Exclusion of RNA strands from a purine motif triple helix

Craig L. Semerad and L. James Maher III*

Eppey 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; Accepted October 19, 1994

ABSTRACT

Research concerning oligonucleotide-directed triple helix formation has mainly focused on the binding of DNA oligonucleotides to duplex DNA. The participation of RNA strands in triple helices is also of interest. For the pyrimidine motif (pyrimidine·purine·pyrimidine triplets), systematic substitution of RNA for DNA in one, two, or all three triplex strands has previously been reported. For the purine motif (purine·purine·pyrimidine triplets), studies have shown only that RNA cannot bind to duplex DNA. To extend this result, we created a DNA triple helix in the purine motif and systematically replaced one, two, or all three strands with RNA. In dramatic contrast to the general accommodation of RNA strands in the pyrimidine triple helix motif, a stable triplex forms in the purine motif only when all three of the substituent strands are DNA. The lack of triplex formation among any of the other seven possible strand combinations involving RNA suggests that: (i) duplex structures containing RNA cannot be targeted by DNA oligonucleotides in the purine motif; (ii) RNA strands cannot be employed to recognize duplex DNA in the purine motif; and (iii) RNA tertiary structures are likely to contain only isolated base triplets in the purine motif.

INTRODUCTION

Besides the canonical double helix, certain nucleic acid polymers have been observed to form three- and four-stranded complexes (1–3; for reviews see refs 4 and 5). For example, two strands of poly(U) RNA bind to a strand of poly(A) RNA to form a stable triplex (1). At acidic pH, two DNA strands composed of alternating dT and dC residues form a stable triplex with a DNA strand of alternating dG and dA residues (6). These polymer interactions suggested an approach to duplex DNA recognition termed oligonucleotide-directed triple helix formation (4,7–9). Sequences of purine bases in duplex DNA are bound by a third DNA strand (typically an oligonucleotide) that occupies the major groove. Triple helix formation arises in two patterns termed the pyrimidine motif and the purine motif (4). In the pyrimidine motif, oligonucleotides bind parallel to the purine strand of the duplex by Hoogsteen hydrogen bonds (T·A·T and C⁺·G·C triplets; 7). In the purine motif (Fig. 1), oligonucleotides bind

antiparallel to the purine strand of the duplex by reverse Hoogsteen hydrogen bonds (T·A·T or A·A·T and G·G·C triplets; 10).

Research in the area of oligonucleotide-directed triple helix formation often focuses on the binding of DNA oligonucleotides to duplex DNA. Possible applications include the design of artificial gene repressors, where duplex DNA is the appropriate target (4). However, the participation of RNA strands in triple helices is of interest for four important reasons. First, RNA can be considered as a possible ligand for duplex DNA, suggesting the concept of artificial, RNA-based gene repressors expressed from therapeutic transgenes (4,11). Second, potential duplex nucleic acid targets for therapeutic intervention might also include RNA duplexes (such as stem-loop structures in folded RNA), or RNA/DNA heteroduplexes. Third, intermolecular triple helix formation involving RNA might play a role in some cases of natural gene regulation (12,13). Fourth, some folded RNAs appear to be stabilized by tertiary contacts involving base triplets (2,14–16).

A few previous studies have explored the ability of RNA strands to participate in oligonucleotide-directed triple helix formation in the pyrimidine and purine motifs. For the pyrimidine motif, systematic substitution of RNA for DNA in one, two, or all three triplex strands has been reported (17–19). Although these authors present results that differ somewhat in detail, the emerging perspective is that the pyrimidine motif is rather promiscuous in its ability to accommodate RNA strands. For example, Roberts and Crothers detected six of the eight possible triplexes involving different combinations of DNA and RNA strands (17). For the purine motif, studies have been limited to exploring whether RNA can act as the third strand in recognition of duplex DNA (11,19). Both previous reports on this subject conclude that the purine motif does not accommodate RNA in triple helices with duplex DNA.

To extend this result, we created a DNA triple helix in the purine motif and systematically replaced one, two, or all three of the substituent strands with RNA. We report that, in dramatic contrast to the general accommodation of RNA strands in the pyrimidine triple helix motif, only when all three nucleic acid strands are composed of DNA is a stable triple helix detected in the purine motif. The lack of triplex formation among any of the other seven possible strand combinations involving RNA suggests that: (i) duplex structures containing RNA cannot be

*To whom correspondence should be addressed

Purine Motif

Figure 1. Purine triple helix motif. Purine bases in homopurine duplex DNA sequences can be recognized by oligodeoxyribonucleotide-directed triple helix formation in the major groove. Specificity arises from the formation of specific base triplets dG·dG·dC and dT·dA·dT (or dA·dA·dT). dR indicates deoxyribose. Strand orientations are indicated by filled circles.

targeted by DNA oligonucleotides in the purine motif; (ii) RNA strands cannot be employed to recognize duplex DNA in the purine motif; and (iii) RNA tertiary structures are likely to contain only isolated base triplets in the purine motif.

