RNA aptamers to the peptidyl transferase inhibitor chloramphenicol

Donald H Burke[1], David C Hoffman[1,*], Analisa Brown[2], Mark Hansen[2], Arthur Pardi[2] and Larry Gold[1,†]

Background: The problem of how macromolecules adopt specific shapes to recognize small molecules in their environment is readily addressed through *in vitro* selections (the SELEX protocol). RNA–antibiotic interactions are particularly attractive systems for study because they provide an opportunity to expand our understanding of molecular recognition by RNA and to facilitate ribosomal modeling. Specifically, the antibiotic chloramphenicol (Cam) naturally binds bacterial ribosomes in the 'peptidyl transferase loop' of 23S ribosomal RNA to inhibit peptide bond formation.

Results: We identified Cam-binding RNA molecules ('aptamers') from two independent initial random RNA populations. Boundary determinations, ribonuclease S_1 sensitivity analyses and the activity of truncated minimal RNAs identified a structural motif that is shared by sequences from both selections. The pseudosymmetric motif consists of a highly conserved central helix of five to six base pairs flanked by A-rich bulges and additional helices. Addition of Cam prior to ribonuclease S_1 protected nucleotides in the conserved cores from cleavage. Reselection from a pool of mutated variants of the minimal aptamer further refined the sequence requirements for binding. Finally, we used proton nuclear magnetic resonance (NMR) to establish a 1:1 RNA : Cam stoichiometry of the complex. Both the protection and NMR data both show that Cam stabilizes the active fold of this aptamer.

Conclusions: There are many different RNA sequences that can bind Cam. The Cam aptamers that we examined have a well-defined secondary structure with a binding pocket that appears to be stabilized by Cam. This RNA motif superficially resembles the Cam-binding site in 23S rRNA, although further work is needed to establish the significance of these similarities.

Addresses: [1]Department of Molecular, Cellular and Development Biology, and [2]Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0347, USA.

Present addresses: *INSERM U386, Université Victor Segalen Bordeaux 2, 146 Rue Leo Saignat, F-33076 Bordeaux, France. †NeXstar Pharmaceuticals, Inc., 2860 Wilderness Place, Boulder, CO 80301, USA.

Correspondence: Donald H Burke and Larry Gold
E-mail: dhburke@beagle.colorado.edu
lgold@nexstar.com

Key words: chloramphenicol, peptidyl transfer, ribosomal RNA structure, SELEX, translation

Received: **28 April 1997**
Revisions requested: **10 May 1997**
Revisions received: **22 September 1997**
Accepted: **23 September 1997**

Chemistry & Biology November 1997, 4:833–843
http://biomednet.com/elecref/1074552100400833

Introduction

The RNA aptamers isolated through *in vitro* selections (SELEX) [1–3] are exceptionally attractive targets for studies of RNA structure and molecular recognition by nucleic acids [4]. Where the structures of aptamer–target complexes have been determined, they have stretched our understanding of both the RNA folding process and the formation of binding pockets within nucleic acids, as seen in the elegant structures of aptamers to adenosine [5,6], citrulline and arginine [7], theophylline [8] and thrombin [9]. In each case, the binding pockets are found in irregular structural features, such as non-sequential and cross-strand base stacks, 'U-turns', 'zeta-folds', non-canonical base pairs, the loops of G-quartets and other unusual arrangements. Theoretical understanding of nucleic acid structure is inadequate to allow reliable predictions of these structural features, even when substantial sequence data sets are available from functional variants. To complicate structural predictions further, the irregular features within the binding pockets are often stabilized by interactions with the bound target. As more aptamer structures are elucidated, some of the general rules of RNA folding beyond the Watson–Crick base pair may begin to emerge.

The best studied nucleic acid–small molecule interactions in biology are those that inhibit or destroy their nucleic acid targets: ricin cleaves rRNA [10], psoralens covalently cross-link double-stranded RNA and DNA [11], bleomycin cleaves DNA [12,13], distamycin binds in the minor groove of AT-rich DNA [14] and a host of antibiotics bind rRNA and inhibit translation. Few natural productive interactions have been studied, such as guanosine binding by self-splicing group I introns [15], although there may be many such interactions that have yet to be discovered. Interactions between antibiotics and their ribosomal RNA (rRNA) targets have received a surge of attention in recent years, fueled both by the availability of new experimental approaches and by the increasingly urgent need to find new antibiotics to treat resistant bacterial strains. For the experimentalist, the association of

antibiotics with particular RNA motifs offers the opportunity to model specific elements of ribosome structure and to probe the basis of molecular recognition by RNA.

Peptidyl transferase activity, by which an activated amino acid is added to the carboxyl terminus of a nascent peptide chain, resides on the large ribosomal subunit (50S particles in bacteria) and is inhibited in prokaryotes by many antibiotics, including chloramphenicol (Cam). The rRNA segment most closely associated with peptide bond formation and with Cam binding is the central loop of domain V in 23S rRNA, also known as the peptidyl transferase loop. Sequences in this loop are highly conserved among all extant forms of life [16,17]. Several bases in the loop are protected by CCdApPuro (a peptidyl transferase inhibitor) [18] and/or cross-link to benzophenone-derivatized aminoacyl tRNA [19]. At least 11 bases in the central loop of domain V are implicated in Cam binding on the basis of the locations of Cam-resistant mutations [20–22] and altered chemical reactivity in the presence of Cam [20,23,24]. Many of the positions associated with Cam binding are also protected by bound tRNA in the A- or P-site [25] or by other antibiotics that interfere with peptide bond formation [24]. It is generally believed that the loop is intimately associated with peptide bond formation and that Cam recognizes some component of rRNA structure critical to this process [21,24,26,27]. To understand how rRNA contributes to peptide bond formation, it would be useful to establish the general nature of RNA–Cam interactions.

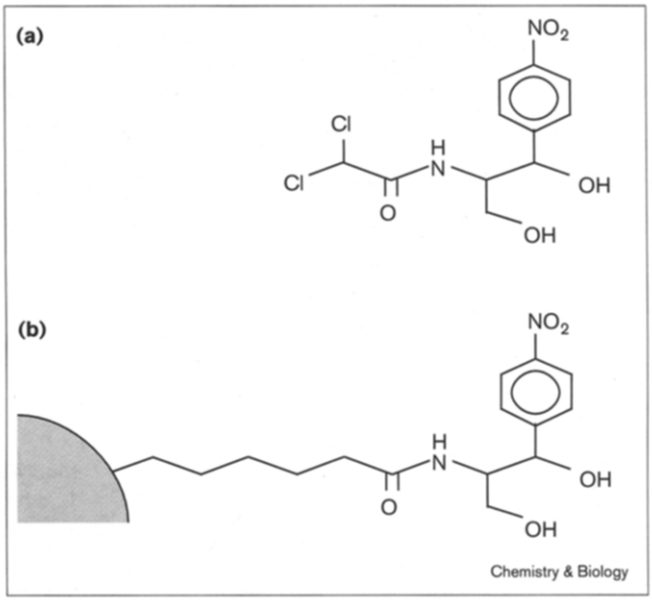
We selected two RNA populations that bind to Cam. The binding characteristics of various isolates and the structural attributes of the dominant sequence motif were determined. Although there is great diversity among the aptamer sequences, several share a sequence motif containing two symmetrically arranged A-rich bulges. Several lines of biochemical evidence suggest that the binding pocket is formed from tertiary interactions between the A-rich bulges, stabilized by Cam binding. Finally, we discuss the possibility that there may be parallels between the arrangement of the bases in the aptamers and possible interactions within the peptidyl transferase loop.

