500–506 Nucleic Acids Research, 1995, Vol. 23, No. 3

Recognition of duplex DNA by RNA polynucleotides

Claudia D. McDonald and L. James Maher, III*

Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, 600 S. 42nd Street, Omaha, NE 68198-6805, USA

Received August 30, 1994; Revised and Accepted December 15, 1994

ABSTRACT

We are interested in creating artificial gene repressors based on duplex DNA recognition by nucleic acids. Homopyrimidine RNA oligonucleotides bind to duplex DNA at homopurine/homopyrimidine sequences under slightly acidic conditions. Recognition is sequence-specific, involving rU·dA·dT and rC+·dG·dC base triplets. Affinities were determined for folded polymeric RNAs (ca. 100–200 nt) containing 0, 1 or 3 copies of a 21 nt RNA sequence that binds duplex DNA by triple helix formation. When this recognition sequence was inserted into the larger folded RNAs, micromolar concentrations of the resulting RNA ligands bound a duplex DNA target at pH 5. However, these binding affinities were at least 20-fold lower than the affinity of an RNA oligonucleotide containing only the recognition sequence. Enzymatic probing of folded RNAs suggests that reduced affinity arises from unfavorable electrostatic, structural and topological considerations. The affinity of a polymeric RNA with three copies of the recognition sequence was greater than that of a polymeric RNA with a single copy of the sequence. This affinity difference ranged from 2.6- to 13-fold, depending on pH. Binding of duplex DNA by polymeric RNA might be improved by optimizing the RNA structure to efficiently present the recognition sequence.

INTRODUCTION

We are exploring the design of site-specific transcriptional repressors based on nucleic acids (DNA or RNA) that can recognize regulatory sequences in duplex DNA. Such artificial repressors are of interest both for their therapeutic potential (1–3) and as possible models for ribonucleoproteins in natural transcriptional regulation (4,5).

One approach to the rational design of artificial repressors involves oligonucleotide-directed triple helix formation. Triple helix formation can arise in two patterns termed the pyrimidine motif and the purine motif (1). In the pyrimidine motif, oligonucleotides bind parallel to the purine strand of the DNA duplex by Hoogsteen hydrogen bonds [T·A·T and C+·G·C triplets, see Fig. 1 (6)]. In the purine motif, oligonucleotides bind antiparallel to the purine strand of the DNA duplex by reverse Hoogsteen hydrogen bonds [T·A·T (or A·A·T) and G·G·C triplets (7)].

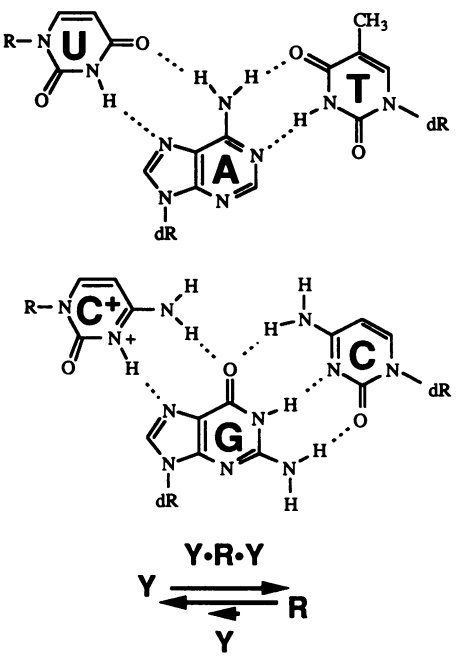

Figure 1. DNA recognition by RNA in the pyrimidine triple helix motif. Purine bases in homopurine DNA sequences can be recognized by oligonucleotide-directed triple helix formation in the major groove. Specificity arises from the formation of specific base triplets (rU·dA·dT and rC+·dG·dC). Strand orientations are indicated below (Y, strand containing homopyrimidine sequence; R, strand containing homopurine sequence).

Triple-helical complexes are thermodynamically stable near physiological conditions and have half-lives of several hours (8–10). Such complexes can inhibit DNA binding proteins (11–13) and can repress eukaryotic promoters *in vitro* (14,15). Indirect evidence has been presented to suggest that triple helices can form and alter gene expression after exposure of intact cells to oligonucleotides (16–18).

We are considering the possibility that RNA transcripts might be engineered to act as gene repressors, perhaps via triple helix formation with duplex DNA. Because the pyrimidine motif

* To whom correspondence should be addressed

requires slightly acidic pH to promote protonation of oligonucleotide cytosines (6,19,20), the pH-independent purine motif has been thought to provide a more viable approach for *in vivo* applications involving DNA oligonucleotides (14,17,18,21). However, whereas RNA can participate as the third strand in triple helices of the pyrimidine motif (22–24), it is not accommodated in the purine motif (24–26). Therefore, our initial studies of duplex DNA recognition by oligomeric RNA (24) and polymeric RNA (this report) employ the pyrimidine triple helix motif and are performed under mildly acidic conditions.

The present study measures triple helix formation by one or three copies of a short (21-nt) homopyrimidine RNA sequence inserted into a larger folded RNA. Affinities of the resulting folded polymeric RNAs for a duplex DNA target are quantitated using a gel mobility shift assay and compared with the affinity of an oligomeric RNA containing only the recognition sequence. The results emphasize the need for strategies to optimize the presentation of the recognition sequence within a larger folded RNA polymer.