MATERIALS AND METHODS

Materials

Radiochemicals were purchased from Amersham. T4 polynucleotide kinase was purchased from New England Biolabs. Monomers for RNA synthesis (base-protected *tert*-butyl dimethylsilyl β -cyanoethyl phosphoramidites) and protected nucleosides linked to polystyrene supports were purchased from Applied Biosystems. Tetrabutylammonium fluoride (1 M in tetrahydrofuran) was purchased from Aldrich.

Oligonucleotides

Oligodeoxyribonucleotides were prepared, purified, and quantitated as previously described (20). Oligoribonucleotides were synthesized at 0.2 μ mol 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 3 h treatment with concentrated ammonia/ethanol (3:1) at 55°C, and then dried *in vacuo*. The residue was treated for 48 h with 1 M tetrabutylammonium fluoride in tetrahydrofuran (0.5 ml). After addition of 1 ml H₂O, 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 (19:1 acrylamide/bisacrylamide) prepared in 1×TBE buffer containing 7 M urea, followed by ultraviolet shadowing and band excision. After elution overnight into 0.3 M ammonium acetate, oligonucleotides were desalted by precipitation from ethanol or by using Sep-pak cartridges, as directed by the manufacturer (Waters). RNA and DNA concentrations were calculated using the following molar extinction coefficients at 260 nm ($M^{-1} \text{ cm}^{-1}$): 15400 (dA), 9800 (rA), 7300 (dC), 6200 (rC), 11 700 (dG), 10 400 (rG), 8800 (dT) and 9350 (rU). Oligonucleotides were characterized by matrix-assisted laser desorption ionization time-of-flight mass spectroscopy (21).

Oligonucleotides were labeled by treatment with T4 polynucleotide kinase in the presence of [γ -³²P]ATP. Labeled oligonucleotides were desalted by precipitation from ethanol in the presence of ammonium acetate. Labeled oligonucleotides were subsequently purified by electrophoresis through 20% native polyacrylamide gels (19:1 acrylamide/bisacrylamide) prepared in 1×TBE buffer supplemented with 8 mM magnesium chloride. The electrophoresis buffer was recirculated. Radiolabeled oligonucleotides were excised from the gel, eluted into 200 μ l elution buffer (200 mM Tris–hydrochloride pH 7.6, 2.5 mM EDTA, 300 mM sodium chloride, 2% sodium dodecyl sulfate), extracted with phenol/chloroform/isoamyl alcohol (24:24:1), and precipitated from ethanol.

Electrophoretic mobility shift titrations

Binding reaction mixtures contained (in order of addition) H₂O, labeled hairpin (1×10^5 c.p.m.; ca. 0.4 pmol), and 1 μ l of 10× binding buffer (250 mM Tris–hydrochloride pH 8, 60 mM MgCl₂) in a final volume of 8 μ l. This solution was heated to 70°C for 5 min, and then cooled rapidly to 22°C. Yeast tRNA (1 μ l of 1 mg/ml solution) was then added, followed by 1 μ l of concentrated DNA or RNA oligonucleotide solution to give the indicated final oligonucleotide concentration in a volume of 10 μ l. Reaction mixtures were incubated at 22°C for 2.5 h and were then supplemented with 1 μ l of an 80% glycerol solution containing bromophenol blue. Reactions were analyzed by electrophoresis through 20% native polyacrylamide gels (19:1 acrylamide/bisacrylamide) prepared in 1×TBE buffer supplemented with 8 mM magnesium chloride. Electrophoresis was performed in this buffer (with recirculation) at 4°C overnight (9 V/cm). The resulting gel was imaged and analyzed by storage phosphor technology using a Molecular Dynamics Phosphor-Imager.

Analysis of gel mobility shift titrations

The apparent fraction, θ , of labeled hairpin bound by oligonucleotides was calculated for each gel lane using the definition:

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

where S_{triplex} and S_{duplex} represent the storage phosphor signal for triplex and hairpin 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 single-stranded oligonucleotide concentration, and n is the Hill coefficient (22).