Results

The selected Cam-binding populations

Two independent Cam-binding RNA populations with 70 or 80 random positions (70Cm and 80Cm, respectively) were isolated using the SELEX protocol [1–3] from random starting pools containing 10^{14}–10^{15} sequences. Cam-binding RNA transcripts were affinity eluted from a Cam-derivatized agarose resin, amplified, and re-transcribed for additional partitions over the resin (Figure 1). The elutions included long incubations in order to encourage selection of high-affinity aptamers with long off rates (see the Materials and methods section). Specific elution from the resin with free Cam was low in the first seven

Figure 1



Structures of **(a)** free Cam and **(b)** Cam-derivatized agarose.

cycles of selection/amplification (less than 0.2% of the input RNA in the seventh cycle). This number climbed to 2–6% in the eighth cycle and around 20% in the ninth. After three more selection cycles at elevated wash stringency, 40–50% of the RNA eluted with Cam. RNA recovered from the 12th cycle was converted to double-stranded DNA for cloning and sequencing. There were 74 distinct sequences among the 96 isolates sequenced (Figure 2; 45 of 52 80Cm sequences and 29 of 44 70Cm sequences). Twelve sequences were found in pairs of identical or nearly identical isolates, two in three different isolates, and two in four different isolates. All other sequences were found only once. Subsets of various sequences contained short runs of four to ten nucleotides that were shared

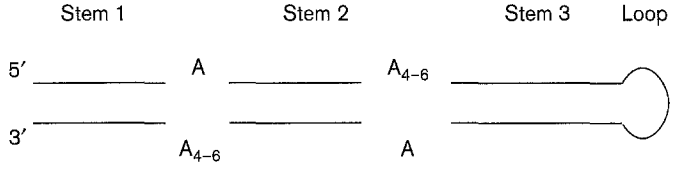
Figure 2 ➤

Sequences of Cam aptamers. **(a)** Sequences with the pseudosymmetric (Cm1-like) motif. Base pairing schema are shown above the alignment. Gaps in the pairing are indicated by ^; bulged or mispaired bases within helices are underlined. **(b)** Schematic depiction of proposed secondary structure of sequences shown in (a). **(c)** Sequences of other Cam aptamers. 15 of these sequences (from 17 isolates, marked with †) could be folded into structures with single asymmetric bulges that weakly resemble half of the Cm1-like binding element. Fixed sequences used in reverse transcription and PCR primer binding are shown in lower case. Numbers in parentheses indicate multiple isolations of the same or nearly identical sequences. *Denote empirical boundaries (see Figure 3). 22 sequences are not shown (10 70Cm and 12 80Cm); these appeared only once, showed no similarities to other sequences, and were not characterized further. All full-length aptamer sequences are available at http://beagle.colorado.edu/~dhburke/all_seqs.html.

(a)

	Stem 1	A	Stem 2	A_{4-6}	Stem 3	Loop	Stem 3'	A	Stem 2'	A_{4-6}	Stem 1'
70Cm15 (2)	*ag	a	GAUGUG	AAAAAA	CUGUGUCCG	ACU	GGACUCAG	A	CACAUU	AAA	CU*
70Cm20r	26nt...GCCCACG	A	CACUGU	AAA	AUGGCU	23nt	AGUAAU	C	ACAGUG	AAAAAC	CGCgggc...20nt
70Cm42	15nt.auacacaag	a	CACUG	AAAAA	UCC	GGCGAA	GGA	A	CAGUG	AAAAA	CGCAUGUAU.51nt
70Cm55	33nt.....CAG		CAGUG	GAAAAG	GAG	22nt	CUC	G	CACUG	GAAAA	UUG.....28nt
80Cm1 (2)	*UGCGA	A	CAGUG	UAAAA	UCGAC	AAU	GUCGA		CACUG	AAAAA	UCGCG*
80Cm12	27nt....UA^AGC	A	CAGUG	GAAAA	UGAA	26nt	UUCG		CACUG	CAAAA	GCUGUA....42nt
80Cm21 (2)	24nt.....AUAGU	A	CAGUG	CAAAA	GAUAAG	CACUAAGA	CGCAUC	A	CACUG	CAAAAC	ACUGU.....46nt
80Cm29	42nt....GAUUCU	A	CAGUG	CAAAACGCA	GGUCGUGUGAGG	ACUAA	CUUCACG^GGCU	CA	CACUG	AAAAA	uugauu....17nt
80Cm33	*GGCA	A	CAGUG	AAAA	ACACCUUUGGUU	UAAAC	AACCC^^UUCGU	A	CACUG	CAAAAC	UGCC*

(b)



(c)

```
fixed sequences: gggaaaagcgaaucauacacaaga ... gggcauaagguauuuaauuccaua
 70Cm3  (2) ACAGAGUCGAGGCUAUUGUGGAUGAUGUGCCCGAACAGACAUGCAUCAAUGUCUUGCCGAAUACCGCAAC
 70Cm4      ACUAAUGAUGUGCCCGCAGACUCAGUUACAGAAUGGUAACUAAUGCGAGACUGCCAAUACCCGGUGU
 70Cm6  (4) AUGAAAAGGGCUGGCGAGACAUAUCCGCUGGGCAAUCAGAUUCGGAGCCGCACCACCCUCGAAGUAGACA
 70Cm9  (3) AGAGCUUGACGGUCCCGAGAGUCGAGCCCAAGCUGACACUGGACCUUUGCGGACCACGUGUUGAUCGUCG
 70Cm11     GCGUUUGUGUUGGUCUACCACAACACGAAAGAUAUACGAUAACGUGAUGAAGCUGCAUACCGUAGUGCGA
 70Cm12 (3) UCGAAAAUCUAGGUGUGAAGAAAGCUUAAUACACUCCGAAAAGCAGUGAUGGAAGCUGCAGUCCAACGUU
 70Cm14     AGAGAACCAAGAAGAAUAUUGGAAAUACACGUUAUCGCUAUAGCGUGCCACGGAAUCGAAAUGACCA
 70Cm19 (4) ACGCAGAUCGAUGCGAGACUUGUGAAGAGAUGCCCGUAAUCGACUAACUGGACACACGAAUACUAACCAC
 70Cm21 (2) AGCUAUGAUCAGAGAUGCAAUAAACCGUCAAGCAGUGAAUCAGCUGUUGCAUCAACUGGAUACCACGGAC
 70Cm24     CGCGGCCCAAGACCUACACCUAAGUGGGCAAGAGAAAGCCAAAACUAUUACAACACAAGUUCACGGAGCC
†70Cm26     AGUUGUGCAGACCGAAUCCUCAGUCCUAGAAUUCCGGACUGUGUAAAAAAACCAAGCCGUACGCAAGU
 70Cm29     GAAUGCUAAUCACCAAUUUCUAACCCGACAAUGUAGAAAUACAUGGAAAUUGCGGAUUGCAUGUGAGGCUCG
 70Cm32     UGAGUGAAACACAGGAGCUACGGACUGAACCUAGCUGACAGGAAACCAUCAAAGUGAUGAUGUGCCAUGGUG
†70Cm45 (2) UUAGUGUAAGAUGUGCCCGCGGACUCAAACUGGAGUCUGGAGAAAAAAGUUCUGUNNCGCAAACCCGACC
 70Cm53 (2) GGCACCAAAGCUGAAGUAGCGGGAUAACUCAAAUUACUUUAGGUGUAUGAAGGUGAAACUAGCAAUGAA


fixed sequences: gggcauaagguauuuaauuccaua ... uugauucggaugcuccgguagcucaacucg
†80Cm3      CACCCGCACCAGAAGCCGAAGAAAAGAGACCGGAAAGUACCACCGGUUGGUGGACCCUGCGGAAACACnnCGCACAACUC
†80Cm4      CGCGACACCGUAACUGCGUCAUAGGCAUCAGGUGGAAAAACGCGGCUAUGUACCAUAACUACCCGACCGUAAACAACUC
†80Cm5 (2)  GCCGUAAACCCAAAAGCCAUGAAUUGAGAAGGAAUGGUUUCGCAUCCCAGCCCGGCUUGCUCAAACUGGUCAUUCGAAAACU
 80Cm7      GCAGACGACGAGGGCACACCGCCUAUGUCACUGUGCUAGGCACACCCUUCCCCCAGGACCACGCACAUUCGUCGCAUGA
 80Cm9      CGCGGAUCCAUAAUGACAUAGGAGAUGGCAACAGUCAAUUGGGGUUUGUGAUGCGGAUUGAAACUUCAAUAACCGGAUGG
 80Cm10     AGCCGAAGCGGGAUGACGCCAAAGCAUUAUCGACCUCCCGGUAUCGUAGCCGGCUCACUAACUGCAAGGUUCAAAAAGA
 80Cm13 (2) GCCGUCCGACGAGGAUUCAAGCGGUUGGAUUCACCAAGUCUUCACCCUUGAGUCUAGGUCACUCAACCAAAGGUACUGCG
†80Cm15     CCCAUGCGUUUUGGGCCGCUCCGGACUCCAGUUAUCAAACAGCCCAACUACGCAGCAUCUUACGUAUGCUGAAAAA
†80Cm16     UGUGCACCUAUCUACUCGCAACCGACAUGCCGCGCUAAUGGUAAAAUCCUUUAGCUCGGGACCGUCACUCAAUUUCGCUA
 80Cm17 (2) CUGACACGACACAGGGCAUGAUGCACCAUCUGUAUCGAAAUUGACUUAACCGCCUAUUCUUACAUAAACUGAACGC
 80Cm19     CCUACACUGUAGCACCACAAACUCCUUGCUCCCUGAAAUUAUUACCCUGGGACAGUGGGGAUGGCUCGUGCCAUGAUCGA
 80Cm23 (2) CUAGACACUCCGAACGACACCCAACAAGCUACGGUAAGCGCUCACCCUGGGGGAUUUUCACGCCCCCGCAUGAGGUGAAG
 80Cm31 (2) GCCUGGCACAGAGUGACGAGUCAAGAUCAGACAGGACACGCUCAUGCCGAUGACCCUUUAGCUUCCCGUCCCGAUUGAG
†80Cm30     AGCCACAUCCCAUCCACCCGUAGGAGAAAAAACCGUCUCUUUACACAUCCUAACGGGCGCUCGACGGUAACAAUUUGCGU
 80Cm38     GCCAAAAGGCAUAAACCACGACGGAGAUUCGGUGGUACGCUGAAUACGUUAGUUAACAACUCCGCUUACGGCAACAGUC
†80Cm42     CUCAAUGCUGUAAAAAGCUGUGUACCGCGCUUGGGUCUUGCGCCACCGAUGCACGUGAUGCAGCGCAGCAAAC
†80Cm43     CACACGACAGUAAAAGUCACUGCUAAAUGAACUAGCCUUCCUUCUAGCAUCUUCUCACUCUAAGGGGAAGCAAGACUGU
†80Cm44     GGCAGAAACAUGUACUCAUACCUUGGGAUCAAUAUGUCACAGACGCGAUUGCUGCUUUACCCAUCACGUGGAAAAUGAAU
 80Cm45     CAAGGCGUCCUGCACAGACAGAACAGCACCAGUUACGCGGCAGAAUCCCCGCGGCUUGAGUUCCUCCUAUGAGA
 80Cm46     GCCGACAUCCGACGAGGAUUUUAAACGAGGGCUCACGAAAAGACCCUCCGCUGCAGUCUAUGUCACUCAUCUUAGAA
 80Cm47     CGGAAUCUGAAGGAGGAAUUUCAGAGGAACCUUAAAUAACCGGACAGCUCGCCUCAUUCCUGCGUCGAAUCUUAAACUUU
 80Cm50     CGGCAGACGAGCCUUGACGAGCCAAUCUACACUUGGCGAUGACCGAUGGGCCCCAGCUACUUCUGGCAGUUUCGAUUCGU
 80Cm51     CCCGgCUAGCCGAUACAUCCAUUCGGAACUGCUGACCGUUAGGUGAAUAUCGCCAGUCCUACACUGGGAGCUCAUAAGCC
 80Cm54     CUCGGGAUACGGCCUCCCGCGAUCGGACAUUCUCCGGCGGCUCUCAGUACUGGAUAAAAGGGCGGUUUCCCGGCUCAGAG
 80Cm55     CCAGCCGCGACCUUCUACUUCAAACCCAGAUACGCUGCAAAACAGAUACUCCUCAGUAUCACAGCCUAACUGUUUGAGUC
†80Cm59     CCGCCGGUCCGCGCCAAUCCAGCAGCCGCACCGUAACUUGUCGUCUGAAACAACCAGGUGUAAAAGGGCGUAAU
†80Cm61     GCCGUGCUGCCGACGAGUACAGCCCUAUCGGCACAAAUGCUGAAAAAGGCAUUCGCUCUUGCCACAGCCCAACUUUGUU
†80Cm62     GUACUGCGAUCUGCAAAACCUCUCGGUGAGACGCAAGUGGCAUUCAUCGAUUGACAGAUUAAACGCACCCUAUUGUGA
†80Cm64     NNGCUACGGAGCCCACGGUUCGUCCAAGAAACACUAGAGCUUCUUAAGUCCUCUGUCUGAACCUGCACCUCUAGAUA
```