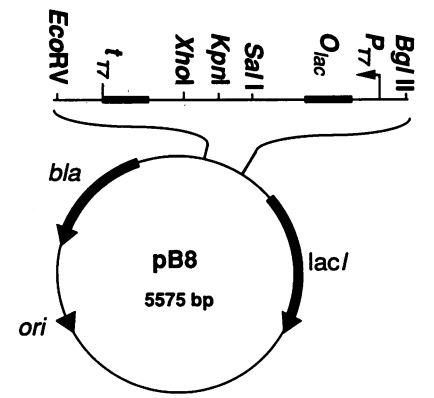

Figure 2. Templates for *in vitro* transcription. Plasmid pB8 was constructed from pET-15b as described in Materials and Methods. Transcription by T7 RNA polymerase produces a 113 nt RNA. One or three copies of a homopyrimidine sequence for DNA recognition were cloned at the *SalI* site (see enlargement above) to allow synthesis of RNA polynucleotides containing homopyrimidine sequences for duplex DNA recognition.

MATERIALS AND METHODS

Materials

Radiochemicals were purchased from Amersham. *E.coli* DNA polymerase I Klenow fragment, T4 DNA ligase, T4 RNA ligase, polynucleotide kinase and restriction endonucleases were purchased from New England Biolabs. T7 RNA polymerase was purchased from Epicentre Technologies. RNase T1 was purchased from Boehringer Mannheim. Mung bean nuclease was purchased from Pharmacia. Monomers for RNA synthesis (base-protected *tert*-butyl dimethylsilyl β-cyanoethyl phosphoramidites) were purchased from MilliGen/Biosearch. Solid supports for RNA synthesis were purchased from Glen Research. Tetrabutylammonium fluoride (1 M in tetrahydrofuran) was purchased from Aldrich.

Transcription templates

Plasmid pB8 (Fig. 2) was derived from pET-15b (Novagen) as follows. Plasmid pET-15b was digested with *Bpu*1102I and *Xba*I. The synthetic duplex:

5′-CTAGTGTCGACAGGGTACCAGCTCGAGT
ACAGCTGTCCCATGGTCGAGCTCAACT-3′

was ligated between the *Bpu*1102I and *Xba*I sites to introduce unique *Sal*I, *Kpn*I and *Xho*I sites in pB8. Plasmid pB8 was transcribed by T7 RNA polymerase to produce **L0** (Fig. 3). One or three copies of the synthetic duplex:

5′-TCGAGTCCTTTTCTTTCTTTTTCCTTC
CAGGAAAAGAAAGAAAAAGGAAGAGCT-3′

were ligated into the *Sal*I site of pB8 to produce plasmids pB9 and pB11, respectively. The inserts were oriented such that transcription produced RNA containing the homopyrimidine sequence. **L1** and **L3** were transcribed from plasmids pB9 and pB11, respectively, using T7 RNA polymerase (Fig. 3).

Oligonucleotides

Oligodeoxyribonucleotides were prepared, purified and quantitated as previously described (14). Oligoribonucleotides were synthesized at 1 μmol scale by phosphoramidite chemistry on an Applied Biosystems 380B synthesizer, with cycle modifications as suggested by the instrument manufacturer. Oligomers were removed from the solid support and partially deprotected by 8 h treatment with concentrated ammonia:ethanol (3:1) at 55°C. The dried residue was treated for 48 h with a 0.5 ml solution of 1 M tetrabutylammonium fluoride in tetrahydrofuran. After addition of 0.5 ml of 2 M triethylammonium acetate (pH 7.0) and 1 ml H_2O, the oligomers were desalted by chromatography over P6 gel (BioRad). Fractions containing the oligomer were dried. The oligomers were then purified by electrophoresis through a 20% acrylamide gel containing 7 M urea, followed by ultraviolet shadowing and band excision. After elution overnight into 0.3 M ammonium acetate, oligonucleotides were desalted using Sep-pak cartridges, as directed by the manufacturer (Waters). RNA oligonucleotide concentrations were calculated using the following molar extinction coefficients at 260 nm ($M^{-1}\cdot cm^{-1}$): 15 400 (A), 7300 (C), 11 700 (G), 9900 (U).

In vitro transcription

Polymeric RNAs were synthesized from 2.5 μg plasmid DNA (linearized by digestion with *Eco*RV) using a T7 RNA polymerase transcription kit (Epicentre Technologies). Transcription reactions (20 μl) were prepared as suggested by the manufacturer and incubated at 37°C for 2 h. *In vitro* transcripts were efficiently terminated at the T7 terminator just upstream of the *Eco*RV site. Transcription reactions were treated with DNase I and stopped by addition of EDTA and formamide sample buffer. RNA was heated to 90°C for 5 min, and purified by electrophoresis through a 5% denaturing polyacrylamide gel. RNA was located by shadowing with ultraviolet light and excised. After elution overnight into 200 μl elution buffer (200 mM Tris–HCl pH 7.6, 2.5 mM EDTA, 300 mM sodium chloride, 2% sodium dodecyl sulfate), the RNA was extracted with phenol:chloroform:isoamyl alcohol (24:24:1) and precipitated from ethanol. RNA was resuspended in H_2O and concentrations were calculated using the relationship 1 $O.D._{260}$ U/ml = 40 μg/ml.