Figure 2. Experimental design. Four synthetic hairpin structures (**DD**, **DR**, **RD**, **RR**) and three single strands (**D**, **R**, **N**) were used in these studies. The first letter of the hairpin designation refers to the homopurine strand, and the second letter to the homopyrimidine strand (**D**, DNA; **R**, RNA). Single strands **D** and **R** contain the proper nucleotide sequence for recognition of the hairpin structures as predicted by the purine triple helix motif. Strand **N** contains a nonspecific sequence. RNA sequences are shown in bold type.

RESULTS AND DISCUSSION

Experimental design

We adopted an experimental design that had proven fruitful for analyzing the effects of RNA strands on stabilities of triple helices in the pyrimidine motif (17). The synthetic oligonucleotides designed for this purpose are illustrated in Fig. 2. Four intramolecular hairpin structures were synthesized using DNA phosphoramidites (**DD**), RNA phosphoramidites (**RR**), or both (**DR**, **RD**). These hairpin structures represent the four possible dispositions of DNA and/or RNA strands in a Watson-Crick duplex. The hairpins were designed to create a guanosine-rich, homopurine sequence suitable for recognition by a guanosine-rich oligonucleotide in the purine triple helix motif (4,23). The two-letter designations for these hairpins reflect the DNA (**D**) or RNA (**R**) character of the strands, where the first letter describes the 5' (homopurine) half of the sequence and the second letter describes the 3' (homopyrimidine) half of the sequence. Hairpin loops were composed of four thymidine residues, except for **RR**, where uridines comprised the loop.

Three additional single-stranded oligonucleotides were synthesized to explore triple helix formation with the synthetic hairpins. DNA oligonucleotide **D** and RNA oligonucleotide **R** are composed of the guanosine-rich sequences predicted to allow duplex recognition in the purine triple helix motif. DNA oligonucleotide **N** represents a nonspecific sequence that serves as a negative control.

The triplex propensity of each of the eight possible combinations of DNA and RNA strands was measured using an electrophoretic gel mobility shift assay (8,24,25). Radiolabeled

Figure 3. Formation of the DDD triplex. (A) Gel mobility shift titration. Radiolabeled **DD** was incubated alone (lane 1) or with increasing μ M concentrations (0.03, 0.06, 0.12, 0.25, 0.5, 1.0) of **N** (■; nonspecific) or **D** (●; specific). The position of the **DDD** complex is indicated. (B) Quantitation of **DDD** stability. Values of θ (as defined by eqn 1 in Materials and Methods) from several experiments are plotted against the concentration of **N** or **D**. The data are fitted to eqn 2 (Materials and Methods), yielding a calculated dissociation constant, K_d , of 1.1×10^{-7} M.

triplexes migrate more slowly through native polyacrylamide gels (containing Mg^{2+}) than the corresponding duplexes. Equilibrium dissociation constants for site-specific triple helices are estimated by monitoring the dependence of the duplex/triplex equilibrium (involving trace concentrations of a labeled duplex) on the concentration of the unlabeled third strand (25).

Stability of the DDD complex

Several studies have shown that guanosine-rich oligodeoxyribonucleotides can bind with sequence-specificity to duplex DNA in the presence of Mg^{2+} (8,10,11,24,26,27). To confirm this interaction in the present system, radiolabeled samples of **DD** were incubated with increasing concentrations of specific **D** or nonspecific **N**. Triple helix formation was monitored by gel mobility. The results of such an experiment are shown in Fig. 3A.

Figure 4. Exclusion of RNA strands from a purine motif triple helix. (A) Radiolabeled hairpins (DD, DR, RD, RR) were incubated with increasing μM concentrations (0, 0.125, 1.0) of unlabeled D. (B) Radiolabeled hairpins (DD, DR, RD, RR) were incubated with increasing μM concentrations (0, 0.125, 1.0) of unlabeled R.

When tested alone, DD migrated as a single band (Fig. 3A, lane 1). The mobility of DD was unaffected by incubation in the presence of up to 1 μM N (Fig. 3A, lanes 2–7). In contrast, incubation with increasing concentrations of unlabeled D produced a DDD complex, as expected (Fig. 3A, lanes 8–13).

The dependence of DDD formation on the concentration of D was determined by pooling data from three similar experiments. These results are shown in Fig. 3B. The fraction of labeled DD in triple helical form, θ (eqn 1, Materials and Methods), is plotted as a function of the concentration of unlabeled third strand D or N. The data were fit to a binding equation (eqn 2, Materials and Methods). The resulting binding curve yields an apparent equilibrium dissociation constant, K_d , of 1.1×10^{-7} M for the DDD complex. This result is comparable to the observed affinity of D for an oligodeoxyribonucleotide duplex that lacks the hairpin loop (28), suggesting that the loop nucleotides do not contribute significantly to the interaction.

