Template Recognition by an RNA-Dependent RNA Polymerase: Identification and Characterization of Two RNA Binding Sites on Qβ Replicase†

David Brown* and Larry Gold

Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, Colorado 80309−0347

Received August 9, 1995⊗

ABSTRACT: Two different SELEX protocols were used to generate two classes of RNA ligands that bound Qβ replicase with nanomolar equilibrium dissociation constants. One set of RNAs appeared to exist as pseudoknots with conserved loop sequences. These ligands bound Qβ replicase and ribosomal protein S1 with equal affinities, indicating that the RNAs bind the replicase through its S1 subunit. The second class of ligands bound the replicase via a pyrimidine rich region. The two sets of ligands did not compete for binding to Qβ replicase, demonstrating that the two RNA families bind unique sites on the phage polymerase. Because the second class of ligands bound ribosomal protein S1 very poorly, it is likely that the second RNA binding site is located on one of the three remaining replicase subunits. Published sequences of RNAs replicated by Qβ replicase possess similarities to the two classes of RNA ligands, providing a possible solution to the question of template recognition by the phage polymerase.

Qβ replicase is the RNA-dependent RNA polymerase charged with replicating the single-stranded RNA genome of the coliphage Qβ. The enzyme consists of four subunits, only one of which is coded for by the viral genome (Kamen, 1970; Kondo *et al.*, 1970). The other three subunits are the host encoded ribosomal protein S1 (Inouye *et al.*, 1974; Wahba *et al.*, 1974) and elongation factors Tu and Ts (EF-Tu and EF-Ts) (Blumenthal *et al.*, 1972).

The mechanism for replication has been clearly established (Dobkin *et al.*, 1979). The plus strand of the viral genome is bound by Qβ replicase and used as a template for the production of a complementary strand, referred to as the minus strand. The plus and minus strands dissociate and each strand is bound by a replicase and used as a template for complementary strand synthesis in a second replication cycle. Because both strands act as templates for replication, the net result of each replicative cycle is the doubling of the previous amount of RNA providing for the exponential amplification of the plus and minus strands.

Replication by Qβ replicase requires that the RNA be bound in such a manner as to introduce the 3′ end of the template into the polymerization site of the replicase. One model holds that an internal site on the RNA molecule is recognized by an RNA binding site on the enzyme and that the tertiary structure of the RNA specifically introduces the 3′ end of the RNA into the polymerization site of the replicase allowing for the initiation of complementary strand synthesis (Weissmann, 1974). In support of this model are experiments that have shown that internal regions of both the plus and minus strands of the phage genome are essential for *in vitro* replication (Schwyzer *et al.*, 1972; Schuppli *et al.*, 1994).

The plus and minus strands of the Qβ genome have different enzyme requirements for replication. Utilization of the plus strand as a template requires the intact holoenzyme (S1, EF-Tu, EF-Ts, and the virus encoded subunit) plus a fifth protein known simply as host factor (Spiegelman *et al.*, 1968). Because S1 alone has been shown to bind a site on the plus strand that seems to be essential for replication, it has been proposed that S1 possesses the RNA binding site that provides template recognition prior to complementary strand synthesis (Weber *et al.*, 1974; Senear & Steitz, 1975). In contrast, the minus strand of Qβ can be replicated *in vitro* by a replicase preparation lacking not only the host factor but ribosomal protein S1 as well (Kamen *et al.*, 1972). This observation suggests that an RNA binding site for minus strand recognition exists on one of the three remaining subunits.

The specific RNA features that are recognized by Qβ replicase are unknown; thus, it is not clear how the replicase selects templates for replication. SELEX[1] (Systematic Evolution of Ligands by EXponential enrichment) (Tuerk & Gold, 1990b), a procedure that has been used to isolate high affinity nucleic acid ligands to a variety of targets, was used to characterize the two putative RNA binding sites of Qβ replicase. Two variations of the SELEX protocol generated two classes of RNA ligands that bound with high affinities to the multisubunit polymerase. The selections were done in the absence of nucleotides so as to select RNAs on the basis of their ability to bind the replicase and not on their capacity to act as replication templates. Comparison of the two classes of SELEX-generated ligands to RNAs that serve as templates for replication reveal similarities that provide insights into template recognition by the phage enzyme.

MATERIALS AND METHODS

Materials. Qβ replicase, at a concentration of 200 μg/mL, was a generous gift from Gene-Trak, Inc. T7 RNA

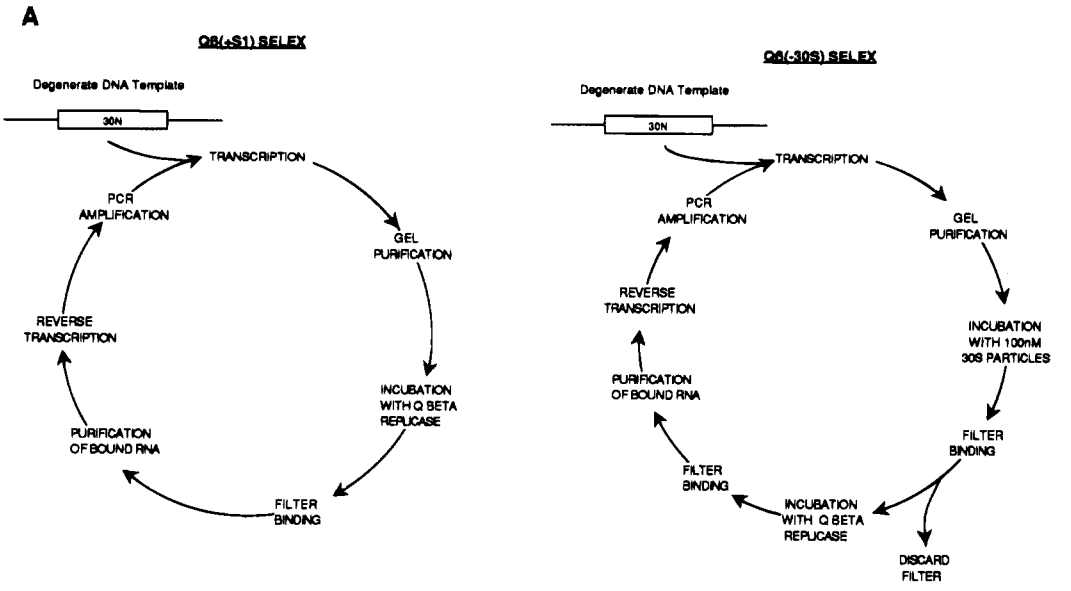
† This research was funded by research grants from National Institutes of Health to L.G. (GM28685 and GM19963).

* Corresponding author.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1995.

[1] Abbreviations: SELEX, Systematic Evolution of Ligands by EXponential enrichment; S1, ribosomal protein S1; K_d, equilibrium dissociation constant; PCR, polymerase chain reaction.

0006-2960/95/0434-14765$09.00/0



B

DNA template
5'-GCCGGATCCGGGCCTCATGTCGAA [30N] TTGAGCGTTTATTCTGAGCTCCC-3'

5' primer
5'-CCGAAGCTTAATACGACTCACTATAGGGAGCTCAGAATAAACGCTCAA-3'

3' primer
5'-GCCGGATCCGGGCCTCATGTCGAA-3'

RNA
GGGAGCUCAGAAUAAACGCUCAA [30N] UUCGACAUGAGGCCCGGAUCCGC

FIGURE 1: Overview of the SELEX protocols and molecules used in the experiments. (A) Schematics of the two SELEX protocols used for isolating high affinity RNA ligands to the two nucleic acid binding sites on Qβ replicase. (B) DNA template used to generate the initial degenerate RNA populations used in the selections. Positions denoted by N indicate that, on the basis of the method used for synthesizing the deoxyoligonucleotide, there is an equal likelihood of that position being any one of the four nucleosides, thus resulting in a randomized position. The 3′ primer was used for reverse transcription of the RNA populations and for PCR amplification in combination with the 5′ primer. The 5′ primer contains a T7 RNA polymerase promoter used for transcription of the PCR product.

polymerase was supplied by NeXstar Pharmaceuticals, Inc. Ribosomal protein S1 was kindly provided by Dr. A. R. Subramanian at the Max-Planck Institute in Berlin. The 30S subunits of *Escherichia coli* ribosomes were purified as described (Boni *et al.*, 1982). Oligonucleotides were purchased from Operon Technologies, Inc.