between two or more individuals, but no single sequence element was present throughout all the isolates.

Cam-binding on the column and in solution

We measured the ability of transcripts from 35 isolates (19 80Cm and 16 70Cm) to be specifically eluted from the Cam affinity resin (Table 1). To expedite the survey, the long elutions used during the selection were replaced with simple flows. Of these 35 RNAs, 26 bound at least 2.5-fold over background. For seven RNAs, including two that did not elute significantly with the simple flow elution (80Cm50 and 70Cm4), much of the input RNA remained bound to the column. When these RNAs were eluted with long incubations, significantly more RNA was released. Their slow elution behavior may have been due to long dissociation off rates, and were probably a required component in their recovery from the selection. Some of the RNAs that eluted poorly in this assay may have had slow on rates and required pre-incubation with the resin (as was done during the selection), whereas some may have been non-specific sequences that had been carried through the selection (80Cm51 showed no appreciable Cam binding in solution, see below).

Dissociation constants (K_d values) from Cam in solution were measured at room temperature using equilibrium filtration [28]. K_d values for the 12th round 70Cm and 80Cm pools were 25 ± 3 and 65 ± 4 μM, respectively, whereas the initial random pools showed no appreciable binding ($K_d > 1000$ μM). K_d values for transcripts from 19 isolates cloned from the 12th round pools ranged from more than 200 μM (barely detectable by this assay) to less than 2 μM (nearly 1000-fold tighter than the starting random pools). The mean value of approximately 60 μM is close to that measured for the aggregate behavior of the selected pools. The best binders were also the best represented sequences in the final selected pools, although a wide range of K_d values appears to be compatible with efficient column retention and specific elution. Further cycles would probably have pushed the selections toward the optimal aptamers and away from the conserved structural motif. For example, the 70Cm population bound with a higher aggregate affinity than the 80Cm population, and contained more individuals that were isolated multiple times, implying that it had been driven closer to completion, but it also contained fewer representatives of the dominant element.

Structure of the dominant Cam aptamer motif from biochemical probing

Nine sequences (from 12 isolates, 12% of the sampled sequence set) shared the potential to fold into helices with two asymmetric bulges, usually with four to six adenosines (A_{4-6}) across from a single adenosine (Figure 2a,b). Both helix 1 and helix 3 were of highly variable sequence with extensive covariation supporting pairing. The bulges and helix 2 were nearly invariant, although the several instances of covariation all support pairing within the central helix. We found it striking and significant that this structural core was conserved in both populations.

Table 1

Cm-binding behavior of 40 RNAs.