Exclusion of RNA strands

Having detected the conventional DDD triplex, experiments were performed to assess the effects of substituting RNA in place of one, two, or all three of the strands of the triplex. This analysis was performed by radiolabeling the four possible hairpin duplexes (DD, DR, RD, RR), and incubating each with increasing concentrations of D or R. Results of these experiments are shown in Fig. 4. Three key observations can be made from these data.

First, the electrophoretic mobilities of hairpin duplexes DD, DR, RD, and RR are distinctly different, showing slower migration with increasing RNA content. Although the physical basis for these mobility differences is unknown, this effect exactly parallels what had been observed for a similar series of homopurine/homopyrimidine hairpins in a previous study (17). Possible explanations include differential counterion condensation due to differences in helical charge density, or differences in helical shape.

Second, as seen previously in Fig. 3, incubation of labeled DD with unlabeled D gave rise to the expected DDD triplex with decreased mobility (Fig. 4A, lanes 1–3). Most of the labeled DD was shifted to DDD in the presence of 1 μM D (Fig. 4A, compare lanes 1 and 3).

Third, stable triplexes were not observed for any of the seven possible complexes involving one or more RNA strands (Fig. 4A, lanes 4–12; Fig. 4B lanes 1–12). Traces of a possible shifted complex were only visible in the case of DR binding by D (Fig. 4A, lane 6) and by R (Fig. 4B, lane 6). These trace complexes may represent triplexes with very low stabilities.

Structural issues

The basis for exclusion of RNA from triplexes in the purine motif is unknown. Skoog and Maher previously used transcriptional repression and electrophoretic mobility shift assays to show that DNA oligonucleotides, but not RNA oligonucleotides, could bind to duplex DNA in the purine motif (11). Possible triplex destabilization due to the absence of a 5-methyl group from uridine in RNA was ruled out because DNA oligomers containing deoxyuridine formed stable triple helices. Triple helix formation improved as the number of deoxyribonucleotide substitutions in the RNA strand was increased.

Escudé *et al.* employed electrophoretic mobility shift and footprinting assays to show a similar result for a different DNA sequence (19). These authors then compared molecular models of dG·dG·dC and rG·dG·dC triplets. Modeling suggested that the 2'-hydroxyl groups in the reverse Hoogsteen strand of the latter complex were unfavorably oriented with respect to the adjacent (5') sugar residues. This reasoning might explain the present results if similar unfavorable interactions can arise in any of the three strands of the triplex.

Comparison of the purine motif (this study) and the pyrimidine motif (17–19) suggests the existence of significant structural differences between the corresponding triplexes. While RNA is excluded from the purine motif, RNA strands are accommodated in a variety of combinations within the pyrimidine motif. Although some disagreement exists as to the relative stabilities of the resulting structures, Roberts and Crothers detected stable complexes for six of the eight possible combinations of RNA and DNA strands (17). Triplex stability was inversely related to duplex instability in this report, and triplexes involving RNA bound to a DNA duplex were among the most stable of the six observed complexes.

Implications for natural RNA recognition and folding

Isolated base triplets exist in the folded structures of tRNA, 5S RNA, and the *Tetrahymena* self-splicing intron (14–16). Some evidence supports the existence of *intramolecular* DNA triple-helices in living bacterial cells (29). Although triple helices might permit gene regulation through the site-specific binding of ribonucleic acid or ribonucleoprotein to duplex DNA (30,31), natural examples of *intermolecular* RNA or DNA triple-helix formation have not been discovered. DNA recognition by ribonucleoproteins has been reported (12,13), but the role of the associated RNA remains unclear.

Prior to the present experimental results, it would have appeared possible that folded RNAs might be stabilized by stacked base triplets in the purine motif. In principle, such structural elements could have produced pseudoknots in which a homopurine/homopyrimidine duplex in a stem-loop structure formed a triple helix with an adjacent guanosine-rich segment of RNA. However, the studies presented here suggest that stacked base triplets of RNA (e.g. **RRR**) cannot give rise to stable triple helices in the purine motif. Such triplets may therefore exist only as isolated structural elements in folded RNAs. It remains possible that some folded RNA structures are stabilized by local triple helices involving stacked base triplets in the pyrimidine motif, but the requirement for cytosine protonation limits the likelihood of such interactions.