Production of the Randomized RNA Library. Fifty picomoles of the DNA oligonucleotide shown in Figure 1B was converted to double-stranded DNA by the polymerase chain reaction (PCR) using the 5′ and 3′ primers (Figure 1B). The DNA template was transcribed by T7 RNA polymerase, yielding the degenerate RNA population used for the selections.

Qβ(+S1) Selection. One hundred picomoles of RNA from the initial, mixed sequence population was used to initiate the Qβ(+S1) selection (Figure 1A). The RNA was incubated for 5 min at 37 °C in a 100 μL reaction containing 100 nM Qβ replicase and replicase buffer [84 mM Tris (pH 7.4), 12 mM $MgCl_2$]. Those RNAs that bound the replicase were separated from the nonbound RNAs by filtering the binding reaction through a prewet nitrocellulose filter (Millipore 25-mm HAWP). The filter was washed with 5 mL of Qβ replicase wash buffer [84 mM Tris (pH 7.4), 12 mM $MgCl_2$, 200 mM NaCl]. The selected RNAs were removed from the filter, reverse transcribed, PCR amplified, and transcribed as described (Tuerk & Gold, 1990b) using the 5′ and 3′ primers shown in Figure 1B. This procedure was repeated for 12 rounds, with the stringency of the selection being increased in successive rounds by decreasing the protein concentration. The PCR product from the twelfth round was ligated into pUC 18 using the *Bam*HI and *Hin*dIII sites, the plasmids were used to transform DH5α, and the resultant clones were sequenced (Chen *et al.*, 1996).

Qβ(−30S) Selection. The second selection, referred to as the Qβ(−30S) selection, was identical to the Qβ(+S1) SELEX protocol except that, prior to binding the RNA to Qβ replicase, the RNA was incubated with 100 nM 30S particles in a total volume of 300 μL for 5 min and the ribosome-bound RNAs were removed from the population via nitrocellulose filtration (MFS-25 purchased from Life Science Products, Inc.). The RNAs that remained in the filtrate were then used in a cycle of replicase binding, reverse transcription, PCR amplification, and transcription as with the Qβ(+S1) selection above. The counterselection/selection protocol was repeated for 11 cycles. The selected populations were cloned and sequenced as above.

Equilibrium Binding Assays. Equilibrium dissociation constants (K_d's) of the various RNA populations and individual clones were determined using the filter binding assay (Carey *et al.*, 1983). γ-^{32}P-Labeled RNA (50 pM) was incubated in 20 μL reactions containing replicase buffer and Qβ replicase varying in concentration from 300 nM to 40

pM for 5 min at 37 °C. The binding reactions were vacuumed through nitrocellulose filters and the filters were washed with 5 mL of Qβ replicase wash buffer. The filters were dried under an infrared lamp, and the protein-bound RNA was quantitated by scintillation counting.

Boundary Determination. 5′ and 3′ boundaries for the Qβ-(+S1) ligands were determined essentially as described (Tuerk *et al.*, 1990a). Five picomoles of partially hydrolyzed 5′-(kinased with [γ-^{32}P]ATP) or 3′-(^{32}P pCp ligated) end-labeled RNA was incubated with 7 pmol of Qβ replicase and 250 pmol of tRNA for 5 min at 37 °C in either 350 μL or 1.05 mL of replicase buffer. The two volumes were used to provide two different protein concentrations without altering the total RNA or replicase available for binding. The bound RNA was separated from the unbound species by nitrocellulose filter binding. The RNA was extracted, precipitated, and lyophilized. The pelleted RNA was dissolved in 10 μL of formamide, and the RNA fragments were separated on an 8% acrylamide/7 M urea gel. The nuclease digested RNA used for aligning the sequence in the boundary experiments was incubated in nuclease buffer [20 mM sodium citrate (pH 5.0), 1 mM EDTA, 7 M urea] with 1 unit of RNase T1 for 10 min at 60 °C (Donis-Keller *et al.*, 1977).

Deletion Analysis of Qβ(−30S) Ligands. A series of short oligonucleotides were used in concert with either the 5′ or 3′ primer presented in Figure 1B in PCR reactions to yield truncated DNA templates. Transcription of the templates by T7 RNA polymerase produced Qβ(−30S)a and Qβ(−30S)b ligands with terminal deletions. The templates were transcribed as described except that 200 μM NTPs plus 2.5 μM [α-^{32}P]GTP were used. The transcripts were purified as described (Chen *et al.*, 1996), and their K_d's for Qβ replicase were measured as above.

Structure Probing: Nuclease Sensitivity. 5′-end-labeled Qβ(−30S)a and Qβ(−30S)b were incubated at 37 °C in replicase buffer for 5 min in the presence of 1 unit of either ribonuclease S1 or T1. The RNA products were then separated on an 8% acrylamide/7 M urea gel and visualized by autoradiography. An RNA sequencing lane was produced for each of the RNAs as described in the *Boundary Determination* section.

Chemical Modification. Unlabeled Qβ(−30S)a and Qβ-(−30S)b were incubated in replicase buffer in the presence of the single-strand specific base-modifying reagents 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide metho-*p*-toluenesulfonate (CMCT), dimethylsulfate (DMS), and kethoxal for 7 min at 37 °C. The reactions were stopped as described (Stern *et al.*, 1988), and the RNAs were precipitated, centrifuged, and lyophilized. The modified RNAs were reverse transcribed by AMV-RT using the 5′-end-labeled 3′ primer shown in Figure 1B. The reverse transcription products were separated on an 8% acrylamide/7 M urea gel and visualized by autoradiography.

Equilibrium Competition Experiments. Competition for binding between the various ligands was measured by observing the radioactively labeled RNA bound to Qβ replicase as a function of the quantity of cold competitor RNA present. The pseudoknot truncate (PT), Qβ(−30S)a-[5′c3′b], and Qβ(−30S)b[5′b3′a] were transcribed in the absence of radiolabeled NTPs. The RNAs were purified as above. Approximately 50 pmol of each RNA was dephosphorylated by calf intestinal phosphatase and 5′-end-labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase. The labeled RNAs were then size purified by PAGE.