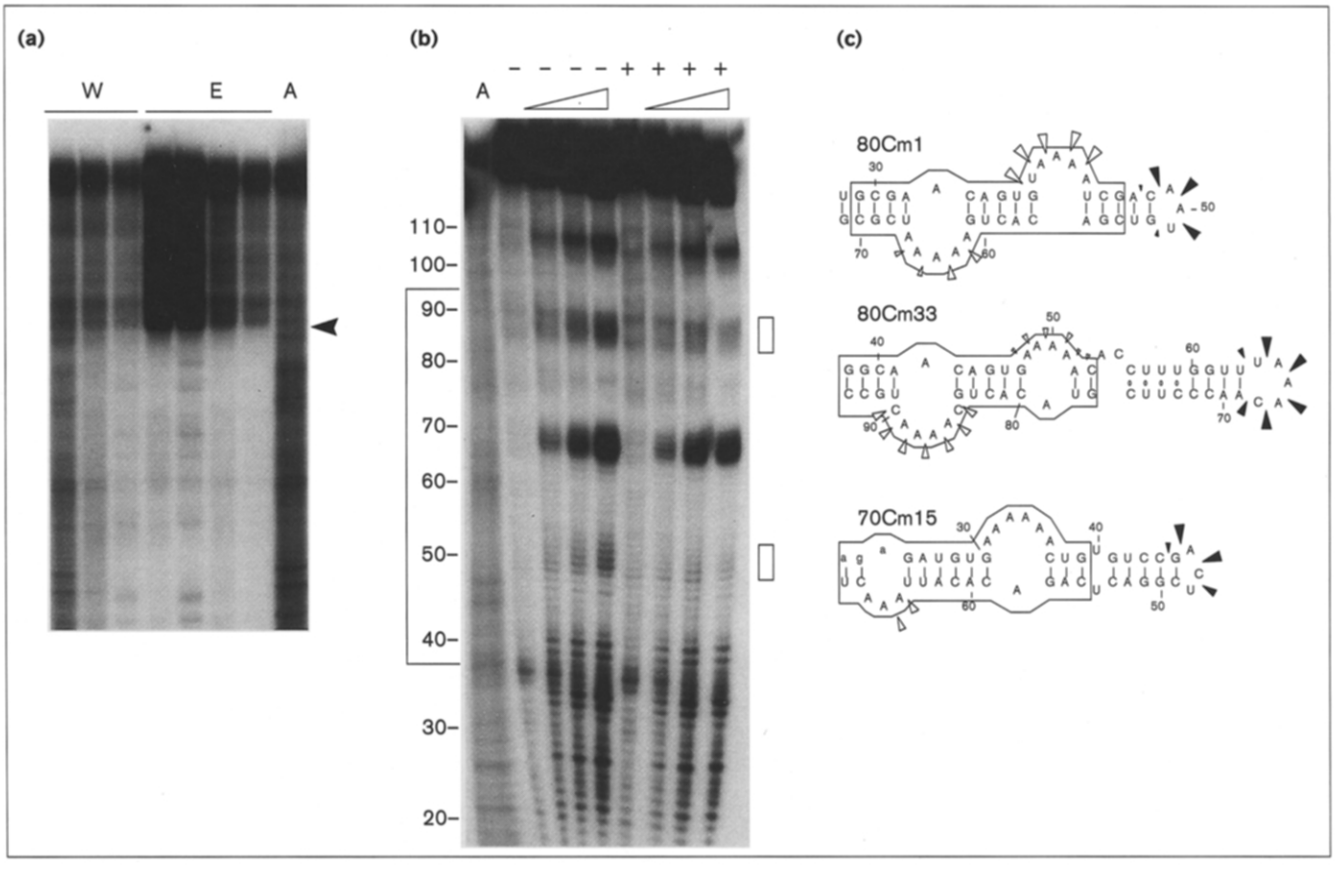
	Retained*	% Eluted	K_d (μM)
Pools			
70Cm pool 0			≈2000
80Cm pool 0			> 1000
70Cm pool12			25 ± 3
80Cm pool12			65 ± 4
Active isolates			
70Cm3 (2)	1.0	7.1	
70Cm4†	18.3	6.2	
70Cm4‡	0.2‡	20.1‡	
70Cm6 (4)	9.8	18.9	2.1 ± 0.3
70Cm9 (3)†	12.0	25.4	8 ± 4
70Cm12 (3)	0.9	23.4	50 ± 10
70Cm12§	5.2§	75.5§	
70Cm14	0.8	5.8	
70Cm15 (2)	2.7	20.3	140 ± 50
70Cm19 (4)‡	5.2	4.1	66 ± 10
70Cm19§	13.7§	20.1§	
70Cm24	1.9	8.1	
70Cm29 (2)	2.3	34.0	130 ± 30
70Cm42	1.8	17.4	275 ± 150
70Cm45 (2)	1.2	21.6	17 ± 10
70Cm53 (2)	1.0	20.2	7 ± 2
70Cm55†	12.1	18.8	
80Cm1 (2)	1.8	20.3	110 ± 60
80Cm7	4.3	12.5	
80Cm10	0.9	17.2	
80Cm15	0.9	9.8	
80Cm19	1.7	28.2	110 ± 60
80Cm30	2.0	20.2	30 ± 5
80Cm31 (2)†	10.9	7.2	
80Cm33	0.7	19.0	165 ± 50
80Cm45	1.8	10.6	
80Cm46	4.0	16.7	216 ± 75
80Cm47†	6.7	2.5	
80Cm50†	22.4	10.0	
80Cm50‡	1.3‡	29.1‡	25 ± 3
80Cm55	1.1	21.0	195 ± 75
80Cm59	3.6	20.9	96 ± 20
Inactive isolates			
70Cm11	0.1	1.7	
70Cm32	1.3	1.6	
80Cm9	2.8	2.0	
80Cm29	0.9	4.9	
80Cm38	1.2	1.9	
80Cm51	2.9	2.1	>1500
80Cm54	1.1	4.6	
Negative control			
5t83	1.4	2.1	

Uncertainties reflect precision of measured K_d after at least three measurements, usually at two or three RNA concentrations. Numbers in parentheses indicate the number of individual isolates with identical or near identical sequences. *% input RNA remaining on column after elution; †significant retention on column following elution; ‡slow elution (see text for details); §500 μl bed.

To determine the 5′ and 3′ boundaries of the sequences required for Cam binding, we performed a deletion–selection analysis [1,29,30] of three of the isolates shown in Figure 2a. The earliest wash fraction shown contains an even distribution of all fragments that is similar to the initial alkaline hydrolysis ladder (Figure 3a). The later wash fractions are increasingly dominated by the same RNAs recovered during the elution. By this point in the wash, all non-functional RNAs have been removed from the column, leaving the RNAs with intact Cam-binding elements to leach off slowly. The smallest transcripts, labeled at the 3′ end and deleted from the 5′ end, that eluted specifically from the column define the 5′ boundaries of the minimal binding elements. Similar experiments were carried out with 5′ end-labeled RNA to determine the 3′ boundaries (not shown). This approach localized the binding elements to regions 44–57 nucleotides in length, in each case exactly flanking the conserved structural core. All of the sequence elements required for Cam binding, along with a few inessential bases, are located between the boundaries.

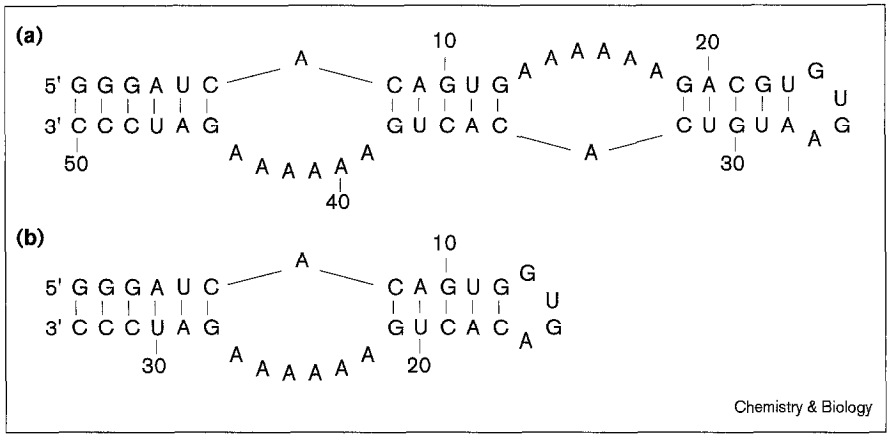
Ribonuclease S_1 (RNase S_1) preferentially cleaves single-stranded RNA. When RNA from isolate 80Cm33 was treated with ribonuclease S_1, the predicted helices were resistant to cleavage by RNase S_1, while the terminal loop and large internal bulges were, for the most part, readily cleaved (Figure 3b). Among the six A_{4-6} loops in the three isolates shown in Figure 3c, five showed reduced sensitivities to RNase S_1 in the presence of 1 mM Cam. These partial protections may be due to stabilization of RNA structure or to direct protection by Cam, either of which would suggest a role for these positions in Cam binding. The large A-rich bulge in 70Cm15 was resistant to cleavage

Figure 3



Structural analysis. **(a)** 5′ Boundary analysis for isolate 80Cm33 is indicated with an arrowhead. W, progressive fractions of RNA washed from column just prior to elution (early on the left, later on the right); E, progressive fractions of RNA eluted with free Cam; A, alkaline hydrolysis ladder prior to column partitioning. **(b)** RNase S_1 sensitivity for 80Cm33 as a function of time (shown as wedges above the autoradiogram) and in the absence (–) or presence (+) of 1 mM Cam added prior to cleavage. Sites of protection are highlighted with the open boxes to the right, and the region between the functional boundaries is indicated on the left. **(c)** Inferred secondary structures of three isolates, with sites of RNase S_1 reactivity denoted by triangles of different sizes to indicate the relative intensities of the cleavage in the absence of Cam. Open triangles indicate sites of protection by Cam, and closed triangles are cleavage sites that do not show protection. Conserved cores are boxed. A run of U·U and C·U pairs may form in isolate 80Cm33 (dots), given the lack of RNase S_1 cleavages in this region and the demonstrated stability of tandem U·U pairs [41].