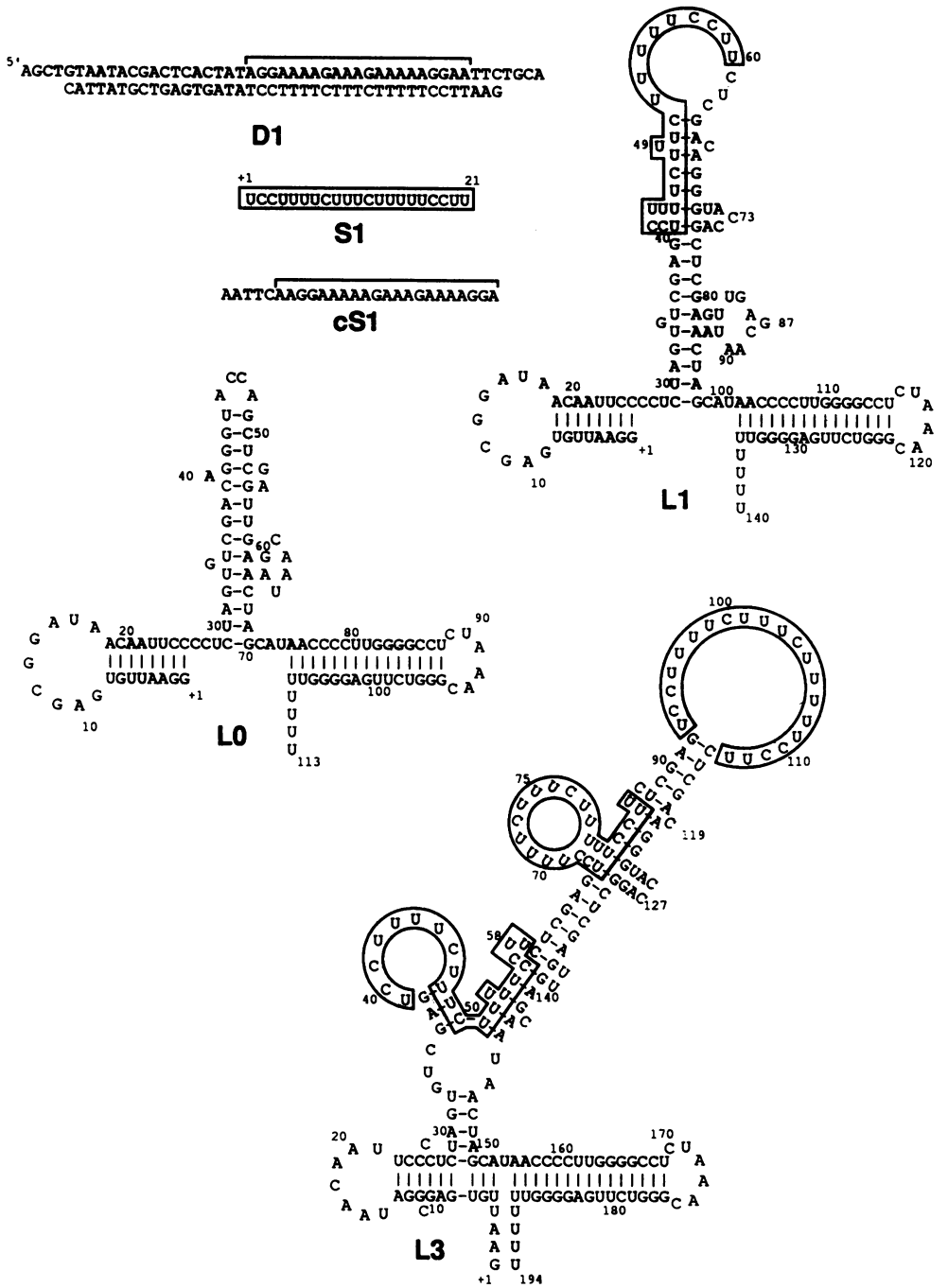

Figure 3. Duplex DNA target and RNA ligands. Duplex DNA **D1** contains a homopurine target for RNA-directed triple helix formation. RNA ligands assayed for recognition of **D1** include 21 nt **S1**, 113 nt **L0** (no recognition sequence), 140 nt **L1** (one copy of recognition sequence) and 195 nt **L3** (three copies of recognition sequence). **L0**, **L1** and **L3** are depicted in the secondary structures predicted by the folding method of Zuker (35). The homopurine recognition target of **D1** is indicated by a bracket. Cognate homopyrimidine sequences on RNA ligands are boxed. DNA oligonucleotide **cS1** is complementary (Watson–Crick) to the recognition sequences of **S1**, **L1** and **L3**.

Electrophoretic mobility shift titrations

Binding reaction mixtures contained (in order of addition) H_2O, labeled **D1** (50 000 c.p.m.; ca. 0.2 pmol), 1 µl of 10 × binding buffer (1 M sodium acetate pH 5 or 0.9 M Tris acetate pH 6 or 0.9 M Tris acetate pH 6.5, 100 mM $MgCl_2$), 1 µl of 1 mg/ml yeast tRNA, 1 µl of 5 mM spermine tetrahydrochloride (0.5 mM final concentration) and RNA ligand (to yield the indicated final concentration) in a final volume of 10 µl. Reaction mixtures were incubated at 22°C for 2 h and were then supplemented with 1 µl of an 80% glycerol solution containing bromophenol blue. Reactions were analyzed by electrophoresis through 5% native polyacrylamide gels (19:1 acrylamide:bisacrylamide) prepared in the appropriate buffer (100 mM sodium acetate pH 5 or 90 mM Tris–acetate pH 6 or 90 mM Tris–acetate pH 6.5) supplemented with 1 mM magnesium chloride. Electrophoresis was performed

(with buffer recirculation) at 4°C overnight (9 V/cm). The resulting gel was imaged and analyzed by storage phosphor technology using a Molecular Dynamics PhosphorImager.