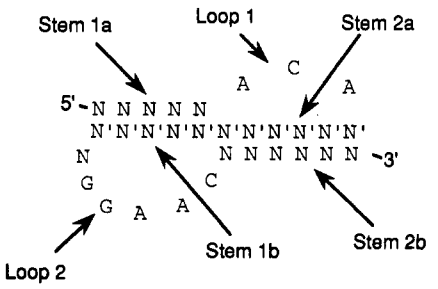
A

Class I

# of Clones	Stem 1a	Loop 1	Stem 2a	Stem 1b	Loop 2	Stem 2b
14	CACAG	ACA	CGAAUA	CUGUG	UGGAAC	UGuucg
12	UGCUG	ACA	CGAAGC	CAGCA	CGGAAC	GCuucg
11	gcucaa	CCC	AUGGAC	UUGAGU	AGGACC	GUCCGU
5	UGCUG	ACA	CGAACC	CAGCA	UGGAAC	GGuucg
4	UGGAG	ACA	GAGUGC	CUCCA	AGGAAC	GCAuuc
3	GCAG	ACA	GAACAC	CUGC	AGGAAC	GUGuuc
3	GCAG	ACU	GAAUCA	CUGC	UGGAAU	UGAuuc
3	GCAG	ACU	GAAGCUA	CUGC	GAUGG	UAGCuuc
2	CCUAG	ACA	CGAACA	CUAGG	AGGAAC	UGuucg
2	GACUG	ACA	AAUCGA	CAGUC	UGGAAC	UCGAuu
2	UAGUG	ACA	GAACAG	CACUA	UGGAAC	CUGuuc
2	ACUG	ACA	AAUGA	CAGU	CUGGAC	UCAUU

B

Class I Consensus Structure



C

Class II

# of clones	
2	CA**UCUCCUUCUCACCUUC**GGCACCAUCGAG
2	UGCCUCUUGAGUA**CUUUGAUCGUCGU**AGUG
2	GGGAAAGUCGAUUUGA**CUGCUUUUGCUUC**

FIGURE 2: Two classes of ligands resulting from the Qβ(+S1) SELEX protocol. Those sequences that appeared in multiple clones of the 100 that were sequenced are shown with the number of clones possessing each sequence shown in the left hand column. Only the portions of the selected RNA ligands that are predicted to form the pseudoknot structure present in the class I molecules are shown in panel A. Capital letters indicate positions that were in the random region; small letters are from the fixed regions. (B) diagram of the predicted consensus pseudoknot. The various components used to describe the class I sequences in panel A are labeled on the pseudoknot structure in panel B. The top two molecules were used for further characterization. The top molecule is Qβ(+S1)a and the second molecule is Qβ(+S1)b. (C) 30N region of the three class II ligands that were present in multiple clones. The C/U-rich regions used to classify the molecules are in bold-face letters.

For competition with the pseudoknot truncate, 40 nM radiolabeled PT was mixed with each of the nonlabeled truncates whose concentrations ranged from 640 to 5 nM in 10 μL of replicase buffer. Ten microliters of 20 nM Qβ replicase in replicase buffer was then added to each of the reactions above, resulting in the following final concentrations: 20 nM radioactive PT, 2.5−320 nM cold competitor, and 10 nM Qβ replicase. The 20 μL reactions were incubated at 37 °C for 10 min and then vacuum filtered through prewet nitrocellulose discs. The filters were washed with 5 mL of replicase wash buffer and dried under an infrared lamp, and the bound radiolabeled RNA was quantitated by scintillation counting. Competition experiments with labeled Qβ(−30S)a[5′c3′b] and Qβ(−30S)b[5′b3′a] were performed exactly as above except that the final concentrations were 40 nM radiolabeled Qβ(−30S)a[5′c3′b] or Qβ(−30S)b[5′b3′a], 5−640 nM cold competitor RNA, and 10 nM Qβ replicase.

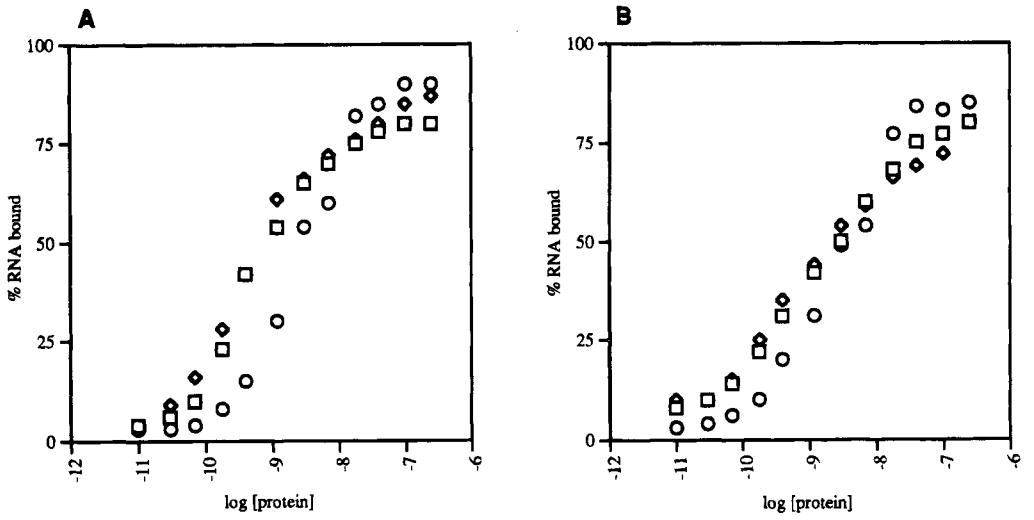


FIGURE 3: Binding properties of Qβ(+S1)a and Qβ(+S1)b. Boxes indicate binding to Qβ replicase, diamonds to the 30S subunit of *E. coli* ribosomes, and circles to ribosomal protein S1. Panel A shows the binding of Qβ(+S1)a; panel B, the binding of Qβ(+S1)b.

RESULTS

Qβ(+S1) Selection. The first SELEX experiment, referred to as the Qβ(+S1) selection, was used to identify RNA ligands that bound with the highest affinity to Qβ replicase. A population of 10^{14} molecules comprising approximately 10^{13} unique sequences was challenged with binding the multisubunit replicase. Twelve rounds of selection were used to reduce the RNA population to those sequences of the 10^{13} that had the highest affinity for the polymerase. The stringency of the selection was increased by stepwise decreases in the replicase concentration from 100 nM in round one down to 9 nM by round 11. Binding curves for several of the rounds indicated that the K_d's of the populations for the replicase dropped consistently from rounds six through ten, going from 300 nM in the early rounds to 5 nM for the final two rounds (data not shown). The PCR product from round 12 was cloned, and 100 of the clones were sequenced.

Careful analysis of the primary sequences of the selected molecules revealed that 83% of the ligands could form a pseudoknot with conserved primary sequences in both loop 1 and loop 2 (class I molecules shown in Figure 2). The nucleotides in the two helical regions showed no conservation, indicating that this portion of the molecules was probably not involved in making base-specific contacts with the protein but was merely a structural component. The observed consensus structure was also seen in two previous SELEX experiments, one whose target was the 30S subunit of *E. coli* ribosomes (30S particles) and the second whose target was ribosomal protein S1 (S1) (Ringquist *et al.*, 1995). A second family of RNAs, referred to as the class II ligands, was also identified. These ligands possessed no distinct structural features but were grouped based on the presence of extended C/U rich regions (Figure 2C).

The two most prevalent class I molecules [Qβ(+S1)a and Qβ(+S1)b] were chosen for characterization (Figure 2A). Filter binding assays were done to determine the dissociation constants of these molecules for Qβ replicase. Binding curves to ribosomal protein S1 and 30S particles were performed as well. K_d's of both RNAs to the replicase were approximately 1 nM (Figure 3). That was also the K_d of the ligands for 30S particles and for S1 alone (Figure 3). The fact that the lone unifying feature among the three macromolecule preparations is the presence of ribosomal protein S1 implies that these RNA molecules bind to Qβ replicase via an RNA binding site located on the S1 subunit.

Boundary experiments were performed to determine if the proposed pseudoknot observed in the sequence comparative analysis was actually the element required for binding. Figure 4 shows the results of the 5′ and 3′ boundaries of Qβ(+S1)b. The two boundaries fall precisely at the ends of the pseudoknot. Likewise, boundary experiments on Qβ-(+S1)a revealed that the putative pseudoknot was indeed the structure within the RNA molecule that was necessary for binding (data not shown).

