Triplex Tethered Oligonucleotide Probes

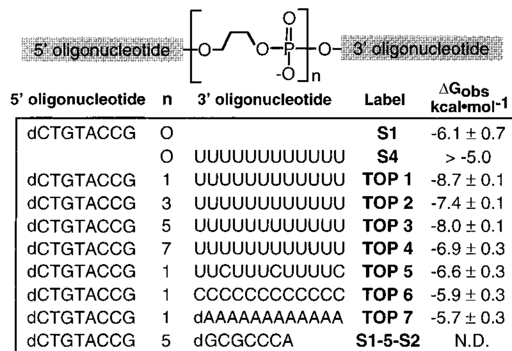
Arikha C. Moses and Alanna Schepartz*

Department of Chemistry, Yale University
P.O. Box 208107, New Haven, Connecticut 06520-8107

Received April 3, 1996

RNA molecules are believed to possess highly complex structures containing single- and double-stranded regions as well as loops, bulges, and pseudoknots.[1] This structural complexity severely complicates efforts to design structure- and/or sequence-selective RNA-binding molecules.[2,3] Recently we described a family of molecules for the sequence- and structure-specific recognition of RNA.[3] These molecules, called tethered oligonucleotide probes (TOPs), are comprised of two short oligonucleotides separated by a flexible, synthetic tether (Figure 1). These oligonucleotides hybridize in a Watson−Crick[4] sense to two noncontiguous, single-stranded regions of a target RNA; the tether traverses the distance between the two regions. TOPs designed to recognize the *L. collosoma* spliced leader RNA and the HIV-1 Rev Response Element RNA (RRE) bind their targets with nanomolar affinity and high specificity.[3d] Here we describe TOPs that hybridize to single- and double-stranded regions of a target RNA in a Watson−Crick[4] and Hoogsteen[5] sense, respectively.[6] These triplex TOPs form high affinity RNA complexes whose stabilities are considerably less sensitive than model triple helices to changes in temperature and monovalent cation concentration.

Because the wtRRE (Figure 2A) lacks an extended polypurine-polypyrimidine site suitable for triple helix formation, we designed a variant RRE containing a 12 base-pair third-strand recognition site (RRE^{AU}, Figure 2B). The natural RRE ligand Rev binds to discrete residues in stem IIB of the wtRRE.[7] We reasoned that lengthening stem IIB while maintaining these residues would retain RRE fold and function. RRE^{AU} and wtRRE bind Rev with similar affinities, with half maximal binding at 20 nM Rev. TOPs designed to recognize RRE^{AU} consist of a 5′ oligodeoxyribonucleotide complementary in a Watson−Crick sense to the bases of RRE^{AU} site 1 and a 3′ oligoribonucleotide complementary in a Hoogsteen sense to the extended region of stem IIB (site 4).[3d]



5′ oligonucleotide	n	3′ oligonucleotide	Label	ΔG_{obs} kcal•mol-1
dCTGTACCG	O		**S1**	-6.1 ± 0.7
	O	UUUUUUUUUUUU	**S4**	> -5.0
dCTGTACCG	1	UUUUUUUUUUUU	**TOP 1**	-8.7 ± 0.1
dCTGTACCG	3	UUUUUUUUUUUU	**TOP 2**	-7.4 ± 0.1
dCTGTACCG	5	UUUUUUUUUUUU	**TOP 3**	-8.0 ± 0.1
dCTGTACCG	7	UUUUUUUUUUUU	**TOP 4**	-6.9 ± 0.3
dCTGTACCG	1	UUCUUUCUUUUC	**TOP 5**	-6.6 ± 0.3
dCTGTACCG	1	CCCCCCCCCCCC	**TOP 6**	-5.9 ± 0.3
dCTGTACCG	1	dAAAAAAAAAAAA	**TOP 7**	-5.7 ± 0.3
dCTGTACCG	5	dGCGCCCA	**S1-5-S2**	N.D.

Figure 1. Sequences of TOPs and oligonucleotides used in this study.