Analysis of gel mobility shift titrations

The apparent fraction, θ, of **D1** bound by RNA ligands was calculated for each gel lane using the definition:

$$\theta = S_{triplex}/(S_{triplex}+S_{duplex}) \quad (1)$$

where $S_{triplex}$ and S_{duplex} represent the storage phosphor signal for triplex and duplex complexes, respectively. Values of the apparent triplex dissociation constant, K_d, were obtained by least squares fitting of the data to the binding isotherm:

$$\theta = ([R]^n/K_d^n)/(1+[R]^n/K_d^n \quad (2)$$

where [R] is the total RNA concentration and n is the Hill coefficient (27).

Enzymatic probing of RNA structure

RNA transcripts (1–2 μg) were labeled at their 3′ termini using RNA ligase and [5′-^{32}P]pCp in an overnight incubation at 16°C, as described by the enzyme supplier. Labeled RNA was purified by denaturing polyacrylamide gel electrophoresis. After elution overnight into 200 μl elution buffer (200 mM Tris hydrochloride pH 7.6, 2.5 mM EDTA, 300 mM sodium chloride, 2% sodium dodecyl sulfate), the RNA was extracted with phenol:chloroform:isoamyl alcohol (24:24:1) and precipitated from ethanol. Aliquots of labeled RNA (30 000 c.p.m.; ca. 1 pmol) were incubated for 30 min at 25°C in 5 μl reactions containing 0.1 M sodium acetate pH 5, 10 mM $MgCl_2$, 0.1 mg/ml yeast tRNA and 0.5 mM spermine tetrahydrochloride. Some reactions contained 2 μM concentrations of either unlabeled duplex **D1** or DNA oligomer **cS1**. Mung bean nuclease (4 U) was then added at 25°C for 2 min. Nuclease treatment was terminated by the addition of an equal volume of urea loading dyes and freezing on dry ice. Separate aliquots of labeled RNA were also treated either at 90°C for 15 min in 50 mM sodium bicarbonate pH 9 (under paraffin oil), or for 10 min at 55°C in 10 M urea with 10 U RNase T1, followed by freezing. RNA samples were heated to 90°C for 30 s, chilled on ice and then analyzed on 8% denaturing polyacrylamide sequencing gels (19:1 acrylamide: bisacrylamide). The resulting gels were dried and imaged by storage phosphor technology using a Molecular Dynamics PhosphorImager.

RESULTS AND DISCUSSION

Experimental design

The duplex DNA target and potential RNA ligands are shown in Figure 3. Duplex DNA **D1** contains a 21 bp homopurine/homopyrimidine sequence that can be recognized by DNA or RNA oligonucleotides in the pyrimidine triple helix motif (24,28). **D1** was labeled by filling the recessed left terminus using Klenow fragment of DNA polymerase I in the presence of radioactive deoxynucleoside triphosphates. Four potential RNA ligands were compared for their ability to bind labeled **D1** in electrophoretic gel mobility shift titrations. RNA **S1** (21 nt) contains only the appropriate homopyrimidine sequence for recognition of **D1** in the pyrimidine triple helix motif. DNA **cS1** contains the Watson–Crick complement to **S1**. RNAs **L0**, **L1** and **L3** contain 0, 1 or 3 copies (respectively) of this homopyrimidine sequence inserted into a 113 nt RNA. Increasing concentrations of unlabeled RNA ligands were incubated with labeled **D1** under slightly acidic pH conditions (pH 5, 6 or 6.5) in the presence of magnesium chloride (10 mM) and spermine tetrahydrochloride (0.5 mM). Electrophoretic separation of the resulting complexes allowed determination of the relative affinities of the RNA ligands as a function of pH.

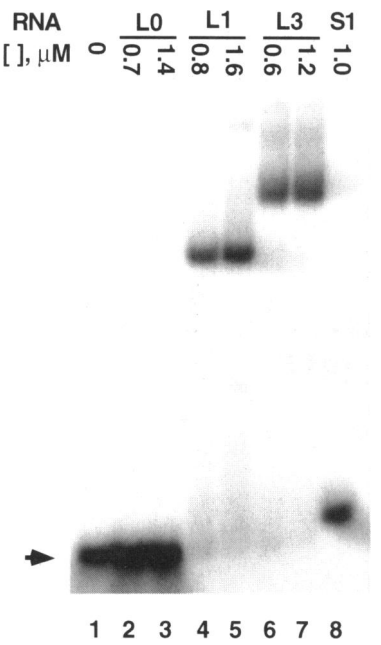

Figure 4. Electrophoretic assay of triple helix formation. Binding of RNA ligands to labeled **D1** as detected after electrophoresis through a native 5% polyacrylamide gel at pH 5 in the presence of 1 mM $MgCl_2$. The indicated (μM) concentrations of RNA ligands **L0**, **L1**, **L3** and **S1** were incubated with labeled **D1** prior to electrophoresis and quantitation. Position of free **D1** is indicated by the arrow.