An RNA molecule possessing the consensus pseudoknot structure, but lacking the flanking regions present in the selected molecules, was produced to determine if the pseudoknot alone was sufficient for binding of Qβ replicase. Figure 5A displays the sequence and predicted structure of the truncated RNA (referred to as the pseudoknot truncate, or simply PT), and Figure 5B shows its binding profile to both Qβ replicase and 30S particles. The K_d of the pseudoknot was 3 nM, revealing that the pseudoknot alone was able to bind with high affinity to Qβ replicase.

The binding properties of the three class II ligands shown in Figure 2C were tested. Each of these molecules bound Qβ replicase with a K_d of approximately 20 nM (data not shown). In contrast to the class I ligands, these RNAs bound 30S particles and ribosomal protein S1 with K_d's of greater than 200 nM (data not shown), suggesting that they were binding at a second site on the replicase. Because the minus strand of the Qβ genome does not require the S1 subunit for replication, it was hoped that the class II ligands were binding to the site on the replicase required for minus strand recognition. There were too few of these molecules to do a thorough comparative analysis; thus a second selection was designed to enrich for the class II RNA ligands.

Qβ(−30S) Selection. The second selection relied on the same protocol as the first, except that a counterselection step was added to reduce the fraction of the population that was binding ribosomal protein S1. The 30S subunits of *E. coli* ribosomes were used to remove the S1 binding RNAs from the SELEX populations. The RNA molecules were incu-

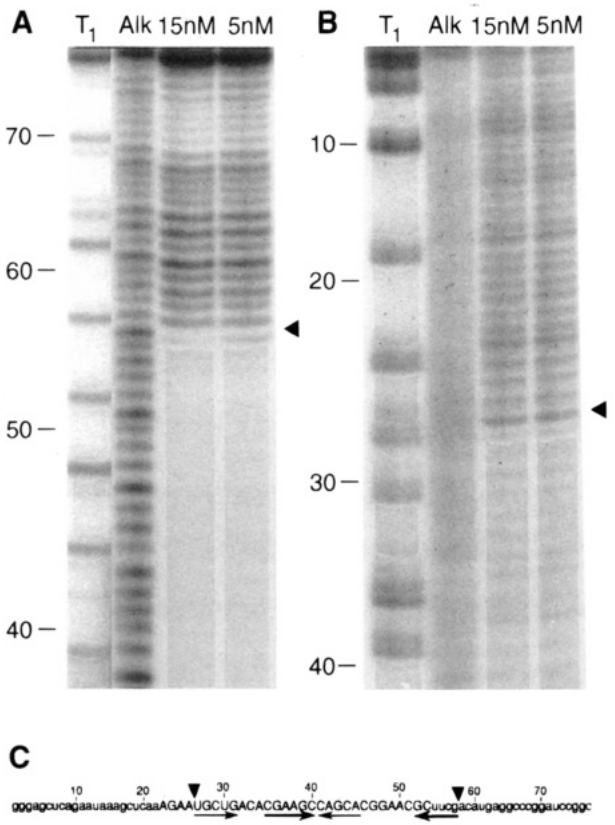


FIGURE 4: 5′ and 3′ boundaries of Qβ(+S1)b. (A) 3′ boundary analysis; (B) 5′ boundary analysis. The lanes for both gels are the same: T1 is the nuclease-digested RNA providing the positions in the RNA that are G's, alk is the alkaline digested RNA, and 15 nM and 5 nM are the lanes of alkaline digested fragments that were bound by 15 nM and 5 nM Qβ replicase, respectively. The numbers to the right of each gel represent the nucleotide position in the RNA located at that site in the gel. (C) Sequence of Qβ(+S1)b with the 5′ and 3′ boundaries indicated by triangles above the sequence. The arrows indicate the proposed base pairing between nucleotides in the RNA yielding the putative pseudoknot. The nucleotides above the broad arrows base-pair with one another, and the nucleotides above the narrow arrows base-pair.

bated with the 30S subunits, and the bound RNAs were removed from the population by nitrocellulose filtration. The RNAs in the filtrate were then used to select the highest affinity binders to Qβ replicase. Although the 30S particles were expected to remove RNAs in addition to the S1 binding ligands, the ribosomal particles were used instead of S1 alone because 30S particles were more stable and consistent in binding the RNAs (personal observation).

The RNA from the second round of the Qβ(+S1) selection was used to initiate the Qβ(−30S) SELEX protocol. Eleven additional rounds of counterselection/selection were used to enrich the RNA pool for molecules that bound to Qβ replicase without using the S1 binding site previously identified. The K_d's of the pools for replicase dropped from approximately 200 nM in the early rounds down to 20 nM in round 13 (the same RNA populations displayed K_d's to 30S particles of 200 nM in the earlier rounds to greater than 300 nM in the latter rounds, indicating that the protocol was successful) (data not shown). Because it was unclear how the counterselection step would affect the RNA population during the SELEX experiment, several different rounds were chosen for cloning. Sixteen clones from round nine, and 32 clones from each of rounds 10, 12, and 13 were sequenced.

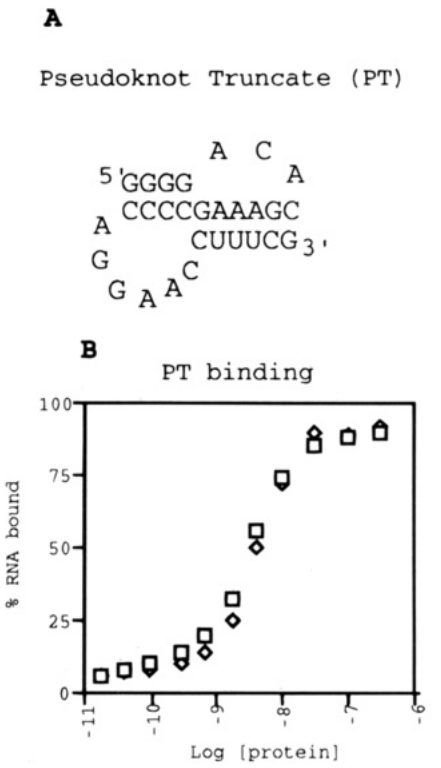


FIGURE 5: Binding properties of the truncated pseudoknot used to show that the structure predicted by sequence analysis was sufficient for binding Qβ replicase. (A) Sequence and predicted structure of the molecule designed to mimic the pseudoknots of the full-length selected RNA ligands isolated in the Qβ(+S1) SELEX protocol. Panel B shows that the binding of the small molecule is very similar to that of its full-length counterparts. Squares indicate binding to Qβ replicase; diamonds, binding to the 30S subunit of *E. coli* ribosomes.

Class I

# of Clones			
1	UG	ACCACUAAUCAAACGGAUUAGCCCA	
1	GU	AGCAAACGGUUACGACGUGCCCAUA	UGUC
1	GCAUG	AAUACAAACACACGUCGCUGACAAC	U

Class II

13	UCGUUG	CCUCUUGUU	AUGAUCGUGGUAUGC
2	UCCCCGAAA	CCCUAUUUCUU	GCUGCUA
2	UUUGCG	UUUGCGCUU	GUACGCCUUUAAUGU
1	CCUCUA	CUCCAUCUUUCCCC	GAUUGCAAGCC
1	CCAA	UUUUUAUUACACC	AGAAGGCGCCAG
1	UCA	UCUCUCCAUGUU	AUGACUGUAGUAUUC
1	UCCAG	UCUGCUCCAUUGUU	AGACUGUGGUAUGC
1	AAGUUG	CCCACCGUUGUCC	AAUUGAUCGUC
1	AGCCAG	UUUUCAUUGCUGUUCC	ACUGUAAA
1	UGCA	CUCUCCUCUUCCUC	AACACUACUCUC
1	CUG	CCGUUCUCCGCACUC	AGUCUGCGGACAUC
1	CACAUCGACA	UUACUGUACCCC	AGGCCUG
1	ACCAACGUCGA	UGUUUACAACCCU	GAGUA
1	CGCAACUGAGUA	CCACCCUUC	AA
1	UCGAUAUGACCG	CUUUUGCUUC	

FIGURE 6: Sequences of the 30N regions of the molecules cloned from round 13 of the Qβ(−30S) selection. The column on the left provides the number of clones of the 32 that were analyzed that possessed the particular sequence. Class I molecules are dominated by adenosines and cytosines. Class II molecules are distinguished by an extended polypyrimidine tract. The top three sequences of class II were characterized. They are, from the top to the bottom, Qβ(−30S)a, Qβ(−30S)b, and Qβ(−30S)c.

