Brosius *et al.*, 1981) under transcriptional regulation of the *lacUV5* promoter; (ii) it contains a copy of the lactose repressor allele *lacI^q* (Calos, 1978); (iii) the chloramphenicol acetyltransferase gene is present and transcribed constitutively from a mutant tryptophan promoter, *trp^c* (de Boer *et al.*, 1983; Hui, 1987); (iv) the RBS of the CAT message has been changed from the wild-type, 5'-GGAGG to 5'-AUGCC, and the MBS of the 16 S rRNA gene has been changed to 5'-GGGAU; and (v) the β -lactamase gene is present to allow maintenance of plasmids in the host strain. pRNA16 ST is similar to pRNA122, but the 23 S and 5 S rRNA genes have been deleted.

Bacterial strains and media

All plasmids were maintained and expressed in *E. coli* DH5 (*supE44*, *hsdR17*, *recA1*, *endA1*, *gyrA96*, *thi-1*, *relA1*) (Hanahan, 1983). Cultures were grown in LB medium (Luria & Burrous, 1957) or LB medium containing 100 μ g/ml ampicillin (LB-Amp100). To induce synthesis of plasmid-derived rRNA from the *lacUV5* promoter, IPTG was added to a final concentration of 1 mM. Strains were transformed by electroporation (Dower *et al.*, 1988) using a Gibco-BRL Cell Electroporator. Transformants were grown in SOC medium (Hanahan, 1983) for one hour prior to plating on selective medium to allow expression of plasmid-encoded genes.

Random mutagenesis and selection

All PCR-directed mutagenesis experiments were performed essentially by the method described by Higuchi (1989). The PCR products were first cloned in pRNA16 ST using the unique *Bgl*III and *Sac*II restriction sites. The PCR upper primer, 688N-691N + 697N-701N, was 5'-CGGTATTCCTCCAGAT-CTCTNNNNNTTCA NNNNTACACCTGGAATTCTA-3' (N = A, T, C, and G) and the lower primer, 16 S-*Avr*II, was 5'-ACGTCGCAA-GACCAAAGAGG-3'. The upper and lower primers were designed to bind to the *Bgl*III restriction site and outside of the *Sac*II restriction site, respectively. A library of random mutants in pRNA16 ST was then digested

with *Bst*EII and cloned into pRNA122. Transformants were incubated in SOC medium containing 1 mM IPTG for four hours to induce rRNA synthesis and then plated on LB-Amp100 agar with and without 100 μ g/ml chloramphenicol. A total of 2.3×10^6 transformants were obtained, yielding approximately one chloramphenicol-resistant survivor per 1000 plated cells. Next, 96 transformants with MIC \geq 100 μ g/ml were randomly chosen and sequenced to identify mutations in the randomized positions. Each of the mutants was sequenced between the *Bgl*III and *Sac*II ligation sites to verify the absence of unprogrammed mutations. None of the isolates contained unprogrammed mutations.

Minimum inhibitory concentration

MICs were determined in microtiter plates or on solid medium. Overnight cultures grown in LB-Amp100 were diluted and induced in the same medium containing 1 mM IPTG for two to three hours. Approximately 10⁶ induced cells were then added to wells (or spotted onto solid medium) containing LB-Amp100 + IPTG (1 mM) and chloramphenicol at increasing concentrations. Cultures were grown 24 hours and the lowest concentration of chloramphenicol that completely inhibited growth was designated as the MIC.

Oligoribonucleotide synthesis

13mer, 14mer and 16mer oligoribonucleotides were synthesized on solid support with the phosphoramidite method (Capaldi & Reese, 1994) on a Cruachem PS 250 DNA/RNA synthesizer. Oligomers were removed from the solid support and deprotected by treatment with ammonia and acid following the manufacturer's protocol. 31mer oligoribonucleotides (Figures 5 and 8) were synthesized by *in vitro* run-off transcription with T7 RNA polymerase from DNA templates containing an 18 nt double-stranded promoter region (Milligan *et al.*, 1987). After transcription, Mg²⁺-pyrophosphate precipitate was removed by centrifugation, and the RNA was concentrated by ethanol precipitation. After synthesis, all RNA oligomers were purified by electrophoresis on a 20% (w/w) polyacrylamide denaturing gel containing 7 M urea. Bands were visualized with an ultraviolet lamp and the corresponding bands were cut out and eluted by soaking crushed gel pieces in 50 mM Tris-OH, 0.5 mM Na₂EDTA buffer (pH 7.5) for six to eight hours at 37 °C. RNA oligos were further purified by ethanol precipitation, desalted by continuous-flow dialysis (BRL), and lyophilized to dryness. NMR samples were prepared by dissolving RNA oligomers in 290 μ l of buffer containing 40 mM NaCl, 10 mM disodium phosphate, and 0.5 mM Na₂EDTA in 10% ²H₂O, 90% H₂O at pH 6.8 and treated with Chelex-100 resin (Sigma) to remove residual divalent metal ions. After resin particles were removed by brief centrifugation, samples were transferred into microvolume NMR tubes (Shigemi Inc.). Concentrations of the NMR samples were 2.5 mM for 14mer, 1.8 mM for 13mer, 1.4 mM for 16mer, 1.2 for wt31mer, and 1.1 mM for mut31mer.