Affinities of oligonucleotide and polynucleotide complexes

An example of the electrophoretic separation of triple helices involving RNA ligands is shown in Figure 4. Incubation of labeled **D1** with micromolar concentrations of oligomeric and polymeric RNAs at pH 5 produced stable complexes that migrated more slowly in the gel (Fig. 4, compare lane 1 with lanes 4–8). As expected, no complex was observed with **L0**, which lacks a homopyrimidine insert (Fig. 4, lanes 2–3). The mobilities of the complexes corresponded to the size of the RNA ligand involved. Thus, compact triplexes involving **S1** were only slightly retarded (Fig. 4, lane 8), whereas complexes involving **L1** and **L3** exhibited greatly reduced mobilities (Fig. 4, lanes 4–7).

Similar experiments were performed for a range of RNA ligand concentrations between ~10 nM and 1 μM in solutions buffered at pH 5, 6 and 6.5. For each ligand concentration, the fraction of labeled **D1** in triplex form, θ, was measured as described in Materials and Methods. Data from these experiments are shown in Figure 5. Least-squares fitting of these data to a binding equation (dotted lines) allowed estimation of the equilibrium dissociation constant, K_d, for each ligand at each pH. These

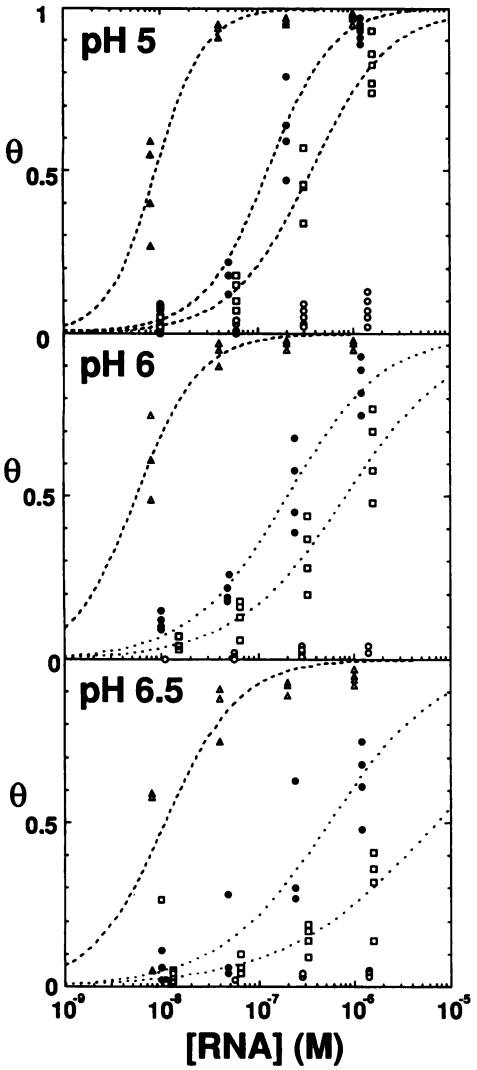

Figure 5. Estimation of RNA binding affinities. The fraction of **D1** in triple-helical form, θ, is plotted as a function of RNA concentration. (Δ) **S1**; (○) **L0**; (□) **L1**; (●) **L3**. Binding curves (dotted lines) correspond to least squares fits to the binding equation given in Materials and Methods.

results are shown in Table 1 and in Figure 6. Several conclusions can be drawn from these data.

Table 1. Dissociation constants for RNA-directed triple helix formation on DNA

RNA	K_d(M) pH 5	pH 6	pH 6.5
L0	–	–	–
L1	3.6×10^{-7}	8×10^{-7}	7.1×10^{-6}
L3	1.4×10^{-9}	2×10^{-7}	5.4×10^{-7}
S1	8.9×10^{-9}	5.6×10^{-9}	1.1×10^{-8}

(–): $K_d >> 10^{-5}$ M.

First, although micromolar concentrations of all ligands carrying at least one copy of the recognition sequence saturate **D1**

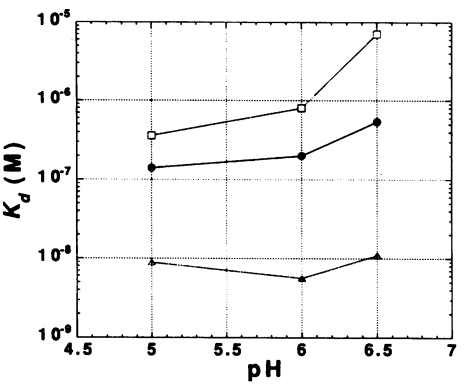

Figure 6. Equilibrium dissociation constants for RNA-directed triple helix formation. Dissociation constants (K_d) estimated from data in Figure 5 are plotted as a function of pH. Standard errors for the K_d estimates are approximately 10%. (Δ) **S1**; (□) **L1**; (●) **L3**.

when tested at pH 5 (e.g. Fig. 4, lanes 4–8), the affinities of **L1** and **L3** are significantly reduced relative to that of oligomer **S1**. This reduced affinity ranges between 20-fold (**L3** versus **S1**, pH 5) and 650-fold (**L1** versus **S1**, pH 6.5). This effect presumably arises from a combination of electrostatic, structural and topological considerations. Electrostatic repulsion may inhibit triple helix formation by polymeric RNAs where additional phosphate charges must approach the DNA duplex without a compensatory increase in stacking and hydrogen-bonding interactions. The tendency of polymeric RNAs to adopt folded structures in which the homopyrimidine sequence is sequestered may also drastically reduce the effective concentration of this sequence. Another critical consideration concerns whether the homopyrimidine domain is topologically disposed to wrap twice around the DNA duplex. Internal loops of RNA are ill-suited for this purpose, especially if adjacent stems cannot be readily unpaired. Under such circumstances, perhaps only a portion of the homopyrimidine domain can engage the duplex DNA target, resulting in a sub-optimal number of specific contacts.