Figure 4



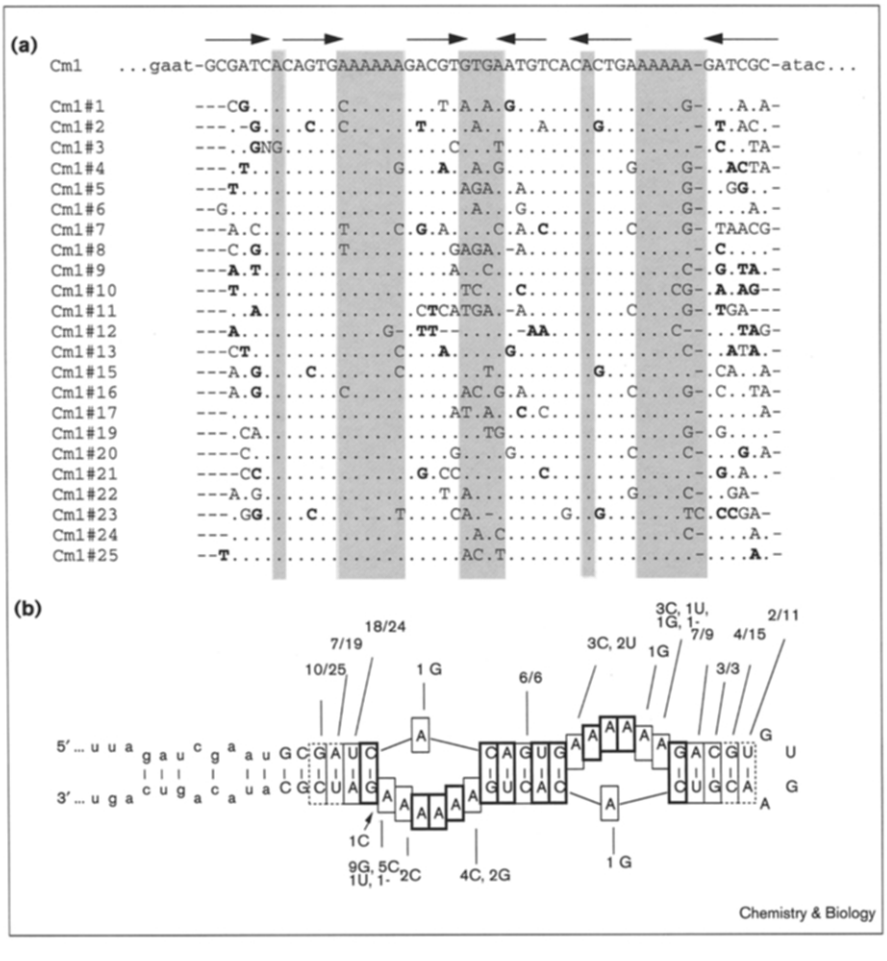
Minimal model Cam aptamers. **(a)** Cm1, a 50-nucleotide RNA with the structural features shown in Figure 3c. **(b)** Cm2, a 33-nucleotide RNA with one asymmetric bulge element.

by S_1 in the absence of Cam, suggesting that it may be ordered in the RNA alone. The boundary and nuclease sensitivity data are summarized in the structural models of these three Cam-binding RNAs shown in Figure 3c.

Refining the sequence requirements for the dominant Cam aptamer motif

A 50-nucleotide RNA containing only the minimal elements suggested by the initial selection, boundary and

Figure 5



(a) Sequences of Cm1 variants, with the base pairing scheme shown above. Unpaired regions are shaded. Aligned nucleotides that are identical to Cm1 (top) are shown as dots; gaps are shown as dashes. Lower-case letters are portions of the appended fixed sequences. Bases in bold represent mutations that support base pairing. Most of the Cm1-derived isolates lacked the first two nucleotides, perhaps as the result of a synthesis error in the starting pool. **(b)** Summary of reselection data. Invariant nucleotides are shown in boxes with thick lines. Base pairs that are supported by covariations are boxed in solid lines, those that are discounted by covariations are boxed with dashed lines. Covariation data are shown as fractions: numerator, total number of mutations that support pairing; denominator, total number of mutations at that position. Mutations at unpaired positions are enumerated, with insertions shown with an arrow and gaps shown with dashes. Lower-case letters are the appended fixed sequences. Upper-case letters with no mutational data shown are highly variable (compare (a)).

RNase S_1 data was synthesized for further analysis (Figure 4a, 'Cm1'). Cm1 RNA bound to and eluted from the Cam-derivatized column with an activity equivalent to that of the full-length RNAs. We detected little column-binding activity from a similar RNA with only one asymmetric bulge (Figure 4b, 'Cm2'). To determine the sequence requirements more precisely, we resynthesized Cm1 with 15% mutagenesis per position. Cam-binding activity returned to the Cm1-derived pool after five cycles of SELEX. The 23 sequences of active isolates analyzed from the Cm1-derived pool contained, on average, 7.9 mutations from the parental sequence, corresponding to a net mutagenesis rate of 16% per position (Figure 5). The patterns of conservation and covariation were similar to those observed for the original selection. Eighteen positions were invariant, including three symmetrically arranged adenosines in each of the A_{4-6} loops, four of the five central base pairs, and the C:G pairs at each end of the binding element (shown as dark boxes in Figure 5b). The central G:C pair mutated on three occasions to a C:G pair, effectively inverting the topology of the element (as in isolate 70Cm42). The 5′ and 3′ positions of each of the two A_{4-6} elements were the most variable (22-70% mutated). The remaining four adenosines in the bulged loops were nearly invariant, with only one or two mutations observed in each (5% mutated). The most proximal nucleotides within the flanking helices showed strong covariations supporting base-pairing interactions, whereas the more distal positions were highly variable, with no preference for Watson–Crick pairs. The structural core appears to require at least two base pairs on either side.

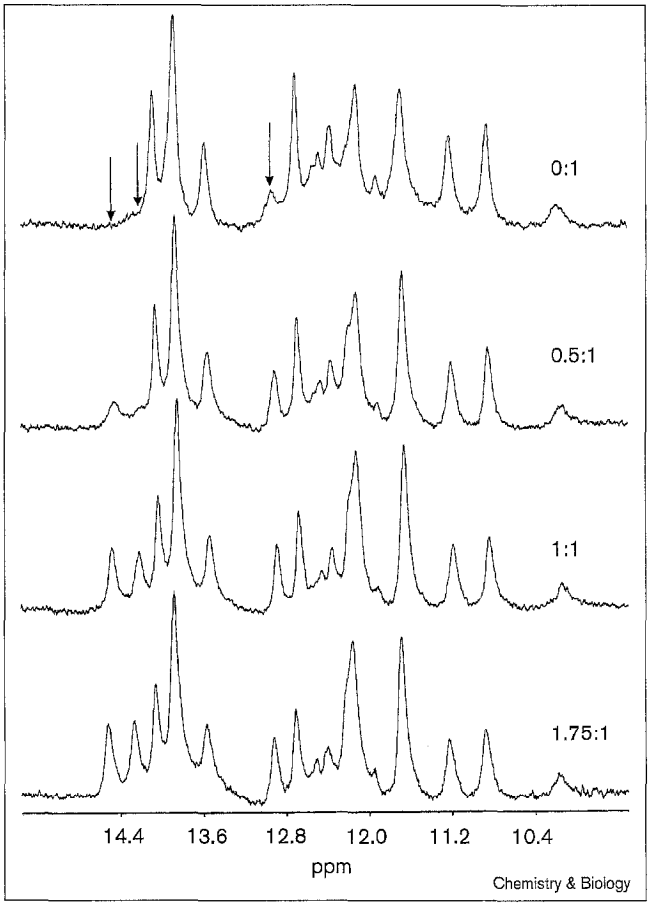
Many RNA sequences can bind Cam besides those in the Cm1-like family, from which we conclude that Cam binding is easy for RNA. Nevertheless, Cm1-like sequences are estimated to have been rare in the initial random pool. One random sequence in 4^{18} ($= 7 \times 10^{10}$) could be expected to contain the 18 most highly conserved nucleotides. Allowing for an additional 4 base pairs (two on each side), no more than one sequence in approximately 3×10^{13} is expected to contain all of the sequence information in the Cm1 element.

We performed a similar reselection using a pool constructed by resynthesizing Cm2 with 15% mutagenesis per position. Strong activity was apparent after six cycles. The sequences from this pool bore little resemblance to Cm2, having an average of approximately 40% mutations per position, and they were not analyzed further. These sequences appear not to have optimized the Cm2 sequence by exploring local sequence space, but rather to have diverged to other motifs.