Two families of RNA molecules were identified by the Qβ(−30S) selection (Figure 6 provides the sequences of the round 13 clones). The class I molecules displayed a preference for A's and C's, a trait not unlike that of the loops of the class I pseudoknots from the Qβ(+S1) selection.

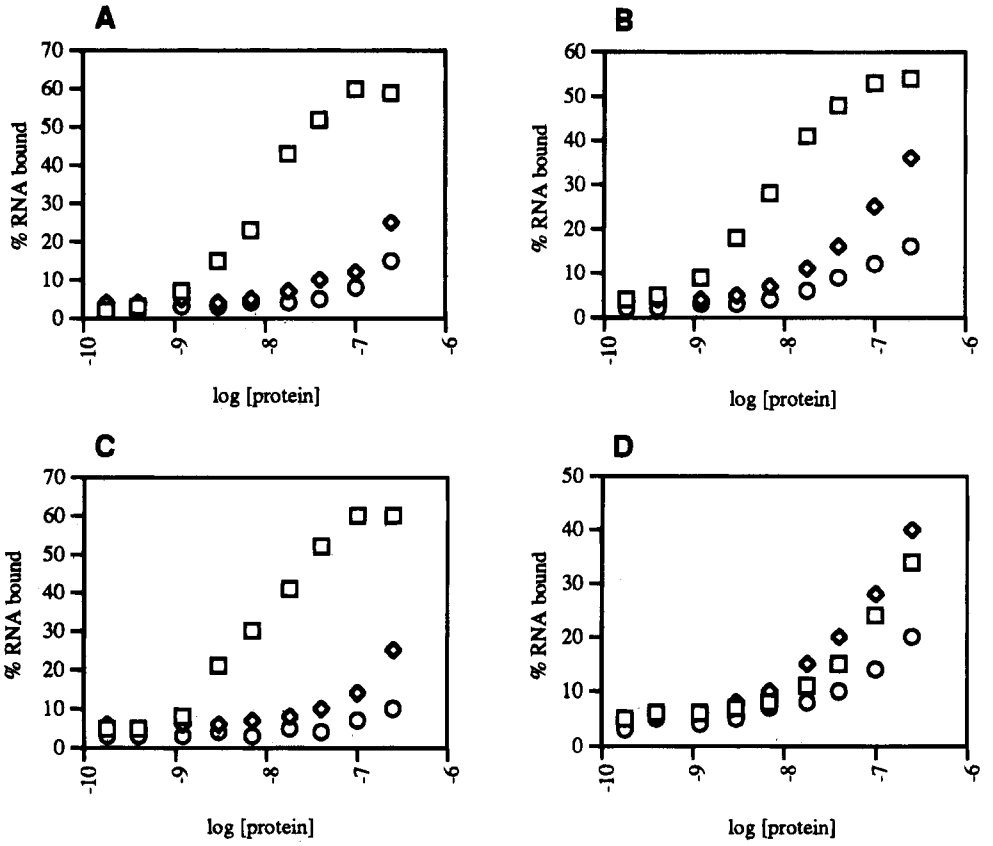


FIGURE 7: Binding properties of the Qβ(−30S) ligands and the initial, randomized population from which both the Qβ(+S1) ligands and Qβ(−30S) ligands were isolated. (A) Binding of Qβ(−30S)a; (B) binding of Qβ(−30S)b; (C) binding of Qβ(−30S)c; (D) binding of the 30N RNA population. Squares indicate binding of the labeled RNAs to Qβ replicase, diamonds to the 30S subunit of *E. coli* ribosomes, and circles to ribosomal protein S1.

Unlike the pseudoknot family, the class I ligands from the Qβ(−30S) selection did not possess a consensus secondary structure. The distinguishing characteristic of the class II RNAs was the presence of a polypyrimidine tract of at least nine nucleotides. A similar motif existed in the class II ligands from the Qβ(+S1) selection (Figure 2C). All four of the rounds that were sequenced provided ligands from both classes, but the overall percentage of class II ligands increased in the latter rounds (data not shown).

The binding of the three class I molecules from round 13 of the Qβ(−30S) selection was characterized. Each of the molecules bound to both Qβ replicase and 30S particles with K_d's of approximately 20 nM (data not shown). That, coupled to the fact that the class I molecules of the Qβ(−30S) and Qβ(+S1) selections competed for binding to Qβ replicase (data not shown), implied that the two sets of ligands bound the same site on the replicase. The difference in binding affinity between the ligands probably resulted from their respective structural stabilities (Gold *et al.*, 1995).

Three of the class II molecules from the Qβ(−30S) selection, Qβ(−30S)a, Qβ(−30S)b, and Qβ(−30S)c, were transcribed and radioactively labeled. The equilibrium dissociation constants of the RNAs for Qβ replicase, the 30S ribosomal subunit of *E. coli*, and ribosomal protein S1 were determined using the filter binding assay. The K_d's of the three RNAs for Qβ replicase were approximately 10 nM (Figure 7). As with the class II ligands from the Qβ(+S1) selection, these three molecules displayed little affinity for ribosomal protein S1, either alone or in the context of the 30S particles. Although the conformation of S1 in Qβ replicase might be different than when it is free or as a part of 30S particles providing an RNA binding site that is unique to the phage protein, that scenario is improbable. A more likely explanation is that the class II ligands are binding at a second RNA binding site on Qβ replicase that is provided by the viral subunit, or one of the two elongation factors, Tu or Ts.

Most of the currently known RNA binding proteins recognize their targets by specific structures that include vital nonpaired nucleotides. Such a structure was anticipated to be a part of the recognition element for the class II RNA binding site. A computer algorithm (Zuker, 1989) was used to predict secondary structures of the class II RNAs sequenced from the round 13 pool. The predicted structures were compared, but, surprisingly, there were no extensively shared structural motifs among the ligands. The conserved C/U rich regions were predicted to have a wide variety of structural contexts: at the bases of helices, at the tops of helices, as single-stranded loops, shared between two helices, and in single-stranded regions separating hairpins. The only shared observation was that at least part of the polypyrimidine tract was predicted to be in a single-stranded region.

Single-strand specific nucleases and base-modifying chemicals were used to study the secondary structures of Qβ(−30S)a and Qβ(−30S)b. As predicted by the folding programs, the two structures were quite dissimilar (Figure 8). Qβ(−30S)a has a fairly stable hairpin that dominates the random region of the RNA. The first four nucleotides of the C/U motif are unpaired. In contrast, most of the random region of Qβ(−30S)b was sensitive to the single-strand-specific reagents used in the structure probing experiments.