Secondly, the apparent pH-dependence of ligand binding between pH 5 and 6.5 is small for **S1**, but larger for **L3** (4-fold decrease in affinity) and **L1** (20-fold decrease in affinity). Cytosine protonation may contribute to triplex stability both by creating a new hydrogen-bond donor within the C·G·C triplet and by reducing the net charge of the complex. Perhaps the electrostatic component of this stabilization is more significant for complexes involving polymeric RNAs. Thus, for a given number of specific contacts, loss of bound protons at higher pH values might selectively destabilize complexes of polymeric RNAs relative to an oligomeric RNA with no excess phosphate charges. In addition, presentation of the recognition sequence may require some unfolding of RNA secondary structures. This unfolding process might be facilitated at lower pH. Thus the substantial pH dependence of **L1** binding could be attributable to a greater requirement for releasing inhibitory secondary structures.

Thirdly, **L3**, which contains three copies of the homopyrimidine recognition sequence, binds between 2.6- (pH 5) and 13-fold (pH 6.5) more tightly to **D1** than **L1** (single copy of homopyrimidine recognition sequence). In the absence of structural considerations, the effective concentration of homopyrimidine sequences contributed by **L3** (per mole of ligand) is 3-fold higher than for **L1**.

Therefore, **L3** may exhibit a lower dissociation constant because any of its three recognition sequences can equally initiate triplex formation. Alternatively, the tertiary structure of **L3** may result in a more favorable presentation of one homopyrimidine domain (see below). The latter possibility is also suggested by the observation that a single predominant **L3–D1** complex is observed in gels (e.g. Fig. 4, lanes 6 and 7). If a **D1** molecule could be bound by **L3** in more than one way, or if multiple **D1** molecules could be simultaneously bound by **L3**, a more complex pattern of products might be expected. Comparison of predicted secondary structures also emphasizes the differing manner in which the homopyrimidine domain(s) might be disposed within the folded RNAs (Fig. 3).

RNA structures

Enzymatic probing of folded RNAs **L0**, **L1** and **L3** was performed for comparison with the predicted structures (Fig. 3) and to detect any RNA structural changes that occur upon triple helix formation with **D1**. Results of these experiments are shown in Figure 7. Structural information was obtained by limited digestion of end-labeled RNAs with mung bean nuclease (single strand specific) under standard triple helix binding conditions (pH 5). Reference ladders were obtained by partial alkaline hydrolysis or limited digestion with RNase T1 (specific for single-stranded guanines). Note that although RNase T1 digestions were performed at elevated temperature in the presence of denaturant, guanines in the very stable stem of the phage T7 transcription terminator were uniformly protected (lane 3, nt 75–95; lane 9, nt 100–125; lane 15, nt 135–180).

The major features of predicted secondary structure for RNAs **L0**, **L1** and **L3** were confirmed in these experiments. These features include the stable 5′ and 3′ hairpins present in each RNA (accessible loops near nt 10 and nt 90 for **L0**, nt 10 and nt 118 for **L1**, nt 10 and nt 167 for **L3**). Although details of the secondary structures between these terminal loops are generally supported by the nuclease mapping, the accessibility of recognition sequences in RNAs **L1** and **L3** (bars in Fig. 7) is notably limited relative to what had been predicted. Lanes 10 and 16 show mung bean nuclease reactivities of folded RNAs **L1** and **L3**. The homopyrimidine recognition sequences are indicated by bars. Nuclease cleavage in these regions is irregular and modest relative to other sites within the molecules. Similar data were obtained with nuclease P1 (data not shown). These results emphasize the extent to which recognition sequences are obscured in these folded RNAs and suggest that these sequences may be involved in more intramolecular structures than had been predicted.

Of additional interest is the effect of triple helix formation on RNA structure. Folded RNAs were incubated with 2 μM **D1** (calculated from K_d data to saturate **L1** and **L3** under these conditions). In the presence of **D1**, the pattern of nuclease reactivity for **L0** (no recognition sequence) was unchanged as expected (compare lanes 4 and 5). Despite the documented ability of **L1** to bind **D1** under these conditions, only a slight footprint was observed over the recognition sequence, with an induced hyperreactivity near nt 25 (compare lanes 10 and 11). Structural changes were similarly modest upon incubation with **cS1**, which is complementary to the recognition sequence (lane 12). These data suggest that the recognition sequence in **L1** is only partially accessible. For **L3**, triple helix formation with **D1** caused a detectable footprint at the recognition sequence predicted to be most accessible (lane 17, black bar). Only modest reductions in nuclease reactivity were observed in the other two recognition sequences (shaded and open bars adjacent to lane 17). It is particularly notable that little RNA unfolding accompanies binding to **D1**.