Stoichiometry of the bound complex

The two symmetrical elements of the Cam binding motif could potentially form a complex with either one or two Cam molecules per RNA. To better understand the stoichiometry of binding, the RNA oligomers Cm1 and Cm2 (see Figure 4) were titrated with Cam and monitored by one-dimensional ^{1}H NMR at 5°C. The ^{1}H region of the NMR spectrum corresponding to the guanine and uridine imino resonances of the RNA aptamer Cm1 is shown in Figure 6. As the Cam concentration was increased, three new proton resonances increased in intensity, saturating at a ratio of 1.25:1 Cam:RNA. These three new proton signals resonate at chemical shifts that are consistent with the formation and/or stabilization of three Watson-Crick base pairs. In contrast, similar NMR studies of the titration of the Cm2 aptamer with Cam gave no evidence of binding (not shown). On the basis of these results, we conclude that tight binding both requires symmetrical elements and occurs in a 1:1 stoichiometry. Finally, at a 1:1 ratio of Cam:Cm1 RNA, we estimate that more than 80% of the

Figure 6



One-dimensional ^{1}H NMR spectra of a titration of the RNA aptamer Cm1 with free Cam at 5°C. Shown are the imino proton resonances, ranging in chemical shift between 10 and 15 ppm. The Cam concentrations are 0 mM, 0.23 mM, 0.46 mM and 0.81 mM (from top to bottom). Cm1 aptamer concentration is 0.45 mM. The Cam:RNA ratios are given above each spectrum. Three new imino proton resonances that appear with increasing concentrations of Cam are designated with arrows.

RNA is complexed, indicating a dissociation constant at 5°C of ≤100 μM.

Discussion

Cam binding by Cm1-like aptamers

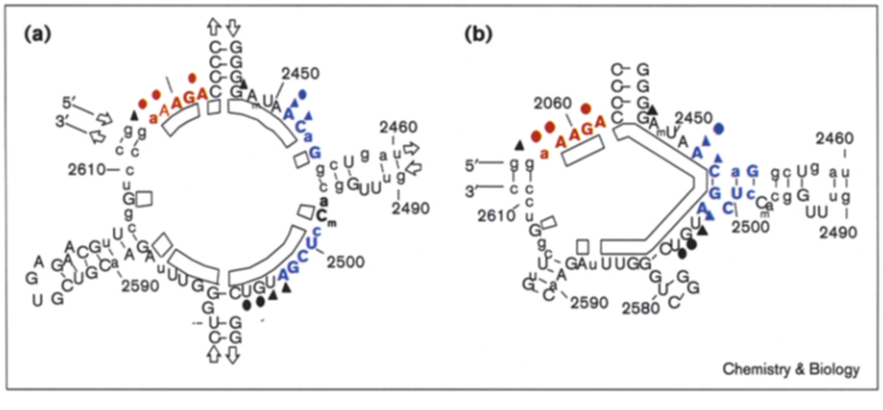
We have identified a Cam-binding RNA motif that contains two large internal bulged loops and was present in two independent selections. All of the required sequence elements are contained within the 50 nucleotide Cm1 RNA. Most of the features of Cm1 were retained upon reselection following mutagenesis. The unpaired bulges are protected from cleavage with RNase S_1 by bound Cam, which forms a 1:1 complex with the RNA according to proton NMR titrations.

It is curious that Cm1 RNA should require two A_{4-6} elements with such strong symmetry about the central helix, even though Cam is not itself symmetrical and the complex contains only one molecule each of Cam and RNA. The 5 base pairs of the central helix are expected to introduce one half turn into a normal A-form helix. Both A_{4-6} elements would then lie on the same face of the aptamer, where they may form a single binding site through tertiary interactions. On the other hand, we cannot rule out the possibility that each molecule contains two binding sites that cannot be occupied simultaneously. The lack of activity in Cm2 argues against this possibility. In further support of there being a single binding site built from tertiary interactions, we note a weak correlation in mutations at the 3′-most positions of the two A_{4-6} elements. Further experiments are needed to assess the separability of the two halves of the aptamer and the possibility of tertiary interactions within Cm1.

Are the Cam aptamers models for the peptidyl transferase loop?

Cam is usually thought to inhibit peptide bond formation by acting as a conformational analog of the 3′-*O*(aminoacyl)adenosine unit of the incoming aminoacyl-tRNA [31–34], or the transition state for peptide bond formation [34], or a newly formed peptide bond [34,35]. Underlying each hypothesis is the assumption that the Cam–ribosome interaction is intrinsically similar to the interaction between the ribosome and some component of the translational apparatus. There are some similarities between Cm1 and the peptidyl transferase loop of 23S rRNA (Figure 7a,b). The sequence ACAG (corresponding to positions 7–10 in Cm1) is present in the 2450 arc of the loop, while the sequence CUcGA (corresponding to positions 36–39 in Cm1, omitting the lower case C) is present in the 2500 arc of 23S. These two arcs may come together with Watson–Crick pairings C2452:G2502, A2453:U2500, and G2454:C2499, bulging out nucleotides A2497, C2498 and C2501. In this arrangement, six nucleotides implicated in Cam binding by ribosomes that are widely dispersed in the standard secondary structural representation (A2451, C2452, A2503, U2504, G2505 and U2506) are brought into proximity. Among these, A2451 and G2505 are both protected from chemical modification upon binding of Cam, carbomycin, vernamycin B [24] and P-site-bound *N*-acetylphenylalanyl-tRNA^{Phe} [25]. Interactions in this region have not been detected in published phylogenetic analyses because there is little natural variation in these positions that could serve as the basis for assessing covariation [16,17]. Nevertheless, ongoing analyses with additional 23S-like sequences suggest interactions between the 2450 arc and the 2500 arc (R. Gutell, personal communication). In addition, A2453, which is highly conserved among bacterial rRNA sequences, is replaced by a highly conserved U in eukaryotes. The resulting U·U pair could undermine formation of the Cam-binding structure and account for the specificity of Cam for bacterial ribosomes. Three additional bases in 23S rRNA whose chemical reactivities are altered upon Cam binding (A2058, A2059 and A2062) form part of an unpaired purine-rich stretch (2058-AAAGA-2062) that is particularly reactive to single-strand-specific probes [20,23,24] (Figure 7). If positioned close to the paired region noted above, the AAAGA sequence could form a

Figure 7



(a) Standard depiction of the peptidyl transferase loop of 23S rRNA using *Escherichia coli* numbering [16]. **(b)** Possible interactions in the 23S rRNA peptidyl transferase loop forming the ribosomal Cam-binding site. Highly conserved positions are shown in upper case with universally conserved positions around the center of the loop indicated with open boxes [17,23]. Less strongly conserved positions are in lower case. Circles, positions with altered chemical reactivity upon Cam binding; triangles, locations of Cam-resistant mutations. Regions discussed in the text as possibly being analogous to Cm1 are shown in red and blue.

structure analogous to the four to six unpaired adenosines in the asymmetric bulges of the Cam-binding isolates (e.g. positions 62–66 of 80Cm1), which are themselves sensitive to RNase S_1 cleavage and protected upon Cam binding. In this model, the ribosomal Cam-binding site in 23S rRNA is formed by juxtaposing the three arcs of the peptidyl transferase loop that have been implicated in Cam binding.

The observations noted above make it tempting to see the selections for Cam aptamers as having converged upon rRNA mimics. However, this model is premised on the independence of the two halves of Cm1, each of which may be analogous to the Cam-binding conformation of 23S rRNA. The model is made tenuous by the lack of activity in Cm2, which appears to contain an intact half-site. If the two oligo-A regions come together through tertiary interactions to form a single binding site, as suggested by the 1:1 stoichiometry, the apparent parallels between these Cam aptamers and 23S rRNA would disappear. Future work will be directed towards resolving these questions.