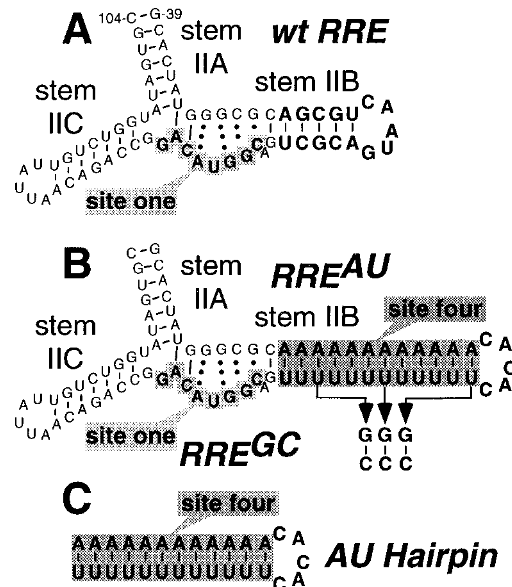


Figure 2. Sequences of **(A)** residues 39−104 of wtRRE, **(B)** RRE^{AU} and RRE^{GC}, and **(C)** the AU hairpin.

To determine whether RRE^{AU} could be recognized with high affinity through the formation of Watson−Crick and Hoogsteen base pairs, we measured the equilibrium dissociation constants of its complexes with TOPs **1**−**4** through the use of a competition electrophoretic mobility shift assay (Figure 3).[3c,8] TOP•RRE^{AU} stabilities at 4 °C ranged from −8.7 kcal•mol^{-1} (K_d = 113 nM) for TOP **1** to −6.9 kcal•mol^{-1} (K_d = 3.6 μM) for TOP **4**. All TOPs tested bound RRE^{AU} with higher affinities than oligonucleotides that recognized RRE^{AU} through Watson−Crick or Hoogsteen base pairs alone: **S1**•RRE^{AU} exhibited a binding free energy of −6.1 kcal•mol^{-1} (K_d = 14 μM) and **S4**•RRE^{AU} exhibited a binding free energy that could only be estimated with this assay (ΔG_{obs} > −5 kcal•mol^{-1}, K_d > 100 μM). TOPs **5** or **6**, containing three or 12 mutations in the 3′ oligonucleotide, bound RRE^{AU} with approximately 2.2 and 2.9 kcal•mol^{-1} lower affinity, respectively, than TOP **1**. TOP **1** bound poorly (ΔG_{obs} = −6.5 kcal•mol^{-1}, K_d = 7.8 μM) to an

† Phone: 203-432-5094. Fax: 203-432-6144. email: alanna@milan.chem.yale.edu.