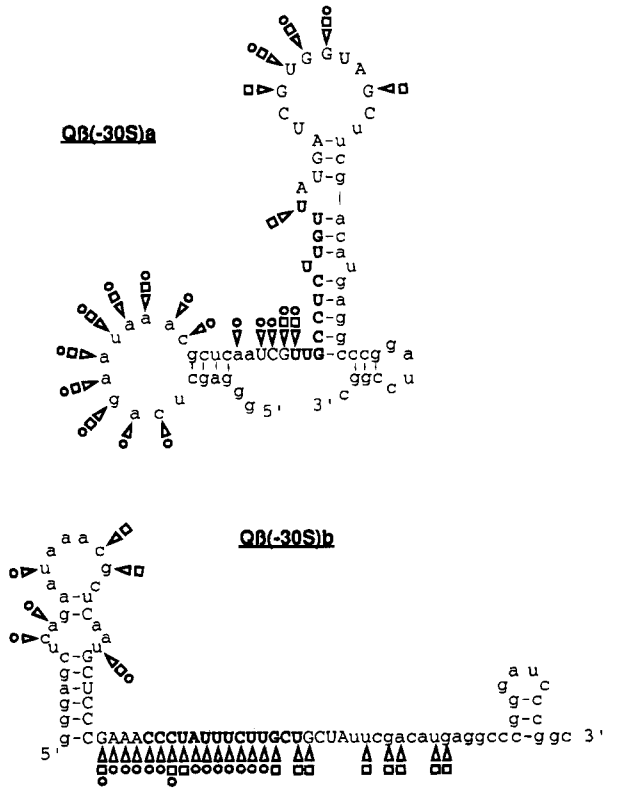


FIGURE 8: Secondary structures of the Qβ(−30S) ligands that most completely accommodates the nuclease and chemical modification data. Triangles represent nucleotide positions that were sensitive to any one of the single strand specific reagents that were used. Circles represent bases that were sensitive to CMCT, DMS, or kethoxal. Squares indicate the positions that were immediately 5′ to phosphate backbone breaks induced by either ribonuclease S1 or T1. Nucleotides from the 5′ and 3′ fixed regions are represented by small letters; nucleotides from the random region are indicated by capital letters. The C/U-rich regions are in bold type.

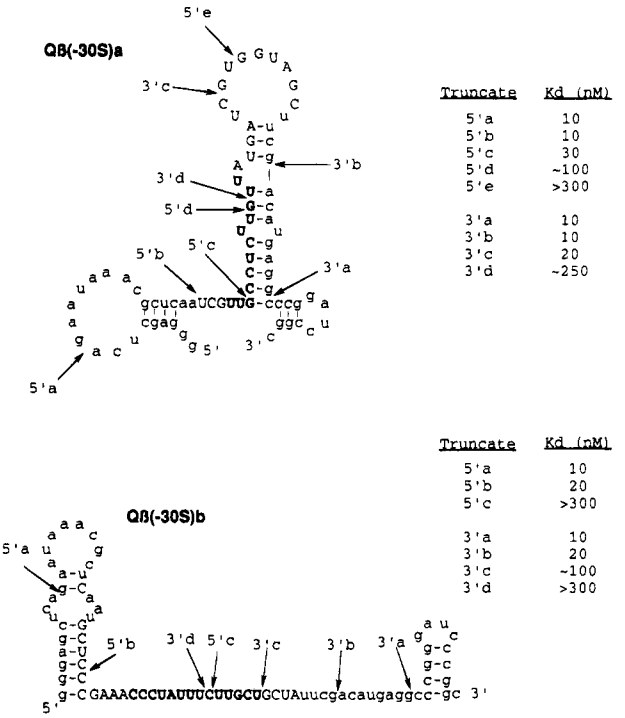


FIGURE 9: Terminal deletion analyses of Qβ(−30S)a and Qβ(−30S)b. 5′ deletions (depicted by arrows with 5′a−e) and 3′ deletions (arrows with 3′a−d) were used to determine those regions of the RNA ligands that were essential for binding to Qβ replicase. Each of the deletions is from the nucleotide at the end of the arrow back to either the 5′ or 3′ end (the 3′ designation indicates that the 3′ end was deleted up to the site; 5′ indicates that the 5′ end was deleted). For the 5′ deletions, the molecules include a GGG at the 5′ terminus to aid transcription by T7 RNA polymerase. The K_d's of the various truncated RNAs are presented in the columns to the right of each structure. The C/U-rich regions are in bold type.

The polypyrimidine region of the second RNA was predicted to be completely unpaired.

A deletion analysis of Qβ(−30S)a and Qβ(−30S)b was undertaken to determine the minimal RNA domain required for binding. For both ligands, it was immediately apparent that the binding domains on the RNAs were independent of the secondary structures predicted by the nuclease and chemical modification experiments (Figure 9). The only reduction in binding on either of the two ligands occurred when the C/U rich regions were partially or completely deleted. Thus the second RNA binding site of Qβ replicase identified by the SELEX experiments appears to be a very simple one, displaying a preference for pyrimidines.

Further characterization of the Qβ(−30S) ligands was done with truncated versions of Qβ(−30S)a and Qβ(−30S)b. The Qβ(−30S)a truncate, referred to as Qβ(−30S)a[5′c3′b], spanned the region from 5′c to 3′b on the molecule (Figure 8). Qβ(−30S)b[5′b3′a], the truncated version of the Qβ(−30S)b ligand, consisted of the region between the 5′b and 3′a deletion sites (Figure 8). The binding curves of the two truncated molecules are nearly identical to those of the full-length RNAs, indicating that, like the Qβ(+S1) pseudoknot truncate, the deletion of the 5′ and 3′ flanking regions of the Qβ(−30S) ligands has little effect on the binding affinity of the RNAs (data not shown).

It was important to understand if the class I and class II ligands bound to unique sites on the replicase, or if they merely represented two families of molecules that bound the same site. To determine this, competition experiments were done. The design involved measuring the binding of a radioactively labeled RNA to Qβ replicase in the presence of varying amounts of nonlabeled competitor RNA. If the two molecules bound at the same site on the protein, then the amount of labeled RNA bound to the replicase would decrease with increasing amounts of competitor. Figure 10 shows the results of binding competitions between truncated versions of the pseudoknot and the two Qβ(−30S) class II ligands. Increasing levels of the unlabeled pseudoknot truncate (PT) decreased the amount of labeled PT that was bound by the polymerase, indicating that the ligand competes with itself for binding (Figure 10A). Inclusion of the unlabeled class II truncated ligands had no affect on the binding of the labeled PT (Figure 10A). Qβ(−30S)a[5′c3′b] and Qβ(−30S)b[5′b3′a] competed with themselves and one another for binding to the protein, but PT had little effect on the binding of either of the ligands (Figure 10B,C). Comparison of the theoretical curves for same-site binding between the two ligand classes and the actual curves that were measured display a disparity that can only be explained by the two families binding at different sites on the polymerase.