NMR methods

Imino proton spectra were acquired at 1 °C or 15 °C on a Varian UNITY 500 MHz. NMR spectrometer using WATERGATE with "flip-back" solvent suppression (Piotto *et al.*, 1992). Spectra were recorded with the car-

rier placed at the solvent frequency and with high-power and low-power pulse widths of 8.5 ms and 1700 ms, sweep width of 12 kHz and gradient field strength of 10.0 G/cm and duration of 1 ms. A total of 512 scans with 8192 points were collected for each spectrum. The data were multiplied by a 4.0 Hz line-broadening exponential function and Fourier transformed with Varian VNMR software. No baseline correction or solvent subtraction was applied. 3-trimethylsilyl propionic-2,2,3,3-d⁴ acid was used as the internal standard for chemical shift reference. The 1D-NOE difference spectra were acquired as described above, but with selective decoupling of individual resonances during the one second recycle delay. Each resonance was decoupled with a power sufficient to saturate <80% of the signal intensity so that spill-over artifacts would be minimized. A total of 4000 scans were collected for each FID in an interleaved fashion in blocks of 16 scans to minimize subtraction errors due to long-term instrumental drift. Assignments of the 14mer and wt31mer were confirmed by recording 2D NOESY with WATERGATE solvent suppression (Figure 7) and by natural abundance ¹H-¹⁵N-gradient-HMQC (data not shown).

Acknowledgments

We thank Dr HyunDae Cho for technical assistance and useful discussions, Dr Christine Chow for help with RNA synthesis and Dr Allen Nicholson for critical review of the manuscript. This work was supported by NIH grants GM55745 and GM52896.