(1) Tinoco, I. J.; Davis, P. W.; Hardin, C. C.; Puglisi, J. D.; Walker, G. T.; Wyatt, J. *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 135.
(2) (a) Bass, B. L.; Cech, T. R. *Nature (London)* **1984**, *308*, 820. (b) Bass, B. L.; Cech, T. R. *Biochemistry* **1986**, *25*, 4473. (c) Moazed, D.; Noller, H. F. *Nature (London)* **1987**, *327*, 389. (d) Murakawa, G. J.; Chem, C.-h., B.; Kuwabare, M. D.; Nierlich, D. P.; Sigman, D. S. *Nucleic Acids Res.* **1989**, *17*, 5361. (e) Buck, H. M.; Koole, L. H.; van Genderen, M. H. P.; Smit, L.; Geelen, J. L. M. C.; Jurriaans, S.; Goudsmit, J. *Science* **1990**, *248*, 208. (f) Saison-Behmoaras, T.; Tocque, B.; Rey, I.; Chassignol, M.; Thuong, N. T.; Helene, C. *EMBO J.* **1991**, *10*, 1111. (g) Chow, C. S.; Behlen, L. S.; Uhlenbeck, O. C.; Barton, J. K. *Biochemistry* **1992**, *31*, 972. (h) Chow, C. S.; Hartmann, K. M.; Rawlings, S. L.; Huber, P. W.; Barton, J. K. *Biochemistry* **1992**, *31*, 3534. (i) Lim, A. C.; Barton, J. K. *Biochemistry* **1993**, *32*, 11029. (j) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3806. (k) Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, *74*, 969. (l) McConnaughie, A. W.; Spychala, J.; Zhao, M.; Boykin, D.; Wilson, W. D. *J. Med. Chem.* **1994**, *37*, 1063. (m) Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1994**, *22*, 2326. (n) Fenster, S. D.; Wagner, R. W.; Froehler, B. C.; Chin, D. J. *Biochemistry* **1994**, *33*, 8391. (o) Vo, T.; Wang, S.; Kool, E. T. *Nucleic Acids Res.* **1995**, *23*, 2937. (p) McDonald, C. D.; Maher, L. J., III *Nucleic Acids Res.* **1995**, *23*, 500. (q) Jensen, K. B.; Atkinson, B. L.; Willis, M. C.; Koch, T. H.; Gold, L. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12220.
(3) (a) Richardson, P. L.; Schepartz, A. *J. Am. Chem. Soc.* **1991**, *113*, 5109. (b) Cload, S. T.; Schepartz, A. *J. Am. Chem. Soc.* **1991**, *113*, 6324. (c) Cload, S. T.; Richardson, P. L.; Huang, Y.-H.; Schepartz, A. *J. Am. Chem. Soc.* **1993**, *115*, 5005. (d) Cload, S. T.; Schepartz, A. *J. Am. Chem. Soc.* **1994**, *116*, 437. This manuscript discusses the recognition of sites 2 and 3.
(4) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
(5) Hoogsteen, K. *Acta Crystallogr.* **1963**, *16*, 907.
(6) Kessler, D. J.; Pettitt, B. M.; Cheng, Y.-K.; Smith, S. R.; Jayaraman, K.; Vu, H. M.; Hogan, M. E. *Nucleic Acids Res.* **1993**, *21*, 4810.
(7) (a) Heaphy, S.; Dingwall, C.; Ernberg, I.; Gait, M. J.; Green, S. M.; Karn, J.; Lowe, A. D.; Singh, M.; Skinner, M. A. *Cell* **1990**, *60*, 685. (b) Malim, M. H.; Tiley, L. S.; McCarn, D. F.; Rusche, J. R.; Hauber, J.; Cullen, B. R. *Cell* **1990**, *60*, 675. (c) Olsen, H. S.; Nelbock, P.; Cochrane, A. W.; Rosen, C. A. *Science* **1990**, *247*, 845. (d) Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W. *Cell* **1991**, *67*, 529. (e) Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. *EMBO J.* **1992**, *11*, 1119. (f) Tiley, L. S.; Malim, M. H.; Tewary, H. K.; Stockley, P. G.; Cullen, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 758. (g) Tan, R.; Chen, L.; Buettner, J. A.; Hudson, D.; Frankel, A. D. *Cell* **1993**, *73*, 1031. (h) Battiste, J. L.; Tan, R.; Frankel, A. D.; Williamson, J. R. *Biochemistry* **1994**, *33*, 2741.
(8) See supporting information for details.