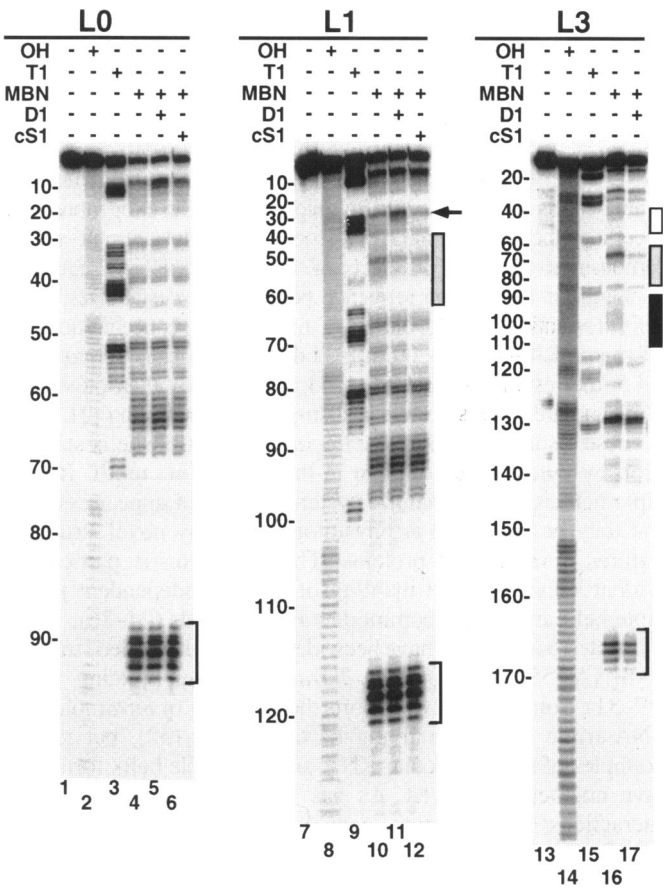

Figure 7. Enzymatic probing of RNA structures. RNAs **L0**, **L1** and **L3** were labeled at their 3′ termini (lanes 1, 7 and 13) and subjected to either partial alkaline hydrolysis (OH; lanes 2, 8 and 14), partial digestion with RNase T1 under denaturing conditions (T1; lanes 3, 9 and 15), or partial digestion with mung bean nuclease (MBN; lanes 4, 10 and 16). Mung bean nuclease digestions were also performed after incubation of labeled RNA with 2 μM unlabeled DNA duplex **D1** (lanes 5, 11 and 17) or complementary DNA **cS1** (lanes 6 and 12). Panels depict different gels. Nucleotides exposed in the loop of the T7 terminator stem–loop structure (see Fig. 3) are indicated by brackets. Homopyrimidine recognition sequences are indicated by bars. Arrow indicates site of increased mung bean nuclease cleavage in lane 11. Some cleavages appear as doublets and triplets due to microheterogeneity in T7 RNA polymerase termination sites in the labeled RNA.

Together these results suggest that homopyrimidine recognition loops within these folded RNAs are structurally constrained so that they are not highly accessible to nucleases or duplex DNA. It appears likely that only a portion of the recognition sequence can form base triplets and/or that only a small fraction of RNAs unfold in a productive manner. The more obvious induced nuclease protection of one recognition sequence within **L3** (relative to **L1**) supports the interpretation that **L3** binds more tightly to **D1** because of the increased availability of this recognition sequence. The folded structures of **L1** and **L3** appear to be relatively rigid because only local nuclease sensitivities (rather than global structures) are altered in the presence of **D1**.

Implications for RNA-based repressors

DNA-directed triple helix formation has been suggested as an approach to the design of artificial gene repressors (1–3). Conventional pharmaceutical strategies involve exogenous addition of DNA oligonucleotides or analogs with the goal of nuclear delivery. We are interested in RNA-directed triple helix formation as an alternative strategy. This approach would involve transfer of an exogenous gene encoding the appropriate repressor RNA. The current results demonstrate that the presence of nonspecific RNA sequences flanking the recognition sequence permits, but destabilizes, triple helix formation. Thus, although micromolar concentrations of polymeric RNAs (100–200 nt) still direct triple helix formation at pH 5, the resulting complexes bind more weakly than oligomeric RNA and are more sensitive to disruption by increasing pH.

These results emphasize the importance of triple helix stabilization by cytosine protonation in the pyrimidine motif. If such triple helices exist in natural systems, it would appear essential that they be stabilized at physiological pH by novel structural features or accessory proteins. This limitation is particularly evident because of the inability of the pH-independent purine triple helix motif to accommodate RNA strands (24–26).

Isolated base triplets have been identified in the folded structure of tRNA, 5S RNA and the *Tetrahymena* self-splicing intron (29–31). Some evidence supports the existence of *intra*molecular DNA triple helices in living bacterial cells (32), but natural examples of *inter*molecular RNA or DNA triple helix formation have not been observed. As has long been speculated, such interactions might permit gene regulation through the site-specific binding of ribonucleic acid or ribonucleoprotein to duplex DNA (33,34). Indeed, DNA recognition by ribonucleoproteins has been reported, although the role of the associated RNA remains unclear (4,5). If found, natural examples of duplex DNA recognition by RNA could yield important design principles for repressor engineering.

The polymeric RNAs in this study were not designed according to any presentation strategy except that 1 or 3 copies of the homopyrimidine domain were inserted between relatively stable stem–loop structures involving the RNA termini. Much of the reduced binding affinity for duplex DNA appears to be attributable to unfavorable folding or topology. We are currently optimizing the presentation of a recognition sequence in the context of a folded RNA by *in vitro* selection from a combinatorial RNA library. These studies will begin to address the role of RNA folding in promoting both the formation and stability of triple helices involving duplex DNA.