References

- Agalarov, S. C., Prasad, G. S., Funke, P. M., Stout, C. D. & Williamson, J. R. (2000). Structure of the S15, S6, S18-rRNA complex: assembly of the 30 S ribosome central domain. *Science*, **288**, 107-112.
- Brosius, J., Ullrich, A., Raker, M. A., Gray, A., Dull, T. J., Gutell, R. R. & Noller, H. F. (1981). Construction and fine mapping of recombinant plasmids containing the *rrnB* ribosomal RNA operon of *E. coli*. *Plasmid*, **6**, 112-118.
- Burns, D. K. & Crowl, R. M. (1987). *In vitro* mutagenesis of chloramphenicol acetyl transferase to investigate structure/function relationship. *Protein Struct. Fold. Des.* **2**, 375-384.
- Calos, M. P. (1978). DNA sequence for a low-level promoter of the *lac* repressor gene and an "up" promoter mutation. *Nature*, **274**, 762-765.
- Capaldi, D. C. & Reese, C. B. (1994). Use of the 1-(2-fluorophenyl)-4-methoxypiperidin-4-yl (Fpmp) and related protecting groups in oligoribonucleotide synthesis: stability of internucleotide linkages to aqueous acid. *Nucl. Acids Res.* **22**, 2209-2216.
- Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. (1999). X-ray crystal structures of 70 S ribosome functional complexes. *Science*, **285**, 2095-2104.
- Clarke, L. & Carbon, J. (1976). A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. *Cell*, **9**, 91-99.
- Clemons, W. M., Jr., May, J. L., Wimberly, B. T., McCutcheon, J. P., Capel, M. S. & Ramakrishnan, V. (1999). Structure of a bacterial 30 S ribosomal subunit at 5.5 Å resolution. *Nature*, **400**, 833-840.
- Cunningham, P. R., Nurse, K., Weitzmann, C. J. & Ofengand, J. (1993). Functional effects of base changes which further define the decoding center of *Escherichia coli*: 16 S ribosomal RNA: mutation of C1404, G1405, C1496, G1497, and U1498. *Biochemistry*, **32**, 7172-7180.
- Dahlberg, A. E. (1989). The functional role of ribosomal RNA in protein synthesis. *Cell*, **57**, 525-529.
- de Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. *Proc. Natl Acad. Sci. USA*, **80**, 21-25.
- Doring, T., Mitchell, P., Osswald, M., Bochkariov, D. & Brimacombe, R. (1994). The decoding region of 16 S RNA; a cross-linking study of the ribosomal A, P and E sites using tRNA derivatized at position 32 in the anticodon loop. *EMBO J.* **13**, 2677-2685.
- Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucl. Acids Res.* **16**, 6127-6145.
- Egebjerg, J. & Garrett, R. A. (1991). Binding sites of the antibiotics pactamycin and celesticetin on ribosomal RNAs. *Biochimie*, **73**, 1145-1149.
- Gautheret, D., Konings, D. & Gutell, R. R. (1995). G:U base-pairing motifs in ribosomal RNA. *RNA*, **1**, 807-814.
- Gutell, R. R. (1994). Collection of small subunit (16 S and 16 S-like) ribosomal RNA structures: 1994. *Nucl. Acids Res.* **22**, 3502-3507.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557-580.
- Higuchi, R. (1989). Using PCR to engineer DNA. In *PCR Technology* (Erlich, H. A., ed.), pp. 61-70, Stockton Press, New York.
- Hui, A. S. & deBoer, H. A. (1987). Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **84**, 4762-4766.
- Joseph, S., Weiser, B. & Noller, H. F. (1997). Mapping the inside of the ribosome with an RNA helical ruler. *Science*, **278**, 1093-1098.
- Kalurachchi, K. & Nikonowicz, E. P. (1998). NMR structure determination of the binding site for ribosomal protein S8 from *Escherichia coli* 16 S rRNA. *J. Mol. Biol.* **280**, 639-654.
- Kean, J. M. & Draper, D. E. (1985). Secondary structure of a 345-base RNA fragment covering the S8/S15 protein binding domain of *Escherichia coli* 16 S ribosomal RNA. *Biochemistry*, **24**, 5052-5061.
- Kierzek, R., Burkard, M. E. & Turner, D. H. (1999). Thermodynamics of single mismatches in RNA duplexes. *Biochemistry*, **38**, 14214-14223.
- Lee, K., Holland-Staley, C. A. & Cunningham, P. R. (1996). Genetic analysis of the Shine-Dalgarno interaction: selection of alternative functional mRNA-rRNA combinations. *RNA*, **2**, 1270-1285.
- Lee, K., Varma, S., SantaLucia, J., Jr & Cunningham, P. R. (1997). *In vivo* determination of RNA structure-function relationships: analysis of the 790 loop in ribosomal RNA. *J. Mol. Biol.* **269**, 732-743.
- Luria, S. E. & Burrous, J. W. (1957). Hybridization between *Escherichia coli* and *Shigella*. *J. Bacteriol.* **74**, 461-476.
- Mankin, A. S. (1997). Pactamycin resistance mutations in functional sites of 16 S rRNA. *J. Mol. Biol.* **274**, 8-15.
- Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999). Expanded sequence dependence of thermo-

- dynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911-940.
- Merryman, C., Moazed, D., McWhirter, J. & Noller, H. F. (1999). Nucleotides in 16 S rRNA protected by the association of 30 S and 50 S ribosomal subunits. *J. Mol. Biol.* **285**, 97-105.
- Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucl. Acids Res.* **15**, 8783-8798.
- Moazed, D. & Noller, H. F. (1986). Transfer RNA shields specific nucleotides in 16 S ribosomal RNA from attack by chemical probes. *Cell*, **47**, 985-994.
- Moazed, D. & Noller, H. F. (1989). Intermediate states in the movement of transfer RNA in the ribosome. *Nature*, **342**, 142-148.
- Moazed, D. & Noller, H. F. (1990). Binding of tRNA to the ribosomal A and P sites protects two distinct sets of nucleotides in 16 S rRNA. *J. Mol. Biol.* **211**, 135-145.
- Moazed, D., Samaha, R. R., Gualerzi, C. & Noller, H. F. (1995). Specific protection of 16 S rRNA by translational initiation factors. *J. Mol. Biol.* **248**, 207-210.
- Morse, S. E. & Draper, D. E. (1995). Purine-purine mismatches in RNA helices: evidence for protonated G:A pairs and next-nearest neighbor effects. *Nucl. Acids Res.* **23**, 302-306.
- Mueller, F., Stark, H., van Heel, M., Rinke-Appel, J. & Brimacombe, R. (1997). A new model for the three-dimensional folding of *Escherichia coli* 16 S ribosomal RNA. III. The topography of the functional centre. *J. Mol. Biol.* **271**, 566-587.
- Muralikrishna, P. & Wickstrom, E. (1989). *Escherichia coli* initiation factor 3 protein binding to 30 S ribosomal subunits alters the accessibility of nucleotides within the conserved central region of 16 S rRNA. *Biochemistry*, **28**, 7505-7510.
- Noller, H. F. (1991). Ribosomal RNA and translation. *Annu. Rev. Biochem.* **60**, 191-227.
- Noller, H. F., Hoffarth, V. & Zimniak, L. (1992). Unusual resistance of peptidyltransferase to protein extraction procedures. *Science*, **256**, 1416-1419.
- Oehler, R., Polacek, N., Steiner, G. & Barta, A. (1997). Interaction of tetracycline with RNA: photoincorporation into ribosomal RNA of *Escherichia coli*. *Nucl. Acids Res.* **25**, 1219-1224.
- Osswald, M., Doring, T. & Brimacombe, R. (1995). The ribosomal neighborhood of the central fold of tRNA: cross-links from position 47 of tRNA located at the A, P or E site. *Nucl. Acids Res.* **23**, 4635-4641.
- Piotto, M., Saudek, V. & Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, **2**, 661-665.
- Rinke-Appel, J., Junke, N., Osswald, M. & Brimacombe, R. (1995). The ribosomal environment of tRNA: crosslinks to rRNA from positions 8 and 20:1 in the central fold of tRNA located at the A, P, or E site. *RNA*, **1**, 1018-1028.
- SantaLucia, J., Jr, Kierzek, R. & Turner, D. H. (1990). Effects of G:A mismatches on the structure and thermodynamics of RNA internal loops. *Biochemistry*, **29**, 8813-8819.
- SantaLucia, J., Jr, Kierzek, R. & Turner, D. H. (1991). Stabilities of consecutive A:C, C:C, G:G, U:C, and U:U mismatches in RNA internal loops: evidence for stable hydrogen-bonded U:U and C:C+ pairs. *Biochemistry*, **30**, 8242-8251.
- SantaLucia, J., Jr, Kierzek, R. & Turner, D. H. (1992). Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin. *Science*, **256**, 217-219.
- Serganov, A. A., Masquida, B., Westhof, E., Cachia, C., Portier, C., Garber, M., Ehresmann, B. & Ehresmann, C. (1996). The 16 S rRNA binding site of *Thermus thermophilus* ribosomal protein S15: comparison with *Escherichia coli* S15, minimum site and structure. *RNA*, **2**, 1124-1138.
- Serra, M. J. & Turner, D. H. (1995). Predicting thermodynamic properties of RNA. *Methods Enzymol.* **259**, 242-261.
- Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A. & Turner, D. H. (1993). RNA hairpin loop stability depends on closing base-pair. *Nucl. Acids Res.* **21**, 3845-3849.
- Serra, M. J., Axenson, T. J. & Turner, D. H. (1994). A model for the stabilities of RNA hairpins based on a study of the sequence dependence of stability for hairpins of six nucleotides. *Biochemistry*, **33**, 14289-14296.
- Serra, M. J., Barnes, T. W., Betschart, K., Gutierrez, M. J., Sprouse, K. J., Riley, C. K., Stewart, L. & Temel, R. E. (1997). Improved parameters for the prediction of RNA hairpin stability. *Biochemistry*, **36**, 4844-4851.
- Stark, H., Orlova, E. V., Rinke-Appel, J., Junke, N., Mueller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R. & van Heel, M. (1997). Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. *Cell*, **88**, 19-28.
- Van de Peer, Y., Robbrecht, E., de Hoog, S., Caers, A., De Rijk, P. & De Wachter, R. (1999). Database on the structure of small subunit ribosomal RNA. *Nucl. Acids Res.* **27**, 179-183.
- Woodcock, J., Moazed, D., Cannon, M., Davies, J. & Noller, H. F. (1991). Interaction of antibiotics with A- and P-site-specific bases in 16 S ribosomal RNA. *EMBO J.* **10**, 3099-3103.
- Xia, T., SantaLucia, J., Jr, Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C. & Turner, D. H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base-pairs. *Biochemistry*, **37**, 14719-14735.
- Yoshizawa, S., Fourmy, D. & Puglisi, J. D. (1999). Recognition of the codon-anticodon helix by ribosomal RNA. *Science*, **285**, 1722-1725.
- Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48-52.

Edited by D. E. Draper

(Received 21 January 2000; received in revised form 2 May 2000; accepted 6 May 2000)