S0002-7863(96)01096-7 CCC: $12.00 © 1996 American Chemical Society

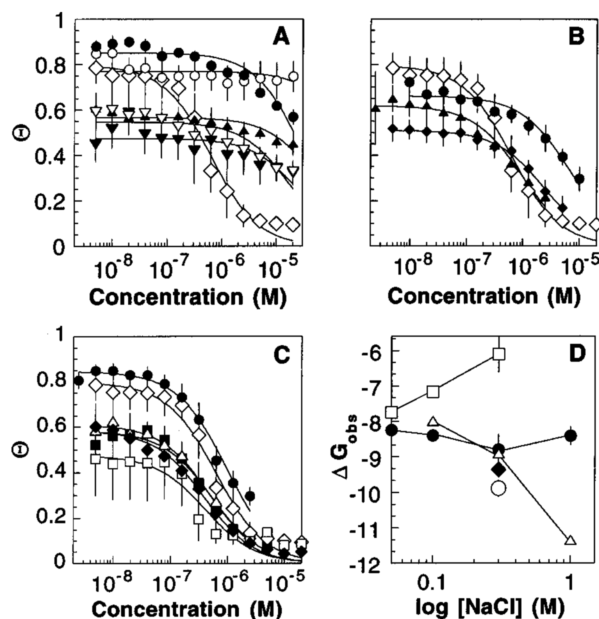


Figure 3. Competition electrophoretic mobility shift experiments[8] were performed at 4 °C unless otherwise indicated. θ = fraction of S1−5-S2 bound to RRE^{AU}. The *x*-axis represents the concentration of competitor TOP or oligonucleotide. K_d values in the high micromolar range are imprecise due to insufficient data. **(A)** ▼, S1; (◇), TOP **1**; ●, TOP **5**; ▲, TOP **6**; ○, TOP **7**; ▽, TOP **1**·RRE^{GC}. **(B)** (◇), TOP **1**; ◆, TOP **2**; ▲, TOP **3**; ●, TOP **4**. **(C)** ■, TOP **1**, 50 mM NaCl; △, TOP **1**, 100 mM NaCl; (◇), TOP **1**, 300 mM NaCl; □, TOP **1**, 1 M NaCl; ●, TOP **1**, 300 mM NaCl, 25 °C incubation, ◆, TOP **1**, 300 mM NaCl, 37 °C incubation. **(D)** Plot of [NaCl] versus ΔG_{obs} for ●, TOP **1**·RRE^{AU}, 4 °C; (◆), TOP **1**·RRE^{AU}, 25 °C, (○), TOP **1**·RRE^{AU}, 37 °C; □, **S1**·RRE^{AU}, 4 °C; and △, **S4**·AU hairpin, 4 °C.

RRE^{AU} analog (RRE^{GC}) containing three GC base pairs within site 4 (Figure 2B). A TOP **1** variant engineered to recognize site 4 through Watson−Crick base pairs (TOP **7**) also displayed low RRE^{AU} affinity ($\Delta G_{obs} = -5.7$ kcal·mol^{-1}, $K_d = 32$ μM), presumably because it is unable to invade the stable site 4 duplex. Our data indicate that TOP·RRE stability is affected in a predictable way by mutations in the 3′ oligonucleotide and site 4. This data is most consistent with a model in which TOPs **1**−**4** recognize site 4 via Hoogsteen interactions and not via nonspecific interactions elsewhere on the RRE; RNase H mapping experiments provide evidence for recognition of site 1 via Watson−Crick interactions.[8]

Comparison of **S1**·RRE^{AU}, **S4**·RRE^{AU}, and TOP **1**·RRE^{AU} stabilities reveals that TOP **1** binds RRE^{AU} with higher affinity than either of its constituent oligonucleotides and that formation of the site 4 triple helix contributes a net 2.7 kcal·mol^{-1} to the complex. At 4 °C, this value is 6.2 kcal·mol^{-1} lower than the −8.9 kcal·mol^{-1} binding free energy of the model **S4**·AU triple helix (Figure 2C) under identical buffer conditions.[8] The dramatic context dependence of triple helix stability suggests that site 4, RRE^{AU}, or the TOP must undergo costly structural deformations upon binding and emphasizes the challenge of structured RNA recognition. Notably, even the modest 2.7 kcal·mol^{-1} contribution of triple helix formation to TOP **1**·RRE^{AU} stability converts a micromolar ligand (**S1**) into a nanomolar ligand (TOP **1**) and permits selective recognition of two RRE^{AU} regions that alone are not recognized above the level of a nonspecific interaction.