ACKNOWLEDGEMENTS

We acknowledge the excellent technical assistance of D. Eicher, C. Mountjoy, W. Olivas, G. Soukup and Junque Sun. We thank B. Gmeiner for comments on the manuscript. Supported in part by Grant 5 P30 CA36727-08 from the National Cancer Institute and Grant GM 47814 from the National Institutes of Health. LJM is a recipient of a Junior Faculty Research Award from the American Cancer Society and a Young Investigator's Award from Abbott Laboratories.

REFERENCES

1 Maher, L.J. (1992) *BioEssays* **14**, 807–815.
2 Hélène, C. (1991) *Eur. J. Cancer* **27**, 1466–1471.
3 Durland, R.H., Kessler, D.J. and Hogan, M. (1991) in *Prospects for Antisense Nucleic Acid Therapy of Cancer and AIDS*, ed. Wickstrom, E. (Alan R. Liss, Inc., New York, NY), pp. 219–226.
4 Davis, T.L., Firulli, A.B. and Kinniburgh, A.J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9682–9686.
5 Beru, N., Smith, D. and Goldwasser, E. (1990) *J. Biol. Chem.* **265**, 14100–14104.
6 Moser, H.E. and Dervan, P.B. (1987) *Science* **238**, 645–650.
7 Beal, P.A. and Dervan, P.B. (1991) *Science* **251**, 1360–1363.
8 Maher, L.J., Dervan, P.B. and Wold, B.J. (1990) *Biochemistry* **29**, 8820–8826.
9 Plum, G.E., Park, Y.-W., Singleton, S.F., Dervan, P.B. and Breslauer, K.J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9436–9440.
10 Durland, R.H., Kessler, D.J., Gunnell, S., Duvic, M., Pettitt, B.M. and Hogan, M.E. (1991) *Biochemistry* **30**, 9246–9255.
11 Maher, L.J., Wold, B. and Dervan, P.B. (1989) *Science* **245**, 725–730.
12 François, J.-C., Saison-Behmoaras, T., Thuong, N.T. and Hélène, C. (1989) *Biochemistry* **28**, 9617–9619.
13 Hanvey, J.C., Shimizu, M. and Wells, R.D. (1989) *Nucleic Acids Res.* **18**, 157–161.
14 Cooney, M., Czernuszewicz, G., Postel, E.H., Flint, S.J. and Hogan, M.E. (1988) *Science* **241**, 456–459.
15 Maher, L.J., Dervan, P.B. and Wold, B. (1992) *Biochemistry* **31**, 70–81.
16 Birg, F., Praseuth, D., Zerial, A., Thuong, N.T., Asseline, U., Le Doan, T. and Hélène, C. (1990) *Nucleic Acids Res.* **18**, 2901–2907.
17 Postel, E.H., Flint, S.J., Kessler, D.J. and Hogan, M.E. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8227–8231.
18 Orson, F.M., Thomas, D.W., McShan, W.M., Kessler, D.J. and Hogan, M.E. (1991) *Nucleic Acids Res.* **19**, 3435–3441.
19 Povsic, T.J. and Dervan, P.B. (1989) *J. Am. Chem. Soc.* **111**, 3059–3061.
20 Singleton, S.F. and Dervan, P.B. (1992) *Biochemistry* **31**, 10995–11003.
21 Ing, N.H., Beekman, J.M., Kessler, D.J., Murphy, M., Jayaraman, K., Zendegui, J.G., Hogan, M.E., O'Malley, B.W. and Tsai, M.-J. (1993) *Nucleic Acids Res.* **21**, 2789–2796.
22 Roberts, R.W. and Crothers, D.M. (1992) *Science* **258**, 1463–1466.
23 Han, H. and Dervan, P.B. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3806–3810.
24 Skoog, J.U. and Maher, L.J. (1993) *Nucleic Acids Res.* **21**, 2131–2138.
25 Escudé, C., François, J.-C., Sun, J.-s., Ott, G., Sprinzl, M., Garestier, T. and Hélène, C. (1993) *Nucleic Acids Res.* **21**, 5547–5553.
26 Semerad, C. and Maher, L.J. (1994) *Nucleic Acids Res.*, **22**, 5321–5325.
27 Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry* Vol. III: The Behavior of Biological Macromolecules (Freeman, San Francisco), p. 864.
28 Maher, L.J. (1992) *Biochemistry* **31**, 7587–7594.
29 Goddard, J.P. (1977) *Prog. Biophys. Molec. Biol.* **32**, 233–308.
30 Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.-P., Ehresmann, C. and Ehresmann, B. (1989) *J. Mol. Biol.* **207**, 417–431.
31 Michel, F. and Westhof, E. (1990) *J. Mol. Biol.* **216**, 585–610.
32 Kohwi, Y., Malkhosyan, S. and Kohwi-Shigematsu, T. (1992) *J. Mol. Biol.* **223**, 817–822.
33 Miller, J.H. and Sobell, H.M. (1966) *Proc. Natl. Acad. Sci. USA* **55**, 1201–1205.
34 Minton, K.W. (1985) *J. Exp. Path.* **2**, 135–148.
35 Program Manual for the GCG Package, Version 7 (1991) Genetics Computer Group, Madison, Wisconsin.