Welch *et al.* [36] have reported two aptamers (designated FA-1 and FB-1) selected to bind the peptidyl transfer inhibitor CCdApPuro, which was designed to mimic the 3′ end of tRNA joined to an amino acid through a transition state analog for aminoacyl transfer [18]. These aptamers predominately recognize CCdA (K_d = 170 nM), with an additional 1.7 kcal mol^{-1} of binding conferred by the inclusion of puromycin ($K_d \approx 10$ nM). They contain a consecutive octamer motif matching the 2450 arc of the peptidyl transferase loop (2447-AUAACAGG-2454) [36]. Both CCdApPuro and Cam interact with the same portion of 23S ribosomal RNA [18,24]. Nevertheless, the CCdApPuro aptamers and the Cam aptamers are not obviously similar. It remains to be demonstrated conclusively whether any of these aptamers is a true mimic of the peptidyl transferase loop. If they are both mimicking some active state of rRNA, their differences may reflect alternate conformations assumed by 23S rRNA at different stages during the translation cycle.

Significance

Ribosomal RNA is essential for translation, but its detailed structure is poorly understood. Knowledge of how RNA interacts with translation-inhibiting antibiotics may aid in formulating models of ribosomal structure, lead to the development of new antibiotics and deepen our understanding of molecular recognition by RNA. We used *in vitro* selections (SELEX) to identify RNA molecules (aptamers) that bind chloramphenicol (Cam), an antibiotic that inhibits translation by binding prokaryotic ribosomes and blocking peptide bond formation. Various genetic and biochemical approaches converged on a structural model for these aptamers, which includes two symmetrically arranged A-rich bulges. Proton nuclear magnetic resonance spectroscopy established that the folded RNA is stabilized upon Cam binding and that there is a stoichiometry of one Cam molecule per aptamer molecule in the bound complex. The two halves of the symmetrically arranged motif appear to come together through tertiary interactions to form a single binding site, although we cannot yet rule out the possibility that the aptamer contains two separable sites with mutually exclusive binding. In the light of this latter possibility, the detailed structures of these aptamers may illuminate the Cam-binding conformation of 23S rRNA. Formation of an analogous structure in 23S rRNA would bring into proximity all three regions of the loop that are implicated in Cam binding by ribosomes. Further analysis of the aptamer–Cam complex will be needed to determine whether there are analogies to be drawn with 23S rRNA.

Materials and methods

Starting nucleic acid populations and amplifications

All deoxyoligonucleotides were synthesized by automated phosphoramidite chemistry. RNA was transcribed from two independent DNA pools (each containing 10^{14}–10^{15} different sequences) with 70 or 80 random positions and distinct fixed sequences for reverse transcription and amplification as described elsewhere [37]. Each transcript is approximately 134 nucleotides (80Cm series) or 118 nucleotides (70Cm series) in length. 1 nmole RNA was subjected to the first cycle of selection (two to five pool equivalents), and 125 pmol RNA were used in subsequent cycles. All RNAs were synthesized by *in vitro* transcription with T7 RNA polymerase [38] and purified by denaturing polyacrylamide gel electrophoresis.

Column binding

RNA was unfolded at 70°C for 2 min in water then refolded by addition of binding buffer and quick cooling on ice. Final buffer concentrations were 20 mM $MgCl_2$, 400 mM NaCl, 100 mM bis-Tris pH 6.4. After equilibrating to room temperature, folded RNA was mixed with 250 µl of a 50% slurry of pre-equilibrated, Cam-derivatized resin (Sigma, 4% beaded agarose derivatized with approximately 7.5 mM available Cam, as measured by its ability to retain CAT protein). The RNA–resin slurry was filtered through glass wool in a 1 ml syringe and washed with 10 column volumes of binding buffer to remove weakly bound RNAs (40 volumes for cycles 10–12). Rigorous elution was used to remove all specifically bound RNAs, including those with long off rates: columns were first flushed with three column volumes of elution buffer (binding buffer + 10 mM Cam), then incubated three times with 1½ volumes elution buffer for 15 to 30 min with rocking and finally flushed with an additional three volumes of elution buffer. Eluted RNA from all nine fractions was pooled, precipitated with glycogen as a carrier and amplified for subsequent cycles. Nonspecific retention on the column was very high (20–50%) during the first three cycles (during which binding buffer was used at half strength), due to anionic exchange properties of the resin. During the post-SELEX survey of the isolates, refolded RNA was layered directly onto a preformed bed of resin, and eluted with eight column volumes, usually as a simple flow elution (without the long incubations noted above). An 83 nucleotide aptamer to coenzyme A (5t83, , D.H.B. and D.C.H., unpublished observations) was used as the negative control.

Mutagenic reselection

Mutagenized pools were synthesized as bottom strand by Genemed Synthesis, Inc. (South San Francisco) to yield template with the format 5′-CGGCTTaagct*tAATACGACTCACTATA*GGGTTAAGTTAGATCG-AAT-(mutagenized aptamer)-ATACAGTCAGTCGAACCCATACTAA-TC-3′. *Hin*dIII cloning site is shown in lower case and the T7 RNA polymerase promoter is italicized. Gel purification of the synthesis product yielded approximately 170 pmol each for Cm1- and Cm2-

derived pools, corresponding to approximately 10^{14} sequences. Primers complementary to the new fixed sequences were used to amplify the pool for transcription. RNA (1 nmol) was used in the first cycle and 250 pmol in each subsequent cycle. Binding buffer was used at full strength throughout, and the elutions included three 10 min incubations with free Cam, as above.

Quantification of binding affinities

Dissociation constants (K_d) in solution were measured using ^{14}C-labeled Cam in a microconcentrator spin assay [28], in which RNA•Cam complex is retained on one side of a size-fractionating membrane. Refolded RNA at various concentrations was mixed with 5–10 μM ^{14}C-labeled Cam (New England Nuclear) on ice in binding buffer, in a total volume of 390 μl. After 15–20 min at room temperature, three 120 μl aliquots of the mixture were spun for 40 s through individual microcon-30 concentrator filters (Amicon), during which 40 ± 7 μl of the 120 μl reaction filtered through. Samples (30 μl) from the bottom (filtrate) and top chambers were counted separately in 4 ml EcoLume scintillation cocktail. A slight RNA-independent accumulation of Cam in the top reservoir was subtracted from each set of data before calculating the concentration of RNA-Cam complex as follows:

[RNA·Cam] = 2/3 (specific activity of Cam) × [(^{14}C counts retained)–(^{14}C counts in filtrate)].

Specific activity was determined by measuring the radioactivity in a [^{14}C] Cam sample of known concentration. The 2/3 factor takes into consideration that the complex is, by the end of the spin, approximately 3/2 as concentrated as at the beginning of the spin, assuming slow re-equilibration. If the re-equilibration is rapid, this equation underestimates the amount of complex formed initially and overestimates K_d. The calculated concentration of bound complex was averaged over three assays for each concentration of RNA used, and separate bind reactions using different RNA concentrations were used in calculating the final K_d.

Boundary determinations

Transcripts were labeled at the 5′ end with [γ-^{32}P]ATP and polynucleotide kinase, or at the 3′ end with [5′-^{32}P]pCp and RNA ligase. Partially hydrolyzed, end-labeled RNA was incubated with resin, then layered onto a bed of preformed resin. RNA from the wash fractions or specifically eluted with free Cam was precipitated and analyzed by gel electrophoresis on a 10% polyacrylamide gel.

Ribonuclease S_1 sensitivities

RNA (5′ labeled) was folded in binding buffer supplemented with 30 μM Zn^{2+}. After removing an aliquot for the zero time point, the remainder was incubated with 10 units of RNase S_1 at room temperature. Aliquots were removed from the reaction after 3, 10 or 30 min, mixed with gel loading buffer, and quickly frozen on dry ice/ethanol to stop the reaction. Cleavage positions were determined by gel electrophoresis on a 10% polyacrylamide by comparing the RNase S_1 cleavage pattern with that of RNase T_1 and the partial alkaline hydrolysis ladder.

NMR spectroscopy

Gel-purified RNA oligomers Cm1 and Cm2 were further purified using a DEAE–sephacel column and then dialyzed into 5.0 mM $MgCl_2$, 100 mM NaCl, 25 mM Tris pH 6.4 (0.25 times the buffer concentration used in column assays) using centricon-3 microconcentrators (Amicon). The approximate Cm2 concentration was calculated by assuming a concentration of 40 μg ml^{-1} for an absorbance of 1 at 260 nm. The extinction coefficient for Cm1 was calculated on the basis of the nucleotide composition at 25°C [39]. The UV melting profile for Cm1 was used to determine the degree of hypochromicity due to secondary structural interactions. The extinction coefficient was then adjusted accordingly to yield a concentration for Cm1 of 0.45 mM. One-dimensional NMR experiments were performed at 5°C and used gradient-tailored excitation for water suppression [40].