Because the site 4 triple helix contributes only a fraction of the TOP **1**·RRE^{AU} binding energy, we wondered whether triplex TOP·RRE^{AU} complexes would be less sensitive to monovalent cation concentration than model RNA triple helices.[9] Comparison of TOP **1**·RRE^{AU}, **S4**·AU hairpin, and **S1**·RRE^{AU} stabilities between 0.05 and 1 M NaCl reveals that **S4**·AU hairpin stability decreases by >3 kcal·mol^{-1} between 1 and 0.05 M NaCl (Figure 3D). TOP **1**·RRE^{AU} stability remains constant within this range, which includes physiological conditions. Not only is the TOP **1**·RRE^{AU} complex less sensitive to salt than the model triple helix, it is also less sensitive to temperature. TOP **1**·RRE^{AU} stability increases by 1.1 kcal·mol^{-1} between 4 and 37 °C, despite the fact that the **S4**·AU hairpin triple helix is completely dissociated at 25 °C under identical buffer conditions (T_m of 12 °C at 2 μM total strand concentration). The increase in triple helix stability within the context of RRE^{AU} presumably results from the increased effective concentration of the site 4 oligonucleotide (approximately 12 mM)[10] within the context of a TOP bound at site 1. The **S1**·RRE^{AU} complex increases in stability by 1.3 kcal·mol^{-1} between 4 and 25 °C and decreases significantly at 37 °C. We conclude that triple helix formation within the context of RRE^{AU} is more favorable at 37 °C than at 4 °C, perhaps because of the increased population of a suitably structured target site.[11]

In summary, we demonstrate that TOPs are able to simultaneously recognize single- and double-stranded regions within a large, structured RNA. This property suggests that TOPs may have application in the functional manipulation of RNAs *in vitro*, and, with suitable backbone modifications, *in vivo*.[12] Moreover, the observation that short RNA triple helices can form readily under physiological conditions of pH, salt, and temperature suggests that these structures could play important roles in RNA folding.

Acknowledgment. This work was supported by the NIH (GM43501). A.C.M. thanks Hoffmann-La Roche for a graduate fellowship.

Supporting Information Available: Experimental details (9 pages). See any current masthead page for ordering and Internet access instructions.

JA961096T

(9) (a) Felsenfeld, G.; Rich, A. *Biochim. Biophys. Acta* **1957**, *26*, 457. (b) Stevens, C. L.; Felsenfeld, G. *Biopolymers* **1964**, *2*, 293. (c) Riley, M.; Maling, B. *J. Mol. Biol.* **1966**, *20*, 359. (d) Singleton, S. F.; Dervan, P. B. *Biochemistry* **1992**, *32*, 13171. (d) Rougee, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochemistry* **1992**, *31*, 9269. (e) Escude, C.; Francois, J.-C.; Sun, J.; Ott, G.; Sprinzl, M.; Garestier, T.; Helene, C. *Nucleic Acids Res.* **1993**, *21*, 5547. (f) Bond, J.; Anderson, C. F.; Record, M. T. *Biophys. J.* **1994**, *67*, 825. (g) For general references on nucleic acid triple helices, see: Dervan, P. B. *Top. Mol. Struct. Biol.* **1989**, *12*, 197. Strobel, S. A.; Dervan, P. B. *Meth. Enzymol.* **1992**. Cheng, Y.-K.; Pettitt, B. M. *Prog. Biophys. Molec. Biol.* **1992**, *58*, 225. Sun, J.-S.; Helene, C. *Curr. Opin. Struct. Biol.* **1993**, *3*, 345. Plum, G. E.; Pilch, D. S.; Singleton, S. F.; Breslauer, K. J. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 319.

(10) Effective molarity was calculated as described: Kim, B.; Little, J. W. *Science* **1992**, *255*, 203.

(11) Battiste, J. L.; Tan, R.; Frankel, A. D.; Williamson, J. R. *Biochemistry* **1994**, *33*, 2741.

(12) Mesmaeker, A. D.; Häner, R.; Martin, P.; Moser, H. E. *Accts. Chem. Res.* **1995**, *28*, 366.