DISCUSSION

Experimental evidence has indicated that Qβ replicase specifically recognizes internal regions of the RNA molecules that it replicates. We used SELEX to generate high affinity

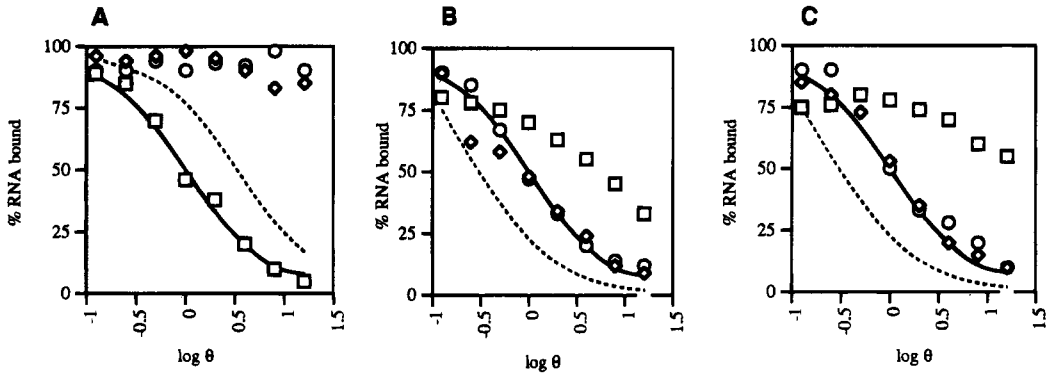


FIGURE 10: Binding competition experiments used to determine if the two ligand families, Qβ(+S1) and Qβ(−30S), bind to the same site on Qβ replicase. Panel A is a measure of the labeled PT bound in the presence of increasing amounts of competitor RNA. Panel B shows a similar curve for labeled Qβ(−30S)a[5′c3′b], and panel C is the binding of labeled Qβ(−30S)b[5′b3′a] versus competitor RNAs. The squares indicate addition of unlabeled PT; the diamonds, unlabeled Qβ(−30S)a[5′c3′b]; and the circles, unlabeled Qβ(−30S)b[5′b3′a]. The solid lines in each of the graphs are theoretical curves for competition between ligands binding at the same site with identical K_d's. The dashed line in panel A is the theoretical binding of labeled PT in the presence of either of the unlabeled Qβ(−30S) truncates if the two bound at the same site based on K_d's of 3 nM for PT and 10 nM for the two competing RNAs. The dashed lines in panels B and C are theoretical curves for same-site competition between the labeled Qβ(−30S) ligands and unlabeled PT using the same K_d's as above.

RNA ligands to the phage protein to better understand the specific requirements for recognition. Two classes of RNA ligands that bound Qβ replicase with nanomolar dissociation constants resulted from two different SELEX procedures. The class I ligands, which have a high fraction of unpaired A's and C's, apparently bind the S1 subunit of the replicase. The class II ligands were dominated by pyrimidine-rich regions that were essential for binding. The two classes did not compete for binding to Qβ replicase, indicating that the two sets of ligands bind unique sites on the replicase.

Although deletion analyses indicated that the class II ligands bound Qβ replicase in a base-specific, structure-independent manner, many of the putative binding regions of the RNAs were predicted to be involved in intramolecular base pairing. Two mechanisms for disrupting the helices and allowing for binding of the replicase to the polypyrimidine motifs could be imagined. Because at least part of each of the C/U rich regions are predicted to be single-stranded, it is conceivable that the replicase could initiate binding to the RNA at the single-stranded portion of the polypyrimidine tract and then increase its contact with the C's and U's as the base-pairs in the helix relax and unpair. Ultimately, the entire hairpin could be melted, leaving the C/U rich region bound by the replicase. A second possible means for converting the helical regions into "binding friendly" single strands involves the base unstacking activity observed in the S1 subunit (Bear *et al.*, 1976). The S1 subunit could unwind the helical regions of the RNAs and allow the second RNA binding site of Qβ replicase to bind the polypyrimidines. Weak evidence supporting the second mechanism comes from the variable effects that the S1 targeted ligand had on the binding of the two class II ligands in the competition experiments shown in Figure 10. In these experiments, PT (the ligand that bound the S1 subunit) had a greater effect on the binding of the structured class II ligand (Figure 10B) than it had on the binding of the unstructured ligand (Figure 10C).

Significance for Replication. Comparison of the published replicatable sequences with both sets of ligands isolated by the SELEX experiments provides interesting similarities. The plus strand of the Qβ genome, which requires S1 for replication, has a primary sequence that is very similar to the loop sequences of the Qβ(+S1) pseudoknot ligands, as

A

Loop consensus from Qß(+S1) SELEX		loop2 loop1 NGGAAC ACA
Class I ligands from Qß(-30S) SELEX		ACAAAC ACAC AAUC AAACGGAUUA
Qß plus strand		AAUAAAUUAUC ACAAUUACUCUUACG

B

Qß minus strand		^{4166}CCUUCGUGCCCU^{4177}
MDV-1	(+)	^{77}CUUUCGCGCUCUCCC^{91}
	(-)	^{103}UCGCGCCUCUUC^{114}
MICROVARIANT	(+)	^{34}UCUCCUCCU^{42}
	(-)	^{26}CUUUGGUUUU^{35}
NANOVARIANT	(+)	^{30}UUUUCUCACCUCUCU^{44}
	(-)	^{28}UCCUCUCUAACUUUC^{42}
RQ 120	(+)	^{76}CUCGUUCCUU^{85}
	(-)	^{45}CUUUUGUUCGU^{55}
MNV-11	(+)	^{31}CCUUUUUCCCUC^{42}
	(-)	^{7}CCCCCCUUC^{15}
CT	(+)	^{38}UCCUCCU^{44}
	(-)	^{39}UCUCCGUC^{46}
RQ 135	(+)	^{37}UCCCCCGU^{45} ^{84}CCCCCGCGCGCCGGUUUC^{101}
	(-)	^{64}UCCCUCAAACCCUCGU^{79}
Mar 57	(+)	^{4}CUUCCUCCCUU^{14} ^{58}CUUCCUCCCUU^{68}
	(-)	^{18}CCCCCUCCUUGU^{29} ^{72}CCCCCUCCUUGU^{83}
Mar 72	(+)	^{28}UCCUCUCUAACUUUC^{42}
	(-)	^{31}UUUUCUCACCUCUCU^{45}

FIGURE 11: Comparative analysis between the SELEX ligands and the published sequences of RNAs that are replicated by Qβ replicase. (A) Portion of the Qβ plus strand that is bound by Qβ replicase as well as by S1 alone. For comparison, the sequences of the loop consensus for the Qβ(+S1) ligands and two of the class II Qβ(−30S) ligands are shown. (B) C/U-rich regions of the Qβ minus strand and the plus and minus strands of all of the published replicatable RNAs that are reminiscent of the class II ligands observed in the Qβ(−30S) selection.

well as the lower affinity nonstructured RNAs seen in class I of the Qβ(−30S) selection (Figure 11A). This region of the plus strand has previously been shown to bind specifically to Qβ replicase as well as to ribosomal protein S1 alone (Senear & Steitz, 1975). Deletion experiments on the plus strand indicate that a region that includes the sequence shown in Figure 11A is required for replication (Weber *et al.*, 1974).

50N Replicatable Sequences	77N Replicatable Sequences
UUUCGCCCUU	CUCGCCUUUUCUC
CCUAUUUUU	UUUCCUCC
CUGCUUCUCGUU	CCUUGUUGCCUC
UUUGUUUUUCGUUU	CUAUCUC
UUCGUCCGUCCC	CUAUCGCCUUCCCC
CUUCUGCGUU	UUUCACAUCU
UUCGCCCUACCCU	CCCCAACUUCGUUC
CCGUCUCGUACCUCC	UCUUUUGCUCCCC
UUCUUCCAC	CUUCCAAUCUUC
UUACAUUCCCU	UGCGUUUCU
UUCGUUCCU	CUUACCACUUU
UUUCGGCUC	UUACCUUGUUUC
CCUCGUCCCC	UUCUUUUACC
CUCGCCUUUUCUC	CUUCCCUU
CCUUUCC	CUCCUUCCUACC
UCUUGUUU	CCUACUCCCU
	UUUAUUAGUUU
	UCGUUCUCUUCU
	UUUCUCUCAUCU
	CCCCCAACUUCGUUC
	CUCCCCAGCCUUC
	UUAUUCGCUUUU
	CCUACUCCCU
	CCUACCCUCC
	CCCCCUCCGU
	CUCGCCUUUUCUC
	CUACCUCC
	CUCUCCCAUCU
	UUUAUUCU
	CCCCCAACUUC

FIGURE 12: Pyrimidine-rich domains of the RNAs described by Brown and Gold (1995). The sequences are divided by the selections in which they were isolated.