Acknowledgements

We thank R. Gutell, T. Shtatland, and Y.-Y. He for comments on the manuscript, M. Welch and M. Yarus for sharing their data prior to publication, and D. Lorenz for assistance in preparing figures. Funding for this work was provided by NIH grants GM19963 to L.G. and AI33098 to A.P., by postdoctoral fellowships 1F32AI09361 from NIH and CHE-9302453 from NSF to D.B. and by additional support from NeXstar Pharmaceuticals, Inc.

References

1. Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510.
2. Ellington, A.D. & Szostak, J.W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822.
3. Gold, L., Polisky, B., Uhlenbeck, O. & Yarus, M. (1995). Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**, 763-797.
4. Cech, T.R. & Szewczak, A.A. (1996). Selecting apt RNAs for NMR. *RNA* **2**: 625-627.
5. Dieckmann, T.S., Suzuki, E., Nakamura, G.K., & Feigon, J. (1996). Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* **2**, 628-640.
6. Jiang, F., Kumar, R.A., Jones, R.A. & Patel, D.J. (1996). Structural basis of RNA folding and recognition in an AMP–RNA aptamer complex. *Nature* **382**, 183-186.
7. Yang, Y.K., M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science* **272**, 1343-1347.
8. Zimmermann, G.R., Jenison, R.D., Wick, C.L., Simorre, J.P. & Pardi, A. (1997). Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat. Struct. Biol.* **4**, 644-649.
9. Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. & Feigon, J. (1993). Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl Acad. Sci. USA* **90**, 3745-3749.
10. Morris, K.N. & Wool, I.G. (1994) Analysis of the contribution of an amphiphilic alpha-helix to the structure and to the function of ricin A chain. *Proc. Natl. Acad. Sci. USA* **91**, 7530-7533.
11. Cimino, G.D., Gamper, H.B., Isaacs, S.T. & Hearst, J.E. (1985). Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.* **54**, 1151-1194.
12. Absalon, M.J., Wu, W., Kozarich, J.W. & Stubbe, J. (1995). Sequence-specific double-strand cleavage of DNA by Fe-bleomycin. 2. Mechanism and dynamics. *Biochemistry* **34**, 2076.
13. Burger, R.M., Drlica, K., & Birdsall, B. (1994). The DNA cleavage pathway of iron bleomycin. Strand scission precedes deoxyribose 3-phosphate bond cleavage. *J. Biol. Chem.* **269**, 25978-25985.
14. Rentzeperis, D., Marky, L. A., Dwyer, T. J., Geierstanger, B. H., Pelton, J.G. & Wemmer, D.E. (1995). Interaction of minor groove ligands to an AAATT/AATTT site: correlation of thermodynamic characterization and solution structure. *Biochemistry* **34**, 2937-2945.
15. Zaug, A.J., Been, M.D., & Cech, T.R. (1986). The *Tetrahymena* ribozyme acts like an RNA restriction endonuclease. *Nature* **324**, 429-433.
16. Schnare, M.N., Damberger, S.H., Gray, M.W. & Gutell, R.R. (1996). Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. *J. Mol. Biol.* **256**, 701-719.
17. Egebjerg, J., Larsen, N. & Garrett, R.A. (1990). Structural map of 23S rRNA, in *The Ribosome: Structure, Function and Evolution.* (Hill, W. E. *et al.* eds). pp. 168-179, American Society for Microbiology Press, Washington, DC.
18. Welch, M., Chastang, J. & Yarus, M. (1995). An inhibitor of ribosomal peptidyl transferase using transition-state analogy. *Biochemistry* **34**, 385-390.
19. Barta, A., Steiner, G., Brosius, J. & Noller, H.F. (1984) Identification of a site on 23S ribosomal RNA located at the peptidyl transferase center. *Proc. Natl Acad. Sci. USA* **81**, 3607-3611.
20. Vester, B. & Garrett, R.A. (1988). The importance of highly conserved nucleotides in the binding region of chloramphenicol at the peptidyl transfer centre of *Escherichia coli* 23S ribosomal RNA. *EMBO J.* **7**, 3577-3587.
21. Douthwaite, S. (1992). Functional interactions within 23S rRNA involving the peptidyltransferase center. *J. Bacteriol.* **174**, 1333-1338.
22. Ofengand, J., Ciesiolka, J., Denman, R. & Nurse, K. (1986). Structural and functional interactions of the tRNA–ribosome complex, in *Structure, Function, and Genetics of Ribosomes.* (Hardesty, B. & Kramer, G., eds) pp. 473-494, Springer-Verlag, New York.

23. Rodriguez-Fonseca, C., Amils, R. & Garrett, R.A. (1995). Fine structure of the peptidyl transferase centre on 23S-like rRNAs deduced from chemical probing of antibiotic-ribosome complexes. *J. Mol. Biol.* **247**, 224-235.
24. Moazed, D. & Noller, H.F. (1987).Chloramphenicol, erythromycin, carbomycin, and vernamycin B protect overlapping sites in the peptidyl transferase region of 23S ribosomal RNA. *Biochimie* **69**, 879-884.
25. Moazed, D. & Noller, H.F. (1989). Interaction of tRNA with 23S rRNA in the ribosomal A, P, and E. sites. *Cell* **57**, 585-597.
26. Noller, H.F. (1991). Ribosomal RNA and translation. *Annu. Rev. Biochem.* **60**, 191-227.
27. Leviev, I.G., *et al.* & Mankin, S. (1994). A conserved secondary structural motif in 23S rRNA defines the site of interaction of amicetin, a universal inhibitor of peptide bond formation. *EMBO J.* **13**, 1682-1686.
28. Jenison, R.D., Gill, S.C., Pardi, A. & Polisky, B. (1994). High-resolution molecular discrimination by RNA. *Science* **263**, 1425-1429.
29. Connell, G.J., Illangesekare, M. & Yarus, M. (1993). Three small ribooligonucleotides with specific arginine sites. *Biochemistry* **32**, 5497-5502.
30. Wallis, M.G., von Ahsen, U., Schroeder, R. & Famulok, M. (1995). A novel RNA motif for neomycin recognition. *Chem. Biol.* **2**, 543-552.
31. Coutsogeorgopoulus, C. (1966). On the mechanism of action of chloramphenicol in protein synthesis. *Biochim. Biophys. Acta* **129**, 214-217.
32. Zemlicka, J., Fernando-Moyano, M.C., Ariatti, M., Zurenko, G.E., Grady, J.E. & Ballesta, J.P.G. (1993). Hybrids of antibiotics inhibiting protein synthesis: synthesis and biological activity. *J. Med. Chem.* **36**, 1239-1244.
33. Drainas, D., Mamos, P. & Coutsogeorgopolous, C. (1993). Aminoacyl analogs by chloramphenicol: examination of the kinetics of inhibition of peptide bond formation. *J. Med. Chem.* **36**, 3542-3545.
34. Bhuta, P., Chung, H. L., Hwang, J.-S. & Zemlicka, J. (1980). Analogs of chloramphenicol: circular dichroism spectra, inhibition of ribosomal peptidyltransferase, and possible mechanism of action. *J. Med. Chem.* **23**, 1299-1305.
35. Spirin, A.S. & Lim, V.I. (1986). Stereochemical analysis of ribosomal transpeptidation, translocation and nacent peptide folding. In *Structure, Function and Genetics of Ribozymes* (Hardesty, B. & Kramer, G., eds), Springer-Verlag, New York.
36. Welch, M., Majerfeld, I. & Yarus, M. (1997). 23S rRNA similarity from selection for peptidyl transferase mimicry. *Biochemistry* **36:** 6614-6623.
37. Burke, D.H., Scates, L.A., Andrews, K. & Gold, L. (1996). Bent pseudoknots and novel RNA inhibitors of type 1 human immunodeficiency virus (HIV-1) reverse transcriptase. *J. Mol. Biol.* **264**, 650-666.
38. Milligan, J.F., Groebe, D.R., Witherel, G.W. & Uhlenbeck, O.D. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic. Acids. Res.* **15**, 8783.
39. Fasman, G.D. (1975). *Handbook of Biochemistry and Molecular Biology, Volume 1: Nucleic Acids.* CRC Press Inc., Cleveland.
40. Piotto, M., Saudek, V. and Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661-665.
41. Wu, M., McDowell, J.A. & Turner, D.H. (1995). A periodic table of symmetric tandem mismatches in RNA. *Biochemistry* **34**, 3204-3211.