Implications for artificial nucleic acid recognition

The requirement for cytosine protonation in the pyrimidine triple helix motif has suggested that the pH-independent purine motif might provide a superior approach to recognition of nucleic acid duplexes for therapeutic purposes. However, triple helix formation in the purine motif is subject to important constraints involving permissive ionic conditions (32,33), and the present results show that the purine triple helix motif cannot be applied to the design of triplexes involving RNA.

ACKNOWLEDGEMENTS

We acknowledge the excellent technical assistance of R.Cerny, D.Eicher, C.Mountjoy, W.Olivas, C.McDonald, G.Soukup and J.Strauss. This work was supported in part by grant 5 P30 C-A36727-08 from the National Cancer Institute and grant GM 47814 from the National Institutes of Health. The Midwest Center for Mass Spectrometry is partially supported by the National Science Foundation, Biology Division (grant no. DIR9017262). C.L.S. was supported by the Summer Research Fellowship Program of the Eppley Institute. L.J.M. 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. Felsenfeld, G., Davies, D.R. and Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024.
2. Chastain, M. and Tinoco, I. (1992) *Nucleic Acids Res.* **20**, 315–318.
3. Sen, D. and Gilbert, W. (1990) *Nature* **344**, 410–414.
4. Maher, L.J. (1992) *BioEssays* **14**, 807–815.
5. Guschlbauer, W., Chantot, J-F. and Thiele, D. (1990) *J. Biomol. Struct. Dyns.* **8**, 491–571.
6. Lee, J.S., Johnson, D.A. and Morgan, A.R. (1979) *Nucleic Acids Res.* **6**, 3073–3091.
7. Moser, H.E. and Dervan, P.B. (1987) *Science* **238**, 645–650.
8. Cooney, M., et al. (1988) *Science* **241**, 456–459.
9. Hélène, C. (1991) *Anti-Cancer Drug Design* **6**, 569–584.
10. Beal, P.A. and Dervan, P.B. (1991) *Science* **251**, 1360–1363.
11. Skoog, J.U. and Maher, L.J. (1993) *Nucleic Acids Res.* **21**, 2131–2138.
12. Davis, T.L., Firulli, A.B. and Kinniburgh, A.J. (1989) *Proc. Natl Acad. Sci. U.S.A* **86**, 9682–9686.
13. Beru, N., Smith, D. and Goldwasser, E. (1990) *J. Biol. Chem.* **265**, 14100–14104.
14. Goddard, J.P. (1977) *Prog. Biophys. Mol. Biol.* **32**, 233–308.
15. Westhof, E., et al. (1989) *J. Mol. Biol.* **207**, 417–431.
16. Michel, F. and Westhof, E. (1990) *J. Mol. Biol.* **216**, 585–610.
17. Roberts, R.W. and Crothers, D.M. (1992) *Science* **258**, 1463–1466.
18. Han, H. and Dervan, P.B. (1993) *Proc. Natl Acad. Sci. U.S.A* **90**, 3806–3810.
19. Escudé, C., et al. (1993) *Nucleic Acids Res.* **21**, 5547–5553.
20. Maher, L.J., Dervan, P.B. and Wold, B.J. (1990) *Biochemistry* **29**, 8820–8826.
21. Piesles, U., et al. (1993) *Nucleic Acids Res.* **21**, 3191–3196.
22. Cantor, C.R. and Schimmel, P.R. (1980) *Biophysical Chemistry*, Vol. III. W.H. Freeman, New York, p. 864.
23. Beal, P.A. and Dervan, P.B. (1991) *Science* **251**, 1360–1363.
24. Durland, R.H., et al. (1991) *Biochemistry* **30**, 9246–9255.
25. Olivas, W.M. and Maher, L.J. (1994) *Biochemistry* **33**, 983–991.
26. Maher, L.J. (1992) *Biochemistry* **31**, 7587–7594.
27. Ing, N.H., et al. (1993) *Nucleic Acids Res.* **21**, 2789–2796.
28. Olivas, W.M. and Maher, L.J. *Biochemistry* (in press).
29. Kohwi, Y., Malkhosyan, S. and Kohwi-Shigematsu, T. (1992) *J. Mol. Biol.* **223**, 817–822.
30. Miller, J.H. and Sobell, H.M. (1966) *Proc. Natl Acad. Sci. U.S.A* **55**, 1201–1205.
31. Minton, K.W. (1985) *J. Exp. Pathol.* **2**, 135–148.
32. Milligan, J.F., et al. (1993) *Nucleic Acids Res.* **21**, 327–333.
33. Cheng, A-J. and Van Dyke, M.W. (1993) *Nucleic Acids Res.* **21**, 5630–5635.