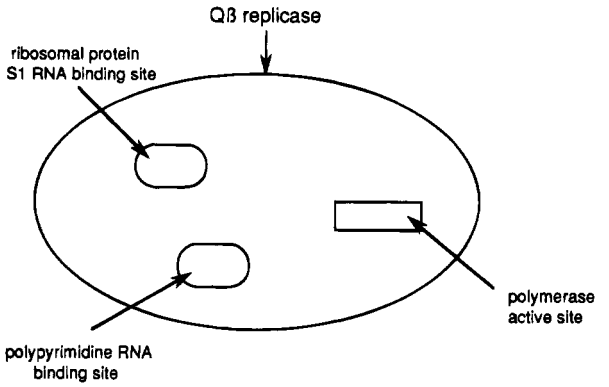


FIGURE 13: Schematic diagram of Qβ replicase with the two proposed RNA binding sites and the polymerase active site. Binding of an RNA at either of the binding sites would increase the probability of introducing the 3′ end of the RNA molecule into the polymerization site.

A deletion analysis of the minus strand of the Qβ genome has isolated an 80 nucleotide region that is required for template activity (Schuppli *et al.*, 1994). This region has a 12 nucleotide pyrimidine-rich tract that is reminiscent of the class II ligands (Figure 11B). Likewise, the plus and minus strands of each of the small replicatable sequences that have been published possess a region with a similar C/U-rich domain (Figures 11B and 12) (Schaffner *et al.*, 1977; Biebricher, 1987; Priano *et al.*, 1987; Munishkin *et al.*, 1988; Munishkin *et al.*, 1991; Moody *et al.*, 1994; Brown & Gold, in press). As with the Qβ minus strand, regions that include the sequences shown for the MDV-1 plus strand and both the plus and minus strands of nanovariant have been shown to either be required for replication (Nishihara *et al.*, 1983) or binding (Schaffner *et al.*, 1977) by Qβ replicase.

The similarities between the replicatable molecules and the SELEX-generated RNA ligands provide an obvious means for template recognition by the phage polymerase (Figure 13). The replicated molecules could be bound by one of the two RNA binding sites on Qβ replicase and either specifically, through RNA tertiary structure, or nonspecifically, by increased localization of the RNAs to the enzyme, introduce their 3′ ends into the polymerase active site of the replicase. Priming of complementary strand synthesis could then initiate the replication cycle.

An interesting consequence of having two binding sites is that the amount of plus and minus strand can be regulated during infection. In the early stages of replication, an equivalent amount of plus and minus strands would be advantageous to maintain exponential amplification and thus speed the process of replication. But the final product should reflect an excess of plus strands, because only the plus strands are encapsidated in the product of viral progeny. If the two strands were using the same binding site, regulation of this type would be difficult because asymmetric replication at the end would necessitate asymmetric replication at the beginning. Having two sites provides the potential for reducing the efficiency of replication at one site while not affecting the replication of the second site. The putative S1 binding site mechanism used for producing minus strands requires the presence of a fifth subunit, host factor. Host factor is the least available protein among the host subunits required for replication (Carmichael *et al.*, 1975). Extensive production of the viral subunit could lead to an excess of replicases lacking the host factor. The production of minus strands would thus be reduced while the production of plus strands could continue unabated.

ACKNOWLEDGMENT

We thank Fred Kramer and David Lane for their advice during the early stages of this research, Patrick Allen for his comments on the research and manuscript, and Steve Ringquist for his critical reading of the manuscript. We thank Gene-Trak, Inc. for their generous contribution of Qβ replicase. We also thank the W. M. Keck Foundation and NeXstar Pharmaceuticals, Inc. for their generous support of RNA science on the Boulder campus.

REFERENCES

Bear, D. G., Ng, R., Van der Veer, D., Johnson, N., Thomas, G., Schleich, T., & Noller, H. F. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1824−1828.

Biebricher, C. (1987) *Cold Springs Harbor Symp. Quant. Biol. 52*, 299−306.

Blumenthal, T., Landers, T., & Weber, K. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1313−1317.

Boni, I. V., Zlatkin, I. V., & Budowsky, E. I. (1982) *Eur. J. Biochem. 121*, 371−376.

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

Carmichael, G. G., Weber, K., Niveleau, A., & Wahba, A. J. (1975) *J. Biol. Chem. 250*, 3607−3612.

Chen, H., Brown, D. M., & Gold, L. (1996) *Methods Enzymol.* (in press).

Dobkin, C., Mills, D., Kramer, F., & Spiegelman, S. (1979) *Biochemistry 18*, 2038−2044.

Donis-Keller, H., Maxam, A., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527−2538.

Gold, L., Polisky, B., Uhlenbeck, O., & Yarus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.

Inouye, H., Pollack, Y., & Petre, J. (1974) *Eur. J. Biochem. 45*, 109−117.

Kamen, R. (1970) *Nature 228*, 527−533.

Kamen, R., Kondo, M., Romer, W., & Weissmann, C. (1972) *Eur. J. Biochem. 31*, 44−51.

Kondo, M., Gallerani, R., & Weissmann, C. (1970) *Nature 228*, 525−527.

Moody, M. D., Burg, J. L., DiFrancesco, R., Lovern, D., Stanick, W., Lin-Goerke, J., Mahdavi, K., Wu, Y., & Farrell, M. P. (1994) *Biochemistry 33*, 13836−13847.

Munishkin, A., Voronin, L., & Chetverin, A. (1988) *Nature 333*, 473−475.

Munishkin, A., Voronin, L., Ugarov, V., Bondareva, L., Chetverina, H., & Chetverin, A. (1991) *J. Mol. Biol. 221*, 463−472.

Nishihara, T., Mills, D., & Kramer, F. (1983) *J. Biochem. (Tokyo) 93*, 669−674.

Priano, C., Kramer, F., & Mills, D. (1987) *Cold Springs Harbor Symp. Quant. Biol. 52*, 321−330.

Ringquist, S., Jones, T., Snyder, E. E., Gibson, T., Boni, I., & Gold, L. (1995) *Biochemistry 34*, 3640−3648.

Schaffner, W., Ruego, K., & Weissmann, C. (1977) *J. Mol. Biol. 117*, 877−907.

Schuppli, D., Barrera, I., & Weber, H. (1994) *J. Mol. Biol. 243*, 811−815.

Schwyzer, M., Billeter, M. A., & Weissmann, C. (1972) *Experentia 28*, 750.

Senear, A., & Steitz, J. (1975) *J. Biol. Chem. 251*, 1902−1912.

Spiegelman, S., Pace, N., Mills, D., Levisohn, R., T., E., Taylor, M., Peterson, R., & Bishop, D. (1968) *Cold Springs Harbor Symp. Quant. Biol. 33*, 101−124.

Stern, S., Moazed, D., & Noller, H. F. (1988) *Methods Enzymol. 164*, 481−489.

Tuerk, C., & Gold, L. (1990) *Science 249*, 505−510.

Tuerk, C., Eddy, S., Parma, D., & Gold, L. (1990) *J. Mol. Biol. 213*, 749−761.

Wahba, A., Miller, M., Niveleau, A., Landers, T., Carmichael, G., Weber, K., Hawley, P., & Slobin, L. (1974) *J. Biol. Chem. 294*, 3314−3316.

Weber, H., Kamen, R., Meyer, F., & Weissmann, C. (1974) *Experentia 30*, 711.

Weissmann, C. (1974) *FEBS Lett. 43*, s10−s18.

Zuker, M. (1989) *Science 244*, 48.

BI951854C